USGS
science for a changing world

Technical Manual for the Geospatial Stream Flow Model (GeoSFM)

By Kwabena O. Asante, Guleid A. Artan, Shahriar Pervez, Christina Bandaragoda, and James P. Verdin

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
DIRK KEMPTHORNE, Secretary

U.S. Geological Survey
Mark D. Myers, Director

U.S. Geological Survey, Reston, Virginia 2008

Contents

Figures

Tables

Conversion Factors

SI to Inch/Pound

Multiply	By	To obtain
Length		
centimeter (cm)	0.3937	inch (in.)
millimeter (mm)	0.03937	inch (in.)
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
meter (m)	1.094	yard (yd)
Area		
square meter (m^2)	0.0002471	acre
square kilometer (km^2)	247.1	acre
square meter (m^2)	10.76	square foot (ft^2)
square kilometer (km^2)	0.3861	square mile (mi^2)
Flow rate		
cubic meter per second (m^3/s)	70.07	acre-foot per day (acre-ft/d)
meter per second (m/s)	3.281	foot per second (ft/s)
cubic meter per second (m^3/s)	35.31	cubic foot per second (ft^3/s)
cubic meter per second (m^3/s)	22.83	million gallons per day (Mgal/d)

Technical Manual for the Geospatial Stream Flow Model (GeoSFM)

By Kwabena O. Asante[1], Guleid A. Artan[1], Shahriar Pervez[1], Christina Bandaragoda[2], and James P. Verdin[3]

[1] Science Applications International Corporation, Contractor to the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD, work performed under USGS contract 03CRCN0001

[2] Water Resource Research Institute, Utah State University, Logan, UT

[3] U.S. Geological Survey

Abstract

The monitoring of wide-area hydrologic events requires the use of geospatial and time series data available in near-real time. These data sets must be manipulated into information products that speak to the location and magnitude of the event. Scientists at the U.S. Geological Survey Earth Resources Observation and Science (USGS EROS) Center have implemented a hydrologic modeling system which consists of an operational data processing system and the Geospatial Stream Flow Model (GeoSFM). The data processing system generates daily forcing evapotranspiration and precipitation data from various remotely sensed and ground-based data sources. To allow for rapid implementation in data scarce environments, widely available terrain, soil, and land cover data sets are used for model setup and initial parameter estimation. GeoSFM performs geospatial preprocessing and postprocessing tasks as well as hydrologic modeling tasks within an ArcView GIS environment. The integration of GIS routines and time series processing routines is achieved seamlessly through the use of dynamically linked libraries (DLLs) embedded within Avenue scripts. GeoSFM is run operationally to identify and map wide-area streamflow anomalies. Daily model results including daily streamflow and soil water maps are disseminated through Internet map servers, flood hazard bulletins and other media.

Introduction

During the past two decades, a number of major flooding events, such as the Midwest floods of 1993 (Kunkel and others, 1994) and the southern African floods of 2000 (Christie and Hanlon, 2001), have drawn attention to the need for wide-area monitoring systems for hydrologic

hazards. The goal of such monitoring is to provide early warning to at-risk populations and decision makers, and to direct the efforts of relief agencies involved in humanitarian assistance and sustainable development activities. In many of these areas, poor ground-based monitoring or data transmission systems hamper the efforts of local water authorities to issue spatially specific flood warnings in a timely fashion. Consequently, hydrologic modeling has been adopted as the main approach for estimating flows for warning generation. While many hydrologic models are in common use today, most of these models are unsuitable for wide-area implementation in data scarce environments. The models may not be designed to ingest geospatial data sets, an important requirement for wide-area models. The estimation of model parameters often also requires observed streamflow data, which is often not available in many parts of the world. Another common problem is the absence of a steady and reliable stream of daily forcing data sets to facilitate near-real time operation.

To help address these information needs, scientists at the U.S. Geological Survey Earth Resources Observation and Science (USGS EROS) Center, with the cooperation of the U.S. Agency for International Development, Office of Foreign Disaster Assistance (USAID/OFDA), have developed a system for operational monitoring of wide-area hydrologic hazards in data scarce environments. The system consists of a Geospatial Stream Flow Model (GeoSFM) for flow simulation and a supporting data processing system for generating daily forcing data (Artan and others, 2004). GeoSFM uses spatially distributed terrain, soil, and land cover data sets for model parameterization. It also includes hydrologic routine algorithms for data preprocessing, postprocessing, and time series manipulation. The data processing system handles the acquisition, transformation, and staging of data sets required for initial model parameterization and for daily model simulations. The integration of these spatially distributed time series of evapotranspiration and precipitation and static grids of surface elevation, soil, and land cover characteristics into an operational monitoring system is described in this document and in Asante and others (2007). This technical manual also describes the various algorithms and routines which GeoSFM uses for runoff estimation and transport.

GeoSFM Architecture

The development of the model was driven by the need to establish a common visual environment for the topographic analysis, geospatial data processing, time series manipulation, and results presentation activities that go into the monitoring of hydrologic conditions over large areas. The spatially distributed nature of the raster grids used in many of these processes pointed to the adoption of a customizable geographic information system with excellent raster functionality. The ArcView GIS software was adopted for the implementation because it provided a visual, customizable development environment with excellent support of raster operations. Figure 1 presents an overview of GeoSFM processes showing the relationship among model inputs and outputs.

Figure 1. Process map and system diagram for the Geospatial Stream Flow Model.
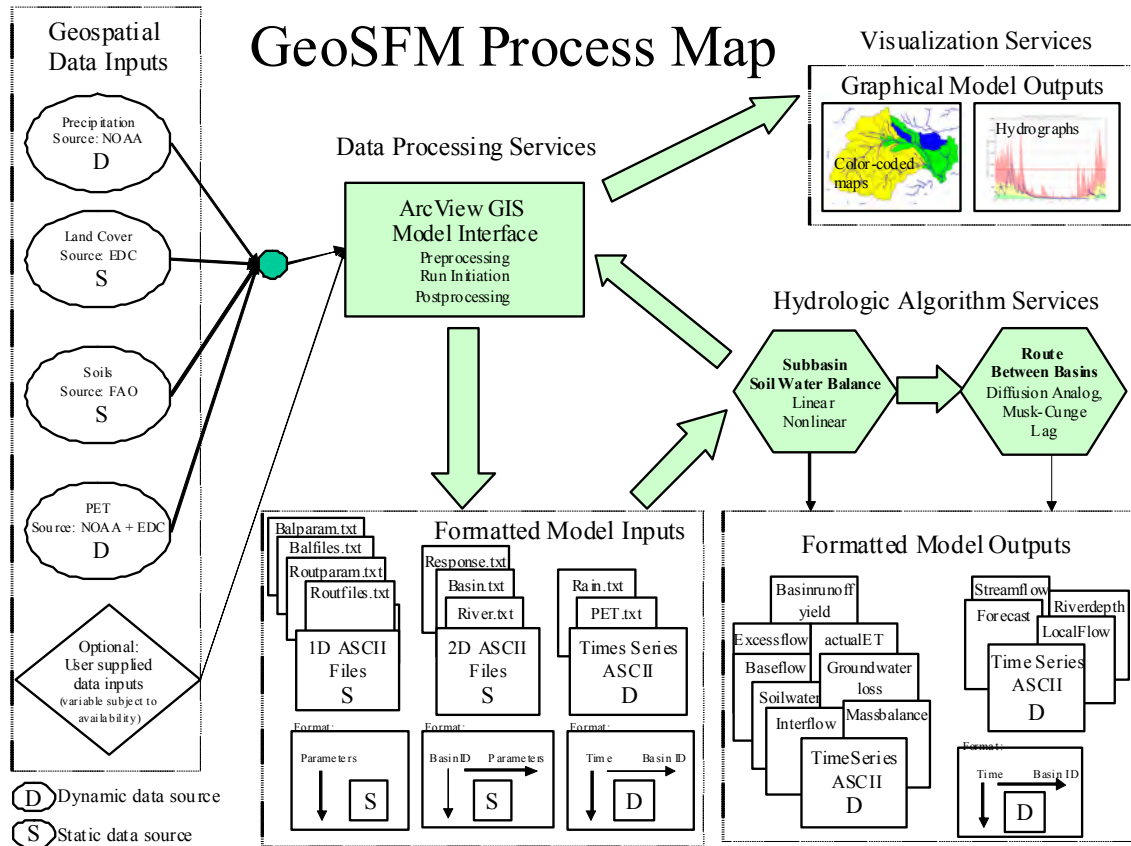
An ArcView extension was developed in Avenue (ArcView's in-built customization language) for the geospatial processing operations. GeoSFM contains preprocessing and postprocessing modules to perform such tasks. It also includes a hydrologic algorithm module for the initiation of time series analysis tasks and a calibration module which assesses the outputs from model runs for proximity to observations, modifies input files, and initiates new model runs in an iterative effort to reduce difference between model results and observations. All of these modules are initiated within ArcView to take advantage of its interface, associated tools, and input/output forms. However, the manipulation of rainfall and evapotranspiration time series to generate streamflow is better handled with compiled programming languages than with interpreted languages that are available in the GIS environment. For this reason, routines for performing the hydrologic computations involved in flow mass balance and routing operations were developed in a mixed programming environment (C/C++ and FORTRAN) and compiled as dynamically linked libraries (DLLs) using Microsoft's Visual Development Studio. The calibration module also includes both executable (*.exe) programs and DLLs.

The Visual Studio environment allows codes created by various developers in programming languages of their preference to be compiled together into a single DLL file. Procedures for calling up DLL routines were implemented from within Avenue scripts, allowing the ArcView extension to serve as a common model interface for both GIS and time series analysis. These procedures eliminate the need to develop a separate interface for hydrologic routing operations. The DLLs and

ArcView scripts work together seamlessly; the user cannot tell that a different program has been initiated. The only major problem encountered in this implementation is that errors generated within the DLL will cause the parent program (ArcView in this case) to shut down. Three measures were implemented in GeoSFM to overcome this problem. First, all data files to be used in the DLL program are screened for completeness using a series of validation checks in ArcView before a routine call is initiated to the DLL. Second, any errors occurring within the DLL are reported to test files instead of through error windows. Third, explicit exit routines are implemented within the DLL to close any open data files prior to returning program control to ArcView.

Preprocessing Module

The GeoSFM preprocessing module contains several routines for deriving flat ASCII files required by other modules from geospatial data sets available for most regions of the world. Detailed procedures for downloading and processing the input geospatial data sets from their original sources are presented in the User's Manual, which accompanies this publication. The routines in this section extract modeling units and parameters from the geospatial data sets and transcribe the results into text files for use in hydrologic computations.

Terrain Analysis Routine

The analysis of topographic data for hydrologic modeling applications relies on the simple principle that water flows in the direction of steepest descent. Hence, by comparing the elevation of a given cell with that of the eight surrounding cells, it is possible to determine which direction incident drops of water would flow. Flow direction in GeoSFM is assigned using the eight direction pour point model (Jenson and Dominque, 1988) in which each grid cell is assigned one of eight compass directions depending on which of its eight neighboring cells its discharges flow to. The computation of a flow direction grid paves the way for the determination of other parameters of hydrologic interest such as upstream contributing area, distance to the basin outlet, and the slope of the land surface.

The division of river basins into catchments with associated river reaches, and the parameterization of these modeling units from Digital Elevation Model (DEM) data have been automated in many hydrologic models including GeoSFM. While the model supports the use of DEMs of any resolution, the 1-kilometer resolution HYDRO1k DEM produced at USGS EROS is used in GeoSFM because of its global coverage. HYDRO1k is a hydrologically corrected DEM, which implies that it is devoid of spurious pits that interrupt hydraulic connectivity over the land surface. These pits are artifacts of the interpolation procedure used in the creation of DEMs, and they result in breaks in the flow network unless they are removed. HYDRO1k data in which pits have been removed using the ArcInfo Fill command are available for download at *http://edcdaac.usgs.gov/gtopo30/hydro.* The terrain analysis proceeds from the hydrologically corrected DEM to produce the following grids:

- Flow Direction Grid (FLOWDIR): A measure of direction of steepest descent from each cell location, computed using ESRI's *flowdirection* function with no special treatments of edges.

- Flow Accumulation Grid (FLOWACC): A count of cells upstream of a given cell location, computed using ESRI's *flowaccumulation* function with no weights.

- Flow Length Grid (FLOWLEN): A measure of distance from a given cell location to the nearest sink cell along the path defined by the flow direction grid, computed using ESRI's *flowlength* function with no weights.

- Stream Grid (STREAMS): A grid identifying the location of river cells, computed by assigning a value of 1 to river cells with flow accumulation grid values greater than a user-defined threshold and NODATA to all other cells.

- Stream Link Grid (STRLINKS): A grid grouping cells belonging to the same river reach with a unique value, computed using ESRI's *streamlink* function with the flow direction and stream grids as inputs.

- Outlet Grid (OUTLETS): A grid identifying the most downstream cell (a cell with the highest flow accumulation value) within each river reach in the stream link grid, computed using ESRI's *zonalmax* and *equality test* (==) grid functions.

- Basin Grid (BASINS): A grid identifying the subbasins associated with each river reach in the stream link grid, computed using ESRI's *watershed* function with the flow direction and outlet grids as inputs.

- Hill Length Grid (HILLLENGTH): A measure of distance from a given cell location to the nearest river cell along the path defined by the flow direction grid, computed using ESRI's *flowlength* function with the flow direction grid modified by assigning NODATA values to river cells.

- Slope Grid (SLOPE): A measure of the maximum rate of change of elevation around a given cell location, computed using ESRI's *slope* function and the input elevation grid with results expressed as percent rise.

- Downstream Grid (DOWNSTREAM): A grid identifying the subbasin immediately downstream of each river reach, computed using ESRI's grid neighborhood notation to query the cell next to each outlet grid cell in the direction defined by the flow direction grid.

The results of the terrain analysis are stored in ArcInfo grid format for access by other routines.

Basin Characterization Routine

In addition to the elevation derivatives, GeoSFM contains algorithms for estimating flow simulation parameters for each catchment and river reach from land cover and soil data sets. Land cover influences both the rate of runoff generation and the rate of overland flow transport. The USGS Global Land Cover Characterization (GLCC) database (Loveland and others, 2000) is used in GeoSFM because of its global coverage. The GLCC data were derived from 1-km AVHRR data and are available in the Interrupted Goode Homolosine and Lambert Equal-Area Azimuthal projections. The data are available for download at *http://edcdaac.usgs.gov/glcc/.* Land cover data are used to compute an impervious area grid (MAXCOVER) to account for the presence of water bodies in a subbasin. Land cover is also used in assigning vegetation roughness, which is subsequently used in conjunction with land surface slope to estimate overland velocity.

For estimating soil parameters such as water holding capacity (WHC), hydrological active depth (DEPTH), texture (TEXTURE), and saturated hydraulic conductivity (KC), GeoSFM uses data from the Digital Soil Map of the World (DSMW) jointly produced by the United Nations Food and Agriculture Organization (FAO) and the United Nations Educational, Scientific and Cultural Organization (UNESCO) (FAO, 1971–1981) because of its global coverage. The DSMW has 10 map sheets covering Africa, North America, Central America, South America, Europe, central Asia, northeast Asia, the Far East, Southeast Asia, and Oceania with data sets available in ArcInfo Vector Export, ERDAS Raster, and IDRISI Raster formats. All the data are in geographic

projection and are produced from original map sheets with a scale of 1:5,000,000. The Zobler data set (Zobler, 1986) is used for characterizing soil type and texture. For soil depth characterization, the Global Data Set of Soil Particle Size Properties produced by Webb and others (1993) was used. The data set specifies the top and bottom depths and the percentage of the texture type (sand, silt, and clay) of soil horizons in 106 different soil types cataloged for nine regions. Detailed procedures for extracting the required grids from these data sets are described in the User Manual, which accompanies this document, because the procedures are performed outside of GeoSFM. The expectation is that, where possible, individual users can use more accurate or alternate data sets to generate the input grids for use in GeoSFM.

The soil and land cover data sets are also used together to determine Soil Conservation Service (SCS) runoff curve numbers in GeoSFM. The curve numbers are used by some routines in GeoSFM to determine the amount of incident precipitation that becomes surface runoff. As shown in table 1, SCS curve numbers are assigned based on land cover type and the soil hydraulic classes (A, B, C, or D) of the U.S. Department of Agriculture's (USDA) soil classification system. The coarse and medium/coarse textures in the FAO data set are matched with the loam and sandy loam (USDA group A) classes. Likewise, medium and medium/fine FAO classes are matched to the USDA soil classes B and C respectively while FAO fine class is matched to the USDA class D. The intersection of the land cover and soil texture grids yields a runoff curve number (RCN) grid, which is placed in the SOILDATA directory of GeoSFM for user access.

Table 1. SCS runoff curve numbers used for various soil and land cover classes in GeoSFM.

Anderson Code	Land Cover Description	Soil Class A	Soil Class B	Soil Class C	Soil Class D
0	Unclassified	54.0	70.0	80.0	85.0
100	Urban and Built-Up Land	81.0	88.0	91.0	93.0
211	Dryland Cropland and Pasture	68.0	79.0	86.0	89.0
212	Irrigated Cropland and Pasture	62.0	71.0	78.0	81.0
213	Mixed Cropland and Pasture	65.0	75.0	82.0	85.0
280	Cropland/Grassland Mosaic	65.0	75.0	82.0	85.0
290	Cropland/Woodland Mosaic	45.0	66.0	77.0	83.0
311	Grassland	54.0	70.0	80.0	85.0
321	Shrubland	45.0	66.0	77.0	83.0
330	Mixed Shrubland/Grassland	49.5	68.0	78.5	84.0
332	Savanna	57.0	73.0	82.0	86.0
411	Deciduous Broadleaf Forest	45.0	66.0	77.0	83.0
412	Deciduous Needleleaf Forest	45.0	66.0	77.0	83.0
421	Evergreen Broadleaf Forest	25.0	55.0	70.0	77.0
422	Evergreen Needleleaf Forest	25.0	55.0	70.0	77.0
430	Mixed Forest	35.0	60.5	73.5	80.0
500	Water Bodies	98.0	98.0	98.0	98.0
620	Herbaceous Wetland	30.0	58.0	71.0	78.0
610	Wooded Wetland	25.0	55.0	70.0	77.0
770	Barren or Sparsely Vegetated	68.0	79.0	86.0	89.0
820	Herbaceous Tundra	98.0	98.0	98.0	98.0
810	Wooded Tundra	98.0	98.0	98.0	98.0
850	Mixed Tundra	98.0	98.0	98.0	98.0
830	Bare Ground Tundra	98.0	98.0	98.0	98.0
900	Snow or Ice	98.0	98.0	98.0	98.0

The basin characterization routine summarizes the elevation derivatives and soil and land cover parameters over each subbasin and exports the results into ASCII files formatted to GeoSFM specifications. For any river basin application, the basin characterization is only performed once during the model creation phase, and the resulting files are stored for access by subsequent simulation routines. The routine produces two basin characteristics files that can be manually altered as necessary to enhance model performance using any text editor. The first of these files is the basin file (basin.txt). This file contains the following characteristics required for soil moisture accounting in each subbasin:

- Soil Water Holding Capacity Parameter (SOILWHC): Contains a value in mm/m of the mean soil water holding capacity for each subbasin, computed using ESRI's *zonalmean* function with the WHCGRID and BASINS grids as inputs.

- Soil Depth Parameter (SOILDEPTH): Contains a value in cm of the mean soil depth for each subbasin, computed using ESRI's *zonalmean* function with the DEPTH and BASINS grids as inputs.

- Texture Parameter (TEXTURE): Contains a value indicating the predominant soil type (1 = sandy soil, 2 = loam, and 3 = clay) or water body (5 = water) in each subbasin, computed using ESRI's *zonalmajority* function with the TEXTURE and BASINS grids as inputs.

