River network routing on the NHDPlus dataset

Cédric David, David Maidment, Guo-Yue Niu, Zong-Liang Yang, Florence Habets, Victor Eijkhout

River network routing on the NHDPlus dataset

Cédric H. David[1,2], David R. Maidment[2], Guo-Yue Niu[1,3], Zong-Liang Yang[1], Florence Habets[4] and Victor Eijkhout[5]

1. Department of Geological Sciences, Jackson School of Geosciences, University of Texas at Austin, Austin, Texas, USA.

2. Center for Research in Water Resources, University of Texas at Austin, Austin, Texas, USA.

3. Biosphere 2, University of Arizona, Tucson, Arizona, USA

4. UMR-Sisyphe 7619, CNRS, UPMC, Mines-Paristech, Paris, France.

5. Texas Advanced Computing Center, University of Texas at Austin, Austin, Texas, USA.

Key words

RAPID, streamflow, network, matrix, parallel computing, PETSc, TAO

Corresponding author

Cédric H. David

Department of Geological Sciences, Jackson School of Geosciences

The University of Texas at Austin

1 University Station C1160

Austin, TX 78712

cedric.david@mail.utexas.edu

Abstract

The mapped rivers and streams of the contiguous United States are available in a geographic information system (GIS) dataset called NHDPlus. This hydrographic dataset has about 3 million river and water body reaches along with information on how they are connected into networks. The USGS National Water Information System provides stream flow observations at about 20 thousand gages located on the NHDPlus river network. A river network model called RAPID is developed for the NHDPlus river network whose lateral inflow to the river network is calculated by a land surface model. A matrix-based version of the Muskingum method is developed herein which RAPID uses to calculate flow and volume of water in all reaches of a river network with many thousands of reaches, including at ungaged locations. Gages situated across river basins (not only at basin outlets) are used to automatically optimize the Muskingum parameters and to assess river flow computations; hence allowing the diagnosis of runoff computations provided by land surface models. RAPID is applied to the Guadalupe and San Antonio River Basins in Texas, where flow wave celerities are estimated at multiple locations using 15-minute data and can be reproduced reasonably with RAPID. This river model can be adapted for parallel computing and although the matrix method initially adds a large overhead, river flow results can be obtained faster than with the traditional Muskingum method when using a few processing cores, as demonstrated in a synthetic study using the Upper Mississippi River Basin.

1. Introduction

Land surface models (LSMs) have been developed by the atmospheric science community to provide atmospheric models with bottom boundary conditions (water and energy balance) and to serve as the land base for hydrologic modeling. Over the past two decades, overland and subsurface runoff calculations done by LSMs have extensively been used to provide water inflow to river routing models that calculate river discharge [*De Roo, et al.*, 2003; *Habets, et al.*, 1999a; 1999b; 1999c; 2008; *Lohmann, et al.*, 1998a; 1998b; 2004; *Maurer, et al.*, 2001; *Oki, et al.*, 2001; *Olivera, et al.*, 2000]. However, river routing within LSMs has traditionally been done using gridded river networks that best fit the computational domain used in LSMs. Today, geographic information system (GIS) hydrographic datasets are increasingly becoming available at the continental scale such as NHDPlus [*USEPA and USGS*, 2007] and the global scale such as HydroSHEDS [*Lehner, et al.*, 2006]. These datasets provide a vector-based representation of the river network using the "blue line" mapped rivers and streams. Furthermore, observations of the river systems are now widely available in databases such as the USGS National Water Information System for the United States in which thousands of gages are available along with their exact location on the NHDPlus river network. Most studies mentioned above – with the exception of Habets et al. [2008] – use a limited number of gages throughout large river basins, often focusing on gages located at river mouths. As the spatial and temporal resolutions of weather and climate models and their underlying land surface models increase, using gages located across basins would help diagnosing the quality of LSM computations. The latest work on general circulation models by the international scientific community, especially by the intergovernmental panel on climate change

[*Solomon, et al.*, 2007], opens potential studies of the evolution of water resources with global change. Using mapped streams and water bodies in LSMs could benefit the resulting assessment of the impact of global change in water resources by providing estimation of changes at the "blue line" level. Furthermore, the use of parallel computing is quite common in regional- to global-scale atmospheric and ocean modeling, but comparatively infrequent in modeling of large river networks. Generally, parallel computing can be utilized to either solve problems of increasing size [as done with the ParFlow model: *Jones and Woodward*, 2001; *Kollet and Maxwell*, 2006; *Kollet, et al.*, 2010] or to decrease computation time [see, for example: *Apostolopoulos and Georgakakos*, 1997; *Larson, et al.*, 2007; *Leopold, et al.*, 2006; *von Bloh, et al.*, 2010]. These two types of approaches to parallel computing are respectively referred to as scalability and speedup of calculations and the work presented herein focuses on the latter. Apostolopoulos and Georgakakos [1997] investigated the speedup of streamflow computations using hydrologic models in river networks as a function of network decomposition and of the computing time ratio between vertical and horizontal water balance calculations. Simple river routing within LSMs being traditionally performed by carrying computations from upstream to downstream, one way to speedup river flow modeling is to use a sequential river routing code to compute independent basins on different processing cores, as done in Leopold et al. [2006] and in Larson et al. [2007]. Such methods allow avoiding inter-processor communication but result in imbalanced computing loads when some basins are much larger than others. Leopold et al. [2006] partly addressed load imbalance by using parallel computing for surface water balance, but the river routing part remains sequential. von Blow et al. [2010] implemented a routing method in which computations do not have to be carried in order from upstream to downstream, therefore obtaining almost perfect speedup. The work developed herein investigates a way to obtain speedup while retaining traditional upstream-to-downstream computations which are used in most river routing schemes.

The present study links a land surface model with a new river network model called RAPID using NHDPlus for the representation of the river network and USGS National Water Information System (NWIS) gages for the optimization of model parameters and the assessment of river flow computations. All models and datasets used herein are available at least for the contiguous United States. The work presented here focuses first on the Guadalupe and San Antonio Basins in Texas (see Figure 1) together covering a surface area of about 26,000 km^2. These basins have about 5,000 river reaches and their corresponding catchments in the NHDPlus dataset (see Figure 2) out of 3 million for the United States. These two basins are also chosen for study because of significant contributions to surface water flow from groundwater sources, because of a large reservoir, at Canyon Lake, where the impacts of constructed infrastructure on flow dynamics have to be considered, and because these rivers flow out into an estuarine system at San Antonio Bay. A synthetic study of the performance of RAPID in a parallel computing environment is also presented using the Upper Mississippi River Basin (see Figure 3), which has about 180,000 river reaches in NHDPlus and covers an area of about 490,000 km^2.

The research presented in this paper aims at answering the following questions: how can a river model be developed for calculation of flow and volume of water in a river network of thousands of “blue-line” river reaches? How can the connectivity information in

NHDPlus be used to run a river network model in part of the United States? How can flow at ungaged locations be reconstructed? How can model computations be assessed and optimized based on all available measurements? How can parallel computing be used to speedup upstream-to-downstream computations of river flow within a large river network?

First, the development of the RAPID model presented. Then, the modeling framework for calculation of river flow in the Guadalupe and San Antonio River Basins using runoff data from a land surface model is developed, followed by results. Finally, the speedup of RAPID in a parallel computing environment is assessed.

2. Model development

The model presented here is named RAPID (Routing Application for Parallel computatIon of Discharge - http://www.geo.utexas.edu/scientist/david/rapid.htm). RAPID is based on the traditional Muskingum method that was first introduced by McCarthy [1938] and has been extensively studied in the literature in the past 70 years. The Muskingum method has two parameters, k and x, respectively a time and a dimensionless parameter. Among the most noteworthy papers related to the Muskingum method, Cunge [1969] showed the Muskingum method is a first-order approximation of the kinematic and diffusive wave equation and proposed a method known as the Muskingum-Cunge method – a second-order approximation of the kinematic and diffusive wave equation – in which the Muskingum parameters are computed based on mean physical characteristics of the river channel and of the flow wave. Koussis [1978] proposed a variable-parameter Muskingum method based on the Muskingum-Cunge method where k varies with the flow but x remains constant on the grounds that the Muskingum method is relatively insensitive to this parameter. Other variable-parameter Muskingum methods allow both k and x to vary [see, for example: *Miller and Cunge*, 1975; *Ponce and Yevjevich*, 1978], although these variable-parameter methods fail to conserve mass [*Ponce and Yevjevich*, 1978]. Notable large-scale uses of the variable-parameter Muskingum-Cunge method include Orlandini and Rosso [1998] and Orlandini et al. [2003]. More recently, Todini [2007] developed a mass-conservative variable-parameter Muskingum method known as the Muskingum-Cunge-Todini method.

As a first step, the traditional Muskingum method with temporally-constant parameters calculated partly based on the work of Cunge [1969] is used in this study because there are significant challenges to overcome in adapting the Muskingum method for river networks, in efficiently running it within a parallel computing environment and in developing an automated parameter estimation procedure before more sophisticated flow equations are used. However, the physics of flow could be improved with many variations based on the Muskingum method or adapted to the Saint Venant equations.

2.1 Calculation of flow and volume of water in a river network

In a network of thousands of reaches, matrices are needed for network connectivity and flow computation. The backbone of RAPID is a vector-matrix version of the Muskingum method shown in Equation (1) and derived subsequently in this section.

$$\left(\mathbf{I}-\mathbf{C_1}\cdot\mathbf{N}\right)\cdot\mathbf{Q}\left(t+\Delta t\right)=\mathbf{C_1}\cdot\mathbf{Q^e}\left(t\right)+\mathbf{C_2}\cdot\left[\mathbf{N}\cdot\mathbf{Q}\left(t\right)+\mathbf{Q^e}\left(t\right)\right]+\mathbf{C_3}\cdot\mathbf{Q}\left(t\right) \tag{1}$$

where t is time and Δt is the river routing time step. The bolded notation is used for vectors and matrices. $\mathbf{I}$ is the identity matrix. $\mathbf{N}$ is the river network matrix. $\mathbf{C_1}$, $\mathbf{C_2}$ and $\mathbf{C_3}$ are parameter matrices. $\mathbf{Q}$ is a vector of outflows from each reach, and $\mathbf{Q^e}$ is a vector of lateral inflows for each reach. Such a vector-matrix formulation of the Muskingum method has to our knowledge never been previously published.

Equation (1) is used for river network routing and can be solved using a linear system solver. The vector-matrix notation provides one flow equation for the entire river network, therefore avoiding spatial iterations. For a river network with m river reaches, all vectors are of size m and all matrices are square of size m. Each element of a vector corresponds to one river reach in the network. For performance purposes, all matrices are stored as sparse matrices (only the non-zero values are recorded). A five-reach, two-node and two-gage river network is used here to clarify the mathematical formulation of the river network model and is shown in Figure 4a). The river network is made up of a combination of river reaches similar to that of Figure 4b). The model formulation is presented here for a small river network but can be generalized to any size of river network.

$\mathbf{Q}$ is a vector of the outflows Q_j of all reaches of the river network, where j is the index of a river reach within the network:

$$\mathbf{Q}(t)=\begin{bmatrix} Q_1(t) \\ Q_2(t) \\ Q_3(t) \\ Q_4(t) \\ Q_5(t) \end{bmatrix}=\left[Q_j(t)\right]_{j\in[1,m]} \quad (2)$$

$\mathbf{Q^e}$ is a vector of flows Q_j^e that are lateral inflows to the river network. Lateral inflows include runoff, groundwater or any type of forced inflow (outflow at a dam, pumping, etc.):

$$\mathbf{Q^e}(t)=\left[Q_j^e(t)\right]_{j\in[1,m]} \quad (3)$$

$\mathbf{Q^e}$ is provided by a land surface model, whose time step is coarser than the river routing time step. Two assumptions are made in the development of RAPID, one regarding the temporal variability of $\mathbf{Q^e}$ and one regarding the location at which $\mathbf{Q^e}$ enters the river network. In this study, the river routing time step is 15 minutes and inflow from land surface runoff is available every 3 hours. In the derivation of Equation (1), $\mathbf{Q^e}$ is assumed constant (i.e. $\mathbf{Q^e}(t+\Delta t)=\mathbf{Q^e}(t)$) over all 15-minute river routing time steps included within a given land surface model 3-hour time step. This partial temporal uniformity simplifies the river network model formulation, limits the quantity of input data and facilitates the coupling with land surface models. This assumption is valid at all times except at the last routing time steps before a new $\mathbf{Q^e}$ is made available by the land surface model. Also, the external inflow $\mathbf{Q^e}$ is assumed to enter the network as an addition to the upstream flow. With these two assumptions, the Muskingum method applied to reach 5 in Figure 4b) gives the following:

$$\begin{aligned} Q_5(t+\Delta t) = C_1 \cdot \left[Q_3(t+\Delta t) + Q_4(t+\Delta t) + Q^e{}_5(t) \right] \\ + C_2 \cdot \left[Q_3(t) + Q_4(t) + Q^e{}_5(t) \right] \\ + C_3 \cdot Q_5(t) \end{aligned} \quad (4)$$

where C_1, C_2 and C_3 are the Muskingum parameters that are stated in Equation (6). The reader should note that these two assumptions are equivalent to using a unit-width lateral inflow along with a term C_4 as found in available literature [see, for example: *Fread*, 1993; *NERC*, 1975; *Orlandini and Rosso*, 1998; *Ponce*, 1986]. Equation (1) is a generalization of Equation (4) using a vector-matrix notation.

$\mathbf{N}$ is a network connectivity matrix. Berge [1958] proposed the concept of matrices associated with graphs. This concept can be applied to the river network in Figure 4a) in order to create the network matrix $\mathbf{N}$ given in Equation (5) in both full and sparse formats. The network connectivity matrix is a square matrix whose dimension is the total number of reaches in the network. A value of one is used at row i and column j if reach j flows into reach i and zero is used everywhere else.

$$\mathbf{N} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} & & & & \\ & & & & \\ 1 & 1 & & & \\ & & & & \\ & & 1 & 1 & \end{bmatrix} \tag{5}$$

The upstream inflow to the network can therefore be computed by multiplying the network connectivity matrix $\mathbf{N}$ by the vector of outflows $\mathbf{Q}$. In case of a divergence in the river network (when going downstream) or in case of a loop, a unique reach (the major divergence) is used to carry all the upstream flow and the other reaches (minor divergences) carry only the flow that results from their lateral inflow. This formulation could be modified to take into account given fractions of flows that separate into different parts of a divergence if that information is available.

