Scale-Dependency and Sensitivity of Hydrological Estimations to Land Use and Topography for a Coastal Watershed in Mississippi

Vladimir J. Alarcon and Charles G. O'Hara

GeoSystems Research Institute, 2 Research Blvd., Starkville MS 39759, USA
{alarcon,cgohara}@gri.msstate.edu

Abstract. This paper investigates the effect of land use and digital elevation models spatial resolution and scale on the simulation of stream flow in two coastal watersheds located in the Mississippi Gulf Coast (USA). Four elevation datasets were used: USGS DEM, NED, NASA's SRTM and IFSAR (300, 30, 30, and 5 meter resolution, respectively). Three land use datasets were included in this study: USGS GIRAS, NLCD, and NASA MODIS MOD12Q1 (400, 30, and 1000 m resolution, correspondingly). The Hydrological Program Fortran (HSPF) was used for estimating stream flow in the two watersheds. Results showed that swapping datasets in a factorial design experiment produce equivalent statistical fit of measured and simulated stream flow data. The results also showed that HSPF-estimated stream flows are not sensitive to scale and spatial resolution of the datasets included in the study.

Keywords: Watershed hydrology, hydrological simulation, HSPF, MODIS, DEM, land use.

1   Introduction

Hydrologic modeling at the watershed scale involves managing large volumes of meteorological, topographical, land use, and water quality data. The management of these large data volumes usually requires linking Geographical Information Systems (GIS) and hydrological models. GIS programs are used for extracting and summarizing weather and physiographic information from digital datasets, and to set-up initial hydrological model's applications that are further refined in later steps of the hydrological modeling process. With the advent of Internet, the ever-growing availability of computational resources, and the demand of users, private or public-domain databases can be accessed directly (from within the GIS programs) and different datasets can be easily downloaded for geo-processing. In particular, topographical and land-use/land-cover datasets are downloaded for the purposes of watershed delineation, land use characterization, geographical positioning of hydro-chemical point sources, etc.

The spatial variability of the physical characteristics of the terrain influences the flow regime in watersheds. It has been investigated from the theoretical and the numerical point of view [1]. Several other researches have explored the sensitivity of
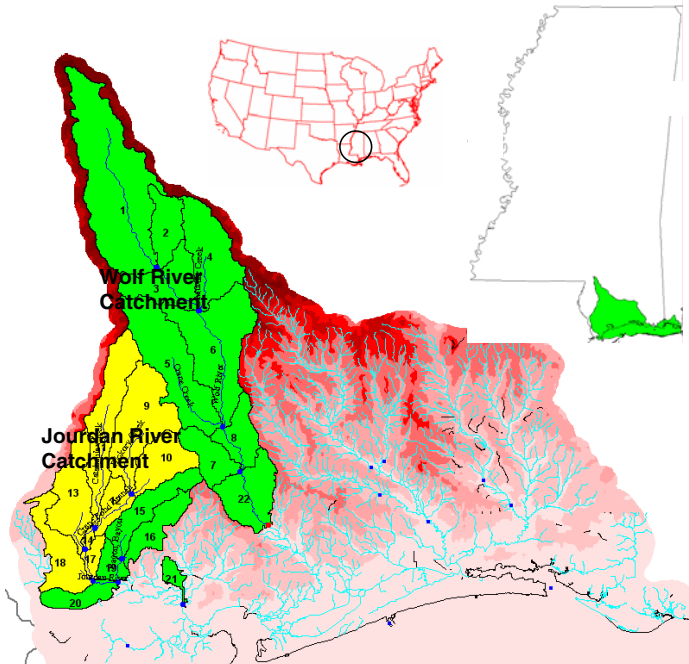
hydrological estimations to topographical or land use datasets. Few, however, have assessed the sensitivity of hydrological estimations to combined swapping of topographical and land use datasets. This paper explores the effects of using several different combinations of topographical and land use datasets in hydrological estimations of stream flow. The Hydrological Simulation Program Fortran (HSPF, [2]) is used for modeling the hydrological processes. The Better Assessment Science Integrating Point & Nonpoint Sources (BASINS) GIS system [3] is used to perform most of the geospatial operations although GEOLEM [4], ArcGis and ArcInfo were also used to complement geospatial processing that was not available in BASINS. This research is part of several studies being undertaken in the Northern Gulf of Mexico region [5][6].