- Saturated Hydraulic Conductivity Parameter (KS): Contains a value of saturated hydraulic conductivity in cm/hr for the predominant soil type in each subbasin, computed using ESRI's *zonalmajority* function with the KS and the BASINS grids as inputs.

- Area of the Subbasin (AREA): Contains the area of each subbasin in km^2, computed from the BASINS grid using the count and unit area of cells within each subbasin.

- Residence Time for the Interflow Reservoir (INTERFLOWLAG): Contains an interflow residence time in days for each subbasin, computed as the time it would take water to drain from the subbasin AREA to the river if drainage occurred along the entire length of the river at the KS rate. Limits of between 2 and 120 days are imposed on the interflow residence times.

- Hill Slope Parameter (HSLOPE): Contains the average subbasin slope for each subbasin, computed as the average change in elevation between subbasin cells and their associated outlet divided by the average flow length between the same cells and outlets.

- Residence Time for the Baseflow Reservoir (BASEFLOWLAG): Contains a baseflow residence time in days for each subbasin, initially estimated as three times the residence time of the baseflow reservoir.

- Runoff Curve Number (RCNUMBER): Contains a SCS runoff curve number for each subbasin, computed using ESRI's *zonalmean* function with the RCN and the BASINS grids as inputs.

- Percent Impervious Cover (MAXCOVER): Contains a fraction of subbasin covered by a water body or other impervious area, computed by applying ESRI's *zonalsum* function with BASINS grid as the zonal grid over the STREAMS grid and the LANDCOV grid and dividing the result by the total count of cells of each subbasin.

- Local Moisture Loss Adjustment (BASINLOSS): Contains a fraction applied to subbasin soil moisture to account for local losses to regional groundwater, assigned based on predominant soil type as 0.95 for sandy soil, 0.97 for loam, 0.99 for clay, 0.98 for water bodies, and 0.99 for other undefined classes.

- Local Evapotranspiration Adjustment (PANCOEFF): Contains a coefficient applied to reference evapotranspiration to obtain potential evapotranspiration; a default value of 0.95 is assumed for all subbasins.

- Topsoil Fraction (TOPSOIL): Contains a fraction of the total soil layer that is classified as topsoil; a default value of 0.1 is assumed for all subbasins.

- Aridity Index (ARIDITY): Contains an index which is used by the Nonlinear Soil Moisture Accounting model for selecting among surface runoff mechanisms; values can range from 1 to 3 (1 = dry, 2 = normal, 3 = dry); a default value of 2 is assumed for all subbasins during initial parameterization.

The basin characterization routine also produces the river file (river.txt), which contains the following characteristics required for river flow transport in each river reach:

- Area of the River Subbasin (AREA): Contains the area in km^2 of the subbasin around each river reach, computed from the BASINS grid using the count and unit area of cells within each subbasin.

- Drainage Area of River Reach Outlet (UPAREA): Contains the drainage area in km^2 upstream of each river reach including the local subbasin, computed using ESRI's *zonalmax* function with the FLOWACC and BASINS grids as inputs.

- River Slope Parameter (RIVSLOPE): Contains the slope of the river channel, computed with ESRI's *zonalrange* function as the difference between the maximum and minimum elevation within each reach in the DEM grid divided by the difference between the maximum and minimum flow length in the FLOWLEN grid with STRLINKS defining the zones for both computations.

- River Length Parameter (RIVLENGTH): Contains the length of the river channel, computed with ESRI's *zonalrange* function as the difference between the maximum and minimum value in the FLOWLEN grid with zones defined by the STRLINKS grid.

- Downstream Reach Identifier (DOWNID): Contains the identity of the downstream river reach, computed with ESRI's *zonalmajority* function from the DOWNSTREAM grid with zones defined by the STRLINKS grid.

- Mannings Roughness Coefficient (RIVMANNING): Contains a value of Manning's roughness coefficient for the reach; a default value of 0.035 is assumed for all reaches.

- River Loss Factor (RIVLOSS): Contains a channel infiltration loss factor for each river; a default value of 1.0 (implying no loss) is assumed for all reaches.

- River Floodplain Loss Factor (RIVPLOSS): Contains a floodplain infiltration loss factor to be used for each river for flows exceeding bankfull flow; a default value of 1.0 (implying no loss) is assumed for all reaches.

- Dam Indicator (HASDAM): Contains a value that indicates the absence (0) or presence (1) of a dam within the river reach; a default value of 0 (implying no dam) is assumed for all reaches.

- Rating Curve Indicator (HASRATING): Contains a value which indicates the absence (0) or presence (1) of a flow rating curve for any location within the river reach; a default value of 0 (implying no rating curve) is assumed for all reaches.

- Flow Data Indicator (HASFLOWDATA): Contains a value which indicates the absence (0) or presence (1) of observed flow data for any location within the river reach; a default value of 0 (implying no flow data) is assumed for all reaches.

- Kinematic Wave Celerity (CELERITY): Contains a value of celerity (in m/s), which is the rate at which a flood wave is propagated through the reach; values between 0.3 m/s and 5 m/s are assigned based on RIVSLOPE and UPSTREAM (used as a surrogate for hydraulic radius).

- Dispersion Coefficient of the Diffusion Analog Equation (DIFFUSION): Contains a value of the dispersion coefficient (in m^2/s), which accounts for flow attenuation in the diffusion analog equation; values between 100 and 10,000 m^2/s are computed based on reach length and celerity by equating the second moments of the Muskingum equation (with X = 0.35) and the diffusion analog equation (Dooge, 1973).

- River Width (RIVWIDTH): Contains a value of river channel width in m, computed based on drainage area using the empirical equation of Dunne and Leopold (1978).

- Reference Flow (FLOWREF): Contains a value of reference flow for the Muskingum Cunge equation, computed based on upstream drainage area using the empirical equations of Dunne and Leopold (1978).

- Run Type Indicator (RUNTYPE): Indicates whether the simulation should proceed as a new run (0) or as a continuation of an old run (1); a default value of 0 (implying new run) is assumed for all reaches.

Unit Hydrograph Generation Routine

As a semidistributed hydrologic model, GeoSFM requires a single input value of precipitation and other forcing data for each catchment during each modeling time step. The distribution of water within the catchment is not modeled explicitly during each time step. Instead, a unit hydrograph is developed to simulate the typical response of the catchment to a uniformly distributed water input event. The unit hydrograph is developed for each catchment during the model preprocessing phase. GeoSFM contains automated GIS-based algorithms for generating these responses. The default procedure involves computing a uniform overland velocity for each catchment based on the mean slope of the catchment and dominant land cover type present. The algorithm also supports the use of nonuniform velocity grids or predetermined uniform velocity values supplied by the user. The distance along the flow path from each grid cell in the catchment to the catchment outlet is also computed. The travel time from each grid cell to the catchment outlet is also computed using ArcView's flowlength function with the FLOWDIR grid defining flow connectivity and the inverse of the VELOCITY grid serving as a weight grid. The default approach for estimating overland velocity from land cover uses Manning's equation with values of hydraulic radius assigned to each cell based on drainage area, as shown in equation 1.

$$VELOCITY = \frac{1}{MANNINGN} \cdot R_H^{2/3} \cdot \sqrt{\text{HILLSLOPE}} \tag{1}$$

where R_H is the hydraulic radius
HILLSLOPE is the average elevation change divided by the average flow length from each cell to the catchment outlet
MANNINGN is the Manning roughness for the dominant land cover in the subbasin
VELOCITY is the average overland velocity in the subbasin

The hydraulic radius values are 0.002 m, 0.005 m, 0.01 m, 0.02 m, and 0.05 m for drainage areas up to 1,000 km^2, 2,000 km^2, 3,000 km^2, 4,000 km^2, and 5,000 km^2, respectively. Cells with drainage areas greater than 5,000 km^2 are assumed to be river cells, and velocities ranging from 0.3 m/s to 1.5 m/s are directly assigned based on drainage and slope. For nonriver cells, Manning's roughness values are estimated based on the land cover type as shown in table 2.

Table 2. Manning's roughness values used for various land cover classes in GeoSFM.

Anderson Code	Description	Manning Roughness
100	Urban and Built-Up Land	0.03
211	Dryland Cropland and Pasture	0.03
212	Irrigated Cropland and Pasture	0.035
213	Mixed Dryland/Irrigated Cropland and Pasture	0.033
280	Cropland/Grassland Mosaic	0.035
290	Cropland/Woodland Mosaic	0.04
311	Grassland	0.05
321	Shrubland	0.05
330	Mixed Shrubland/Grassland	0.05
332	Savanna	0.06
411	Deciduous Broadleaf Forest	0.1
412	Deciduous Needleleaf Forest	0.1
421	Evergreen Broadleaf Forest	0.12
422	Evergreen Needleleaf Forest	0.12
430	Mixed Forest	0.1
500	Water Bodies	0.035
620	Herbaceous Wetland	0.05
610	Wooded Wetland	0.05
770	Barren or Sparsely Vegetated	0.03
>800	Tundra, Snow or Ice	0.05

The distribution of discharge at the catchment outlet is given by the probability density function (PDF) of travel times in the catchment. The PDF is analogous to the geomorphologic instantaneous unit hydrograph (GIUH) of Rodriguez-Iturbe and Valdes (1979). The unit hydrograph which is the probability mass function of flow travel times is obtained by discretizing the PDF over the routing interval. The resulting unit hydrograph is only applicable for the particular routing interval for which it was defined. For example, a change from a daily routing interval to a six hourly interval would require the definition of a new unit hydrograph response. The unit hydrograph response is computed for each catchment during preprocessing and stored in an ASCII file for subsequent use.

Weather Data Processing Routine

Precipitation data is the most essential input to any hydrologic model, and in the case of the GeoSFM, it was the driving factor in the design of the model. While ground-based gauges are

widely considered to provide the most accurate measure of precipitation at any location, the spatial coverage of such gauges is very poor in many regions of the world, and access to data from those gauges is very limited and rarely timely. Consequently, a number of satellite precipitation products are generated by meteorological agencies around the world to meet user needs for timely data. The primary rainfall data used for daily simulations in GeoSFM are satellite-derived rainfall estimates (RFE) produced by the Climate Prediction Center (CPC) of the National Oceanic and Atmospheric Administration (NOAA) (Xie and Arkin, 1997).

The NOAA RFE product uses satellite images from a variety of sources in conjunction with ground-based gauges to process rainfall estimates daily. The theoretical basis for satellite rainfall estimation derives from the fact that rainfall forms when moisture in the atmosphere is cooled to condensation. In the absence of condensation nuclei, moisture condenses at temperatures of 235 K and below. By monitoring the temperature of cloud tops from infrared imagery from geostationary satellites such as METEOSAT 7, scientists can determine areas where moisture condensation, and consequently rainfall, is occurring. Gauge measurements at experimental sites are used to obtain a first estimate of rainfall rates. Microwave imagery from sensors, such as the Special Sensor Microwave/Imager (SSM/I) of the Defense Meteorological Satellite Program and the Advanced Microwave Sounding Unit (AMSU) from NOAA, are also incorporated to estimate nonconvective rainfall. The observed rainfall data used in the NOAA RFE comes from the Global Telecommunications System (GTS) operated by the World Meteorological Organization (WMO). The system aggregates ground-based data collected by national meteorological agencies around the world. While there are not enough stations in the GTS system to support hydrologic modeling on their own, they facilitate daily ground truthing and adjustment of the satellite-derived products.

Another important model input is the daily potential evapotranspiration (PET) data produced by the Early Warning group at the USGS EROS (Verdin and Klaver, 2002). Data from the Global Data Assimilation System (GDAS) (Kanamitsu, 1989) is used to solve the Penman-Monteith equation to generate grids of PET at a daily time step. GeoSFM contains procedures for ingesting the resulting PET grids and computing actual daily evapotranspiration based on antecedent soil moisture conditions. The Penman-Monteith equation takes into consideration weather conditions including air temperature, fluxes of long and short wave radiation, actual and saturation vapor pressure, relative humidity, and wind speed at 2 m above the ground. A reference crop height of 0.12 m, a fixed surface resistance of 70 s m^{-1}, and an albedo of 0.23 are assumed. A daily PET grid with a spatial resolution of 100 by 100 km and global coverage is produced and distributed via the USGS EROS FTP site at ftp://edcftp.cr.usgs.gov/pub/edcuser/fewsips/global.

GeoSFM contains algorithms for computing mean areal precipitation (MAR) and mean areal evapotranspiration (MAE) values for each catchment from these grids. To correctly associate a given grid with the time interval for which it was computed, the algorithms rely on a grid file naming convention which includes the year and Julian day. GIS routines are used to identify and average the grid cells within each catchment, and the results are stored in ASCII files. For areas over which spatially distributed precipitation data is not available, data interpolation routines are provided for converting station readings into a continuous surface. The routines include inverse distance weighting, spline, and kriging interpolation methods. A generation of evapotranspiration time series in ASCII format is the final step before the initiation of hydrologic model runs.

Soil Moisture Accounting Module

As a continuous simulation model, GeoSFM contains routines for computing runoff and soil moisture conditions daily. The two options provided for performing this accounting in GeoSFM are the Linear Soil Moisture Accounting (LSMA) and the Nonlinear Soil Moisture Accounting

(NSMA) modules. The two modules have been coded in a mixed programming environment and compiled as dynamically linked libraries (DLLs). Avenue code has been developed to call up and execute the DLL routines from within ArcView. Consequently, the user is able to perform either of the SMA runs from within the GIS environment. Irrespective of the choice of soil moisture accounting model, GeoSFM generates a series of ASCII files containing the surface, interflow, baseflow, and percolation fluxes as well as the soil moisture storage. The output files are formatted for ingestion by the channel routing routines in GeoSFM. The features of the models are described below.

Linear Soil Moisture Accounting (LSMA) Routine

The LSMA routine is a bucket model in which surface runoff is generated with a partial contributing area formulation, interflow is generated from the bucket, and baseflow is generated from an unbounded storage below the bucket. The soil layer acts as a single control volume with rainfall as the only input, and evapotranspiration, surface runoff, subsurface runoff, and deep groundwater percolation as outputs. The relative simplicity of this model makes it easy to implement and parameterize in a wide variety of settings with little or no calibration. However, this model is less well suited to high spatial or temporal resolution applications or to settings where a model well calibrated with observed data is required. In such settings, the nonlinear module may be a better option.

Table 3. List of key variables for the Linear Soil Moisture Accounting routine.

Item	Type	Variables Description	Comments
B	Integer	Number of subbasins	Source = balparam.txt
T	Integer	Simulation time steps	Source = balparam.txt
R	Integer	Response time steps	Source = balparam.txt
STARTDAY	Real	Simulation start day	Source = balparam.txt
STARTYR	Real	Simulation start year	Source = balparam.txt
LDAY	Integer	Simulation end day	Source = balparam.txt
LYEAR	Integer	Simulation end year	Source = balparam.txt
INIFRACTION	Real	Infiltration fraction	Source = balparam.txt
BALMODE	Real	Run initialization mode	Source = balparam.txt
INTERVAL	Real	Interval	Source = balparam.txt
INHOURS	Real	Simulation time step duration in hours	Source = balparam.txt
HSLOPE	Real	Hill slope	Source = basin.txt
KS	Real	Hydraulic conductivity	Source = basin.txt
TOPSOIL	Real	Topsoil as fraction of full soil layer	Source = basin.txt
PANCOEF	Real	Evapotranspiration Pan coefficient	Source = basin.txt
IMPRATIO	Real	Impervious cover ratio	Source = basin.txt
SOILLAG	Real	Interflow residence time in days	Source = basin.txt
GWLOSSTOTAL	Real	Total groundwater loss	Source = basin.txt
GWLOSSFACTOR	Real	Groundwater loss factor	Source = basin.txt
GWLAG	Real	Groundwater residence time in days	Source = basin.txt
STMAX	Real	Maximum storage	Source = Computations
STTEMP	Real	Temporary storage	Source = Computations
AEVAL	Real	Evapotranspiration	Source = Computations
REVAL	Real	Residual evapotranspiration	Source = Computations

RESPLIST	Character	Response file header	Source = response.txt
BASLIST	Integer	Basin file header	Source = basin.txt
TIMELIST	Integer	Time step listing	Source = rain.txt
WORKDIR	Character	Working directory	Source = balfiles.txt
		Arrays	
RAIN(B,T)	Real	Mean Areal Precipitation	Input from rain.txt
EVAP(B,T)	Real	Mean Areal Evapotranspiration	Input from evap.txt
BASIN(B,12)	Real	Subbasin Characteristics	Input from basin.txt
RESPONSE(B,R)	Real	Unit Hydrograph File	Input from response.txt
LOCALFLOW(B,T)	Real	Local contribution to flow	Output to localflow.txt
AEVAP(B,T)	Real	Actual evapotranspiration	Output to aevap.txt
GWLOSS(B,T)	Real	Water loss to groundwater	Output to gwloss.txt
SURP(B,T)	Real	Total runoff generated daily	Output to basinrunoffyield.txt
EXCESSFLOW(B,T)	Real	Surface runoff contribution	Output to excess.txt
BASEFLOW(B,T)	Real	Baseflow runoff contribution	Output to baseflow.txt
INTFLOW(B,T)	Real	Interflow runoff contribution	Output to interflow.txt
STORE(B,T)	Real	Total soil water storage	Output to soilwater.txt
GWSTORE(B,T)	Real	Groundwater store	Temporary Variable
REVAP(B,T)	Real	Residual Evapotranspiration	Temporary Variable
		Input-Output Files	
RAINFN	Input	Subbasin mean areal rainfall file	Default = Rain.txt, Source = balfiles.txt
EVAPFN	Input	Mean areal evapotranspiration file	Default = Evap.txt, Source = balfiles.txt
BASINFN	Input	Subbasin characteristics file	Default = Basin.txt, Source = balfiles.txt
RESPFN	Input	Subbasin unit hydrograph file	Default = Response.txt, Source = balfiles.txt
SURPFN	Output	Total runoff yield file	Default = Basinrunoffyied.txt, Source = balfiles.txt
STOREFN	Output	Soil water storage file	Default = Soilwater.txt, Source = balfiles.txt
AEVAPFN	Output	Actual evapotranspiration file	Default = Actualevap.txt, Source = balfiles.txt
GWLOSSFN	Output	Regional groundwater loss file	Default = Gwloss.txt, Source = balfiles.txt
OUTSWFN	Output	Final soil moisture file	Default = Cswater.txt, Source = balfiles.txt
LOGFN	Output	Runtime messages and errors file	Default = Logfilesoil.txt, Source = balfiles.txt
EXCESSFN	Output	Excess precipitation file	Default = Excessflow.txt, Source = balfiles.txt
BASEFN	Output	Baseflow file	Default = Baseflow.txt, Source = balfiles.txt
INTFN	Output	Interflow file	Default = Interflow.txt, Source = balfiles.txt
PARAMFN	Input	Listing of simulation parameters	Default = Balparam.txt, Fixed File Name
BALFILESFN	Input	Listing of simulation files	Default = Balfiles.txt, Fixed File Name
DEFAULTFN	Output	Default methods file	Default = Default.txt, Fixed File Name

The FORTRAN language script, 1DBalance.for, computes the LSMA as documented below. The listing of variables and input and output files used in this script are presented in table 3. At the start of a soil moisture accounting simulation, initiate values of storage state variables are required for each subbasin. The maximum storage of each subbasin, STMAX(B), is computed by multiplying its soil water holding capacity, WHC(B), by the depth of the soil column, SOILDEPTH(B), as shown in equation 2. A conversion factor of 0.01 is applied to obtain a result in mm because WHC(B) is in "mm/m" and SOILDEPTH(B) is in "cm" in the basin file (basin.txt).

$$STMAX(B) = WHC(B) \cdot SOILDEPTH(B) \cdot 0.01 \qquad\qquad (2)$$

An initial estimate of moisture storage at the beginning of the simulation run is required. In GeoSFM, a user-supplied estimate of initial soil moisture fraction is used to make this determination. A software default of 0.1 (implying 10 percent moisture content) is provided, but the

user may change this estimate by providing a value between 0 and 1 prior to initiation of the simulation run. Given this fraction, the initial soil moisture storage is computed by multiplying maximum storage by the infiltration fraction (equation 3).

$$\text{STORE(B,0)} = \text{STMAX(B)} \cdot \text{INIFRACTION} \tag{3}$$

Note that a single value of initial fraction is currently applied throughout the simulation area. Initial storage in the groundwater reservoir, GWSTORE(B,0), is assumed to be zero. Some units are converted during model initiation to ensure consistency of computations. For example, the saturated hydraulic conductivity of the soil in each subbasin, KS(B), is provided in the basin file, basin.txt, in cm/h. It is converted to m/day by multiplying by 0.024. Likewise, the Hill Slope parameter provided in the basin file is converted from a percent slope to a fraction by dividing by 100.

