

Documentation for the  
Global Man-made Impervious Surface (GMIS) Dataset From  
Landsat

**November 2017**

Eric C. Brown de Colstoun<sup>1</sup>, Chengquan Huang<sup>2</sup>, Panshi Wang<sup>2</sup>, James C. Tilton<sup>3</sup>, Bin Tan<sup>4,6</sup>, Jacqueline Phillips<sup>5</sup>, Sarah Niemczura<sup>5</sup>, Pui-Yu Ling<sup>2</sup> and Robert Wolfe<sup>6</sup>

<sup>1</sup> Biospheric Sciences Laboratory, NASA Goddard Space Flight Center, Mailstop 618, NASA/GSFC, Greenbelt, MD 20771 USA

<sup>2</sup> Department of Geographical Sciences, University of Maryland, College Park, MD USA

<sup>3</sup> Computational and Information Sciences and Technology Office, NASA/GSFC, Greenbelt, MD USA

<sup>4</sup> Science Systems and Applications, Inc., Lanham, MD USA

<sup>5</sup> Universities Space Research Association, Columbia MD USA\*

<sup>6</sup> Terrestrial Information Systems Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD USA

**Abstract**

Urbanization is an important driver of change across our home planet. With over half of the world's population living in urban areas today, the mapping and monitoring of urbanization is critical to understanding these changes and their potential impacts. The availability of high resolution, free satellite imagery at multiple epochs from the Global Land Survey (GLS) Landsat archive provides great opportunities to map global man-made surfaces and extent in unprecedented detail. The Global Man-made Impervious Surface (GMIS) Dataset From Landsat consists of global estimates of fractional impervious cover derived from the GLS Landsat dataset for the target year 2010. The GMIS dataset consists of two components: 1) global percent of impervious cover; and 2) per-pixel associated uncertainty for global impervious cover. These layers are co-registered to the same spatial extent at a common 30m spatial resolution. The spatial extent covers the entire globe except Antarctica and some small islands. This dataset is the first global, 30m dataset of man-made impervious cover to be derived from the GLS data for 2010 and is a companion dataset to the Global Human Built-up And Settlement Extent (HBASE) dataset.

---

\*J. Phillips is currently with Alion Science and Technology (McLean, VA) and S. Niemczura is with MacDonald, Dettwiler and Associates Information Systems LLC (Gaithersburg, MD)

**Data set citation:**

Brown de Colstoun, E. C., C. Huang, P. Wang, J. C. Tilton, B. Tan. J. Phillips, S. Niemczura, P.-Y. Ling, and R. E. Wolfe. 2017. Global Man-made Impervious Surface (GMIS) Dataset from Landsat. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <https://doi.org/10.7927/H4P55KKF>. Accessed DAY MONTH YEAR.

**Suggested citation for this document:**

Brown de Colstoun, E. C., C. Huang, P. Wang, J. C. Tilton, B. Tan. J. Phillips, S. Niemczura, P.-Y. Ling, and R. E. Wolfe. 2017. Documentation for the Global Man-made Impervious Surface (GMIS) Dataset from Landsat. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <https://doi.org/10.7927/H4JD4TVQ>. Accessed DAY MONTH YEAR.

We appreciate feedback regarding this dataset, such as suggestions, discovery of errors, difficulties in using the data, and format preferences. Please contact:

NASA Socioeconomic Data and Applications Center (SEDAC)  
Center for International Earth Science Information Network (CIESIN)  
Columbia University  
Phone: 1 (845) 365-8920  
Email: [ciesin.info@ciesin.columbia.edu](mailto:ciesin.info@ciesin.columbia.edu)  
Author Email: [eric.c.browndecolsto@nasa.gov](mailto:eric.c.browndecolsto@nasa.gov)  
Author Email: [pswang@umd.edu](mailto:pswang@umd.edu)

