1. TITLE

1.1 Data Set Identification

ISLSCP II River Routing Data (STN-30p)

1.2 Database Table Name(s)

Not applicable to this data set.

1.3 File Name(s)

The Simulated Topological Network (STN) data set for the International Satellite Land Surface Climatology Project (ISLSCP) Initiative II data collection is made up of two data files at spatial resolutions of 0.5 and 1.0 degree in both latitude and longitude, denoted as hd and 1d, respectively, in each file name: **river_routing_stn_1deg.zip**, and **river_routing_stn_hdeg.zip**. When expanded, each data file contains a number of global, gridded data layers and two associated attribute tables named as follows:

stn_basin_id_XX.asc: Basin grid, which assigns unique identifiers to groups of gridcells forming individual river basins. XX can be hd or 1d, depending on the spatial resolution of the data set.

stn_flow_accum_XX.asc: Flow accumulation grid or upstream area grid (km^2), which provides the catchment area estimate for each gridcell according to the STN-30p network.

stn_flow_direction_XX.asc: Flow direction grid, which represents the horizontal connectivity of the continental land mass.

stn_dist2mouth_XX.asc: Distance [km] to the outlet of river basins.

stn_mainstem_length_XX.asc: Mainstem length [km] grid.

stn_stream_order_XX.asc: Strahler stream order grid.

stn_elevation_hd.asc: Adjusted elevation (m) grid, which combines HYDRO1k aggregated elevation at 30-minute resolution with STN-30p, where the inconsistencies between the elevation and the flow direction data sets (i.e. increasing elevation along downstream flow path) were eliminated. ****NOTE: this data layer is only available at 0.5 degree resolution.

stn_slope_hd.asc: Slope (degree) along the STN-30p grid, provides slope information for each gridcell along the cell's flow direction. ****NOTE: this data layer is only available at 0.5 degree resolution.

stn_basin_attribute_XX.dat (table): Basin attribute table contains attribute information such as basin name, catchment area, mainstem length, etc. for each individual basin. See below for table descriptions.

stn_cell_attribute_XX.dat (table): Cell attribute table contains attribute information for each land cell. See section 8.2 for table descriptions.

1.4 Revision Date of this Document

March 30, 2011

2. INVESTIGATOR(S)

2.1 Investigator(s) Name and Title

Dr. Charles J. Vörösmarty, associate professor
Head of Water Systems Analysis Group
Complex Systems Research Center,
Institute for the Study of Earth, Oceans, and Space
University of New Hampshire

2.2 Title of Investigation

The Simulated Topological Network (STN).

2.3 Contacts (For Data Production Information)

	Contact 1	**Contact 2**
2.3.1 Name	Dr. Charles J. Vörösmarty	Dr. Balázs M. Fekete
2.3.2 Address	CUNY Energy Institute The City University of New York 140th St and Convent Ave	CUNY Energy Institute The City University of New York 140th St and Convent Ave
City/St.	New York, NY	New York, NY
Zip Code	10031	10031
Country	USA	USA
2.3.3 Tel. No.	(212)-650-7042	(212)-650-7042

Fax No.		
2.3.4 E-mail	cvorosmarty@ccny.cuny.edu	bfekete@ccny.cuny.edu

2.4 Data Set Citation

Vörösmarty, C. J., and B. Fekete. 2011. ISLSCP II River Routing Data (STN-30p). 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/1005

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.

Vörösmarty C. J., B. M. Fekete, M. Meybeck and R.B. Lammers: Global Systems of Rivers: Its role in organizing continental land mass and defining land-to-ocean linkages, *Global Biochemical Cycles*, 14(2): 599-621, 2000.

Vörösmarty, C. J., B. M. Fekete, M. Meybeck and R. B. Lammers: Geomorphometric attributes of the global system of rivers at 30-minute spatial resolution, *Journal of Hydrology*, 237: 17-39, 2000

3. INTRODUCTION

3.1 Objective/Purpose

The proper representation of the global river system is essential for modeling horizontal water transfers. Gridded river networks derived from digital elevation models suffer from various errors, which are particularly severe at coarse resolution. The goal in developing the Simulated Topological Network, or STN-30p, was to provide the large-scale hydrological modeling community an accurate representation of the global river system. STN-30p was developed before the HYDRO1k Digital Elevation Model (DEM) become available, therefore its river network topology and all the derived information from the network such as basin delineation, upstream area, distance to oceans, etc. are independent from HYDRO1k (which is part of the present **ISLSCP II** initiative), however a corrected DEM is also provided as part of the data set, which applied a correction to the aggregated HYDRO1k elevation at 30' resolution to eliminate the inconsistencies between the network and the DEM. Per request of the **ISLSCP II** scientific advisory board a 1.0 degree version of the STN-30p network was developed specifically for the **ISLSCP II** data archive using network re-gridding algorithms.

