1. TITLE
2. INVESTIGATOR(S)
3. INTRODUCTION
4. THEORY OF ALGORITHM/MEASUREMENTS
5. EQUIPMENT
6. PROCEDURE
7. OBSERVATIONS
8. DATA DESCRIPTION
9. DATA MANIPULATIONS
10. ERRORS
11. NOTES
12. REFERENCES
13. DATA ACCESS
14. GLOSSARY OF ACRONYMS

1. TITLE

1.1 Data Set Identification

ISLSCP II UNH/GRDC Composite Monthly Runoff

1.2 Database Table Name(s)

Not applicable to this data set.

1.3 File Name(s)

The files in this data set comprise river discharge point data from the Global Runoff Data Centre (GRDC) for all 390 stations with measurements in the period 1986-1995 and 2) global, gridded, monthly composite runoff time series for the period 1986-1995, produced at UNH using the GRDC river discharge data. The two river discharge files contain the monthly river discharge for appropriate GRDC gauging stations and ancillary information on the stations (e.g. location, river basin name, etc…), respectively. The various files of the UNH-GRDC composite runoff fields are named using the following naming convention:

- **comp_runoff_1deg.zip**: When expanded, this file contains 120 .asc (ASCII) files (1986-1995, 1 file for each month) as 1.0 degree spatial resolution in latitude and longitude showing the mean monthly composite runoff fields. Example: **comp_runoff_1d_199406.asc** is the composite runoff from June 1994 at 1.0 degree resolution.

- **comp_runoff_hdeg.zip**: When expanded, this file contains 120 .asc (ASCII) files (1986-1995, 1 file for each month) as 0.5 degree spatial resolution in latitude and longitude showing the mean monthly composite runoff fields. Example: **comp_runoff_hd_199406.asc** is the composite runoff from June 1994 at 0.5 degree resolution.

- **runoff_hd_changemap.asc**: shows the differences between the ISLSCP II land/water mask and the original data set (see Sections 8.4 and 9.2.3).

- **runoff_subbasins_1deg.zip:** When extrapolated this file contains 10.asc files as 1.0 degree spatial resolution in latitude and longitude showing the subbasins of the time series discharge gauging stations which operated in year 19*yy* out of the 390 selected stations with observations (1986-1995). Example: **runoff_subbasins_1d_1986.asc**

- **runoff_subbasins_hdeg.zip:** When extrapolated this file contains 10.asc files at 0.5 degree spatial resolution in latitude and longitude showing the subbasins of the time series discharge gauging stations which operated in year 19*yy* out of the 390 selected stations with observations (1986-1995). Example: **runoff_subbasins_hd_1986.asc**
- **wbm_runoff_1deg.zip:** When extrapolated contains 120.asc files (1986-1995, 1 file for each month) showing the mean monthly water balance model runoff, as 1.0 degree spatial resolution in latitude and longitude.

- **wbm_runoff_hdeg.zip:** When extrapolated contains 120.asc files (1986-1995, 1 file for each month) showing the mean monthly water balance model runoff, as 0.5 degree spatial resolution in latitude and longitude.

- **runoff_subbasins_data.zip:** 1 file which when extrapolated contains 4 files: a readme file, a text file, a comma-separated data file with station ID, river names, areas, and data values subbasins of the discharge gauging stations which operated out of the 390 selected stations, and a .def file.

- **grdc_river_disch1986-1995.dat :** Contains the discharge data by station ID and collection year.

- **grdc_stations1986-1995.dat:** Contains the data on the station ID number, river name, station name, country, area, start month, end month, end year, lat, lon.

- **runoff_corrections_1deg.zip:** Gridded ASCII file showing the long-term mean annual correction coefficient based on climatological discharge gauges.When expanded,contains 10.asc files at 1.0 degree spatial resolution.

- **runoff_corrections_hdeg.zip:** Gridded ASCII file showing the annual correction coefficients based on the applicable discharge monitoring stations out of the selected gauges.When expanded,contains 10.asc files at 0.5 degree spatial resolution.

1.4 Revision Date of this Document

February 14, 2011

2. INVESTIGATOR(S)

2.1 Investigator(s) Name and Title

Dr. Balázs M. Fekete
Research Scientist
Complex Systems Research Center
Institute for the Study of Earth, Oceans, and Space
University of New Hampshire

Dr. Thomas Maurer
Head
Global Runoff Data Centre (GRDC)
Federal Institute of Hydrology (Bundesanstalt für Gewässerkunde)
Koblenz, Germany

2.2 Title of Investigation

Database development for the Global Hydrological Archive and Analysis System (GHAAS)

2.3 Contacts (For Data Production Information)

	Contact 1 (UNH)	Contact 2 (UNH)
2.3.1 Name	Dr. Balázs M. Fekete	Dr. Charles J. Vörösmarty
2.3.2 Address	CUNY Energy Institute The City University of New York	CUNY Energy Institute The City University of New York
City/St.		
Zip Code		
Country	USA	USA
2.3.3 Tel. No.	212-650-8146	212-650-7042
Fax No.		
2.3.4 E-mail	bfekete@ccny.cuny.edu	cvorosmarty@ccny.cuny.edu

	Contact 3 (GRDC)
2.3.1 Name	Dr. Thomas Maurer
2.3.2 Address	Global Runoff Data Centre in the Federal Institute of Hydrology (BfG) Am Mainzer Tor 1
City/St.	56068 Koblenz
Zip Code	
Country	Germany
2.3.3 Tel. No.	49 261 1306-5224
Fax No.	49 261 1306-5280

2.3.4 E-mail	thomas.maurer@bafg.de

2.5 Data Set Citation

Fekete, B., and C. J. Vörösmarty. 2011. ISLSCP II UNH/GRDC Composite Monthly Runoff. 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/994

2.4 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.