A fraction of incident rainfall becomes surface runoff before entering the control volume. The magnitude of this fraction is directly proportional to the ratio of the fraction of the subbasin that is impervious. The partial contributing area fraction (PARATIO) is updated at every simulation time step based on total soil moisture storage at the beginning of the time step by assuming a linear relationship between the depth of water in subsurface storage exceeding the soil depth and the partial contributing area, subject to a maximum of 1.0, as shown in equation 4.

$$\text{PARATIO(B,T)} = \text{MIN} \left\{ \begin{array}{c} \text{Max}\left[\left(\dfrac{\text{GWSTORE(B,T-1)} - \text{SOILDEPTH}}{\text{HILLSLOPE} \cdot \text{HILLLENGTH}}\right), 0\right] \\ \\ (1 - \text{IMPRATIO}) \end{array} \right\} \tag{4}$$

This relationship is based on two assumptions. The first assumption that the land surface is V-shaped allows the hill slope and hill length to be used in determining the average elevation of the watershed divide above the river. The second assumption is that the groundwater table is directly below the soil layer used for moisture accounting. Hence, by comparing the average elevation with the subsurface storage less than the soil depth, it is possible to determine what portion of the subbasin surface area is being saturated by subsurface storage.

Once impervious area ratio has been updated, the excess precipitation (EXCESSRAIN) is computed as the fraction of rainfall (RAIN) landing on the permanent impervious area and the partial contributing area created by the saturation from subsurface storage. The remainder of the rainfall is allowed to enter the soil layer where it either meets the evapotranspiration requirements or augments the subsurface reservoir, STORE(B,T), from which interflow is generated. Additional excess precipitation is generated when the soil storage capacity, STMAX(B), is exceeded, as presented in equation 5.

$$\text{EXCESSRAIN(B,T)} = \text{MAX} \left\{ \begin{array}{c} \left[\text{STORE(B,T-1)} + \text{RAIN(B,T)} - \text{STMAX(B)}\right] \\ \\ \text{RAIN(B,T)} \cdot \left[\text{PARATIO(B,T)} + \text{IMPRATIO}\right] \end{array} \right\} \tag{5}$$

Potential evapotranspiration is first adjusted by a pan coefficient (PANCOEF) to account for known biases in satellite or pan-based estimates supplied in the evap.txt file. Where no known biases exist, a pan coefficient of 1 is used. Values of pan coefficients for each subbasin are

specified in the basin characteristic file, basin.txt. The conversion of potential to actual evapotranspiration is based on the availability of moisture from rainfall, soil moisture, or runoff sources to satisfy this demand (equation 6).

$$AEVAP(B,T) = MIN\begin{Bmatrix} EVAP(B,T) \\ STORE(B,T-1) + RAIN(B,T) - EXCESSRAIN(B,T) \end{Bmatrix} \quad (6)$$

To ensure conservation of mass is maintained throughout the computation, a temporary storage term, STORETEMP, is used to track the moisture remaining in storage after each flux extraction or addition during the simulation time step, as shown in equation 7.

$$STORETEMP = STORE(B,T-1) + RAIN(B,T) - EXCESSRAIN(B,T) - AEVAP(B,T) \quad (7)$$

A first-order exponential loss function is used to compute the amount of water percolating from the bounded soil bucket, BFLOSS(B,T). The loss rate is computed based on storage, soil depth, and the saturated hydraulic conductivity (equation 8).

$$BFLOSS(B,T) = \left(\frac{STORETEMP \cdot KS(B)}{SOILDEPTH(B)} \right) \cdot \exp\left(- \frac{KS(B)}{SOILDEPTH(B)} \right) \quad (8)$$

The water enters the groundwater storage from which baseflow is generated. After this extraction, the residual storage within the soil layer is updated to reflect current moisture conditions (equation 9).

$$STORETEMP = \begin{pmatrix} STORE(B,T-1) + RAIN(B,T) - EXCESSRAIN(B,T) \\ - AEVAP(B,T) - BFLOSS(B,T) \end{pmatrix} \quad (9)$$

Percolation to regional groundwater is similarly computed using a linear loss function but with the groundwater residence time supplied in the basin file (equation 10).

$$GWLOSS(B,T) = \left(\frac{(GWSTORE(B,T-1) + BFLOSS(B,T))}{GWLAG(B)} \right) \cdot \exp\left(- \frac{1}{GWLAG(B)} \right) \quad (10)$$

A linear reservoir formulation is used to compute the amount of water from the bounded soil bucket that is released to the surface water network of rivers and lakes as interflow, INTERFLOW(B,T). The residence time required for this computation is supplied in the basin file as the interflow lag time (equation 11).

$$INTERFLOW(B,T) = \left(\frac{STORETEMP}{INTLAG(B)} \right) \cdot \exp\left(- \frac{1}{INTLAG(B)} \right) \quad (11)$$

After estimating interflow, residual moisture in the soil layer is computed for use in the next simulation time step, using equation 12:

$$STORE(B,T) = \begin{pmatrix} STORE(B,T-1) + RAIN(B,T) - EXCESSRAIN(B,T) \\ - AEVAP(B,T) - BFLOSS(B,T) - INTERFLOW(B,T) \end{pmatrix} \quad (12)$$

Baseflow contribution to surface runoff is also estimated with a linear reservoir function. As with the interflow, the residence time is supplied in the basin file as the groundwater lag time (equation 13).

$$BASEFLOW(B,T) = \left(\frac{(GWSTORE(B,T-1) + BFLOSS(B,T) - GWLOSS(B,T))}{GWLAG(B)} \right) \cdot \exp\left(-\frac{1}{GWLAG(B)} \right) \quad (13)$$

After estimating interflow, residual moisture in the groundwater reservoir is computed for use in the next simulation time step (equation 14).

$$GWSTORE(B,T) = \begin{pmatrix} GWSTORE(B,T-1) + BFLOSS(B,T) \\ - BASEFLOW(B,T) - GWLOSS(B,T) \end{pmatrix} \quad (14)$$

The moisture accounting for the time step ends with the fluxes and storage computed. The entire processing sequence is repeated for subsequent simulation time steps until the final rainfall input has been processed. An additional 100 time steps of simulations is performed assuming no additional rainfall input. The additional simulations allow runoff to be generated from the moisture in soil storage for flow forecasting applications. The resulting moisture fluxes and storage are written to ASCII files for postprocessing or ingestion into the flow routing modules.

Nonlinear Soil Moisture Accounting (NSMA) Routine

The NSMA module provides a more complete representation of subsurface processes by creating separate soil layers within which interflow and baseflow processes occur. The SCS runoff curve number method, which takes into account vegetation, soil type, and antecedent moisture, is used for separating runoff into surface and interflow components, while the Green-Ampt equation (Green and Ampt, 1911) is used for extracting water from the interflow soil layer to feed the baseflow soil layer. In addition to the parameters from the single-layer model, parameters describing the fraction of the soil column that is attributable to the respective interflow and baseflow layers, SCS runoff curve numbers, and Green-Ampt parameters are required for the two-layer model. The superior representation of land surface and soil moisture processes makes the two-layer model better suited for application in areas where finer resolution data is available to support its parameterization. The variables used in this script can be grouped under three broad categories: input variables, model state variables, and model output variables, as presented in table 4.

Table 4. List of key variables for the Nonlinear Soil Moisture Accounting routine.

Variable Group	Variable Name	Type	Description
		Input	
PERIOD	DAY	Integer	Day of the month
	DOY	Integer	Julian day of the year
	YEAR	Integer	Year of the dam release
	MONTH	Integer	Month of the dam release
	HOUR	Integer	Hour of the release

SOIL	SWCX	Float	Saturation soil water content (mm)
	SWCM	Float	Minimum soil water content (mm)
	SWCF	Float	Soil water content at field capacity (mm)
	WHC	Float	Water holding capacity (mm)
	USD	Float	Upper soil layer depth, percent of the total active soil
	LSD	Float	Lower soil layer depth, percent of the total active soil
	TSD	Float	Total soil, upper and lower, layer depth (cm)
	Ks	Float	Saturation soil hydraulic conductivity (cm)
	Bp	Float	Bubbling pressure (m)
	GAMMA	Float	Pore-size index
	TEXTURE	Integer	Soil texture; 1 = sandy soil, 2 = loam, and 3 = clay
LANDFORM	HLSLP	Float	Mean slope of the basin
	AREA	Float	Basin area (km²)
	UPAREA	Float	Basin outlet upstream watershed area (km²)
LANDUSE	CN	Float	Basin mean SCS runoff curve number
	XIA	Float	Maximum fraction of the basin area that can be impervious
	MIA	Float	Fraction of the basin area that is permanently impervious
METEOROLOGY	PRECP	Float	Time series of observed basin daily mean areal rainfall (mm)
	PET	Float	Time series of observed basin daily evapotranspiration (mm)
	Kc	Float	Crop coefficient, between 0 and 1
MODELPARAM	LAGSTR	Float	Linear reservoir routing constant for overland runoff
	LAGGWT	Float	Linear reservoir routing constant for groundwater runoff
	LAGIRF	Float	Linear reservoir routing constant for interflow runoff
	LGRW	Float	Fraction of the baseflow loss to the regional groundwater flow
	GUH	Float	Unit hydrograph parameters
	MOISTDISTR	Float	Represents the degree of areal nonuniform distribution of tension water capacity, Xinanjiang model
	RMODE	Integer	Excess rainfall calculation mode, 1 = Philips equation, 2 = SCS CN method, and 3 = excess moisture to soil water deficit
	NDAY	Integer	Maximum convolution days for the unit response function
		State	
PROGNOSTIC	USWC	Float	Upper soil water content (mm)
	dUSWC	Float	Change in upper soil water content (mm)
	LSWC	Float	Lower soil water content (mm)
	dLSWC	Float	Change in lower soil water content (mm)
	X	Float	Retention rainfall (mm)
	FGWS	Float	Conceptual reservoir for 1st groundwater reservoir level (mm)
	SGWS	Float	Conceptual reservoir for 2nd groundwater reservoir level (mm)
	INTERFLOWSTOR1	Float	1st conceptual reservoir for soil moisture storage (mm)
	INTERFLOWSTOR2	Float	2nd conceptual reservoir to describe soil moisture storage (mm)
	BASINIMPERVAREA	Float	Percent of basin area impervious at any time
	SURFACEWATER	Float	Excess precipitation of the past days (mm)
		Output	
RUNOFF	EXPRP	Float	Excess rainfall (mm)
	DRNFF	Float	Surface runoff from excess rainfall and impervious area (mm)
	IFLW	Float	Interflow (mm)
	GWL	Float	Losses to the regional groundwater flow (mm)
	FBS	Float	Fast baseflow flux (mm)
	SBS	Float	Slow baseflow flux (mm)

	TRFLW	Float	Total local runoff at the basin outlet (mm)
	AET	Float	Actual evapotranspiration (mm)
STATISTICS	FLOW [12] [500]	Float	Streamflow sorted by month (m³/sec)
	DAYS [12]	Integer	Number of days in each month of the runs
TOTALFLUXES	SUMET	Float	Total Evapotranspiration
	SUMRAIN	Float	Total precipitation
	SUMRUNOFF	Float	Total runoff
	SUMBASEFLOW	Float	Total baseflow
	SUMLOSSGROUND	Float	Total groundwater loss
	SUMINTERFLOW	Float	Total interflow

Excess precipitation (EXPRP) is computed based on precipitation input (PRECP) using one of three infiltration equations (Philip's infiltration equation, the SCS curve number method, and a simple bucket model) depending on what is specified in the basin file. The default is the SCS curve number method. At each time step, the default curve number is adjusted based on the percentage saturation of the upper soil layer ($SWCX_{UpperLayer}$), as shown in equation 15 (Woolhizer, 1976).

$$
CN = \begin{cases}
1.95 \cdot CN_0 \cdot 2.718282^{(-0.00663 \cdot CN_0)} & (0.9 < SWCX_{UpperLayer} < 1) \\
CN_0 & (0.6 <= SWCX_{UpperLayer} <= 0.9) \\
0.39 \cdot CN_0 \cdot 2.718282^{(0.009 \cdot CN_0)} & (0 < SWCX_{UpperLayer} < 0.6)
\end{cases} \tag{15}
$$

The adjusted curve number is used to generate excess precipitation when the daily rainfall (PRECP) is higher than 20 percent of SCS runoff, using equation 16.

$$
RUNOFF.EXPRP = \frac{\left(PRECIP - 0.2 \cdot \left[\left(\frac{1000}{CN} \right) - 10 \right] \right)^2}{\left(PRECIP + 0.8 \cdot \left[\left(\frac{1000}{CN} \right) - 10 \right] \right)} \tag{16}
$$

To simulate variable source area for surface runoff abstraction, the impervious area is allowed to vary during each time step between a minimum that equals the stream and open water areas and a predetermined maximum. Impervious runoff area (BASINIMPERVAREA) is computed as a function of the upper zone water content (USWC). If the USWC is less then 90 percent of the field capacity of the topsoil ($SWCF_{UpperLayer}$), then the impervious runoff generating area is equal to the minimum impervious area in the basin (MIA). Otherwise, the impervious runoff area is a function of the maximum impervious area (XIA), the saturation capacity of the topsoil ($SWCX_{UpperLayer}$), and the USWC, as shown in equation 17.

$$
BASINIMPERVAREA = MIA + \left((XIA - MIA) \cdot \left(\frac{SWC_{UpperLayer} - (0.9 \cdot SWC_{UpperLayer} \cdot USD)}{SWCX_{UpperLayer} - (0.9 \cdot SWCF_{UpperLayer})} \right) \right) \tag{17}
$$

Evapotranspiration from the upper soil layer is limited to 20 percent or less of the potential evapotranspiration. When the water content in the upper soil layer is less than 70 percent of field capacity, the evapotranspiration is further reduced by a ratio of the current and maximum residual water holding capacity (equation 18).

$$AET_{UpperSoil} = \begin{cases} \left(\dfrac{(USWC - (SWCM \cdot USD)) \cdot PET \cdot 0.2}{(0.7 \cdot SWCF \cdot USD) - (SWCM \cdot USD)} \right) & USWC < 0.7 \\[2em] PET \cdot 0.2 & USWC >= 0.7 \end{cases} \quad (18)$$

After the actual evapotranspiration is removed from the upper layer, the soil water content in the layer is updated by adding nonexcess precipitation and subtracting evapotranspiration (equation 19).

$$USWC = USWC + PRECP - AET_{UpperSoil} - EXPRP \qquad (19)$$

Validation checks are included to ensure that USWC is greater than or equal to the upper soil layer wilting point but less than its field capacity. Any excess moisture above the upper soil layer field capacity becomes available for infiltration or percolation using equation 20.

$$ESW_{UpperLayer} = \begin{cases} (USWC - (WHC \cdot USD)) & (USWC > (SWHC \cdot USD)) \\[1.5em] 0 & (USWC <= (WHC \cdot USD)) \end{cases} \quad (20)$$

Water in the topsoil layer exceeding the upper soil layer field capacity is available for either interflow or percolation. The rate at which water moves out of the upper soil layer is characterized by the sorptivity (SORP), which is computed as a function of soil properties and water content (equation 21) (Brooks and Corey, 1964).

$$SORP_{UpperLayer} = \sqrt{2 \cdot K_s \cdot (SWCX_{UpperLayer} - USWC) \cdot \left(\frac{(2 + 3 \cdot GAMMA)}{((1 + 3 \cdot GAMMA) \cdot Bp \cdot 0.5)} \right)} \quad (21)$$

The fractions that actually will be interflow (equation 22) and percolation (equation 23) are computed based on sorptivity SORP (from Green-Ampt), hydraulic conductivity KS, and basin slope HlSlp. In general, basins with little slope have only a small fraction of interflow and a high fraction of percolation. In most cases, the fraction for percolation will be larger than the fraction for interflow because slopes are typically much smaller than 45° (both fractions will be 0.5 for a slope of 45°).

$$INTERFLOW_{UpperLayer} = \left(\frac{\sin(HlSlp)}{(\sin(HlSlp)) + (\cos(HlSlp))} \right) \cdot ESW_{UpperLayer} \qquad (22)$$

$$PERCOLATION_{UpperLayer} = \left(\frac{\cos(HlSlp)}{(\sin(HlSlp)) + (\cos(HlSlp))} \right) \cdot ESW_{UpperLayer} \qquad (23)$$

Transpiration from the lower soil layer amounts to 80 percent or less of the potential evapotranspiration (PET). The potential evapotranspiration is further reduced by the ratio of current to maximum available water content in the lower layer if the current lower soil water content (LSWC) is less than 70 percent of field capacity (equation 24).

$$
AET_{LowerSoil} = \begin{cases} \left(\dfrac{(LSWC - (WHC \cdot LSD) \cdot PET \cdot 0.8}{(0.7 \cdot SWCF \cdot LSD) - (SWCM \cdot LSD)} \right) & LSWC < 0.7 \\[2em] PET \cdot 0.8 & LSWC >= 0.7 \end{cases} \quad (24)
$$

The lower soil moisture content is updated by adding percolating flows from the upper layer and subtracting transpiration from the lower soil layer, with equation 25.

$$
LSWC = LSWC + PERCOLATION_{UpperLayer} - AET_{LowerSoil} \quad (25)
$$

As with the upper layer, validation checks are used to ensure that lower soil layer moisture content does not fall below its soil wilting point but stays below field capacity. Any moisture in excess of field capacity of the lower soil layer goes into one of three storage components: groundwater, fast interflow storage, and slow interflow storage. A simple groundwater loss fraction (LGRW) is used to determine how much water of the excess is released to increase regional groundwater flow (equation 26). The loss only occurs when water content in the lower soil layer exceeds its water holding capacity. For time periods when no excess is available, the groundwater storage remains unchanged.

$$
GWSTORE = GWSTORE + \left(MAX \begin{cases} ((LSWC - (WHC \cdot LSD)) \cdot LGRW) \\[1.5em] 0 \end{cases} \right) \quad (26)
$$

Of the remainder, 25 percent enters a slow response interflow storage ($INTERFLOWSTOR_{LowerLayer}^{Slow}$), while the remaining 75 percent enters a fast response interflow storage ($INTERFLOWSTOR_{LowerLayer}^{Fast}$), as shown in equations 27 and 28.

$$
INTERFLOWSTOR_{LowerLayer}^{Fast} = INTERFLOWSTOR_{LowerLayer}^{Fast} + (EWC_{LowerLayer} - GWLOSS) \cdot 0.75 \quad (27)
$$

$$
INTERFLOWSTOR_{LowerLayer}^{Slow} = INTERFLOWSTOR_{LowerLayer}^{Slow} + (EWC_{LowerLayer} - GWLOSS) \cdot 0.25 \quad (28)
$$

Water is gradually released from the two interflow and baseflow reservoirs for transfer to the nearest stream, downstream through the stream network, and eventually to a terminal point in an ocean or inland lake. The rate of release of water from each of these reservoirs is governed by a linear response function with residence times obtained from the input river characteristics file. The responses for the fast and slow interflow and baseflow reservoirs are given in equations 29, 30, and 31.

$$
INTERFLOW_{LowerLayer}^{Fast} = (INTERFLOWSTOR_{LowerLayer}^{Fast} \ / \ LAGIRF) \quad (29)
$$

$$
INTERFLOW_{LowerLayer}^{Slow} = (INTERFLOWSTOR_{LowerLayer}^{Slow} \ / \ (0.5 \cdot LAGIRF)) \quad (30)
$$

$$
BASEFLOW = (GWSTORE \ / \ LAGGWT) \quad (31)
$$

Descriptions of the routines available in GeoSFM for river transport are provided in the following sections.