**Contents**

I.	Introduction.....	3
II.	Data and Methodology.....	4
III.	Data Set Description(s).....	8
IV.	How to Use the Data.....	9
V.	Potential Use Cases.....	9
VI.	Limitations.....	9
VII.	Acknowledgments.....	10
VIII.	Disclaimer.....	10
IX.	Use Constraints.....	10
X.	Recommended Citation(s).....	11
XI.	Source Code.....	11
XII.	References.....	11
XIII.	Documentation Copyright and License.....	13
	Appendix 1. Revision History.....	13
	Appendix 2. Contributing Authors & Documentation Revision History.....	13

## I. Introduction

Urban areas represent one of the most important forms of land use and the changes associated with their expansion are also some of the most important forms of land cover/use changes. These changes are typically permanent and at least proportional to the increase in worldwide population. While urban areas still represent today a small proportion of the Earth's surface (~3%, Balk et al., 2004; Imhoff et al., 2004), their impacts on hydrology, weather, resource demand and utilization, and emissions, for example, are increasingly felt from regional to continental, and even global scales (Quattrochi and Ridd, 1994; Bonan 1997; Shepherd, 2005; Jin et al., 2005). As we look to the not too distant future with a rapidly urbanizing planet (United Nations, 2008), it becomes imperative to develop the tools with which we can accurately measure and monitor urban expansion and its characteristics. These are central to a better understanding of the impact and consequences of this change, the impacts and consequences of a changing climate on future cities and settlements, but also any potential adaptation and/or mitigation strategies for climate change that could be developed for the future.

While there have been successes in mapping global urban areas from satellites (e.g. Imhoff et al., 1997; Elvidge et al., 1999; Schneider et al., 2003, 2009) and/or gridded socioeconomic data (e.g. Balk et al., 2004; Salvatore et al., 2005), the accuracy of these products has been limited by the coarse resolution of the sensors (~500m - 1km) and/or ancillary data, but also by the difficulties in accurately mapping a locally and globally heterogeneous 'urban' environment. The recent availability of consistent, global scale datasets at ~30m resolution such as the *Geocover* (Tucker et al., 2004) and its successor Global Land Survey (GLS) (Gutman et al., 2008, 2013) from the Landsat satellites provides an unprecedented opportunity to map global urbanization at this resolution for the first time, with unprecedented detail and accuracy. Moreover, the spatial resolution of Landsat is absolutely essential to accurately resolve urban targets such as buildings, roads and parking lots (Small, 2005). Finally, with GLS data now available for the 1975, 1990, 2000, 2005 and 2010 time periods, the land cover/use changes due to urbanization over 35+ years can also be quantified at this spatial scale.

The successes of the U.S. Geological Survey's (USGS) National Land Cover Dataset (NLCD) (Homer et al., 2004) and the NASA-funded Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Masek et al., 2006a, 2008) have demonstrated the feasibility of routinely processing continental-scale Landsat data, for multiple time periods, and to derive products (e.g. land cover, impervious cover, forest disturbance) from these data. Approximately 30m global Landsat surface reflectance and forest change products have been generated at the Global Land Cover Facility (GLCF) at the University of Maryland at College Park (UMCP) through projects funded under NASA's Land Cover-Land Use Change (LCLUC) and Making Earth System data records for Use in Research Environments (MEASURES) Programs. Finally, the general availability, free-of-charge, very high resolution commercial satellite data through the unclassified Web-based Access and Retrieval Portal (WARP) of the National Geospatial-Intelligence

Agency (NGA) makes it possible to develop 1-4m impervious cover training data that can be used to create and validate such a global product. Our methods build upon the NLCD approach, and were applied in much the same way Landsat data are used to train regression tree algorithms for the Vegetation Continuous Fields (VCF) products being produced from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Hansen et al., 2004).

In addition to producing 30m impervious cover using methods that have been proven effective for continuous fields land cover mapping, we also developed a companion Global Human Built-up And Settlements Extent (HBASE) mask through this project. The purpose of doing this was twofold: 1) to address the need of a 30m urban area and settlements extent map; and 2) to address the problem of over-prediction of impervious cover in areas covered by dry soil, sands, rocks, etc. Using the HBASE mask, we can remove retrievals of impervious surface in such areas. These two datasets are available as individual layers here but are intimately connected.

