Geospatial representation of stream channels

By Venkatesh Merwade and David R. Maidment

Submitted to the Texas Water Development Board
In fulfillment of Grant # 2002-483-439

January 2003

Center for Research in Water Resources
J. J. Pickle Research Campus
University of Texas at Austin
Austin TX 78712 – 4497
Introduction

Accurate representation of stream channels in three-dimensions plays a major role in hydraulic modeling (Crowder and Diplas, 2000; French and Clifford, 2000). The data requirement of all hydraulic models is well established and documented, but the procedures by which the composite terrain data are used to create the digital terrain model (DTM), and by which this DTM is used to define the channel geometry, receive insufficient attention given their significance in the overall modeling process (French and Clifford, 2000).

In geographic information system (GIS) spatial data are usually organized into vector and raster forms. In the vector form, the data are represented by points, lines, and polygons. The raster structure stores data in a two-dimensional grid of cells, where each cell contains a value for the attribute being mapped. Another commonly used GIS dataset is a triangulated irregular network (TIN). A TIN is a model for surface representation wherein the data are stored as an integrated set of nodes with elevations and triangles with edges. With the advancement of GIS technology and the availability of spatial data, both raster and vector, commonly used hydraulic models such as HEC-RAS and MIKE 11 are linked with GIS to extract geometry data such as river banks, centerline, and cross-sections. For example HEC-GeoRAS 3.0 is an ArcView extension designed to process geospatial data for use with HEC-RAS. HEC-GeoRAS creates an import file containing geometric data from an existing TIN of the stream channel. In order to extract the geometric attributes from the TIN, the user has to create ArcView shapefiles for the channel banks, centerline and cross-sections manually (USACE, 2000). Creating the
geometric attributes of stream channels is prone to error depending on the nature of the study area and details required to accurately represent the channel.

Therefore there are two important issues with hydraulic modeling in GIS: accurate digital terrain model, and the extraction of geometric data from the digital terrain model. This paper addresses the second issue and presents an automated way of extracting the thalweg, cross-sections, and profile lines, for stream channels using three-dimensional measurement points. The thalweg is the line connecting the points of lowest bed elevation in a natural river. Thus the thalweg follows the deepest part of the river. Profile lines define the longitudinal profile of a river channel parallel to the direction of flow, and cross-section refers to the section of a channel taken transverse to the direction of the flow (Noman and Nelson, 2002).

Motivation for the study

The Texas Water Development Board (TWDB) conducts instream flow studies using two-dimensional hydrodynamic model. Instream flow studies are undertaken to establish the relationship between the flow and fish populations, thus enabling water resource planners to make decisions on water distribution policies taking into account the instream flow requirements within the river. The hydrodynamic modeling results are dependent on accurate river channel bathymetry and this study addresses the issue of extracting accurate geometry information from the field collected bathymetry data.

Previous work

As mentioned previously, GIS data are usually organized into vector and raster forms. The geometric description of stream channels using thalweg, banklines, cross-sections, and profile lines are therefore vector data. There is not much research done on
representing the channel geometry in vector form. However, there have been some efforts notably by Tate et. al. (2002) and Wadzuk and Hodges (2001) that are helpful in developing the concepts presented in this paper.

Tate et al. (2002) presented a GIS approach for the development of a digital terrain model based on stream channel representation in HEC-RAS. The authors used the detailed stream cross-sections from HEC-RAS and combined them with a digital terrain model grid to produce a surface model that provided a good representation of the general landscape with additional detail in the stream channel. HEC-RAS uses a relative coordinate system in which the coordinate of any given point is defined by its location along the centerline (stream measure, $M_s$), its location along a given cross section line (cross section measure, $M_{cs}$) and elevation ($z$). Thus the cross-sections are indexed along the centerline and then all the points on the cross-sections are measured across the centerline. In order to merge the data from HEC-RAS and the grid, they developed a methodology to transfer the data from HEC-RAS coordinate system to the regular Cartesian coordinate system.