$\mathbf{C_1}$, $\mathbf{C_2}$ and $\mathbf{C_3}$ are diagonal matrices with their diagonal elements being the coefficients used in the Muskingum method [*McCarthy*, 1938], respectively C_{1j}, C_{2j} and C_{3j} such that:

$$C_{1j}=\frac{\frac{\Delta t}{2}-k_j\cdot x_j}{k_j\cdot\left(1-x_j\right)+\frac{\Delta t}{2}},\quad C_{2j}=\frac{\frac{\Delta t}{2}+k_j\cdot x_j}{k_j\cdot\left(1-x_j\right)+\frac{\Delta t}{2}},\quad C_{3j}=\frac{k_j\cdot\left(1-x_j\right)-\frac{\Delta t}{2}}{k_j\cdot\left(1-x_j\right)+\frac{\Delta t}{2}} \tag{6}$$

where k_j is a storage constant (with dimension of a time) and x_j a dimensionless weighting factor characterizing the relative influence of the inflow and the outflow on the volume of the reach j. The Muskingum method is stable for any $x\in[0,0.5]$, regardless of the value of k and Δt [*Cunge*, 1969]. For any j: $C_{1j}+C_{2j}+C_{3j}=1$.

In RAPID, the parameters k and x of the Muskingum method are allowed to differ from one river reach to another, and corresponding vectors are defined in Equation (7):

$$\mathbf{k}=\left[k_j\right]_{j\in[1,m]}\quad,\quad\mathbf{x}=\left[x_j\right]_{j\in[1,m]} \tag{7}$$

The constants defined in Equation (6) are used as the diagonal elements of the matrices $\mathbf{C_1}$, $\mathbf{C_2}$ and $\mathbf{C_3}$. Equation (8) shows an example for $\mathbf{C_1}$. $\mathbf{C_2}$ and $\mathbf{C_3}$ are treated similarly.

$$\mathbf{C_1}=\begin{bmatrix} C_{1_1} & & & & \\ & C_{1_2} & & & \\ & & C_{1_3} & & \\ & & & C_{1_4} & \\ & & & & C_{1_5} \end{bmatrix} \tag{8}$$

The sum $\mathbf{C_1} + \mathbf{C_2} + \mathbf{C_3}$ equals the identity matrix.

The calculation of the volume of water in a given reach can be needed for coupling with groundwater models. Here, the first order, explicit, forward Euler method is applied to the continuity equation to calculate the volume of water in each river reach of the network, as shown in Equation (9) where the first, second and third terms of the right-hand-side are the volume of water that respectively were in the river reach, flowed into the reach, and discharged from the reach:

$$\mathbf{V}(t+\Delta t) = \mathbf{V}(t) + \left[\mathbf{N}\cdot\mathbf{Q}(t) + \mathbf{Q^e}(t)\right]\cdot\Delta t - \mathbf{Q}(t)\cdot\Delta t \tag{9}$$

where $\mathbf{V}$ is a vector of the volume of water V_j in each river reach j:

$$\mathbf{V}(t) = \left[V_j(t)\right]_{j\in[1,m]} \tag{10}$$

Details on the massively-parallel implementation of the matrix-based Muskingum method presented in this section, and of the automated parameter estimation presented in the section below are given in Appendix A.

2.2 Parameter estimation

In order to estimate the parameters $\mathbf{k}$ and $\mathbf{x}$ to be used in RAPID, an inverse method is developed. The principle of an inverse method is to optimize the parameters of a model so that the outputs of the model approach observations. A cost function reflecting the difference between model calculations and observations is needed to assess the quality of a set of model parameters. The best set of parameters is chosen as the set that minimizes the cost function, and is determined through optimization. A square-error cost function ϕ is chosen:

$$\phi(\mathbf{k},\mathbf{x})=\sum_{t=t_o}^{t=t_f}\left[\left(\frac{\overline{\mathbf{Q}}(t)-\mathbf{Q}^{\mathbf{g}}(t)}{f}\right)^T\cdot\mathbf{G}\cdot\left(\frac{\overline{\mathbf{Q}}(t)-\mathbf{Q}^{\mathbf{g}}(t)}{f}\right)\right] \tag{11}$$

where the summation is made daily. The T in exponent is for vector transpose. t_o and t_f are respectively the first day and last day used for the calculation of ϕ. The model parameter vectors $\mathbf{k}$ and $\mathbf{x}$ are kept constant within the temporal interval $[t_o, t_f]$, and the cost function is calculated several times with different sets of parameters during the optimization procedure. f is a scalar that allows ϕ to be on the order of magnitude of 10^1 which is helpful for automated optimization procedures. $\overline{\mathbf{Q}}(t)$ is the daily-average outflow vector, calculated based on the mean of all routing time steps in a given day. $\mathbf{Q}^{\mathbf{g}}(t)$ is a vector with the total number of river reaches for dimension, with the daily value observed $Q_j^g(t)$ corresponding to reach j where gage measurements are available, and zero where no gage is available. $\mathbf{G}$ is a sparse diagonal matrix that allows the dot-product to survive only where gages are available, so that $\mathbf{G}$ has a value of one on the diagonal element of index j if a gage is available on reach j and zero everywhere else. Using the example network given in Figure 4a), $\mathbf{G}$ and $\mathbf{Q^g}(t)$ take the following form:

$$\mathbf{G}=\begin{bmatrix} & & & & \\ & & & & \\ & & 1 & & \\ & & & & \\ & & & & 1 \end{bmatrix} \quad , \quad \mathbf{Q^g}(t)=\begin{bmatrix} 0 \\ 0 \\ \mathrm{Q}_3^{\mathrm{g}}(t) \\ 0 \\ \mathrm{Q}_5^{\mathrm{g}}(t) \end{bmatrix} \tag{12}$$

According to Fread [1993], $x \in [0.1;0.3]$ in most streams. By analogy with the kinematic wave equation, Cunge [1969] showed that the parameter k of the Muskingum method is the travel time of a flow wave through a river reach. For a given river reach j of length L_j where a flow wave of celerity c_j travels, k_j is obtained by dividing the length by the celerity of the wave, as shown in Equation (13):

$$k_j = \frac{L_j}{c_j} \tag{13}$$

Although the routing model defined by Equation (1) allows for variability of the parameters (k_j, x_j) on a reach-to-reach basis, attempting to automatically estimate model parameters independently for all the reaches of a basin would be a costly undertaking. Therefore, the search for optimal parameters is limited to determining two multiplying factors λ_k and λ_x such that:

$$k_j = \lambda_k \cdot \frac{L_j}{c_j} \quad , \quad x_j = \lambda_x \cdot 0.1 \tag{14}$$

To minimize the influence of the initial guess on the optimization procedure, three different initial guesses for (λ_k, λ_x) are used. Out of the three corresponding optimal (λ_k, λ_x) obtained, only the one couple leading to the minimum value of the cost function ϕ is kept. Therefore, the optimization procedure leads to only one optimal couple (λ_k, λ_x) for a given basin in the network. Note that – as a first step – $\mathbf{x}$ is here constant over a given basin on the grounds that the Muskingum method is relatively insensitive to this parameter [*Koussis*, 1978]. Some data available in NHDPlus (such as mean flow, mean velocity, slope, etc.) associated with available formulations for $\mathbf{x}$ [for example: *Cunge*, 1969; *Orlandini and Rosso*, 1998] could be used to improve the proposed method.

3. Application

RAPID is designed to handle large routing problems. Given a river network and connectivity information as well as lateral inflow to the river network, RAPID can run on any river network. In this study, a framework for computation of river flow in the Guadalupe and San Antonio River Basins is developed that uses a one-way modeling framework with an atmospheric dataset, a land surface model and RAPID as the river model. This section presents how the Guadalupe and San Antonio River Basins are described in the NHDPlus dataset, how a land surface model is used to provide lateral inflow to the river network, and how the meteorological forcing is prepared.

3.1. RAPID used on NHDPlus

There are a total of 5175 river reaches with known direction and connectivity within the NHDPlus description of the Guadalupe and San Antonio river basins (as shown in Figure 2). These 5175 reaches have an average length of 3.00 km and the average catchment defined around them is 5.11 km^2 in area; all are used for this study. Details on the fields used in the NHDPlus dataset including the unique identifier COMID used for all river reaches and their corresponding catchments; and on how NHDPlus is used with RAPID are given in Appendix B. In this study, the vector of outflows in all river reaches $\mathbf{Q}$ was arbitrarily initialized to the uniform value of 0 m^3s^{-1} prior to running RAPID.

3.2. Land surface model and coupling with RAPID

Within this study, the core physical model governing the one-dimensional vertical fluxes of energy and moisture is the Community Noah Land Surface Model with Multi-Physics Options, hereafter referred to as Noah-MP [*Niu, et al.*, 2010]. Noah-MP offers multiple options for choosing the modeling of certain physical phenomena. In this study, the soil moisture factor for stomatal resistance is of “Noah type” [*Niu, et al.*, 2010] and the runoff scheme is from “SIMGM” [*Niu, et al.*, 2007]. The soil column is 2 meter deep, below which is an unconfined aquifer. In order to represent the characteristics of the structural soil over the model domain, the saturated hydraulic conductivity, which is determined by the soil texture data, is enlarged by factor of ten (through calibration). The soil hydrology of Noah (soil moisture) is run at an hourly time step and runoff data are produced every three hours. In this study, the state variables of Noah were initialized through a spin-up method.

Noah-MP calculates the amount of water that runs off on and below the land surface. This quantity is used to provide RAPID with the water inflow from outside of the river network. David et al. [2009] presented a coupling technique using a hydrologically enhanced version of the Noah LSM called Noah-distributed [*Gochis and Chen*, 2003] that allows physically-based modeling of the horizontal movement of surface and subsurface water from the land surface to a river reach. In interest of a simpler coupling scheme, the work of David et al. [2009] has been modified. In this study, a flux coupler between Noah and RAPID is developed using the catchments available in the NHDPlus dataset.

The NHDPlus catchments contributing runoff to each river reach were determined as part of the NHDPlus development using a digital elevation model and its associated flow accumulation and flow direction grids. These grids have a native resolution of 30 m. The map of catchments is available in NHDPlus in both gridded (at 30-m resolution) and vector formats in a shapefile. Running a land surface model at a 30-m resolution is very resource demanding. Therefore, a coarser resolution of 900 m cell size is chosen. The shapefile of NHDPlus catchment boundaries is converted to a grid of size 900 m. Within this conversion process, the accuracy of the boundaries of the catchments is lowered but the catchment boundaries are reasonably respected and the computational cost of the land surface model calculations is reasonable. For each 3-hour output of the Noah model, surface and subsurface runoff data is superimposed onto the catchment grid, and all runoff that corresponds to the catchment of each river reach is summed and used as the water inflow to the river reach. Figure 5 shows the principle of the flux coupler in which the 900-m runoff data generated by the Noah model is superposed to the 900-m map of NHDPlus catchment COMIDs to determine the lateral inflow for NHDPlus reaches used by RAPID.

Therefore, no horizontal routing is used between the land surface and the river network in the proposed scheme. This differs from some other models that use runoff from a one-dimensional model to force a river routing model. For instance, the two dimensional wave equation is used in Gochis and Chen [2003] or the linear reservoir equation is used in Ledoux et al. [1989].

The coupling method used here can be adapted to any land surface model that computes surface and subsurface runoff on a grid. This coupling technique is automated in a Fortran program.

3.3. Meteorological forcing

Land surface models need meteorological forcing in order to compute the water and the energy balance at the surface. The Noah LSM requires seven meteorological parameters: precipitation, specific humidity, air temperature, air pressure, wind speed, downward shortwave and downward longwave radiation. Hourly precipitation is obtained from

NEXRAD and downscaled from its original resolution (4.763 km) to 900 m using the method developed in Guan et al. [2009]. All other meteorological parameters are downloaded from the 3-hourly North American Regional Reanalysis (NARR) and converted from its original resolution (32.463 km) to 900 m using a simple triangle-base linear interpolation. All meteorological data are prepared for four years (01 January 2004 – 31 December 2007).

4. Calibration and results for the Guadalupe and San Antonio River Basins

The framework for computation of river flow that is developed in the previous section is used to calculate river flow in all 5175 river reaches of the Guadalupe and San Antonio River Basins for four years (01 January 2004 – 31 December 2007). In this section, flow wave celerities in several sub-basins are estimated from measurements, the model parameters used in RAPID are presented, and flows computed are compared to observed flows. Issues related to the time step used in RAPID and to the simulated wave celerities are also presented.

4.1. Estimation of wave celerities

The USGS Instantaneous Data Archive (http://ida.water.usgs.gov/ida/) provides 15-minute flow data that can be used to determine the flow wave celerity. Data at fifteen gaging stations within the two basins studied are obtained from IDA over two time periods (01 January 2004 – 30 June 2004 and for 01 January 2007 – 30 June 2007). The maximum lagged cross-correlation between hydrographs at two consecutive gaging stations is used to determine the flow wave celerity. The lagged cross-correlation ρ is a measure of similarity between two wave forms as a function of a lag time τ_{lag} applied to one of them, as shown in Equation (15).

$$\rho = \frac{\sum\left(Q^a(t)-\overline{Q^a}\right)\cdot\left(Q^b\left(t+\tau_{lag}\right)-\overline{Q^b}\right)}{\sqrt{\sum\left(Q^a(t)-\overline{Q^a}\right)^2\cdot\sum\left(Q^b\left(t+\tau_{lag}\right)-\overline{Q^b}\right)^2}} \tag{15}$$

where Q^a and Q^b are the flows measured at the upstream and downstream station, respectively; and the summation is here made every 15 minutes for 6 months. Figure 6 shows the correlation as a function of increasing lag time between three different sets of consecutive gaging stations. The lag time giving the maximum correlation is taken as the travel time τ_{travel} for the flow wave between the two stations. The travel times are estimated for eleven sets of two stations and are shown on Table 1. Travel times of 0 s are reported at two stations, where the flow wave is probably too fast to be captured by 15-minute measurements. The wave celerity c is then computed using Equation (16)

$$c = \frac{d}{\tau_{travel}} \tag{16}$$

where d is the distance between two stations. The NHDPlus Flow Table Navigator Tool (http://www.horizon-systems.com/nhdplus/tools.php) is used to estimate the curvilinear distance between two stations along the NHDPlus river network that are shown on Table 1. The wave celerity has been estimated for eleven sub-basins within the Guadalupe and San Antonio river basins. Table 2 shows the values that are obtained for the two time periods considered, as well as their average. Figure 7 shows the corresponding sub-basins as well as the locations of all gaging stations.

