Large scale hydrologic and hydrodynamic modeling using limited data and a GIS based approach

Rodrigo C.D. Paiva *, Walter Collischonn, Carlos E.M. Tucci

Instituto de Pesquisas Hidráulicas – IPH, Universidade Federal do Rio Grande do Sul – UFRGS, Av. Bento Gonçalves, 9500 – Porto Alegre, 90050-260 RS, Brazil

A R T I C L E  I N F O

Article history:
Received 5 October 2010
Received in revised form 6 June 2011
Accepted 12 June 2011
Available online 25 June 2011
This manuscript was handled by K. Georgakakos, Editor-in-Chief, with the assistance of Enrique R. Vivoni, Associate Editor

Keywords:
Large scale hydrodynamic model
Amazon
Flow routing
Flood inundation
STRM DEM
GIS algorithm

S U M M A R Y

In this paper, we present a large-scale hydrologic model with a full one-dimensional hydrodynamic module to calculate flow propagation on a complex river network. The model uses the full Saint–Venant equations and a simple floodplain storage model, and therefore is capable of simulating a wide range of fluvial processes such as flood wave delay and attenuation, backwater effects, flood inundation and its effects on flood waves. We present the model basic equations and GIS algorithms to extract model parameters from relatively limited data, which is globally available, such as the SRTM DEM. GIS based algorithms include the estimation of river width and depth using geomorphological relations, river cross section bottom level and floodplain geometry extracted from DEM, etc. We also show a case study on one of the major tributaries of the Amazon, the Purus River basin. A model validation using discharge and water level data shows that the model is capable of reproducing the main hydrological features of the Purus River basin. Also, realistic floodplain inundation maps were derived from the results of the model. Our main conclusion is that it is possible to employ full hydrodynamic models within large-scale hydrological models even using limited data for river geometry and floodplain characterization.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Large-scale hydrological models have been used for the impact assessment of land use change (e.g. Matheussen et al., 2000) and climate change (e.g. Nijssen et al., 2001; Xu et al., 2005) on water resources, for the study of hydrobiogeochemical processes (e.g. Foley et al., 1996; Zhang et al., 2002), as a basis for hydrological forecast systems (e.g. Wood et al., 2002; Collischonn et al., 2005; Thielen et al., 2009) and also as the land surface component of climate models (Pitman, 2003; Wood et al., 1992; Liang et al., 1994). These models are in constant development, partly as a consequence of a growing understanding of the hydrological processes, partly due to increasing computational capacity, and also due to the growing availability of remote sensing data sources (Schmugge et al., 2002), such as energy fluxes estimates (Bastiaanssen et al., 1998) and river water level (Frappart et al., 2006; Getirana et al., 2010).

Most of the large-scale hydrological models are conceptual models having at least a small physical basis. Some of them try to represent the role of different vegetation and soil types on the streamflow generation processes and on water and energy budgets of the basin. These features can be found in such models as VIC (Liang et al., 1994), LASCAM (Viney et al., 2000), TOPKAPI model (Liu and Todini, 2002), and IBIS–THMB (Coe et al., 2007). However, one key aspect in the behavior of large-scale river basins that has been partially neglected in this generation of large-scale hydrological models is river hydrodynamics.

Large-scale hydrological models usually use simple flow routing models that focus only on the flood wave delay and attenuation, such as storage or linear reservoir based models (Coe, 2000, 1997; Arora, 2001; Liston et al., 1994; Decharme et al., 2008), linearized Saint–Venant equations representing wave advection and diffusion (Lohmann et al., 1996), kinematic wave approximations (Liu and Todini, 2002; De Roo et al., 2000), or the Muskingum–Cunge model (Collischonn et al., 2007) and its variations (Beighley et al., 2009; Todini, 2007; Koussis, 2009).

While these simplified flow routing methods can in principle adequately represent flood wave delay and attenuation, they cannot deal with some of the hydrodynamic processes that are often present in large-scale river systems, notably backwater and floodplain storage effects. This is the case in the Amazon basin, where flow of several rivers is controlled by backwater effects, as reported by Meade et al. (1991), Trigg et al. (2009) and Tomasella et al. (2010), and the floodplains are a significant component of the fluvial system (Bonnet et al., 2008). Floodplains also play an important role in many other South American river basins (Hamilton...
Hydrodynamic flow routing in large-scale hydrological models could also provide results of other flow related variables, such as river stage, velocity and slope. These models could also be used as a basis for transport models by solving the advection–diffusion equation and sediment transport equations.

One dimensional (1D) hydrodynamic river models are well developed since the late 1970s (Cunge et al., 1980), and there are even well known packages for 1D hydrodynamic modeling (USACE, 2002; DHI, 2003). These models have been applied in relatively large-scale problems (Paz et al., 2010; Remo and Pinter, 2007; Biancamaria et al., 2009; Lian et al., 2007), however their use within large-scale distributed hydrological models, which also represent rainfall-runoff processes, is uncommon. There are also approaches for flood inundation simulation ranging from simplified models such as a composite cross section (Cunge et al., 1980) to 2D models (Galland et al., 1991; Bates and De Roo, 2000).