## **II. Data and Methodology**

The methodology described below is only for the GMIS dataset. For the HBASE dataset, we refer users to its appropriate documentation for details on that methodology. As mentioned above, HBASE was produced as an effective way to remove or mask out errors of commission in GMIS, typically over bare soil surfaces. Figure 1 gives the overall structure of our methodology.

### **Step 1: GLS 2010 Surface Reflectance Dataset**

The 2010 Global Land Survey (GLS) Landsat image archive is used in this study and was provided by the USGS. The GLS series of datasets were developed through a joint NASA-USGS collaboration to provide a convenient basis for developing global land-cover and change products (Gutman et al. 2008; Gutman et al. 2013). The GLS goal was to provide one cloud free image acquired during the peak growing season per epoch per location, although sub-optimal images (e.g. cloudy) were also used when a qualified image was not available. The GLS 2010 dataset used in this study consisted of about 53% Landsat 5 Thematic Mapper (TM) images, 42% gap-filled Landsat 7 Enhanced Thematic Mapper (ETM+) images, and 5% Earth Observer 1 (EO-1) Advanced Land Imaging (ALI) images. Here, the “gap-filled” Landsat 7 data have been nominally corrected to account for the failure of the ETM+ Scan-Line Corrector (SLC) using data adjacent in either time and/or space (Gutman et al. 2013). The datasets used in this study have also been atmospherically corrected and converted to surface reflectance (SR) using the LEDAPS algorithms (Masek et al. 2006) at the Global Land Cover Facility (GLCF) at the University of Maryland. Assessments of similar datasets for 2000 and 2005 have shown that GLS SR has high agreement with MODIS SR and has satisfactory radiometric quality (Feng et al. 2013).

## Step 2: Global Urban Extent Mapping

One of the main challenges in mapping subpixel urban imperviousness is that some non-urban areas such as bare surfaces and fallow fields may be spectrally confused with urban impervious materials. As a result, subpixel estimation algorithms often over predict urban imperviousness in such areas. A common approach to mitigate this problem is to define a non-urban area mask and set the impervious value to 0 for all pixels within this mask. Such a mask is often developed based on coarse resolution nightlight data and other ancillary data sources (Homer et al., 2004). In this study, we developed a novel approach that integrates spatial texture and contextual information to map the spatial extent of an urban area (i.e. HBASE). Subpixel imperviousness was calculated only for pixels within the urban extent mask. Those outside of this mask were assigned an impervious value of 0.