River Transport Module

After runoff generated within a given catchment travels to the catchment outlet, it enters the river network and works its way downstream to the basin outlet. GeoSFM supports two linear routines (pure lag and the diffusion analog) and one nonlinear method, (the Muskingum Cunge) for simulating the in-channel phase of flow.

Pure Lag Flow Routine

Lag routing is the simplest of the river transport approaches. It accounts for the advection of flow but does not include any attenuation or deformation of the input. Consequently, the input flow remains unaltered in magnitude at the discharge point, and a single flow parameter, travel time, between the input and discharge locations is required for the computation. The FORTRAN script, Lagroute.for, retrieves runoff generated by the soil water balance in each subbasin and routes it through the river network by pure translation as documented below. The variables, input files, and output files used in this program are presented in table 5.

Table 5. List of key variables for the pure lag flow routine.

Name	Type	Description	Comments
		Variables	
B	Integer	Number of subbasins/rivers	Source = routparam.txt
T	Integer	Simulation time steps	Source = routparam.txt
R	Integer	Response time steps	Source = routparam.txt
F	Integer	Forecast time steps	Source = routparam.txt
STARTDAY	Real	Simulation start day	Source = routparam.txt
STARTYR	Real	Simulation start year	Source = routparam.txt
LDAY	Integer	Simulation end day	Source = routparam.txt
LYEAR	Integer	Simulation end year	Source = routparam.txt
INHOURS	Real	Simulation time step duration in hours	Source = routparam.txt
AREA	Real	River reach subbasin area	Source = river.txt
UPAREA	Real	Total drainage area at river reach outlet	Source = river.txt
DOWNID	Integer	Subbasin downstream of river reach	Source = river.txt
RIVLENGTH	Real	Length of the river reach	Source = river.txt
RIVMANNING	Real	Manning roughness coefficient for reach	Source = river.txt
RIVERLOSS	Real	In-Streamflow loss coefficient	Source = river.txt
RIVERPLOSS	Real	Overbank flow loss coefficient	Source = river.txt
HASDAM	Integer	River reach dam flag	Source = river.txt
CELERITY	Real	Flow celerity for river reach	Source = river.txt
		Arrays	
SURP(B, T)	Real	Total runoff generated daily	Input from basinrunoffyield.txt
RIVER(B, 12)	Real	River Characteristics File	Input from river.txt
RESPONSE(B, R)	Real	Unit Hydrograph File	Input from response.txt
STREAMFLOW(B, T+F)	Real	Streamflow File	Output to streamflow.txt
LOCALFLOW(B,T+F)	Real	Local contribution to flow	Output to localflow.txt

INFLOW(B, T+F)	Real	UpStreamflow into a subbasin	Output to inflow.txt
FORECAST(B, F)	Real	Streamflow Forecast File	Output to forecast.txt

Input and Output Files

SURPFN	Output	Total runoff yield file	Default = Basinrunoffyied.txt, Source = balfiles.txt
RIVERFN	Input	River characteristics file	Default = River.txt, Source = balfiles.txt
RESPFN	Input	Subbasin unit hydrograph file	Default = Response.txt, Source = balfiles.txt
RESERVOIRFN	Output	File containing reservoir listing	Default = Basinrunoffyied.txt, Source = balfiles.txt
STOREFN	Output	Soil water storage file	Default = Soilwater.txt, Source = balfiles.txt
AEVAPFN	Output	Actual evapotranspiration file	Default = Actualevap.txt, Source = balfiles.txt
GWLOSSFN	Output	Regional groundwater loss file	Default = Gwloss.txt, Source = balfiles.txt
OUTSWFN	Output	Final soil moisture file	Default = Cswater.txt, Source = balfiles.txt
LOGFN	Output	Runtime messages and errors file	Default = Logfilesoil.txt, Source = balfiles.txt
EXCESSFN	Output	Excess precipitation file	Default = Excessflow.txt, Source = balfiles.txt
BASEFN	Output	Baseflow file	Default = Baseflow.txt, Source = balfiles.txt
INTFN	Output	Interflow file	Default = Interflow.txt, Source = balfiles.txt
PARAMFN	Input	Listing of simulation parameters	Default = Balparam.txt, Fixed File Name
BALFILESFN	Input	Listing of simulation files	Default = Balfiles.txt, Fixed File Name
DEFAULTFN	Output	Default methods file	Default = Default.txt, Fixed File Name

In lag routing, inflow into the upstream end of a river reach is translated with a time delay but without change in amplitude to the downstream end of the same reach. The flow is then transferred to the next downstream river reach. This routing approach is easily implemented in data sparse settings because it only requires a single parameter, the lag time. Mathematically, lag routing can be expressed as equation 32.

$$Q(t) = I(t - t') \tag{32}$$

Where $Q(t)$ is the discharge at time, t
t' is the travel time between input and discharge locations
$I(t-t')$ is the input at time $(t-t')$

The routing process begins by reading in the runoff generated within the subbasin associated with each river reach from the runoff yield file (SURPFN) into the runoff yield array (SURP(B, T)). A set of parameters associated with each river reach is also read from the river characteristics file (RIVERFN) into the river characteristics array (RIVER(B, 12)). The characteristics contained in this array include RIVERID of the next downstream reach, indicators of the presence of dams along the reach. If a dam is present in the basin, additional files—a dam link file and a reservoir characteristics file, or a file containing prespecified discharge values—are opened for each dam. The surplus runoff from the various sources, including excess precipitation, interflow, and baseflow, are aggregated (equation 33).

$$SURP(B, T) = EXCESSRAIN(B, T) + INTERFLOW(B, T) + BASEFLOW(B, T) \tag{33}$$

The surplus runoff is converted from depth units such as millimeters to flow units such as cubic meters per second using equation 34.

$$INRUNOFF(B,T) = \frac{SURP(B,T) \cdot BAREA \cdot POLYLOSS}{1000 \cdot 3600 \cdot STEPSIZE} \tag{34}$$

The result is convolved with the basin response function and transferred to the subbasin outlet as the local contribution to river flow (equation 35).

$$LOCALFLOW(B,T+R) = [INRUNOFF(B,T)] * [RESPONSE(B,R)] \tag{35}$$

where * is the convolution integral

UpStreamflow contribution entering a river reach must also be translated to the downstream end of the reach and added to the local flow contribution, allowing for translational losses along the river reach. For reaches without dams, this translation can be represented using equation 36.

$$OUTFLOW(B,T) = LOCALFLOW(B,T) + (INFLOW(B,T-T') \cdot RIVLOSSCOEFF) \tag{36}$$

If the river reach contains a DAM, the preexisting reservoir water storage is updated by adding the upStreamflow. Reservoir discharge is then computed from the updated storage by a user-defined storage function (equation 37).

$$OUTFLOW(B,T) = LOCALFLOW(B,T) + \frac{\partial}{\partial T}(DAMSTORE(B,T-1) + INFLOW(B,T-T')) \tag{37}$$

The discharge from the river reach is transferred to the next downstream reach, D, as presented in equation 38.

$$INFLOW(D,T) = \sum_{n} OUTFLOW(B,T) \tag{38}$$

where *n* is the number of reaches immediately upstream of reach *D*.

The discharge from each river reach is then written to the streamflow file.

Diffusion Analog Routine

The diffusion analog method is another linear transport routine available in GeoSFM. This method goes beyond the lag routing by accounting for both flow advection (using flow time or celerity) and attenuation (using a flow dispersion coefficient). In fact, the diffusion analog equation is the linear solution of the advection-dispersion equation (also known as the Navier-Stokes equation) for a plane rectangular source (Maidment, 1993). Mathematically, the diffusion analog equation can be expressed as equation 39.

$$Q(t) = I(t_o) \cdot \left(\frac{x}{\sqrt{4 \cdot \pi \cdot D(t-t_o)^3}} \right) \cdot \exp\left[-\left(\frac{(V \cdot (t-t_o) - x)^2}{4 \cdot D(t-t_o)} \right) \right] \tag{39}$$

where

D is the dispersion coefficient in m^2/s
V is the flow celerity in m/s
x is the length of the river reach in m
t_o is the time of the input event in seconds
i is the present time in seconds
$Q(t)$ is the discharge at the downstream end of the river reach
$I(t_o)$ is the inflow at the upstream end of the river reach

In its implementation in GeoSFM, the diffusion analog routine is essentially the lag routing with additional functionality to account for the deformation or attenuation of the input. The attenuation is controlled by a time invariant parameter called the dispersion coefficient. Consequently, one parameter controls how quickly water travels through the river reach while a second parameter controls the spreading that occurs within the reach. The FORTRAN program, diffroute.for, retrieves runoff generated by the soil water balance in each subbasin and routes it through the river network by diffusion analog as documented below. The variables, arrays, input files, and output files used in this program are presented in table 6.

Table 6. List of key variables for the diffusion analog routine.

Name	Type	Description	Comments
		Variables	
B	Integer	Number of subbasins/rivers	Source = routparam.txt
T	Integer	Simulation time steps	Source = routparam.txt
R	Integer	Response time steps	Source = routparam.txt
F	Integer	Forecast time steps	Source = routparam.txt
D	Integer	Number of river response ordinates	Internal computation
STARTDAY	Real	Simulation start day	Source = routparam.txt
STARTYR	Real	Simulation start year	Source = routparam.txt
LDAY	Integer	Simulation end day	Source = routparam.txt
LYEAR	Integer	Simulation end year	Source = routparam.txt
INHOURS	Real	Simulation time step in hours	Source = routparam.txt
AREA	Real	River reach subbasin area	Source = river.txt
UPAREA	Real	Total drainage area at reach outlet	Source = river.txt
DOWNID	Integer	Subbasin downstream of river reach	Source = river.txt
RIVLENGTH	Real	Length of the river reach	Source = river.txt
RIVMANNING	Real	Manning roughness coefficient	Source = river.txt
RIVERLOSS	Real	In-Streamflow loss coefficient	Source = river.txt
RIVERPLOSS	Real	Overbank flow loss coefficient	Source = river.txt
HASDAM	Integer	River reach dam flag	Source = river.txt
CELERITY	Real	Flow celerity for river reach	Source = river.txt
		Arrays	
SURP(B, T)	Real	Total runoff generated daily	Input from basinrunoffyield.txt
RIVER(B, 12)	Real	River Characteristics File	Input from river.txt
RESPONSE(B, R)	Real	Unit Hydrograph File	Input from response.txt
STREAMFLOW(B,T+F)	Real	Streamflow File	Output to streamflow.txt
LOCALFLOW(B, T+F)	Real	Local contribution to flow	Output to localflow.txt
INFLOW(B, T+F)	Real	UpStreamflow into a subbasin	Output to inflow.txt
FORECAST(B, F)	Real	Streamflow Forecast File	Output to forecast.txt
DIFFRESP(B, D)	Real	River Response Ordinates	Internally Computed

SURPFN	Output	Total runoff yield file	Default = Basinrunoffyied.txt, Source = balfiles.txt
RIVERFN	Input	River characteristics file	Default = River.txt, Source = balfiles.txt
RESPFN	Input	Subbasin unit hydrograph file	Default = Response.txt, Source = balfiles.txt
RESERVOIRFN	Output	File containing reservoir listing	Default = Reservoir.txt, Source = balfiles.txt
STOREFN	Output	Soil water storage file	Default = Soilwater.txt, Source = balfiles.txt
AEVAPFN	Output	Actual evapotranspiration file	Default = Actualevap.txt, Source = balfiles.txt
GWLOSSFN	Output	Regional groundwater loss file	Default = Gwloss.txt, Source = balfiles.txt
OUTSWFN	Output	Final soil moisture file	Default = Cswater.txt, Source = balfiles.txt
LOGFN	Output	Runtime messages and errors file	Default = Logfilesoil.txt, Source = balfiles.txt
EXCESSFN	Output	Excess precipitation file	Default = Excessflow.txt, Source = balfiles.txt
BASEFN	Output	Baseflow file	Default = Baseflow.txt, Source = balfiles.txt
INTFN	Output	Interflow file	Default = Interflow.txt, Source = balfiles.txt
PARAMFN	Input	Listing of simulation parameters	Default = Balparam.txt, Fixed File Name
BALFILESFN	Input	Listing of simulation files	Default = Balfiles.txt, Fixed File Name
DEFAULTFN	Output	Default methods file	Default = Default.txt, Fixed File Name

The surplus runoff from excess precipitation, interflow, and baseflow are aggregated (equation 40).

$$\text{SURP}(B, T) = \text{EXCESSRAIN}(B, T) + \text{INTERFLOW}(B, T) + \text{BASEFLOW}(B, T) \quad (40)$$

The surplus runoff is converted from depth units such as millimeters to flow units such as cubic meters per second using equation 41.

$$\text{INRUNOFF}(B, T) = \frac{\text{SURP}(B, T) \cdot \text{BAREA} \cdot \text{POLYLOSS}}{1000 \cdot 3600 \cdot \text{STEPSIZE}} \quad (41)$$

The result is convolved with the basin response function and transferred to the subbasin outlet as the local contribution to river flow, represented with equation 42.

$$\text{LOCALFLOW}(B, T + R) = \left[\text{INRUNOFF}(B, T)\right] * \left[\text{RESPONSE}(B, R)\right] \quad (42)$$

where * is the convolution integral

UpStreamflow contribution entering a river reach must also be transferred to the downstream end of the reach with a convolution integral to account for the spreading out of flow along the reach. Since the dispersion coefficient in the diffusion analog equation is time invariant, the ordinates associated with its response function along any particular river reach are also time invariant. Consequently, the ordinates are estimated once at the beginning of the simulation and applied to the flow entering the upstream end of the reach at all subsequent simulation time steps. The resulting discharge is added to the local flow contribution after allowing for translational losses along the river reach. For reaches without dams, this translation can be presented by equation 43.

$$\text{OUTFLOW}(B, T) = \left(\text{LOCALFLOW}(B, T) + (\text{INFLOW}(B, T - T') * \text{DIFFRESP}(B, D)) \right) \cdot \text{RIVLOSSCOEFF} \quad (43)$$

where * is the convolution integral

   D is the number of ordinates of the river response function

   DIFFRESP(B, D) is the matrix containing the river response function ordinates

If the river reach contains a dam, the preexisting reservoir water storage is updated by adding the upStreamflow. Reservoir discharge is computed as the updated storage by a user-defined storage function represented by the differential term in equation 37. The discharge from the river reach is also transferred to the next downstream reach, *D*, using equation 38, and the discharge from each river reach is then written to the streamflow file.

Muskingum Cunge Routine

The Muskingum Cunge is a nonlinear, variable parameter routing method. Like the linear Muskingum routing method, it uses the continuity equation and an empirical storage equation. It relies on the Muskingum K coefficient (analogous to flow time) to control the rate of advection and the Muskingum X coefficient to control the rate of attenuation, or spreading out. Cunge proposed to amend the method by allowing the Muskingum X coefficient to vary during each time step based on the flow at the previous time step. In effect, the rate of attenuation of flow is dependent on the condition of flow. Mathematically, Muskingum Cunge routing can be expressed using equations 44 and 45.

$$\frac{dS}{dt} = I - Q \tag{44}$$

$$S = K\left[XI + (1 - X)Q\right] \tag{45}$$

The Muskingum K and X coefficients are estimated using equations 46 and 47, respectively.

$$K = \frac{\Delta x}{\bar{c}} \tag{46}$$

$$X = \frac{1}{2} - \left(\frac{\overline{Q}}{2\bar{c}\,\overline{B}\,S_e\,\Delta x}\right) \tag{47}$$

where

   K is the Muskingum *K* coefficient in seconds

   X is the dimensionless Muskingum *X* coefficient

   c is the average flood wave celerity in m/s

   B is the average channel width in m

   S_e is the dimensionless friction slope

   x is the length of the channel

   S is the storage in the channel at time *t* in m/s

   I is the inflow in m^3/s

   Q is the discharge in m^3/s

All catchments within the analysis area are parameterized for the three routing methods during preprocessing such that the user can opt for any of the three methods just prior to execution without additional preprocessing. The variables used in this script are described in table 7.

Table 7. List of key variables for the Muskingum Cunge routine.