The reason why simple flow routing methods are still preferred within large-scale hydrological models is probably related to the lower input data needs and the lower computational burden they impose, while still providing reasonable results in several cases. Hydrodynamic river models need river cross section data that are usually surveyed only for relatively short reaches. Hydrodynamic models are also more computational intensive than simplified flow routing methods, and require more detailed information on boundary and initial conditions.

However, computing capacity keeps growing, and also new and higher accuracy remote sensing data such as the Digital Elevation Model (DEM) from the Shuttle Radar Topography Mission (SRTM) (Rabus et al., 2003; Farr et al., 2007) can be used to provide some of the information that is needed to compute parameters for large scale hydrodynamic models. In a recent example, Trigg et al. (2009) showed that the use of a simplified bathymetry, i.e. a wide rectangular channel approximation and a simple bed slope, does not introduce significant errors in the hydrodynamic simulation results in part of the central Amazon.

In the present paper, we describe the development of a model for large scale hydrologic and hydrodynamic modeling. The model uses limited data that can be usually found in global scale for most river basins, mainly SRTM DEM. GIS methods are used to extract the needed information from SRTM DEM. The proposed model is basically an improvement of the MGB-IPH large scale hydrological model (Collischonn et al., 2007). First, a brief description of the MGB-IPH hydrological model is presented (Section 2). Later, details concerning the hydrodynamic model equations, then a representation of the floodplains and the numerical scheme are shown (Section 3). We also present the GIS based algorithms developed to extract the hydrodynamic model parameters from the DEM (Section 4) and a simple approach for 2D flood inundation simulation (Section 5). The model is tested in a case study of the river Purus, one of the major tributaries of the Amazon (Section 6). Other results for an even larger region, and comparisons with simplified flow routing methods are presented by Paiva et al. (in review).

2. The hydrological model

The MGB-IPH model (“Modelo Hidrológico de Grandes Bacias”) is a large scale distributed hydrological model. It is similar to other large scale hydrological models such as LARSIM (Ludwig and Bremicker, 2006) and VIC (Liang et al., 1994; Nijssen et al., 1997). It is a process based model that uses physical and conceptual equations to simulate the terrestrial hydrological cycle: soil water budget, energy budget and evapotranspiration, interception, superficial, sub-superficial and groundwater flow generation and routing and river flow routing. It uses a daily or hourly simulation time step.

The early version of the MGB-IPH model was based on a square cell discretization of the river basin (Collischonn et al., 2007), while the version described in the present paper uses a division of the basin in small catchments using the ArcHydro methods (Maidment, 2002). Each catchment is subdivided into Hydrological Response Units (HRUs) which are areas with similar hydrological behavior and defined by a combination of soil and land cover maps (Beven, 2001; Kouwen et al., 1993). Vertical water and energy budgets are computed independently for each HRU in each catchment. Canopy interception is represented by a reservoir with maximum storage as a function of vegetation leaf area index. Soil water budget is computed simulating the soil as a single water reservoir. Soil infiltration and runoff are computed based on the variable contributing areal concept of the ARNO model (Todini, 1996), which is also used in the PDM (Moore and Clarke, 1981), VIC2L and LARSIM models. Energy budget is computed using basic surface meteorological variables and evapotranspiration from soil, vegetation and canopy into the atmosphere is estimated based on the Penman Monteith approach (Monteith, 1965; Shuttleworth, 1993; Allen et al., 1998). Subsurface flow is computed using an equation similar to the Brooks and Corey unsaturated hydraulic conductivity equation (Rawls et al., 1993). Percolation from soil layer to groundwater is calculated according to a simple linear relation between soil water storage and maximum soil water storage. Then, the flow generated within each catchment is routed to the stream network using three linear reservoirs (base flow, subsurface flow and surface flow). A detailed description of the model is given by Collischonn et al. (2007), and recent applications of the model are presented by Getirana et al. (2010), Collischonn et al. (2008), and Collischonn et al. (2005).

Until now, river flow routing within the MGB-IPH model used to be calculated using the Muskingum Cunge method, but in this paper we describe a mixed method, which employs a hydrodynamic model in flat reaches of the main rivers and the Muskingum–Cunge method in the headwater parts of the drainage network, where slope is usually higher.

3. The hydrodynamic model

The hydrodynamic model is based on the IPH-JV model, first developed by Tucci (1978). The model solves the full Saint Venant equations (Cunge et al., 1980):

\[
\frac{\partial Q}{\partial x} + b \frac{\partial h}{\partial t} = q_{\text{cut}} - q_{\text{fl}}
\]

(1)

\[
\frac{\partial h}{\partial t} + \frac{1}{2} \frac{\partial Q}{\partial x} + \left( gA - v^2 b \right) \frac{\partial h}{\partial x} - v \frac{\partial A}{\partial x} = gA(S_0 - S_f)
\]