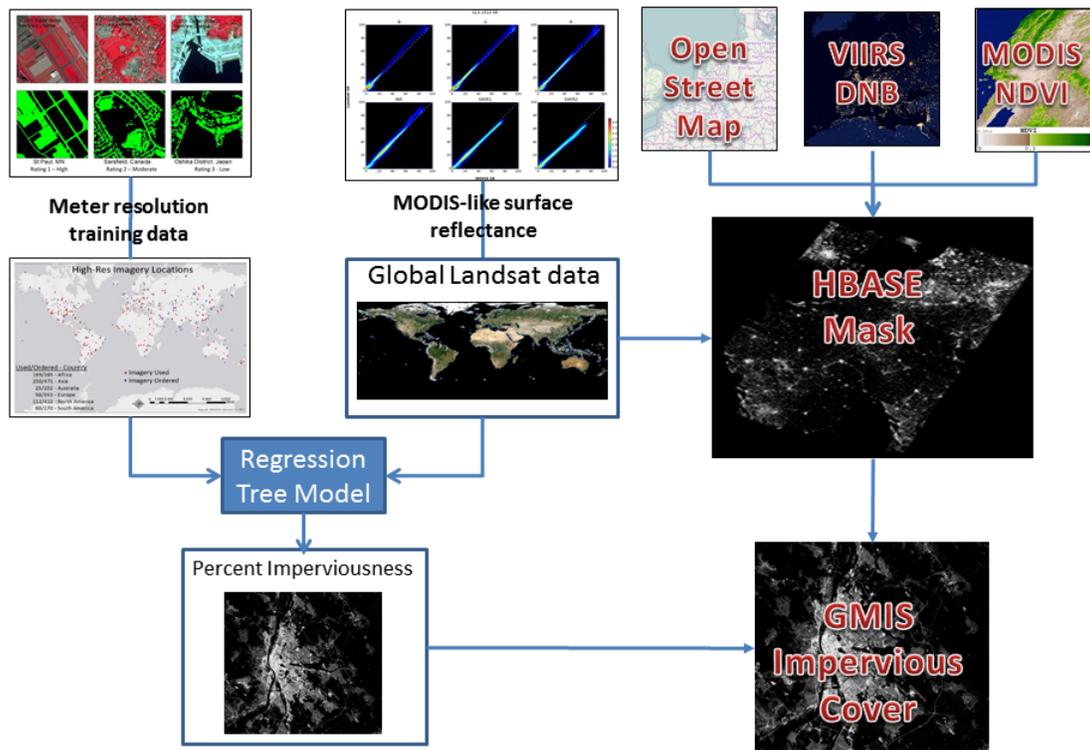


Figure 1. Overall flowchart for the GMIS product methodology

### Step 3: Development of A Global Urban Imperviousness Reference Dataset

Our overall strategy for deriving global impervious cover from the GLS data was based on the approach used by Homer et al., (2004) and Yang et al., (2003) in developing the urban imperviousness product of the NLCD. It works across spatial scales by using the very high spatial resolution commercial satellite data acquired from the National Geospatial-Intelligence Agency (NGA) to produce training data at the 30m spatial resolution. These training data are then used to train a regression tree algorithm (Cubist) to generate continental scale products on a scene-by-scene basis. This type of approach was first established in the remote sensing community to derive land cover continuous field products at coarse spatial resolutions (e.g. 500m or 1km) from AVHRR (Advanced Very High Resolution Radiometer) and MODIS (DeFries, Townshend, and Hansen, 1999; DeFries et al., 1997; Hansen et al., 2005; Huang and Townshend, 2003). These studies used the finer scale Landsat data for global training and validation of the coarser AVHRR/MODIS results.

The training process began with the acquisition of high-resolution imagery from the NGA WARP website. We collected and reviewed over 1,800 high-resolution scenes. In the reviewing process, we looked for cloud free images that represent the various landscapes throughout the globe. These areas ranged from the city center of the United Kingdom to the rural country sides of the Appalachian Mountains. After applying an automated orthorectification algorithm based on Gao et al., (2009), each image was then divided into three smaller subsets at a size of 2500 X 2500 pixels. These subset images are then individually classified as either impervious or non-impervious using object-based image interpretation software called HSeg (Hierarchical Image Segmentation) Learn, which allowed accurate and efficient image classification through interaction with a trained interpreter. The derived high resolution classifications went through a QA process, where low quality classifications were excluded from further analysis. Finally, the high resolution classifications of satisfactory quality were overlaid on top of the 30m grids of the Landsat data to calculate percent imperviousness at the 30m resolution. Overall, a total of 27 million 30m reference pixels were derived based on about 2,000 high resolution images (Table 1). These reference pixels were distributed across all continents except Antarctica (Figure 2).

**Table 1.** Number of reference pixels derived from NGA data by continent.

<b>Continent</b>	<b>Number of 30m Training Pixels</b>
1) Europe	3,490,323
2) North America	5,099,105
3) Africa	3,603,191
4) South America	3,566,212
5) Australia	1,691,799
6) Asia	9,313,222
Totals	26,763,852

#### Step 4: Subpixel Urban Imperviousness Mapping

The reference imperviousness datasets derived above were used to train the Cubist regression tree algorithm to establish relationships between subpixel imperviousness and the spectral values of Landsat data. The established relationships were then applied to all GLS 2010 images to produce subpixel imperviousness products. Despite careful quality control in deriving the reference imperviousness datasets, they needed to be filtered before they could be used to train the regression tree algorithm. Due to temporal mismatches between some high resolution images and the GLS 2010 Landsat images, misclassification errors, cloud cover, and other image quality issues, some reference pixels with very high percent impervious values also had high greenness values, which were indicative of quality problems with these pixels. Based on a comprehensive assessment of the reference data, a set of greenness based threshold values were used to exclude reference pixels that likely had errors from being used in further analysis.