Variable Group	Variable Name	Type	Description
		Input	
PERIOD	DAY	Integer	Day of the month
	DOY	Integer	Julian day of the year
	YEAR	Integer	Four-digit year
	MONTH	Integer	Two-digit month
	HOUR	Integer	Two-digit hour
CHANNEL	ROUGH	Float	Basin mainstream channel mean Manning n coefficient
	RIVLENGTH	Float	Basin mainstream channel length (m)
	SLOPE	Float	Basin mainstream channel mean slope
	WIDTH	Float	Basin mainstream channel mean width (m)
	BANKFULLFLOW	Float	Basin mainstream channel bankfull flow (m^3/sec)
	LOSSFLDP	Float	Floodplain to regional groundwater water loss fraction
	LOSSRIV	Float	In-stream infiltration and evaporation water loss fraction
	HASDAM	Integer	Flag for presence of a dam in river reach, 1 = yes, 0 = no
	HASRATING	Integer	Flag for presence of rating curve for river reach, 1 = yes, 0 = no
	RATINGTYPE	Integer	Type of rating curve, 1 = exponential, 2 = polynomial
LANDFORM	HLLNG	Float	Mean length of the mainstream tributaries (m)
	AREA	Float	Basin area (km^2)
	UPAREA	Float	Basin outlet upstream watershed area (km^2)
RESERVOIRS	ISREGULATE	Integer	Flag indicating if the reservoir is regulated, 1 = yes, 0 = no
	RINFLOW	Float	Reservoir inflow (m^3/sec)
	ROUTFLOW	Float	Reservoir outflow (m^3/sec)
	RSTORAGE	Float	Amount of water storage in the reservoir (m^3)
	RLEVEL	Float	Reservoir water level (m)
	RAREA	Float	Reservoir surface area (km^2)
	LAGRSV	Float	Reservoir flow lag time (sec)
		State	
PROGNOSTIC	DEPTH	Float	Depth of the water in mainstream channel (m)
	ROUTSTOR	Float	conceptual channel reservoir storage (m^3/sec)
	DSTORSTART	Float	Initial conceptual channel reservoir storage capacity (m³/sec)
	FLDSTOR1	Float	1st conceptual reservoir to describe flood plain storage (mm)
	FLDSTOR2	Float	2nd conceptual reservoir to describe flood plain storage (mm)
River Routing	COEFF[4]	Float	Muskingum Cunge routing scheme coefficients
	NSPACE	Integer	Number of Muskingum Cunge routing segments in a reach
	NTIME	Integer	Number of Muskingum Cunge routing interval per time step
	dx	Float	Routing segment length for Muskingum Cunge scheme
	dt	Float	Routing interval for Muskingum Cunge flow segment
	IFLOW	Float	Initial streamflow for Muskingum Cunge flow segments
	LAG	Float	Travel for water passing through channel segment
	CELERITY	Float	celerity of the flow
		Output	
STREAMFLOW	OUTFLOW(T)	Float	present streamflow at the subbasin outlet (m^3/sec)
	INFLOW(T)	Float	present streamflow at the subbasin inlet (m^3/sec)
	LOCALFLOW	Float	Flow generated within the local subbasin (m^3/sec)
	OBS	Float	observed streamflow (m^3/sec)
	INFLOW(T-1)	Float	past streamflow at the basin headwater inlet (m^3/sec)
	OUTFLOW(T-1)	Float	past streamflow at the basin headwater outlet (m^3/sec)
	Qref	Float	Reference flow or theoretical bankfull flow (m^3/sec)
	HASOBS	Integer	flag that indicates if basin stream has observed data

| RATING | RATPARAM[3] | Float | Parameters of the rating curve |
| | HMAX | Float | Bankfull stage (m) |

A theoretical bankfull flow is estimated from upstream contributing area for each river to allow for the separation of overbank flow from channel flow. The relationship is presented as equation 48.

$$BANKFULLFLOW = 36.0 \cdot (UPAREA / 2)^{0.68} \cdot 0.02832 \tag{48}$$

While the input runoff data have a predefined time interval, subscale timing and routing intervals must be defined for the Muskingum Cunge routing scheme to ensure numerical stability of computations. The space and time subinterval are based on reference flow and celerity (from the river file) and fixed throughout the simulation, even though the actual flow celerity varies during each simulation time step. The size of the time subinterval is computed as the smaller of the time it takes for the kinematic wave to travel through the river reach or one-sixth of the input data time interval (equation 49).

$$\Delta t = MIN \begin{cases} \dfrac{RIVLENGTH / WAVECELERITY}{6.0} \\[2em] \dfrac{INTERVAL}{6.0} \end{cases} \tag{49}$$

The routing space subinterval is similarly estimated as equation 50.

$$\Delta x = MIN \begin{cases} MAX \begin{cases} (0.6 \cdot \Delta t \cdot WAVECELERITY) \\[1.5em] (0.4 \cdot \Delta t \cdot WAVECELERITY) - \left(\dfrac{Q_{ref}}{WIDTH \cdot SLOPE \cdot WAVECELERITY} \right) \end{cases} \\[3em] (0.4 \cdot \Delta t \cdot WAVECELERITY) + \left(\dfrac{Q_{ref}}{WIDTH \cdot SLOPE \cdot WAVECELERITY} \right) \end{cases} \tag{50}$$

During each time step, a new wave celerity is computed based on the water surface slope and shape during the preceding routing time subinterval (equation 51).

$$WAVECELERITY = \left(\frac{SLOPE}{ROUGH} \right)^{0.5} \cdot 1.6666667 \cdot \left(\frac{Q_d \cdot ROUGH}{\sqrt{SLOPE \cdot WIDTH}} \right)^{0.4} \tag{51}$$

Muskingum K and X parameters are also recomputed to reflect the new celerity using equations 52 and 53, respectively.

$$MUSKK = \frac{\Delta x}{WAVECELERITY} \tag{52}$$

$$MUSKX = 0.5 \cdot \frac{(1 - (Q_d / WIDTH))}{WAVECELERITY \cdot SLOPE \cdot \Delta x} \tag{53}$$

After calculating K and X parameters, Reynolds (D) and courant (C) numbers are also calculated to verify numerical stability with equations 54 and 55, respectively.

$$REYNOLDSD = (1 - 2 \cdot MUSKX) \tag{54}$$

$$COURANTC = \frac{WAVECELERITY \cdot \Delta t}{\Delta x} \tag{55}$$

If the sum of these two numbers (REYNOLDSD + COURANTC) is less than 1.05 or if COURANTC is greater than 2, then the routing time and space subinterval (*dt* and *dx*) are adjusted until the two conditions are met. For propagation of flow within the river channel, a set of Muskingum coefficients are computed using the four coefficients in equations 55A, 55B, 55C, and 55D.

$$COEF0 = \frac{\Delta t - (2.0 * MUSKK * MUSKX)}{2 * MUSKK * (1 - MUSKX) + \Delta t} \tag{55A}$$

$$COEF1 = \frac{\Delta t + (2.0 * MUSKK * MUSKX)}{2 * MUSKK * (1 - MUSKX) + \Delta t} \tag{55B}$$

$$COEF2 = \frac{2 * MUSKK * (1 - MUSKX) - \Delta t}{2 * MUSKK * (1 - MUSKX) + \Delta t} \tag{55C}$$

$$COEF3 = \frac{(Q_{local} / RIVLENGTH) * \Delta x * \Delta t}{2 * MUSKK * (1 - MUSKX) + \Delta t} \tag{55D}$$

The coefficients are applied to the flow at the upstream section for the current and previous time step and to the flow at the current section for the previous time step to estimate current flow at the current section with equation 56.

$$CFLOW(n\Delta x, n\Delta t) = \begin{pmatrix} COEF0 \cdot CFLOW(n\Delta x - 1, n\Delta t) \\ + COEF1 \cdot CFLOW(n\Delta x - 1, n\Delta t - 1) \\ + COEF1 \cdot CFLOW(n\Delta x, n\Delta t - 1) \end{pmatrix} \tag{56}$$

The flow in the floodplain outside the river channel is simulated using two conceptual linear reservoirs in series, the lag time for each of which is computed using equation 57.

$$\text{LAGFLD} = \frac{0.42}{p^{0.5}} \cdot \left(\frac{(\text{RIVLENGTH}/0.3048) \cdot \text{ROUGH} \cdot 1.2}{\text{SLOPE}^{0.5}} \right)^{0.8} \tag{57}$$

Since the two reservoirs have the same residence time, their net effect can be represented by the 2 parameter gamma distribution. Discharge from river reaches containing reservoirs is handled in one of two ways. Where discharges are supplied through an input file, releases are made according to such specifications. Otherwise, reservoir releases are made using a conceptual level-pool reservoir. This involves computing and applying three coefficients to current and previous reservoir inflows and discharge with equations 58, 59, and 60.

$$\text{RCOEF1} = \text{RCOEF2} = \frac{\left(\dfrac{1}{\text{LAGRSV}}\right)}{2 + \left(\dfrac{1}{\text{LAGRSV}}\right)} \tag{58}$$

$$\text{RCOEF3} = \frac{2 - \left(\dfrac{1}{\text{LAGRSV}}\right)}{2 + \left(\dfrac{1}{\text{LAGRSV}}\right)} \tag{59}$$

$$\text{ROUTFLOW(T)} = \left(\begin{array}{l} \text{RCOEF1} \cdot \text{RINFLOW(T -1)} \\ + \text{RCOEF2} \cdot \text{RINFLOW(T)} \\ + \text{RCOEF3} \cdot \text{ROUTFLOW(T -1)} \end{array} \right) \tag{60}$$

The reservoir discharge is then added to streamflow at the downstream end of the routing reach and propagated downstream through the stream network.

Calibration Module

GeoSFM includes an automatic calibration method intended to improve the application of the rainfall-runoff model. Because of uncertainties in the input and output data sets, model structure error, and the quantity or quality of data, it is difficult to determine a best set of model parameters (Duan and others, 1992). The calibration algorithm performs a global search of the specified parameter space, which has been shown to be effective in dealing with difficulties in searching such as multiple local optima and derivative discontinuities. The algorithm is also efficient because it needs relatively few model runs to obtain the optimum parameter set.

The relationship between the streamflow model and the calibration algorithm is presented in figure 2. By the end of a successful calibration, the streamflow model will have been run many times (iterations) while tracking information from the parameter values input for each of the model runs. The parameter values that give the best model performance become the optimal parameter set. The GeoSFMCalib is linked to the spatial input data through basin and river watershed attribute files that list parameter values for each subwatershed. These attribute files (basin.txt and river.txt) are rewritten with new parameter values each time the model is run. The GeoSFM.dll is used to call the streamflow model. The modeled output is compared to the observations using an aggregated statistic over the time series. The user can choose one of six statistical tests to make this

comparison. The statistical test used to compare the model results to the observation can also be called an objective function. At the time of writing this report, the software connection between the calibration results and the ArcView interface had not been completed and is shown with a dashed connector.

Figure 2. Calibration module relationships between the Geospatial Stream Flow Model and the Multi-Objective Shuffled Complex Evolution Metropolis algorithm.
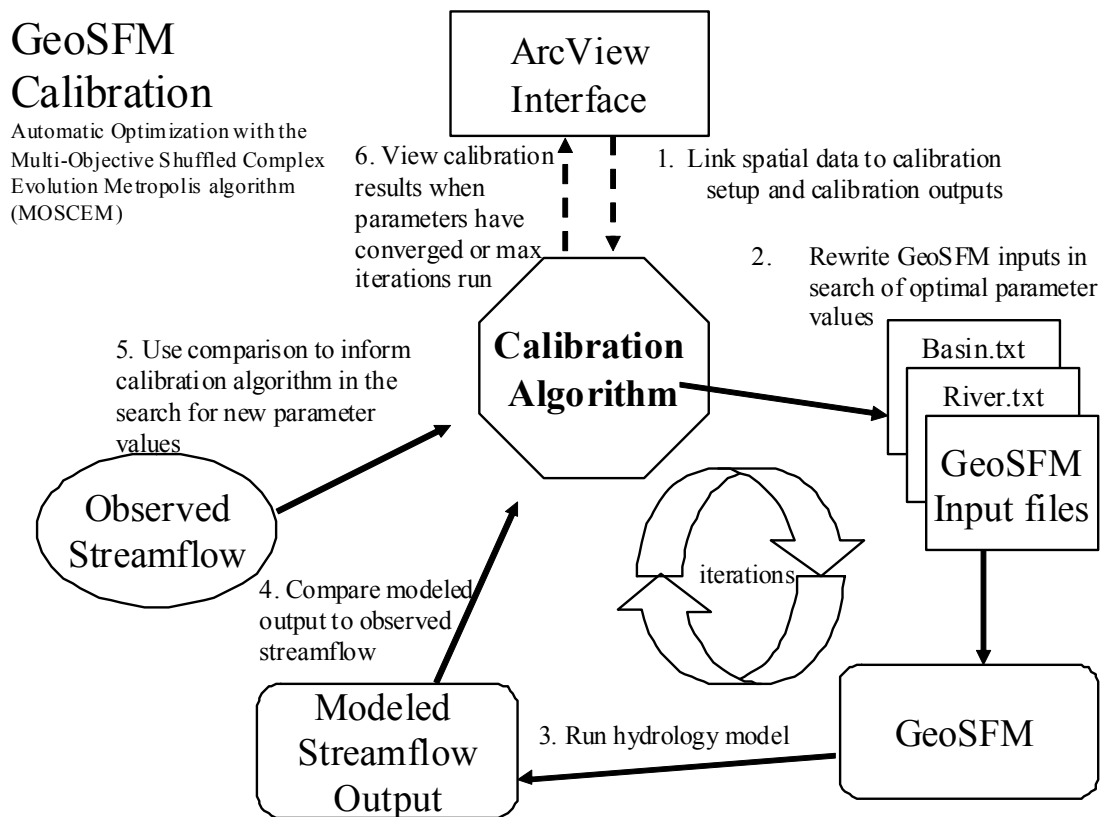
A flowchart of the calibration process is presented in figure 3. The calibration process begins with the model setup. The terrain analysis extracts model parameters and creates basin.txt and river.txt watershed attributes. Copies of basin.txt and river.txt are saved under different names (basin_original.txt and river_original.txt) because the calibration process overwrites the former pair of files with each program iteration. Next, the user must assess the uncalibrated model. The best calibration results can only be achieved with realistic parameter inputs and an appropriate model structure for the watershed that is being modeled. The uncalibrated model results should be analyzed before beginning the calibration. The program geosfmcalib.exe consists of a caller program (dllcaller.exe) which calls the GeoSFM and runs the model one time using the basin.txt and river.txt inputs and a statistical analysis program (UncalibStats.exe) which reports basic model performance by comparing model output with observed streamflow. The user should also plot observed and modeled streamflow time series to visually check and ensure that the calibration results are reasonable. While subjective, visual analysis may help to identify large errors in input

data, parameter ranges, or observed streamflow, which the calibration process cannot be expected to overcome.

Figure 3. Flowchart of the calibration process.

```
┌─────────────────────────────┐
│  1. Use Terrain Analysis to │
│  extract Model Parameters   │
│  from land surface data     │
└─────────────────────────────┘
              │
              ▼
     ⬭ 2. Test Uncalibrated Model ⬭
              │
              ▼
          ⬭ 3. Perform SA
          400 model runs ⬭
              │
              ▼
     ⬭ 4.  Calibrate sensitive parameters
        5,000-10,000 model runs ⬭
              │
              ▼
     ⬭ 5.  Run Postprocessing
     Returns uncertainty boundaries on streamflow
        and model performance statistics ⬭
              │
              ▼
┌─────────────────────────────────┐
│ 6.    Calibrated model parameters│
│       Can be used for            │
│       operational model runs     │
└─────────────────────────────────┘
```
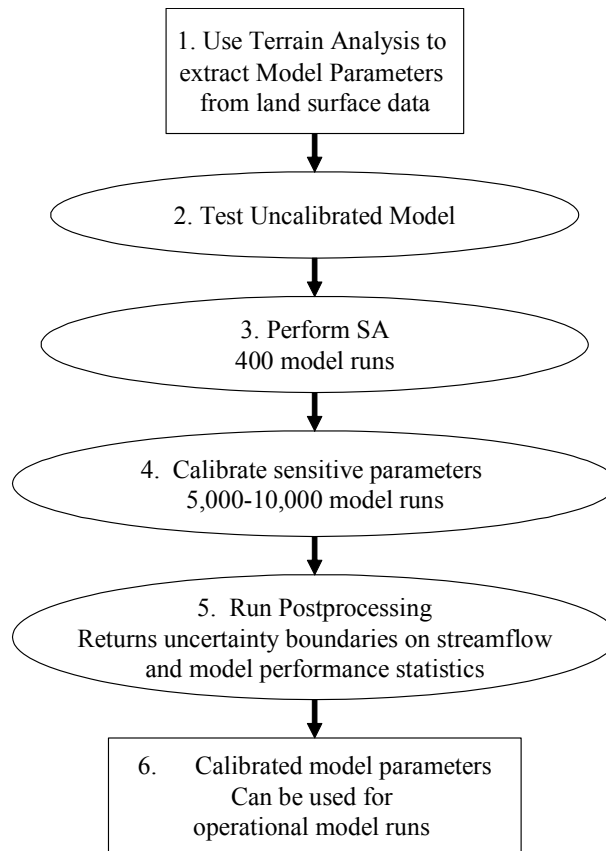
Next, sensitivity analysis should be performed. Sensitivity analysis (SA) serves the dual purpose of testing which sensitive parameters should be used for calibration and analyzing feasible parameter ranges. Simply put, the goal of SA is to see if there is a change in model results when the parameter values are changed. The program geosfmsarun.exe tests 20 input parameters over user-specified ranges. A total of 400 model runs are used to complete the SA test. The SA result can be used to test the model structure; if parameters assumed to have a strong impact on model results do not show any sensitivity, the model structure should be reassessed. Likewise, the SA can give a clear understanding of the feasible parameter space. If the SA results show a parameter to be too sensitive or not sensitive at all, the parameter range may need to be refined. Getting the minimum and maximum parameter space refined as much as possible using realistic values for the specific watersheds being modeled is imperative for obtaining realistic calibration results.

For calibration, a subset of the 20 parameters tested in SA should be selected. For the parameters not selected for calibration, the uncalibrated values in the watershed attribute files are used. The program geosfmcalib.exe is the main calibration program. Although many calibration algorithms have been developed for hydrologic modeling over the years, the MOSCEM algorithm has been chosen for its ability to consistently find optimum parameter sets with the least number of model runs. The number of model runs necessary to converge to the optimum parameter values

depends on the length of streamflow record, the number of parameters being tested, and the complexity of the relationship between model parameters and the watershed response being modeled. The range will generally be 5,000–10,000 model runs to converge to a global minimum.

One-At-a-Time (OAT) Sensitivity Analysis

The sensitivity analysis method is intended to give an efficient look at model parameter sensitivity so that sensitive parameters can be chosen for calibration in the least number of model runs. The One-At-a-Time (OAT) method (Morris, 1991) is a local test because in each model run, only one parameter is changed while all other parameters are held constant. With this method, changes in the output for each model run can be unambiguously attributed to the change of the parameter. To ensure that parameters are tested over the full range, parameter values are taken at a 20 equal interval sample for each of the 20 parameters. This sampling results in a total of 400 model runs. This efficient method for this application ensures the full range of parameters is tested while attributing parameter effects on simulations with relatively few model runs. However, this approach does not examine interactions among parameters.

Because the GeoSFM has spatially distributed parameter information, care was taken to retain the spatial pattern of parameter values derived from land surface data and terrain analysis. The minimum and maximum ranges to test for each parameter (given in range.txt) and the basin average value for each parameter (derived from subbasin parameter values in basin_original.txt and river_original.txt) are used to derive minimum and maximum multiplier values. Values at equal intervals between the minimum and maximum multiplier values for each parameter are used to write model inputs (basin.txt and river.txt) one parameter by one interval at a time. Intervals are calculated using a fraction (f) of the distance between the minimum and maximum parameter multiplier value. The range of multiplier values used to define parameter search space is calculated with equation 61.

$$OATmultiplier = \frac{f \cdot Max_{range.txt}}{Max_{watershed}} + \frac{Min_{range.txt}}{Min_{watershed}} \tag{61}$$

where $Min_{range.txt}$ and $Max_{range.txt}$ are user-specified feasible upper and lower bounds for the parameter space given in range.txt, and where $Min_{watershed}$ and $Max_{watershed}$ are the upper and lower bounds extracted from the subbasin parameter values given in the watershed attribute files (basin.txt and river.txt). The parameters written for each subbasin in the watershed attribute files are calculated using equation 62.

$$Subbasin\ parameter_{OAT} = OATmultiplier \cdot Subbasin\ parameter_{original} \tag{62}$$

The OAT sensitivity test changes the parameter values at all subbasins for each OAT parameter test. For the remaining subbasin parameter values, original subbasin parameter values are used (from basin_original.txt and river_original.txt).

For any subbasin, if the result from equation 62 is greater than the maximum set in range.txt, the OAT parameter is set to the maximum, calculated using equation 63.

$$OATmultiplier_{max} = \frac{Max_{range.txt}}{Max_{watershed}} \tag{63}$$

When the basin characteristics files (basin.txt and river.txt) are written, the GeoSFM.dll is used to run the model. A model configuration consisting of one soil moisture accounting routine (linear or nonlinear) and one river transport routine (lag, diffusion analog, or Muskingum Cunge) is specified in an input file (whichModel.txt). For each OAT test, the model results (streamflow.txt) are analyzed using three statistical tests (stest): maximum, mean, and standard deviation. These tests give different information on the effect that a parameter change may have on model results. Generally, parameter changes have the greatest numerical effect on the peaks or the maximum streamflow modeled. The sensitivity for each of 20 parameters is calculated for three statistical tests by averaging the equal interval model results (found in intermediate file, SensAnalysis.txt), as shown in equation 64.

$$S = \frac{\sum_{i=1}^{20} \left| Stest_i - Stest_{i-1} \right|}{20} \tag{64}$$

Results can be charted in order to choose which parameters are most important in the calibration (see SarunOutput.txt). Sensitive parameters may be model dependent, watershed dependent, or time period dependent. The sensitivity must be tested while preparing to calibrate so that resources are not wasted optimizing parameters that have little or no effect on model output.