2 Methods

2.1 Study Area

Two main river catchments in Saint Louis Bay watershed at the Mississippi Gulf Coast were the focus of this study (Figure 1).

The Jourdan River catchment drains approximately 88220 ha and it is the largest contributor of flow to the St. Louis Bay, with an average stream flow of 24.5 m^3/s. The

Fig. 1. Wolf River and Jourdan River catchments. Both catchments are located in Saint Louis Bay watershed, Mississippi Gulf Coast, USA.

Wolf River flows into St. Louis Bay from the east (Figure 1). The Wolf River catchment drains slightly more than 98350 ha with an average stream flow of 20.1 m^3/s.

2.1   Topographical Datasets

The topographical digital datasets (also known as Digital Elevation Models, DEM) used in this research were: USGS DEM, NED, and NASA's SRTM and IFSAR. USGS produces five types of elevation data: 7.5-minute, 7.5-minute-Alaska, 15-minute-Alaska, 30-minute, and 1-degree DEMs. In this research, 1-Degree DEMs corresponding to the 3 arc-second (or 1:250,000-scale) USGS topographic map series were used. The National Elevation Data (NED) database provides a seamless mosaic elevation data having as primary initial data source the 7.5-minute elevation data for the conterminous United States [7]. NED has a consistent projection (geographic), resolution (1 arc second, approximately 30 m), and metric elevation units [8]. The NASA Shuttle Radar Topography Mission (SRTM), in collaboration between the National Aeronautics and Space Administration (NASA) and the National Imagery and Mapping Agency (NIMA), collected interferometric radar data, which has been used by the Jet Propulsion Laboratory (JPL) to generate a near-global topography data product for latitudes smaller than 60 degrees. Intermap's Interferometric Synthetic Aperture Radar (IfSAR) Digital Elevation Models [9] are topographical data for 7.5-minute by 7.5-minute units (corresponding to the USGS 1:24,000 scale topographic quadrangle map series) comprising elevations at 5 meter postings. DEMs are intensively used in water resources modeling. Watershed delineation and digital stream (rivers) definition depend heavily on those DEMs. Figure 2 shows all topographical datasets after being geo-processed for the study area. The theoretical link between the digital elevation model and the parameters of the Hydrological Simulation Program Fortran (HSPF) is well described in [10].
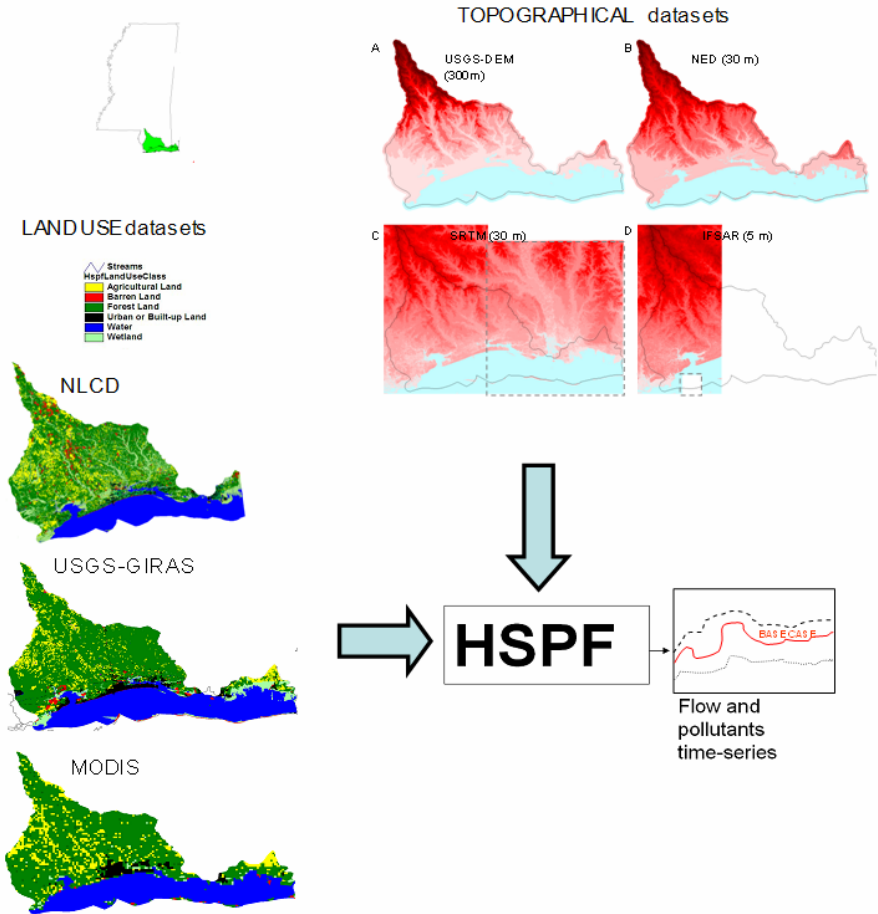
2.2   Land Use Datasets