Because urban impervious surface is a relatively rare class at continental to global scales, the reference data were dominated by pixels with low or no impervious cover. Regression tree models trained using such highly unbalanced training data tend to over predict impervious values in the low end and under predict in the high end. In order to create a more balanced training dataset for each continent, we subsampled the reference data over that continent in the low end to reduce the amount of training samples with low or 0 impervious cover. In the high end, the reference samples with high impervious

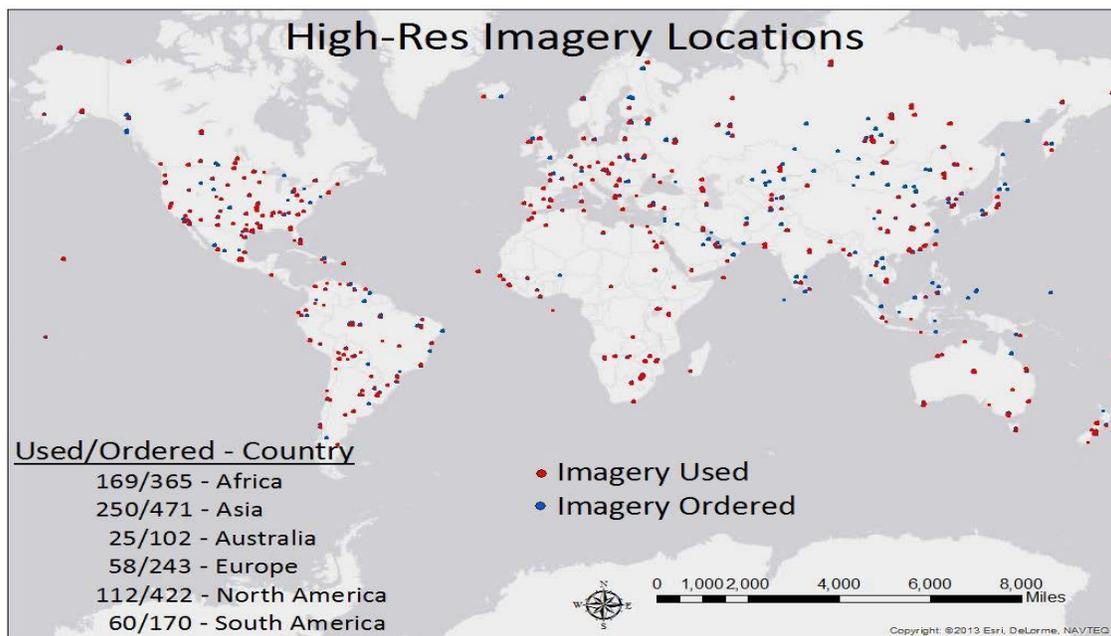


Figure 2. Distribution of high resolution images obtained from NGA and those used to develop impervious surface training data.

values from all continents were pooled together to form the high end of the training data for each continent.

Another issue to consider in developing a regression tree model was to determine the number of rules generated from the tree. Based on a series of experiments where we tested rule numbers from 500 to 10,000, we found that models with ~2,000 rules provided satisfactory results with acceptable processing speed. Models with too few rules were likely too simple and had many obvious errors, while those with too many rules were too slow when used to make predictions for each continent. As mentioned earlier, impervious surface models typically over predict imperviousness in non-vegetated non-urban areas like dry soil, sands, and rocks. To address this problem, the derived regression tree models were only applied to pixels within the HBASE urban extent mask.