Sensitivity analysis measures the impact on model outputs due to changes in inputs. Sensitivity analysis gives a picture of the importance of uncertainty in model inputs and can help refine model structure. OAT is one of the simplest of many techniques for sensitivity analysis. Other techniques include Latin-Hypercube sampling, Monte Carlo sampling, Generalized Sensitivity Analysis (GSA), Multi-Objective GSA (MOGSA), and the Fourier Amplitude Sensitivity Test (FAST).

The Multi-Objective Shuffled Complex Evolution Metropolis (MOSCEM) Algorithm

The purpose of calibration is to adjust model parameters so that model results closely match real observations. Although many GeoSFM model parameters are derived from spatially distributed observed data, uncertainties in initial parameter estimates due to differences input data scale, errors in observed data, and inaccuracies in the structure of the model require that parameters be adjusted. As with all hydrologic models, parameter adjustment allows the GeoSFM model to overcome what is unknown as well as what cannot be measured about a watershed.

Much research has been done on automatic optimization strategies. These strategies take advantage of computers to objectively find an optimal set of parameters that minimizes the difference between the model-simulated output and the observed output values. The alternative is to use manual trial-and-error calibration, which is time consuming, subjective, and limited in search capabilities to a small parameter space. The research in automatic optimization strategies seeks to create algorithms that can consistently find the optimum parameters (effective) in as few model runs as possible (efficient). The shuffled complex evolution (SCE) (Duan and others, 1992) is the most popular algorithm for searching a global parameter space without having to run a random Monte Carlo search, which would make the problem several orders of magnitude larger. Vrugt and others (2003a) improved on the SCE by replacing the downhill simplex search aspect of the algorithm with a Metropolis-annealing scheme and incorporating a multiobjective framework (Vrugt and others, 2003b). The resulting MOSCEM algorithm combines complex shuffling, controlled random search, and probabilistic search to give an estimate of probabilistic uncertainty

for a set of parameters that minimizes the difference between modeled and observed values, given multiple objectives.

An important aspect of modeling hydrologic systems is capturing multiple characteristics of the watershed. For a given model, multiple characteristics may include 1) the driven versus nondriven portions of the hydrograph, 2) surface water, groundwater, and evapotranspiration or other modeled fluxes, 3) streamflow at multiple locations within the watershed, or 4) characteristics representing model performance using different statistical tests or objective functions such as RMS, bias, or a maximum likelihood estimator. To capture more than one characteristic in an automatic optimization, multiobjective methods have been developed (Vrugt and others, 2003b). Using multiple sources of observed data can improve the model calibration.

Postprocessing of Calibration Results

The program GeoSFMpostproc.exe is used to process the calibration results. The postprocessing program reads in calibration output files. The calibration algorithm does not save the streamflow results for each of the thousands of model runs due to computer memory and space restrictions, but it does retain the parameter values that created the best model results. The postprocessing runs the model using the best parameter values so that the streamflow results can be plotted and analyzed.

The calibration algorithm works on the assumption that there is no "best" parameter set due to the uncertainties in model structure and data inputs. The postprocessing program runs the model for the entire set of multiple "best" parameter values and returns upper and lower bounds to enable plotting of the uncertainty boundaries. The one parameter set which shows the most balanced tradeoff between the minimized objectives is chosen as the singular "best" parameter set. Six difference model performance statistics are reported for this model run. However, this parameter set (the final basin.txt and river.txt in the working directory) should be used with the understanding that this is only one of many equivalent solutions for the global optimum.

Postprocessing Utilities

A number of postprocessing utilities included in the GeoSFM system to facilitate ease of use are described in this section. Because input precipitation data are spatially distributed, it is often difficult to fully discern the evolution of rainfall events over time. Precipitation and evapotranspiration data may also be required for longer accumulation periods, such as 10-day and monthly totals. GeoSFM includes a routine for vertically integrating grid layers to perform pixel level accumulations and other temporal statistics such as maxima, minima, standard deviations, medians, modes, variety, range, and nonzero values for each pixel. The results of the analyses are presented as grids. While not directly related to the computation of streamflow, the spatial statistics routines are useful for running verification checks on input data and for studying sources of anomalously high streamflow or soil water levels in simulation results.

At the end of a simulation run, a summary of the results can be computed. GeoSFM includes a tool for computing a variety of flow statistics including the maximum, minimum, mean, standard deviation, median, 25th, 33rd, 66th, and 75th percentile flows for each catchment. The 33rd and 66th percentile flows may be adapted as the respective low- and high-flow thresholds. These thresholds can be updated in the catchment data layer and subsequently used in the generation flow status maps. The results of the flow statistics computations are stored in the form of tables linked to the catchment data layer. The user can access the values for each catchment by clicking on the catchment in the visual interface or by modifying the legend of the associated data layer in the visual interface to display the computed values.

Visual maps are considerably easier to interpret than tabular time series data, particularly when dealing with large river basins. GeoSFM contains a tool for displaying the results of simulations for any given date in a visual map. The streamflow values on a user-selected date are presented in the form of indices which present the values in the context of predefined criteria. The default criterion for differentiating between low and normal flow is the 33rd percentile flow for the analysis period, while the 66th percentile flow is the minimum threshold separating normal flows and high flows. However, the user can define other criteria such as return period flow or predetermined drought and flood warning levels for the classification of flows. Each catchment is assigned a flow status index of 1, 2, or 3 to signify the respective low, normal, and high flow conditions. A color-coded map is then produced showing the flow status of each catchment for the selected day.

GeoSFM contains a graphing tool for plotting hydrographs at the completion of a simulation run. The tool can be activated from either a program menu or the tool menu bar. The user can then visually select the catchment for which a hydrograph is required by clicking on the catchment in the visual interface. The program automatically matches the spatial information with the time series and produces a hydrograph. The flexibility of this tool is limited by ArcView's rather limited charting capabilities. The user is consequently encouraged to import the ASCII files from flow simulations into spreadsheet programs such as Microsoft's Excel for more sophisticated time series plotting capabilities.

Conclusions

The geospatial streamflow modeling system developed at the USGS EROS is a semidistributed, catchment-based hydrologic model, complete with supporting parameterization and forcing data for monitoring wide-area hydrologic events. It is designed to run operationally using widely available remotely sensed data sets and ground observations. The preprocessing modules used for processing the input data within a GIS environment have been briefly presented in this manual. Next, the hydrologic analysis module, which consists of a linear soil moisture accounting routine, a more complex nonlinear soil moisture accounting routine, and a choice of three river flow transport routines, namely pure lag, diffusion analog, and Muskingum Cunge, is presented. The hydrologic analysis routines are written as external DLLs, which are initiated seamlessly from ArcView. A calibration module is also presented which links the hydrologic analysis routines to a One-At-a-Time sensitivity analysis routine and a Multi-Objective Shuffled Complex Evolution Metropolis algorithm for parameter calibration. Finally, a postprocessing module, which contains various routines for analyzing and presenting model output in visual formats, is also presented. Taken together, the modules allow users to undertake basic hydrologic assessments in locations with limited in situ information. It is anticipated that GeoSFM will be used to preserve life, property, and the environment through flood hazard and water resource monitoring applications around the world, particularly in locations with limited in situ monitoring infrastructure.

References Cited

Artan, G.A., Asante, K.O., and Verdin, J.P., 2004, A continental scale flood hazard monitoring system, Eos Trans. AGU, v. 85, no. 17, Jt. Assem. Suppl.

Asante, K.O., Macuacua, R.D., Artan, G.A., Lietzow, R., and Verdin, J.P., 2007, Developing a flood monitoring system from remotely sensed data for the Limpopo Basin, IEEE Transactions on Geoscience and Remote Sensing, v. 45, no. 6, p. 1709–1714.

Brooks, R.H., and Corey, A.J., 1964, Hydraulic properties of porous media, Hydrology Paper 3, Colorado State University, Fort Collins, CO.

Christie, F., and Hanlon, J., 2001, Mozambique and the Great Flood of 2000. Indiana University Press, Bloomington, IN.

Dooge, J.C.I., 1973, Linear theory of hydrologic systems. Tech.Bull. no. 1468, US Agricultural Research Service, Washington, D.C.

Duan, Q., Gupta, V.K., and Sorooshian, S., 1992, Effective and efficient global optimization for conceptual rainfall-runoff models, Water Resources Research, v. 28, p. 1015–1031.

Dunne, T., and Leopold, L.B., 1978, Water in Environmental Planning. Freeman, San Francisco, CA.

Food and Agriculture Organization (FAO), 1971–1981, Soil Map of the World. UNESCO, Paris, vol. 1-10 (1:5M scale maps and accompanying texts).

Green W.H., and Ampt, G.A., 1911, Studies on soil physics, part I, the flow of air and water through soils, Journal of Agricultural Sciences, vol. 4, no. 1, p. 1–24.

Jenson, S.K., and Dominque, J.O., 1988, Extracting topographic structure from digital elevation data for geographic information system analysis, Photogrammetric Engineering and Remote Sensing, vol. 54, no. 11, p. 1593–1600.

Kanamitsu, M., 1989, Description of the NMC global data assimilation and forecast system, Weather Forecasting, vol. 4, p. 335–342.

Kunkel, K.E., Changnon, S.A., and Angel, J.R., 1994, Climatic aspects of the 1993 Upper Mississippi River Basin flood, Bulletin of American Meteorological Society, vol. 75, no. 5, p. 811–974.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, J., Yang, L., and Merchant, J.W., 2000, Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1-km AVHRR Data, International Journal of Remote Sensing, vol. 21, no. 6, p. 1303–1330.

Maidment, D., ed., 1993, Handbook of Hydrology. McGraw Hill, New York, NY.

Morris, M.D., 1991, Factorial sampling plans for preliminary computational experiments, Technometrics, vol. 33, no. 2, p. 161–174.

Rodriguez-Iturbe, I., and Valdez, J.B., 1979, The geomorphological structure of unit response, Water Resources Research, v. 15, p. 1490–1520.

Verdin, J.P., and Klaver, R.W., 2002, Grid cell based crop water accounting for the Famine Early Warning System, Hydrologic Processes, vol. 16, p. 1617–1630.

Vrugt, J.A., Gupta, H.V., Bouten, W., and Sorooshian, S., 2003a, A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, Water Resources Research, v. 39, no. 8, p. 1201.

Vrugt, J.A., Gupta, H.V., Bastidas, L.A., Bouten, W., and Sorooshian, S., 2003b, Effective and efficient algorithm for multiobjective optimization of hydrologic models, Water Resources Research, v. 39, no. 8, p. 1214.

Webb, R.S., Rosenzweig, C.E., and Levine, E.R., 1993, Specifying land surface characteristics in general circulation models: soil profile data set and derived water-holding capacities, Global Biogeochemical Cycles, vol. 7, p. 97–108.

Xie, P., and Arkin, P.A., 1997, A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, Bulletin of the American Meteorological Society, vol. 78, no. 11, p. 2539–2558.

Zobler, L., 1986, A world soil file for global climate modeling. NASA Tech. Memo. 87802, NASA Goddard Institute for Space Studies, New York, NY.

Appendix 1: GeoSFM Data Specifications

Preprocessing Module: Performing Terrain Analysis

Input Grid	Description	Units	Values
Elevations	Hydrological corrected Digital Elevation Model – source HYDRO1k DEM.	m	

Input Shapefile

Limpbas.shp	Shapefile defining analysis extent.		

Input Parameters

Analysis Extent	Boundary defined for analysis area (same as Limpbas.shp).	m	
Analysis Cell Size	Boundary defined for cell area (same as elevations).		
Stream Delineation Threshold	Minimum number of cells upstream of a given location for initiating stream. Default = 1,000, minimum = 500.		

Output Grids

FlowDir	Flow direction grid consists of numeric values assigned using the eight direction pour point model. Each cell is assigned one of eight compass directions.		
FlowAcc	Flow accumulation grid defines the number of contributing cells draining into each cell. Values range from 0 at topographic highs to large numbers at the mouths of rivers.		
FlowLen	DownStreamflow length grid defines the distance from the cell to the basin outlet.	m	
Streams	Stream grid defines stream cells.		1
StrLinks	Stream link grid assigns a unique value to each stream ID.		
Outlets	Outlet grid defines number of outlets; one outlet per subbasin, each outlet assigned a unique value.		
Basins	Subbasin grid assigns a unique ID value to each subbasin.		
HillLength	Hill length grid defines distance to stream, from cell to nearest stream.	m	
Slope	Hill slope grid describes the maximum change in elevation between each cell and its eight neighbors. Expressed in percentage.		
Downstream	Downstream grid assigns a unique ID number of the subbasin immediately downstream.		
BigBasins	Basin boundary polygon.		
Basply.shp	Shapefile defining all subbasins with unique ID and gridcodes.		
Rivline.shp	Shapefile defining streams with unique IDs, gridcodes, and from_node and to_node.		

Preprocessing Module: Generating Basin Characteristics Files

Input Grids	Description	Units	Values

Rcn	SCS runoff curve number grid uses hydraulic soil classes together with the GLCC classes to derive an RCN value.	
Whc	Soil water holding capacity grid is calculated from the DSMW soil class and the total available water median soil moisture value.	mm/m
Soildepth	Hydrological active soil layer depth grid created from DSMW soil code value and Webb's soil depth values.	cm
Texture	Texture grid created from Zobler's seven classes converted to three classes—coarse, medium, and fine.	
Ks	Saturated soil hydraulic conductivity grid created on Ks values, which are based on Zobler's texture classes.	cm/hr
Maxcover	Maximum impervious cover grid based on wetland areas and water bodies found in the GLCC data set and stream cells defined in the flowacc grid.	

Output Files

Basin.txt *see .txt file detail below*	File containing each subbasin characteristics. This file is rewritten during the calibration process for each model run.	
Basin_original.txt	Copy of basin.txt. Retains original basin characteristic data.	
Order.txt *see .txt file detail below*	Computational order—used in subsequent program operations.	
river.txt *see .txt file detail below*	File containing river characteristics. This file is rewritten during the calibration process for each model run.	
river_original.txt	Copy of river.txt. Retains original river characteristic data.	
describe.txt *see .txt file detail below*	File descriptions for 12 different files generated in modeling process.	

Basin.txt

BasinID	Subbasin identification number.		
SoilWHC	Soil water holding capacity calculated from the DSMW soil class and the total available water median soil moisture value.	mm/m	
SoilDepth	Hydrological active soil layer depth created from DSMW soil code value and Webb's soil depth values.	cm	
Texture	Texure value created from Zobler's seven classes and converted to three classes—coarse, medium, and fine.		1,2,3
Ks	Saturated soil hydraulic conductivity value created based on Zobler's Ks and texture values.	cm/hr	
Area	Drainage area of the subbasin.	km^2	
Interflowlag	Residence time for the interflow reservoir.	days	
Hslope	Average subbasin slope.		
Baseflowlag	Residence time for the baseflow reservoir.	days	

RCNumber	SCS runoff curve number uses hydraulic soil classes together with GLCC classes to derive a RCN value.		0–100
MaxCover	Maximum impervious cover value based on wetland area and water bodies found in the GLCC data set a long with the stream cells defined in the flow accumulation data.		0–1.0
BasinLoss	Fraction of the soil water infiltrating to groundwater.		0–1.0
Pancoeff	Pan coefficient for correcting PET readings.		0–1.0
Topsoil	Fraction of soil layer classified as hydrologically active topsoil layer.		0–1.0
Aridity	Aridity index range (1 = arid, 2 = wet).		1,2
Order.txt			
BasinID	Most downstream subbasin Number.		
	2nd most downstream subbasin Number.		
	3rd most downstream subbasin Number.		
	↓		
	Most upstream subbasin Number.		
River.txt			
RiverID	Identification number of the river reach.		
Area	Local drainage area of the river reach.	km^2	
UpArea	Total area upstream of the river reach.	km^2	
RivSlope	Average slope of the river reach/basin mainstream channel mean slope.		
RivLength	Length of the river reach.	m	
DownID	Identification number of the downstream river reach.		
RivManning	Value of Mannings n for the river reach, Default = 0.035.		
RiverLoss	Fraction of river water lost to infiltration.		0–1.0
RiverPloss	Fraction of the river water lost to evaporation.		0–1.0
HasDam	Indicates whether the river reach contains a dam (0 = no, 1 = yes).		0,1
HasRating	Indicates whether the reach has a rating curve (0 = no, 1 = yes).		0,1
HasFlowData	Indicates whether the river reach has observed flow data (0 = no, 1 = yes).		0,1
Celerity	Velocity of the flood wave through the river reach.	m/s	
Diffusion	Flow attenuation (or dispersion) coefficient of the reach.	m^2/s	
RivWidth	Average channel width.	m	
FlowRef	Reference flow for the section.	m^3/s	
RunType	Indicates if simulation should proceed from an existing or a new run (0 = new, 1 = existing).		0,1
describe.txt			
Basin Characteristics	List and description of the 12 different characteristics generated in the basin.txt.		
River Characteristics	List and description of the 17 different characteristics generated in the river.txt.		
Response Characteristics	Explanation of the time step runoff data generated in the response.txt.		

Computation Order	Explanation of the subbasin downstream/upStreamflow in the order.txt.
Rainfall Characteristics	Explanation of the time step rainfall data generated in the rain.txt.
Evapotranspiration Characteristics	Explanation of the time step PET data generated in the evap.txt.
Balance Parameters	List and description of the nine different characteristics generated in the balparam.txt.
Balance File Listing	List and description of the 17 different input/output files in the balfiles.txt.
Route Parameters	List and description of the nine different characteristics generated in the routparam.txt.
Route File Listing	List and description of the 13 different input/output files in the routfiles.txt.
Total Runoff from each Subbasin	Explanation of the time step total runoff data generated in the basinrunoffyield.txt.
Total Streamflow from each Subbasin	Explanation of the time step total discharge generated in the streamflow.txt.

Preprocessing Module: Basin Response Computation

Input Grids/Files	Description	Units	Values
Basins	Subbasin grid assigns a unique ID value to each subbasin.		
Flowdir	Flow direction grid consists of numeric values assigned using the eight direction pour point model. Each cell is assigned one of eight compass directions.		
Flowlen	DownStreamflow length grid defines the distance from the cell to the basin outlet.		
Outlets	Outlet grid defines number of outlets, one outlet per subbasin, each assigned a unique value.		
Usgslandcov	USGS Land Cover grid.		
Velocity	Overland flow velocity value or user-defined velocity grid.		
Elevations	Elevation grid.	m	
Flowacc	Flow accumulation grid defines the number of contributing cells draining into each cell. Values range from 0 at topographic highs to large numbers at the mouths of rivers.		
Order.txt *see .txt file detail -1.3*	Computational order file.		

Output Grids/Files

Traveltime	Travel time from each grid cell to the catchment outlet based on slope, watershed size, flow resistance (Manning's roughness coefficient), and water input.
Velocity	Option 1: Estimated uniform overland velocity for each catchment based on the mean slope of the catchment and dominant land cover.

	Option 2: Velocity values for each catchment defined by user.		
	Option 3: Velocity default value for each catchment defined by user.		
response.txt *see .txt file detail below*	File containing fraction of excess runoff arriving at subbasin outlet in each subsequent time step, similar to a unit hydrograph for each subbasin.		

response.txt

BasinID	BasinID of subbasin beginning with the most downstream.		
	↓		
	Most upstream subbasin.		
Day 0	Fraction of excess runoff arriving at subbasin outlet in each subsequent time step.		0.0–1.0
Day 1	↓		
Day 2			
↓			
Day 20			

Preprocessing Module: Extracting Rain/Evap Data

Input Grid/File	Description	Units	Values
Basins	Basin Grid.		
order.txt *see .txt file detail -1.3*	Computational order file.		

