

A GIS-based relational data model for multi-dimensional representation of river hydrodynamics and morphodynamics



Dongsu Kim ^{a,*}, Marian Muste ^{b,1}, Venkatesh Merwade ^{c,2}

^a Civil and Environmental Engineering, Dankook University, 126 Jukjeon-dong, Yongin-si, Gyeonggi-do, South Korea

^b IHR-Hydroscience & Engineering, The University of Iowa, 323E Hydraulics Lab, Iowa City, IA 52242, USA

^c Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN 47907, USA

ARTICLE INFO

Article history:

Received 21 March 2014

Received in revised form

26 November 2014

Accepted 2 December 2014

Available online 27 December 2014

Keywords:

Arc River

River data model

River hydrodynamics

River morphodynamics

Spatio-temporal GIS

ABSTRACT

The emerging capabilities of the geo-based information systems to integrate spatial and temporal attributes of in-situ measurements is a long-awaited solution to efficiently organize, visualize, and analyze the vast amount of data produced by the new generations of river instruments. This paper describes the construct of a river data model linked to a relational database that can be populated with both measured and simulated river data to facilitate descriptions of river features and processes using hydraulic/hydrologic terminology. The proposed model, labeled Arc River, is built in close connection with the existing Arc Hydro data model developed for water-related features to ensure the connection of the river characteristics with their floodplains and watersheds. This paper illustrates Arc River data model capabilities in conjunction with Acoustic Doppler Current Profiler measurements to demonstrate that essential river morphodynamics and hydrodynamics aspects can be described using data on the flow and its boundaries.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The traditional approach for experimentally describing river hydrology is based on the one-dimensional river network representation, with streamflow and cross-section specifications (geometry, bed and bank materials, other local aspects) as main hydrodynamic and morphologic quantities. The conventional instruments used for the one dimensional representation include a variety of instruments that separately measure flow or bathymetric properties (e.g., current meters and sounders). These data and information has been extensively used for river monitoring purposes and for calibrating and/or validating not only one (1D) dimensional numerical simulation models, but also of two (2D) or three dimensional (3D) ones. Given that the capabilities of the simulation models to replicate complex flow features are still limited irrespective of their dimensionality, the 1D measurements acquired in situ have played a critical role in the search for solutions to

emerging riverine problems such as sediment and pollutant transport, floods, and eco-habitat restoration (Fischer et al., 1979; Chaudhry, 1993; Julien, 1998; Chanson, 2004).

The advent of the new generation of acoustic and optical instruments has marked a great leap in our capability to acquire data in rivers. For example, the Acoustic Doppler Current Profilers (ADCPs) measure instantaneously 3-dimensional velocity components in verticals along with the river cross-section geometry from a single measurement across the rivers (i.e., transect). Not only that ADCPs revolutionized the routine river measurements by enabling efficient and safe acquisition of discharges in a fraction of the time needed by conventional instruments but they also document river behavior through comprehensive, multi-dimensional field datasets with increased spatial and temporal resolutions. For example, ADCP measurements acquired in less than 30 min in a 5-m deep, 300-m wide cross-section in Pool 15 of the Mississippi river entail approximately 8000 velocity vectors and bathymetry profiling (Kim and Muste, 2012). Acquisition of transects over river segments allow to quantify the hydrodynamics and morphodynamics at the reach scale (up to 5 km per day), which was unthinkable with the previous generation of river instruments (Muste et al., 2012). Consequently, the ADCPs are rapidly emerging as tools of choice for monitoring, exploring, and understanding river features through in-situ measurements (Rennie, 2002; Kostaschuk et al., 2004;

* Corresponding author. Tel.: +82 31 8005 3611.

E-mail addresses: dongsu-kim@dankook.ac.kr (D. Kim), marian-muste@uiowa.edu (M. Muste), vmerwade@purdue.edu (V. Merwade).

¹ Tel.: +1 319 384 0624; fax: +1 319 335 5238.

² Tel.: +1 765 494 2176.

Muste et al., 2004; Dinehart and Burau, 2005; Nystrom et al., 2007; Szupiany et al., 2007; Sime et al., 2007; Kim and Muste, 2012).

The measurement of stream discharges has been, and continues to be, in most cases, the primary use of the ADCPs (RDI, 1996; SonTek, 2000). Consequently, the software currently associated with ADCPs is designed to control the instrument operation and estimate flow rates. The ADCPs raw data contain, however, a wealth of velocity-derived (dynamic) and river geometry (static) information that is currently not fully exploited because the lack of adequate tools for data processing and information extraction. Handling, processing, and extracting information and knowledge from this vast amount of raw data with the available instrument-associated software is quite limited, therefore recently efforts have been made to develop new and efficient approaches for storing, visualizing, and retrieval of the features of interest for analyzing river processes (Kim and Muste, 2012; Parsons et al., 2013). This paper explores the capabilities of the geographic information system (GIS) technology to integrate multi-dimensional static and dynamic data on river processes in one representation framework. Combining the productive power of the new generation of instruments with the efficiency of the GIS tools for representation and analysis of the data leads to tools with a convenient visualization and synthesis of data over a range of spatio-temporal scales which brings in a paradigm shift in understanding and managing river related issues by handling comprehensive and multilayered information.

Currently, there are several GIS-based data models designed for water resources area. They include Arc Hydro (Maidment, 2002), Arc Hydro Groundwater (Strassberg et al., 2011), and Arc Marine (Wright et al., 2007) developed for surface water hydrology, groundwater and ocean datasets, respectively. These data models are gaining popularity in their respective communities for both incorporating observed and simulated multi-dimensional datasets. Notably, the Arc Marine and Arc Hydro Groundwater models accommodate 3D dynamic datasets and track features and processes in space and time by using the underlying geodatabase model. These GIS-based data models enable efficient integration of diverse multidimensional features and use of a variety of GIS functionalities for advanced data analyses. Another essential feature of the above-mentioned data models is that they are built using the RDBMS (Relational Database Management System) structure. The RDBMS feature distinguishes the GIS data models from the traditional file-based data storage through their capabilities to efficiently store the information in hierarchical tables connected through internal and external relationships, quick search, retrieval, and distribution of data. Collectively, the creation of these data models has laid groundwork for more efficient and intuitive organization of the data and information characterizing aquatic environment.

The popular Arc Hydro data model for hydrology can only represent streams as a 1D network, and river-bank and cross-section geometry as 2D features. Despite that rivers play a key role in many hydraulic and hydrologic studies, there is no a similar comprehensive multi-dimensional data model for the riverine environment. Motivated by the recent availability of multidimensional hydrodynamic data collected with new instruments and techniques (e.g., ADCPs, Large-Scale Particle Image Velocimetry), advances in GIS technologies along with the calls for building comprehensive hydrologic information systems, this paper presents a customized GIS-based river data model called 'Arc River'. Creation of the Arc River model enables cross-disciplinary network analysis along the lines discussed by Hodges (2013) and Liu and Hodges (2014) by creating a river data model whereby data and models interact seamlessly irrespective of the river process that is observed. This paper illustrates a use case for the Arc River data

model in conjunction with Acoustic Doppler Current Profiler data to describe essential river morphodynamics and hydrodynamics aspects that can be obtained from direct field measurements. Arc River is, however, a generic model that can be applied to a plethora of morpho-hydrodynamic processes if data acquired with other instruments are ingested in the model. The model can also be easily extended to any river-related processes by adding them into its classes concepts and terminology associated with the additional observed features.

2. Arc river data model structure

Arc River is a customized GIS-based data repository built on top of ESRI's geodatabase technology (Arctur and Zeiler, 2004). Arc River is designed to: (i) represent river data in a curvilinear coordinate system to support river channel oriented spatial analyses; (ii) represent multidimensional river features through points, lines, polygons, and volumes; (iii) represent simulated gridded data for river channels that can be readily coupled with observed data; (iv) represent spatio-temporal evolution of dynamic river objects (such as bedforms) within single or a series of events using Eulerian or Lagrangian observational frameworks (Currie, 1993), and (v) efficiently store and retrieve data acquired in-situ along with the ancillary metadata. In the Eulerian framework, the river characteristics are observed from a fixed location through a well-defined observation window, whereas in the Lagrangian framework the river characteristics are observed by following their evolution in space and time as individual entities.

A typical procedure for data model development involves the following steps:

- (i) development of a conceptual design and specifications to account for the physical objects and their natural relationships. These domain-specific features will provide the foundation for data representation on the data model;
- (ii) conversion of the conceptual design to a graphical representation by using design tools such as the ArcInfo® Unified Modeling Language (UML) enclosed in Microsoft Visio (ESRI, 2007);
- (iii) conversion of the UML diagram to a computer readable format, such as Extensible Markup Language (XML) obtained with a CASE tool (ESRI, 2007); and,
- (iv) instantiation of a geodatabase by reading the XML file of UML diagram using ArcCatalog®, where the internal structures of the geodatabase (classes) are created. Once a geodatabase is created, all the feature classes are populated with the actual data.

Essential elements of the UML diagram for the Arc River model are shown in Fig. 1. The major conceptual groups of the Arc River data model are three feature datasets (RiverHydroFeatures, RiverFeatureSeries, RiverGrid) and a raster dataset (RiverRasterCatalog). RiverHydroFeatures data group consists of various feature classes that systematically stores fundamental river objects such as cross-sections, river points, and river geometry. RiverFeatureSeries contains a series of features that represent hydrodynamic variation of river objects in both space and time. RiverGrid feature dataset handles 2D or 3D meshes and grid point, and RiverRasterCatalog accumulates stacks of raster data. All features are supplemented with a group of tables as shown in 'Data' box in Fig. 1. The data information tables contain ancillary information such as the data provenance (measured or simulated), variable types, instrument used for the measurements, methods, and group. The time series of spatially distributed rasters, labeled RiverRasterCatalog, are separately recorded.

