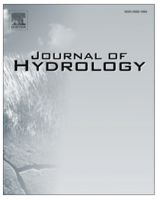
Journal of Hydrology 530 (2015) 206–215

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol



A conceptual data model coupling with physically-based distributed hydrological models based on catchment discretization schemas

Yuanming Liu [a], Wanchang Zhang [a,b,*], Zhijie Zhang [c]

[a] State Key Laboratory of Pollution Control & Resource Reuse, School of Environment, Nanjing University, Nanjing 210093, China
[b] Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China
[c] School of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210093, China

ARTICLE INFO

Article history:
Received 23 July 2015
Received in revised form 19 September 2015
Accepted 21 September 2015
Available online 30 September 2015
This manuscript was handled by Geoff Syme, Editor-in-Chief

Keywords:
Conceptual data model
Geodatabase
Discretization schemas
HRU
Model coupling
DHM

SUMMARY

In hydrology, the data types, spatio-temporal scales and formats for physically-based distributed hydrological models and the distributed data or parameters may be different before significant data pre-processing or may change during hydrological simulation run time. A data model is devoted to these problems for sophisticated numerical hydrological modeling procedures. In this paper, we propose a conceptual data model to interpret the comprehensive, universal and complex water environmental entities. We also present an innovative integration methodology to couple the data model with physically-based distributed hydrological models (DHMs) based on catchment discretization schemas. The data model provides a reasonable framework for researchers of organizing and pre-processing water environmental spatio-temporal datasets. It also facilitates seamless data flow fluid and dynamic by hydrological response units (HRUs) as the core between the object-oriented databases and physically-based distributed hydrological models.

1. Introduction

In hydrology, the physically-based distributed hydrological models (DHMs) is employed to address a wide spectrum of environmental and water resources problems. Mathematical modeling throughout the DHM is closely related to data acquisition and processing by sensor and internet of things technology, remote sensing and space technology, geographic information systems (GISs), digital elevation models (DEMs), topographic representation and object-oriented design, etc., as shown in Table 1 (Singh and Woolhiser, 2002). The data types, spatio-temporal scales and formats for the DHMs and the distributed data or parameters may be different before significant data pre-processing or may change during hydrological simulations. The DHMs simulate hydrological state variables in space and time while using heterogeneous datasets such as the DEM, climate, land use, topography, and hydrogeology. However, the input data may directly affect the hydrological simulations. The DEM spatial resolution might have minor impacts on flood simulation results, but inundation areas from flood simulations varied significantly across different DEM data sources (Li and Wong, 2010). It was proved that the accuracy and resolution of the input DEM had serious implications on the values of important spatial indices in hydrology derived from the DEM (Vaze et al., 2010). The latest study on spatio-temporal soil moisture patterns showed that the soil moisture influenced by many factors (e.g., topography, soil properties, vegetation types, management, and meteorological conditions), and the negative linear relationships between the coefficient of variation and the mean soil moisture indicated lower spatial variability at higher mean soil moisture (Korres et al., 2015). Moreover, because the interactions between the traditional spatial database management systems (DBMSs) and GIS-centric hydrological models lead to tight coupling, the time necessary to prepare data with a particular format, conFigure data sources, perform computations and analyse results prevents experimenters from performing sophisticated modeling procedures that use extensive data (Bray, 2008). The data requirements and the difficulty of constructing input parameters, however, have long been obstacles to cost-effective professional simulations. These limitations can be largely offset using an object-oriented model that addresses the definition of data in space and time, such as the common coordinate reference systems (CRSs), resolution and bi-temporal.

---

* Corresponding author at: State Key Laboratory of Pollution Control & Resource Reuse, School of Environment, Nanjing University, Nanjing 210093, China.

E-mail address: wczhang@ceode.ac.cn (W. Zhang).

http://dx.doi.org/10.1016/j.jhydrol.2015.09.049
0022-1694/© 2015 Elsevier B.V. All rights reserved.

Table 1
Data requirements for the calculation of each physical state (modified from Kumar et al., 2010).

Process	Data support
Channel runoff	River segment, DEM, precipitation, evaporation, Manning's coefficient, coefficient of discharge for weir flow across river banks, leakage coefficient, boundary conditions
Interception	Interception storage capacity, precipitation, Leaf Area Index (LAI), Normalized Difference Vegetation Index (NDVI), evapotranspiration, initial interception
Evapotranspiration	Wind speed, humidity, net radiation, soil heat flux, vapor pressure deficit, mean air density, interception storage capacity, LAI, NDVI, soil saturation, atmospheric resistance, stomatal resistance, vegetation fraction, unsaturated zone saturation
Snow melt	Initial snow depth, initial snow density, initial snow surface layer temperature, initial average snow cover temperature, average snow liquid water content, net solar radiation, incoming thermal radiation, air temperature, vapor pressure, wind speed, soil temperature, precipitation
Infiltration	River segment, unsaturated soil moisture, hydraulic conductivity, porosity, macropore density, precipitation rate, maximum infiltration capacity
Surface runoff	River segment, net precipitation, evapotranspiration, DEM, boundary conditions
Underground runoff	DEM, capillary flow, hydraulic conductivity, bedrock depth, soil porosity, soil parameters, boundary conditions