3.2 Summary of Parameters

Two data files at spatial resolutions of 0.5 and 1.0 degrees are provided based on the original STN30p (V5.12) network from the University of New Hampshire (UNH). The core components of the data set are:

- Flow direction grid, which represents the horizontal connectivity of the continental landmass.
- Basin grid, which assigns unique identifier groups of gridcells forming individual river basins.
- Flow accumulation grid or upstream area grid, which provides the catchment area estimate for each gridcell according to the STN-30p network.
- Stream order grid, which assigns stream order to every network gridcell according to the Strahler ordering scheme.
- Mainstem length upstream. The mainstem is defined in STN-30p as the path following the maximum catchment area. This definition is applied consistently, therefore the mainstem according to STN-30p might deviate from the usual definition (e.g. the mainstem of the Mississippi, follows the Missouri). This distance takes into account both the Earth curvature and the orientation of individual gridcells along the path to the basin outlet, but does not take into account the sinuosity of the real river courses, therefore it tends to be 10-20 % shorter than the reported river length (Vörösmarty et al. 2000, Fekete et. al 2001).
- Distance to ocean grid, provides the distance to the outlet of each basin along the river network. Similarly to the mainstem length, this distance is calculated in projected space and takes into account the Earth curvature and the cell orientation but does not include sinuosity.
- Adjusted elevation grid, which combines HYDRO1k aggregated elevation at 30-minute resolution with STN-30p, where the inconsistencies between the elevation and the flow direction data sets (i.e. increasing elevation along downstream flow path) were eliminated.
- Slope along the STN-30p grid, provides slope information for each gridcell along the cell's flow direction.
- Basin attribute table contains attribute information such as, basin name, catchment area, mainstem length, receiving ocean, etc… for each individual basin in STN-30p.
- Cell attribute table contains attribute information for each land cell. This information includes basin identifier which links the gridcells to the corresponding basin record entry in the basin table, number of gridcells in the tributary, catchment area upstream, cell area (which takes into account the Earth curvature), mainstem length upstream, number of gridcells downstream. The cell table is sorted by basin ID and catchment area upstream, therefore the first record entry in the cell table represents the mouth cell of the Amazon basin. Records 1 to 1916 contain the gridcells representing the Amazon basin. The 1917th record in the cell table is the mouth cell of the Nile, which is the second largest river system according to STN-30p. The ordered cell table makes it ideal in water transport simulations since the position in the table defines the gridcell's place in the hierarchy in the river systems.

3.3 Discussion

STN-30p represents the potential connectivity of the continental land mass by assigning one of eight (E, SE, S, SW, W, NW, N, NE) possible flow directions to each continental gridcell

(Jenson 1988, Band 1993). The potentiality of STN-30p reflects the fact that flow direction is assigned to every land cell regardless of the existence of actively flowing rivers. STN-30p can be viewed as a river network, which would exist if sufficient surface runoff was available to form river channels everywhere.

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.

The STN-30p flow direction grid was derived from a Digital Elevation Model ETOPO5 (Edwards, 1989) and edited manually by considering various digital maps such as ARC/World (ESRI, 1992a), the Digital Chart of the World (ESRI, 1992b) and discharge gauges. The estimated catchment area at discharge gauging stations was compared to reported values from discharge gauging station data sets (Vörösmarty, 1996; UNESCO-IHP, 1969-84; Lammers 2001).

The data set contains not only the core directionality grid, but several network derived data layers are also included such as basin grid and flow accumulation grids, adjusted elevation and slopes along the flow direction. The adjusted elevation was based on HYDRO1k, which was aggregated to the STN-30p resolution. The inconsistencies between the elevation and the STN-30p were eliminated by lowering the elevation in those gridcells where the elevation increased downstream along the simulated network. The elevation was further adjusted by an 11 gridcells running mean average filter along the river network (considering 5 gridcells upstream and 5 gridcells downstream along the river channel). Finally, the 1.0 degree product was derived from STN-30p applying flow aggregation algorithm developed at UNH (Fekete et al., 2001).