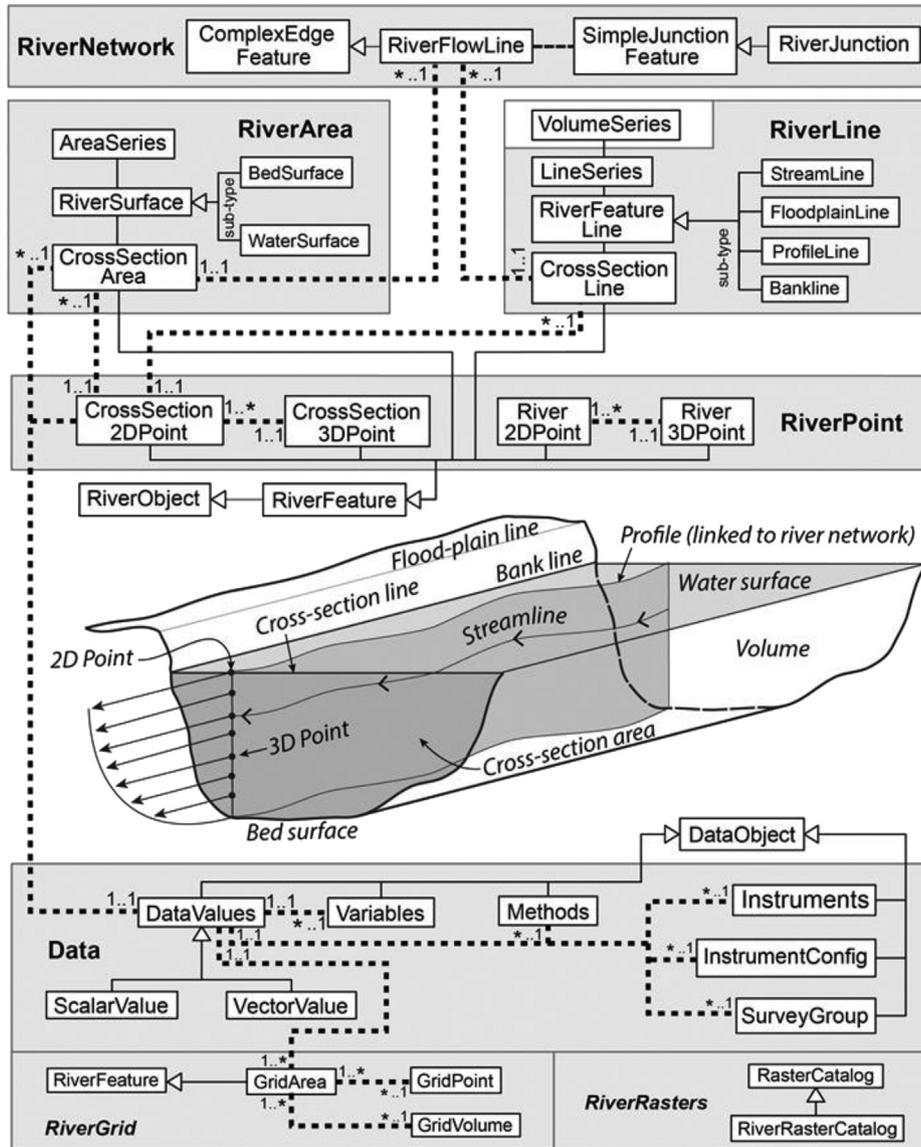


Fig. 1. Classes and their relationships in the Arc River data model framework.

3. Arc river functionality

3.1. Representation of river data in a curvilinear coordinate system

In the standard Cartesian coordinate system, GIS objects are represented in space by $x, y,$ and z -coordinates. ArcGIS also includes linear referencing where objects are represented with reference to a linear line such as a river or street. For example, a boat ramp along a river can be referenced as 100 m downstream from the mouth of the river instead of giving its (x,y) coordinates. This approach to represent objects along a river is referred to as flow oriented or curvilinear orthogonal coordinate system (Merwade et al., 2005). The system uses (s,n,z) coordinates for object location, where s is distance along the centerline of the river, n is perpendicular to river centerline, and z is the vertical distance with reference to some datum (see Fig. 2). ESRI ArcGIS® supports curvilinear coordinates such as *Point ZM*, *Polyline ZM*, and *PolygonZM*.

The Arc River data model assigns both curvilinear coordinates as well as Cartesian coordinates for object representation. The flow-oriented coordinate system has been traditionally used in river

applications such as river mile tracking for navigation purposes, or for locating various features along the river channels. This approach enables accurate identification of the distance between two points along the river channel as being different from the absolute distance between the points (see Fig. 2). This representation is easily understood and used for various river specialized analyses.

3.2. Representation of multidimensional river features through points, lines, polygons, and volumes

Arc River is capable of storing data in 1, 2, and 3 dimensions to represent various river features as presented in Fig. 3. In river analysis, 1D information consists of averaged hydrodynamic quantities over a cross-section such as stream discharge, mean bulk flow velocity, Froude number, and longitudinal dispersion coefficient. In this context, the cross-section line or area containing the information is regarded as a 1D object. Information such as depth averaged distributions or horizontal distributions of velocities or bed shear stress are, however, 2D, with the cross-section points considered as their attributes. 3D objects are introduced to

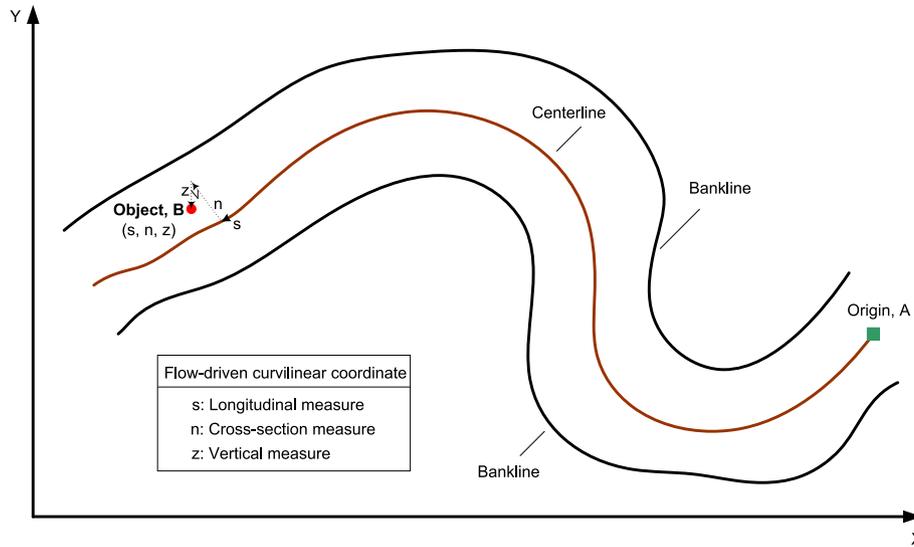


Fig. 2. Three-dimensional curvilinear flow-oriented coordinate system.

accommodate information such as velocity distributions along vertical, lateral and longitudinal directions. Arc River data model links 1D, 2D and 3D river objects through a common base: the river network (see Fig. 3).

For compatibility purposes, the connection between multidimensional objects through the river network in Arc River is established using the terminology of the Arc Hydro data model (Maidment, 2002). In other words, a river network is the backbone of Arc River, and it is related to all other river objects through a common identifier. A river network can be extracted from existing datasets such as the National Hydrography Dataset (NHD) (<http://nhd.usgs.gov>) or it can be custom created. For example, for a relatively small scale river reach analysis (not so accurately represented in NHD), where detailed bathymetry data are available, the river network can be locally constructed using thalweg line generation algorithms such as the one developed by Merwade et al. (2005). However, for most practical purposes, public domain data such as NHD is a good starting point.

Fig. 3 illustrate the connectivity between Arc River multidimensional objects (point features in this case) defined in a cross-

section and the associated river network defined in Arc Hydro whereby river lines are described through HydroEdge feature class. In the context of ADCP measurements, Arc Hydro's HydroEdge feature is associated to an ADCP transect or fixed-point measurement depending on the procedure used to acquire the data. These measurements are then used to generate Arc Hydro's HydroNetwork which topologically connects stream lines and points in conjunction with the one-dimensional description of the river network.

Connectivity between Arc River objects and the underlying network is based on both spatial proximity and feature dimensions. For example, if a river object (e.g., point) is closer to one HydroEdge feature compared to all other HydroEdge features in the river network, then the object becomes connected to the nearest HydroEdge. In addition, river objects themselves are related to one another based on their dimensions. For instance, the CrossSectionLine shown in Fig. 4a is identified as a polyline connecting point measurements across a cross-section and classified as a 1D object, because it contains cross-section averaged 1D quantities such as mean velocity and discharge. The CrossSectionLine object contains multiple 2D points (having a one to many relationship), which are called as CrossSection2DPoint. So by relating key identifiers, the CrossSection2DPoints can be connected to a CrossSectionLine. Similarly, 3D objects, such as CrossSection3DPoint along the vertical direction of the cross-section, can be related to a 2D object, CrossSection2DPoint (see Fig. 4b).

While typically river measurements are acquired along cross-sections, point measurements (defined as river points in this context) are also often collected such as for water quality monitoring and ADCP campaigns at fixed locations (Muste et al., 2004). Such isolated points are defined in Arc River as River2DPoint and River3DPoint objects. The connectivity rule between River2DPoint and River3DPoint is similar to the one used for cross-sectional objects. The River2DPoint is directly connected to the river network even if it is not exactly located along the network. The River3DPoint is connected to the river network through the River2DPoint object.