Wadzuk and Hodges (2001) presented a method for straightening the bathymetry of sinuous, dendritic reservoirs. Their main goal of straightening the bathymetry was to reduce the number of computational cells in hydrodynamic computations. They used two extra coordinates, one measured along the centerline and the other measured transverse to the centerline, to represent bathymetry in a flow-oriented coordinate system. The centerline used by Wadzuk and Hodges to define this new coordinate system was neither the true geometric center of the channel nor the line following the greatest depth. The
centerline was just an arbitrary smooth line, where the cross-sections did not cross each other.

A linear referencing system is a method of identifying a location on a stream network by reference to the known positions of spatial objects. For example reach 110033, distance 100 miles from the upstream end, uniquely identifies a position in geographic space without having to express it in (x,y) coordinates (ESRI, 2001). Referencing the stream channels depending on the location along the stream is important and is also recognized by national agencies such as the USGS and the EPA. For example the National Hydrography Dataset (NHD) developed by the USGS and the EPA has measures assigned to its reaches and these measures can be used to define locations along NHD for geo-referencing points, such as point discharges, or they can be used to define the start and end points of areas along a linear reach, like a study area for fish habitat studies (EPA, 2002).

ArcHydro (Maidment et. al., 2002) is a geospatial and temporal data model for water resources that operates within ArcGIS environment. ArcHydro is a data model that can be used for hydrologic simulation by exchanging data between ArcHydro and any independent hydrologic model. ArcHydro divides water resources data into five components: network, drainage, channel, hydrography, and time series. The channel component of Arc Hydro includes three-dimensional representation of the shape of river channels and provides a standardized way of representing the profile lines and cross sections.

It is therefore recognized that having some sort of methodology to address the stream channels along the flow rather than using the Cartesian coordinates to find the
exact location in the stream channels is required. This paper uses this concept of introducing the new coordinate system and extends this approach to develop a methodology of creating geometry of stream channels automatically using GIS. The geometric description of stream channels extracted by this method can be used to populate the channel component of Arc Hydro.

Study area and Data

![Figure 1. Study area and data. (a) Texas state; (b) Brazos network for the study area; (c) Bathymetry data collected along the Brazos river; (d) Small section of the Brazos river showing measurement points. Each point has (x,y,z) associated with it.]

Bathymetry data collected by the Texas Water Development Board (TWDB) along the Brazos River are used for the present study. The study reach is about 4.6 miles long and about 100m wide. As shown in Figure 1, the study reach is located due NE of Wallis County, TX on FM 1093, and adjacent to the SW of Simonton County, TX. It
starts from the confluence of Allens Creek with the Brazos River and extends about 4.6 miles downstream. The bathymetry data are collected using a boat-mounted acoustic depth sounder linked to a Global Positioning System (GPS). The data are recorded on a computer by moving the boat along the study reach. As the boat moves along, the depth sounder provides the depth of water and the GPS gives the location of each point in terms of latitude and longitude. All the data are recorded in a computer that is connected to both the depth sounder and the GPS. Thus, the whole setup provides $x,y$ locations and depth of water. Obtaining the channel bathymetry requires knowledge of water surface elevation (WSE). Pressure transducers are used at key points to measure the water surface elevation and interpolated in between to get the water surface profile. Spatial variation of the WSE is small relative to the channel bathymetry and therefore the measurement of WSE at key points with linear interpolation between is justified (Austin and Wentzel, 2001). The depth of water is subtracted from the water surface elevation to get the riverbed elevation, $z$ value.

**ArcGIS and ArcObjects**

ArcGIS is GIS software from Environmental Science Research Institute (ESRI) that stores spatial and temporal data in a relational database. ArcGIS also offers customization capabilities using Visual Basic and ArcObjects. ArcObjects is the development platform for ArcGIS and provides an infrastructure to build customized applications based on existing components (Waltuch et al., 2001). In other words, the user can create his/her own functions if they are not available with the standard software. ArcHydro (Maidment et al., 2002) is an excellent example of how ArcObjects can be
used to build customized water resources data model. ArcGIS 8.2 is used for the present work and the computer programs are developed using Visual Basic and ArcObjects.

