1. TITLE

1.1 Data Set Identification

ISLSCP II HYDRO1k Elevation-derived Products.

1.2 Database Table Name(s)

Not applicable to this data set.

1.3File Name(s)

There are 2 data files (.zip) in 1 degree (1d) and half degree (hd) resolution.:

* **hydro1k_elevation_1deg.zip**
* **hydro1k_elevation_hdeg.zip**

When extrapolated, these files contain 24 data files pertaining to 4 elevation and elevation-based products (see section 3.2) and are named using the following naming convention:

hydro1k_*descriptor*_*statistic*_XX.asc

Here *descriptor* is a short name or abbreviation describing the file and can be elevation (elev), slope, aspect, or Compound Topographic Index (cti) (See section 3.2). *Statistic* is a short name or abbreviation for the statistical parameter provided in the particular file. These statistics can be maximum (max), minimum (min), range, median, mean, standard deviation (stdev), skewness (skew), and kurtosis (kurt). Note that not all statistics are provided for every parameter. The **XX** in the file name can be hd or 1d and stands for a spatial resolution of one degree or 0.5 degree in both latitude and longitude, respectively. As an example, the file that contains the global elevation range for each cell at 0.5 degree spatial resolution is named **hydro1k_elev_range_hd.asc**.

Two additional files at each resolution provide land/water boundary information for users: **hydro1k_changemap_xd.asc** and **hydro1k_landmask_xd.asc** (**xd** represents hd and 1d).

1.4 Revision Date of this Document

April 25, 2011

2. INVESTIGATOR(S)

2.1 Investigator(s) Name and Title

Kristine L. Verdin, Senior Scientist
Raytheon / EROS Data Center

2.2 Title of Investigation

HYDRO1k Elevation Derivative Database.

2.3 Contacts (For Data Production Information)

	Contact 1	Contact 2 (ISLSCP)
2.3.1 Name	Kristine L. Verdin	Eric Brown de Colstoun
2.3.2 Address	EROS Data Center	NASA/GSFC Code 614.4
City/St.	Sioux Falls, SD	Greenbelt, MD
Zip Code	57198	20771
Country	USA	USA
2.3.3 Tel. No.	(605) 594-6002	(301) 614-6597
Fax No.	(605) 594-6529	(301) 614-6695
2.3.4 E-mail	kverdin@usgs.gov	ericbdc@ltpmail.gsfc.nasa.gov

2.4 Data Set Citation

Verdin, K. L. 2011. ISLSCP II HYDRO1k Elevation-derived Products. In Hall, Forrest G., G. Collatz, B. Meeson, S. Los, E. Brown de Colstoun, and D. Landis (eds.). ISLSCP Initiative II Collection. Data set. Available on-line [http://daac.ornl.gov/] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. doi: 10.3334/ORNLDAAC/1007

2.5 Requested Form of Acknowledgment

Users of the International Satellite Land Surface Climatology (ISLSCP) Initiative II data collection are requested to cite the collection as a whole (Hall et al. 2006) as well as the individual data sets. Please cite the following publications when these data are used:

Hall, F.G., E. Brown de Colstoun, G. J. Collatz, D. Landis, P. Dirmeyer, A. Betts, G. Huffman, L. Bounoua, and B. Meeson, The ISLSCP Initiative II Global Data sets: Surface Boundary Conditions and Atmospheric Forcings for Land-Atmosphere Studies, *J. Geophys. Res.*, 111, doi:10.1029/2006JD007366, 2006.

Please acknowledge Kristine Verdin and the United States Geological Survey (USGS) EROS Data Center (EDC) when these data are used.