4.2. Parameters used in RAPID

RAPID needs two vectors of parameters $\mathbf{k}$ and $\mathbf{x}$ that can either be determined using physically-based equations, through optimization, or a combination of both. In this study, daily stream flow data are obtained from the USGS National Water Information

System (http://waterdata.usgs.gov/nwis) in order to use the built-in parameter estimation. Within the Guadalupe and San Antonio river basins, NWIS has 74 gages that measure flow, 36 of them having full records of daily measurements the four years studied (01 January 2004 – 31 December 2007). These 36 stations are used for parameter estimation. Four sets of model parameters – denoted by the superscripts α, β, γ and δ – are used in this study. These sets of parameters are all based on Equation (14) which is used with a uniform wave celerity of $c^0 = 1km \cdot h^{-1} = 0.28 m \cdot s^{-1}$ throughout the basin or with the celerities c_j determined based on the IDA lagged cross-correlation study.

The first set, $(\mathbf{k}^{\boldsymbol{\alpha}}, \mathbf{x}^{\boldsymbol{\alpha}})$ is obtained from parameter estimation shown in Equation (11) using the uniform wave celerity $c^0 = 0.28 m \cdot s^{-1}$ and the resulting values of the two multiplying factors λ_k and λ_x of Equation (14) are:

$$\begin{aligned} k_j^{\alpha} &= \lambda_k^{\alpha} \cdot \frac{L_j}{c^0} \quad , \quad x_j^{\alpha} = \lambda_x^{\alpha} \cdot 0.1 \\ \lambda_k^{\alpha} &= 0.131042 \ , \quad \lambda_x^{\alpha} = 2.58128 \end{aligned} \tag{17}$$

The parameters $(\mathbf{k}^{\boldsymbol{\beta}}, \mathbf{x}^{\boldsymbol{\beta}})$ are determined without optimization using the celerities c_j determined based on the IDA lagged cross-correlation study and set to:

$$\begin{aligned} k_j^{\beta} &= \lambda_k^{\beta} \cdot \frac{L_j}{c_j} \quad , \quad x_j^{\beta} = \lambda_x^{\beta} \cdot 0.1 \\ \lambda_k^{\beta} &= 1 \quad , \quad \lambda_x^{\beta} = 1 \end{aligned} \tag{18}$$

The third set of parameters $(\mathbf{k}^{\gamma}, \mathbf{x}^{\gamma})$ is obtained through optimization using the celerities c_j determined based on the IDA lagged cross-correlation study and the resulting values are:

$$\begin{aligned} k_j^{\gamma} &= \lambda_k^{\gamma} \cdot \frac{L_j}{c_j} \quad , \quad x_j^{\gamma} = \lambda_x^{\gamma} \cdot 0.1 \\ \lambda_k^{\gamma} &= 0.617188 \,, \quad \lambda_x^{\gamma} = 1.95898 \end{aligned} \tag{19}$$

The optimization converges to a value of **k** that is 38% smaller than that estimated with the IDA lagged cross-correlation, suggesting that a faster flow wave in the river network produces better flow calculations. In the present study, routing on the land surface from the catchment to its corresponding reach is not modeled. Therefore, one would expect that the optimized flow celerity in the river network would be slower than that estimated from river flow observations, which is not the case here. This suggests that runoff is either produced too slowly or too far upstream of each gage; maybe because runoff in land surface models is often calibrated based on a lumped value at the downstream gage of a basin, as was done here with Noah-MP. Further details on the quality of runoff simulations are given in Section 4.4.

The fourth set of parameters $(\mathbf{k}^{\delta}, \mathbf{x}^{\delta})$ is determined for a better match of celerity calculations, as explained later in this paper.

4.3. Time step of RAPID simulation

Cunge [1969] showed that the Muskingum method is stable for any $x \in [0, 0.5]$ and that the wave celerity computed by the Muskingum method approaches the theoretical wave celerity of the kinematic wave equation if the time step of the river routing equals the travel time of the wave (for $x = 0.5$), as shown in Equation (20):

$$\forall j \in [1, m] \qquad c_j \simeq \frac{L_j}{\Delta t} \tag{20}$$

However, both the celerity of flow and the length of river reaches vary along the network; and the model formulation of RAPID allows for only one unique value of the time step Δt be chosen. In the Guadalupe and San Antonio River Basins, the mean length is 3 km and the median length is 2.4 km. The probability density function and the cumulative density functions for the lengths of river reaches are shown in Figure 8. The celerities estimated earlier are on the order of $c = 2.5 m \cdot s^{-1}$. Using the median value of the reach length along with $c = 2.5 m \cdot s^{-1}$, Equation (20) gives $\Delta t = 960 s$. In order to have an integer conversion between the river routing time step and the land surface model time step (3 hours), a value of $\Delta t = 900 s = 15 \min$ is chosen.

4.4. Analysis of the quality of river flow computation

For various model simulations, the average and the root mean square error (RMSE) of computed flow rate are calculated using daily data and are given in Table 3. The Nash efficiency [*Nash and Sutcliffe*, 1970] is bounded by the interval $\left]-\infty, 1\right]$ and gives an estimate of the quality of modeled river flow computations when compared to observations; and is also given in Table 3. An efficiency of 1 corresponds to a perfect model and 0 corresponds to a model producing the mean of observations. The results shown for a lumped model correspond to when runoff from Noah is accumulated at the gage directly without any routing. The average values of flow in RAPID simulations are tied to the amount of runoff water calculated by the Noah LSM and the bias generated by the land surface model cannot be fixed by RAPID. However, the internal connectivity of the NHDPlus river network is well translated in RAPID and mass is conserved within RAPID since the flow rates in the lumped simulation and in all four simulations of RAPID are the same. Figure 9 shows the ratio between observed and lumped stream flow at 17 gages located across the Guadalupe and San Antonio River Basins. This ratio is around unity downstream of the Guadalupe and San Antonio Rivers, but is greater than 7 upstream; suggesting that runoff is most likely overestimated at the center of the basin. Additionally, runoff is largely underestimated at two stations just downstream of the outcrop area of the Edwards Aquifer: the Comal River at New Braunfels and the San Marcos River at San Marcos. These stations measure large average stream flow (respectively 10.59 m^3/s and 5.9 m^3/s) although draining a relatively small area (respectively 336 km^2 and 129 km^2), and are actually two of the largest springs in Texas. These flows are much larger than the lumped runoff (respectively 0.67 m^3/s and 0.26 m^3/s), which is expected because the modeling framework presented herein does not does not explicitly simulate aquifers.

However, the RAPID simulations $(\mathbf{k}^{\alpha},\mathbf{x}^{\alpha})$, $(\mathbf{k}^{\beta},\mathbf{x}^{\beta})$ and $(\mathbf{k}^{\gamma},\mathbf{x}^{\gamma})$ lead to a smaller RMSE and a higher Nash Efficiency than the lumped runoff. This shows that an explicit river routing scheme with carefully-chosen parameters allows obtaining better stream flow calculations than a simple lumped runoff scheme, as expected.

Within the different RAPID simulations, the set of parameters $(\mathbf{k}^{\gamma}, \mathbf{x}^{\gamma})$ gives the best results for RMSE and Nash efficiency, followed by $(\mathbf{k}^{\beta}, \mathbf{x}^{\beta})$, $(\mathbf{k}^{\alpha}, \mathbf{x}^{\alpha})$ and $(\mathbf{k}^{\delta}, \mathbf{x}^{\delta})$. Therefore, a greater spatial variability in the values of k contributes to the quality of model results, and the built-in optimization in RAPID further enhances these model results. An example hydrograph for the Guadalupe River near Victoria TX is shown in Figure 10, and is computed using $(\mathbf{k}^{\gamma}, \mathbf{x}^{\gamma})$.

4.5. Comparison between estimated and computed wave celerities

In order to assess the capacity of the modeling framework to reproduce surface flow dynamics, the celerity of the flow wave in outputs from RAPID are computed. Fifteen-minute river flow is computed with RAPID, and the lagged cross-correlation presented earlier is used to calculate the wave celerity within the RAPID simulation. Table 2 shows the celerities that are computed from RAPID outputs. In the first three sets of model parameters used, the wave celerities simulated in RAPID are greater than those observed. One can also notice than even for $(\mathbf{k}^{\beta}, \mathbf{x}^{\beta})$, the model-simulated celerities are different than the observed celerities which are used to determine the vector $\mathbf{k}^{\beta}$ itself. This was predicted by Cunge [1969] who showed that the difference between the celerity of the kinematic wave equation and that computed using the Muskingum method is a function of both x and the quotient $\Delta t / L_j$. Only the specific values $x = 0.5$ and Δt verifying $\Delta t / L_j = c_j$ allow obtaining the same celerity. Furthermore, the work herein is done in a river network, and the celerity estimated between two points does not correspond only to the main river stem but rather to a combination of all river reaches present in the network in between the two points. The ratio of the average celerities from RAPID using $(\mathbf{k}^{\boldsymbol{\beta}}, \mathbf{x}^{\boldsymbol{\beta}})$ over the average observed celerities is 1.54. As a final experiment, a new set of parameters $(\mathbf{k}^{\boldsymbol{\delta}}, \mathbf{x}^{\boldsymbol{\delta}})$ is created to account for the faster waves in RAPID.

$$\begin{aligned} k_j^{\delta} &= \lambda_k^{\delta} \cdot \frac{L_j}{c_j} \quad , \quad x_j^{\delta} = \lambda_x^{\delta} \cdot 0.1 \\ \lambda_k^{\delta} &= 1.54 \quad , \quad \lambda_x^{\delta} = 1 \end{aligned} \tag{21}$$

Table 2 shows that the parameters $(\mathbf{k}^{\boldsymbol{\delta}}, \mathbf{x}^{\boldsymbol{\delta}})$ allow for wave celerities that are closer to the observed ones than the celerities obtained with the other sets of parameters. The average flow wave celerity over the 11 calculations in RAPID is within 3% of that estimated with IDA flows. Unfortunately, these closer wave celerities also lead to a decrease in the quality of RMSE and Nash Efficiency. Therefore, model celerities closer to celerities estimated from observations can be obtained, but generally deteriorate other statistics of calculations. Again, this might be due to runoff being produced too slowly or too far upstream of each gage.

4.6 Potential improvement of spatial variability in RAPID parameters

In the work presented here, the parameter x is spatially and temporally constant over the modeling domain and the parameter k is temporally constant but varies at the river reach level based on the length of each reach and on the celerity of the flow wave going through it. Flow wave celerities are estimated for 11 sub-basins based on flow observations and the spatial variability of k presented in this study is therefore partly limited by the size of the sub-basins used for flow wave estimation. However such an approach for computation of RAPID parameters allows taking into account wave celerities that are estimated based on observations made at high temporal resolution as well as verifying the modeling framework through reproduction of estimated wave celerities. In a separate study applying RAPID to all rivers of Metropolitan France, David et al. [2011] present a physically-based formulation of k and a sub-basin optimization for both k and x, therefore allowing further spatial variability of parameters. David et al. [2011] show that using a combination of reach length, river bed slope and basin residence time for the parameter k and applying the optimization procedure to sub-basins both improve the efficiency and the RMSE of RAPID flow computations. Such work could be adapted to the study herein based on information provided in the NHDPlus dataset – for example reach length, mean annual flow velocity and river bed slope – which would be advantageous when applying RAPID to domains larger than the Guadalupe and San Antonio River Basins where estimation of wave celerities everywhere may require excessive amounts of computations.

4.7 Statistical Significance

Changes in the routing procedure – i.e. no routing or routing using various RAPID parameters – lead to various changes in the values of efficiency and RMSE, as shown in Section 4.2. The statistical significance of the changes can be assessed in order to determine whether or not various routing experiments are effective. For two different routing procedures used, the efficiency (respectively RMSE) at one gage can be compared to the efficiency (respectively RMSE) at the same gage, although variability of efficiency (respectively RMSE) between independent gages can be large. Therefore, there is a logical pairing of efficiency and RMSE calculated at a given gage between two experiments and hence matched pair tests are appropriate to assess the statistical significance. Several common options are available for matched pair tests (with increasing level of complexity): the sign test, the Wilcoxon signed-ranked test [*Wilcoxon*, 1945] and the paired t-test. The sign test has no assumption on the shape of probability distributions of samples used but is quite simple since only the sign of differences between two paired samples is accounted for. The Wilcoxon signed-ranks test incorporates the magnitude of differences between paired samples under the assumption that differences between pairs are symmetrically distributed. The paired t-test may be used when the differences between pairs are known to be normally distributed. The assumption of the Wilcoxon signed-ranks test (symmetry) is not as restrictive as that of the paired t-test (normality). In case where small sample sizes are used – as done in this study – testing for symmetry or normality may not be meaningful. Additionally, violations of the symmetry assumption in the Wilcoxon signed-ranks test have minimal influence on the corresponding p-values [*Helsel and Hirsch*, 2002]. These two reasons motivate the use of the Wilcoxon signed-ranks test in the study herein. The null hypothesis H_0 for this test is that the median of differences between two populations is zero. The purpose of changes in the routing procedure being to improve results by increasing the efficiency and decreasing the RMSE, alternate hypotheses can assume that one population tends to be generally either larger (H_1) or smaller (H_2) than the other.

Therefore, p-values corresponding to one-sided tests are used in this study. Low significance levels mean that H_0 is unlikely, hence that a significant change is observed. The Wilcoxon signed-ranks test sorts pairs with nonzero difference based on the absolute value of the differences and sums all positive (respectively negative) ranks in a variable named W^+ (respectively W^-). The corresponding p-values vary with the number of nonzero differences and with the value of W^+ and W^-. Fortran programs were created to compute the exact value of the test statistic (not using a large-sample approximation) as well as the corresponding p-values. Table 4 shows the results of the Wilcoxon signed-ranks test for both efficiency and RMSE and for several paired experiments using two different routing procedures. The same 15 stations named on Figure 7 and used in Table 3 serve here for statistical significance assessment and the corresponding 15 values of efficiency and of RMSE are utilized as sample values.

Several conclusions can be drawn from Table 4. First, the Wilcoxon signed-ranks tests comparing results obtained by RAPID with parameters α, β and γ to a lumped runoff approach show that the null hypothesis can be rejected for a one-sided test at a 10% level of significance in all cases, except for the efficiency between RAPID with β parameters and a lumped approach at a 13% level of significance. All these tests validate that the improvements mentioned in Section 4.2 (increased efficiency and decreased RMSE) are statistically significant and confirm that an explicit river routing scheme allows obtaining better stream flow calculations than a simple lumped runoff scheme, as expected.