(2)

where the first and second equations are the 1D channel mass and momentum conservation laws, Q (m³ s⁻¹) is river discharge, f (s) is time, x (m) is river longitudinal space coordinate, b (m) is river cross section width at free surface elevation, q_{\text{cut}} (m³ s⁻¹) is local channel lateral inflow (the sum of the surface, subsurface and base flow from the catchment), q_{\text{fl}} (m³ s⁻¹) is the river-floodplain flow exchange, h (m) is water depth, v (m s⁻¹) is flow velocity averaged over the cross section, g (m s⁻²) is acceleration due to gravity, A (m²) is the cross sectional flow area perpendicular to the flow direction and S_{0} (m m⁻¹) and S_{f} (m m⁻¹) are the bed slope and friction
slope in the x-direction. Friction slope is estimated using Manning’s equation. Flow at river confluences is modeled using a simple mass continuity equation and the energy equation discarding energy losses and the kinetic term (Cunge et al., 1980). It is also possible to model hydraulic river singularities such as dams or other hydraulic structures using internal boundary conditions.

The river reaches are discretized into several river cross sections (Fig. 1) where the hydraulic variables are computed. The model also divides the catchments into floodplain units (Fig. 1). These are areas between two river cross sections where the river-floodplain flow exchange and floodplain water storage are computed.

Flood inundation is simulated using a simple storage model (Cunge et al., 1980), which assumes that (i) the flow velocity parallel to the river direction is null on floodplains, (ii) the floodplain units act only as storage areas and (iii) the floodplain water level equals the water level at the main channel. Considering the model basic assumptions and the mass conservation law, the river-floodplain flow exchange \( q_f \) equals:

\[
q_f = \frac{A_f(z)}{dx} \frac{\partial h}{\partial t} = \frac{L(z)}{dx} \frac{\partial h}{\partial t}
\]

where \( A_f (m^2) \) is the flooded area and \( L (m) \) is the floodplain equivalent width, measured for each floodplain unit, as described in Section 4.5.

This is a simplistic approach that cannot fully represent all aspects of floodplain hydrodynamics such as floodplain water levels that are not equal to the adjacent main channel, floodplain flows subparallel to the main river channel, and bidirectional floodplain flows, that where observed by Alsdorf et al. (2007) using remote sensing data in the Amazon. Additionally, as observed by Alsdorf et al. (2003), water level changes in the main channel may not instantly propagate across the entire floodplain, as is implicitly considered in the model described here. However, the proposed model, which is relatively complex in terms of river hydraulics, but somewhat simplified in terms of floodplain simulation, may be appropriate because (i) this model is oriented to large scale applications and aimed at providing discharge, water level and flood extent results, by simulating translation and diffusion of flood waves, backwater effects and the influence of floodplain storage on flood waves, (ii) it is not oriented for studying smaller scale floodplain hydraulics, (iii) it is compatible with large scale applications, different from complex floodplain models which require higher computational resources and detailed floodplain bathymetry data, that cannot be provided by current global scale DEMs and (iv) it is in accordance with Trigg et al. (2009), who stated that “it is important to get the hydraulics of the main channel right before tackling the more complex interactions with the floodplain”.

The partial differential equations of the model are solved using a linear and implicit finite difference numerical scheme developed by Chen (1973), similar to the Preissman scheme (Cunge et al., 1980). Since the model simulates a river network with lots of confluences, the set of discretized equations forms a non symmetric sparse linear system. Then, for better computational efficiency, the matrix solver uses a modified Gauss elimination procedure based on a skyline storage method, avoiding the storage of null elements of the linear system of equations. This method was developed by Tucci (1978) and is similar to methods used by HEC-RAS (USACE, 2002). In addition, a modification to the method developed by Tucci (1978) is proposed. An additional pointer is set in the beginning of the simulation to store only algebraic operations that are actually needed for the Gauss elimination algorithm, i.e. it avoids operations with null elements, and then computes only those during the rest of the simulation. This improvement increased the speed of calculations but it will not be fully described here.

4. GIS based algorithms

Any hydrodynamic model should be applied using data from detailed surveys of river cross sections and, when necessary, floodplain extent. This kind of data is usually available only at small scale, for relatively short river reaches. For large scale hydrologic modeling the necessary information has to be derived from globally available data sets.

In our case, which is focused on applications of the model in large river basins, most of the necessary data is extracted from Digital Elevation Models (DEM) (e.g. the SRTM DEM described in Rabus et al., 2003; Farr et al., 2007) and some GIS algorithms. The algorithms use a raster representation of topography and all other spatial variables. The spatial variables are described through matrices of r rows and c columns, which store gridded values in \( r \times c \) positions defined as pixels. The pixel \((i,j)\) corresponds to the element of a matrix located at row \(i\) and column \(j\). The Digital Elevation Model is denoted by \(Z(i,j)\) and provides the terrain elevation in each point \((i,j)\). The algorithms for model parameter extraction include, as presented in Fig. 2, (1) DEM preprocessing, (2) river network and catchment discretization, (3) river cross section discretization, location and topology, (4) river cross section geometry, (5) river cross section bottom level, (6) floodplain geometry. These procedures are presented in the following sections. Examples of results from these algorithms are presented in Section 6.3.