The physical, bio-chemical, and morphologic observations in rivers are represented in the model through a collection of feature classes grouped under the 'RiverHydroFeatures' feature dataset. The river objects in the RiverHydroFeatures are defined based on their spatial locations, measurement type, and shapes. The shapes

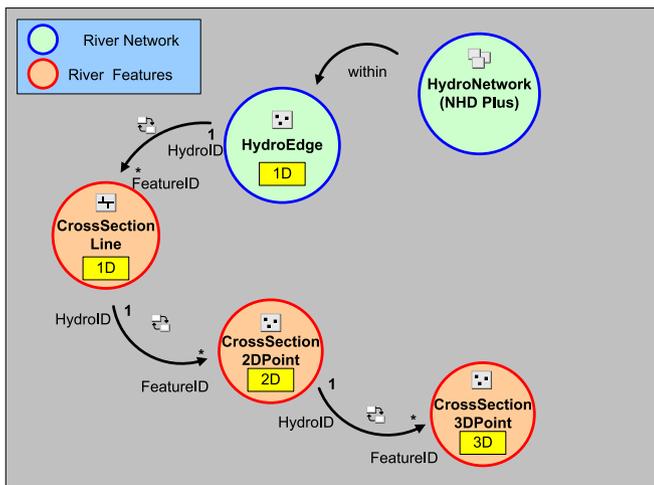


Fig. 3. Connectivity between the multidimensional cross-sectional river objects and the river network.

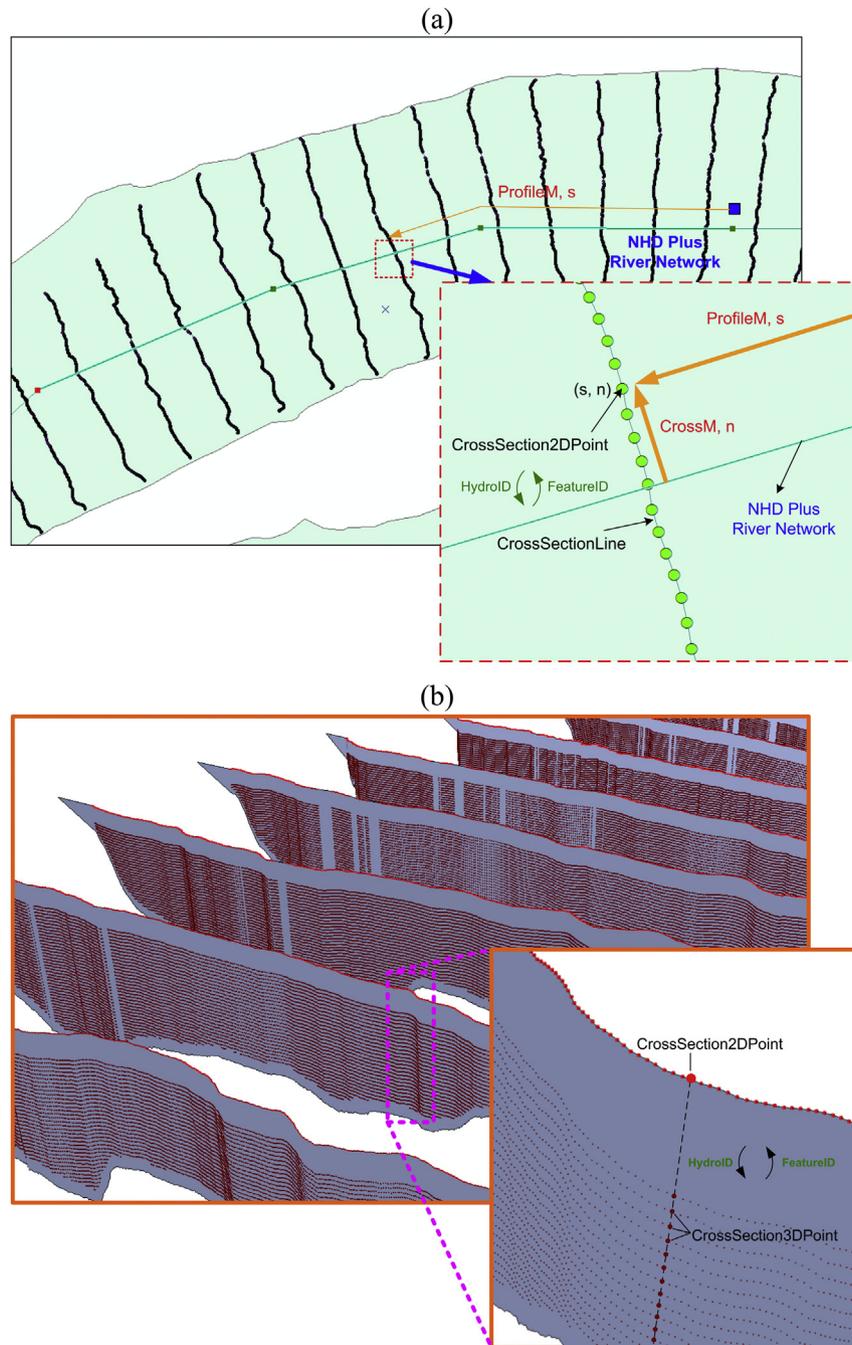


Fig. 4. Diagram of the connectivity between multidimensional river objects in a cross-section and the river network: a) Schematic of the CrossSection2DPoint feature class and its relationship with the river network in plan view; b) Relationship between the CrossSection3DPoint and CrossSection2DPoint in 3D cross-sections.

themselves entail a number of spatial forms such as cross-section, bankline, water surface area, monitoring site and channel volume. RiverHydroFeatures can be built using 1D, 2D or 3D data acquired by a variety of instruments. The data model objects describing the river hydrodynamics and morphology are represented through cross-sections, river points, and river geometry elements. The cross-section objects can be used to represent 1D, 2D or 3D geometric features. Fig. 5 shows an example of a 3D cross-section. River point features can be utilized in a variety of ways to represent river data, but essentially they represent information at isolated points in the flow. The geometry elements objects represent rivers' hydrodynamic and morphodynamic characteristics such as banklines,

free water surface, flow volume, and the thalweg line. All river objects in the RiverHydroFeatures class are linked to river network (e.g., NHD Plus and see Figs. 1, 4 and 11), and flow-oriented coordinates (s, n, z) are accordingly assigned for grounding the features at known domain representation standards. Feature classes are also categorized based on their measurement types (i.e., cross-section and point) and are internally related to one another. For instance, a CrossSectionLine feature class contains many points defined within the CrossSection2DPoint feature class. The relationships between feature classes facilitate connecting all the river objects in the RiverHydroFeatures component and their referencing to the river network.

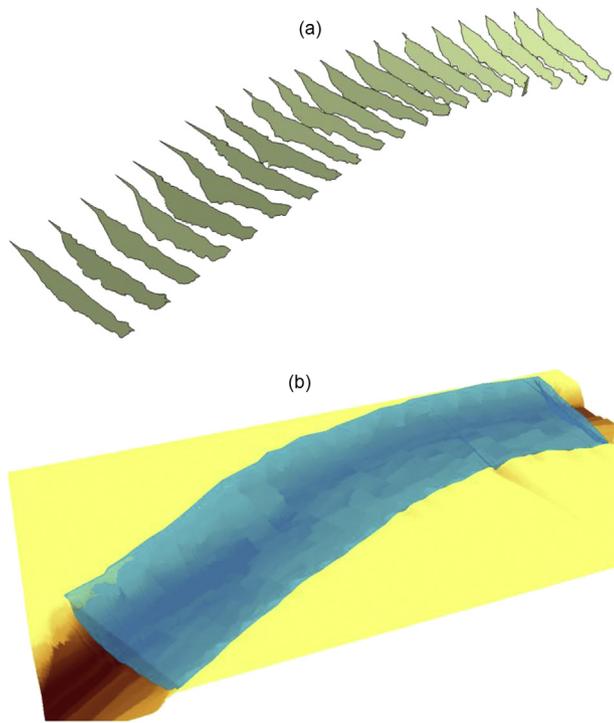


Fig. 5. Geometric features for representing river objects in addition to point features: a) 3D representation of the cross sections using the CrossSectionArea feature class; b) volumetric description of a river reach using RiverVolume feature class.

The RiverHydroFeatures dataset describes the change in hydrodynamics over time using an Eulerian framework, whereby sets of spatially fixed objects are successively represented to capture the dynamics of hydrodynamic attributes passing through observing window. In fact, the feature classes in the RiverHydroFeatures component include only the spatial location and shape. The observed data describing hydrodynamic variation (e.g., velocity time series) are separately stored into time series tables (ScalarValue or VectorValues table) located outside of the RiverHydroFeatures feature dataset. Connecting the feature classes with time series tables ensures that the hydrodynamic values are attached to river objects to describe their variation over time. Analyses of river processes that require representations of other than those described by point or areas related to cross-sections can be accomplished through RiverFeatureLine, RiverFeatureArea, and RiverFeatureVolume. RiverFeatureLine is a line feature class that represents the geometry and location of river reaches. RiverFeatureArea is a polygon feature class for representing area river features such as water surface area, and RiverFeatureVolume is a volume feature class representing the volume of a river reach, as illustrated in Fig. 5a and b, respectively. The attribute of these feature classes contains FeatureTypeName for recording the name of the geometric river object. The RiverFeatureLine has additional descriptors (subtypes) for identification of other line features such as Bankline, Profileline, Streamline, and FloodPlain line. A subtype is a special integer attribute used to define different classifications of features (Arctur and Zeiler, 2004).