4. THEORY OF ALGORITHM/MEASUREMENTS

Gridded networks are regarded as an efficient mechanism to describe the linkage between the landmass and river systems and ultimately the oceans. Such a network defines the connectivity of the individual gridcells by specifying the flow direction from one cell to another. STN-30p allows tree-like connectivity (i.e. individual gridcells could receive from up to eight of its neighbors and have only one outlet). This tree like structure defines an unambiguous linkage between any gridcell and all downstream gridcells since no loops or multiple routes between two gridcells are allowed.

The tree structure is sufficient for most parts of the global river systems, but puts a limitation on its ability to represent true flow pathways, when the rivers split along their course (such as river deltas). This limitation is less severe at the relatively coarse resolution of STN-30p since most river deltas at the 30' resolution are smaller than the size of a gridcell.

5. EQUIPMENT

The STN data set is derived from vector data and not from actual measurements.

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

Not available at this revision.

6.2 Spatial Characteristics

6.2.1 Spatial Coverage

The original STN-30p data set covers the continental landmasses between 55 degrees S and 83 degrees N excluding Greenland and the Arctic Archipelago. The data sets provided here are global but also exclude Antarctica. 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 89.75 degrees N and 179.75 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 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

STN-30p represents the static geomorphology of the continental landmass therefore it can be viewed as one snapshot in time of the contemporary river systems, but it has flow directions for inactive (non-flowing) regions under the present climate. This allows the application of STN-30p beyond contemporary climate since it provides connectivity for all the continental land cells, however this potential is limited since the network is static and does not allow the reconfiguration of the connectivity as it might occur due under different climate regime (e.g. past connection between the Aral, the Caspian and the Black seas).

6.3.2 Temporal Resolution

STN-30p, similarly to elevation, represents the geomorphology of the Earth. However, while the geomorphology might change in geological time scales, it is typically treated as static or fixed information.

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
(1) Global Grid Files (*.asc)				
Flow direction	Flow direction grid, which represents the horizontal connectivity of the continental land mass	1 – 128 Oceans = -99 No Data over Land = -88	ARC/INFO flow directions (See Section 8.4)	STN-30p
Basin ID	Basin grid, which assigns unique identifiers to groups of gridcells forming individual river basins. This ID can be used to mach the corresponding basin attribute record in basin table below.	1 – 6152 Oceans = -99 No Data over Land = -88	N/A	
Stream Order	Stream order of every network gridcell according to the Strahler ordering scheme.	1 – 6 Oceans = -99 No Data over Land = -88	Strahler Scheme	
Flow Accumulation	The upstream catchment area for each land cell.	390 – 5,853,804 Oceans = -99 No Data over Land = -88	km^2	
Mainstem length	Mainstem length upstream following the maximum catchment area.	7 – 5909 Oceans = -99 No Data over Land = -88	km	
Distance to mouth	Distance to river basin outlet.	7 – 5909 Oceans = -99 No Data over Land = -88	km	
Adjusted Elevation*	Adjusted HYDRO1k elevation consistent with STN-30p.	-72 to 5719	m	HYDRO1k DEM
Slope*	Slope along STN-30p.	0.001 – 79.33 Oceans = -99 No Data over Land = -88	m/km	HYDRO1k DEM
(2) Basin Attributes Tables (stn_basin_attribute_xx.dat)				
ID	Basin ID number.	1 - 6152	N/A	STN-30p
NAME	Basin Name (GHAASBasin or real basin name from GEMS/GLORI).	N/A	N/A	

BASINORDER	Strahler stream order at mouth (1-6).	1 - 6	Strahler Scheme	
COLOR	Optimized color (7 - 11) to display basins.	7 - 11	N/A	
BASINLENGTH	Mainstem length.	7 – 5909	km	
BASINAREA	Basin area.	390 – 5,853,804	km^2	
CONTINENTNAME	Continent name.	N/A	N/A	
SEANAME	Name of the recipient sea.	N/A	N/A	
OCEANNAME	Name of the recipient ocean.	N/A	N/A	
(3) Cell Attribute Tables (stn_cell_attribute_xx.dat)				
ID	Cell ID number.	1 - 59132	N/A	STN-30p
NAME	Basin name.	N/A	N/A	
ORDER	Strahler stream order.	1 - 6	Strahler Scheme	
BASINID	Integer basin identifier.	1 - 6152	N/A	
BASINCELLS	Number of upstream cells.	1 - 1916	N/A/	
TRAVEL	Number of cells downstream.	0 - 101	N/A	
CELLAREA	Cell area.	390 - 3091	km^2	
SUBBASINLENGTH	Mainstem length upstream.	7 – 5909	km	
SUBBASINAREA	Upstream area.	390 – 5,853,804	km^2	
DISTTOMOUTH	Distance to river mouth defined as the confluence with an equal or higher order	7 – 5909	km	

	stream.			
DISTTOOCEAN	Distance to the outlet of river basins.	7 – 5909	km	
CELLLENGTH	Length of 30 minute or 1.0 degreeriver segments.	7 – 78.629	km	

*Note: The adjusted elevation and slope grids are only available at 1/2 degree resolution. Ranges are given for the 1/2 degree resolution files and may be different for 1.0 degreeresolution files.