Fekete, B. M., C. J. Vörösmarty, W. Grabs (2002) High Resolution Fields of Global Runoff Combining Observed River Discharge and Simulated Water Balances, *Global Biogeochemical Cycles*, 16(3).

Fekete, B. M., C. J. Vörösmarty, W. Grabs (1999) *Global, Composite Runoff Fields Based on Observed River Discharge and Simulated Water Balances*, GRDC Report 22, Global Runoff Data Center, Koblenz, Germany.

3. INTRODUCTION

The UNH/GRDC composite runoff data combines simulated water balance model runoff estimates derived from climate forcing with monitored river discharge from the GRDC. It can be viewed as data assimilation applied in a water balance model context (conceptually similar to the commonly used 4DDA techniques used in meteorological modeling).

This data set received mixed reviews during the peer-review process of deciding the inclusion of various data sets into the International Satellite Land Surface Climatology Project (ISLSCP) Initiative II data collection. The majority of the concerns regarding this data set stemmed from being a combination of model results and observations. Some reviewers expressed concerns about potential misuse if the users are not fully aware of the nature of the compositing process and do not realize how much is model output and how much is based on observation. The ISLSCP Initiative II Science Working Group opted for the inclusion of this data set due to its uniqueness. Users should read this document carefully before using this data set.

3.1 Objective/Purpose

River discharge, which is the integrated signal of runoff, is perhaps the most accurately measured component of the hydrological cycle (Hagemann and Dümenil, 1998; Krahe et al. 1996; Rantz, 1982) but it is not yet fully utilized in Earth system studies partly due to the limited access to such data and the difficulties relating the point measurement of discharge to the other components of the hydrological cycle. This situation is changing by the increasing number of nations publishing discharge related information via the World Wide Web, and the World Meteorological Organization's recent effort to establish the hydrological component of their Global Terrestrial Network (Cihlar 2001). Recently released gridded networks with global extent (Vörösmarty et al. 2000a, 2000b, Oki and Sud, 1998, Graham et al., 1998) provided the means to establish the linkage between the landmass and the river channels (and ultimately the oceans).

The purpose of the present data set is to demonstrate the value of combining river discharge observations with spatially distributed runoff estimates from water balance calculations. Such a data assimilation scheme preserves the spatial specificity of the water balance calculations while constrained by the more accurate discharge measurement. The resulting composite runoff estimates are useful for numerous applications ranging from water resource assessments (Vörösmarty 2000c; Revenga 2001; UNEP/FAO/OSU, 2002) to validation of atmospheric models (Dai and Trenberth, 2002).

3.2 Summary of Parameters

- 10 x 12 layers (ten years with 12 monthly values), Composite runoff grid representing the January, 1986 - December, 1995 period.
- 1 layer, Climatological annual long-term average water balance model runoff.
- 10 layers, Annual water balance model runoff (January, 1986 - December, 1985)
- 10 x 12 layers, Monthly water balance runoff grid (January, 1986 - December, 1995).
- 1 layer, Climatological correction coefficient
- 10 layers, Annual correction coefficients.
- 1 layer, Climatological station subbasin grid
- 1 layer, Subbasin grid of the climatological stations
- 10 layers, Subbasin grids of the operating time series stations for each year in the 1986 - 1995 period.
- Monthly river discharge measured at 390 gauging stations for the period 1986-1995.
- List of the 390 gauging stations with attributes.

The UNH/GRDC runoff data were originally developed at a spatial resolution of 1/2 degree in both latitude and longitude. The ISLSCP II staff has aggregated the 1/2 degree to 1 degree resolution for this collection.

3.3 Discussion

UNH-GRDC composite runoff fields represent corrected estimates of runoff over the continental landmass by combining water balance model runoff estimates (Vörösmarty et al., 1989, 1998) with observed discharge from the archives of the Global Runoff Data Centre, Koblenz, Germany (Grabs et al., 1996, http://www.bafg.de/GRDC/EN/Home/homepage__node.html). The resulting composite runoff fields preserve the spatial patterns of the water balance model runoff and yet are constrained by the observed discharge.

Monthly water balance model estimates for the 1986-1995 time period were calculated using climate forcing data (i.e. air temperature, precipitation, wind speed, cloud coverage, vapor pressure deficit) from the Climate Research Unit (CRU) of the University East Anglia (New, 1999 and 2000). Most of the climate data (air temperature, precipitation, cloud coverage, vapor pressure deficit) were available as a time series for the 1901-95 period. The wind speed was only available as a climatology of the 1961-1995 period; therefore, in the water balance calculations the climatological wind speed repeated for every year during the simulated period. The water balance model from Vörösmarty et al. (1989, 1998) was configured with Shuttleworth and Wallace (1985) potential evaporation function using land cover characterization from the Terrestrial Ecosystem Model (TEM) (Melillo, 1993) and soil characterization from FAO/UNESCO (1986).