3.3. Data and process representation

Observed or simulated data values that describe characteristics of river objects can be expressed as scalars or vectors quantities. In mathematical terms, scalar quantities only describe magnitude, and vector quantities describe both magnitude and direction of the

variable. Most of the data models in water resources handle scalar quantities such as water quality and stream discharge measurements. Hydrodynamic variables such as velocity and shear stress should be represented as vector variables. One approach to accommodate vector variables is to consider the components of a vector as independent values, and store them as different observations. This approach is, however, not adequate because a vector value describes one variable with spatial attributes. Another approach, proposed herein, is to store vector components in the same table of the data model (e.g., using a single row within a table). The Arc River data model stores vectors and scalars in separate tables. The scalar value table is the same as the data value of time series table used in other data models such as Arc Hydro. The vector value table is designed to accommodate multiple components of the vector value, e.g., *DataValueX*, *DataValueY*, and *DataValueZ* for a three-dimensional velocity vector representation. Vector table stores only river properties (not its location) such as velocity components, and it is externally connected to its spatial coordinates such as CrossSectionLine (see Fig. 11).

In contrast with the in-situ measurements where data are typically collected at irregular discrete points, results from simulation models and spatial interpolation are typically available as a grid or mesh with variables associated to grid elements. Comparison of observed and simulated data in the same flow domain requires assigning the quasi-randomly located measurements (e.g., ADCP verticals) to the numerical simulation grid. Another practical situation where data need to be represented in a grid form are observations acquired with image-based techniques where the data are available over a grid pre-established by the user (Creutin et al., 2003). To answer these needs, Arc River contains a RiverGrid feature dataset that can store gridded information using points, lines, polygons or multipatch. The design of the RiverGrid feature dataset is similar to that developed in previous studies (Strassberg et al., 2011; Wright et al., 2007). The Arc River data model separates gridded information from the observation-based RiverHydroFeaturefeature dataset to effectively handle data resulting from numerical simulation or spatial interpolation on in-situ measurements. In principle, Arc River RiverGrid feature dataset can be represented in two or three dimensions as point (GridPoint), area (GridArea), and volume (GridVolume), as illustrated in Fig. 6a. As seen in Fig. 6a, grid point or mesh (or grid area) can be located in either two or three dimensional space. Representations such as those in Fig. 6a and c are especially useful when the 3D data scattered along the cross-section are to be reported onto a specified grid in the cross-section.

Fig. 6b shows attributes of the GridPoint feature class, with HydroID being a unique identifier for each grid point. This class relates grid points with their values in the time series table in order to represent hydrodynamic variation over time. As seen in Fig. 6d, the velocity vector is visualized for each grid point through a relationship involving GroupID attribute. GroupID integrates relevant grid points for spatially distributed representation. For example, grid points in a specific cross-section are grouped as a single group. The GroupID helps to easily extract grouped grid points that are located in the relevant geometric location such as cross-section. GridAreaID is used when grid points represent either the center or nodes of a grid area.

The RiverGrid feature dataset enables time varying description of hydrodynamic or morphodynamic processes using an Eulerian approach. As illustrated in Fig. 6d, velocity vectors can be visualized for each grid point by using such an approach. GridArea feature class integrates a set of polygon features that represent the face of regular or irregular meshes (e.g., cell). Similar to grid points, the grid areas can be located in vertical and horizontal planes. The grid areas illustrated in Fig. 6c represent a collection of cells placed

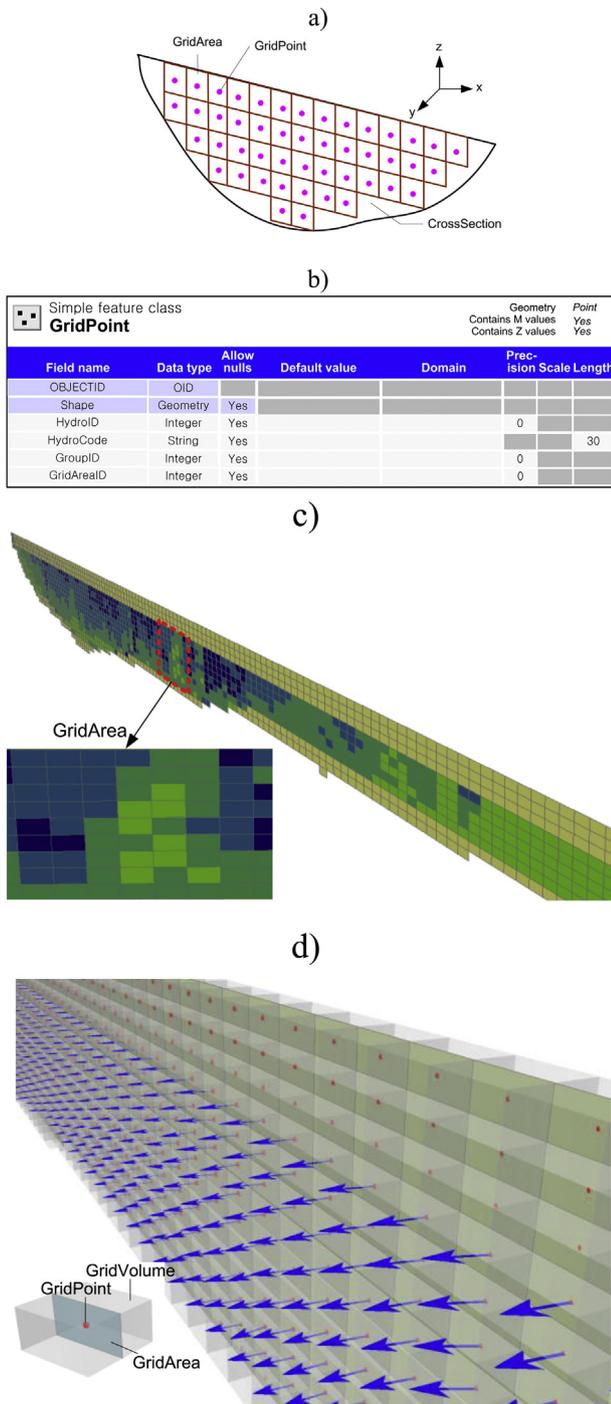


Fig. 6. Grid features in the Arc River data model; a) grid point and area in the vertical plane; b) an example of Grid attribute table (GridPoint); c) grid area in the vertical plane; d) three-dimensional gridded volume where Grid Features are associated instantaneous velocity vectors.

vertically in a cross-section. Each cell in the cross-section contains hydrodynamic quantities that are obtained through spatial interpolation of neighboring point measurements, and visualized with distinct colors commensurate with their values. A group of polygon cells can illustrate the variation of a hydrodynamic parameter over time through changing color. Basically, the color scheme can be recognized as a function of multiple cross-sections. GridVolume feature class integrates a set of multipatch features that represent volumetric finite element cells (e.g., cuboids) in linear or curvilinear flow-oriented coordinate system. The collection of volumetric

GridVolume features can represent either the static or dynamic state of a river volume (see Fig. 6d).

Grid features (point, cell, and volume) can be grouped to represent the hydrodynamic information at a given time. The identifier for each grid feature is relationally linked to its value stored in the time series table (vector or scalar). Once the time series is populated and related to all grid features in a group, the simultaneous representation of two- or three-dimensionally distributed quantities (e.g., velocity vector) in space and time is enabled. In the Arc River data model, the group concept used in grid features is capable of representing any grid shape in two- or three-dimensions. Fig. 6d illustrates grid point (with vector attached), grid area and grid volume. Velocity vectors at each grid point can vary over time to capture the dynamics of the flow over a cross-section. While grid features are spatial representations in 2D or 3D, their position and shape are not changing over time, thus providing typical examples for the Eulerian type framework. A related representation approach was previously developed in the Arc Hydro data model using the 'HydroResponseUnits' concept (Maidment, 2002).

Many observations and measurements acquired as spatially distributed datasets can be represented as raster data in GIS. Raster datasets are commonly used to describe pixel-based distribution of variables in the form on an image or a grid. The raster representation enables the data to be described as a matrix of cells, where each cell represents a pixel with some value stored at its center (Zeir, 1999). Examples of raster surfaces in river applications include elevation of river beds or spatial distribution of hydrodynamic quantities over a portion of the river reach. Similar to grid feature representation, a series of rasters indexed by time stamp can describe the spatial variation of a specific layer over time by stacking the time series of rasters for a given domain. Arc River accomplishes this capability using the raster catalog in ArcGIS® in conjunction with their ancillary information. ESRI geodatabase technology supports RasterCatalog dataset which can store a series of rasters and their attributed information (e.g., time and name) in tables. The RiverRasterCatalog in the Arc River data model is designed to accommodate rasters which describe spatial and temporal variation of river hydrodynamic or morphologic objects in a given river reach.

Fig. 7 represents 3D view of a stack of rasters within the RiverRasterCatalog which changes over time, thus providing an Eulerian view of the represented data or process. The attributed information regarding a series of raster is stored in a table within the RasterCatalog. Each row in an attribute table is linked to a raster. In addition, the attribute table stores temporal information as well as the type of variable that the raster represents such as elevation and bed shear stress distribution. The structure of the attribute table for the RiverRasterCatalog is as follows (Fig. 7a): The 'shape' is geometric type of raster dataset. While the polygon (or cell) is a typical raster shape, other shapes such as point and line are also applicable as raster shape. 'Raster' field is a raster data type attribute, which is unique among rasters and links to a raster dataset stored in the raster catalog. In addition, 'Name' field indicates the simplified attribute for describing a raster. Both Raster and Name are default attributes in a raster catalog, and are automatically created by ArcGIS when a raster catalog is instantiated. 'TSDatTime' records a time stamp when field data for generating a raster dataset was collected, which is added to attribute table purposed to represent temporal variation of the stacked rasters. Through VariableID, RiverRasterCatalog is linked to the Variables table which contains a variety of information for variables. The 'GroupID' integrates relevant rasters inside RiverRasterCatalog. For example, GroupID = 1 represents rasters for river bed surface in the Iowa River and GroupID = 2 can represent rasters for the bed shear stress distribution in Wyoming Slough on Mississippi