**Polyline in ArcGIS**

In ArcGIS, a polyline is a set of line segments that may or may not be connected (ESRI, 1999). The polyline discussed in this paper refers to connected line segments as shown in Figure 2. Traditionally a polyline stores its \((x,y)\) coordinates at all of its vertices. In ArcGIS, besides these two basic coordinates, a polyline can also store two extra coordinates, \(m\), and \(z\). The \(m\) coordinate refers to measure, and \(z\) coordinate refers to elevation. When a polyline stores \((x,y,m)\) it is called *Polyline M* and when it stores all the four coordinates \((x,y,z,m)\), it is called *Polyline ZM*. A three-dimensional polyline in this paper refers to a Polyline ZM.

**Figure 2.** Polyline ZM in ArcGIS. (a) A polyline ZM with its vertices and segments; (b) Properties of the each vertex of the polyline ZM.

There are two types of measure co-ordinates that can be assigned to any polyline: relative or absolute. As the name suggests, in the case of relative measure, the user assigns values at the two end points and the intermediate values are interpolated relative to these values. Relative measures are therefore percentage values. In case of absolute measure, the polyline has measure values that are assigned with reference to a fixed
point. For example, if the upstream end of the channel is the reference point then the measures are equal to the length of the polyline at any particular point along it. Absolute measures are assigned in meters.

**Trends in River Bathymetry**

![Diagram showing river bathymetry](image)

**Figure 3.** Trends in river bathymetry. (a) Small hypothetical river channel; (b) Longitudinal profile of a typical river channel; (c) cross-section view of a typical river channel.

In order to understand the trends in river bathymetry consider Figure 3, which shows a hypothetical reach of a river channel. If we take a section along the flow and across the flow, we get the longitudinal profile of the riverbed and the cross-section, respectively. We can see from Figure 3b and 3c that the river has typical trends in the bathymetry along the flow and across the flow. Therefore, in order to extract any information from the bathymetry data it is very important to first extract these trends. Figure 4 shows elevation data, z value for each point, plotted for a small section of the Brazos River shown in Figure 1d.
Figure 4. Elevation data plotted against Easting and Northing for the study area shown in Figure 1d.

In Figure 4, the data do not show any trends as observed in Figure 3. This is because of the global trends that are imposed by the Cartesian coordinate system. In order to remove the global trend it is necessary to make the data independent of Easting and Northing coordinate axes.

A new coordinate system called Hydraulic coordinate system \((m,r)\), similar to Wadzuk and Hodges (2001), is introduced here. In the new coordinate system the data are stored according to their location with respect to the centerline of the channel. In the Hydraulic coordinate system, shown in Figure 5, for any point, \(m\) is the distance along the centerline and \(r\) is the perpendicular distance from the centerline. The centerline runs in the direction of the flow. Therefore the \(m\) coordinate for the centerline is zero at the beginning (upstream end) of the channel and is equal to the length of the centerline at the downstream end of the channel. The \(m\) coordinate for any point is always positive. The points lying on the left side of the centerline have positive \(r\) coordinates and the points on the right hand side of the centerline have negative \(r\) coordinates.
Figure 5. Hydraulic $(m,r)$ coordinate system for a stream channel

The Cartesian coordinate system uses the intersection of $x$ and $y$ axes $(0,0)$ as the origin that is unique for all data. In order to have consistent $(m,r)$ coordinates, it is necessary to establish a unique origin for the Hydraulic coordinate system as well. Since $(m,r)$ coordinates are measured with respect to the centerline, it is necessary to define a unique centerline. The thalweg is the traditional definition of the centerline and can also be used in the Hydraulic coordinate system. The thalweg is important not only as the centerline, but it also plays an important role in the digital terrain model of stream channels. It is important to understand that the hydraulic coordinate system can have its origin $(m=0, r=0)$ either at the upstream end of the channel or at the downstream end of the channel. In this study the upstream end is taken as the origin, but the same analysis could be applied with the origin at the downstream end if desired.
GIS methodology for Thalweg