The discharge monitoring stations were co-registered to a 30-minute simulated topological network (STN-30p) from Vörösmarty et al. (2000a, 2000b). The co-registration of the discharge gauging stations allowed for the delineation of inter-station regions (Figure 1, Section 9.4) between the gauging stations. These inter-station regions established the basis for the correction of the water balance model estimated runoff. Mean annual runoff over the inter-station regions was calculated by dividing the inter-station discharge (the discharge generated within the inter-station region, i.e. the difference between the discharge at the outlet of the inter-station region and the sum of the discharge values entering from upstream) by the area of the inter-station region. Mean annual water balance model runoff was computed by averaging the simulated water balance model runoff over the inter-station regions of the discharge monitoring stations. The ratio of the observed vs. simulated average annual runoff was applied as a correction coefficient to the water balance simulated runoff.

This technique (Fekete et al., 1999 and 2002) is fairly simple to implement when discharge for the same monitoring stations are available for the simulated period. Unfortunately, it was not the case for the 1986-1995 period, therefore the reporting stations for every year had to be identified and corresponding tributaries and inter-station discharges calculated. Since large portions of the otherwise monitored land mass had no observational record during the 1986-1995 period, a long term mean annual correction (correction coefficient calculated using climatological average forcings related to climatological mean discharge) was applied when no time series observation was available. Regions with no observation at all (neither long-term mean nor time series records) were left without any correction (i.e. water balance model only).

Extreme values of the correction coefficient (< 0.5 or > 2.0) were rejected. Such values of the correction coefficients are indicators of inconsistencies in the discharge data or the deficiency of STN-30p in representing sub-basins associated with discharge gauging stations. A typical case of extreme values is when the correction coefficient is negative as a result of decreasing discharge downstream (i.e. negative inter-station discharge). This situation could occur on "losing" streams, where the river discharge either naturally seeps into the groundwater or due to intensive human water uptake (e.g. for irrigation). Extreme values, besides likely being erroneous or unnatural, would cause sudden changes in the corrected runoff fields. To avoid such sudden changes the correction coefficient was applied only when its value was in the 0.5 -- 2.0 range. Otherwise the original water balance runoff estimate was unchanged.

The present data set contains not only the composite runoff fields but all the intermediate products such as the water balance model runoff estimate, the inter-station regions of the selected stations year by year (Figure 2, Section 9.4), the monthly climatology, and the 10 year time series for the ISLSCP II period (1986-95) of the discharge gauges.

All the data (input forcings, water balance model runoff estimates, gridded network, discharge gauges along with the associated discharge records) applied in the development the composite runoff fields are included in the ISLSCP Initiative II data compilation either as part of the present data set or separate component of the ISLSCP Initiative II data holdings. The runoff files have been checked against the ISLSCP II land/water mask and have been made consistent with the ISLSCP II land/water mask. Files containing any cells that have been removed in this phase are provided in separate files.

4. THEORY OF ALGORITHM/MEASUREMENTS

The classical representation of the hydrological processes distinguishes two phases of the runoff generation processes. Water balance calculation represent the surface processes as the precipitation is translated to spatially distributed runoff which is the excess water generated on unit land surface area as a residual from evapotranspiration and changes in water stock size (soil moisture, shallow groundwater).

5. EQUIPMENT

The present data set combines river discharge measurement with water balance model estimates. As such, this data set is a result of a large array of measurements of the climate forcing (e.g. precipitation, air temperature, vapor pressure, solar radiation, wind speed) land surface characteristics (e.g. elevation, land-cover, etc.) and river discharge. Since river discharge is the key element of the present data set, in this section only the discharge gauging will be discussed. See the documentation for the CRU data set in this ISLSCP II data collection for more information on that data set.

5.1 Instrument Description

River discharge is typically inferred from stage height (flow height) measured by hydrograph (a floating device that can record the water level at any given time). The discharge is estimated from stage height using “rating curves” that relates the stage height to discharge (Rantz, 1982). See http://grdc.bafg.de/ for more information.

5.1.1 Platform (Satellite, Aircraft, Ground, Person)

Hydrogaph.

5.1.2 Mission Objectives

To measure river discharge.

5.1.3 Key Variables

River flow height (stage height).

5.1.4 Principles of Operation

See http://grdc.bafg.de/ for more information.

5.1.5 Instrument Measurement Geometry

See http://grdc.bafg.de/ for more information.

5.1.6 Manufacturer of Instrument

Varies by country.

5.2 Calibration

The key to accurate discharge measurement is establishing the rating curve (relating stage height to discharge). This calibration of the stage height records is carried out through field surveys when the discharge is accurately measured for given stage heights by surveying the riverbed geometry, and the flow velocity distribution within the riverbed cross-section. The discharge surveys represent individual points on the rating curve. A series of field surveys is necessary at different stage heights to establish the rating curve accurately. Furthermore, the rating curve could change over time as the riverbed changes, therefore repeated field surveys are necessary to capture the potential changes of the rating curves.

5.2.1 Specifications

5.2.1.1 Tolerance

The overall accuracy of the discharge measurement (the accuracy of the stage height record and the uncertainties in the rating curves) is typically in the 5 - 10% according to the United States Geological Survey (USGS) (Rantz, 1982). This probably could go up to 20 % (Hagemann and Dümenil, 1998, Dingman, 2001).

5.2.2 Frequency of Calibration

The frequency of field surveys varies from station to station. The regular field surveys are necessary not only to capture more points along the rating curve (the discharge at different stage heights) but also to track the potential changes in the riverbed that can cause the alteration of the rating curve. On average, USGS carries out 6 - 7 field surveys per station each year (Rantz, 1982).