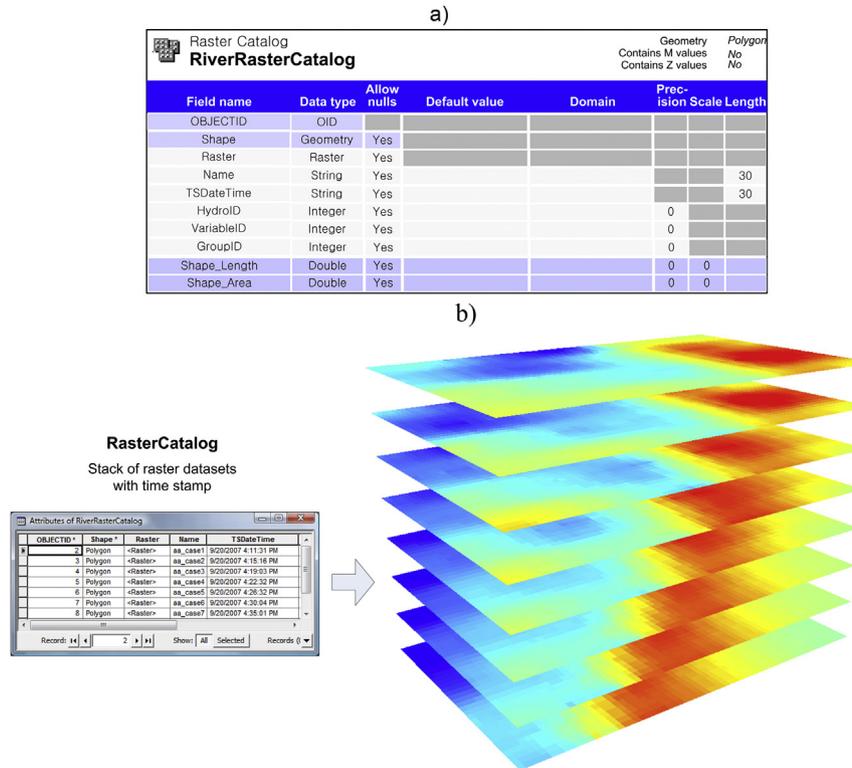


Fig. 7. Representation of RiverRasterCatalog: a) structure of RiverRasterCatalog attribute table indexing relevant rasters; b) Stack of gridded rasters representing the time variation of a spatial attribute (e.g., bedform height).

River. This approach helps in differentiating diverse group of rasters when heterogeneous variables characterizing the river reach are stored in the same raster catalog. TSDatetime, VariableID, and GroupID are not default attributes of the raster catalog table, and should be externally populated when a series of rasters are imported into the RiverRasterCatalog.

3.4. Representation of dynamic river objects in Eulerian and Lagrangian frameworks

The spatio-temporal evolution of the river hydrodynamic and morphologic characteristics can be tracked using either the Eulerian or the Lagrangian framework (Currie, 1993). In the Eulerian framework, the flow is observed from a fixed location within an observation window (point, line, volume). For example, velocity variations can be continuously sampled over time from a fixed location resulting in a time series of the simulated or measured velocity components. Most of the Arc River objects, as well as other GIS representations use the Eulerian approach for representing data. This framework is adequate for characterizing steady state processes or providing a first order interpretation of the data. The Eulerian framework is, however, not best suited for riverine processes that evolve both in space and time (e.g., bedforms, coherent structures, plumes), and thus observing from one location might hinder capturing the dynamics of the process as it unfolds.

The Lagrangian framework observes the hydrodynamic or morphologic entity by moving with the object through space and time. While the movement might change the shape or location of an object that is usually treated as different objects in the conventional GIS framework the object should be represented as the same object in the Lagrangian framework. Arc River model accommodates representation of the hydrodynamic or morphodynamic objects in Lagrangian as well as Eulerian framework. There

are situations whereby the boundary of a given object is changing its shape in time. Fig. 8a exemplifies such a case where the shape of a river bedform moving through a fixed observational area is changing as the flow passes over it. Though this case is similar to the Eulerian description, it should be treated differently from it to store changing boundary. A more complex situation is when the object is actually moving and it is also changing its shape over time such as downstream migration of a pollutant cloud generated by a

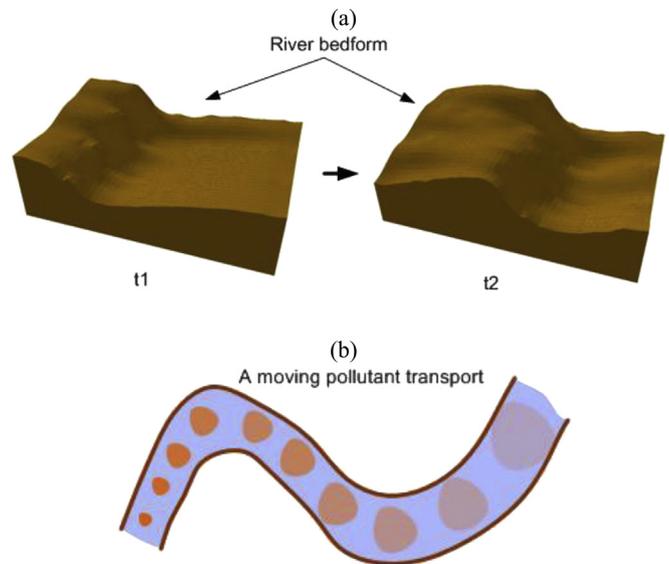


Fig. 8. Stationary and moving objects with deformable boundaries in a Lagrangian framework: a) stationary bedform with the deformable river bed; b) moving objects with the deformable boundary.

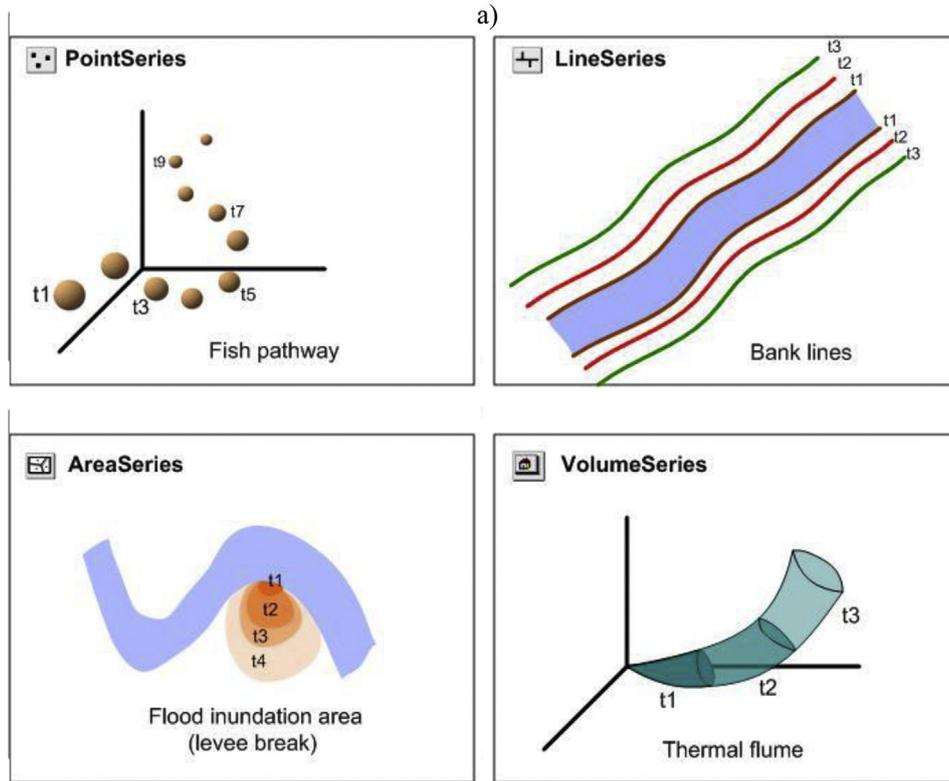
point source. In the initial time steps, the point source grows in the vertical and lateral direction until “fully mixed” conditions along those directions are attained. While moving downstream, the cloud is continuously dispersed through mixing by the shear driven by the spanwise velocity gradient. This evolution is captured by colors in Fig. 8b. Colors are associated to magnitudes of pollutant concentration in the cloud. In those cases, the Lagrangian framework is well suited to describe such processes, where boundaries are spatially changing.

Implementation of the Lagrangian framework in GIS is challenging as the model describes objects mainly through their static locations and shapes. If the shape or location of a single object changes, GIS actually records the new shape and location as a new GIS object for recording. For example, although a single pollutant object is moving downstream over time, it appears to have multiple objects (features) in GIS software. The successful solution of the problem is based on the Feature Series concept developed by Arctur and Zeiler (2004) and Goodall (2005).

For storing dynamic objects, Arc River uses ‘RiverFeatureSeries’ in the data model. A unique identifier (SeriesID) is assigned for a given moving object, and is sustained for a series of features of the same object, thus capturing their changing shape over time. The

temporal sequence of dynamically varying features is captured by providing a time index (or time stamp) to each feature. Events (i.e., variable) or results of event (i.e., observed or simulated data values) are independently stored into various ancillary tables, and are uniquely related to the object for representing the current state of the object. The recent advances in GIS support the representation of the dynamic variation of the moving objects through animation or tracking tools (e.g., Tracking Analyst or Animation tool in ArcGIS). These tools facilitate the selective visualization of the shape and location of a moving object over time.

In addition to describing moving objects for supporting Lagrangian observations, the RiverFeatureSeries component can also handle changing shapes of the stationary objects such as bed surface shape in a given reach when Eulerian view is used. Most of stationary objects represented in the RiverHydroFeature dataset, can be recorded as a single feature in a feature class because of their fixed shape (or location). However, if the shape of a stationary object is changing, they become a series of features recorded in the feature classes. Consequently, the special case of describing stationary objects in the Eulerian framework can be also represented by the RiverFeatureSeries.



a)

b)

Simple feature class						
Point / Line / Area / VolumeSeries						Geometry
						Contains M values
						Contains Z values
						Yes
						Yes
Field name	Data type	Allow nulls	Default value	Domain	Precision	Scale Length
OBJECTID	OID					
Shape	Geometry	Yes				
HydroID	Integer	Yes			0	
SeriesID	Integer	Yes			0	
SeriesName	String	Yes				30
GroupID	Integer	Yes			0	

Fig. 9. RiverFeatureSeries: a) schematic drawings and examples of the FeatureSeries component; b) structure of Point, Line, Area, and VolumeSeries.