8.3 Sample Data Record

The first few records for the basin attribute table called **stn_basin_attributes_hd.dat** are shown below:

```
"ID"   "Name"      "BasinOrder"      "Color"     "BasinLength"     "BasinArea"
       "ContinentName"   "SeaName"   "OceanName" "Esri"      "Hydro"
       "ClimateZones"    "MouthXCoord"     "MouthYCoord"
1      "Amazon"    6     7     4327.070312 5853804.000000    "South America"
       "Atlantic Ocean"  "Atlantic Ocean"  "South America"   "South America"
       "Tropics"   -51.750000  -1.250000
2      "Nile"      5     7     5908.688477 3826122.000000    "Africa"
       "Mediterranean Sea"     "Mediterranean+Black Sea"     "Africa"
       "Africa"    "Temperate" 31.250000   31.250000
3      "Zaire"     5     8     4338.881836 3698802.750000    "Africa"
       "Atlantic Ocean"  "Atlantic Ocean"  "Africa"    "Africa"    "Tropics"
       12.750000   -5.750000
```

The first few records for the cell attribute table called **stn_cell_attributes_hd.dat** are shown below:

```
"ID"   "Name"      "ToCell"    "FromCell"  "Order"     "BasinID"   "BasinCells"
       "Travel"    "CellArea"  "SubbasinLength"  "SubbasinArea"
       "DistToMouth"     "DistToOcean"     "CellLength"      "CellXCoord"
       "CellYCoord"
1      "GHAASCell:1"     128   88    6     1     1916  0     3090.547119
       4327.070312 5853804.000000    78.623306   78.623306   78.623306   -
51.750000   -1.250000
2      "GHAASCell:2"     128   20    6     1     1895  1     3089.841064
       4248.446777 5788899.500000    157.239128  157.239128  78.615822   -
52.250000   -1.750000
3      "GHAASCell:3"     1     116   6     1     1731  2     3089.841064
       4169.831055 5288402.000000    212.812592  212.812592  55.573463   -
52.750000   -1.750000
```

8.4 Data Format

All of the files in the ISLSCP Initiative II data collection are in the ASCII, or text 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. Values in the files for flow accumulation, distance to river mouth, mainstem length, slope and elevation are written as real numbers while files for basin Ids, flow direction, and stream order are written as integers. Missing values over water are assigned the value of -99. Missing values over land are assigned the value of -88.

The flow direction grid is given as the standard ARC/INFO representation of the 8 flow direction grid using the standard encoding scheme shown below:

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

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

Users of the ISLSCP Initiative II data collection will find the following related data sets in this collection: Global Elevation and Elevation-Based Derivatives from HYDRO1k and the UNH/Global Runoff Data Center (GRDC) Composite Runoff Fields Time Series and River Discharge data. Users should check the following web site for other hydrological data sets available from the Water Systems Analysis Group at UNH: http://www.watsys.unh.edu/

9. DATA MANIPULATIONS

STN-30p was derived from the ETOPO5 DEM (Edwards, 1989) applying standard maximum gradient search (Jenson, 1988, Band, 1993). The resulting initial network was manually edited considering digital river maps such as ARC/World (ESRI, 1992a), and the Digital Chart of the World (ESRI, 1992b). The upstream catchment area estimates from STN-30p were compared to reported catchment area values at discharge gauging stations from various sources (Vörösmarty, 1996; Lammers, 2001) and further manual editing was applied as necessary to adjust STN-30p so the STN-30p derived catchment area matches the reported values as closely as possible. This step not only helped the reduce the potential errors in STN-30p but also was important to identify potential errors in the reported catchment areas according to the different discharge gauging records.

Once the flow direction grid was completed individual basins were identified by consulting Global River Discharge Database (GLORI) (Meybeck, 1995). As a result of this step, 397 river basins were named in STN-30p according to the GLORI database. Various network characteristics (catchment area, mainstem length, distance to basin outlet, Strahler stream order, etc.) were calculated at the gridcells and the basin level.