Second, comparisons between RAPID using α and γ parameters show that sub-basin variability in wave celerities is advantageous to a spatially uniform wave celerity approach at a 19% level of significance for efficiency and at a 7% level for RMSE. Third, comparisons between RAPID using γ and δ parameters confirms that wave celerities close to those determined from observations deteriorate results at a 3% level of significance for both efficiency and RMSE. Finally, one cannot conclude on the statistical significance of the comparison between RAPID using β and γ parameters concerning the improvement of optimization procedure. However, since RAPID using γ parameters produce better average values than RAPID using β parameters and since the statistical significance of RAPID using γ parameters compared to a lumped approach is better than that of RAPID using β parameters compared to lumped approach, the optimization can still considered advantageous.

5. Synthetic study of the Upper Mississippi River Basin, speedup of parallel computations

Through the use of mathematical and optimization libraries that run in a parallel computing environment, RAPID can be applied on several processing cores. The work presented above focuses on the Guadalupe and San Antonio River basins together forming a river network with 5,175 river and water body reaches, which size do not justify the use of parallel computing. However, all the tools and datasets used are available for the Contiguous United States where the NHDPlus dataset has about 3 million reaches. Adapting the proposed framework to simultaneously compute flow and volume of water in all mapped water bodies of the contiguous United States would require solving matrix equations of size 3 million. For such a large scientific problem, parallel computing can be helpful if speedup can be achieved, i.e. if increasing the number of processing cores decreases the total computing time.

5.1 Synthetic study used for assessment of parallel performance

As a proof of concept, the evaluation of the parallel computing capabilities of RAPID is presented here using the Upper Mississippi River Basin (shown on Figure 3) which has 182,240 river and water body reaches available as Region 07 in the NHDPlus dataset. The number of computational elements for the Upper Mississippi River Basin is about 35 times larger than the combination of the Guadalupe and San Antonio River Basins, and about 16 times smaller than the entire Contiguous United States. The river network of the Upper Mississippi River Basin is fully interconnected, all water eventually flowing to a unique outlet.

In order to assess the performance of RAPID, the same problem consisting in the computation of river flow in all reaches of the Upper Mississippi River Basin, over 100 days, at a 900-second time step is solved for all results reported in Section 5.3. For this performance study, the runoff data symbolized by vector $\mathbf{Q^e}$ in Equation (1) are synthetically generated and set to $1\ \mathrm{m}^3$ every 3 hours for all reaches and all time steps and the vectors of parameters $\mathbf{k}$ and $\mathbf{x}$ are temporally and spatially uniform as shown in Equation (22):

$$k_j = \frac{L_j}{2.5 m \cdot s^{-1}} \quad , \quad x_j = 0.3 \qquad (22)$$

5.2 Basics of solving a linear system on computers

Numerically solving a linear system is typically an iterative process mainly involving two-steps at each iteration: preconditioning followed by applying a linear solver. Preconditioning is a procedure that transforms a given linear system through matrix multiplication into one that is more easily solved by linear solvers, hence decreasing the total number of iterations to find the solution and saving time. If the linear system is triangular, preconditioning is sufficient to solve the problem, and a linear solver is not needed. In a parallel computing environment, a matrix is separated into diagonal and off-diagonal blocks, each processing core being assigned one diagonal block and its adjacent off-diagonal block. Solving a linear system in parallel is made using blocks and parallel preconditioning is determined based on elements in the diagonal blocks. Preconditioning is sufficient to solve a given parallel linear system if the system is diagonal by blocks –

i.e. all off-diagonal blocks are empty – and if each diagonal block is triangular; in most other cases iterations of preconditioning and applying a linear solver are needed.

5.3 Parallel speedup of the synthetic study

For comparison purposes, the traditional Muskingum method was also implemented in RAPID in order to assess the performance of the matrix-based Muskingum method developed herein. Figure 11 shows a comparison of computing time between the traditional Muskingum method shown in Equation (4) and applied consecutively from upstream to downstream and the Matrix-based Muskingum method used in RAPID. Only one processor in used for all results in Figure 11 but the computation method differs. The matrix $\mathbf{I}-\mathbf{C_1}\cdot\mathbf{N}$ being triangular (see Appendix B), solving the linear system of Equation (1) can be limited to matrix preconditioning if using only one processing core. In a parallel computing environment, $\mathbf{I}-\mathbf{C_1}\cdot\mathbf{N}$ is separated in blocks, each diagonal block corresponding to a sub-basin. With several processing cores, matrix preconditioning would be sufficient to solve Equation (1) if $\mathbf{I}-\mathbf{C_1}\cdot\mathbf{N}$ could be made diagonal by blocks, each diagonal block being a triangular matrix. In a river network that is fully interconnected such as that of the Upper Mississippi River Basin $\mathbf{I}-\mathbf{C_1}\cdot\mathbf{N}$ cannot be made diagonal by blocks because the connectivity between adjacent sub-basins would always appear as an element in an off-diagonal block matrix (cf. Equation (23) when i and j are connected but belong to different sub-basins). This limitation would not apply if one was to compute the Mississippi River basin on one (or on one set of) processing core(s) and the Colorado River Basin on another (or on another set of) processing core(s) for example. Therefore, when solving Equation (1) on several processing cores for the Upper Mississippi River Basin, preconditioning is not sufficient and iterative methods need be used. An iterative method implies several computations including preconditioning, matrix-vector multiplication and calculation of residual norm at each iteration.

On one processing core, solving the matrix-based Muskingum method with preconditioning only is about twice as long as solving the traditional Muskingum method, as shown in Figure 11. This extra time can be explained because the computation of the right-hand-side of Equation (1) is approximately as expensive as solving the traditional Muskingum method and approximately as expensive as preconditioning. However, the computation of the right-hand-side is done only once per time step regardless of the number of iterations if using an iterative linear solver and scales very well because all operations require no communication except for the product $\mathbf{N} \cdot \mathbf{Q}$ which involves little communication. Figure 11 also shows the computing time when using an iterative solver. The sole purpose of the first iteration in an iterative solver is to determine an initial residual error that is to be used as a criterion for convergence in following iterations. This first iteration mainly involves preconditioning and calculation of a residual norm. On one processing core only, the second iteration converges because preconditioning is sufficient. The two iterations and calculations of norms explain the doubling of computing time between preconditioning only and an iterative solver on one unique processing core that is shown in Figure 11. Overall, the overhead created by an iterative solver over the traditional Muskingum method is about a factor of four. Again, both preconditioning and calculation of residual norms scale well although the latter can be limited by communications. Therefore, the main issue with using a matrix method is the number of iterations needed before the iterative solver converges because all other overhead dissipates with an increasing number of processing cores used. Surprisingly, the number of iterations needed for the iterative solver to converge increases much less quickly than the number of processing cores used, hence allowing to gain total computation time with increased number of processing cores and to produce results faster than the traditional Muskingum method as shown on Figure 12. This suggests that even in a basin where all river reaches are interdependent, some upstream and downstream sub-basins can be computed separately in an iterative scheme given that they are distant enough from each other. The physical explanation is that flow waves are not fast enough to travel across the entire basin within one 15-minute time step. This de-coupling of computations could not be achieved by using the traditional version of the Muskingum method, since computations are not iterative and have to be performed going from upstream to downstream. Figure 12 shows that the total computing time with an iterative matrix solver on 16 processing cores is almost a third of the time needed by the traditional Muskingum method and keeps decreasing further with more processing cores. However, as the number of cores increase, the relative importance of the computation of residual norms within the iterative solver increases up to taking almost half of the solving time, as shown in Figure 12. This limitation will most likely disappear as computer technology advances and communication time decreases. One should note that the output files match on a byte-to-byte basis and hence model computations are strictly the same regardless of the method used; i.e. traditional Muskingum method or Matrix-based Muskingum method, iterative or not. This strict similarity between output files and the slow increase in iterations are also verified for the study of the Guadalupe and San

Antonio River Basins presented above; hence the use of synthetic data and simplified model parameters does not influence the trends in speedup.

Computing loads are balanced for all simulations in this study, i.e. the number of river reaches assigned to each processing core is almost identical across cores. Figure 13 shows how sub-basins of the Upper Mississippi River Basin are divided among processing cores as well as the longest river path of the basin. The longest path goes through 8 sub-basins on 8 cores, and 14 sub-basins on 16 cores. If one were to apply the traditional Muskingum method on several processing cores with the division in sub-basins shown in Figure 13, computations would have to be made sequentially from upstream to downstream, each core having to wait for its upstream core to be done prior to starting its work. Hence, assuming that the total computing time can be evenly divided by the total number of nodes and neglecting communication overhead, one could only hope to decrease computing time by a factor of $8/8=1$ (no gain) for 8 cores and by a factor of $16/14=1.14$ for 16 cores. The iterative matrix solver provides much better results (a decrease by a factor of 2.90 for 16 cores).

River flow is a causal phenomenon that mainly goes downstream. Therefore, when using an upstream-to-downstream computation scheme and unless dealing with completely separated river basins, one cannot expect to obtain perfect speedup i.e. decreasing of computing time by a factor equal to the number of cores. However, today's supercomputers having tens of thousands of computing cores, one could leverage such power to save human time. Additionally, the matrix method developed here can be directly applied to a combination of independent river basins in which case speedup would be ideally perfect. Furthermore, matrix methods such as the one developed here could be adapted to more complex river flow equations – like variable-parameter Muskingum methods or schemes allowing for backwater effects – in order to save total computing time. Finally, the splitting up into sub-basins used here is very simple and optimizing this partition by limiting connections between sub-basins or taking into account flow wave celerities relatively to basin sizes could respectively help limit the number of communications and the number of iterations in the linear system solver.

Conclusions

NHDPlus is a GIS dataset that describes the networks of mapped rivers and water bodies of the United States. One of the main advantages of NHDPlus is that connectivity information for the river networks is available. Therefore, this dataset offers possibilities for the development of river routing models that simultaneously calculate flow and volume of water in all water bodies of the nation. Furthermore, the USGS National Water Information System has thousand of gages located on the NHDPlus network which can be used to assess the quality of such river models across river basins (not only at basin outlets). The research presented in this paper investigates how to develop a river network model using NHDPlus networks and how to assess model computations and optimize model parameters with USGS stream flow measurements. All tools and datasets used herein are available for the contiguous United States, but this research addresses two smaller domains. The combination of the Guadalupe and San Antonio River Basins in Texas is used in a 4-year case study, and the Upper Mississippi River Basin is used in a speedup study with synthetic data. Graph theory is applied to a river network to create a network matrix that is used to develop a vector-matrix version of the Muskingum method and applied in a new river network model called RAPID. It has been shown that a GIS-based hydrographic dataset can be used as the river network for a river model to compute flow in large networks of thousands of reaches, including ungaged locations. A simple flux coupler for connecting a land surface model with an NHDPlus river network is presented. No horizontal routing of flow from the land surface to the river network is used in this study, and such an addition would help improve model calculations. An inverse method is developed to estimate model parameters in RAPID using available gage measurements located across the river basins. Wave celerities are estimated in several locations of the basin studied. RMSE and Nash efficiency of computed flow rates in four RAPID simulations are compared with a basic lumped model where runoff is directly accumulated at the gage, with gage measurements and among themselves. RAPID produces better RMSE and Nash efficiency than the lumped model and the improvements are statistically significant. Although the quality of RAPID calculations is tied to the quantity of runoff generated by the land surface model that provides runoff, mass is conserved within RAPID since the average flow rate is conserved. Spatial variability of parameters enhances the RMSE and Nash efficiency of RAPID calculations. Wave celerities are reproduced within a few percents with the model proposed, although wave celerities closer to those estimated from gage data generally deteriorate the other statistics of calculations. This deterioration might be due to runoff being produced too slowly or too far upstream of each gage. The parameters used in this study are simple, but could be improved based on information available in NHDPlus such as slope, mean flow and velocity of all reaches or by using modified versions of the Muskingum method with time-variable parameters although the latter would necessitate modification of the optimization procedure developed herein. The matrix formulation in RAPID can be transferred in a parallel computing environment. A synthetic study of the Upper Mississippi River Basin shows that although a large initial overhead is added by the matrix method, this overhead decreases with increasing number of processing cores. More importantly, an iterative matrix solver allows de-coupling of sub-basins – even if the main river basin is fully interconnected – hence permitting computation of sub-basins separately if they are distant enough from each other. As consequence, while producing the exact same results as the traditional Muskingum method, the matrix-based Muskingum method decreases the total computing time when run on several processing cores. Such a gain in computing time would be highly beneficial if addressing larger scales, like the entire Contiguous United States which would represent a square matrix of size 3 million.

Acknowledgements

This work was partially supported by the U.S. National Aeronautics and Space Administration under the Interdisciplinary Science Project NNX07AL79G, by the U.S. National Science Foundation under project EAR-0413265: CUAHSI Hydrologic Information Systems, by Ecole des Mines de Paris, France, and by the American Geophysical Union under a Horton (Hydrology) Research Grant. The authors wish to thank the PETSc and TAO developers, especially Dr. Barry Smith, Dr. Matthew Knepley, Dr. Satish Balay and Dr. Jason Sarich for their continuous assistance throughout the development of RAPID. Thank you to Dr. Karl Shultz from TACC for his help regarding the handling of inputs and outputs on supercomputers. The computing resources were provided by TACC, which is gratefully acknowledged. Thank you to Dr. Stefano Orlandini for suggestions on an early version of this work. The authors are thankful to the two anonymous reviewers and to the editor for their valuable comments and suggestions that helped improved the original version of this manuscript.

Appendix A – Implementation of RAPID

The river network routing model is coded in Fortran 90 using the Portable, Extensible Toolkit for Scientific Computation (PETSc) mathematical library [*Balay, et al.*, 1997; *Balay, et al.*, 2008; *Balay, et al.*, 2009] and the Toolkit for Advanced Optimization (TAO) optimization library [*McInnes, et al.*, 2009]. PETSc can be used to create matrices and vectors and to apply a variety of linear operations such as matrix-vector multiplications or linear system solving. TAO offers multiple methods for unconstrained and constrained optimization. Both PETSc and TAO are built upon the Message Passing Interface [*Dongarra, et al.*, 1994] – a standard for communications between processing cores – and can seamlessly be run in a sequential or a parallel computing environment. In this study, sparse matrices are stored using the sequential AIJ format when using one processing core and the MPIAIJ format when using several cores. Linear systems are solved within PETSc either by preconditioning only or with preconditioning associated to a Richardson method. The preconditioning methods used herein are ILU on one processing core, and bloc Jacobi on several cores. The optimization method used in TAO is a line search algorithm called the Nelder-Mead method. The netCDF file format [*Rew and Davis*, 1990] is utilized for both inputs and outputs. RAPID is run on single- and multiple-processor workstations as well as on Lonestar (http://www.tacc.utexas.edu/resources/hpcsystems/#lonestar), a supercomputer running at the Texas Advanced Computing Center (TACC). This Dell Linux Cluster has 1,460 nodes, each node with 8 GB of memory and with two dual-core sockets. Lonestar has a total of 5,840 computing cores.