### III. Data Set Description(s)

#### Data set description:

The GMIS product is delivered in the format of GeoTIFF raster files with two layers (bands): percent imperviousness (0-100) and standard error in percent imperviousness (0-100).

The percent imperviousness (0-100) estimate is given by the Cubist algorithm. Note that percent imperviousness for pixels classified as non-HBASE have been set to 200. The standard error (1 sigma) of percent imperviousness is estimated by the Cubist algorithm. Note that error for pixels classified as non-HBASE have been set to zero. The values are coded as follows:

Value	Label
0-100	Percent imperviousness, standard error
200	Non-HBASE
255	No data, clouds, shadows

#### Data set web page:

<http://sedac.ciesin.columbia.edu/data/set/ulandsat-gmis-v1>

#### Data set format:

The data are available in GeoTIFF format as downloadable zip files. Users can access:

- Explore View – <http://sedac.ciesin.columbia.edu/mapping/gmis-hbase/explore-view/>
- Download View – <http://sedac.ciesin.columbia.edu/mapping/gmis-hbase/download-view/>

### **Data set downloads:**

Users can create a downloadable zip file via:

- Explore View by selecting the “Download View” button and will have the option to select the type of region to download (e.g. country, shapefile, rectangle).
- Download View with the option to select the type of region to download (e.g. country, shapefile, rectangle).

The data are provided in individual data files according to Universal Transverse Mercator (UTM) zones at native 30m, 250m, and 1km spatial resolutions.

## **IV. How to Use the Data**

Users are able to download data by country and region. Users should be aware that there are missing data or no data regions because of cloud cover, and Scan Line Corrector (SLC) gaps.

## **V. Potential Use Cases**

The dataset is expected to have a rather broad spectrum of users, from those wishing to examine/study the fine details of urban land cover over the globe at full 30m resolution to global modelers trying to understand the climate/environmental impacts of man-made surfaces at continental to global scales. For example, the data are applicable to local modeling studies of urban impacts on the energy, water, and carbon cycles as well as analyses at the individual country level.

## **VI. Limitations**

The GMIS/HBASE datasets are derived directly from Landsat imagery from the Global Land Survey (GLS) 2010 dataset (Gutman et al. 2013). The quality of the GMIS/HBASE datasets is impacted by limitations of the selected GLS 2010 imagery and still contain residual cloud covered areas and areas where gaps caused by the Landsat 7 Scan Line Corrector (SLC) failure have not been successfully filled. The SLC gaps can cause visual features to appear in the data at different zoom levels. Areas where the gaps still persist or where cloud cover/cloud shadows were found have been coded as “NoData” (i.e. value = 255). Because of our approach using HBASE, it is also possible that small areas with impervious cover have been removed by HBASE. It is also possible that small areas of bare soil within cities have a non-zero impervious cover. Finally, coastlines and water bodies have been masked-in from the best available source which may also contain small errors and/or omit small features.

## **VII. Acknowledgments**

This study was funded by NASA's Land Cover and Land Use Change (LCLUC) Program (grants 09-LCLUC09-2-0136-1 & NNX11AH67G). Additional support for Panshi Wang and Chengquan Huang were provided by NOAA and USGS. The NGA high resolution satellite images were processed through the GMIS project and were originally obtained by NASA under the NGA's NextView license agreement. The authors thank James Zhan, Mike Taylor and Sike Li for their efforts in deriving the impervious surface percentage data that was used for product assessment. The authors would also like to thank the Global Land Cover Facility at University of Maryland for providing the GLS-2010 surface reflectance dataset.

Funding for the dissemination of this dataset was provided under the NASA contract NNG13HQ04C for the continued operation of the Socioeconomic Data and Applications Center (SEDAC), which is operated by the Center for International Earth Science Information Network (CIESIN) of Columbia University.