Because STN-30p has been heavily edited to represent the actual river networks as closely as possible, it is no longer entirely consistent with any digital elevation model (including ETOPO5 from which STN-30p was derived). An adjusted DEM was developed based on the aggregated HYDRO1k DEM at 30' resolution (which is part of the ISLSCP II data collection). The DEM adjustment performed a downstream search along every possible river routes within STN-30p and lowered the elevation as necessary to maintain a minimum slope of 0.001 m/km.

The second step of the DEM adjustment was to apply a running average of the elevation along all possible routes in STN-30p. This step eliminated some of the unrealistic rapid slope changes along STN-30p river pathways (see Figures 1 and 2 in Section 9.4).

The adjusted elevation grid is consistent with STN-30p, but applying a standard maximum gradient search on the elevation grid would not necessary result in the same gridded network as STN-30p. In other words, the elevation adjustment does not ensure that the direction of the maximum gradient is identical to the flow direction given by STN-30p. Since most GIS packages identify slope as the maximum gradient slope, applying STN-30p in a flow routing scheme would potentially require slope along STN-30p flow direction, therefore a separate slope coverage is provided, which is calculated along STN-30p directions rather than representing the maximum gradient in each cell.

A 1.0 degree resolution version of STN-30p was developed specifically for the ISLSCP Initiative II collection. This product was derived from STN-30p using a network re-gridding algorithm developed at UNH (Fekete et al., 2001). This algorithm is based on upstream catchment area that can be aggregated and used to define flow directions similar to DEMs (except that the flow direction is defined as the maximum increasing gradient). The basin names of named basins from the original STN-30p network were transferred via basin centroids (a point coverage of the gravity centers of individual basins derived from STN-30p).

9.1 Formulas

9.1.1 Derivation Techniques/Algorithms

STN-30p was derived from ETOPO5. The original 5' DEM was aggregated to 30' resolution (i.e. every 6x6 gridcells were averaged). The resulting mean elevation at 30' resolution was used in a maximum gradient search algorithm, which defined the flow direction to one of the eight neighboring cells by calculating the slope gradient to each of the neighboring gridcells and selecting the flow direction resulting in the highest downhill gradient. Due to DEM errors and errors introduced by aggregating the 5' DEM to coarser resolution the resulting simulated network had numerous errors in depicting the flow pathways of the real river systems. Manual editing (i.e. rotation of the gridcells individually) was applied to correct the DEM, by overlaying the simulated gridded network with digitized vector representation of river network such as ARC/World and Digital Chart of the World, which guided the rotation of the individual gridcells.

9.2 Data Processing Sequence

9.2.1 Processing Steps and Data Sets

The processes of developing simulated topological river networks starts with deriving initial flow routing from a DEM. This step involves projecting and re-gridding the DEM as necessary. In the case of STN-30p, which was derived from ETOPO5, aggregation of the 5' DEM to 30' resolution was necessary. Initial flow routing was derived from the aggregated 30' DEM using maximum search gradient (see Section 9.1.1)

The time consuming part of developing STN-30p was the manual editing necessary to correct the routing errors inherited from STN-30p. The first path of the network editing was done by overlaying the simulated network with ARC/World rivers vector coverage. ARC/World digital vector coverages of various map objects such as rivers, lakes, roads, railways, cities, etc. were digitized from 1:3 000 000 scale maps, therefore it represents only the major river systems of the Earth. Once the major rivers were corrected in STN-30p, it was overlaid by Digital Chart of the World, which is a similar set of digital vector coverages to ARC/World, but it has significantly more details since it was digitized from 1:1 000 000 scale Operational Navigational Charts (ONC) maps. Overlaying STN-30p by DCW helped to correct the smaller river systems.

After STN-30p was corrected based on vector river networks, the catchment area derived from STN-30p was compared against the catchment area given at discharge monitoring stations from various databases (e.g. “Discharge of Selected Rivers of the World” from UNESCO). The STN-30p was further edited to improve the correspondence between STN-30p derived and reported catchment area at discharge gauging stations.

9.2.2 Processing Changes

None.

9.2.3 Additional Processing by the ISLSCP Staff

The ISLSCP Initiative II staff converted the original data files from the ARC/INFO GRID format into the standard ASCII format used in the collection. During this process, all missing data were made consistent among layers assigning a value of -88 to all points with missing data over land, and a value of -99 to all ocean points. This was done by comparing the data to the ISLSCP II land/water and land/sea masks. The land/water mask was used because the Caspian Sea contained real data while the land/sea mask of ISLSCP II (i.e. no inland water bodies) did not contain the Caspian Sea. It was not possible to separate actual missing data over land (e.g. Greenland) from undefined points. This was done to all the layers except the 1/2 degree elevation grid.