Three land use datasets were included in this study (Figure 2): USGS GIRAS, USGS NLCD, and NASA MODIS MOD12Q1. All the datasets were downloaded from public land use databases in the US that provide several different types of land-use/ land- cover digital maps.

   The USGS GIRAS is a set of maps of land use and land cover for the conterminous U.S. delineated with a minimum mapping unit of 4 hectares and a maximum of 16 hectares (equivalent to 400 m spatial resolution), generated using the Geographic Information Retrieval and Analysis System (GIRAS) software. Today, they are widely known as the USGS GIRAS land use datasets.

   Derived from the early to mid-1990s Landsat Thematic Mapper satellite data, the National Land Cover Data (NLCD) is a 21-class land cover classification scheme applied consistently over the United States. The spatial resolution of the data is 30 meters and mapped in the Albers Conic Equal Area projection, NAD 83.

   The NASA MODIS MOD12Q1 Land Cover Product (MODIS/Terra Land Cover, 1000 m spatial resolution) [11] is provided by NASA through several internet portals. The land use map is classified in 21 land use categories, following the International Geosphere-Biosphere Program land cover classification. The map covers most of the

Fig. 2. Datasets used in the study. Topographical: USGS DEM, NED, NASA's SRTM and IFSAR (300, 30, 30, and 5 meter resolution, respectively). Land use: USGS GIRAS, NLCD, and NASA MODIS MOD12Q1 (400, 30, and 1000 m resolution, correspondingly).

globe and is updated every year. All the land use datasets included in this research were geo-processed for achieving the same number of land use categories (agricultural, barren, forest, urban, water, and wetland land use classes), so that the summarization of land use information was consistent for input into HSPF, and subsequent comparison of HSPF output.
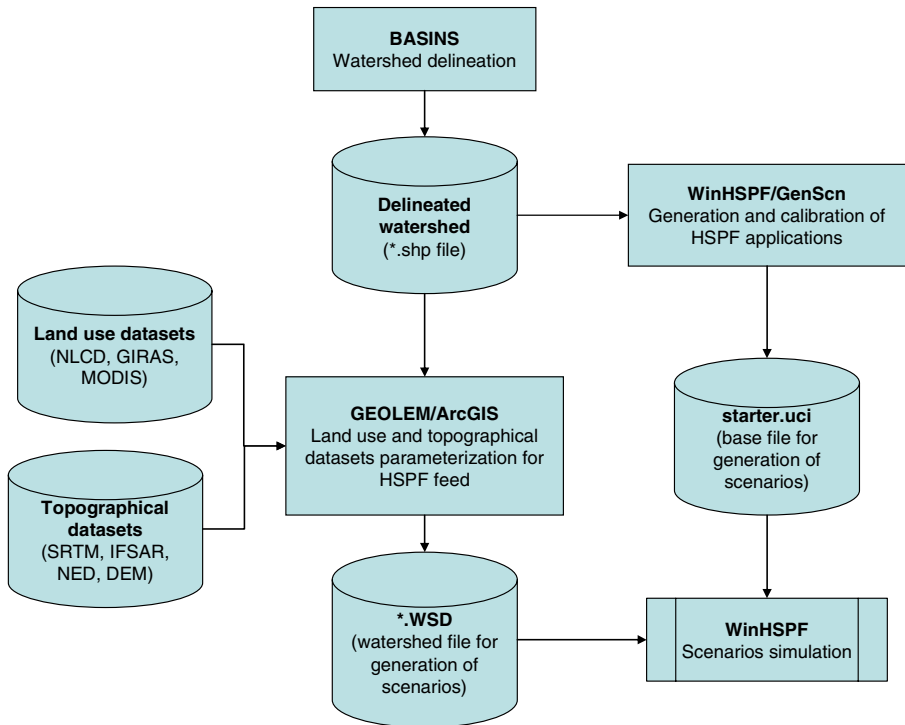
2.3 Hydrological Model