In order to define a unique thalweg, it is necessary to have an accurate bathymetric surface. Creating an accurate bathymetric surface of a stream channel is a challenge by itself. However, it is possible to create a thalweg that is accurate enough to serve the purpose of the Hydraulic coordinate system. A GIS based procedure is introduced to define the thalweg. This procedure uses the measured bathymetry points, an arbitrary centerline defined by the user, and the boundary of the channel to find the thalweg. The arbitrary centerline can be a blue line from a digital map hydrography or interpreted from an aerial image or user can digitize a centerline over the bathymetry points. Similarly the boundary can be digitized using an aerial image. In this study the boundary is first defined using the digital orthophoto quadrangle (DOQ) for the study area. The DOQ available for the study area is over five years old and the river changes over time. Therefore the initial boundary does not cover all the bathymetry points and is modified slightly to cover all the points.
The thalweg flowchart is shown in Figure 6. The flow chart is straightforward and there are eight steps involved as shown in the figure. The most important thing in the whole process is the initial centerline defined by the user. The program basically takes this line and moves it so that it falls along the deepest part of the riverbed. In other words, the channel boundary provides a buffer zone and the arbitrary line is moved in this buffer zone so that it follows the deepest part of the channel.
If the user defined line is not dense i.e. if there are not enough segments then the arbitrary centerline is densified so that each segment of the polyline is approximately 100m long. The arbitrary centerline is also checked for any sharp angles between the segments. If two adjoining segments make a sharp angle with each other, then those segments are filtered out so as to get a smooth line. Normal lines created in step 5 are in fact the cross-sections covering the entire width of the channel at each vertex of the polyline. These normals are three-dimensional polylines (polylineMZ) and have vertices corresponding to the underlying cell of the grid. Each vertex of the normal has a z value extracted from the grid cell. In order to move the initial centerline, all the vertices of every individual normal are evaluated. The vertex that has the maximum z value is picked. This is the deepest point along that normal. Similarly the deepest points along all the normals are found and are then joined to form a three-dimensional polyline, the thalweg. Figure 7 shows these steps in GIS. For explanation purpose, the normal lines shown in Figure 7 are created at 15m interval instead of 100m. Steps 2 to 8 are not visible to user in the GIS.
Figure 7. GIS procedure for the thalweg in GIS. (a) Input data; (b) Raster surface created using bathymetry points; (c) Normal lines created for the initial line at 15m interval; (d) initial line is moved to the deepest points along all the normals; (e) Final result with initial line changed to thalweg; (f) Vertical profile of a normal showing the location of initial line and the deepest point (thalweg).

Assigning \((m,r)\) co-ordinates to bathymetry points

As explained earlier, \(m\) and \(r\) coordinates are assigned to the data points with reference to the centerline. In order to assign an \(m\) coordinate to any data point it is important to assign measure to the centerline. The centerline is therefore assigned absolute measures so that all the points can be referenced to the exact length of the centerline. There are tools available in ArcGIS to assign measures to a polyline. Once measure values are assigned to the centerline, the next step is to assign \((m,r)\) coordinates to all the points. The procedure to assign \((m,r)\) coordinates is described below.
Figure 8. Geometric concepts for defining \((m,r)\) coordinates for a bathymetry point \(P\).

The geometric concepts involved in assigning the \((m,r)\) coordinates are explained with reference to Figure 8. In order to calculate \((m,r)\) coordinate for any bathymetry point, \(P\), in the Figure 8a, it is necessary to find the segment that is closest to that
particular point. In Figure 8b, $ABC$ is a polyline with two segments $AB$ and $BC$. The
criterion that is used states that if the lines connecting the point to the ends a segment
makes an acute angle with both ends of any segment, then the point is closer to that
segment than to any other segment. In order to find out whether point $P$ is closer to $AB$ or
$BC$, $P$ is first connected to points $A$ and $B$ of the segment $AB$. Angle $\alpha$ (angle between $PA$
and $AB$) and angle $\beta$ (angle between $PB$ and $PA$) are calculated. $\alpha$ is an acute angle and $\beta$
is an obtuse angle. Therefore point $P$ is not adjacent to segment $AB$. Point $P$ is then
connected to points $B$ and $C$ of segment $BC$. Angle $\lambda$ (angle between $PB$ and $BC$) and
angle $\theta$ (angle between $PB$ and $BC$) are calculated. Since both $\lambda$ and $\theta$ are acute angles, $P$
is adjacent to $BC$.

The next step is to calculate the perpendicular distance from $P$ to segment $BC$,
$PG$, and the distance $BG$ in Figure 8c. This is calculated as below

$$PG = PB \times \sin(\lambda)$$

$$BG = PB \times \cos(\lambda)$$

$PG$ gives the $r$ coordinate for the point $P$, and the $m$ coordinate is the sum of $BG$ and the
$m$ coordinate of vertex $B$.