5.2.3 Other Calibration Information

None given.

6. PROCEDURE

6.1 Data Acquisition Methods

Water balance calculations were carried out using gridded mean monthly climate input (air temperature, precipitation, cloud cover, vapor pressure and wind) at 30' resolution from the Climate Research Unit of the University of East Anglia (New 1999, 2000). These gridded climate data for the 1901-1995 time period were available on CD-ROM from CRU directly (http://www.cru.uea.ac.uk/), but a subset -- covering the 1986 - 1995 period -- is also included in the ISLSCP Initiative II holdings on the recommendation of the developers of the present composite runoff fields.

Correction coefficients were calculated as the ratio of the observed mean runoff (the inter-station discharge divided by the inter-station area) vs. the average of the mean annual water balance runoff over the inter-station area was calculated on the annual basis (where the inter-station discharge is the difference between the discharge at the upstream gauges and the sub-basin outlet, while the inter-station area is the catchment area between the upstream gauges and the basin outlet).

The individual monthly grid cell values of the water balance model runoff were multiplied by the correction coefficient of the corresponding monitoring station when the correction coefficient fell within the 0.5 - 2.0 range. Otherwise, climatological corrections were calculated similarly from climatological discharge gauging stations. Water balance runoff estimates were left unchanged when no time series or climatologically monitored observed runoff was available or both the time series and the climatological correction would have fallen outside of the 0.5 - 2.0 range.

6.2 Spatial Characteristics

6.2.1 Spatial Coverage

The UNH/GRDC Composite runoff fields cover the same geographic extent (between 55 degrees S to 83 degrees N excluding Greenland and the Arctic Archipelago) as the STN-30p network, which provided the basis for the compositing of river discharge and water balance model runoff. The runoff files provided here are global in nature but do not have actual data over Greenland and Antarctica. The river discharge data from GRDC come from stations located throughout the world.

6.2.2 Spatial Resolution

The UNH/GRDC Composite runoff fields were developed at a 1/2 degree spatial resolution in both latitude and longitude. The ISLSCP II staff has aggregated the 1/2 degree data to 1 degree spatial resolution. The river discharge data are point data.

6.3 Temporal Characteristics

6.3.1 Temporal Coverage

The temporal coverage of the GRDC archive varies by station (see http://grdc.bafg.de/). The temporal coverage of all the data provided here is January, 1986 to December, 1995.

6.3.2 Temporal Resolution

The composite runoff fields are monthly. Most of the river discharge data are monthly. A few stations provide daily records (see **grdc_stations1986-1995.txt**) file.

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
River Discharge Data (grdc_river_disch1986-1995.dat)				
GRDC-ID	Identification number for each discharge gauging station after GRDC.	N/A	N/A	GRDC
Date	Year and month of discharge measurement.	N/A	N/A	GRDC
Discharge	Monthly river discharge measured at a particular station.	Min=0 Max=306317 No Data = -9999	m^3/s	GRDC
Discharge Station Attributes (grdc_stations1986-1995.dat)				
ID	Station number	1-390	N/A	
GRDC-ID	Identification number for each discharge gauging station after GRDC.	N/A	N/A	GRDC
River Name	Name of river where station is located	N/A	N/A	
Station Name	Name of the discharge station	N/A	N/A	
Country	Two letter country code	N/A	N/A	
Area	Reported catchment area	10141 to 4640300	km^2	
StartMonth	Beginning month of observation records.	01 to 12	N/A	
Start Year	Beginning year of observation record.	1807 to 1986	N/A	
End Month	Last month of observation records.	01 to 12	N/A	
End Year	Last year of observation records	1990 to 1999	N/A	
Time Series	Time Series Type:	M or D	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
	"M" = Monthly "D" = Daily			
Percent Record	Percent of missing values in observation records.	0 to 82.9	N/A	
GRDC663	Station used as climatological station ? 0 = No 1 = Yes	0-1	See 8.2.2	
LonDD	Latitude for the center of a cell. South latitudes are negative.	Min=-90 Max=90	Decimal Degrees	
LatDD	Longitude for the center of a cell. West longitudes are negative.	Min=-180 Max=180	Decimal Degrees	
Monthly Composite Runoff Maps (2 files, extrapolated to *.asc)				
Monthly Composite Runoff	Monthly discharge-corrected runoff	Min=0 Max=1000 Water=-99 No Data= -88	[mm/month]	UNH
Monthly Water Balance Model Runoff Maps (2 files, extrapolated to *.asc)				
Monthly WBM Runoff	Modeled water balance runoff	Min=0 Max=1000 Water=-99 No Data= -88	[mm/month]	UNH
Differences Tables (*.dif)				
Lat	Latitude for the center of a cell. South latitudes are negative.	Min=-90 Max=90	Decimal Degrees	Earth Grid
Lon	Longitude for the center of a cell. West longitudes are negative.	Min=-180 Max=180	Decimal Degrees	Earth Grid
Data_Removed	Data value in each cell of the original file that did not match the ISLSCP II land/water mask, and was removed.	Min=0 Max=1000	[mm/month]	Original Data
Change Map (*_changemap.asc)				
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)	Min=-1 Max=1	See 8.2.2	Computed