Appendix B – NHDPlus used in RAPID

NHDPlus [*USEPA and USGS*, 2007] is a geographic information system (GIS) dataset for the hydrography of the United States. This dataset provides the mapped streams and rivers as well as the catchments that surround them. NHDPlus is based on the medium resolution 1:100,000 scale national hydrographic dataset (NHD). One of the main improvements in NHDPlus is the network connectivity available in the value added attributes (VAA) table for the river network. Each NHDPlus reach in the national network is assigned a unique integer identifier called COMID. NHDPlus catchments also have a COMID, the same COMID being used for the reach and its local contributing catchment. Nodes are located at the two ends of each NHDPlus river reach. A unique integer identifier is given to all nodes in the national river reach network. The VAA table includes $FromNode$ and $ToNode$ fields that give which node is upstream and which is downstream of a given reach. Two reaches that are connected in a river network share a node, and the reach j flows into the reach i if $ToNode(j) = FromNode(i)$. The NHDPlus connectivity between reaches, catchments and nodes is illustrated for three catchments of the Guadalupe and San Antonio basins in Figure 14.

In its current formulation, RAPID can handle several upstream reaches but only one unique downstream reach. However, divergences exist in mapped river networks, as they do in NHDPlus. The VAA table offers a $Divergence$ field to each of the river reaches (with values of 0 – not part of a divergence, 1 – main path of a divergence, 2 – minor path of a divergence). In the current formulation of RAPID, the main part of a divergence carries all the upstream flow. The $FromNode$, $ToNode$ and $Divergence$ fields are used to populate the network matrix given in Equation (5), by means of the following logical statement:

$$\forall (i,j) \in [1,m]^2,\ \textit{if}\ \left[FromNode(i) = ToNode(j) \right]\ \textit{and}\ \left[Divergence(j) \neq 2 \right] \Rightarrow N_{i,j} = 1 \quad (23)$$

where $N_{i,j}$ is the element of $\mathbf{N}$ located at row i and column j. Therefore, upstream to downstream connection is conserved if the downstream reach is the major branch of a divergence or if it is not part of a divergence at all, but the connection is not made for a minor branch of a divergence.

The VAA table also has information on the relative location – upstream or downstream – of NHDPlus reaches. This information is available in a field called *Hydroseq* consisting of a unique integer attributed to all NHDPlus reaches. Sorting the *Hydroseq* field in decreasing order prior to computations guarantees that all upstream elements are computed prior to solving the flow equations for any given river reach. This organization of computations allows the matrix $\mathbf{I} - \mathbf{C_1} \cdot \mathbf{N}$ of Equation (1) to be made lower triangular which increases the ease and speed of solving this linear system.

References

Apostolopoulos, T. K., and K. P. Georgakakos (1997), Parallel computation for streamflow prediction with distributed hydrologic models, *Journal of Hydrology*, *197*, 1-24.

Balay, S., W. D. Gropp, L. C. McInnes, and B. F. Smith (1997), Efficient Management of Parallelism in Object Oriented Numerical Software Libraries, in *Modern Software Tools in Scientific Computing*, edited by E. A. a. A. M. B. a. H. P. Langtangen, pp. 163-202.

Balay, S., K. Buschelman, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang (2008), PETSc Users Manual (Revision 3.0.0), 1-191 pp, Argonne National Laboratory.

Balay, S., K. Buschelman, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang (2009), PETSc Web page, Available online at http://www.mcs.anl.gov/petsc.

Berge, C. (1958), Matrice Associée d'un graphe, in *Théorie des Graphes et ses Applications*, edited, pp. 126-128, Dunod, Paris.

Cunge, J. A. (1969), On the subject of a flood propagation computation method (Muskingum method), *Journal of Hydraulic Research*, *7*, 205-230.

David, C. H., D. J. Gochis, D. R. Maidment, W. Yu, D. N. Yates, and Z.-L. Yang (2009), Using NHDPlus as the Land Base for the Noah-distributed Model, *Transactions in GIS*, *13*, 363-377.

David, C. H., F. Habets, D. R. Maidment, and Z.-L. Yang (2011), RAPID applied to the SIM-France model, *Hydrological Processes*, *Accepted Article, doi: 10.1002/hyp.8070*, n/a-n/a.

De Roo, A., B. Gouweleeuw, and J. Thielen (2003), Development of a European flood forecasting system, *International Journal of River Basin Management*, *1*, 49-59.

Dongarra, J., D. Walker, E. Lusk, B. Knighten, M. Snir, A. Geist, S. Otto, R. Hempel, E. Lusk, W. Gropp, J. Cownie, T. Skjellum, L. Clarke, R. Littlefield, M. Sears, S. Husslederman, E. Anderson, S. Berryman, J. Feeney, D. Frye, L. Hart, A. Ho, J. Kohl, P. Madams, C. Mosher, P. Pierce, E. Schikuta, R. G. Voigt, R. Babb, R. Bjornson, V. Fernando, I. Glendinning, T. Haupt, C. T. H. Ho, S. Krauss, A. Mainwaring, D. Nessett, S. Ranka, A. Singh, D. Weeks, J. Baron, N. Doss, S. Fineberg, A. Greenberg, D. Heller, G. Howell, B. Leary, O. McBryan, P. Pacheco, P. Rigsbee, A. Sussman, S. Wheat, E. Barszcz, A. Elster, J. Flower, R. Harrison, T. Henderson, J. Kapenga, A. Maccabe, P. McKinley, H. Palmer, A. Robison, R. Tomlinson, and S. Zenith (1994), Special Issue - MPI - a Message-Passing Interface Standard, *International Journal of Supercomputer Applications and High Performance Computing*, *8*, 159-416.

Fread, D. L. (1993), Flow Routing, in *Handbook of Hydrology*, edited by D. R. Maidment, pp. 10.17-10.18, McGraw-Hill, New York.

Gochis, D. J., and F. Chen (2003), Hydrological Enhancements to the Community Noah Land Surface Model, available online at http://www.ucar.edu/library/collections/technotes/technotes.jsp.

Guan, H., J. L. Wilson, and H. Xie (2009), A cluster-optimizing regression-based approach for precipitation spatial downscaling in mountainous terrain, *Journal of Hydrology*, *375*, 578-588.

Habets, F., J. Noilhan, C. Golaz, J. P. Goutorbe, P. Lacarrere, E. Martin, C. Ottle, and D. Vidal-Madjar (1999a), The ISBA surface scheme in a macroscale hydrological model applied to the Hapex-Mobilhy area - Part I: Model and database, *Journal of Hydrology*, *217*, 75-96.

Habets, F., J. Noilhan, C. Golaz, J. P. Goutorbe, P. Lacarrere, E. Leblois, E. Ledoux, E. Martin, C. Ottle, and D. Vidal-Madjar (1999b), The ISBA surface scheme in a macroscale hydrological model applied to the Hapex-Mobilhy area - Part II: Simulation of streamflows and annual water budget, *Journal of Hydrology*, *217*, 97-118.

Habets, F., P. Etchevers, C. Golaz, E. Leblois, E. Ledoux, E. Martin, J. Noilhan, and C. Ottle (1999c), Simulation of the water budget and the river flows of the Rhone basin, *Journal of Geophysical Research-Atmospheres*, *104*, 31145-31172.

Habets, F., A. Boone, J. L. Champeaux, P. Etchevers, L. Franchisteguy, E. Leblois, E. Ledoux, P. Le Moigne, E. Martin, S. Morel, J. Noilhan, P. Q. Segui, F. Rousset-Regimbeau, and P. Viennot (2008), The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France, *Journal of Geophysical Research-Atmospheres*, *113*.

Helsel, D. R., and R. M. Hirsch (2002), Chapter A3. Statistical Methods in Water Resources, in *Techniques of Water-Resources Investigations of the United States Geological Survey. Book 4, Hydrologic Analysis and Interpretation*, edited, pp. 137-156, United States Geological Survey.

Jones, J. E., and C. S. Woodward (2001), Newton-Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems, *Advances in Water Resources*, *24*, 763-774.

Kollet, S. J., and R. M. Maxwell (2006), Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model, *Advances in Water Resources*, *29*, 945-958.

Kollet, S. J., R. M. Maxwell, C. S. Woodward, S. Smith, J. Vanderborght, H. Vereecken, and C. Simmer (2010), Proof of concept of regional scale hydrologic simulations at hydrologic resolution utilizing massively parallel computer resources, *Water Resour. Res.*, *46*, W04201.

Koussis, A. D. (1978), Theoretical Estimations of Flood Routing Parameters, *J. Hydraul. Div. Am. Soc. Civ. Eng.*, *104*, 109-115.

Larson, J. W., A. P. Craig, J. B. Drake, D. J. I. Erickson, M. Branstetter, and M. W. Ham (2007), A Massively Parallel Dynamical Core for Continental- to Global-Scale River Transport, paper presented at Proceedings of the International Congress and Modelling and Simulation (ModSim 2007).

Ledoux, E., G. Girard, G. de Marsily, J. P. Villeneuve, and J. Deschenes (1989), Spatially Distributed Modeling: Conceptual Approach, Coupling Surface Water and Groundwater, in *Unsaturated Flow in Hydrologic Modeling Theory and Practice*, edited by H. J. Morel-Seytoux, pp. 435-454, Kluwer Academic Publishers.

Lehner, B., K. Verdin, and A. Jarvis (2006), HydroSHEDS Technical Documentation, available online at http://hydrosheds.cr.usgs.gov.

Leopold, C., M. Süß, and J. Breitbart (2006), Programming for Malleability with Hybrid MPI-2 and OpenMP: Experiences with a Simulation Program for Global Water Prognosis, paper presented at European Conference on Modelling and Simulation.

Lohmann, D., E. Raschke, B. Nijssen, and D. P. Lettenmaier (1998a), Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, *43*, 131-141.

Lohmann, D., E. Raschke, B. Nijssen, and D. P. Lettenmaier (1998b), Regional scale hydrology: II. Application of the VIC-2L model to the Weser River, Germany, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, *43*, 143-158.

Lohmann, D., K. E. Mitchell, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, B. A. Cosgrove, J. Sheffield, Q. Y. Duan, L. F. Luo, R. W. Higgins, R. T. Pinker, and J. D. Tarpley (2004), Streamflow and water balance intercomparisons of four land surface models in the North American Land Data Assimilation System project, *Journal of Geophysical Research-Atmospheres*, *109*, 1-22.

Maurer, E. P., G. M. O'Donnell, D. P. Lettenmaier, and J. O. Roads (2001), Evaluation of the land surface water budget in NCEP/NCAR and NCEP/DOE reanalyses using an off-line hydrologic model, *Journal of Geophysical Research-Atmospheres*, *106*, 17841-17862.

McCarthy, G. T. (1938), The Unit Hydrograph and Flood Routing, paper presented at Conference of the North Atlantic Division, U.S. Engineer Department, New London, Connecticut, on June 24, 1938.

McInnes, L. C., J. Moré, T. Munson, and J. Sarich (2009), TAO User Manual (Revision 1.10), 1-64 pp, Mathematics and Computer Science Division, Argonne National Laboratory, Available online at http://www.mcs.anl.gov/tao.

Miller, W. A., and J. A. Cunge (1975), Simplified Equations of Unsteady Flow, in *Unsteady Flow in Open Channels*, edited by K. Mahmood and V. Yevjevich, pp. 216-232, Water Resources Publications, Fort Collins, CO.

Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I -- A discussion of principles, *Journal of Hydrology*, *10*, 282-290.

NERC (1975), Flood Routing Studies, in *Flood Studies Report*, edited, National Environment Research Council, London.

Niu, G. Y., Z. L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007), Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, *Journal of Geophysical Research-Atmospheres*, *112*, 1-14.

Niu, G. Y., Z. L. Yang, K. E. Mitchell, F. Chen, M. Ek, M. Barlage, L. Longuevergne, A. Kumar, K. Manning, D. Niyogi, E. Rosero, M. Tewari, and Y.-L. Xia (2010), The Community Noah Land Surface Model with Multi-Physics Options, *Journal of Geophysical Research-Atmospheres*, (submitted).

Oki, T., Y. Agata, S. Kanae, T. Saruhashi, D. W. Yang, and K. Musiake (2001), Global assessment of current water resources using total runoff integrating pathways, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, *46*, 983-995.

Olivera, F., J. Famiglietti, and K. Asante (2000), Global-scale flow routing using a source-to-sink algorithm, *Water Resources Research*, *36*, 2197-2207.

Orlandini, S., and R. Rosso (1998), Parameterization of stream channel geometry in the distributed modeling of catchment dynamics, *Water Resources Research*, *34*, 1971-1985.

Orlandini, S., G. Moretti, M. Franchini, B. Aldighieri, and B. Testa (2003), Path-based methods for the determination of nondispersive drainage directions in grid-based digital elevation models, *Water Resources Research*, *39*.

Ponce, V. M., and V. Yevjevich (1978), Muskingum-Cunge Method with Variable Parameters, *J. Hydraul. Div. Am. Soc. Civ. Eng.*, *104*, 1663-1667.

Ponce, V. M. (1986), Diffusion Wave Modeling of Catchment Dynamics, *J. Hydraul. Eng.-ASCE*, *112*, 716-727.

Rew, R., and G. Davis (1990), Netcdf - an Interface for Scientific-Data Access, *Ieee Computer Graphics and Applications*, *10*, 76-82.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H. L. Miller (2007), *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)*, 996 pp., Cambridge University Press, Cambridge.

Todini, E. (2007), A mass conservative and water storage consistent variable parameter Muskingum-Cunge approach (vol 11, pg 1645, 2007), *Hydrology and Earth System Sciences*, *11*, 1783-1783.

USEPA, and USGS (2007), NHDPlus User Guide, available online at http://www.horizon-systems.com/nhdplus/documentation.php.

von Bloh, W., S. Rost, D. Gerten, and W. Lucht (2010), Efficient parallelization of a dynamic global vegetation model with river routing, *Environmental Modelling & Software*, *25*, 685-690.

Wilcoxon, F. (1945), Individual Comparisons by Ranking Methods, *Biometrics Bulletin*, *1*, 80-83.