In some cases, Figure 8d, it is possible that $P$ forms acute angles with more than
one segment. In such cases the segment that gives the shortest perpendicular distance to
the point $P$ is considered. For example in Figure 8d, $P$ forms acute angles with both $BC$
and $CD$. However, the perpendicular distance from $P$ to $BC$, $PG$, is smaller than the
perpendicular distance from $P$ to $CD$, $PF$. Therefore, $P$ is referenced to $BC$ and not to
$CD$. 17
A problem arises when bathymetry point lies near the intersection of two segments. In this case it is difficult to associate that point with either segment because it does not form acute angles with any of them and cannot have unique \((m,r)\) coordinates. This problem is illustrated in Figure 8e. In Figure 8e, \(P\) does not form acute angles with both \(AB\) and \(BC\) and therefore, cannot be associated with either of the segments. This issue arises because the point \(B\) has two normals: one perpendicular to \(AB\) and the other one perpendicular to \(BC\). Any point that lies between these two normals cannot be associated to any of these segments. This issue can be resolved by converting the polyline into a smooth curve such as a piecewise Bezier curve. In simple terms, a Bezier curve is a curve which is determined by a set of control points (vertices of polyline) and each point of the curve is calculated from a polynomial function which uses the coordinates of the control points as parameters. There is a function available in ArcObjects that uses a Bezier curve to approximate a polyline by a smooth curve. When the polyline is converted into a piecewise Bezier curve, the transition is smooth at the intersection of segments thus providing a unique normal at any point on the curve as shown in Figure 8f.

Therefore the polyline must be first converted into a Bezier curve before assigning \((m,r)\) coordinates. If the polyline is not smoothen with Bezier curve then the \((m,r)\) coordinates are not assigned properly to the bathymetry points that lie at the intersection of segments. When the points are transferred to the Hydraulic coordinate system, they become independent of the global trend, and the shape of the river becomes straight as shown in Figure 9.
After transferring the data to the \((m,r)\) coordinate system, the bathymetry trends can be seen by plotting the data against the \(r\) axis. Figure 10 shows a channel-cross section when the bathymetry points shown in Figure 4 are plotted against the \(r\) axis. Therefore, by converting the data from Cartesian to a Hydraulic coordinate system, we can remove the global trend and capture the typical trends observed in any river bathymetry.

**Figure 9.** \((m,r)\) coordinate transformation. (a) Measurement points in \((x,y)\); (b) Measurement points in \((m,r)\)

**Figure 10.** Measurement points plotted against \(r\) axis in the \((m,r)\) coordinate system.
FishNet in ArcGIS

The surface FishNet tool available from ESRI creates a three-dimensional mesh of lines from a raster grid and TIN surface. It is a well known fact to GIS users that the raster and TIN surfaces are very slow to render. The FishNet can be considered as a wireframe model that renders itself faster than the original surface. Another advantage of having a FishNet, besides fast rendering, is easy navigation for visualizing the three-dimensional surface. We can create fishnet for the Brazos bathymetry points by first creating a raster surface and then using the FishNet tool to create the FishNet.

Extracting cross-sections and profile lines from the FishNet

Besides rendering and visualization, a FishNet can be utilized to automatically create cross-sections and profile lines from a surface, thus providing a more meaningful wireframe model in the GIS environment. In a FishNet, created in the Cartesian coordinate system, the horizontal lines are parallel to the x-axis and the vertical lines are parallel to the y-axis. So if the FishNet is created in the Hydraulic \((m,r)\) coordinate system, the horizontal lines are parallel to the centerline \((m\text{-axis})\) and vertical lines are perpendicular to the centerline. A FishNet in \((m,r)\) coordinate system is shown in Figure 11.
Therefore, all the lines that are perpendicular to the centerline are cross-sections and all the lines parallel to the centerline are profile lines. When the fishnet created in \((m, r)\) coordinate system is transferred back to the Cartesian coordinate system, we get a framework of cross-sections and profile lines as shown in Figure 12.
Summary and discussion