The RiverFeatureSeries component consists of four feature series classes depending on the shape of the feature: PointSeries, LineSeries, AreaSeries, and VolumeSeries. The classes are required because it is not possible to store different type of shape within a single feature class. Fig. 9 illustrates each feature class and illustration of their use in river studies. These examples substantiate the capability of the data model to represent the diversity of shapes and dimensions associated with river processes as well as the possibility to track their evolution in space and time.

The attributes in the Feature Series classes are designed to represent four different feature classes. Shape field can be point, polyline, polygon, or multipatch, respectively. HydroID is assigned for each feature (row) as a unique identifier for a river object at a given location and time, which is used to connect the object with its hydrodynamic values with time stamps and events separately stored in other tables. SereisID and SeriesName identify which features in the FeatureSeries classes belong to an object. GroupID can be assigned when it is needed to visualize or analyze multiple heterogeneous moving objects for better understanding their interaction over time. Underlying information for each location of moving objects are separately stored in the scalar or vector data table, and connected by exchanging identifiers.

3.5. Tabular and metadata information

Arc River is designed to store events and their related values separately from the river feature object by using a unique identifier in the river object that relates the river object with comprehensive information about the specific events or measurements. Elements of such an approach were used in previously hydrologic data models such as the Arc Hydro (Maidment, 2002) and the Observations Data Model (ODM, Tarboton et al., 2007). For example, Arc Hydro data model integrates the monitoring locations and their time series data within the GIS database. On the other hand, the Observations Data Model extensively documents observations with metadata entailing comprehensive information including both hydrologic observations and ancillary explanations (i.e., data characteristics and observation method). The Arc River data model fuses elements from both Arc Hydro and ODM into its architecture, which can be appropriate for describing multidimensional river processes. In particular, the Arc River includes the following group of information (tables) in conjunction with river objects: time series or stamp of data values for scalar and vector data, variable (event), instrument, methods, and survey group. A river object can be related with many scalar or vector values, and each value can have variable, method, instrument and group.

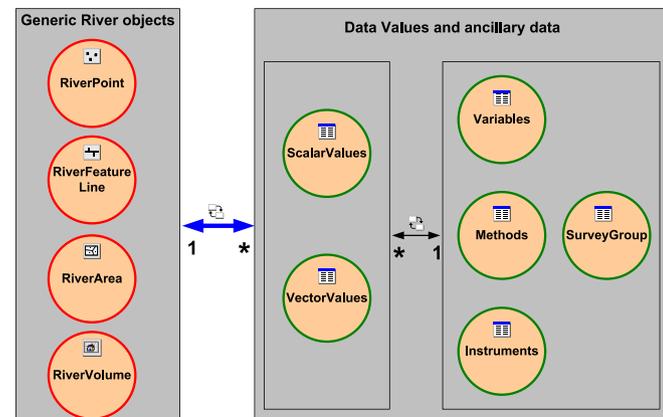


Fig. 10. Connectivity between generic river objects and ancillary information.

Arc River describes events through scalar or vector variables represented by the ScalarValues and VectorValues tables, respectively (see Fig. 11a). In addition, Arc River supports other ancillary information such as variable, method, instrument, and group. The Method table provides information about how a value was obtained, for example, through spatial interpolation or depth averaging. The Variable (event) table classifies the name of a value (event), and the Instrument table records the name of instrument used in the data observation. Fig. 10 illustrates the connectivity between river objects and their ancillary information. For example, each river object (RiverPoint) such as river point contains observed data (e.g., bathymetry in ScalarValue), including the instrument used (e.g., eco-sounder), variable type, and method (e.g., acoustics).

Arc River archives river feature objects separately from the time series hydrodynamic data (time record plus value), but these can be combined through a relationship between the feature class and data value tables. Fig. 11b exemplifies how a feature class (CrossSectionLine) is able to store its hydrodynamic data using this relationship. Through one to many relationships, a feature in the CrossSectionLine feature class (HydroID = 1) has multiple vector and scalar values in both VectorValues and ScalarValues tables (FeatureID = 1).

While the ScalarValues table only stores scalar values such as discharge or cross-section area in the DataValues field with the time record, the VectorValues table is designed to accommodate directional information containing two or three components of a hydrodynamic vector such as velocity or shear stress. The DataValueX, DataValueY, and DataValueZ fields indicate these vector components, respectively. They can be used to calculate the direction and magnitude to enable visualization of vector arrows within GIS applications. In particular, when the uncertainty information for the value is available, it is stored in the Uncertainty field. Currently, the Uncertainty field is only incorporated in ScalarValues table.

Both data value tables store time record (time series or stamp) thus capturing the spatial and temporal changes in the river object characteristics. TSDatetime field stores such time record information, which represents the time when the raw data was collected and when it was interpolated to a raster. In general, the river objects described in the Eulerian framework normally store a time series of hydrodynamic values though one (object) to many (values) relationship. For storage efficiency, the value table is designed to be separated from feature class. For the moving river objects, however, a river feature related to data value table describes only the location of the object at a certain time stamp, not its time series. In such case, the time stamp can be directly stored within the corresponding feature class (Arctur and Zeiler, 2004; Goodall, 2005). Nevertheless, Arc River separates time records from the RiverFeatureSeries by sustaining the same structure as the one used in the description of the stationary objects in order to: (i) keep data structures in the data model simple and consistent; (ii) handle multiple events which have one-to-many relationships; and (iii) support comprehensive set of ancillary information.

Data value tables contain event information (variable) which provides the metadata for the given values, and VariableID field relate the data value with relevant event (see Fig. 12a). MethodID allows linking the data value with the method information describing how the data value was produced (e.g., depth-averaged, time-average, or simulated by models). The Variables table stores information about what type of events occurred in the given river object. For example, the measured or simulated values of velocity and dispersion coefficient are recorded in this table. In addition, the time and duration of this event stored in the data value table is connected to the variable table. As the attributes of Variables table, VariableID and VariableCode are identifiers for each variable or

event, and VariableName records the name of variable. VariableType field provides a method for differentiating which variable supports vector or scalar type of data values. If a unit is available for the variable, it is recorded in the Unit field. A more detailed description regarding the variable or event can be added in the Description field if needed.

Instruments table provides information regarding the provenance of the data such as whether the data is measured or result from a numerical model (Fig. 12b). Even when dealing with just one category such as measurements, different instruments can be used to get the same variable. For instance, discharge can be evaluated from the gage height recorder, ADCP measurements, or weir.

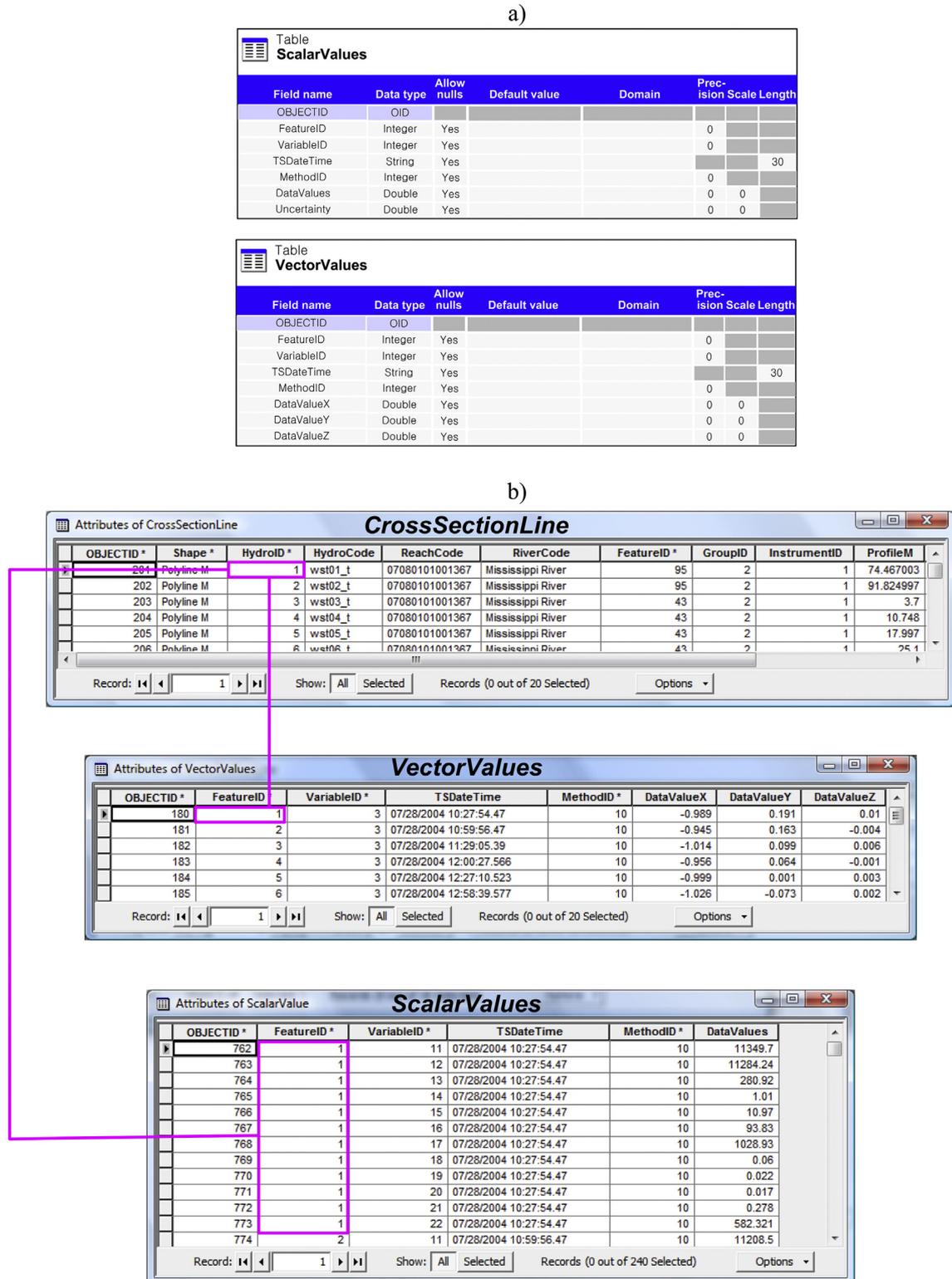


Fig. 11. Metadata for time series tables: a) structure of data value tables (ScalarValues and VectorValues); b) an example of relationship between the feature class and its data values.