Hydrological modeling of the Jourdan and Wolf watersheds is performed using the Hydrological Simulation Program Fortran (HSPF). HSPF is a computer model designed for simulation of non-point source watershed hydrology and water quality. Time-series of meteorological/water-quality data, land use and topographical data are used to estimate

stream flow hydrographs and polluto-graphs. The model simulates interception, soil moisture, surface runoff, interflow, base flow, snowpack depth and water content, snowmelt, evapo-transpiration, and ground-water recharge. Simulation results are provided as time-series of runoff, sediment load, and nutrient and pesticide concentrations, along with time-series of water quantity and quality, at any point in a watershed. Additional software (WDMUtil and GenScn) is used for data pre-processing and post-processing, and for statistical and graphical analysis of input and output data.

2.4   Geo-processing and Hydrological Modeling

Figure 3 summarizes the geo-processing steps and hydrological modeling approach.
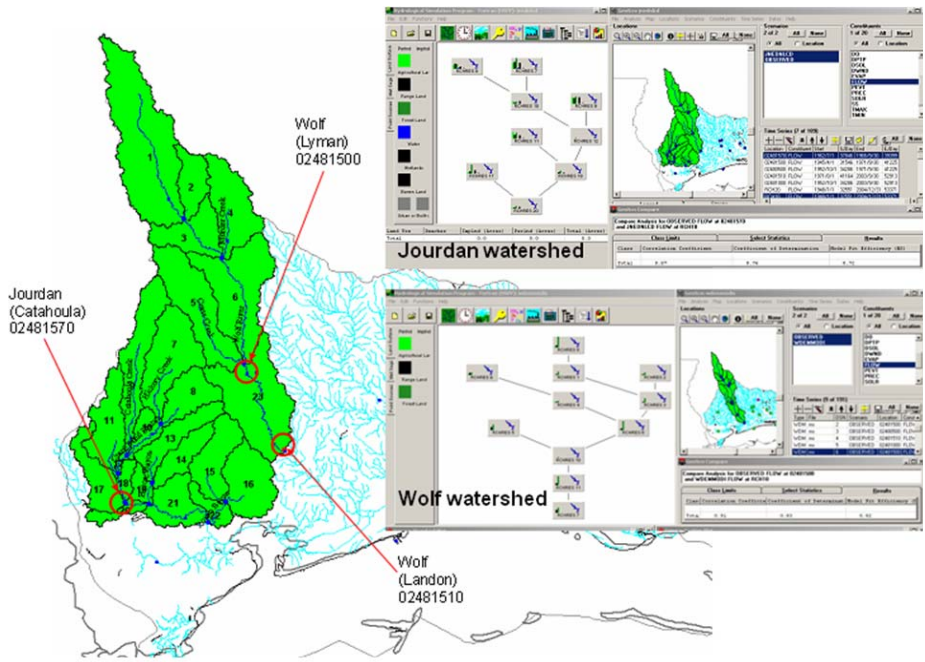

Fig. 3. Flowchart of the geo-processing steps for generating the hydrological model's data input. BASINS, GEOLEM, ArcGIS, and Arc Info were used. These GIS systems extracted land use and topographical information for input into HSPF.