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
	0 = Data sets agree over land or water (data unchanged) 1 = ISLSCP II mask is land and original data is water or missing (data added as -88)			
Runoff Corrections Maps (2 files, extrapolated to *.asc)				
Runoff Correction (long term)	Long-term mean annual correction coefficient based on 660 climatological discharge gauges.	N/A		UNH
Runoff Correction (yearly)	Annual correction coefficients based on the applicable discharge monitoring stations out of the 390 selected gauges.	N/A		UNH
Runoff Subbasins Maps (2 files, extrapolated to *.asc)				
Runoff Subbasin (long term)	Map of the subbasins of the 390 discharge gauging stations with observations in the 1986-95 period.	N/A		UNH
Runoff Subbasin (yearly)	Maps of the subbasins of the time series discharge gauging which operated in year 19*yy* out of the 390 selected stations.	N/A		UHN

8.3 Sample Data Record

The "differences" file is an ASCII table with some header lines, then the Lat and Lon coordinates of each removed point, plus the value of that point. See the sample below.

```
ISLSCP-2 Differences for file 'comp_runoff_hd_198601.asc'.
Contains Lat-Lon coordinates and data for each point in the original file
that differed from the ISLSCP-2 Land/Sea mask, and thus was removed.

Lat,Lon,Data_Removed
82.25,-61.75,0.0
81.75,-72.25,0.0
81.75,-71.75,0.0
81.75,-71.25,0.0
81.75,-65.25,0.0
81.75,-25.25,0.0
81.25,-92.25,0.0
81.25,-89.75,0.0
```

Samples of the river discharge file (**grdc_river_disch1986-1995.dat**) are given below:

```
"GRDC-ID"      "Date"      "Discharge"
"1425500"      "1986-01"   97.440002
```

```
"1425500"      "1986-02"    75.660004
"1425500"      "1986-03"    87.169998
"1425500"      "1986-04"    92.230003
"1425500"      "1986-05"    210.350006
"1425500"      "1986-06"    351.730011
"1425500"      "1986-07"    144.320007
"1425500"      "1986-08"    147.600006
"1425500"      "1986-09"    539.729980
```

Samples of the discharge station attributes file (**grdc_stations1986-1995.dat**) are given below:

```
"ID"      "GRDC-ID"    "RiverName" "StationName"     "Country"    "Area"
   "StartMonth"        "StartYear" "EndMonth"  "EndYear"    "TimeSeries"
   "PercentRecord"     "GRDC663"    "LonDD"      "LatDD"
1  "1112100"    "SENEGAL"    "KAYES"      "MI"  157400.000000     6      1951
   12     1990   "M"    0.000000     0      -12.450000  14.450000
2  "1112200"    "FALEME"     "GOURBASSY" "MI"  15000.000000      4      1954
   12     1990   "M"    0.000000     1      -11.380000  13.230000
3  "1112300"    "SENEGAL"    "GALOUGO"    "MI"  127000.000000     5      1905
   12     1990   "M"    35.000000    1      -11.130000  13.830000
```

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 files at different spatial resolutions each contain the following numbers of columns and rows:

1.0 degree: 360 columns by 180 rows

1/2 degree: 720 columns by 360 rows

(note: the "1d" files were created by the ISLSCP Staff through averaging). All values in these files are written as real numbers. Missing data cells are given the value of –99 over water, and -88 over land (mainly Greenland and Antarctica).

The **grdc_river_disch1986-1995.dat** and **grdc_stations1986-1995.dat** files are text files where each column of data is separated by a “tab” character (i.e. tab-delimited file). These files can be easily imported into standard spreadsheet programs.

The ASCII **comp_runoff** and **wbm_runoff** map files (with the extension of ".asc") have all had the ISLSCP II land/water mask applied to them. All points removed from the original half-degree files are stored in "differences" files (with the extension ".dif"). These ASCII files contain the Latitude and Longitude location of the cell-center of each removed point, and the data value at that point. There is one ".dif" file for each half-degree ASCII map file.

The "change map" file shows the results of applying the land/water mask to the **comp_runoff** and **wbm_runoff** map files, as a viewable ASCII map: all points added ("1"), all points unchanged ("0"), and all points removed ("-1"). There is only a file for the half-degree data, as the 1-degree data was created through averaging.

The **runoff_corrections** and **runoff_subbasins** ASCII map files have not been masked, thus there are no ".dif" files and no change map associated with them.

WARNING: The 1x1 degree map products are recommended for browse use only. These data files are averaged from the original 0.5 x 0.5 degree pixels. Thus the data values at specific pixels are not exact. Use this data with caution and always refer to the original half-degree data files for specific information.

8.5 Related Data Sets

Several related data sets are included in this ISLSCP II data collection, including the CRU climate data, the STN-30p river network and the Hydro1k Digital Elevation data set. Users may wish to examine the following web sites for additional data sets: The Global Runoff Data Centre (http://www.bafg.de/GRDC/EN/Home/homepage__node.html) , the Water System Analysis Group (http://www.wsag.unh.edu/), and the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) http://daac.ornl.gov/ISLSCP.islscpii.html .

9. DATA MANIPULATIONS

9.1 Formulas

The key formula of the water balance calculation is:

$$R = P - E - dS/dT \tag{1}$$

Where R is runoff, P is precipitation, E is estimated evapotranspiration and dS/dt is the soil moisture rate of change.

The runoff correction can be formulated as:

$$\bar{\xi}_{si} = R_{oi}/R_{wi} \tag{2}$$

where R_{oi} is observed runoff calculated from measured discharge, R_{wi} is the water balance model runoff averaged over the tributary of the corresponding discharge monitoring station for discharge station i.