4.1. DEM preprocessing and catchment discretization

DEM preprocessing and the river and catchment discretization algorithms presented in this section are computed using well known algorithms as presented in Jenson and Domingue (1988) and available in several GIS packages such as the ArcHydro tools (Maidment, 2002).

The DEM preprocessing algorithms include first the computation of the flow direction \(\delta(i,j)\) raster. Flow direction is a function of DEM that returns the direction of the largest slope (N, NE, E, SE, S, SW, W, NW or null). Given the flow direction, the coordinates of the downstream pixel hydrologically connected to \((i,j)\) are given by the function \(F\):

\[F(i,j) = (i,j) + \Delta\delta(i,j)\]
considering $\Delta(i,j)$ a function that returns the relative position $(\Delta_i, \Delta_j)$ of the downstream pixel (e.g. $\Delta(N) = (+1, +1)$ or $\Delta(NE) = (+1, -1)$). The flow direction algorithm also includes the removal of pits from the DEM to ensure hydrological continuity (Jenson and Domingue, 1988). The next step is the computation of the flow accumulation raster $A(i,j)$, that returns for each pixel $(i,j)$, the sum of the surface area of all upstream pixels (Burrough and McDonnell, 1998).

Later, using the flow direction and flow accumulation results, the river and catchment discretization is performed. The drainage $D(i,j)$ maps the river network on pixels with flow accumulation $A(i,j)$ larger than a threshold $A_{\text{min}}$ (Burrough and McDonnell, 1998). Then, the river network is segmented into several river reaches, defined by river segments between upstream and downstream confluences (or between a confluence and a river source) and each river reach gets a different code. Finally, catchment discretization is performed by identifying pixels located upstream of each river reach and the result is the catchment raster $B(i,j)$. For each pixel $(i,j)$, the algorithms follow downstream the flow path defined by the flow direction using the function $F$ until it finds a river reach and assigns $B(i,j)$ equal to the river reach code (Maidment, 2002).

4.2. Cross section discretization

A river reach within a catchment is then subdivided in several shorter sub-reaches, as shown in Fig. 1. These sub-reaches are limited by the positioning of model river cross-sections, where hydraulic variables $Q$ and $h$ are computed. For a given river reach $p$, the number of cross sections is defined as $n_p = L_p/\Delta x_p + 1$, where $\Delta x_p$ is the distance between two cross sections and $L_p$ is the length of the reach $p$. A predefined value of the distance $\Delta x$ between the cross sections is defined considering the spatial scale of interest and the performance of the numerical scheme of the hydrodynamic model (Cunge et al., 1980) and can be estimated a priori considering recommendations such as $\Delta x < cT/M$ as presented in Castellarin et al. (2009), where $c = (gh)^{1/2}$ is the flood wave celerity, $T$ is the period of the flood wave and $30 < M < 50$.

Then, the location of model cross sections is defined by dividing the river reach in a number of sub-reaches that results in the sub-reach length ($\Delta x_p$) as close as possible to a predefined value. The algorithm starts from any pixel of the reach $p$ and then follows upstream the river flow path using the flow direction raster, the drainage raster and the inverse of the function $F$ (Eq. (4)) until it finds the upstream extreme of the reach. The coordinates of the first cross section of the reach $p$ equals this current position. Then it sets the coordinates of the second cross section ($m = 2$) by repeating the following steps: (1) follow the reach flow path downstream using Eq. (4) and compute the accumulated distance $x$ from the first cross section; (2) when $x \geq (m - 1) \Delta x_p$ for the first time, set the coordinates of the $m$th cross section equal to the current position; (3) set $m = m + 1$ to compute the next cross section and repeat steps 1–3.

4.3. Cross section geometry

River cross sections are represented by a rectangular shape, with width $B(m)$ and maximum depth $H(m)$. Parameters $B$ and $H$ are estimated through geomorphologic equations of the type $H = a \cdot B^b$, where $a$ and $b$ are parameters and $A_{\text{res}}$ is the upstream drainage area. This kind of approach is used in several large scale simplified flow routing algorithms, such as Coe et al. (2007), Arora (2001) or Decharme et al. (2008). The parameters used in the case study are presented in Section 6.3.

4.4. Cross section bottom level

This section describes the procedure for estimating cross section bottom levels. This procedure takes three steps: (i) first a river longitudinal profile is extracted from the DEM for each river reach. Then, two corrections are applied to this longitudinal profile: (ii) removal of effects of vegetation and water depth and (iii) removal of random noise using an iterative moving average filter. The details of these procedures are described below.