   BASINS was used for downloading basic data and delineating the catchments included in this study. The creation of the initial HSPF models was done using the WinHSPF interface included in BASINS. The generated HSPF models (one for Jourdan River catchment and another for Wolf's, Figure 4) were calibrated for stream flow using the NED and NLCD datasets combination, and the USGS stream flow gauge stations at Catahoula (02481570), Lyman (02481500), and Landon (02481510) (shown in Figure 4). The User Control Input (UCI) files resulting from the calibration

(one per each catchment) were used to modify the HSPF's starter.uci file (this file is used by the HSPF program to generate a HSPF model). In this research, two starter.uci files were created (for each of the watersheds under study). These files allowed preserving the hydrological parameters (resulting from the hydrological calibration process) constant.

The delineated catchments were used as geographical containers for further geo-processing. GEOLEM and ArcGIS were used to parameterize land use and topographical information from the datasets described in sections 2.1 and 2.2. The parameterization was summarized in several *.wsd watershed files (12 per catchment) for different combinations of land use and topographical data. These watershed files and the starter.uci files were used for generating 12 HSPF models per catchment, each of them having the same hydrological parameter values as the corresponding calibrated models. Land use and topographical parameters, however, were different in each of those models, depending on the land use and topographical datasets combined for generating the corresponding WSD file.

Fig. 4. HSPF models for Jourdan and Wolf watersheds. USGS stream flow gauge stations at Catahoula, Lyman, and Landon were used for hydrological calibration of the models.

3   Results

After calibrating and validating the HSPF applications to Jourdan and Wolf watersheds using the finest resolution datasets (NED and NLCD combined), HSPF was used to simulate stream flow hydrographs for each of the 12 combinations shown

in Tables 1 and 2. Those simulated stream flow hydrographs were compared to measured stream flow and the following best-fit coefficients were assessed: coefficient of determination (r^2); Nash-Sutcliff model fit efficiency (NS). The Nash-Sutcliffe coefficient (NS) [12] represents the fraction of the variance in the measured data explained by the model. The NS ranges between minus infinity to one. A NS value of one represents perfect fit. Nash-Sutcliffe coefficient has been used by many researches and is considered one of the best statistical for evaluation of continuous-hydrograph simulation programs [13][14]. The NS is given by the following equation:

$$NS = 1 - \frac{\sum_{j=1}^{n}(O_j - S_j)^2}{\sum_{j=1}^{n}(O_j - \overline{O_j})^2}. \tag{1}$$

In Formula (1), O_j is the observed stream flow at time step j; $\bar{O}$ is the average observed stream flow during the evaluation period; and S_j is the simulated stream flow at time step j.

   The results summarized in Tables 1 and 2, and depicted in Figure 5, show that scale and spatial resolution of topographical and land use datasets do not affect the quality of statistical fit between simulated and measured stream flow data. In fact, the NS and r^2 values for Jourdan River (0.72 and 0.76, respectively) in the base-case scenario (i.e., using NED topography and NLCD land use) are the lowest statistical fit coefficient values. However, the narrow range of NS and r2 values for all the experiments performed for the Jourdan River watershed ($0.72 \leq NS \leq 0.75$; $0.76 \leq r^2 \leq 0.80$) would not allow pointing the best combination of land use and topographical datasets.
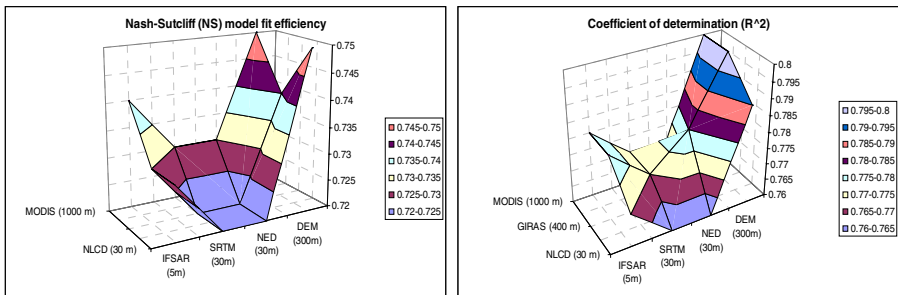
Table 1. Results of statistical fit between simulated and measured stream flow for Jourdan River watershed