3. INTRODUCTION

3.1 Objective/Purpose

The purpose of this data development effort was to make use of high-resolution global elevation data to produce coarse scale elevation and elevation-based parameters at 1.0 and 0.5-degree spatial resolution to support a wide variety of global modeling activities through the International Satellite Land Surface Climatology Project (ISLSCP) Initiative II data collection. The intent was also to produce coarse scale elevation data sets with sufficient statistical information (up to fourth moment) to allow a good statistical description of the sub-cell distribution of any particular elevation parameter (i.e. elevation, slope and aspect). The database used in the data aggregation effort was the HYDRO1k data set with a native spatial resolution of 1 km, the highest resolution database of global coverage of standard elevation-based derivatives (slope, aspect, elevation, compound topographic index, etc.). HYDRO1k, developed at the U.S. Geological Survey's (USGS) EROS Data Center, is a geographic database providing comprehensive and consistent global coverage of topographically derived data sets. Developed from the USGS' recently released 30 arc-second digital elevation model (DEM) of the world (GTOPO30), HYDRO1k provides a standard suite of geo-referenced data sets (at a resolution of 1 km) that will be of value for all users who need to organize, evaluate, or process hydrologic information on a continental scale (see http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30)

3.2 Summary of Parameters

The aggregated elevation and elevation-based products provided for this ISLSCP Initiative II data collection are as follows (also see Section 8.2 for a full listing):

- Elevation:Elevation: For each 1 or 0.5-degree cell, the mean, median, minimum, maximum, range, standard deviation, skewness and kurtosis of the HYDROK1k 1-km elevation values are given.
- Slope: Mean, median, standard deviations, skewness and kurtosis of the 1 km values of slope within the coarser cells are given.
- Aspect: The aspect is also aggregated from the 1-km data and median, mean, standard deviation, skewness and kurtosis in aspect are provided here, expressed in degrees from North.
- Compound Topographic Index (CTI): The maximum, median, mean, standard deviation, skewness and kurtosis of the CTI are provided.

Two additional files at each resolution provide land/water boundary information for users.

3.3 Discussion

Development of the HYDRO1k database was made possible by the completion in 1996 of the 30 arc-second digital elevation model entitled GTOPO30 (Gesch et al, 1999). This data set, with its nominal cell size of 1 km, has been and will continue to be applied by many scientists and researchers to hydrologic and land form studies. Inevitably, these studies require processing of the DEM and extraction of derivative data layers. In order to reduce repetitive data processing

procedures and to provide a globally consistent data set, the HYDRO1k database aims to provide a standard suite of products, developed in a consistent fashion for the entire globe and make them available to the entire user community. The standard HYDRO1k products include, for each continent of the global with the exception of Antarctica and Greenland, six raster layers (elevation, flow direction, flow accumulation, slope, aspect, and compound topographic index) and two vector layers (drainage paths and watershed boundaries) at a 1-km resolution. More information on HYDRO1k is available at http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/README.

The HYDRO1k data sets (Version 1) have been developed on a continent-by-continent basis for all landmasses of the globe, with the exception of Antarctica, Greenland and, for data quality reasons, the continent of Australia. For this ISLSCP II data development effort, however, a preliminary Antarctica, Greenland and mainland Australia portion of the data set were produced. However, the data layers for these three landmasses have not been subjected to the same quality assessment as the other continents. The data layers for each continent were developed in a Lambert-Azimuthal Equal Area projection. The data aggregation was accomplished in the Lambert-Azimuthal framework and the resulting aggregated values transferred onto an equal-angle global grid.

We note that the calculation of the statistical moments for Aspect in this collection have been made with angles that can sometime cause misleading values for the mean aspect, in particular. As an example, if any particular cell contains many 1-km cells with slopes near 0 or 360 degrees, the mean aspect for this cell would be close to 180 degrees, which is in the opposite direction than the original data. At this revision of the ISLSCP Initiative II data collection we have simply not been able to fix this problem in a robust fashion. We caution the user and recommend looking at **all** the statistics for aspect before using these layers.

4. THEORY OF ALGORITHM/MEASUREMENTS

None given at this revision. More information on the HYDRO1k DEM can be found at: http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/README

5. EQUIPMENT

This data set was developed from the HYDRO1k database and did not use any remotely-sensed information.

5.1 Instrument Description

5.1.1 Platform (Satellite, Aircraft, Ground, Person)

Not applicable to this data set.

5.1.2 Mission Objectives

Not applicable to this data set.

5.1.3 Key Variables

Not applicable to this data set.

5.1.4 Principles of Operation

Not applicable to this data set.

