

Web-Enabled Landsat Data (WELD): A Consistent, Long-Term, Large-Area, 30m Data Record For The Terrestrial User Community

D.P. Roy, J. Ju, I. Kommadreddy

Geographic Information Science Center of Excellence, South Dakota State University, Brookings, SD 57007, USA,
david.roy@sdstate.edu, junchang.ju@sdstate.edu, Indrani.Kommareddy@sdstate.edu

Abstract – Since January 2008, the U.S. Geological Survey has been providing free terrain-corrected and radiometrically calibrated Landsat data via the Internet. This revolutionary data policy provides the opportunity to use all the data in the U.S. Landsat archive and to consider the systematic utility of Landsat data for long-term large-area monitoring. The NASA funded Web-enabled Landsat Data (WELD) project is systematically generating 30m weekly, monthly, seasonal, and annual composited Landsat Enhanced Thematic Mapper Plus (ETM+) mosaics of the conterminous United States and Alaska from 2005 to 2012. The WELD products are developed specifically to provide consistent data that can be used to derive land cover as well as geophysical and biophysical products for regional assessment of surface dynamics and to study Earth system functioning. The products are processed so that users do not need to apply the equations and spectral calibration coefficients and solar information to convert the ETM+ digital numbers to reflectance and brightness temperature, and successive products align precisely, making them simple to use for multi-temporal applications. The consistency and quality of the ETM+ data will be improved through a fusion with standard MODIS land products, including the MODIS BRDF reflectance anisotropy product to radiometrically normalize and fill missing (cloudy and SLC-off) Landsat pixels and using MODIS atmospheric characterization data to systematically atmospherically correct the Landsat data. Information on the WELD products, how to access them via the Internet, and preliminary examples of prototype global WELD 30m products are presented.

Key words: Landsat ETM+; Web-enabled Landsat Data

1. INTRODUCTION

The series of existing and planned Landsat satellites, operated by the U.S. Geological Survey (USGS) Landsat project with satellite development and launches supported by NASA, represent the longest temporal record of space-based earth observations (Williams et al. 2006). Landsat data provide a balance between the requirements for localized high-spatial resolution studies and global monitoring (Goward et al. 2001). Since January 2008, the USGS has been providing free terrain corrected (Level 1T) Landsat data via the Internet at no cost (Woodcock et al. 2008). This revolutionary data policy provides the opportunity to use all the data in the US Landsat Data Archive and to consider the systematic utility of Landsat data for long-term large-area monitoring.

The NASA funded Web-enabled Landsat Data (WELD) project is systematically generating 30m composited Landsat Enhanced Thematic Mapper Plus (ETM+) mosaics of the conterminous United States (CONUS) and Alaska from every ETM+ acquisition with cloud cover <80% (Roy et al. 2010). The resulting high spatial resolution 30m Landsat mosaic products are made freely available to the user community.

2. WELD PROCESSING

The WELD processing is applied to all Landsat ETM+ L1T acquisitions with cloud cover <80% sensed over the conterminous United States (CONUS) and Alaska, approximately 8,000 and 1,800 acquisitions per year respectively. A WELD product versioning scheme is used to reflect product reprocessing using improved algorithms, ancillary data and sensor knowledge. The currently available products are Version 1.5 and provide the information summarized in Table 1.

Table 1: WELD Version 1.5 product contents (all defined at 30m)

Science Data Set Name	Data Type	Scale factor	Units
Band1 TOA REF	int16	10000	Unitless
Band2 TOA REF	int16	10000	Unitless
Band3 TOA REF	int16	10000	Unitless
Band4 TOA REF	int16	10000	Unitless
Band5 TOA REF	int16	10000	Unitless
Band61 TOA BT	int16	100	Degrees Celsius
Band62 TOA BT	int16	100	Degrees Celsius
Band7 TOA REF	int16	10000	Unitless
NDVI TOA	int16	10000	Unitless
Day Of Year	int16	1	Day
Saturation Flag	uint8	1	Unitless
DT Cloud State	uint8	1	Unitless
ACCA State	uint8	1	Unitless
Num Of Obs	uint8	1	Unitless

The spectral radiance sensed by each Landsat ETM+ detector is first converted to radiance (units: $W m^{-2} sr^{-1} \mu m^{-1}$) using the calibration information available with each L1T acquisition. The radiance sensed in the reflective and thermal wavelengths are then converted to top of atmosphere (TOA) reflectance (unitless) and brightness temperature (units: kelvins) respectively using standard formulae (Chander et al. 2009). This provides data that has physical meaning and, for example, can be compared with laboratory and ground based measurements, model outputs, and data from other satellite sensors, and importantly can be used by the community to derive higher level geo-physical and bio-physical products. The Version 1.5 products are not corrected for atmospheric effects, although the planned Version 2.0 products will be (Ju et al. 2011). The normalized difference vegetation index (NDVI) is derived from the TOA red and near-infrared reflectance to provide a heritage vegetation index. Per-band radiometric saturation identification and two cloud masks based on heritage ACCA (Irish 2006) and a recent classification tree (Roy et al. 2010) algorithm are also generated.