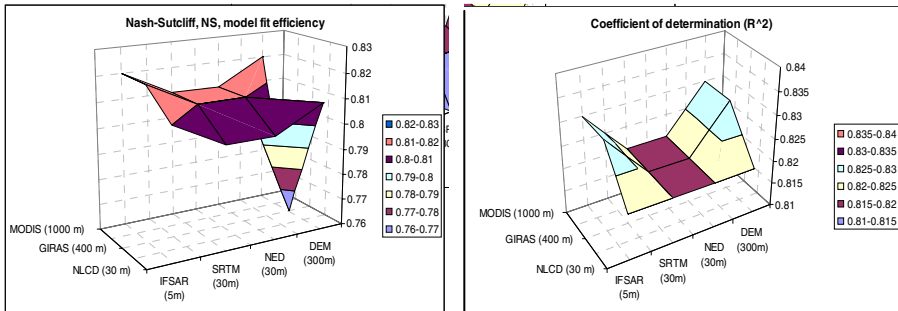| Model fit efficiency (Nash-Sutcliff, NS)			
	MODIS (1000 m)	GIRAS (400 m)	NLCD (30 m)
USGS DEM (300m)	0.75	0.74	0.75
NED (30m)	0.73	0.72	0.72
SRTM (30m)	0.73	0.72	0.72
IFSAR (5m)	0.74	0.73	0.73
Coefficient of determination R^2			
	MODIS (1000 m)	GIRAS (400 m)	NLCD (30 m)
USGS DEM (300m)	0.8	0.8	0.79
NED (30m)	0.77	0.78	0.76
SRTM (30m)	0.77	0.77	0.76
IFSAR (5m)	0.78	0.78	0.77

Table 2. Results of statistical fit between simulated and measured stream flow for Wolf River

Model fit efficiency (Nash-Sutcliff, NS)			
	MODIS (1000 m)	GIRAS (400 m)	NLCD (30 m)
USGS DEM (300m)	0.82	0.76	0.81
NED (30m)	0.81	0.81	0.8
SRTM (30m)	0.81	0.81	0.8
IFSAR (5m)	0.82	0.82	0.81
Coefficient of determination R^2			
	MODIS (1000 m)	GIRAS (400 m)	NLCD (30 m)
USGS DEM (300m)	0.83	0.83	0.82
NED (30m)	0.82	0.82	0.82
SRTM (30m)	0.82	0.82	0.82
IFSAR (5m)	0.83	0.83	0.82

Jourdan River watershed

Wolf River watershed

Fig. 5. Graphical comparison of statistical fit results for Jourdan and Wolf watersheds. Nash-Sutcliff (NS) and r^2 coefficients are show in the vertical axis.

In the case of measured and simulated stream flow time-series comparison for Wolf River watershed, the range of values for NS and r^2 is equivalent to that of Jourdan River watershed: $0.76 \leq NS \leq 0.82$; $0.82 \leq r^2 \leq 0.83$.

In particular, the combination of moderate resolution topographical datasets (such as SRTM, 30 m) and coarse resolution land use datasets (such as MODIS, 1000 m) produce good statistical fit between simulated and measured stream flow hydrographs. Model fit coefficients (r^2 and NS) for the MODIS-SRTM combination range between 0.73 and 0.81. Interestingly, the coarsest resolution datasets (MODIS and DEM) do not provide the worst statistical fit between measured and simulated stream flow.

4 Conclusions

The results from this research show that scale and spatial resolution of digital land use and topography datasets do not affect the statistical fit of measured and simulated stream flow time-series, when using the Hydrological Program Fortran (HSPF) as the hydrological model recipient of land use and topographical data. In fact, this research shows that when the input to HSPF comes from the coarsest spatial resolution datasets (MODIS land use and USGS DEM topography), HSPF simulated stream flow show equivalent statistical fit to measured stream flow as when finer spatial resolution datasets are combined. This suggest that stream flow simulation is not sensitive to scale or spatial resolution of land use and/or topographical datasets.