5.1.5 Instrument Measurement Geometry

Not applicable to this data set.

5.1.6 Manufacturer of Instrument

Not applicable to this data set.

5.2 Calibration

5.2.1 Specifications

5.2.1.1 Tolerance

Not applicable to this data set.

5.2.2 Frequency of Calibration

Not applicable to this data set.

5.2.3 Other Calibration Information

Not applicable to this data set.

6. PROCEDURE

6.1 Data Acquisition Methods

All of the data provided here in the ISLSCP Initiative II data collection are based on the HYDRO1k DEM at 1-km spatial resolution which has been aggregated to spatial resolutions of 0.5 and 1.0 degrees. The methods described below relate to the creation of the HYDRO1k product at its original resolution.

As noted, the basis of all of the HYDRO1k data layers is a hydrologically conditioned DEM. This is, essentially, a quality controlled global DEM called GTOPO30 in a Lambert-Azimuthal Equal Area projection. The processing of the GTOPO30 DEM into HYDRO1k involved:

1. Projection of the DEM data into the continent-specific Lambert-Azimuthal Equal Area project.
2. Filling of the DEM. This involved evaluation of sink features to identify natural sink features, while removing spurious sinks.
3. Development of streamline and basin boundaries.
4. Comparison of derived streamlines and basin boundaries against existing digital data. In most cases, the Digital Chart of the World (DCW) drainage cover was used for comparison (Defense Mapping Agency, 1992; Danko, 1992). However, all available map sources were used. Comparison of the generated streamlines with mapped hydrography allows identification of essentially two types of errors in the DEM: Errors of omission or inclusion of natural sink features and errors in the DEM which prevent proper flow across its surface.

5. Editing of the DEM to correct errors found in Step 4. These editing procedures are repeated until the DEM is able to produce streamlines and basins that adequately match mapped hydrography.

Following generation of the hydrologically correct DEM, the final versions of the additional derivative data layers are produced. Along with the hydrologically correct DEM, the following five raster data layers are developed using standard GIS techniques. The user should refer to http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/README for more details on data processing of the HYDRO1k data sets.

Aspect: The aspect data set describes the direction of maximum rate of change in the elevations between each cell and its eight neighbors. It can essentially be thought of as the slope direction. It is measured in positive degrees from 0 to 360, measured clockwise from north. Aspects of cells of zero slope (flat areas) are assigned values of -1.

Flow Directions: The flow direction data layer defines the direction of flow from each cell in the DEM to its steepest down-slope neighbor. Flow directions follow the convention adopted by ARC/INFO's flow direction implementation and are expressed as integers between 1 and 255. For a cell with simple defined direction of flow to one of its eight neighbors, the convention is as expressed in the following table:

32	**64**	**128**
16		**1**
8	**4**	**2**

For example, if the direction of steepest drop was to the left of the current cell, the flow direction would be recorded as 16. If a cell has the same change in z value in multiple directions and that cell is part of a sink, the flow direction is referred to as undefined. In such cases, the value for that cell in the output flow-direction grid will be the sum of those directions. For example, if the change in z value is the same both to the right (flow direction = 1) and down (flow direction = 4), the flow direction for that cell is 1 + 4 = 5.

Flow Accumulations: The flow accumulation data layer defines the number of upstream pixels flowing into each cell. Since the cell size of the HYDRO1k data set is 1 km, the flow accumulation value translates directly into drainage areas in square kilometers. Values range from 0 at topographic highs to very large numbers (on the order of millions of cells) at the mouths of large rivers.

Slope: The slope data layer describes the maximum change in the elevations between each cell and its eight neighbors. The slope is expressed in degrees of slope between 0 and 90.

Compound Topographic Index: The Compound Topographic Index (CTI), commonly referred to as the Wetness Index, is a function of the upstream contributing area and the slope. The implementation used in the HYDRO1k data set is based on Moore et al (1991). The CTI is calculated using the flow accumulation (FA) layer along with the slope as:

$$CTI = \ln((FA + 1)/ \tan(slope)) \quad (1)$$

In areas of no slope, a CTI value is obtained by substituting a slope of 0.001. This is the smallest slope obtainable from a 1000-m data set with a 1-m vertical resolution. For this ISLSCP II data set, aggregated layers for elevation, slope, aspect and CTI are provided.