The corrected runoff can be expressed as

$$R_c = \xi_{si} R_{wbm} \tag{3}$$

Where R_{wbm} is the water balance model simulated grid cell runoff and R_c is the corrected grid cell runoff.

9.1.1 Derivation Techniques/Algorithms

See Fekete et al. (2002) for more information.

9.2 Data Processing Sequence

9.2.1 Processing Steps and Data Sets

See Fekete et al. (2002) for more information.

9.2.2 Processing Changes

None.

9.2.3 Additional Processing by the ISLSCP Staff

The original runoff data files submitted to the ISLSCP staff were ASCII ARCINFO grid files, with 6 header lines, at the half-degree scale. The data were found to be lacking the first 14 lines of the globe and the last 69 lines of the globe (containing Antarctica). Thus, the images were 277 data lines (should be 360), and 720 values per data line. The files also had missing data encoded as "-9999.000000" and 0 encoded as "0.000000". The data files were reprocessed to remove the 6 header lines add and the top 14 lines and bottom 69 lines from the modified half-degree land/sea mask (-99 water, -88 land). Missing data and zero were changed to "-99" and "0.0". The data files were then renamed to the current naming scheme (see Section 1.3).

The ISLSCP II staff then processed the half-degree ASCII **comp_runoff** and **wbm_runoff** map files by comparing them against the ISLSCP II land/water mask. Missing land data in these files were added as -88. New ASCII table files containing the removed points (points that didn't match the land/water mask), also called "differences" files with the extension ".dif", were also created. These files contain the Latitude and Longitude of the cell-center of each point removed, and the value removed from each map.

A "change map" was also created for the half-degree data, showing the results of applying the land/water mask to the **comp_runoff** and **wbm_runoff** map files. This viewable ASCII map shows all points added ("1"), all points unchanged ("0"), and all points removed ("-1"). Users can use the data, the ".dif" files and the change map to re-create the original data set as it was submitted by the Principal Investigator.

The **runoff_correction** and **runoff_subbasin** ASCII map files were not masked, thus there are no ".dif" files and no change map associated with them.

Seeing a need for a 1-degree product, the ISLSCP Initiative II Staff took the half-degree data files and aggregate them to a 1-degree spatial resolution. For each 1 degree cell, 4 half-degree cells were averaged, ignoring missing data cells, and filling the new cell with -99 if three or more half-degree cells were -99 as well. Then the data files were renamed to the current naming scheme (see Section 1.3). Thus, there are no ".dif" files or "change maps" for the 1-degree data, because it is a "created" product.

WARNING: The 1x1 degree product is recommended for browse use only. These data files are averaged from the original 0.5 x 0.5 degree pixels. Thus the data values at specific pixels are not exact. Use this data with caution and always refer to the original half-degree data files for specific information.

9.3 Calculations

9.3.1 Special Corrections/Adjustments

See Fekete et al. (2002) for more information.

9.4 Graphs and Plots

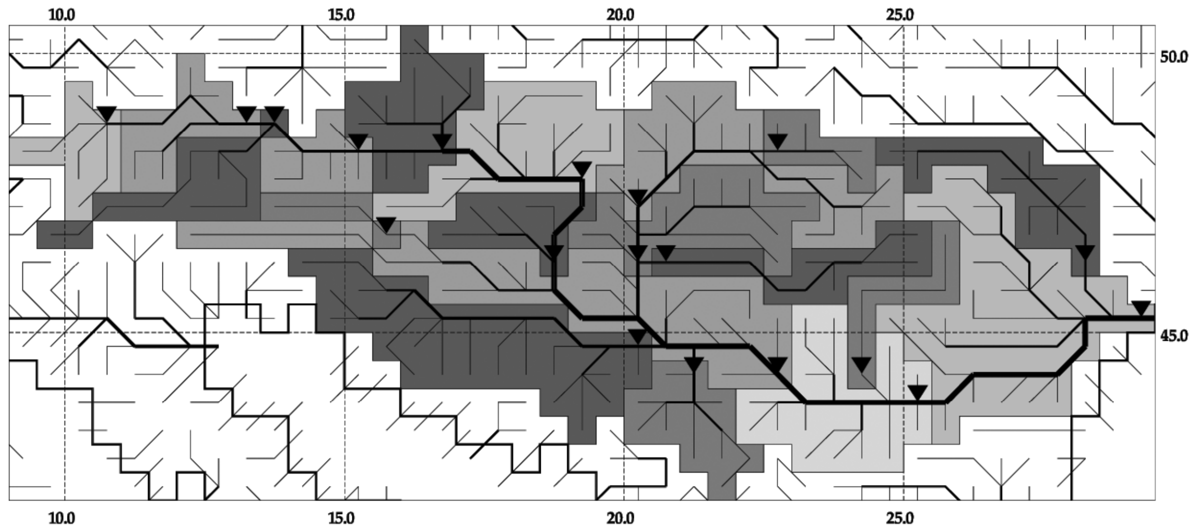

Figure 1: Inter-station regions of the Danube river basin derived from the 30-minute network (STN30p).

Figure 2: Distance to next stations. The figure shows the temporal evolution of the discharge gauging network by shading the continental land-mass according to the distance to the nearest downstream station. The distance to the next station is a good indicator of how well the river systems are monitored. Darker shades represent shorter distance. FIGURE NOT AVAILABLE AT THIS REVISION.

10. ERRORS

10.1 Sources of Error

The UNH/GRDC data set is the result of a complex processing of a wide array of input data, therefore the potential errors in any input source could affect the final product. Errors in the climate forcings (the precipitation in particular) have a significant effect on the water balance calculations.