Table 1 Travel time (s) for the flow waves estimated using the lagged cross-correlation in the Guadalupe and San Antonio River Basins, both from IDA measurements and from RAPID model runs; and distance (km) between gaging stations

		Location of the two consecutive streamflow gages										
		Ingram - Kerrville	Kerrville - Comfort	Comfort - Spring Branch	Sattler - Gonzales	Gonzales - Cuero	Cuero - Victoria	Schroeder - Victoria	Bandera - Macdona	Macdona - Elmendorf	Elmendorf - Falls City	Falls City - Goliad
Travel time (s) from IDA	2004	7200	18900	60300	162900	132300	70200	*0*	20700	*0*	126000	162000
	2007	6300	18900	59400	131400	108900	70200	8100	37800	15300	91800	126000
	average	6750	18900	59850	147150	120600	70200	4050	29250	7650	108900	144000
Travel time (s) from RAPID outputs	RAPID (k^{α},x^{α})	6300	8100	35100	90000	29700	38700	5400	50400	29700	22500	52200
	RAPID (k^{β},x^{β})	6300	8100	46800	128700	84600	36000	4500	24300	8100	91800	124200
	RAPID (k^{γ},x^{γ})	4500	6300	27900	88200	60300	31500	2700	15300	6300	58500	80100
	RAPID (k^{δ},x^{δ})	9000	9000	72900	174600	117900	75600	5400	37800	16200	140400	193500
Distance (km)		13.77	40.40	100.85	203.50	110.72	100.71	24.44	116.36	71.73	79.16	137.16

Table 2 Wave celerities (m/s) estimated using the lagged cross-correlation in the Guadalupe and San Antonio River Basins, both from IDA measurements and from RAPID model runs

Wave celerity (m/s)		Location of the two consecutive streamflow gages										
		Ingram - Kerrville	Kerrville - Comfort	Comfort - Spring Branch	Sattler - Gonzales	Gonzales - Cuero	Cuero - Victoria	Schroeder - Victoria	Bandera - Macdona	Macdona - Elmendorf	Elmendorf - Falls City	Falls City - Goliad
from IDA	2004	1.91	2.14	1.67	1.25	0.84	1.43	∞	5.62	∞	0.63	0.85
	2007	2.19	2.14	1.70	1.55	1.02	1.43	3.02	3.08	4.69	0.86	1.09
	average	2.05	2.14	1.69	1.40	0.93	1.43	3.02	4.35	4.69	0.75	0.97
from RAPID outputs	RAPID (k^{α},x^{α})	2.19	4.99	2.87	2.26	3.73	2.60	4.53	2.31	2.41	3.52	2.63
	RAPID (k^{β},x^{β})	2.19	4.99	2.16	1.58	1.31	2.80	5.43	4.79	8.85	0.86	1.10
	RAPID (k^{γ},x^{γ})	3.06	6.41	3.61	2.31	1.84	3.20	9.05	7.61	11.38	1.35	1.71
	RAPID (k^{δ},x^{δ})	1.53	4.49	1.38	1.17	0.94	1.33	4.53	3.08	4.43	0.56	0.71

Table 3 Comparison of observed and simulated flows at fifteen locations within the Guadalupe and San Antonio River Basins

	Average daily stream flow (m3/s)						Flow ratio	RMS error (m3/s) using daily averages					Nash efficiency using daily averages				
Gaging station	Observed	Lumped	RAPID (k^{α},x^{α})	RAPID (k^{β},x^{β})	RAPID (k^{γ},x^{γ})	RAPID (k^{δ},x^{δ})	Observed/Lumped	Lumped	RAPID (k^{α},x^{α})	RAPID (k^{β},x^{β})	RAPID (k^{γ},x^{γ})	RAPID (k^{δ},x^{δ})	Lumped	RAPID (k^{α},x^{α})	RAPID (k^{β},x^{β})	RAPID (k^{γ},x^{γ})	RAPID (k^{δ},x^{δ})
Johnson Ck nr Ingram, TX	1.16	0.06	0.06	0.06	0.06	0.06	19.33	4.41	4.41	4.41	4.41	4.41	-0.05	-0.05	-0.05	-0.05	-0.05
Guadalupe Rv at Kerrville, TX	4.15	0.14	0.14	0.14	0.14	0.14	29.64	15.04	15.04	15.04	15.04	15.04	-0.06	-0.05	-0.05	-0.05	-0.06
Guadalupe Rv at Comfort, TX	9.97	0.81	0.81	0.81	0.81	0.81	12.31	26.57	26.51	26.51	26.52	26.53	-0.06	-0.06	-0.06	-0.06	-0.06
Guadalupe Rv nr Spring Branch, TX	19.74	5.91	5.91	5.91	5.91	5.91	3.34	42.09	43.06	43.48	42.72	44.80	0.26	0.23	0.21	0.24	0.16
Guadalupe Rv at Sattler, TX	22.04	6.62	6.62	6.62	6.62	6.62	3.33	40.08	39.85	39.77	39.94	39.57	-0.06	-0.04	-0.04	-0.05	-0.03
Guadalupe Rv at Gonzales, TX	64.28	23.27	23.27	23.27	23.27	23.27	2.76	79.83	80.93	86.44	80.40	93.78	0.45	0.44	0.36	0.45	0.25
Guadalupe Rv at Cuero, TX	73.23	52.63	52.62	52.61	52.62	52.60	1.39	76.86	56.41	64.91	55.52	82.74	0.59	0.78	0.71	0.79	0.53
Guadalupe Rv nr Victoria	80.96	61.95	61.93	61.92	61.93	61.91	1.31	93.97	70.11	65.07	68.05	89.06	0.54	0.75	0.78	0.76	0.59
Coleto Ck at Arnold Rd nr Schroeder, TX	3.45	8.78	8.78	8.78	8.78	8.78	0.39	15.43	15.44	15.45	15.46	15.44	0.03	0.03	0.03	0.02	0.03
Coleto Ck nr Victoria, TX	3.99	13.72	13.72	13.72	13.72	13.72	0.29	21.82	22.61	22.46	22.26	22.65	0.10	0.03	0.05	0.06	0.03
Medina Rv at Banderas, TX	5.30	0.75	0.75	0.75	0.75	0.75	7.07	10.78	10.77	10.77	10.77	10.77	0.05	0.05	0.05	0.05	0.05
Medina Rv nr Macdona, TX	8.73	2.09	2.09	2.09	2.09	2.09	4.18	12.89	12.74	12.72	12.74	12.72	0.29	0.31	0.31	0.30	0.31
San Antonio Rv nr Elmendorf, TX	25.05	7.95	7.95	7.95	7.95	7.95	3.15	39.91	39.27	39.23	39.41	39.16	0.34	0.36	0.36	0.36	0.37
San Antonio Rv nr Falls City, TX	25.01	12.36	12.36	12.36	12.36	12.36	2.02	33.23	31.13	30.63	31.26	32.00	0.45	0.52	0.53	0.51	0.49
San Antonio Rv at Goliad, TX	37.54	34.96	34.95	34.95	34.95	34.94	1.07	42.34	37.73	34.58	36.92	39.10	0.56	0.65	0.71	0.67	0.63
Mean	25.64	15.47	15.46	15.46	15.46	15.46		37.02	33.73	34.10	33.43	37.85	0.23	0.26	0.26	0.27	0.22

Table 4 Results of the Wilcoxon signed-ranks test applied to fifteen stations for efficiency and RMSE and to various routing procedures

Efficiency

x	y	Number of non-zero differences	Total rank	W^+ (computed for y-x)	p-value corresponding to W^+	W^- (computed for y-x)	p-value corresponding to W^-
Lumped runoff	RAPID (k^α,x^α)	11	66	51.0	0.06152	15.0	0.94922
Lumped runoff	RAPID (k^β,x^β)	11	66	47.0	0.12012	19.0	0.89697
Lumped runoff	RAPID (k^γ,x^γ)	11	66	51.0	0.06152	15.0	0.94922
Lumped runoff	RAPID (k^δ,x^δ)	10	55	22.5	0.70459	32.5	0.33008
RAPID (k^α,x^α)	RAPID (k^γ,x^γ)	10	55	37.0	0.18750	18.0	0.83887
RAPID (k^β,x^β)	RAPID (k^γ,x^γ)	10	55	28.5	0.48047	26.5	0.55811
RAPID (k^γ,x^γ)	RAPID (k^δ,x^δ)	12	78	13.0	0.98291	65.0	0.02124

RMSE

x	y	Number of non-zero differences	Total rank	W^+ (computed for y-x)	p-value corresponding to W^+	W^- (computed for y-x)	p-value corresponding to W^-
Lumped runoff	RAPID (k^α,x^α)	13	91	25.5	0.92145	65.5	0.08966
Lumped runoff	RAPID (k^β,x^β)	13	91	26.0	0.91614	65.0	0.09546
Lumped runoff	RAPID (k^γ,x^γ)	13	91	25.0	0.92676	66.0	0.08386
Lumped runoff	RAPID (k^δ,x^δ)	13	91	42.5	0.59345	48.5	0.43299
RAPID (k^α,x^α)	RAPID (k^γ,x^γ)	11	66	15.0	0.94922	51.0	0.06152
RAPID (k^β,x^β)	RAPID (k^γ,x^γ)	12	78	41.0	0.45483	37.0	0.57495
RAPID (k^γ,x^γ)	RAPID (k^δ,x^δ)	12	78	64.0	0.02612	14.0	0.97876

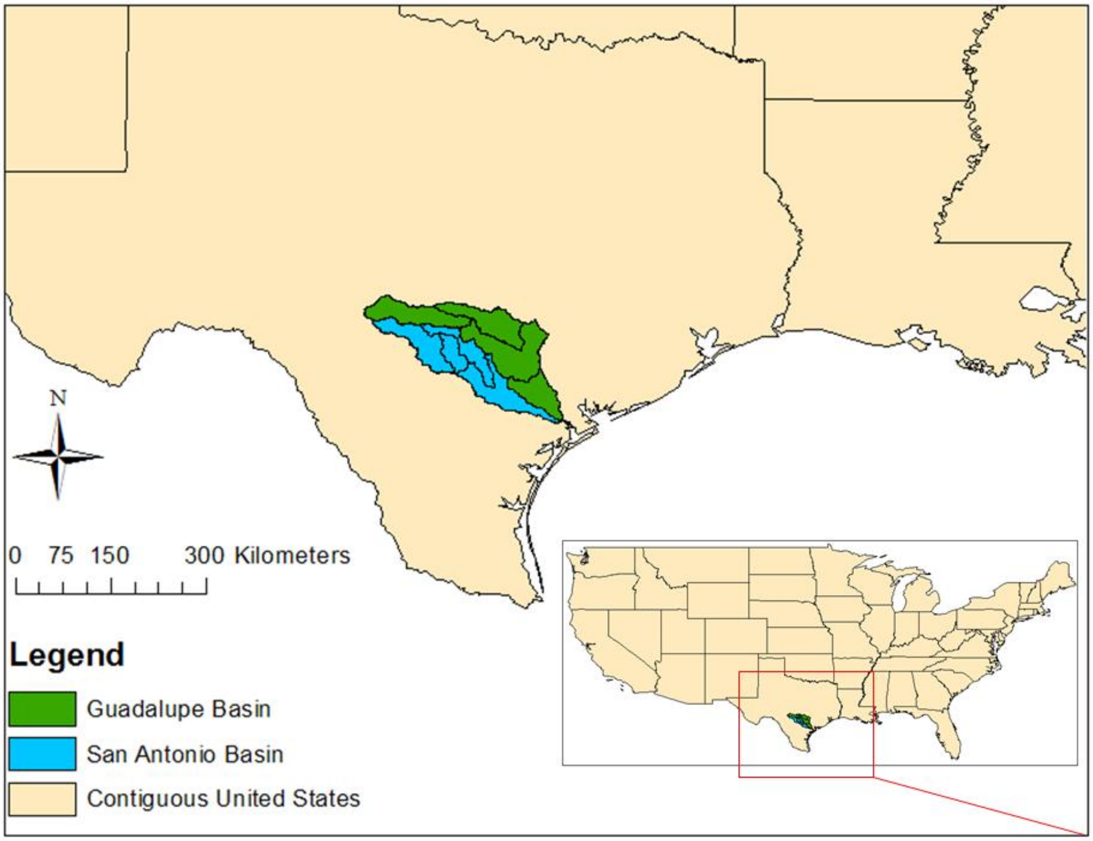

Figure 1 Guadalupe and San Antonio Basins

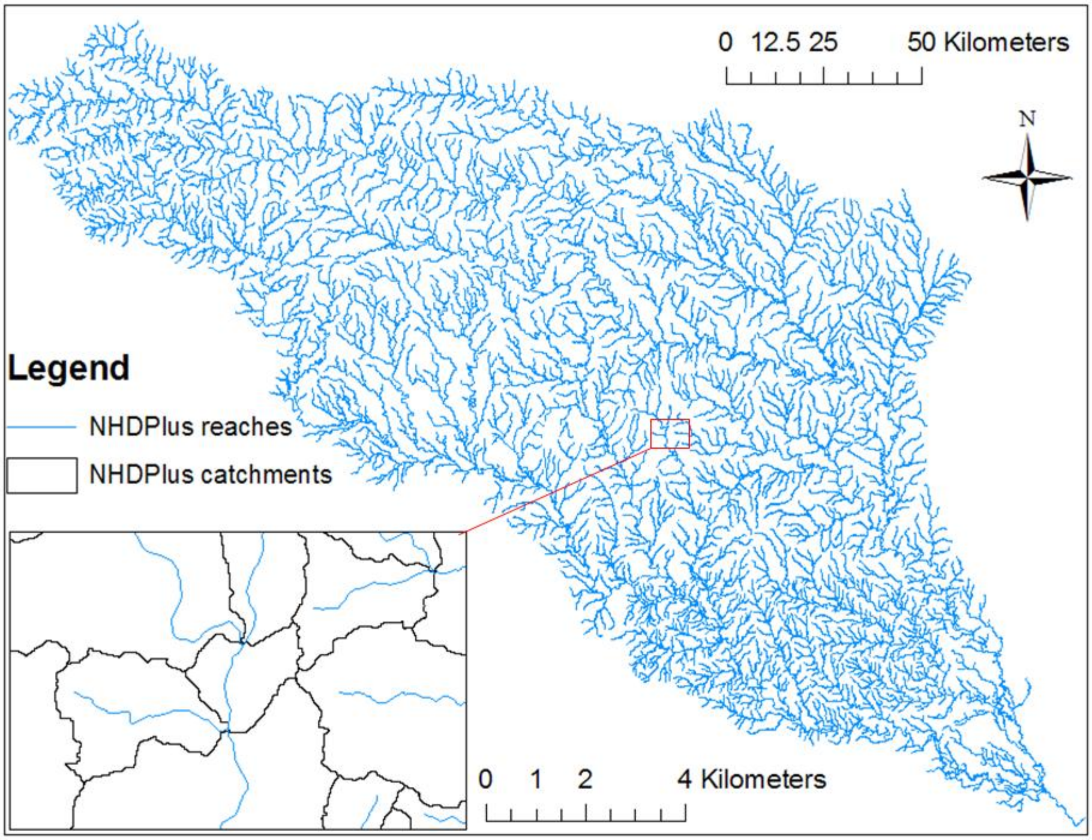

Figure 2 NHDPlus river network and catchments for the Guadalupe and San Antonio Basins