6.2 Spatial Characteristics

6.2.1 Spatial Coverage

The data sets provided here are global in spatial coverage. The data in each file are ordered from North to South and from West to East beginning at 180 degrees W and 90 degrees N. Point (1,1) represents the gridcell centered at 89.75 degrees N and 179.75 degrees W for the 0.5 degree data, and 89.5 degrees N and 179.5 degrees W for the 1.0 degree data.

6.2.2 Spatial Resolution

The data are given in two regular equal-angle lat/long Earth grids that have spatial resolutions of 0.5 degree by of 0.5 degree and 1.0 degree by 1.0 degree in both latitude and longitude.

6.3 Temporal Characteristics

6.3.1 Temporal Coverage

All data layers are time invariant.

6.3.2 Temporal Resolution

All data layers are time invariant.

7. OBSERVATIONS

7.1 Field Notes

Not applicable to this data set.

8. DATA DESCRIPTION

8.1 Table Definition with Comments

Not applicable to this data set.

8.2 Type of Data

8.2.1 Parameter/ Variable Name	8.2.2 Parameter/ Variable Description	8.2.3 Data Range*	8.2.4 Units of Measurement	8.2.5 Data Source
Landmask	Original land/water mask used in the aggregation of HYDRO1k: 0 = Ocean 1 = Land	0 to1	See 8.2.2	

8.2.1 Parameter/ Variable Name	8.2.2 Parameter/ Variable Description	8.2.3 Data Range*	8.2.4 Units of Measurement	8.2.5 Data Source
Point Changed	Differences between the ISLSCP II land/water mask and the original data: -1 = ISLSCP II mask is water and original data is land (data removed) 0 = Data sets agree over land or water (data unchanged) 1 = ISLSCP II mask is land or water and original data is missing (fill value used).	-1 to 1	See 8.2.2	Original mask and ISLSCP II land/water mask
Elev_min	Minimum elevation value within each cell	-405 to 5149	Meters	HYDRO1k
Elev_max	Maximum elevation value within each cell	-22 to 8752		
Elev_range	Range of elevation values within each cell	0 to 7902		
Elev_median	Median elevation value within each cell	-86 to 5747		
Elev_mean	Mean elevation value within each cell	-74.87 to 5724.91		
Elev_stdev	Standard deviation of the elevation values within each cell	0.00 to 1525.66		
Elev_skew	Skewness in elevation values within each cell	-13.90 to 55.69 Undefined=-99		
Elev_kurt	Kurtosis in elevation values within each cell	1 to 3102 Undefined=-99		
Slope_median	Median slope within each cell	0.00 to 21.16	Degrees	
Slope_mean	Mean slope within each cell	0.00 to 21.04		
Slope_stdev	Standard deviation of the slopes within each cell	0.00 to 14.78		
Slope_skew	Skewness of the slopes within each cell	-2.82 to 55.69 Undefined=-99		
Slope_kurt	Kurtosis of the slopes within each cell	1.01 to 3102 Undefined=-99		
Aspect_median	Median Aspect of each cell.	0.00 to 353.03 Flat areas=-1	Degrees from North	
Aspect_mean	Mean Aspect of each cell.	0.005 to 352.21 Flat areas=-1		

8.2.1 Parameter/ Variable Name	8.2.2 Parameter/ Variable Description	8.2.3 Data Range*	8.2.4 Units of Measurement	8.2.5 Data Source
Aspect_stdev	Standard deviation in aspect for each cell	0.0002 to 177.61 Flat areas=-1		
Aspect_skew	Skewness in aspect for each cell	-1.00 to 55.69 Undefined=-99		
Aspect_kurt	Kurtosis in aspect for each cell	1.00 to 3102 Undefined=-99		
Cti_max	Maximum CTI value within each cell	0.00 to 24.23	Unitless	
Cti_median	Median CTI within each cell	0.00 to 13.44		
Cti_mean	Mean CTI within each cell	0.00 to 13.54		
Cti_stdev	Standard deviation in CTI for each cell	0.000 to 5.405		
Cti_skew	Skewness in CTI for each cell	-25.65 to 55.69 Undefined=-99		
Cti_kurt	Kurtosis in CTI for each cell	1.04 to 3102 Undefined=-99		

**** Data ranges are given for the 1/2 degree data and may be different for the 1 degree resolution data.***

8.3 Sample Data Record

Not applicable to this data set.