9.3 Calculations

9.3.1 Special Corrections/Adjustments

A special correction was applied to HYDRO1k derived 30' DEM included in the present ISLSCP II data compilation. This correction eliminated the inconsistencies between STN-30p and the HYDRO1k derived 30' elevation data set. The inconsistencies (i.e. increasing elevation along downstream river routes) occur due to the differences between the initial DEM (ETOPO5) used to derive STN-30p and HYDRO1k and the manual correction of STN-30p. The inconsistencies were eliminated by a two paths processing algorithm, which performs a downhill search starting at the headwater cells (gridcells without any upstream gridcells), and lowers the elevation to maintain an minimum downhill gradient (1 cm/km). The second path of the algorithm starts at the headwater cells and applied a running mean averaging considering five upstream and 5 downhill gridcells. This running averaging was performed only for the mainstems of every possible routes (i.e. the downhill search was stopped whenever the algorithm encountered a confluence with the higher order stream). The resulting DEM and the corresponding slopes along STN-30p flow directions are shown on Figure 1 and Figure 2.

9.4 Graphs and Plots

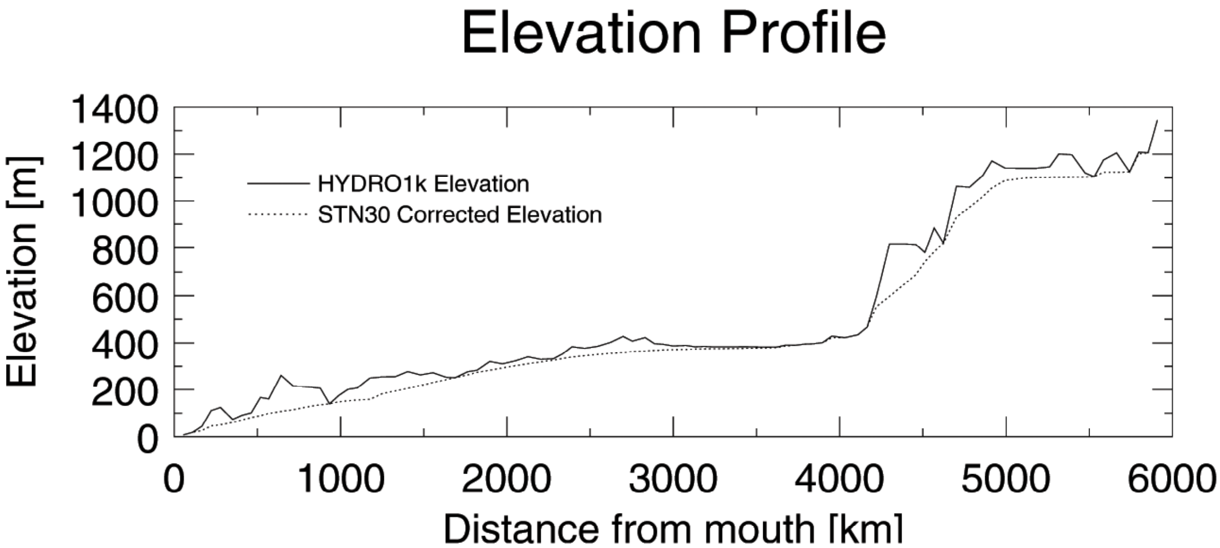

Figure 1. Mainstem elevation profile of the Nile

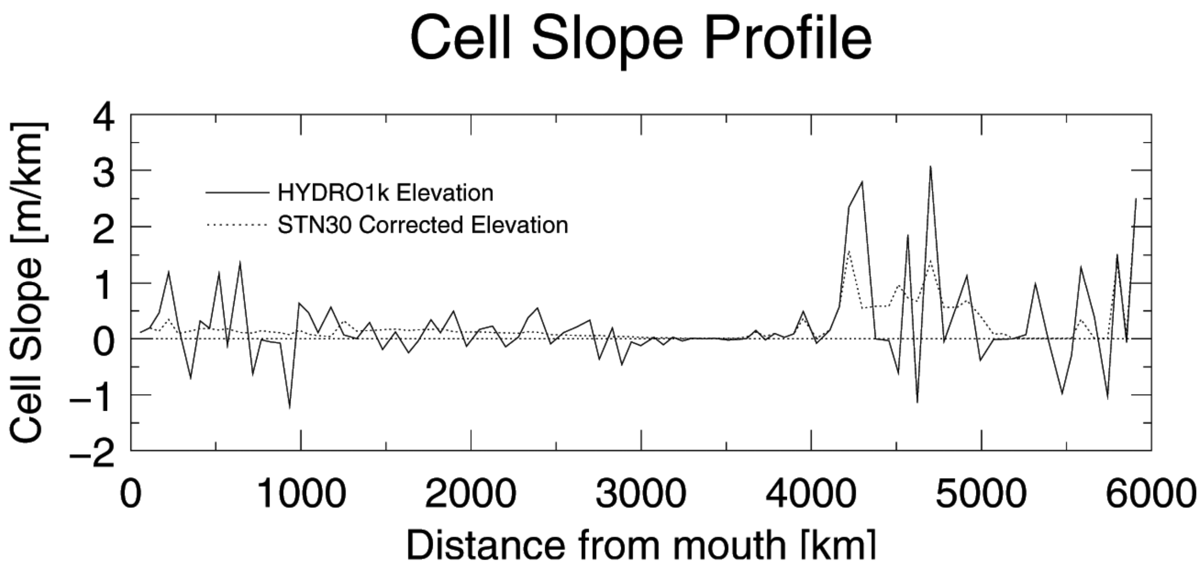

Figure 2. Mainstem slope profile of the Nile

10. ERRORS

10.1 Sources of Error

Gridded networks have numerous sources of error. First and foremost the spatial resolution imposes severe limitations on how accurately gridded networks can represent actual river networks. Vörösmarty et al. (2000a) found that the minimum catchment area that can be represented at a 10% accuracy by a 30-minute resolution network is around 25,000 km^2. Fekete et al. (2001) came to the similar conclusion from theoretical considerations. The accurate representation of river length related characteristics, such as basin shape (the ratio of mainstem length and catchment area) or width function (Veneziano, 2000) is even more limited (Fekete, 2001). Therefore the minimum basin size where these characteristics are at least 10% accurate is on the order of 400,000 km^2. The importance of the accurate representation of the length related characteristics varies by applications. STN-30p is adequate for monthly flow simulations, but for shorter time steps the user should realize the limitations in representing the river lengths accurately.

10.2 Quality Assessment

10.2.1 Data Validation by Source