Input Parameters

raindata	Rain data from directory.	mm	
evapdata	Evap data from directory.	tenths of mm	
Start Year	Start year of data.		
End Year	End year of data.		
Start Day Number	Start day of data –defaults 1– Julian Date.		
End Day Number	End day of data –number of days included for data estimate –defaults -240- entered **10** (for shorter processing time) – Julian Date.		

Output Files

rain.txt see .txt file detail below	File containing time step rainfall for each subbasin (per day).	mm	
evap.txt *see .txt file detail below*	File containing time step PET for each subbasin (per day).	tenths of mm	

rain.txt

Time	Date – example 1999001 – format yyyyddd		
	↓		
	Date 1999010		
BasinID of subbasin beginning with the most downstream	time step rainfall for each subbasin.	mm	
↓			
Most upstream subbasin			

evap.txt

Time	Date – example 1999001 – format yyyyddd	
	↓	
	Date 1999010	
BasinID of subbasin beginning with the most downstream	time step PET for each subbasin.	tenths of mm

↓
Most upstream subbasin

Simulation Module: Computing Soil Water Balance

Input Files	Description	Units	Values
rain.txt *see .txt file detail -2.1*	File containing time step rainfall for each subbasin.	mm	
evap.txt *see .txt file detail -2.1*	File containing time step PET for each subbasin.	tenths of mm	
basin.txt *see .txt file detail -1.3*	File containing each subbasin characteristics.		
response.txt *see .txt file detail -1.4*	File containing time step excess runoff arriving at subbasin outlet.		
Input Parameters			
Computation Start Year	Year entered into field.		
Computation Start Day	Day—default 001— Julian Date		
Number of Rain/Evap Days	Days—default 10		
Number of Response Days	Days—default 21 (how long the hydrograph is stretched)		
Number of Subbasins	Number—default 365		
Initial Soil Moisture	Assume soil is dry containing only 10% of its storage capacity—default 0.1 (decimal).		
Data Format (l for Daily Data)	1 = Day (default), 0 = Hourly		
New Run or Continue Previous Run	0 = New run (default), 1 = Continue previous run		
Basin Polygon Theme	Polygon of analysis area – selected **basply.shp** for exercise.		
Key Field e.g. Grid Code	Selected **gridcode** for exercise–number ID.		
Mode	0 = Simulation, 1 = Calibration -defaults-0		
Soil Model	Option of Single Layer Soil Model or Double Layer Soil Model—selected **Single Layer** for exercise.		
Output Files			
balparam.txt *see .txt file detail below*	File containing: Number of ordinates of unit hydrograph response, number of simulation time steps, simulation start year/day, and number of catchments, simulation interval in hours, data format, model initialization code, and initial soil fraction.		

balfiles.txt *see .txt file detail below*	File listing input and output .txt files.	
Output Files		
basinrunoffyield.txt *see .txt file detail below*	Basin runoff yield—runoff for each subbasin.	
soilwater.txt *see .txt file detail below*	Output soil moisture storage.	
actualevap.txt *see .txt file detail below*	Actual evapotranspiration.	mm
gwloss.txt *see .txt file detail*	Output file water lost to the regional groundwater flow.	mm
cswater.txt *see .txt file detail below*	Output file containing last model time step simulated soil water content.	mm
excessflow.txt *see .txt file detail below*	Output file excess runoff fluxes.	mm
interflow.txt *see .txt file detail below*	Output file interflow or fast base flow fluxes.	mm
baseflow.txt *see .txt file detail below*	Output file base flow fluxes.	mm
logfilesoil.txt *see .txt file detail below*	Process/Error Log File.	
massbalance.txt *see .txt file detail below*	Output file of sum of water cycle components for the total simulation period —only for **Double Layer Soil Model**.	
initial.txt *see .txt file detail below*	Input/output file that contains initial conditions for model state variables—only if **Double Layer Soil Model** is selected.	
default.txt *see .txt file detail below*	Soil model and routing configuration selected for model run.	
balparam.txt		
No header	Number of ordinates of unit hydrograph response.	
	Number of simulation time steps.	
	Simulation start year.	
	Simulation start day.	
	Number of catchments.	
	Simulation interval in hours.	24
	Data format indicator (1/0)—defaults 1.	0–1
	Model initialization mode (1/0)—defaults 0 (i.e. no mode initialization).	0–1
	Initial soil fraction, Default = 0.1.	0–1
	Model Calibration Mode (0/1)—defaults 0 (i.e. noncalibration mode).	0–1
balfiles.txt		
No header	Input rainfall file (rain.txt).	
	Input potential evapotranspiration file (evap.txt).	
	Input basin characteristics file (basin.txt).	
	Input unit hydrograph response file (response.txt).	
	Output balance parameter file (balparam.txt).	

Output runoff yield file (basinrunoffyield.txt).
Output soil moisture storage file (soilwater.txt).
Output actual evapotranspiration file (actualevap.txt).
Output groundwater loss file (gwloss.txt).
Output final soil water storage file (cswater.txt).
Output surface precipitation excess file (gwloss.txt).
Output interflow file (interflow.txt).
Output baseflow file (baseflow.txt).
Output mass balance file (massbalance.txt).
Output log file (logfilesoil.txt).
Output initialization file (initial.txt).
Output workdirectory (<full directory pathname>).

basinrunoffyield.txt

Time	Date – example 1999001 – format yyyyddd	
	↓	
	Date 1999110	
BasinID of subbasin beginning with the most downstream	Time step runoff for each subbasin.	mm
↓		
Most upstream subbasin		

soilwater.txt

Time	Date – example 1999001 – format yyyyddd (first column)	
	↓	
	Date 1999110	
ID of subbasin beginning with the most downstream	Time step soil moisture storage for each subbasin.	
↓		
Most upstream subbasin		

actualevap.txt

Time	Date – example 1999001 – format yyyyddd (first column)	
	↓	
	Date 1999110	
ID of subbasin beginning with the most downstream	Time step for actual evapotranspiration for each subbasin.	mm
↓		
Most upstream subbasin		

gwloss.txt

Time	Date – example 1999001 – format yyyyddd (first column)	
	↓	
	Date 1999110	

| ID of subbasin beginning with the most downstream | Time step water lost to the regional groundwater flow for each subbasin. | |

↓
Most upstream subbasin

cswater.txt

Basin ID	ID of subbasin beginning with the most downstream.	
	↓	
	Most upstream subbasin.	
CurrentSW	Contains last model time step simulated soil water content.	mm

excessflow txt

Time	Date – example 1999001 – format yyyyddd (first column)	
	↓	
	Date 1999110	
ID of subbasin beginning with the most downstream	Time step of excess runoff fluxes.	mm

↓
Most upstream subbasin

interflow.txt

Time	Date – example 1999001 – format yyyyddd (first column)	
	↓	
	Date 1999110	
ID of subbasin beginning with the most downstream	Time step of interflow or fast base flow fluxes.	mm

↓
Most upstream subbasin

baseflow.txt

Time	Date – example 1999001 – format yyyyddd (first column)	
	↓	
	Date 1999110	
ID of subbasin beginning with the most downstream	Time step base flow fluxes.	mm

↓
Most upstream subbasin

logfilesoil.txt

No header	Open output files for calib/sens/standard run modes.
	Read basin parameter file.
	Open output files for standard runs.

	Initialized basin objects.	
	Read basins characteristics.	
	Calculated runoff contributing partial area model parameters.	
	Opened met input rain and evap files.	
	Converted slope from percent to fraction.	

default.txt

No header	Soil configuration selected for model run.	
	Soil1 – Single layer soil model.	
	Soil2 – Double layer soil model.	
	Routing method selected for model run.	
	Rout1 – Simple lag routing method.	
	Rout2 – Diffusion analog routing method.	
	Rout3 – Muskingum Cunge routing method.	

massbalance.txt

basinId	ID of subbasin beginning with the most downstream. ↓ Most upstream subbasin.	
precipitation	Sum of water cycle precipitation for the total simulation period.	mm
surfaceRunoff	Sum of water cycle surface runoff for the total simulation period.	mm
interflow	Sum of water cycle fast base flow flux for the total simulation period.	mm
baseflow	Sum of water cycle base flow flux for the total simulation period.	mm
evaporation	Sum of water cycle evaporation for the total simulation period.	mm
moisture_change	Difference of the water cycle value from start of model to end of model for the total simulation period.	mm
residual	Difference of water cycle (rain in, evaporation/runoff out) for the total simulation period.	mm

initial.txt

Basin ID	ID of subbasin beginning with the most downstream. ↓ Most upstream subbasin.	
upperSoilWater	Upper soil water values for each subbasin.	
lowerSoilWater	Lower soil water values for each subbasin.	
fastBaseflow	Fast base flow values for each subbasin.	
slowBaseflow	Slow base flow values for each subbasin.	
interflowWater	Interflow reservoir water values for each subbasin.	
upperSoilExcessWater	Upper soil excess water values for each subbasin.	
lowerSoilExcessWater	Lower soil excess water values for each subbasin.	
surfaceWaterAtTime0	Surface water value at first time step of hydrograph for each subbasin.	
surfaceWaterAtTime1	Surface water value at next time step of hydrograph for each subbasin.	

surfaceWaterAtTime2	Surface water value at next time step of hydrograph for each subbasin.		
↓	↓		
surfaceWaterAtTime19	Surface water value at next time step of hydrograph for each subbasin.		

Simulation Module: River Flow Transport

Input Files	Description	Units	Values
basinrunoffyield.txt *see .txt file detail -2.2*	Basin Runoff Yield File.		
response.txt *see .txt file detail -1.4*	File containing time step excess runoff arriving at subbasin outlet.		
river.txt *see .txt file detail -1.3*	File containing river characteristics.		
reservoir.txt *see .txt file detail below*	Reservoir Characteristic File.		
rating.txt *see .txt file detail below*	Stream gauge input containing BasinID, type, parameters and maximum river depth.		
Input Parameters			
Basin Polygon Theme	Polygon of analysis area – selected **basply.shp** for exercise.		
Key Field e.g. Grid Code	Selected **gridcode** for exercise.		
Number of Days of Forecast Required	Default - 3		
Mode	0 = Simulation, 1 = Calibration -default 0-		
Routing Method	Needed for the computation: three options—simple lag routing method, diffusion analog routing method, or Muskingum Cunge routing method—selected **simple lag routing method** for exercise.		
reservoir.txt			
DamID	BasinID where dam is located.		
Storage	Storage in 1,000m^3.		
Residence Time	Defaults **30.**		
Operated	0 –No user-defined file.		0, 1
	1 –User-defined file used.		
OpFilName	Path and file name if user-defined file is used. Defaults –**None.**		
rating.txt			
BasinID	BasinID of subbasin beginning with the most downstream.		
	↓		
	Most upstream subbasin.		
Type	Two different types based on the following equations:	Q = m^3/sec	1,2
	Type 1 – Q = A(H+B)^C		
	Type 2 – Q = (A*H^2) + BH + C		
		H = m	
	where: A = Par1		

	B = Par2		
	C = Par3		
Par1	Equation Parameter for "A"		
Par2	Equation Parameter for "B"		
Par3	Equation Parameter for "C"		
hmax	Maximum height where equation is valid.		

Output Files

routparam.txt *see .txt file detail below*	File containing: number of simulation time steps, simulation start year/day, number of catchments, simulation interval in hours, model type, number of forecast days, model initialization mode, number of reservoirs.
routfiles.txt *see .txt file detail below*	File listing input and output .txt files.

Output Files

inflow.txt *see .txt file detail below*	Flow values coming from upstream for each subbasin.
streamflow.txt *see .txt file detail below*	Streamflow file contains a time series of velocity values in cubic meters per second for each subbasin.
localflow.txt *see .txt file detail below*	Flow value generated in each subbasin.
riverdepth.txt *see .txt file detail below*	River depth file populates when **Muskingum Cunge routing method** is selected.
damstatus.txt *see .txt file detail below*	Reservoir status file, present only if there are dams in the basin.
damlink.txt *see .txt file detail below*	Present only if there are dams in the basin.
forecast.txt *see .txt file detail below*	Forecast file contains the streamflow forecast for the number of days that were selected (3 days in exercise).
times.txt *see .txt file detail below*	Start/end time and date of computing streamflow process.
logfileflow.txt *see .txt file detail below*	Process/Error log file.

routparam.txt

No Header	Number of simulation time steps.		
	Simulation start year.	YYYY	
	Simulation start day.	JDAY	
	Number of catchments.		
	Simulation interval.	hours	24
	Model type (0 = Musk-Cunge, 1 = Linear Reservoir).		0,1
	Number of forecast day.		0–100
	Model initialization mode (1/0).		1,0
	Number of reservoirs.		

routfiles.txt

No Header	Route parameter file (routparam.txt).
	Input river characteristics file (river.txt).
	Input river initialization file (initial.txt).
	Input runoff yield file (basinrunoffyield.txt).

	Input reservoir-river link file (damlink.txt).
	Input observed flow file (obsflow.txt).
	Input rating curve flow file (rating.txt).
	Output streamflow file (streamflow.txt).
	Output subbasin flow contribution file (localflow.txt).
	Output average river depth file (riverdepth.txt).
	Output subbasin upStreamflow contribution (inflow.txt).
	Output flow routing log file (logfileflow.txt).
	Output work directory (<full directory pathname>).

inflow.txt

Time	Date – example 1999001 – format yyyyddd
	↓
	Date 1999013
BasinID of subbasin beginning with the most downstream	UpStreamflow values for each subbasin.
↓	
Most upstream subbasin	

streamflow.txt

Time	Date – example 1999001 – format yyyyddd	
	↓	
	Date 1999013	
BasinID of subbasin beginning with the most downstream	Time step total discharge for each subbasin.	m^3/s
↓		
Most upstream subbasin		

localflow.txt

Time	Date – example 1999001 – format yyyyddd	
	↓	
	Date 1999013	
BasinID of subbasin beginning with the most downstream	Flow value generated in each subbasin.	
↓	Flow value generated within each subbasin.	m^3/s
Most upstream subbasin		

forecast.txt

Time	Date – example 1999011 – format yyyyddd	
	↓	
	Date 1999013	
BasinID of subbasin beginning with the most downstream	Forecasted flow value for each subbasin.	m^3/s
↓		
Most upstream subbasin		

times.txt

No header   Starting time: Day of the week, month/day, hour/minute/second, year.

Ending time: Day of the week, month/day, hour/minute/second, year.

Example – Fri Oct 28 16:49:59 2005

logfileflow.txt

No header   File history. Sample messages below.
CREATED LOGFILEFLOW.TXT FILE
 ****READ ROUTFILES.TXT ****
 OPENED ROUTPARAM.TXT FILE
 ****READ ROUTPARAM.TXT ****
 OPENED RIVER.TXT FILE
 OPENED BASINRUNOFFYIELD.TXT FILE
 CREATED INFLOW.TXT FILE
 CREATED STREAMFLOW.TXT FILE
 OPENED DAMLINK.TXT FILE
 CREATED LOCALFLOW.TXT FILE
 ****READ BASIN FILE HEADER****
 ****READ RUNOFF FILE HEADER****
 ****READ BASIN FILE****
 ****READ BASIN RUNOFF YIELD FILE****
 ****INITIATING ROUTING****
 WRITING OUTPUT FILES
 CLOSING TIME SERIES FILES
 RETURNING TO ARCVIEW

riverdepth.txt   Additional Output for Muskingum Cunge routing method.

Time   Date – example 1999001 – format yyyyddd (first column)

↓

Date   1999013

BasinID of subbasin beginning   Time step river depth for each subbasin.   m
with the most downstream

↓

Most upstream subbasin

damstatus.txt

damlink.txt   Present only if dams present.

Calibration Module: Performing Sensitivity Analysis

Input Parameters	Description	Units	Values

| River Reach | Drop-down list of river reach–selected 140 for exercise 4. |
| Model Configuration | Drop-down list of all possible model configurations–soil/routing. Selected one soil layer, lag routing for exercise 4. |

Input Files

range.txt *see above table*	Provides minimum and maximum values of a feasible parameter range for testing model sensitivity.
balparam.txt *see section 2.2*	File containing: Number of ordinates of unit hydrograph response, number of simulation time steps, simulation start year/day, and number of catchments, simulation interval in hours, data format, model initialization code, and initial soil fraction.
basin_original.txt	Original basin characteristics.
basin.txt	File containing each subbasin characteristics. This file is rewritten during the sensitivity analysis process.
river_original.txt	Original river characteristics.
river.txt	File containing river characteristics. This file is rewritten during the sensitivity analysis process.
moscem_param.txt *see .txt file detail below*	Gives location of outlet in model output file.
whichmodel.txt *see .txt file detail below*	Sets choice of water balance and routing model configurations.

Moscem_param.txt
No header

Nflux	Number of output variables (fluxes).	
Ntstep1	Number of time steps of input variables.	
Nstep2	Number of time steps of output variables.	
Obj_func,	Statistical test for optimization:	1,2,3,4,5,6
	1 = "RMSE" root mean square error	
	2 = "STD" standard deviation	
	3 = "HMLE" maximum likelihood error function	
	4 = "NSE" Nash-Sutcliffe efficiency	
	5 = "NSC" Number of sign changes	
	6 = "BIAS"	
Missing Value	Value of missing data in observed_streamflow.txt.	–9999
Nflux_obs	Column in observed_streamflow.txt to test.	
Nflux_model	Column in streamflow.txt to extract modeled data.	

whichModel.txt

Model Index	Models:
	Water balance model
	Routing model
Index Description	Water balance model – 1 = 1D balance, 2 = 2Dbalance
	Routing model –1 = diffusion, 2 = Muskingum Cunge, 3 = lag

Output Files

oat.txt *see .txt file detail below*	One-At-a-Time parameter index and multiplier values used in sensitivity analysis.
SensAnalysis.txt *see .txt file detail below*	Results of sensitivity analysis for each model run, includes parameter index, multiplier value, mean, maximum, and standard deviation of modeled streamflow results.
SArunOutput.txt *see .txt file detail below*	Gives the mean absolute difference of test results over the parameter range for each parameter. The greater the differences, the more sensitive the parameter.

oat.txt

Whichparam	List of 1
	↓
	through 20
	For all 20 runs, total of 400 lines.
OAT Parameter Value	Multiplier values used in sensitivity analysis.

SensAnalysis.txt

whichParam	List of 1
	↓
	through 20
	For all 20 runs, total of 400 lines.
OAT Parameter Value	Multiplier values used in sensitivity analysis.
Max_mod	Maximum of modeled streamflow results.
Mean_mod	Mean of modeled streamflow results.
stddev	Standard deviation of modeled streamflow results.

SArunOutput.txt

whichparm	List of 1
	↓
	through 20
max_mod	The maximum difference of test results over the parameter range for each parameter.
mean_mod	The mean difference of test results over the parameter range for each parameter.
stddev	The standard deviation of test results over the parameter range for each parameter.

Calibration Module: Performing Model Calibration

Input Parameters	Description	Units	Values
Observed Streamflow Stations	The streamflow station ID displayed in a drop-down list. Selected **1** for exercise 4.		
BasinID for Streamflow Station	Drop-down list of basin IDs, select ID that corresponds with the streamflow station selected. Selected **140** for exercise 4.		
Number of Runs	Drop-down list of choices. Selected **54** for exercise 4. (Selected for faster processing time.)		

The number of model runs depends on the length of streamflow record, number of parameters being tested, and the complexity of the parameter dependent model response for the watershed being modeled. The range will generally be 5,000–10,000 model runs for convergence to a global minimum.