The L1T processed data are reprojected into the Albers Equal Area projection (Snyder 1993). Given the high data volume it is not possible to store the reprojected CONUS or Alaska Landsat 30m data in single files. Instead the data are stored in fixed earth-located tiles composed of 5000 x 5000 30m pixels; there are 501 CONUS and 162 Alaskan tiles. Figure 1 illustrates the CONUS tiling scheme.

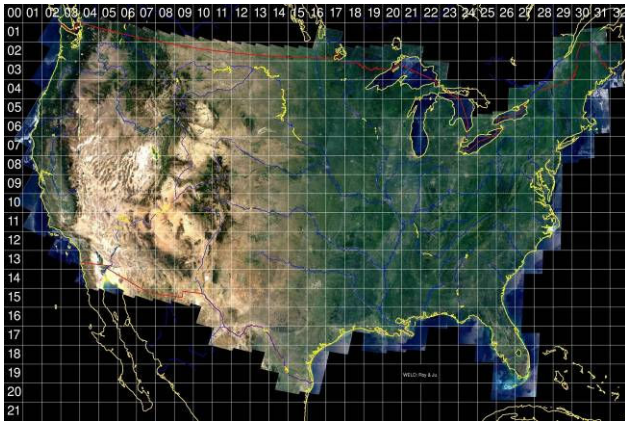


Figure 1: CONUS WELD product tiling scheme (white vectors, each tile composed of 5000 x 5000 30m pixels) shown superimposed on the annual 2008 WELD product true color TOA red (0.63-0.69µm), green (0.53-0.61µm) and blue (0.45-0.52µm) ETM+ reflectance.

Each Albers projected LIT acquisition is composited temporally into *weekly*, *monthly*, *seasonal (3 month)* and *annual* WELD products using criteria developed to select observations with minimal cloud and atmospheric contamination (Roy et al. 2010). The standard seasonal definition adopted by the climate modeling community is used where Winter is defined by the months: December, January and February, and a year is the 12 months from December of one year to the November of the following year (Ju and Roy 2008). The weekly products are defined as consecutive 7-day products starting January 1st. The compositing criteria are based on refinements to heritage maximum NDVI and maximum brightness temperature criteria (Holben 1986, Cihlar et al. 1994, Roy 1997). For each composited pixel the day of the year that the selected pixel was acquired on and the number of different valid acquisitions considered at that pixel over the compositing period are stored (Table 1).

The *weekly*, *monthly*, *seasonal (3 month)* and *annual* WELD products are stored in Hierarchical Data Format (HDF). HDF is a data file format designed by the U.S. National Center for Supercomputing Applications to assist users in the storage and manipulation of scientific data across diverse operating systems and machines. For example, HDF is used to store the standard MODIS Land products (Justice et al. 2002).

3. WELD PRODUCT AVAILABILITY

The WELD products are distributed from the USGS EROS and are freely available via the Internet. Currently the Version 1.5 *weekly*, *monthly*, *seasonal (3 month)* and *annual* WELD products for the CONUS and Alaska are available for 2006, 2007, 2008, 2009 and 2010. This provides a total of about 20TB of data with internal compression on.

The WELD products are available in **HDF** in 5000 x 5000 30m pixel tiles via ftp (<ftp://weldftp.cr.usgs.gov/>). In addition, the products are also available via an intuitive What You See Is What You Get (WYSIWYG) Internet interface developed at SDSU and ported to the USGS EROS (<http://weld.cr.usgs.gov/>). Users of the WYSIWYG require a web browser with Java script enabled. The system allows users to interactively order any rectangular spatial subset of any WELD product, up to 2GB, in a way that the WELD tile structure is transparent to the user. Users are able to interactively select and view any of the CONUS or Alaska

products, and pan and zoom (to a spatial resolution of 210m) within the selected browse imagery. Users may order an arbitrary rectangular geographic area of interest, either interactively, by moving a rubber band box over the displayed browse image, or by specifying geographic coordinates in a text field. Once a region of interest has been selected the appropriate WELD tiles are assembled, subset and mosaiced, and placed on an HTTP site in **GeoTIFF** format. The user is sent an email with the relevant HTTP access information.