8.4 Data Format

All of the files in the ISLSCP Initiative II data collection are in the ASCII grid format. The file format consists of numerical fields of varying length, which are delimited by a single space and arranged in columns and rows. The 0.5 degree files in this data set all contain 720 columns by 360 rows. The 1.0 degree files all contain 360 columns by 180 rows. Flat areas (i.e. slope=0) are assigned the value of –1 in the median, mean and standard deviation in aspect files. Areas with standard deviations equal to 0 in any parameter are undefined in both skewness and kurtosis and are assigned the value –99.

All files are gridded to a common equal-angle lat/long grid, where the coordinates of the upper left corner of the files are located at 180 degrees W, 90 degrees N and the lower right corner coordinates are located at 180 degrees E, 90 degrees S. Data in the map files are ordered from North to South and from West to East beginning at 180 degrees West and 90 degrees North.

8.5 Related Data Sets

ISLSCP II project information and data sets can also be obtained from the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) http://daac.ornl.gov/ISLSCP_II/islscpii.shtml .

9. DATA MANIPULATIONS

9.1 Formulas

9.1.1 Derivation Techniques/Algorithm

See Section 6.1 and http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/README for more information.

9.2 Data Processing Sequence

9.2.1 Processing Steps and Data Sets

Transformation of the HYDRO1k layers, which are developed in the Lambert-Azimuthal Equal Area projection, into a geographic framework was accomplished through the following data processing sequence. All data processing was done with ARC/INFO and ARCVIEW.

- The HYDRO1k data layers were organized by continent. The landmasses of Greenland and Australia (which are not included in Version 1 of HYDRO1k) were processed into HYDRO1k data layers in the Lambert-Azimuthal Equal Area projection. Similarly, Antarctica, which also was not included in Version 1 of HYDRO1k, was processed in the polar stereographic projection.
- The spatial extent of the HYDRO1k data for each continent was projected into geographics – defining the extent of the non-ocean pixels within the geographic framework – for each continent.
- A template was created, as a vector coverage, corresponding to the desired full global coverage (180 E to 180 W, 90 S to 90 N) with a 0.5 degree spacing.
- For each continent, the global template corresponding to HYDRO1k pixels was selected. This template, for each continent, was projected into the continental Lambert-Azimuthal Equal Area projection.
- This projected template was used to define the zones for preparation of the summary statistics. Summary statistics were calculated using ARCVIEW zonal stats functions.

The summary statistics were transferred back onto the global template (in geographic projection) and the resulting aggregated grids were developed. Mean, standard deviation, skewness and kurtosis were calculated using standard statistical formulae.

9.2.2 Processing Changes

None.

9.2.3 Additional Processing by the ISLSCP II Staff

The original data files submitted to the ISLSCP II staff were re-processed into the ASCII format from their original ARC/INFO GRIDASCII format. Based on the water values given in the land/water mask submitted with the data all areas with missing data flags (-9999) were recoded to 0 in the elevation and slope files. Areas with standard deviations equal to 0 were recoded to -99 (undefined) in all skewness and kurtosis files. Values with median slope equal to 0 were set to -1 in the median aspect files. Cells where

the mean slope standard deviation in slope were both equal to 0 were recoded to -1 (flat areas) in the mean and standard deviation in aspect files.

This data set was not made consistent with the ISLSCP II land/water mask because the original data had elevation information where the ISLSCP II mask showed land (i.e. 50% or greater of the cell is land). The ISLSCP II staff compared the original land/water mask with the ISLSCP II mask and produced a change map showing the differences between the two masks. Users should consult this map for questions on land/water boundaries.