A river longitudinal profile is extracted from the DEM for each reach to estimate river bottom level of computational cross sections. The algorithm follows the reach path from upstream to downstream as described in Section 4.2 and computes terrain elevation $z_{\text{DEM}}$ as a function of distance $x$ from the beginning of the reach:

$$z_{\text{DEM}}(x) = Z(i,j)$$

The elevations of the river longitudinal profile extracted from SRTM DEM are not equal to river bottom level $z_0(x)$ (m). Systematic errors related to vegetation, surface water effects and also random error related to noise on DEM data (Rabus et al., 2003; Sun et al., 2003; Farr et al., 2007; Kellndorfer et al., 2004; Lehner et al., 2006) affect the SRTM DEM elevation, which we call $z_{\text{DEM}}(x)$. We consider that the river bottom level $z_0(x)$ can be estimated using the following equation:

$$z_0(x) = z_{\text{DEM}}(x) - H(x) - H_{\text{veg}}(x) - \epsilon(x)$$

where $H$ (m) is the water depth, $H_{\text{veg}}$ (m) is the effective vegetation height (different from real vegetation height since SRTM C-band radar penetrates the canopy to some extent) and $\epsilon$ (m) is a random noise, as illustrated in Figs. 3 and 4.

The effective vegetation height $H_{\text{veg}}$ is expected to depend on the actual vegetation height, and on the ratio between river width $B$ (m) and DEM spatial resolution $res$ (m). For different relations between river width and DEM resolution, DEM elevation may be close to the actual water stage or more perturbed by vegetation height, as shown in Fig. 4. In large rivers, the elevation obtained from SRTM data is close to water level. In narrow rivers, $z_{\text{DEM}}$ is close to the elevation of the vegetation canopy, which covers the river. Considering this, a preliminary longitudinal profile of bottom levels $z_0$ is estimated by:
rivers. At the extreme of the reaches, the filter also takes a special form: \( z_{i,t+1} = 0.95z_{i,t} + 0.05z_{i+1,t} \). The main parameter of the filter is the maximum number of iterations \( l_{\text{max}} \). Its choice is subjective and a verification of the results is needed. A large (or small) value of \( l_{\text{max}} \) gives smooth (or noisy) longitudinal profiles. This iterative moving average filter approach was preferred under a simpler polynomial fit for each river reach (e.g. LeFavour and Alsdorf, 2005) because the latter creates discontinuities in the confluentes and between neighbor river reaches.

4.5. Floodplain geometry

Floodplains are characterized by a function \( A_{\text{fl}}(z,m) \) that relates water level \( z \) with flooded area \( A_{\text{fl}} \) for each sub-catchment connected to a sub-reach of the river. These sub-catchments, which are called here floodplain units, are denoted by \( P(i,j) \) and relate each pixel \( (i,j) \) to a cross section located immediately downstream. These are the areas hydrologically connected to the river segment between two cross sections, i.e. a set of pixels that are immediately upstream of a cross section (Fig. 1). To compute \( P(i,j) \), we use a raster watershed algorithm similar to the one used for catchment discretization (Burrough and McDonnel, 1998). The purpose of defining those floodplain units is only to derive local functions relating water stage and inundated area. Rainfall-runoff calculations are done at the catchment level only.

The function \( A_{\text{fl}} \) is defined using the DEM and floodplain units rasters. The total flooded area between the cross sections \( m \) and \( m-1 \) when the average water level equals \( z \) is computed as the sum of the surface area of the set of pixels inside the floodplain unit \( m \) that are lower than \( z \):

\[
A_{\text{fl}}(z,m) = \sum_{i,j \in S} dA(i,j)
\]

where \( S = \{(k,l) | P(k,l) = m, Z'_{\in (k,l)} \leq z\} \), \( dA(\text{km}^2) \) is the pixel surface area \( (\text{m}^2) \), \( Z'_{i,j} \) \((\text{m}) \) is a corrected DEM. A simple approach to correct the DEM is to consider a constant value of vegetation height \( H_{\text{veg}} \) for the whole DEM and compute \( Z'_{i,j} = Z(i,j) - H_{\text{veg}} \). Other approaches for DEM correction could also be used such as different \( H_{\text{veg}} \) values combined with land cover maps. The floodplain
equivalent width curve is computed as \( L(z, m) = \frac{A}{\Delta m} \), and a correction is made \( L = \max(0, L - B) \) to remove the main river width.

5. Flood inundation model

In this section, we describe two procedures: (i) how the model generates 2D water level results and (ii) how the model considers the flood extent in the water and energy budget computations.

Considering the model hypothesis to simulate the floodplain hydraulics, it is possible to generate inundation results in terms of floodplain water level or depth. The floodplain water depth \( Y_i(j, t) \) in the time interval \( t \) and pixel \( (i, j) \) is computed using previous results of 1D water levels in each cross section \( z \), of floodplain units raster \( P(i, j) \) and of corrected DEM \( Z(i, j) \):

\[
Y_i(j, t) = \max(P_{i(j)} - Z(i, j), 0)
\]

The flood extent is also considered in the surface water and energy budget of the model. To consider this, we use a methodology called here as Dynamic HRU. Basically, the fraction area of the Hydrological Response Units (HRUs) in each catchment is considered to be variable in time. The fraction area of the HRU representative of water surface equals to the total flooded area inside the catchment. Also, the areas of the other HRUs are adjusted to keep the sum of all HRUs equal the catchment area. Since we use a simplistic model for simulating floodplains, the Dynamic HRU approach may also have some limitations, although these might be less severe than not considering the flood extent time variability into water and energy balance computations.