Since HSPF is a lumped-parameter hydrological model, the summarization of physiographic information (before it is input to the model) is a required step. This pre-processing of raw land use and topographical data (re-classification, averaging of data per sub-basin, etc.) may be a factor that explains the insensitivity of simulated stream flow to land use and topographical datasets swapping, when using HSPF. Future studies should explore the effects that this swapping may have when using distributed hydrological models.

All the hydrological models generated in this research were developed using mainly precipitation and evapo-transpiration time-series data, primarily because HSPF only requires those time-series for simulation of stream flow. For future water quality modeling of the Jourdan and Wolf watersheds, other weather time-series data (such as temperature, etc.) will be included.

A thorough study on the relationship of soil moisture and overestimation/ underestimation of stream flow will also be included in future researches.

References

1. Manfreda, S., Rodriguez-Iturbe, I.: On the Spatial and Temporal Sampling of Soil Moisture Fields. Water Resources Research 42(5), W05409.1–W05409.10 (2006)
2. Bicknell, B.R., Imhoff, J.C., Kittle, J.L., Jobes, T.H., Donigian, A.S.: HSPF Version 12 User's Manual. National Exposure Research Laboratory. Office of Research and Development U.S. Environmental Protection Agency (2001)

3. Environmental Protection Agency: BASINS: Better Assessment Science Integrating Point & Nonpoint Sources: A Powerful Tool for Managing Watersheds, `http://www.epa.gov/waterscience/BASINS/`
4. Viger, R.J.: GEOLEM: Improving the Integration of Geographic Information in Environmental Modeling Through Semantic Interoperability. Master's Thesis, University of Toronto, Canada (2004)
5. Alarcon, V.J., McAnally, W., Diaz-Ramirez, J., Martin, J., Cartwright, J.: A Hydrological Model of the Mobile River Watershed, Southeastern USA. In: Maroulis, G., Simos, T.E. (eds.) Computational Methods in Science and Engineering: Advances in Computational Science, vol. 1148, pp. 641–645. American Institute of Physics (2009)
6. Alarcon, V.J., McAnally, W., Wasson, L., Martin, J., Cartwright, J.: Using NEXRAD Precipitation Data for Enriching Hydrological and Hydrodynamic Models in the Northern Gulf of Mexico. In: Maroulis, G., Simos, T.E. (eds.) Computational Methods in Science and Engineering: Advances in Computational Science, vol. 1148, pp. 646–650. American Institute of Physics (2009)
7. Environmental Protection Agency: USGS 30 Meter Resolution, One-Sixtieth Degree National Elevation Dataset for CONUS, Alaska, Hawaii, Puerto Rico, and the U. S. Virgin Islands, `http://www.epa.gov/waterscience/basins/metadata/ned.htm`
8. USGS: Earth Resources Observation and Science, `http://edc.usgs.gov/products/elevation.html`
9. Intermap Technologies Inc.: IFSAR product handbook, `http://www.intermap.com/images/handbook/producthandbook.pdf`
10. Alarcon, V.J., O'Hara, C.G., McAnally, W., Martin, J., Diaz, J., Duan, Z.: Influence of elevation dataset on watershed delineation of three catchments in Mississippi, `http://www.gri.msstate.edu/research/nasa_rpc/papers/AWRA_elevation_2006.pdf`
11. NASA: MODIS MOD12 Land Cover and Land Cover Dynamics Products User Guide, `http://www-modis.bu.edu/landcover/userguidelc/lc.html`
12. Nash, J.E., Sutcliffe, J.V.: River flow forecasting through conceptual models, part I: a discussion of principles. Journal of Hydrology 10, 282–290 (1970)
13. Engelmann, C.J.K., Ward, A.D., Christy, A.D., Bair, E.S.: Application of the BASINS database and NPSM model on a small Ohio watershed. Journal of the American Water Resources Association 38(1), 289–300 (2002)
14. Legates, D.R., McCabe, G.J.: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resources Research 35(1), 233–241 (1999)