The water balance calculation itself has its limitations representing the complexity of the soil-vegetation-atmosphere water exchange processes. The correction of the water balance model runoff with observed discharge improves the runoff estimates since river discharge is generally more accurately measured than any other component of the water cycle, but river discharge measurements have their own errors of 5 to 10% (Hageman and Dümenil, 1998; Rantz, 1982) Besides potential errors in the discharge measurement the discontinuities in the discharge records may introduce serious artifacts as well. Whenever the discharge gauging network configuration changes, the correction coefficient is potentially derived from different stations with different spatial and error characteristics that can cause sudden changes in the composite runoff fields.

Furthermore, the gridded river network (STN-30p) representing the linkage between the landmass and the river systems has its own limitations that can introduce additional errors in the runoff calculation. First of all, any error in depicting the tributaries of the discharge gauges translates to similar error in the inter-station runoff estimate. Furthermore, erroneous network topology could result in errors in the station topology derived from the simulated network.

10.2 Quality Assessment

10.2.1 Data Validation by Source

The UNH/GRDC composite runoff fields incorporate all the available information (climate forcings, river discharge measurements) to provide the most realistic assessment of the continental runoff. Since the goal in developing such a data set is to use all the information for creating the data set no information was left for validation. This is a conceptually different approach than the traditional model calibration and validation, where calibration data are used to adjust the model and validation data are used to test the model. The UNH/GRDC composite runoff field is not meant to be modeled data, it is purposely a combination of the available measurement and observation, which represents our current knowledge about the spatial and temporal distribution of the continental runoff.

However additional data are available in parts of the globe (i.e. United States, Canada and some parts of Europe) the testing of the composite fields against these additional data sources would be misleading since these regions are well represented in the GRDC data archive and in the relevant climate forcing data therefore the composite runoff fields are likely to perform better than in less monitored regions.

10.2.2 Confidence Level/Accuracy Judgment

The accuracy of the composite runoff fields varies spatially and temporally. Well monitored regions (i.e. United States and most parts of Europe) are more reliable, while poorly monitored regions (i.e. most of Africa, Central and South America) are less reliable. The distribution of the inter-station regions accompanied with the composite runoff fields could be consulted to assess the degree of monitoring, which largely dictates the expected accuracy.

10.2.3 Measurement Error for Parameters and Variables

The UNH/GRDC composite runoff fields data set integrates various data. The most critical ones are the precipitation input and the discharge records that were used to correct the runoff estimate. Gauge undercatch in precipitation measurements can exceed 100 % (Legates and Willmott, 1990). Discharge measurements have 5-20 % accuracy (Hageman and Dümenil, 1998; Rantz, 1982).

10.2.4 Additional Quality Assessment Applied

None.

11. NOTES

11.1 Known Problems with the Data

The UNH/GRDC composite fields were developed by applying annual correction coefficients to the monthly data. This is a necessary approach, since many discharge gauging station have large tributaries where the assumption of negligible year-to-year storage is reasonable, but substantial storage change is possible on a monthly basis. Applying annual corrections assumes that the seasonal timing of the water balance runoff captures the seasonal timing correctly

A further limitation is the proper treatment of the so-called losing streams. Some rivers have decreasing discharge going downstream. These losing streams are feeding their surrounding regions either due to human intervention (irrigation) or naturally (riverwater seeping into the groundwater). The current composite fields ignore these regions and maintain the water balance model runoff estimate (since the inter-station discharge comes out negative in these regions the correction coefficient would be negative as well, which does not fall in the 0.5 – 2.0 range).

The limitation of the correction coefficient is also a result of potential problems in the observed discharge records and the river network representation. The 0.5 – 2.0 range was chosen arbitrary without an in depth investigation of the cause of extreme correction coefficient values.

11.2 Usage Guidance

The use of the composite runoff data for validation is not recommended (because it is a mixture of modeled and measured discharge). The composite fields, will be much more reliable in watersheds where observations are present. This is because correction coefficients (a ratio correction of annual-average observations to annual-average WBM model output) have been determined for each grid cell in each basin where observations were available on annual and climatological bases. These coefficients were then used to correct the WBM output to observations and to produce grid-based, spatial patterns of runoff. This will give some indication of the spatial patterns of runoff within the basins (i.e. the basin average discharge will match observations) but the pattern of runoff within the basin will be WBM and forcing dependent. The user should bear in mind that the number of monitored basins decreases significantly from 1986 to 1995. By the end of the ISLSCP Initiative II period, most of the observations are in North and South America, and in Europe.

The monthly, annual-average and climatological global fields of WBM output are model outputs, There is a potential inconsistency when comparing to ISLSCP Initiative II data sets that did not use the climate data set from the Climate Research Unit of the University of East Anglia, Furthermore, the water balance model is affected by not only the climate forcings but the land surface condition inputs such as land use, land cover, vegetation, soil type, and topography, Finally, since the runoff data set is calibrated with observed discharge, the runoff field implicitly includes the loss due to irrigation. The irrigation withdrawal is approximately 10% of river runoff on global average, but sometimes more than 20% particularly in developed countries, and this amount may not be negligible.

In summary, the most reliable portions of this data set are the observed hydrograph data and the time series of gridded templates for the basins. The spatial patterns from the corrected regions in the composite fields should not be used as validation, rather simply as a comparison data set. Points where the correction coefficients fell outside the 0.5 to 2.0 range (i.e. the suspect data) are flagged as being both model output and potentially problematic.