6. The Purus River basin case study

As previously mentioned, the main objective of this work is to develop methods for large scale hydrodynamic modeling considering limited data and GIS based algorithms for parameter extraction from DEMs. In this section, we show an evaluation of the developed methods through an application of the model in the Purus River basin. First, we give a brief description of the Purus River basin and a presentation of the main data used. Subsequently, the evaluation of the developed methods is presented focusing on GIS procedures, model operation and feasibility and also some analyses on hydrological results. Hydrological results are not fully explored here, since further confirmation of the developed model and discussion about advantages of a full hydrodynamic model for large scale hydrologic modeling is presented in Paiva et al. (in review).

6.1. The Purus River basin

The Purus River is one of the main tributaries of the Amazon River (Fig. 5), its drainage area equals 370,000 km² and its average discharge is 11,000 m³ s⁻¹. The Purus River basin has typical characteristics of the Amazon region. It is a very large river, the main land cover of the basin is rain forest and both high rainfall rates and seasonality play an important role in hydrology. Purus River slopes are quite small (<5 cm/km) and there are large floodplains, which become seasonally inundated (Hess et al., 2003). The floodplain width on the lower Purus River is of the order of 30 km, which corresponds to approximately 30 times the main channel width. Due to these features, flood waves travel slowly along the river, taking 2 or 3 months from the headwaters to the river mouth (Paiva et al., in review). Significant backwater effects are also present on the main river (Meade et al., 1991) and its tributaries (as suggested by Paiva et al., in review).

6.2. Data

We used the SRTM DEM (Rabus et al., 2003; Farr et al., 2007) obtained from the Hydrosheds project data (Lehner et al., 2006). We used the DEM15s product, which is a DEM based on SRTM data with some corrections on missing data and 15° resolution (approximately 500 m). An HRU map with 12 classes was developed using soil and vegetation maps. The soil map used is a combination of the Brazilian database RADAMBrasil (Projeto RADAMBRASIL, 1982) and SOTERLAC/ISRIC (Dijkshoorn et al., 2005) for areas outside Brazil. Concerning the land cover, we used the “Vegetation Map of South America” developed by Eva et al. (2002). Hydrologic data (rainfall and river discharge and stage time series) was provided by the Brazilian National Water Agency (ANA – Agência Nacional de Águas). Precipitation data from rain gauges was used as input for the hydrological model. Discharge and water level data from gauging stations were used for calibration and validation of the model (see Fig. 5). We used four ANA stream gauges to validate the model, coded 13710001 (1), 13750000 (2), 13880000 (3) and 13980000 (4).

River cross sections surveys from stream gauge locations were used to develop geomorphologic equations relating river width and depth with drainage area for the Amazon region. Other meteorological input variables of the model – near surface air temperature, pressure, moisture, wind speed and incoming shortwave solar radiation – were obtained from NCEP reanalysis (Kalnay et al., 1996).

6.3. GIS procedures results

In this section we show the main results from the GIS based procedures for the large scale hydrodynamic model discretization and parameter estimation described in Section 4. The aim is to offer a better understanding about the GIS based procedures and to show its feasibility. For a better understanding and visualization, the results are shown (Fig. 6) in the smaller area (~165 x 165 km) highlighted in Fig. 5. This area will be called here as the zoom area. The DEM (figure 6a) shows the Purus River and its large floodplain flowing from the lower left corner of the figure to the northeast. The drainage raster \( D(i, j) \) and catchments raster \( B(i, j) \) (Fig. 6b and c) were created considering a threshold drainage area...
equal to 625 km$^2$, giving rise to 345 catchments with an average drainage area of 1000 km$^2$. The river reaches discretization into sub-reaches and cross sections (Fig. 6d) was performed considering $\Delta x < 10$ km. Fig. 6e shows results of floodplain units raster $P_{i,j}$, together with model river reaches cross sections. One of the floodplain units between two model cross sections is highlighted, and
will be used later to show some of the model simulation results. As an example, Fig. 7 shows the curve of water level versus flooded area \((A_f)\) extracted from DEM (Eq. (9)) of the area highlighted in Fig. 6e.

The parameters of the cross sections, i.e. width \(B (\text{m})\) and maximum depth \(H (\text{m})\), were estimated as functions of drainage area \(A_d (\text{km}^2)\), using the following geomorphologic equations:

\[
B = 0.8054A_d^{0.5289} \quad (R = 0.77) \quad \text{and} \quad H = 1.4351A_d^{0.1901} \quad (R = 0.72).
\]

These equations were developed for the Amazon River basin using cross sections profiles from 341 gauge stations located in the Brazilian part of the Amazon basin, including the rivers Araguaia and Tocantins (Paiva, 2009). These geomorphologic equations provide similar results as those previously developed by Coe et al. (2007) and Beighley et al. (2009).