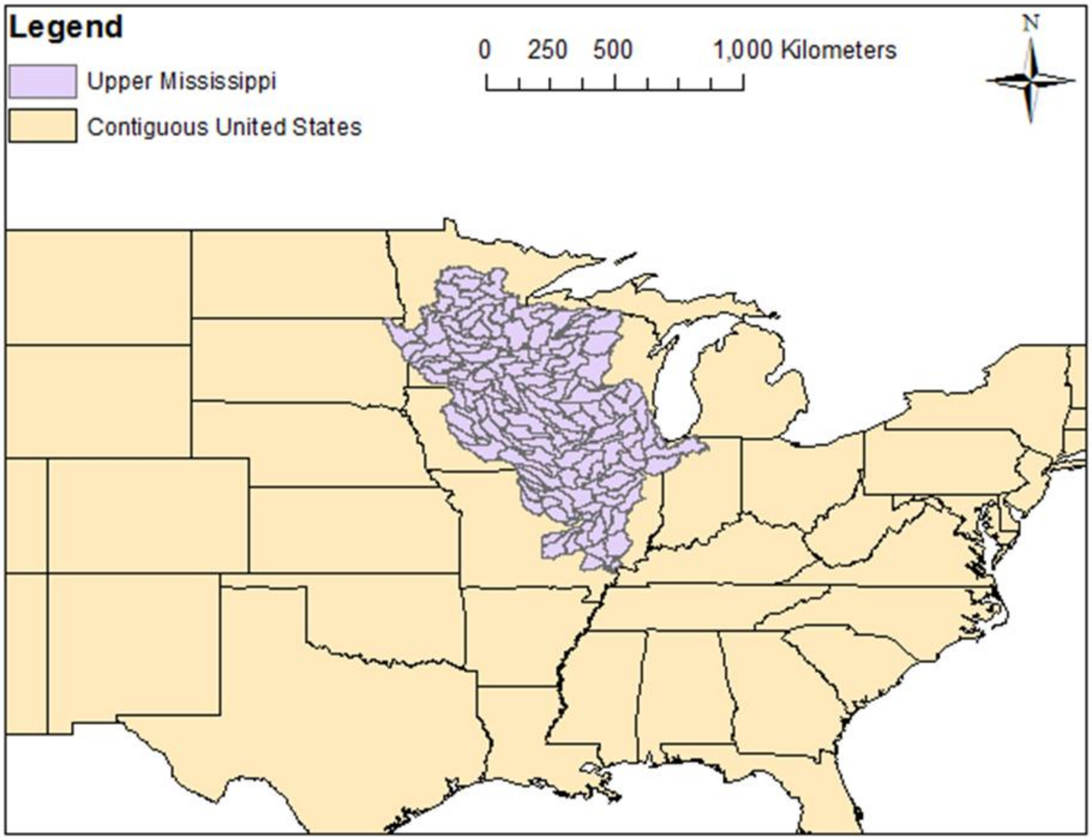

Figure 3 Upper Mississippi River Basin

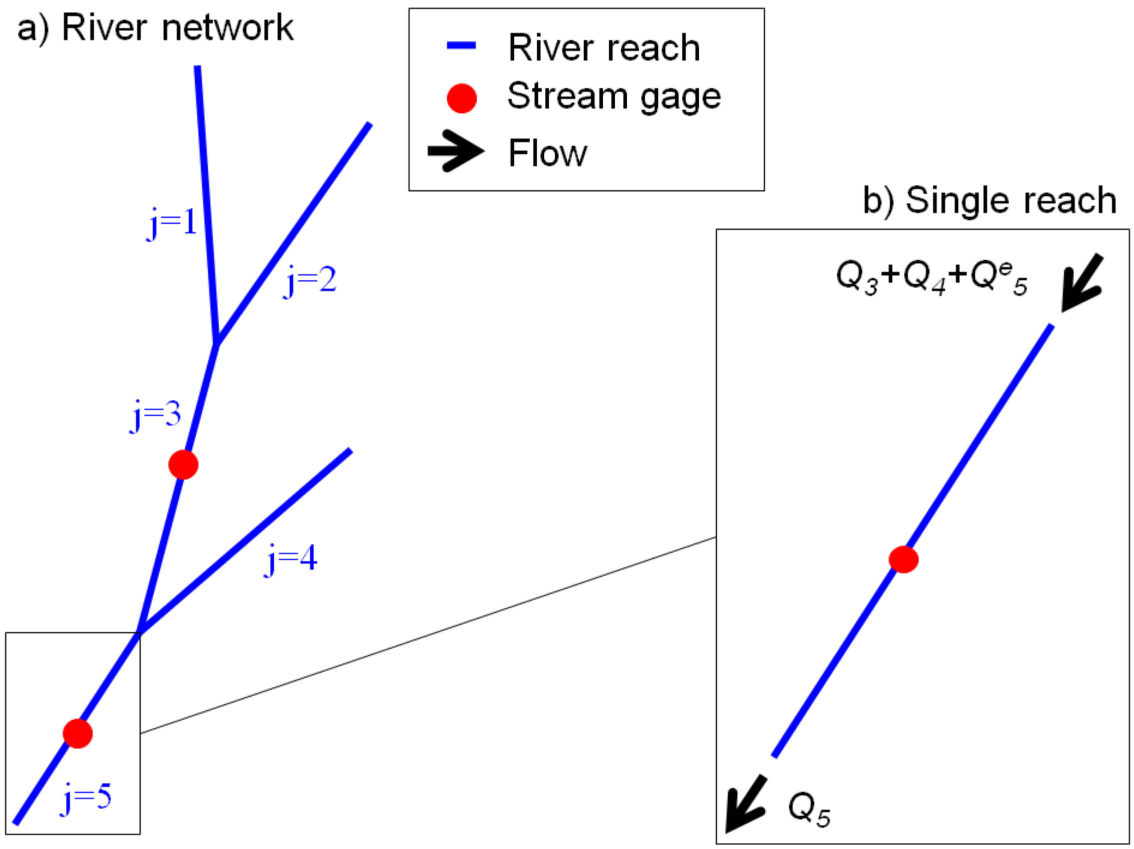

Figure 4 River network

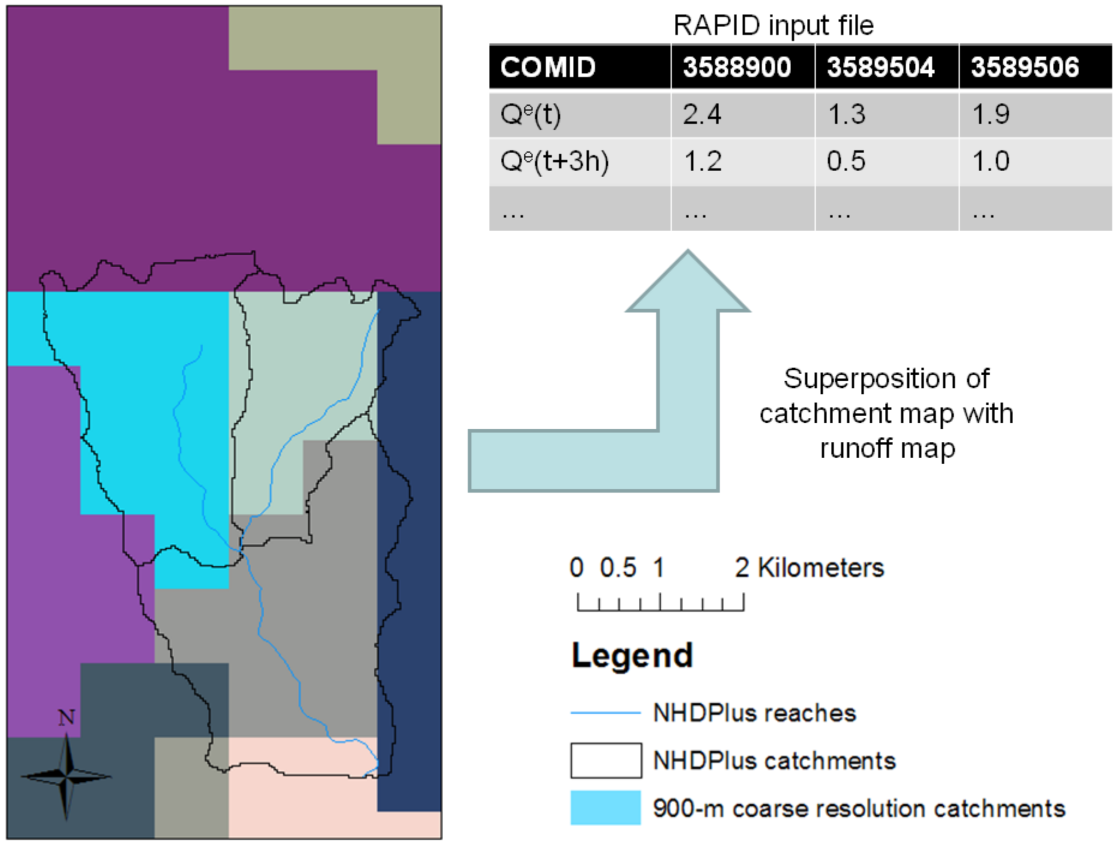

Figure 5 Principle of flux coupler between Noah and RAPID

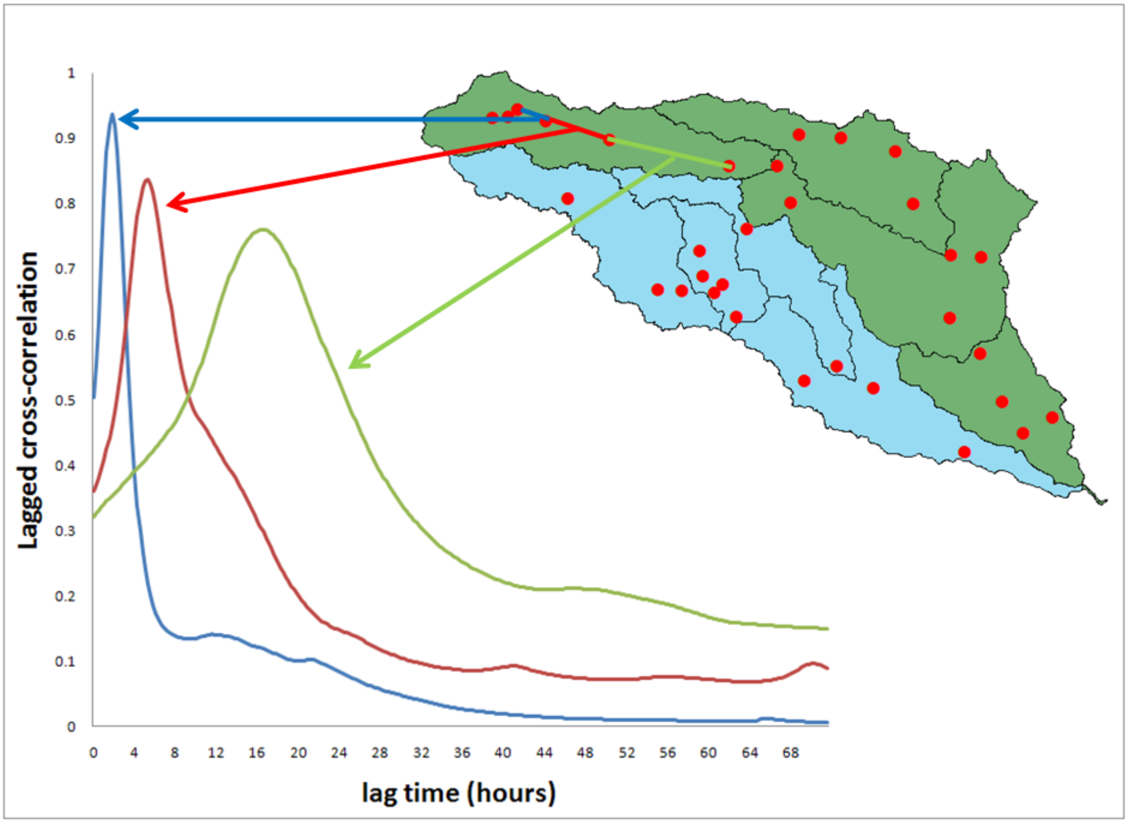

Figure 6 Lagged cross-correlation as a function of lag time

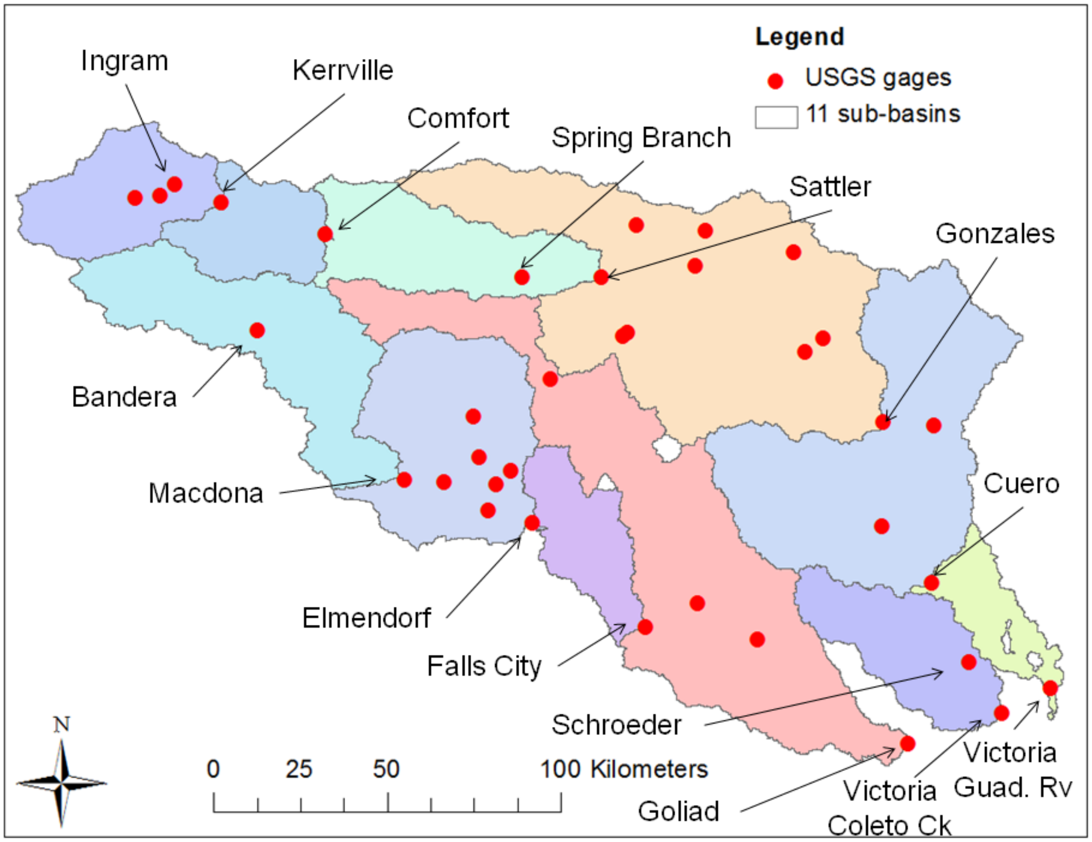

Figure 7 Wave celerities are estimated for eleven different sub-basins within the Guadalupe and San Antonio river basins. Location of 36 gaging stations used for optimization and names of the 15 gaging stations used for estimation of wave celerities. The same sub-basins are used for distributed parameters in RAPID