11.3 Other Relevant Information

None given.

12. REFERENCES

12.1 Satellite/Instrument/Data Processing Documentation

None.

12.2 Journal Articles and Study Reports

Cihlar, J., W. Grabs and J. Landwehr*: Establishment of a Global Hydrological Observation Network for Climate,* Report of the GCOS/GTOS/HWRP Expert Meeting, GCOS 63, GTOS 26, WMO/TD No. 1047, World Meteorological Organization, 2001

Dai, A. and K. E. Trenberth: Estimates of freshwater discharge form continents: Latitudinal and seasonal variations, *Journal of Hydrometeorology*, submitted, 2002

Dingman, S. L.: *Physical Hydrology*, 2nd Edition, Prentice-Hall Inc., Englewood Cliffs, NJ, 2001

FAO/UNESC: FAO *Soil Map of the World in Digital form*, UNEP/GRID, Digital Raster Data on 2-minute Geographic (lat x lon) 5400x10800 grid, Carouge, Switzerland, 1986

Fekete, B. M., C. J. Vörösmarty, W. Grabs: High Resolution Fields of Global Runoff Combining Observed River Discharge and Simulated Water Balances, *Global Biogeochemical Cycles*, 16(3), 15 1-6, 2002

Fekete, B. M., C. J. Vörösmarty, W. Grabs: Global, *Composite Runoff Fields Based on Observed River Discharge and Simulated Water Balances*, GRDC Report 22, Global Runoff Data Centre, Koblenz, Germany, 1999

Grabs, W., T. De Couet and J. Pauler: *Freshwater fluxes from the continents into the world oceans based on data of the Global Runoff Data Base*, GRDC Report 10, Global Runoff Data Centre, Koblenz Germany, 1996

Hagemann, S. and L. Dümenil, L.: A parameterization of the lateral waterflow for the global scale, *Climate Dynamics*, 14, 17-31, 1998

Krahe, P and W. Grabs: *Development of a GIS-supported Water Balance Model as a tool for the Validation of Climate Models and Hydrometeorological Data sets,* WMO-IAHS, Workshop on Continental Scale Hydrological Models: Charting the Future, November 11-13, 1996

Legates, D. R. and C. J. Willmott: Mean seasonal and spatial variability in gauge-corrected, global precipitation, *Journal of Climatology*, 10, 111-127, 1990

Graham, S. T., J. S. Famiglietti and D. R. Maidment: Five-minute, 1/2° and 1° data sets of continental watersheds and river networks for use in regional and global hydrologic and climate system modeling studies, *Water Resources Research*, 35, 583-587, 1998

Melillo, J. M., A. D. McQuire, D.W. Kicklighter, B. Moore III., C. J. Vörösmarty and A.L Schloss: *Nature*, 363, 234-240, 1993

New M., M. Hume and P. Jones: Representing Twentieth Century Space-time Climate Variability: I. Development of a 1961-1990 mean monthly terrestrial climatology, *Journal of Climatology*, 12, 829-856, 1999

New M., M. Hume and P. Jones: Representing Twentieth Century Space-time Climate Variability: II. Development of 1901-1990 monthly terrestrial monthly grids of terrestrial surface, *Journal of Climatology*, 13, 2217-2238, 2000

Oki, T. and Y. C. Sud*: Design of total runoff integrating pathways TRIP: A global river channel network*, Earth Interactions, http://EarthInteractions.org, EIO13, 1998

Rantz, S. E.: *Measurement and Computation of Streamflow: Volume 2. Computation of Discharge*, water-supply paper, 2175, U.S. Geological Survey, Reston, VA., 1982

Revenga, C., J. Brunner, N. Henninger, K. Kassem, and S. Murray: Global Freshwater Ecosystem Assessment, World Resources Institute, Washington, DC, 2000

Shuttleworth, W. J. and J. S. Wallace: Evaporation from sparse crops – an energy combination theory, *Q. J. R. Met. Soc*., 111, 839-855, 1985

UNEP/FAO/OSU: *Atlas of International Freshwater Agreements*, United Nations Environment Programme, Nairobi, Kenya, 2002

Vörösmarty C. J., B. Moore III., A. L. Grace, M. Gildea, J. M. Melillo, P. J. Peterson, E. B. Rastetter and P. A. Steudler: Continental scale models of water balance and fluvial transport: An application to South America, *Global Biogeochemical Cycles*, 3, 241-265, 1989

Vörösmarty C. J., C. A. Federer and A. L. Schloss: Potential evaporation functions compared on US watersheds: Possible implication for global-scale water balance and terrestrial ecosystem modeling, *Journal of Hydrology*, 207, 147-169, 1998

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, 2000a

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, 2000b

Vörösmarty, C. J., P. Green, J. Salisbury and Lammers, R. B.: Global water resources: Vulnerability from climate change and population growth, *Science*, 289, 284-288, 2000c

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

CRU	Climate Research Unit
DAAC	Distributed Active Archive Center
DCW	Digital Chart of the World
DEM	Digital Elevation Model
ISLSCP	International Satellite Land Surface Climatology Project
GHAAS	Global Hydrological Archive and Analysis System
GRDC	Global Runoff Data Centre

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
TEM	Terrestrial Ecosystem Model
UNH	University of New Hampshire
USGS	United States Geological Survey
WBM	Water Balance Model