Fig. 8 shows the longitudinal profile of the Purus River on the zoom area extracted from the DEM (black line) and after correction. The systematic errors related to vegetation and surface water effects were removed (Eq. (7)) considering the following parameters: \(H_{\text{reg}} = 17 \text{ m}, \alpha = 0.8\) and \(\beta = 1.0\). A single value for \(H_{\text{reg}}\) was adopted for the whole DEM, since rainforest land cover predominates in the Purus basin and its value was estimated through an inspection analysis of the DEM at the border of forested and deforested areas. Considering that the number of iterations \(i_{\text{iter}}\) of the iterative moving average filter is a subjective parameter, different values were tested. The filtered longitudinal profile is smoother when larger values of \(i_{\text{iter}}\) are used, however, when \(i_{\text{iter}}\) tends to infinity, the longitudinal profile gets too smooth, and information is lost. Therefore we used \(i_{\text{iter}} = 5000\), but this parameter should be verified in future applications using a similar inspection.

6.4. Streamflow and water level modeling results

In this section, Purus River simulation results are presented. The MGB-IPH parameters related to soil water budget were calibrated, using discharge data from the 1985 to 2005 period. Calibration was performed using a combination of manual and automatic optimization, as described by Collischonn et al. (2007), using the MOCOM-UA algorithm (Yapo et al., 1998). During the calibration phase only gauges located at the headwaters of the basin (shown by the grey squares in Fig. 5) where considered for hydrograph comparisons. The model was then validated using the downstream gauges, identified by numbers 1 to 4 in Fig. 5. Gauges 1–3 provided discharge data and gauges 2–4 water level data. The few hydrodynamic model and data preprocessing parameters were not calibrated and were set to: Manning’s \(n = 0.030\), \(H_{\text{reg}} = 17 \text{ m}\), \(i_{\text{iter}} = 5000\), \(\alpha = 0.8\) and \(\beta = 1.0\). Manning’s value was adopted based on values for natural rivers presented by Chow (1959) and by other authors who applied models in the Amazon region, such as LeFavour and Alsdorf (2005) and Trigg et al. (2009). Further details on model calibration can be found in Paiva et al. (in review) and Paiva (2009).

Fig. 9 shows observed daily stream flow from September 1990 to September 1991 at the stream gauges shown in Fig. 5. Hydrographs in Fig. 9 are arranged from upstream to downstream and the ratio between daily and mean stream flow is shown. Fig. 10 shows observed and simulated daily stream flow in Purus River at the same stream gauges. In these figures, one can see that while the flood peaks rise from more or less 6000 \(\text{m}^3\text{s}^{-1}\) to almost 15000 \(\text{m}^3\text{s}^{-1}\), the flood wave becomes smoother. Seasonal precipitation in this region causes alternated high and low water periods. Hydrographs of the upper part of the basin are noisy, with several peaks related to intense rainfall events. As the flood wave travels to the lower part of Purus River, it is attenuated and delayed due to the storage of high volumes of water on the floodplain. This process produces much smoother hydrographs and flood wave travel times of several months, as can be seen in Fig. 9. These hydrological features were relatively well represented by the model, as can be seen in Fig. 10. Also, error statistics comparing observed and simulated stream flow during the 1985–2005 period corroborate the good model performance, although large errors are found in specific time intervals, such as in high flows in Fig. 10c. Volume error, or bias, at the four gauge stations was: 9%, −1% and −5%. Nash–Sutcliffe efficiency coefficient values for each of the gauge stations were: 0.84, 0.89 and 0.91.

Fig. 11 shows observed and simulated water levels from 1987 to 1991 at stream gauge stations presented in Fig. 5. The main hydrological features observed in the water levels of Purus River are well represented by the model. The range of water levels is quite large—being 15, 15 and 11 m for the observed data in the three gauge stations. The simulated range of the water levels is very similar to the observed one. Large errors were only found at gauge 2 (−27%).
while at the other two gauges river stage amplitudes were correctly reproduced (errors of \(-6\%\) and \(-1\%\), respectively at gauges 3 and 4). As in Fig. 10, water levels shown in Fig. 11 also illustrate the attenuation and delay of the flood wave due to floodplain storage as it travels through the Purus River. Analysis of the hydrographs and relatively high Nash–Sutcliffe efficiency coefficient values at those three gauges (0.84, 0.90 and 0.92) suggest that the model is performing good and that the delay and smoothing of the flood wave is being well represented by the propagation module of the hydrological model.

6.5. Simulated flood inundation dynamics

Based on model water level results maps of floodplain inundation can be obtained by comparing water levels calculated at hydrodynamic cross sections with DEM elevations, according to the procedure presented in Section 5. In this section, we present some flood inundation results and discuss the potential of the model in simulating different types of inundation dynamics.

Fig. 12 shows the simulated water depth on the floodplain at three time intervals throughout the hydrological year 1995/1996 (Fig. 12a–c). It also shows the computational river reaches and cross sections of the model. The first water depth map was derived using water levels calculated at the beginning of the hydrological year, at low water season (15-October-1995), the second is from the high water season (22-April-1996) and the third is from the declining water level period (31-July-1996). Areas appearing white in Fig. 12 are not flooded, while dark areas are flooded to depths of several meters. It can be seen that the flooded areas expand from the dry to the wet season, and that floodplain areas located in the downstream area (right part of Fig. 12) are still inundated after the passing of the flow peak.