## **VIII. Disclaimer**

CIESIN follows procedures designed to ensure that data disseminated by CIESIN are of reasonable quality. If, despite these procedures, users encounter apparent errors or misstatements in the data, they should contact SEDAC User Services at [ciesin.info@ciesin.columbia.edu](mailto:ciesin.info@ciesin.columbia.edu). Neither CIESIN nor NASA verifies or guarantees the accuracy, reliability, or completeness of any data provided. CIESIN provides this data without warranty of any kind whatsoever, either expressed or implied. CIESIN shall not be liable for incidental, consequential, or special damages arising out of the use of any data provided by CIESIN.

## **IX. Use Constraints**

Users are free to use, copy, distribute, transmit, and adapt the work for commercial and non-commercial purposes, without restriction, as long as clear attribution of the source is provided.

## X. Recommended Citation(s)

### Data set(s):

Brown de Colstoun, E. C., C. Huang, P. Wang, J. C. Tilton, B. Tan, J. Phillips, S. Niemczura, P.-Y. Ling, and R. E. Wolfe. 2017. Global Man-made Impervious Surface (GMIS) Dataset from Landsat. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <https://doi.org/10.7927/H4P55KKF>. Accessed DAY MONTH YEAR.

## XI. Source Code

No source code is provided.

## XII. References

- Balk, D., F. Pozzi, G. Yetman, U. Deichmann, and A. Nelson. 2004. *The Distribution of People and the Dimension of Place: Methodologies to Improve the Global Estimation of Urban Extents*. Working Paper, CIESIN, Columbia University. Palisades, NY. Available from <http://sedac.ciesin.columbia.edu/downloads/docs/grump-v1/grump-v1-ur-paper-webdraft1.pdf>.
- Bonan, G. 1997. Effects of land use on the climate of the United States, *Climatic Change*, 37:449-486.
- CCSP (Climate Change Science Program). 2003. *Strategic Plan for the U.S. Climate Change Science Program*, Climate Change Science Program, Washington, D.C. pp. 211.
- Elvidge, C. D., K. E. Baugh, E. A. Kihn, H. W. Kroehl, and E. R. Davis. 1997. Mapping city lights with nighttime data from the DMSP Operational Linescan System. *Photogrammetric Engineering and Remote Sensing*, 63(6): 727–734.
- Feng, M., J. O. Sexton, C. Huang, J. G. Masek, E. F. Vermote, F. Gao, R. Narasimhan, S. Channan, R. E. Wolfe, and J. R. Townshend. 2013. Global surface reflectance products from Landsat: Assessment using coincident MODIS observations. *Remote Sensing of Environment*, 134, 276-293.
- Gao, F., J. Masek and R. E. Wolfe. 2009. An automated registration and orthorectification package for Landsat and Landsat-like data processing. *Journal of Applied Remote Sensing*, 3,033515, <https://doi.org/10.1117/1.3104620>.
- Goetz, S. J., D. Varlyguin, A. J. Smith, R. K. Wright, S. D. Prince and M. E. Mazzacatto, et al. 2004. *Application of multitemporal Landsat data to map and monitor land cover and land use change in the Chesapeake Bay Watershed*, in P. Smit and L. Bruzzone (eds.) *Analysis of Multi-temporal Remote Sensing Images*, World Scientific Publishers, Singapore, pp. 223-232.