STN-30p was tested against various data sets providing catchment area information for certain sub-basins (UNESCO-IP, 1964-84, Lammers, 2001) and for entire river basins (Meybeck, 1995). For instance, the comparison against 663 selected gauges from the Global Runoff Data Center archives showed 7.5 % absolute error with 2-3 % bias (which is largely due to STN-30p representing the potential catchment area). The catchment area of more than 40 % of the 663 gauges were represented within 5 % error.

10.2.2 Confidence Level/Accuracy Judgement

Not available at this revision for the 0.5 degree or 1.0 degree data sets.

10.2.3 Measurement Error for Parameters and Variables

See Section 10.1

10.2.4 Additional Quality Assessment Applied

None.

11. NOTES

11.1 Known Problems with the Data

The most important limitation of STN-30p is its relatively coarse resolution. Vörösmarty et al (2000a) and Lammer et al. (2001) demonstrated that the minimum basin size a 30' resolution network could represent at 10 % accuracy is around 10 000-25 000 km^2 (i.e. 5-10 gridcells). River basins smaller than this threshold are not depicted accurately or at all at 30' resolution.

Fekete et. al. (2001) demonstrated that a minimum of 200 gridcells is necessary for gridded networks to represent the geomorphological characteristics such as basin shape (the ratio of the mainstem length divided by the catchment area) or the width function (the distribution catchment area as a function of distance to mouth). At the resolution of STN-30p, 200 gridcells represent 400 000 -500 000 km^2, approximately the size of the Don or Limpopo river systems. Smaller river systems are still accurately represented in terms of catchment area, but their geomorphological characteristics are likely to be inaccurate. This poor representation means that the travel distances from various parts of the basins are inaccurate therefore the use of STN-30p is limited. STN-30p can be used without hesitation in cases when the network connectivity represents the dominant element of river routing and the travel time delays are negligible (e.g. calculating mean annual discharge). STN-30p can also be used for river routing, when the travel time delays are negligible in river basins smaller than 400 000 -500 000 km^2. Since these basins have a mainstem length in the order of 1000-2000 km (which is about the distance a drop of water can travel at the speed of 1 m/s – a typical velocity in natural rivers – in a month) therefore STN-30p is an adequate network for monthly flow routing. STN-30p can be applied for river routings at any shorter time steps, but the user has to realize that the discharge regimes in smaller basins than the mentioned threshold might be severely affected by the lack of proper representation of the geomorphological characteristics.