Detailed results are given for one of the floodplain units, which is highlighted in Fig. 12. For this floodplain unit we show time series of the flooded area \(A_f\) and exchange flux between river and floodplain \(q_{fl}\), as calculated by the model (Fig. 13a). Fig. 13b shows time series of the main river discharge \(Q\), again compared to \(q_{fl}\). Finally, Fig. 13c shows water level \(z\) in the floodplain and its temporal derivative \(dz/dt\).

The inundation dynamics are characterized by a strong seasonality driven by the main river discharge and water level variations. According to model results, floodplain lateral exchange is driven by the river water level variations. When the water level is rising (lowering), water flows from the main river (floodplain) to the floodplain (main river) and the flooded area also increases (decreases).

Since one of the basic assumptions of the model is that there is no water level difference between the main river and the connected floodplain, total flooded area exhibits a similar behavior of water levels. Also, the model assumes that all marginal lakes and other parts of the floodplain are fully connected to the river. This can be usually considered valid, since normally these marginal lakes are connected to the main river through small channels (e.g. Bonnet et al., 2008).

Simulation results show that the majority of water in the floodplain comes from the main river, and that inflow and outflow
Intensity is defined by variation in time of main river water level. Both features were reported by Bonnet et al. (2008) who analyzed inflow and outflow from a marginal lake of the Amazon River near Óbidos using water level observations and modeling results. Similar main river driven floodplain inundation dynamics were also observed by Alsdorf et al. (2010) in the mainstem Amazon floodplain using gravimetric and imaging satellite methods. In addition, inundation results of the model were validated in terms of flooded areas extent (Paiva et al., in review; Paiva, 2009) and a good performance of the model was found.

It is worth noting that the model presented here is not able to simulate a full 2D inundation dynamic, e.g. when two large rivers can exchange water through floodplain flow. This situation may occur in Amazon rivers, but it is not as common in this area as it is in the Pantanal region (Paz et al., 2010), and probably has relatively low impact on large scale modeling results. To improve the representation of floodplain flow a 2D model should be more appropriate (Bates and De Roo, 2000; Bates et al., 2010). The model presented here only considers the floodplain storage in the 1D hydrodynamic model, taking into account its effects on flood wave delay and attenuation; and then we use the 1D water levels calculated at cross section locations to generate 2D inundation maps. Although it has some limitations, the approach presented here can be considered sufficient for large scale hydrological modeling and has much less computational demands.

---

**Fig. 12.** Simulated floodplain water depth on the zoom area of Purus River on 15-October-1995 (a), 22-April-1996 (b) and 31-July-1996 (c).

**Fig. 13.** (a) Simulated daily flooded area $A_f$ (grey line) and floodplain lateral exchange $q_f$ (black line), (b) simulated daily discharge $Q$ (grey line) and floodplain lateral exchange $q_f$ (black line) and (c) simulated daily water level $z$ (grey line) and its temporal derivative $dz/dt$ (black line) in a floodplain unit of Purus River for the period of 1994–1999.
7. Conclusions

We describe the development of a model for large scale hydrologic and hydrodynamic modeling. The proposed approach is suitable for large scale applications in regions with limited data, since it uses regionalization methods and GIS algorithms that can be applied using globally available datasets.

The model and methods used to derive the necessary information were tested in the Purus River basin, which is a representative part of the Amazon basin in terms of river hydraulics and floodplain abundance. The case study shows the feasibility of the GIS based algorithms for parameter extraction. A comparison between observed and simulated discharges and water levels at gauging stations shows that the model is capable of reproducing the main hydrological features of the Purus River basin. Calibration of parameters related to the hydrodynamic model was not necessary. However, while our modeling approach for flow propagation in rivers is relatively complex and complete, our description of floodplain dynamics is very simple. Our approach does not fully reproduce what is actually happening in the floodplains, nevertheless realistic floodplain inundation maps were derived from the results of the model. Our main concern is large scale applications and the effects that floodplain inundation and emptying has on variables measured at the river gauges.

Model errors may be related to input data, and limitations of the model itself. In terms of the river flow propagation module, we believe that errors may be related to the input data uncertainty, e.g. DEM precision and errors related to vegetation and cross section geometry provided by geomorphologic relations. Nevertheless, results show that it is possible to employ full hydrodynamic models within large-scale hydrological modeling even using limited data for river geometry and floodplain characterization. In comparison to other large scale flow routing algorithms, the model can simulate additional hydrological features such as backwater effects and flood inundation dynamics and can provide other output variables as river water levels, total flooded area and 2D water depth as well.

In a forthcoming work (Paiva et al., in review), we present a model validation including an evaluation of model errors and its sources. We evaluate its advantages if compared with simplified flow routing algorithms and also relate the differences in results from both techniques with the role played by the floodplains and backwater effects.

Acknowledgements

The authors are grateful for the financial and operational support from the Brazilian agencies FINEP and ANA as for constructive comments of Marie-Paule Bonnet, Douglas Alsdorf, and anonymous reviewers.

References