a)

Table Variables						
Field name	Data type	Allow nulls	Default value	Domain	Precision	Scale Length
OBJECTID	OID					
VariableID	Integer	Yes			0	
VariableCode	String	Yes				30
VariableName	String	Yes				30
VariableType	String	Yes				30
Unit	String	Yes				30
Description	String	Yes				30

b)

Table Instruments						
Field name	Data type	Allow nulls	Default value	Domain	Precision	Scale Length
OBJECTID	OID					
InstrumentID	Integer	Yes			0	
InstrumentName	String	Yes				30
Description	String	Yes				30

Table Methods						
Field name	Data type	Allow nulls	Default value	Domain	Precision	Scale Length
OBJECTID	OID					
MethodID	Integer	Yes			0	
MethodName	String	Yes				30
Description	String	Yes				30

c)

Table Group						
Field name	Data type	Allow nulls	Default value	Domain	Precision	Scale Length
OBJECTID	OID					
GroupID	Integer	Yes			0	
GroupName	String	Yes				30
RiverName	String	Yes				30
StartDateTime	Date	Yes				30
EndDateTime	Date	Yes				30
Description	String	Yes				30

Fig. 12. Data information for ancillary information: a) structure of Variables table; b) structure of Instruments and Methods tables; c) structure of Group table.

Methods table refers to method of how the data value was evaluated as well as comments on the strengths and limitations for each instrument (Fig. 12b). Moreover, in some cases, the data is not directly measured or simulated, rather obtained through spatial interpolation or time averaging. They are related to data value tables through InstrumentID and MethodID.

The Group table fulfills the function of grouping river objects according to user defined needs by combining relevant information (Fig. 12c). GroupID is recorded in the feature class to provide any grouped information. GroupName defines the group identity. StartDateTime and EndDateTime field record the time interval for the stored data.

4. Connectivity between Arc River and Arc Hydro data models

As mentioned above, the available ODM and Arc Hydro handle only one- or two-dimensional datasets while Arc River enables representation of 3D vector quantities and river topographic features acquired with any instruments. Given the versatility of

the new generation of acoustic and image-based instruments to provide a multidimensional representation of the river hydrodynamics and morphodynamics (Muste et al., 2012), some of the Arc River data ingestion tools have been tailored to accommodate specific instruments. Illustrated herein are tools that allow ingestion of ADCP raw data in Arc River data model. The connection between the Arc River and Arc Hydro representations is based on concepts and terminology developed for Arc Hydro for describing the river network (Maidment, 2002). Fig. 13 illustrates this connection by overlapping Arc Hydro's HydroEdge feature with the curvilinear river-attached coordinates (s, n, z) that are used by Arc River to represent multidimensional river objects. Data collected with ADCPs in a transect or fixed point are used to generate Arc Hydro's HydroNetwork which topologically connects stream lines and points in conjunction with the one-dimensional description of the river network. The connectivity between multidimensional objects (point features in this case) is defined both with respect to the cross-section as well as river network.

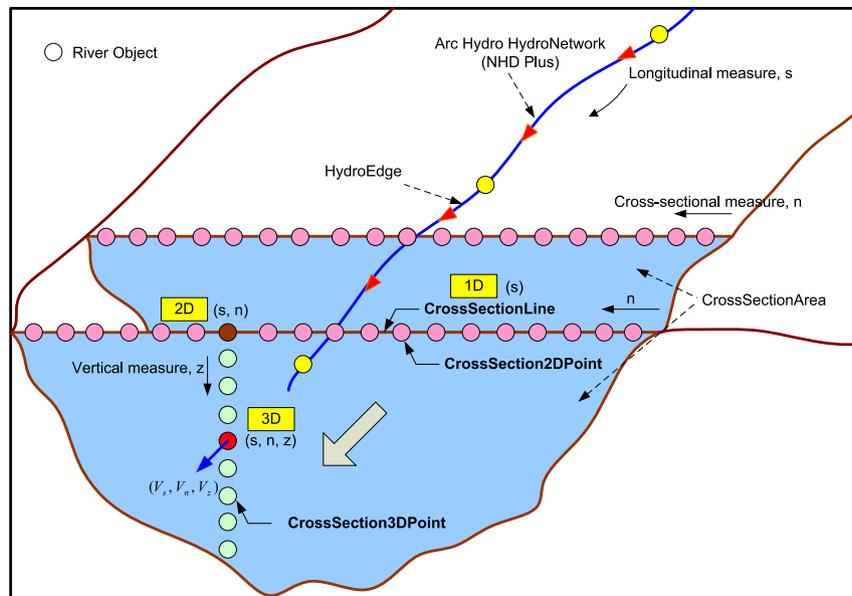


Fig. 13. The connectivity between Arc Hydro and Arc River feature classes.

In addition to the connectivity between a certain river object and the river network, river objects themselves are related to each other based on their dimensions. For instance, as can be seen in the Fig. 13, a CrossSectionLine, named after a polyline connecting point measurements across a cross-section, is classified as a 1D object because it contains various cross-section averaged 1D quantities such as mean velocity and discharge. The CrossSectionLine object contains multiple 2D points (having a one to many relationship), which are called as CrossSection2DPoint. So by exchanging key identifiers between them, CrossSection2DPoints are connected to a CrossSectionLine. Similarly, many 3D objects, such as CrossSection3DPoint along the vertical direction of the cross-section, can be related to a 2D object, CrossSection2DPoint. Fig. 14 shows populated CrossSection2DPoint feature class in the Arc River data model using ADCP measurements, and its connectivity with CrossSectionLine and RiverFlowLine as well as with the Vector and ScalarValue tables. The software used to populate the ADCP dataset into Arc River data model can be obtained from the first author upon request.

5. Concluding remarks

Arc River data model described in this paper enables efficient organization of river field measurements in a multidimensional representation of the river characteristics over a range of spatial and temporal scales. In other words, Arc River connects the instrument and physical river domain areas using well-defined and rigorous model components and relationships that uniquely convert the instrument raw data into practical data that can be understood by a river specialist. This conversion cannot be typically done by the instrumentation, hydraulic, or computer science domain specialist working alone. The raw data files outputted by the instrument report multiple characteristics regarding instrument operation, sensor and signal characteristics along with the actual data associated with the measured variables. Only the latter are actually used in the hydraulics domain where the river hydro- and morphodynamic aspects are of interest. Furthermore, Arc River model is built using computer science knowledge that along with the expertise of the previously-mentioned domains allows to extract and preserve the relevant data from the instrument files as

riverine spatio-temporal features defined with concepts and terminology used in the hydraulics domain. This representation allows the investigator to observe the river data grouped in features such as kinematic or geometrical river lines, bathymetry, shear stress or velocity and velocity-derived quantity mapping in a cross section or over a river reach and their evolution in time.

This new model parallels and complements data models developed in related disciplines, such as Arc Hydro, Arc Marine and Arc Hydro Groundwater. The construction of this data model has taken advantage of the cyberinfrastructure developed through previous in house efforts (Kim and Muste, 2012) as well as concepts and tools developed by river information communities (www.cuahsi.org) being an illustrative example of adaptation and extension of generic cyberinfrastructure concepts and workflows across disciplines involved in watershed resources investigations. The model can be easily extended to accommodate additional data associated with river-related processes including environmental and socio-economic aspects. The data model is currently applicable within the ArcGIS platform but the model can be easily adapted to other GIS platforms or non-spatial databases. More details of the Arc River data model and its applications can be found in Kim (2008).