Although the work presented in this paper is developed within the ArcGIS environment, all the concepts are explained in detail so that they can be applied in other environments as well. This approach is based on transforming the data from the Cartesian coordinates \((x,y)\) to Hydraulic \((m,r)\) coordinates. Since the \((m,r)\) coordinates are defined with reference to the centerline, defining the thalweg as the centerline makes the coordinate transformation unique. If a unique centerline is not defined then different users can have different \((m,r)\) coordinates for the same data, which is not desirable. Input data for defining a thalweg are measured bathymetry points, channel boundary, and a user defined arbitrary centerline. This method has one disadvantage that depending on
how the user digitize the arbitrary centerline, the final result may vary, but it does not affect the coordinate transformation significantly because the main goal is to separate the measured bathymetry points into two parts. The first part, on the right hand side of the centerline, has negative $r$ coordinates and the second part, on the left hand side of the centerline, has positive $r$ coordinates. However, if due care is taken while defining the centerline by avoiding sharp angles and keeping the segments approximately equal in length, the final result provides a good approximation of the thalweg.

The FishNet tool is useful for visualization and if it is generated in $(m,r)$ coordinates and transformed back into the original coordinates, we get a network of cross-sections and profile lines. These cross-sections and profile lines are three-dimensional with $z$ value stored at each vertex. The cross-sections can be used as input to numerical models such as HEC-GeoRAS that read the channel geometry from cross-section data. This method has one disadvantage that: the FishNet tool generates profile lines and cross-sections at equal intervals, which is undesirable in some instances. For example, we need more cross-sections at locations where there are more variations in bathymetry compared to areas with small variations. However, the user can define the distance between the cross-sections small enough to cover all the minute details in the bathymetry and delete extra cross-sections in the areas where they are not required. Deleting some cross-sections is easier than defining extra cross-sections manually. Usually more cross-sections are desired along bends and this theory can be used to densify the fishnet along bends. So if the user generates a dense FishNet in $(m,r)$ coordinates and when the fishnet is transformed back into the original coordinates, some cross-sections can be filtered out depending on their location on the centerline. If the
centerline is straight, the spacing between the cross-sections can be increased by removing intermediate lines. Currently, this is not included in the computer program and the next step will be to include this thinking into the program to make it more meaningful.

Conclusions and further work

An automated approach for producing vector datasets for stream channels namely thalweg, cross-sections, and profile lines using measured bathymetry points is presented. The following conclusions are drawn from this study:

- The \((m,r)\) coordinate system offers an efficient way of treating the stream channel bathymetry by referencing the data in the direction of flow and across the flow.
- An ordinary FishNet with a network of horizontal and vertical lines can be converted into a meaningful dataset of cross-sections and profile lines by using the \((m,r)\) coordinate system.
- The most important conclusion that can be drawn from this study is that the introduction of \((m,r)\) coordinate system for stream channels has significant potential and it opens doors to other areas in stream channels research. For example it can be used in developing schemes for interpolating bathymetry data. Interpolation of bathymetry with detrended data gives better results than the trended data (Carter and Shankar, 1997). Converting the data from Cartesian coordinates to \((m,r)\) coordinates removes the global trend. If the data can be separated into two parts; one with positive \(r\) coordinates and the other with negative \(r\) coordinates, we can see trends even in these datasets. If these trends are captured and then the data is interpolated, the resulting surface is likely to have better representation. The next stage of this
research is to explore this possibility and also use these trends to fit analytic
functions to make prediction about areas where there are no or very sparse data.

References:

Austin Barney, and Wentzel Mark, 2001. Two-dimensional fish habitat modeling for
assessing instream flow requirements. Proceedings of the International
Symposium on Integrated Water Resources Management at the University of
California, Davis (LAWR), April 2000, IAHS Publ. no. 272, 2001, pp. 393–399.

environmental numerical modeling of gravel-bed rivers. Applied Mathematical
Modeling 20(11), 699-708.


Environmental Science Research Institute (ESRI), 1999. Modeling our world. Edited by
Michael Zeiler. ESRI Press, Redlands, California.

Environmental Science Research Institute (ESRI), 2001. Linear referencing and dynamic

estuarine environmental dynamics: some computational and methodological

Maidment David, Morehouse Scott, and Grise Steve. In Chapter 2, Arc Hydro
framework. Arc Hydro-GIS for water resources edited by David R. Maidment.