STN-30p was designed to represent potential flow therefore the catchment area according to STN-30p can be significantly different from reported values (e.g. Nile, Parana), when the reported catchment area reflects only the contributing (actively flowing) basin area (e.g. Nile, Niger, Parana).

11.2 Usage Guidance

STN-30p is ideal for calculating basin-wide characteristics (i.e. mean precipitation, air temperature, distribution of land cover classes, etc.) of the medium to large river basins. It can be used in flow routing simulations at monthly or longer time-steps, but the travel time delays in smaller basins (< 400 000 km^2) at shorter time steps could be severely affected by the lack of STN-30p to represent the routing distances accurately. Clearly, the 1.0 degree version provided here suffers from many of the same shortcomings.

11.3 Other Relevant Information

None at this revision.

12. REFERENCES

12.1 Satellite/Instrument/Data Processing Documentation

None.

12.2 Journal Articles and Study Reports

Band, L. E.: *Extraction of channel networks and topographic parameters from digital elevation data*, in Channel Network Hydrology (Editor: Beven K.), John Wiley & Sons, New York, NY, USA.

Fekete, B. M., C. J. Vörösmarty and R. B. Lammers: Scaling gridded networks for macro-scale hydrology: Development and analysis and control of error, *Water Resources Research*, Vol.: 37(7), pp.: 1955-1968, 2001

Edwards, M. O.: *Global Gridded Elevation and Bathymetry (ETOPO5)*, Digital Raster Data on a 5-minute Geographic (lat □ lon) 2160 □ 4320 (centroid-registered) grid, NOAA National Geophysical Data Center, Boulder CO, 1989

ESRI: *ARC/World*, CD-ROM, Environmental Systems Research Institute Inc, Redlands, CA, 1992

ESRI: *Digital Chart of the World for use with ARC/INFO*, Data Dictionary, 1:1M, CD-ROM, Environmental Systems Research Institute Inc, Redlands, CA, 1992

Jenson, S. K. and Dominque, J. O.: Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis, *Photogrammetric Engineering and Remote Sensing*, Vol: 54, pp: 1593-1600, 1988.

Lammers, R. B., A. I. Shiklomanov, C. J. Vörösmarty, B. M. Fekete: Assessment of contemporary Arctic river runoff based on observational discharge records, *JGR-Atmosphere*, Vol.:104(D4), pp:3321-3334, 2001

Meybeck, M.: *Global River Database (GLORI*), GEMS-Water, Geneva, 1995
Strahler, A.N.: *Quantitative geomorphology of drainage basins and channel networks*, In Chow V.T.: Handbook of Applied Hydrology, McGraw-Hill Book Company, Inc., 1964
UNESCO-IHP: *Discharge of Selected Rivers of the World Vol: I-III*, UNESCO Publication, 1969-84.
Veneziano, D., G. E. Moglen, P. Furcolo and V. Iacobellis: Stochastic model of the width function, *Water Resource Research*, Vol: 36(4), pp.: 1143-1157, 2000
Vörösmarty C. J., B. M. Fekete, M. Meybeck and R.B. Lammers: Global Systems of Rivers: Its role in organizing continental land mass and defining land-to-ocean linkages, *Global Biochemical Cycles*, Vol.: 14(2), pp: 599-621, 2000
Vörösmarty, C. J., B. M. Fekete, M. Meybeck and R. B. Lammers: Geomorphometric attributes of the global system of rivers at 30-minute spatial resolution, *Journal of Hydrology*, Vol.:237,pp.:17-39, 2000
Vörösmarty, C. J., B. M. Fekete and B. A. Tucker: *River Discharge Database, Version 1.0 (RivDIS v1.0)*, Volumes 0 through 6. A contribution to IHP-V Theme 1. Technical Documents Series, UNESCO, Paris, 1996

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

DAAC	Distributed Active Archive Center
DCW	Digital Chart of the World
DEM	Digital Elevation Model
ISLSCP	International Satellite Land Surface Climatology Project
GLORI	Global River Discharge Database
GRDC	Global Runoff Data Center
GSFC	Goddard Space Flight Center (NASA)
NASA	National Aeronautics and Space Administration
ONC	Operational Navigational Charts
ORNL	Oak Ridge National Laboratory

STN	Simulated Topological Network
UNH	University of New Hampshire