The proposed Arc River data model is a demonstration of emerging trend in many natural science areas and it is especially relevant to user groups in the hydrologic and hydraulic community where new generations of instruments raise considerable concerns of how to efficiently and quickly retrieve practical information from big datasets. However, it should be noted that the current data model has lots of room for improvements. For example, the Arc River data model was mostly implemented by utilizing ADCP measurements and a simple one-dimensional simulation model (i.e., ARPTM). Other observation techniques such as LSPIV and multi beam depth sounder, as well as simulation models should be applied to make sure that present design of the data model works properly. Similarly, the data model has limited capability to define cross-sections for the braided river channel or flow confluences, and thereby is unable to relate the multidimensional cross-sectional information with a one-dimensional river network in such cases. Finally, for this data model to be used by the wider community and fully functional, a set of tools need to be developed

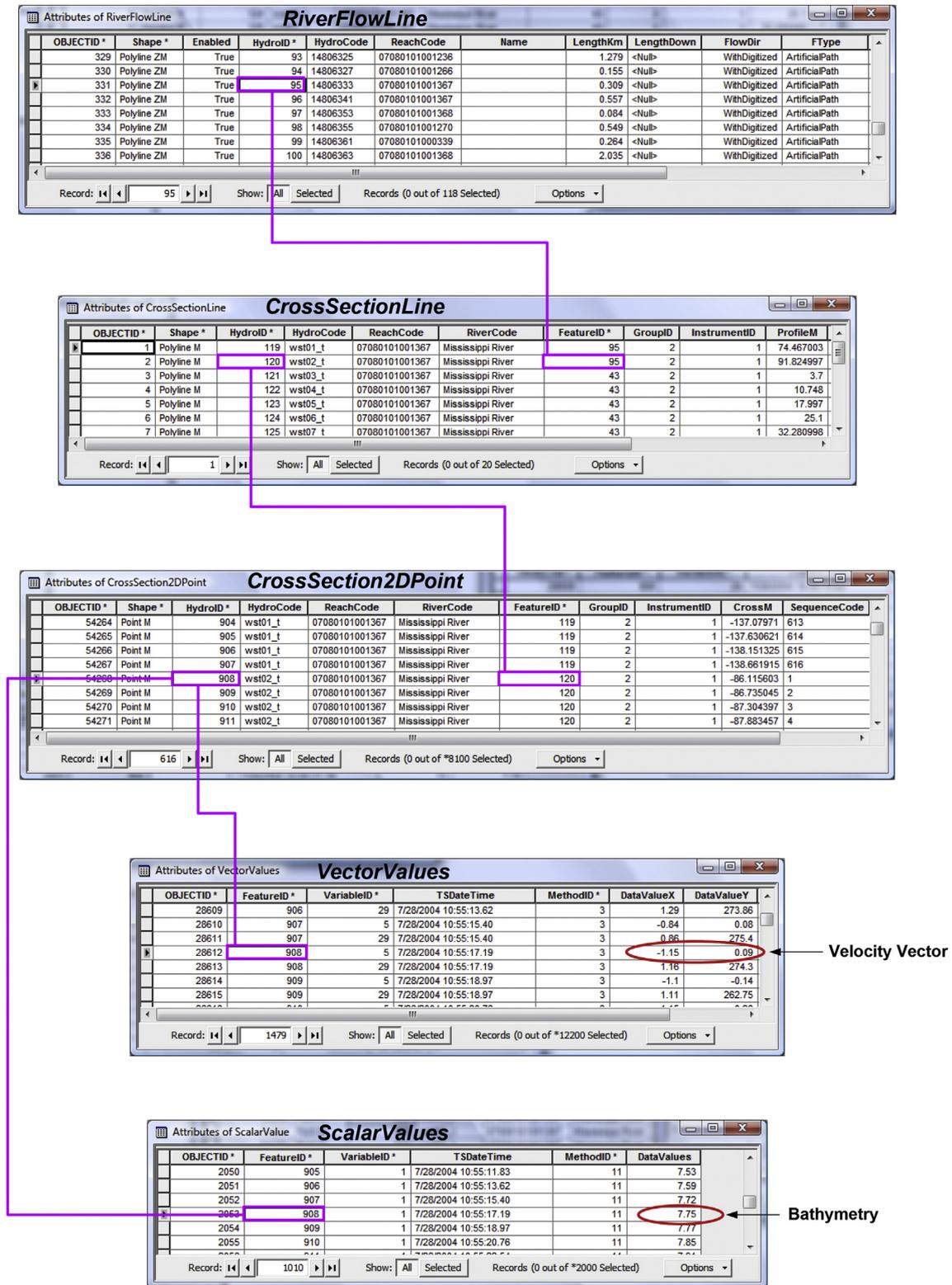


Fig. 14. CrossSection2DPoint feature class defined by the depth-averaged velocity vector and bathymetry obtained with ADCP measurements.

to load data from multiple sources into the model and to visualize 3D dynamic datasets.

These initial developments represent contributions to a larger community effort towards establishing a cyberinfrastructure-based

information system for watershed-scale eco-hydrologic observatories whereby many disciplines are integrated in digital environments using physical understanding of the landscape processes. Development of such a complex system is a long-term process,

ideally developed with ample dialogue with the community. Moreover, it requires an extensive community dialogue throughout the development process such as the concepts, methodologies and outcomes to be acceptable to a wide range of users and eventually to be converted in community practice standards.

Acknowledgments

This research was supported by a grant (10-CHUD-B02) from the Regional Innovative Technology Program funded by the Korean MLTM (Ministry of Land, Transport and Maritime) Affairs. The second author also gratefully acknowledge the partial support provided for this work by the Iowa Flood Center, IHR-Hydroscience & Engineering and the National Science Foundation award BCS-1114978.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.envsoft.2014.12.002>.

References

- Arctur, D., Zeiler, M., 2004. Designing Geodatabases – Case Studies in GIS Data Modeling. ESRI, Redlands, California.
- Chanson, H., 2004. Environmental Hydraulics of Open Channel Flows. Elsevier Butterworth-Heinemann, Burlington, MA, USA.
- Chaudhry, H., 1993. Open-channel Flow. Prentice-Hall, New Jersey, USA.
- Creutin, J.D., Muste, M., Bradley, A.A., Kim, S.C., Kruger, A., 2003. River gauging using PIV technique: proof of concept Experiment on the Iowa river. *J Hydrol.* 227 (3–4), 182–194.
- Currie, I.G., 1993. Fundamental Mechanics of Fluids. McGRAW-HILL, Hightstone, New Jersey, USA.
- Dinehart, R.L., Burau, J.R., 2005. Repeated surveys by acoustic Doppler current profiler for flow and sediment dynamics in a tidal river. *J Hydrol.* 1–21.
- ESRI, 2007. ESRI ArcGIS Diagrammer User Guide. Redlands, CA.
- Fischer, H.B., List, E.J., Koh, R.C., Imberger, J., Brooks, N.H., 1979. Mixing in Inland and Coastal Waters. Academic Press, New York.
- Goodall, J.L., 2005. A Geotemporal Framework for Hydrologic Analysis. Doctoral Dissertation. The University of Texas at Austin, Austin, TX.
- Hodges, B.R., 2013. Challenges in continental river dynamics. *Environ. Model. Softw.* 50, 16–20.
- Julien, P.Y., 1998. Erosion and Sedimentation. Cambridge University Press, Cambridge CB2 2RU, United Kingdom.
- Kim, D., 2008. Arc River – Geotemporal River Hydrodynamic Data Model. Doctoral Dissertation. The University of Iowa.
- Kim, D., Muste, M., 2012. Multi-dimensional representation of river hydrodynamics using ADCP data processing software. *Environ. Model. Softw.* 38 (12), 158–166.
- Kostaschuk, R., Villard, P., Best, J., 2004. Measuring velocity and shear stress over Dunes with acoustic doppler profiler. *J Hydraulic Eng.* 130 (9), 932–936.
- Liu, F., Hodges, B.R., 2014. Applying micro-processor analysis methods to river network modeling. *Environ. Model. Softw.* 52, 234–252.
- Maidment, D.R., 2002. Arc Hydro – GIS for Water Resources. ESRI, Redlands, California, USA.
- Merwade, V.M., Maidment, D.R., Hodges, B.R., 2005. Geospatial representation of river channels. *J Hydrologic Eng.* 10 (3), 243–251.
- Muste, M., Yu, K., Pratt, T., Abraham, D., 2004. Practical aspects of ADCP DataUse for quantification of mean river flow characteristics; Part 2: fixed-vessel measurements. *Flow Meas. Instrum.* 15 (1), 17–28.
- Muste, M., Kim, D., Merwade, V., 2012. Modern digital instruments and techniques for hydrodynamic and morphologic characterization of river channels. In: Church, M., Biron, P., Roy, A. (Eds.), *Gravel Bed Rivers: Processes, Tools, Environments*. Wiley-Blackwell.
- Nystrom, E.A., Rehman, C.R., Oberg, K.A., 2007. Evaluation of mean velocity and turbulence measurements with ADCPs. *J. Hydraul. Eng.* 133 (12), 1310–1318.
- Parsons, D.R., Jackson, P.R., Czuba, J.A., Oberg, K.A., Mueller, D.S., Rhoads, B., Best, J.L., Johnson, K.K., Engel, F., Riley, J., 2013. Velocity Mapping Toolbox (VMT): a processing and visualization suite for moving-vessel ADCP measurements. *Earth Surf. Process. Landforms* 38, 1244–1260. <http://dx.doi.org/10.1002/esp.3367>.
- RDI, 1996. Acoustic Doppler Current Profilers - Principle of Operation, a Practical-Primer. RD Instruments, San Diego, CA.
- Rennie, C.D., 2002. Measurement of bed load velocity using an acoustic DopplerCurrent profiler. *J. Hydraulic Res.* 128 (5), 473–483.
- Sime, L.C., Ferguson, R.I., Church, M., 2007. Estimating shear stress from movingboat acoustic Doppler velocity measurements in a large gravel bed river. *Water Res. Res.* 43, W03418.
- SonTek, 2000. Doppler Velocity Log for ROV/AUV Applications. SonTek, San Diego, CA.
- Strassberg, G., Jones, N., Maidment, D., 2011. Arc Hydro Groundwater: GIS for Hydrogeology. ESRI, Redlands, California, USA.
- Szupiany, R.N., Amsler, M.L., Best, J.L., Parsons, D.R., 2007. A Comparison of fixed-and moving-vessel flow measurements with an acoustic Doppler profiler (aDp) in a large river. *J. Hydraul. Eng.* 133 (12), 1299–1309.
- Tarboton, D.G., Horsburgh, J.S., Maidment, D.R., 2007. CUAHSI Community Observation Data Model (ODM), Version 1.0. Design Specifications.
- Wright, D.J., Blongewicz, M.J., Halpin, P.N., Breman, J., 2007. Arc Marine: GIS for a Blue Planet. ESRI Press, Redlands, California.
- Zeiler, M., 1999. Modeling Our World. ESRI, Redlands, California, USA.