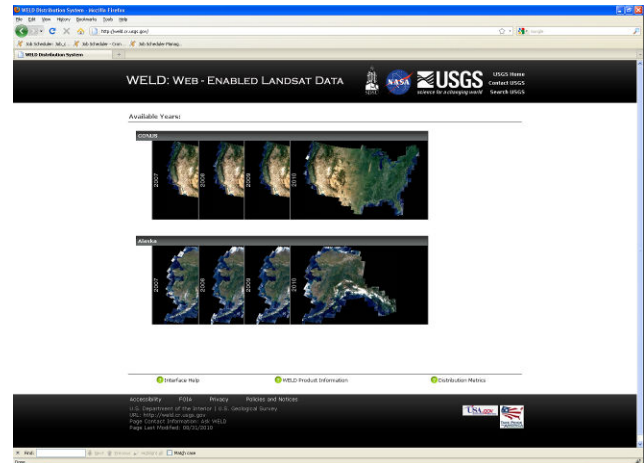


Figure 2: Screen shot of the top web page of the WELD What You See Is What You Get interface, showing the 2007-2010 CONUS and Alaska annual products.

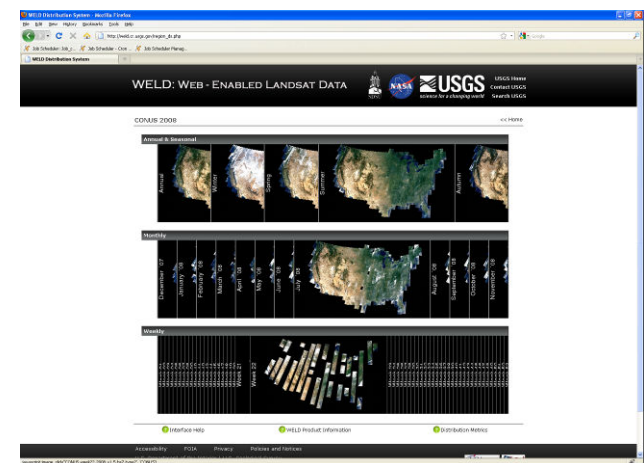


Figure 3: Screen shot of the WELD What You See Is What You Get interface, showing the CONUS 2008, annual and 4 seasonal (top row), 12 monthly (middle row) and 53 weekly (bottom row) products.

More information and documentation, including known product issues, software tools, a Frequently Asked Questions (FAQ) section, sample browse images, and a WELD help email, are available on the WELD Product Documentation web site <http://globalmonitoring.sdstate.edu/projects/weld/>.

4. PLANNED WELD PRODUCT IMPROVEMENTS

All years from 2005 to 2012 will be reprocessed as improved versions of the algorithms are developed. Major and minor version changes are reflected in the first and second digits respectively of the version number. In general it is preferable to use the latest WELD product version which reflects improvements to the WELD processing algorithms and input

data. The next major WELD reprocessing will be Version 2.0 and will have the following elements:

- atmospheric correction of the top of atmosphere reflectance bands (Ju et al. 2011)
- radiometric normalization of the reflectance to nadir view and fixed solar zenith angle (Roy et al. 2008)
- gap filling of the Landsat ETM+ SLC-off gaps in the reflectance and thermal bands (Roy et al. 2008)
- 30m percent land cover: weekly snow/ice, water and bare ground; annual peak greenness tree cover, bare ground and other vegetation (Hansen et al. 2011).

5. GLOBAL WELD PRODUCTS ?

The WELD project has demonstrated the potential of large volume Landsat data processing to provide, on a systematic repeat basis, near continental scale data products that provide a high spatial resolution analogue to the moderate and coarse spatial resolution land products generated from the MODIS and AVHRR data streams (Justice et al. 2002, Tucker et al. 2005).

The challenges to processing global scale Landsat data are considerable. Global Landsat processing is a “Big Data” issue - there are approximately 150 million MODIS 1km land pixels globally and three orders of magnitude more Landsat 30m pixels; the conterminous United States alone is composed of 11,000,000,000 30m pixels (Roy et al. 2010). Figure 4 shows as a proof of concept the first single-month near global coverage Landsat ETM+ composite ever produced. This data set took 2.5 weeks to process on the WELD computing system; 6,796 ETM+ acquisitions with cloud cover < 40% obtained from the USGS EROS archive were used. The processing challenge is not only data volume; global satellite product generation experience has shown a need for algorithm refinement and product reprocessing (Townshend, 1994, Masuoka et al. 2010). For these reasons, generation of global WELD products is now being investigated in a NASA cloud supercomputing environment.



Figure 4: Global WELD proof of concept: single month global coverage Landsat ETM+ composite depicting “leaf-on” conditions generated using ETM+ data acquired in January 2010 (Southern hemisphere <20°), October 2009 (Tropical latitudes ±20°) and July 2009 (Northern hemisphere >20°). True color TOA 30m ETM+ reflectance is shown, defined in the MODIS Land sinusoidal projection.