Objective Function Type	Convergence can be measured in terms of Root Mean Square Error (RMSE) Standard Deviation (STD) Maximum Likelihood Error (MLE) Nash-Sutcliffe Efficiency (NSE) Number of Sign Changes (NSC) BIAS Selected **RMSE** for exercise 4.

Input Files

balparam.txt	File containing: Number of ordinates of unit hydrograph response, number of simulation time steps, simulation start year/day, and number of catchments, simulation interval in hours, data format, model initialization code, and initial soil fraction.
basin_original.txt	Original basin characteristics.
basin.txt	File containing each subbasin characteristics. This file is rewritten during the sensitivity analysis process.
river_original.txt	Original river characteristics.
river.txt	File containing river characteristics. This file is rewritten during the sensitivity analysis process.
streamflow.txt *see .txt detail below*	Streamflow file contains a time series of velocity values in cubic meters per second for each subbasin.
parameter.in	Range of multiplier values used to define parameter search space:
see file detail below	Minparameter.in = Minrange.txt / (basin ave param) Maxparameter.in = Maxrange.txt / (basin ave param)
moscem.in *see file detail below*	Main calibration algorithm control file.
MOSCEMx.txt *see .txt file detail below*	File listing which parameters to calibrate.
moscem_param.txt *see .txt detail below*	Calibration control file.
ObjOptFlag.in *see file detail below*	Flag for number of fluxes to "multi"-calibrate.
whichModel.txt	Sets choice of water balance and routing models.
observed_streamflow.txt	Observations from which to compare model results.

parameter.in

No	Index.
Name	Parameter name *see "Parameters to be Calibrated."*

Default	Dummy variable.	
Lower	Minimum multiplier value.	
Upper	Maximum multiplier value.	
OptIdx	Optimization flag: 0 = do not optimize, 1 = optimize	
Description	Description of parameter *see "Parameters to be Calibrated."*	

moscem.in

No Header

nOptPar	Number of parameters used for optimization.	
nOptObj	Number of fluxes used for optimization.	
nSamples	Number of random samples from nOptPar dimensioned parameter space per shuffle.	
nComplex	Number of complexes.	
nMaxDraw	Maximum number of function evaluations/model runs.	
ParamMult file	Name of the parameter file including path. Default = parameter.in	
whichFlux file	Name of the file specifying which fluxes to be calibrated, including path. Default = objoptflag.in	
Observed streamflow file	Name of the file providing evaluation data, including path. Default = observed_streamflow.txt	
ObjectiveResults output	Name of file storing the parameter values of the final points, including path. Default = objectives.out	
Parameter values output	Name of file storing the objective function values of the final points, including path. Default = parameter_values.out	
Parameter convergence output	Name of the file keeping track of the convergences of each of the parameters, including path. Default = par_convergence.out	

MOSCEMx.txt

Parameter Index	One line for each parameter to be calibrated.	
	↓	
	Last line setting to denote end of life.	–999

ObjOptFlag.in

No Header	Runoff_1	1,0
	Runoff_2	1,0

observed_streamflow.txt

Time	Date – example 1999001 – format yyyyddd	
	↓	
	Date 1999365	
Streamflow Stations	Streamflow values.	
↓		
Last station		

Output Files

Par_convergence.out *see file detail below*	Parameter convergence for best nSamples.	
Objectives.out *see file detail below*	Parameter_values for best nSamples.	

55

| Parameter_values.out *see file detail below* | Objective function results for best nSamples. | | |

Objectives.out

nSampleIndex	Sample number starting with 1		
	↓		
	Last sample 60		
Objective values for each flux	Parameter_values for best nSamples.		

Fitness/Probability for each sample

<div align="center">

Calibration Module: Postprocessing Calibration Results
</div>

Input Files	Description	Units	Values
balparam.txt	File containing: Number of ordinates of unit hydrograph response, number of simulation time steps, simulation start year/day, and number of catchments, simulation interval in hours, data format, model initialization code, and initial soil fraction.		
basin_original.txt	Original basin characteristics.		
basin.txt	File containing each subbasin characteristics. This file is rewritten during the sensitivity analysis process.		
river_original.txt	Original river characteristics.		
river.txt	File containing river characteristics. This file is rewritten during the sensitivity analysis process.		
postproc.in *see file detail below*	File containing nSamples, observed streamflow values, Parameter_values.out, Objectives.out, Timeseries.out, ObjectStats.out and Trdoff_bounds.out.		

postproc.in
No Header

nSamples	Same nSamples from moscem.in, equivalent to number of parameter sets saved in parameter_values.out.		
	Observed streamflow values file. Default = Observed_streamflow.txt		
	Parameter value output file from calibration. Default = Parameter_values.out		
	Objective function results output file from calibration. Default = Objectives.out		
	Timestep, Observed Streamflow, Single timeseries of "best" calibration run. Default = Timeseries.out		
	Objective Statistics for all six objective functions for the streamflow runs presented in timeseries.out. Default = ObjectStats.out		
	Timestep, Lower boundary, upper boundary, and average streamflow over the nSample model runs. Default = Trdoff_bounds.out		

Output Files

timeseries.out *see file detail below*	Objective function results output file from calibration.		
objectiveStats.out *see file detail below*	All statistic tests from "best" calibration run in timeseries.out.		
Trdoff_bounds.out *see file detail below*	Timestep, lower boundary, upper boundary, and average streamflow over the nSample model runs.		

timeseries.out

Timestep	Date	1	1–365
		↓	
	Date	365	
Observed Station 1	Observed Streamflow values.		
Calibrated Station 1	Single time series of "best" calibration run.		

trdoff_bounds.out

Timestep	Date	1	1–365
		↓	
	Date	365	
LowerBound Station1	Lower boundary streamflow values for station 1.		
UppBound Station1	Upper boundary streamflow values for station 1.		
Average Station1	Average streamflow values for station 1.		

Postprocessing : Update Bankfull and Flow Statistics

Input Files	Description	Units	Values
streamflow.txt *see .txt detail below*	Streamflow file contains a time series of velocity values in cubic meters per second for each subbasin.		
Output Files			
monthlyflow.txt *see .txt detail below*	Streamflow file contains a monthly time series of velocity values in cubic meters per second for each subbasin.		
annualflow.txt *see .txt detail below*	Streamflow file contains an annual time series of velocity values in cubic meters per second for each subbasin.		
riverstats.txt *see .txt detail below*	List of streamflow statistic types and corresponding values.		
statsparam.txt *see .txt detail below*	File which contains the start year, start month, start day, end year, number of simulation time steps, number of catchments, statistical type, and a list of input/output file names and paths.		
logfilestats.txt *see .txt detail below*	Log file of statistical parameters and output file names and paths.		

monthlyflow.txt

Time	Date – example 199901 – format yyyymm (year, month)		
BasinID of subbasin beginning with the most downstream	Streamflow file contains a monthly total time series of velocity values in cubic meters per second for each subbasin.	m^3/s	

↓
Most upstream subbasin

annualflow.txt

Time	Date – example 1999 – year	
BasinID of subbasin beginning with the most downstream	Streamflow file contains an annual total time series of velocity values in cubic meters per second for each subbasin.	m^3/s

↓
Most upstream subbasin

riverstats.txt

ID	File record number.
Gridcode	Subbasin ID number.
Soilwater	Soil water value for subbasin.
Currentsw	Current soil water value (same as Soilwater) for subbasin.
Sum	Total Streamflow.
Count	Number of days of data.
Mean	Streamflow mean value for each catchment.
Maximum	Streamflow maximum value for each catchment.
Minimum	Streamflow minimum value for each catchment.
Range	Streamflow range value for each catchment.
Variance	Streamflow variance for each catchment.
Stddev	Streamflow standard variation for each catchment.
Median	Streamflow median value for each catchment.
Quart25	25 percentile flow for each catchment.
Quart75	75 percentile flow for each catchment.
Quart33	33 percentile flow for each catchment.
Quart66	66 percentile flow for each catchment.
Highflow	Streamflow high flow value for each catchment.
Lowflow	Streamflow low flow for each catchment.
Medflow	Streamflow medium flow value for each catchment.

statsparam.txt

No Header	Start Year.
	Start Month.
	Start Day of the Month.
	End Year.
	Number of simulation time steps.
	Number of catchments.
	Statistical Type.
	Input file name and path: Default = streamflow.txt
	Monthly Flow Output file and path: Default = monthlyflow.txt
	Annual Flow Output file and path: Default = annualflow.txt

logfilestats.txt   *sample messages below*

No Header	Start Year.
	Start Day.
	Number of Records.
	Number of Fields.
	Operation (0-MEAN, 1-MEDIAN, 2-MAX).

Input filename e.g.
C:\GeoSFM\workspace\streamflow.txt.

Monthly Output filename e.g.
C:\GeoSFM\workspace\monthlyflow.txt.

Annual Output filename e.g.
C:\GeoSFM\workspace\annualflow.txt.

Calculating monthly max for month 1.
Writing annual values.

Postprocessing: Display Flow Percentile Map

Input File	Description	Units	Values
streamflow.txt *see .txt detail in section 3.2*	Streamflow file contains a time series of velocity values in cubic meters per second for each subbasin.		
Input Parameter			
Date	Selected from drop-down list of dates for flow percentile map —yyyyddd—Julian day		
Output			
Map	Map showing daily variation of streamflow values displayed as low, normal, and high conditions.		

Postprocessing: Display Flow Hydrographs

Input Parameters	Description	Units	Values
Basin Theme	Drop-down list of choices – **basply.shp** selected.		
Subbasin ID	Drop-down list of choices – **Gridcode** selected.		
Output			
Hydrograph	Graph showing daily variation of streamflow at the outlet of the selected subbasin.		

GeoSFM Utilities: Downloading Rain/Evap Grid

Input Parameters	Description	Units	Values
FTP Site Information	FTP Site Location default –**edcftp.cr.usgs.gov** Site Username default –**anonymous** Site Password default –**anonymous@edc**		
Specify Type of Data to Download	Options: **Get Rainfall Grids**, Get Forecast Rainfall Grids, **Get Global Pet Grids**, and Get Global TRMM Rainfall Grids –selections made for exercise 6 are in bold.		
Region of Analysis	Options: Global, **Africa**, SE Asia, Afghanistan, Australia, NAmerica, CAmerica, SAmerica, and Europe –selected Africa for exercise 6 –data is not available for all regions.		
Select Grids to downloaded	A list of 16 days of data is presented –10 days were selected to complete exercise 6. **–rain_yyyyddd.tar.gz or evap_yyyyddd.tar.gz**		
Zip File			
rain_yyyyddd.tar.gz	Zip rainfall data downloaded to local working directory –one file per day.	mm	
evap_yyyyddd.tar.gz	Zip evap data downloaded to local working directory –one file per day.	tenths of mm	

GeoSFM Utilities: Unzip, Untar, Imagegrid

Input Grid	Description	Units	Values
rain_yyyyddd.tar.gz	Zip rainfall data downloaded to local working directory –one file per day.		
evap_yyyyddd.tar.gz	Zip evap data downloaded to local working directory –one file per day.		
rain_yyyyddd.bil	Unzipped rainfall image files processes in local working directory.		
evap_yyyyddd.bil	Unzipped evap image files processes in local working directory.		
Output Grid			
rain_yyyydddg	Grid containing rainfall data for each day – geographic projection, added to ArcView table of contents.		
evap_yyyydddg	Grid containing evap data for each day –geographic projection added to ArcView table of contents.		

GeoSFM Utilities: Projecting Geographic Grids to Lambert Azimuthal

Input Grid	Description	Units	Values
rain_yyyydddg	Grid containing rainfall data for each day – geographic projection, added to ArcView table of contents.		
evap_yyyydddg	Grid containing evap data for each day –geographic projection added to ArcView table of contents.		
Input Parameters			
Select Continent for Analysis	Options: Global, **Africa**, SE Asia, Afghanistan, Australia, NAmerica, CAmerica, SAmerica, and Europe –selected **Africa** for exercise 6.		
Analysis Properties	AnalysisExtent: **Current Value** Analysis Cell Size: **Current Value** Use all default values.		
Projection	Lambert Equal-Area Azimuthal.		
Output Grid			
rain_yyyyddd	Grid containing rainfall data for each day –Lambert Azimuthal projection added to ArcView table of contents.	mm	
Evap_yyyyddd	Grid containing evap data for each day –Lambert Azimuthal projection added to ArcView table of contents.	tenths of mm	

GeoSFM Utilities: Compute Rain/Evap Grid Statistics

Input Parameters	Description	Units	Values
raindata	Rain data from directory.		
evapdata	Evap data from directory.		
Start Year	Start year of data —1999.		
End Year	End year of data —1999.		
Start Day Number	Start day of data –defaults 1-Julian Day		

End Day Number	End day of data –number of days included for calculation of statistics –defaults 365-entered **10** (shorter processing time) –Julian Day

Input Parameter

GRID_STATYPE_SUM	Sum of values for date range.
GRID_STATYPE_MAX	Maximum value for date range.
GRID_STATYPE_MEAN	Average of values for date range.
GRID_STATYPE_STD	Standard Deviation value for date range.
GRID_STATYPE_MEDIAN	Median value for date range.
GRID_STATYPE_RANGE	Range of values for date range.
GRID_STATYPE_MIN	Minimum value for date range.
GRID_STATYPE_MAJORITY	Majority value for date range.
GRID_STATYPE_MINORITY	Minority value for date range.
GRID_STATYPE_VARIETY	Number of different values for date range.
GRID_NONZERO_VALUES	Nonzero values for date range.

Output Grid

R10SUM1-*grid name depends on type and number of days selected*	R = rain, 10 = number of days, SUM = statistical type, 1 = each calculation run is incremented.
E10SUM1-*grid name depends on type and number of days selected*	E = evap, 10 = number of days, SUM = statistical type, 1 = each calculation run is incremented.

GeoSFM Utilities: Pick Grid Values at Point

Point theme	Description	Units	Values
gauges2.shp	Site IDs, latitude and longitude coordinates in decimal degrees.		

Input Parameters

raindata	Rain data from directory.	mm	
evapdata	Evap data from directory.	mm*10	
Start Year	Start year of data –1999.		
End Year	End year of data –1999.		
Start Day Number	Start day of data –defaults 1-Julian Day		
End Day Number	End day of data –number of days included for calculation of statistics –defaults 365-entered 10 (shorter processing time) –Julian Day		

Output File

rainpoint.txt *see .txt detail below*	File containing time step rainfall (mm) for each station ID (per day).
evappoint.txt *see .txt detail below*	File containing time step PET (tenths of mm) for each station ID (per day) File.

rainpoint.txt

Time	Date – example 1999001 — format yyyyddd	
	↓	
	Date 1999010	
Station ID number starting with the smallest number	Time step rain for each point (rain gauge station).	mm
↓		
Largest number		

evappoint.txt

Time	Date – example 1999001 — format yyyyddd	
	↓	
	Date 1999010	
Station ID number starting with the smallest number	Time step PET for each point (evap gauge station).	Tenths of mm
↓		
Largest number		

<div align="center">

GeoSFM Utilities: Interpolate Station Data to Grid
</div>

Input Shapefile	Description	Units	Values
gauges2.shp	Site IDs, latitude and longitude coordinates in decimal degrees.		
Limpbas.shp	Shapefile defining modeled basin region.		
Input Files			
rainstations.txt *see .txt detail below*	File containing time step rainfall (mm) for each station ID.		
evapstations.txt *see .txt detail below*	File containing time step PET (tenths of mm) for each station ID.		
Input Parameters			
Date	Data start date formatted m/d/yyyy		
Interpolate Surface Method	Two interpolation methods are displayed in the drop-down list:		
	Inverse Distance Weighted (IDW) interpolator assumes each input has a local influence that diminishes with distance; the points closer to the processing cell are given a greater weight.		
	Spline interpolator is a general purpose method that fits a minimum-curvature surface through the input points.		
Inverse Distance Weighted (IDW) parameters	This is the default interpolation method.		
Z Value Field	The field containing the values to be used in the interpolation. –**Stations_i**		

	For IDW specify either **Nearest Neighbors** or a **Fixed Radius** in the interpolation. –defaults **Nearest Neighbors.**		
Number of Neighbors / Fixed Radius	**Number of Neighbors** or input points –defaults **12**. For **Fixed Radius** enter the radius to search for points.		
Power	Controls the significance of the surrounding points on the interpolated value – the lower the number the more influence from distance points. Defaults **2** For exercise 6 changed to **1**.		
Barriers	A break that limits the search for input points. Defaults **No Barriers**.		
Spline Parameters			
Z Value Field	The field containing the values to be used in the interpolation. –**Stations_i**		
Input Parameters cont.			
Weight	Weight factor.		
No: of Points	Specifies the number of points to use per region.		
Type	Select from drop-down list: **Regularized** – a smooth surface. **Tension** – tunes the stiffness of the surface according to the character of the modeled phenomenon.		
Output Grids			
Rain_1999001	Grid showing rainfall values interpolated across subbasins originating from ground station data for one day.	mm	
↓			
Rain_1999010	(10-day example)		
Evap_1999001	Grid showing PET values interpolated across subbasins originating from ground station data for one day.	tenths of mm	
↓			
Evap_1999010	(10-day example)		

GeoSFM Utilities: Sink Filling Menu			
Input Grid	**Description**	**Units**	**Values**
DEM	Elevation data requiring fill sink process.		
Output Grid	**Description**		
fill1	Elevation data after fill sink process.		

GeoSFM Utilities: Converting Time Series From Daily to Monthly and Annual			
Input Files	**Description**	**Units**	**Values**
streamflow.txt *see .txt detail section 3.2*	Streamflow file contains a time series of velocity values in cubic meters per second for each subbasin.		
Input Parameter			
Statistic to be computed	List includes max and mean —max chosen for exercise.		

Output Files

monthlyflow.txt *see .txt file detail below*	Streamflow file contains a monthly time series of velocity values in cubic meters per second for each subbasin.	
annualflow.txt *see .txt file detail below*	Streamflow file contains an annual time series of velocity values in cubic meters per second for each subbasin.	
statsparam.txt *see .txt file detail below*	File containing the start year, start month, start day, end year, number of simulation time steps, number of catchments, statistical type, and a list of input/output file names and paths.	

monthlyflow.txt

Time	Date – example 199901 – format yyyymm (year, month)	
BasinID of subbasin beginning with the most downstream	Streamflow file contains a monthly total time series of velocity values in cubic meters per second for each subbasin.	m^3/s

↓
Most upstream subbasin

annualflow.txt

Time	Date – example 1999 –year	
BasinID of subbasin beginning with the most downstream	Streamflow file contains an annual total time series of velocity values in cubic meters per second for each subbasin.	m^3/s

↓
Most upstream subbasin

Statsparm.txt

No Header	Start Year.
	Start Month.
	Start Day of the Month.
	End Year.
	Number of simulation time steps.
	Number of catchments.
	Statistical Type.
	Input file name and path: C:\GeoSFM\workspace\streamflow.txt
	Monthly Flow Output file and path: C:\GeoSFM\workspace\monthlyflow.txt
	Annual Flow Output file and path: C:\GeoSFM\workspace\annualflow.txt

GeoSFM Utilities: Frequency Distribution Computation

Input File	Description	Units	Values
streamflow.txt *see .txt detail section 3.2*	Streamflow file contains a time series of velocity values in cubic meters per second for each subbasin.		

Input Parameters

Basin coverage/grid theme	Selected **basply.shp** for exercise.
Select statistics to be computed	Selected **Total Flow, Mean Flow, Maximum Flow,** and **Minimum Flow** for exercise –multiple choices from a drop-down list.
Select field to display	Selected **Mean** for exercise –multiple choices from a drop-down list.

Output

Map	Map showing daily variation of streamflow values displayed as selected statistic –for the exercise **Mean** was selected –multiple choices are presented in drop-down list.