9.3 Calculations

9.3.1 Special Corrections/Adjustments

See http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/README for more information.

9.4 Graphs and Plots

See http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30

10. ERRORS

10.1 Sources of Error

The exclusion of the Australian portion of the globe in the development of the Version 1 of the HYDRO1k data layers was done based on data quality concerns. Inclusion of the Australian continent may have introduced errors, since this continent was not subjected to the same quality assessment as the rest of HYDRO1k. Similarly, the Greenland and Antarctic portions of the data set have not undergone quality assessment.

Generalization of the CTI layers through statistics summary degrades the value of the index. The CTI is based on flow accumulation values, along with slope, and does not lend itself to generalization. Recalculation of the CTI based on a generalized flow accumulation value (STN-30p) and a generalized slope (from HYDRO1k at 0.5 degrees) should be considered.

The aggregation of the aspect fields is problematic because mean aspect can include slopes with slopes near 0 or 360 degrees, which can yield a mean slope of close to 180 degrees, which is in the opposite direction than the original data. At this revision of the ISLSCP Initiative II data collection we have simply not been able to address this issue.

10.2 Quality Assessment

10.2.1 Data Validation by Source

See http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/README for more information.

10.2.2 Confidence Level/Accuracy Judgment

See http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/README for more information.

10.2.3 Measurement Error for Parameters and Variables

See http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/README for more information.

10.2.4 Additional Quality Assessment Applied

None.

11. NOTES

11.1 Known Problems with the Data

We have noted that the calculation of the statistical moments for Aspect in this collection have been made with angles that can sometime cause misleading values for the mean aspect, in particular. As an example, if any particular cell contains many 1-km cells with slopes near 0 or 360 degrees, the mean aspect for this cell would be close to 180 degrees, which is in the opposite direction than the original data. At this revision of the ISLSCP Initiative II data collection we have simply not been able to fix this problem in a robust fashion. We caution the user and recommend looking at ALL the statistics for aspect before using these layers.

11.2 Usage Guidance

See above.

11.3 Other Relevant Information

None.

12. REFERENCES

12.1 Satellite/Instrument/Data Processing Documentation

http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/README

12.2 Journal Articles and Study Reports

Danko, D.M., 1992. The digital chart of the world. *GeoInfo Systems*, 2:29-36.

Defense Mapping Agency, 1992, *Development of the Digital Chart of the World*: Washington, D.C., U.S. Government Printing Office.

Gesch, D.B., Verdin, K.L., Greenlee, S.K., 1999, New land surface digital elevation model covers the Earth. *Eos. Transactions. American Geophysical Union* 80 (6), pp. 69-70

Moore, I.D., R.B. Grayson and A.R. Ladson, 1991, *Digital Terrain Modelling: A Review of Hydrological, Geomorphological and Biological Applications*. In: Hydrological Processes An International Journal, January - March, 1991, pp. 3 - 30.

Verdin, K.L. and J.P. Verdin, 1999, A topological system for delineation and codification of the Earth's river basins, *Journal of Hydrology*, vol. 218, nos. 1-2, pp. 1-12

13. DATA ACCESS

13.1 Contacts for Archive/Data Access Information

The ISLSCP Initiative II data are available are archived and distributed through the Oak Ridge National Laboratory (ORNL) DAAC for Biogeochemical Dynamics at http://daac.ornl.gov.

13.2 Contacts for Archive

E-mail: uso@daac.ornl.gov
Telephone: +1 (865) 241-3952

13.3 Archive/Status/Plans

The ISLSCP Initiative II data are archived at the ORNL DAAC. There are no plans to update these data.

14. GLOSSARY OF ACRONYMS

CTI	Compound Topographic Index
DAAC	Distributed Active Archive Center
DCW	Digital Chart of the World
DEM	Digital Elevation Model
EDC	EROS Data Center
ISLSCP	International Satellite Land Surface Climatology Project
GSFC	Goddard Space Flight Center (NASA)
NASA	National Aeronautics and Space Administration
ORNL	Oak Ridge National Laboratory
STN	Simulated Topological Network
USGS	U.S. Geological Survey