REFERENCES

G. Chander, B. Markham, D. Helder, “Summary of Current Radiometric Calibration Coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI Sensors”, *Remote Sens. of Env.* Vol. 113, p.p. 893–903, 2009.

J. Cihlar, D. Manak, M. D’Iorio, “Evaluation of compositing algorithms for AVHRR data over land”, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, p.p. 427-437, 1994.

S. Goward, J. Masek, D. Williams, J. Irons, R.J. Thompson, “The Landsat 7 mission, Terrestrial research and applications for the 21st century”, *Remote Sensing of Env.*, Vol. 78, p.p. 3-12, 2001.

M.C., Hansen, A. Egorov, D.P. Roy, P. Potapov, J. Ju, S. Turubanova, I. Kommareddy, T. Loveland, “Continuous fields of land cover for the conterminous United States using Landsat data: First results from the Web-Enabled Landsat Data (WELD) project”, *Remote Sensing Letters*, Vol. 2(4), p.p. 279-288, 2011.

B. Holben, “Characteristics of maximum-value composite images from temporal AVHRR data”, *International Journal of Remote Sensing*, Vol. 7, p.p. 1417-1434, 1986.

R.I. Irish, J.L. Barker, S.N. Goward, T. Arvidson, “Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm”, *Photogrammetric Engineering and Remote Sensing*, Vol. 72, p.p. 1179 – 1188, 2006.

J. Ju, and D.P. Roy, “The Availability of Cloud-free Landsat ETM+ data over the Conterminous United States and Globally”, *Remote Sensing of Environment*, Vol. 112, p.p. 1196-1211, 2008.

J. Ju, E. Vermote, D.P. Roy, J. Masek, “Continental Scale Validation of MODIS-Based and LEDAPS Operational Landsat ETM+ Atmospheric Correction Methods”, This conference, 2011.

C.O. Justice, J.R.G. Townshend, E. Vermote, E. Masuoka, E., R. Wolfe, N. Saleous, D.P. Roy, J. Morisette, “An overview of MODIS Land data processing and product status”, *Remote Sensing of Environment*, Vol. 83, p.p. 3-15, 2002.

E. Masuoka, D.P. Roy, R. Wolfe, J. Morisette, S., Sinno, M. Teague, N. Saleous, S. Devadiga, C.O. Justice, J. Nickeson, MODIS land data products—generation, quality assurance and validation, chapter 22 in Ramachandran, B., Justice, C.O., and Abrams, M.J., eds., *Land remote sensing and global environmental change—NASA’s Earth Observing System and the science of ASTER and MODIS*, New York, Springer, p.p. 511-534, 2010.

D.P. Roy, “Investigation of the maximum normalised difference vegetation index (NDVI) and the maximum surface temperature (Ts) AVHRR compositing procedures for the extraction of NDVI and Ts over forest”, *International Journal of Remote Sensing*, Vol. 18, p.p. 2383-2401, 1997.

D.P. Roy, J. Ju, P. Lewis, C. Schaaf, F. Gao, M. Hansen, E. Lindquist, “Multi-temporal MODIS-Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data”, *Remote Sensing of Environment*, Vol. 112, p.p. 3112-3130, 2008.

D.P. Roy, J. Ju, K. Kline, P. Scaramuzza, V. Kovalsky, M. Hansen, T. Loveland, E. Vermote, C. Zhang, “Web-enabled Landsat Data (WELD): Landsat ETM+ Compositing Mosaics of the Conterminous United States”, *Remote Sensing of Env.*, Vol. 114, p.p. 35-49, 2010.

J.P. Snyder, *Flattening the Earth: Two thousand years of map projections*, The University of Chicago Press, Chicago and London, 1993.

J.R.G. Townshend “Global data sets for land applications from the Advanced Very High-Resolution Radiometer—an introduction”, *Int. Journal of Remote Sensing*, Vol. 15, p.p. 3319–3332, 1994.

C.J. Tucker, J. Pinzon, M.E. Brown, D.A. Slayback, E., Pak, R. Mahoney, “An extended AVHRR 8-km NDVI data set compatible with MODIS and SPOT vegetation NDVI data”, *International Journal of Remote Sensing*, Vol. 26(20), p.p. 4485–4498, 2005.

D.L. Williams, S. Goward, T. Arvidson, “Landsat: yesterday, today, and tomorrow”, *Photogrammetric Engineering and Remote Sensing*, Vol. 72, p.p. 1171-1178, 2006.

Woodcock, C.E. et al. “Free access to Landsat imagery”, *Science*, Vol. 320, p. 1011, 2008.