- Gutman, G., R. Byrnes, J. G. Masek, S. Covington, C. Justice, S. Franks and R. Headley. 2008. Towards monitoring Land-cover and land-use changes at the global scale: The global land survey 2005. *Photogrammetric Engineering and Remote Sensing*, 92:6-10.
- Gutman, G., C. Huang, G. Chander, P. Noojipady, and J. G. Masek. 2013. Assessment of the NASA-USGS Global Land Survey (GLS) datasets. *Remote Sensing of Environment*, 134, 249-265.
- Homer, C., C. Huang, L. Yang, B. Wylie and M. Coan. 2004. Development of a 2001 National Landcover Database for the United States. *Photogrammetric Engineering and Remote Sensing*, 70:829-840.
- Imhoff, M. L., W. T. Lawrence, D. C. Stutzer, and C. D. Elvidge. 1997. A technique for using composite DMSP/OLS “city lights” satellite data to map urban areas. *Remote Sensing of Environment*, 61: 361-370.
- Imhoff, M. L., L. Bounoua, R. DeFries, W. T. Lawrence, D. Stutzer, C. J. Tucker, and T. Ricketts. 2004. The consequences of urban land transformation on net primary productivity in the United States, *Remote Sensing of Environment*, 89:434-443.
- Irons, J. R., and J. G. Masek. 2006. Requirements for a Landsat Data Continuity Mission. *Photogrammetric Engineering & Remote Sensing* 72 (10): 1102-1108.
- Jantz, P., S. Goetz and C. Jantz. 2005. Urbanization and the loss of resource lands in the Chesapeake Bay Watershed, *Environmental Management*, 36:808-825.
- Jin, M., J. M. Shepherd and M. D. King. 2005. Urban aerosols and their variations with clouds and rainfall: A case study for New York and Houston. *Journal of Geophysical Research*, 110, D102S20, <https://doi.org/10.1029/2004JD005081>.
- Kaufman, Y. J., et al. 1997. The MODIS 2.1-micron channel–correlation with visible reflectance for use in remote sensing of aerosol. *IEEE Transactions on Geoscience and Remote Sensing*, 35:1286-1296.
- Masek, J. G., et al. 2006. A landsat surface reflectance dataset for North America, 1990-2000. *IEEE Geoscience and Remote Sensing Letters*, 3:68-72.
- Quattrochi, D. A and M. K. Ridd. 1994. Measurements and analysis of thermal energy responses from discrete urban surfaces using remote sensing data. *International Journal of Remote Sensing*, 15: 1991-2002.
- Salvatore, M., F. Pozzi, E. Ataman, B. Huddleston and M. Bloise. 2005. *Mapping global urban and rural population distributions*. Environment and Natural Resources Series, No. 24 – FAO, Rome. 88p.
- Shepherd, J. M. 2005. A review of current investigations of urban-induced rainfall and recommendations for the future. *Earth Interactions*, 9:1-27.
- Schneider, A., M. A. Friedl, D. K. McIver, and C. E. Woodcock. 2003. Mapping urban areas by fusing multiple sources of coarse resolution remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 69:1377–1386.
- Schneider, A., M. A. Friedl, and D. Potere. 2009. A new map of global urban extent from MODIS satellite data. *Environmental Research Letters*, 4(4):044003.
- Small, C. 2005. A global analysis of urban reflectance. *International Journal of Remote Sensing*, 26:661-681.
- Tucker, C. J., D. M. Grant and J. D. Dykstra. 2004. NASA’s global orthorectified Landsat dataset. *Photogrammetric Engineering & Remote Sensing*, 70(3): 313-322.

United Nations. 2005. *World Urbanization Prospects: The 2005 Revision*. United Nations, Department of Economic and Social Affairs, Population Division (2006). Working Paper No. ESA/P/WP/200.

Vermote E. F., N. ElSaleous, C. O. Justice, Y. J. Kaufman, J. L. Privette, L. Remer, J. C. Roger, and D. Tanre. 1997. Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm and validation. *Journal Of Geophysical Research-Atmospheres*, 102 (D14): 17131-17141.

Yang, L., C. Huang, C. G. Homer, B. K. Wylie, and M. J. Coan. 2003. An approach for mapping large-area impervious surfaces: Synergistic use of Landsat 7 ETM+ and high spatial resolution imagery. *Canadian Journal of Remote Sensing*, 29, 230-240.

### **XIII. Documentation Copyright and License**

Copyright © 2017. The Trustees of Columbia University in the City of New York. This document is licensed under a Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). 

#### **Appendix 1. Revision History**

No revisions have been made to this dataset.

#### **Appendix 2. Contributing Authors & Documentation Revision History**

Revision Date	Contributors	Revisions
November 20, 2017	Eric Brown de Colstoun, Sri Vinay	This document is the 1 <sup>st</sup> instance of documentation.