A data model could be devoted to these problems for sophisticated numerical hydrological modeling procedures, which facilitates the flow of seamless data between the database and hydrological simulations (Kumar et al., 2010). For this purpose, both semantic data models and hydrological models have achieved fruitful research results in their respective fields. There was a massive international effort in the field of data modeling for the water environment to share these data properly among applications within semantic specifications. For example, the WaterML 2.0 schema for time series observations, at specific locations, about climate, river flow and water quality, etc., proposed by the Open Geospatial Consortium (OGC) hydrology domain working group, was used to streamline data collection and reporting worldwide (Botts et al., 2008; Mauree, 2010; Zaslavsky et al., 2007). Recently, under the umbrella of the hydrological information system (HIS) project, the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) developed a variety of web service models providing access to large repositories of observed hydrological data (Ames et al., 2012; Mauree, 2010). Since 2007, to facilitate the interoperability of hydrographic information between member states, the Infrastructure for Spatial Information in the European Community (INSPIRE) Data Specifications for the Hydrography theme provided an outstanding framework for object-oriented data modeling, within which the generic model components were extended to the complex network model (INSPIRE, 2010).

In contrast, previous studies on hydrological simulations focused on the underlying connections between the hydrological model and a Geographic Information System (GIS). The most familiar applications include the Hydrologic Data Development System (HDDS), employing Arc/Info; the Water and Erosion Prediction Project (WEPP), implemented through GIS interfaces; the Better Assessment Science Integrating point & Non-point Sources (BASINS), which integrated an open-source MapWindow GIS interface; and the Soil & Water Assessment Tool (SWAT), based on the ArcGIS. (Cochrane and Flanagan, 1999; Kinerson et al., 2009; Neitsch et al., 2004; Smith, 1997).

Although various DHMs have been developed and have demonstrated great analytical progress, the simulation of any specific catchment is subject to the challenges of data pre-processing. First, the complex structure of the input dataset generates a difficult processing challenge for researchers who are not familiar with all of the DHMs (Refsgaard, 1997; Wang et al., 2009). Second, due to the lack of a solid integration methodology to facilitate seamless data flow between the databases and hydrological applications, researchers are rarely encouraged to consider the organizational optimization of input data for precise results. Fortunately, the current research illustrates the feasibility of semantic models coupled with hydrological process models. For example, the Arc Hydro created cooperatively by the Environmental Systems Research Institute (ESRI) and the Center for Research in Water Resources (CRWR) at the University of Texas, greatly expanded our knowledge of hydrological data modeling and large datasets preprocessing (Maidment, 2002). However, the Arc Hydro itself contained no routines to simulate hydrologic processes (Chinh et al., 2010; Whiteaker et al., 2006). Meanwhile, the increased data modeling convenience and hydrological simulation stability achieved by complete adherence to the commercial ESRI ArcGIS framework came at the price of decreased portability and compatibility, as shown in Table 2.

In this paper, we propose a conceptual data model to reasonably couple with the DHMs. The model is based on analysis of not only the data requirements for expected simulation purposes like other data models above, but also the catchment discretization schemas for hydrological units in the DHMs that other have not considered. The model would be capable of interpreting the comprehensive,

Table 2
Semantic models coupled with hydrological process models.

Semantic models	Features	Shortages
Arc Hydro by ESRI, Inc.	Introduced a major development by defining a solid data model within ESRI ArcGIS (Maidment, 2002)	Only provided a framework for pre-processing datasets without routines to simulate hydrologic processes (Chinh et al., 2010; Whiteaker et al., 2006)
GIS-centric conceptual model by Kumar et al. (Kumar et al., 2010)	A locally scoped GIS-centric conceptual model for coupling a Penn State Integrated Hydrological Model	Without specific model coupling mechanism according to the common hydro-physical characteristics of the DHMs
Data specification on Hydrography by INSPIRE (INSPIRE, 2010)	Provided a solid framework for modeling involved with description of lakes, rivers and other waters, with their phenomena	
Conceptual model by World Meteorological Organization Commission for Hydrology	Reconciled the underlying differences in the representation of hydrological features and levels of detail in typical datasets	Subsequent work was underway to reconcile hydrological models (Atkinson et al., 2012)
Open Modeling Interface	Provided a standardized interface to define, describe and transfer data on a time basis between coupled components (Gregersen et al., 2007)	Frequent data exchange between multiple models would reduce the efficiency based on the request-response mode (Voinov and Cerco, 2010)