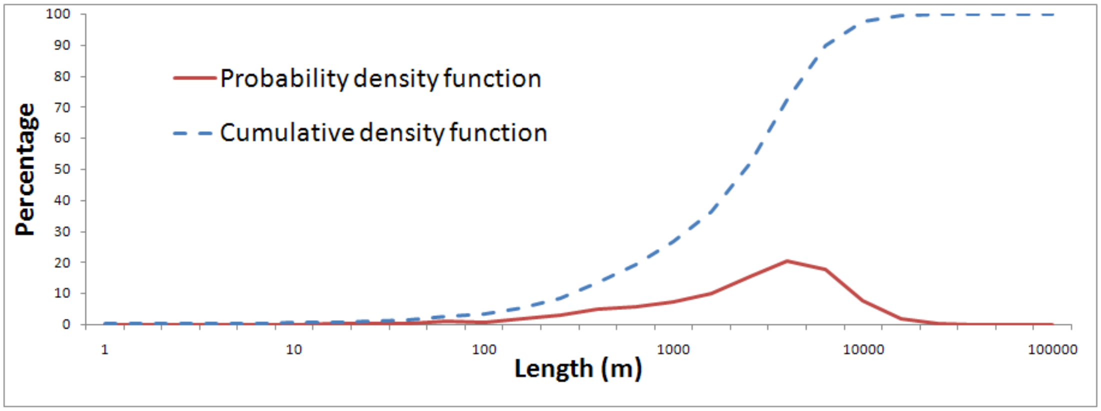

Figure 8 Statistics of river reach lengths in Guadalupe and San Antonio River Basins

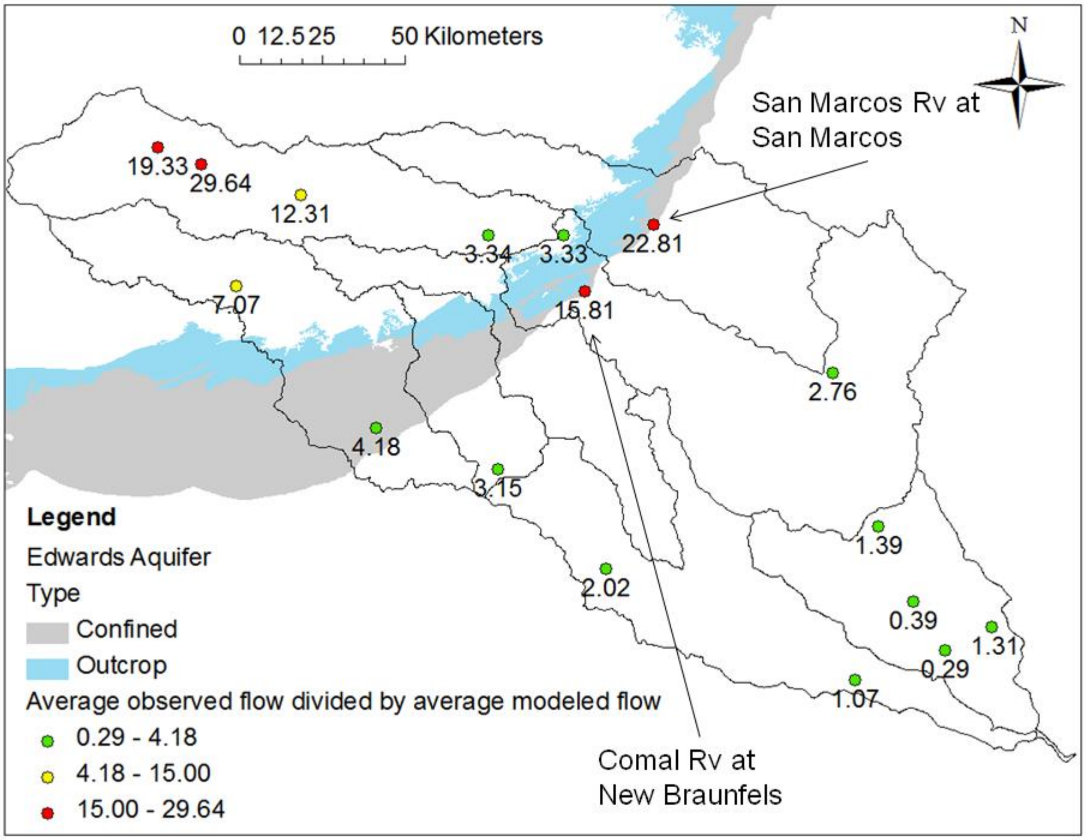

Figure 9 Ratio between observed and modeled stream flow at 16 gages, location of the Edwards Aquifer. Location of the two largest springs in Texas.

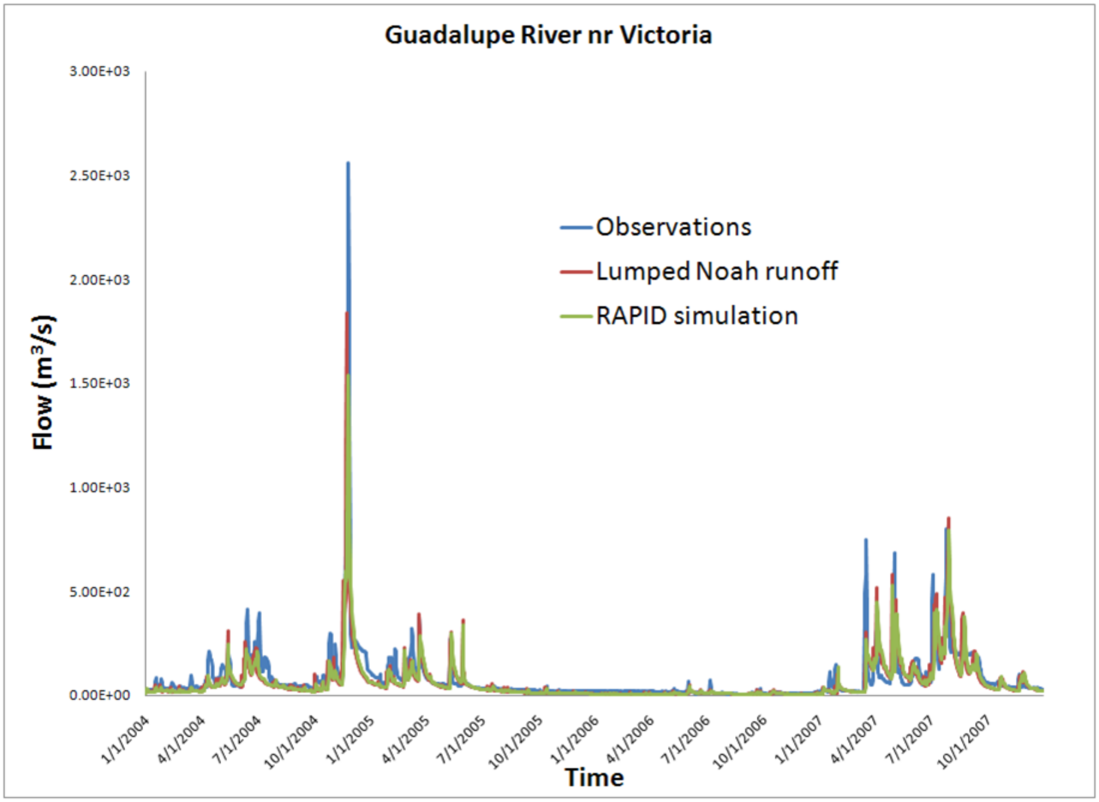

Figure 10 Hydrograph of observed, lumped and routed flows for the Guadalupe River near Victoria, using (k^γ,x^γ)

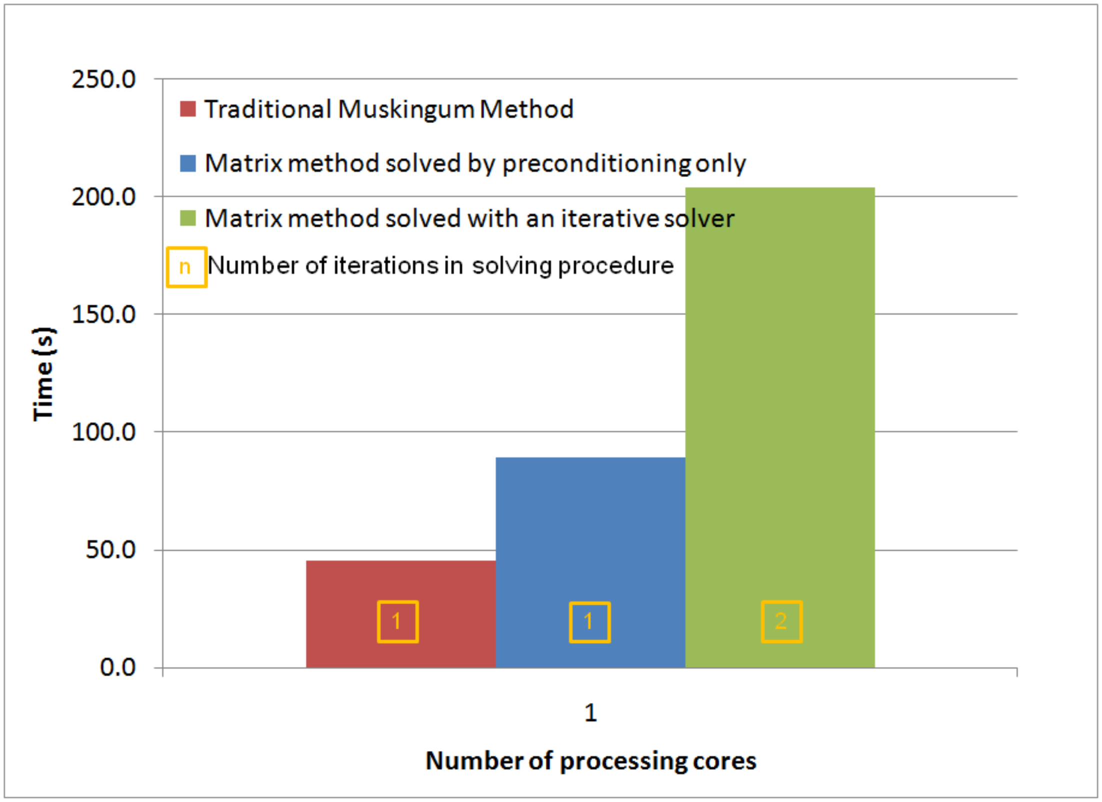

Figure 11 Comparison of computing time between the traditional Muskingum method and matrix methods

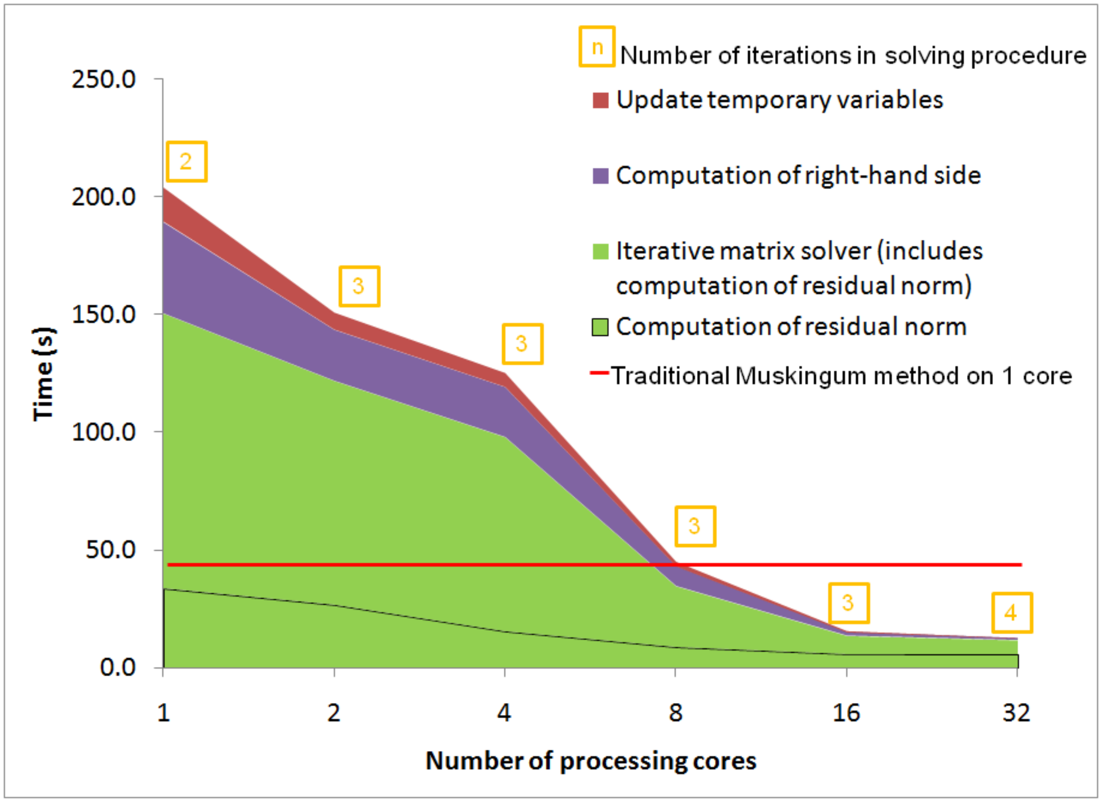

Figure 12 Total computing time for matrix method with an iterative solver as a function of the number of processing cores, number of iterations needed, total computing time for the traditional Muskingum method.

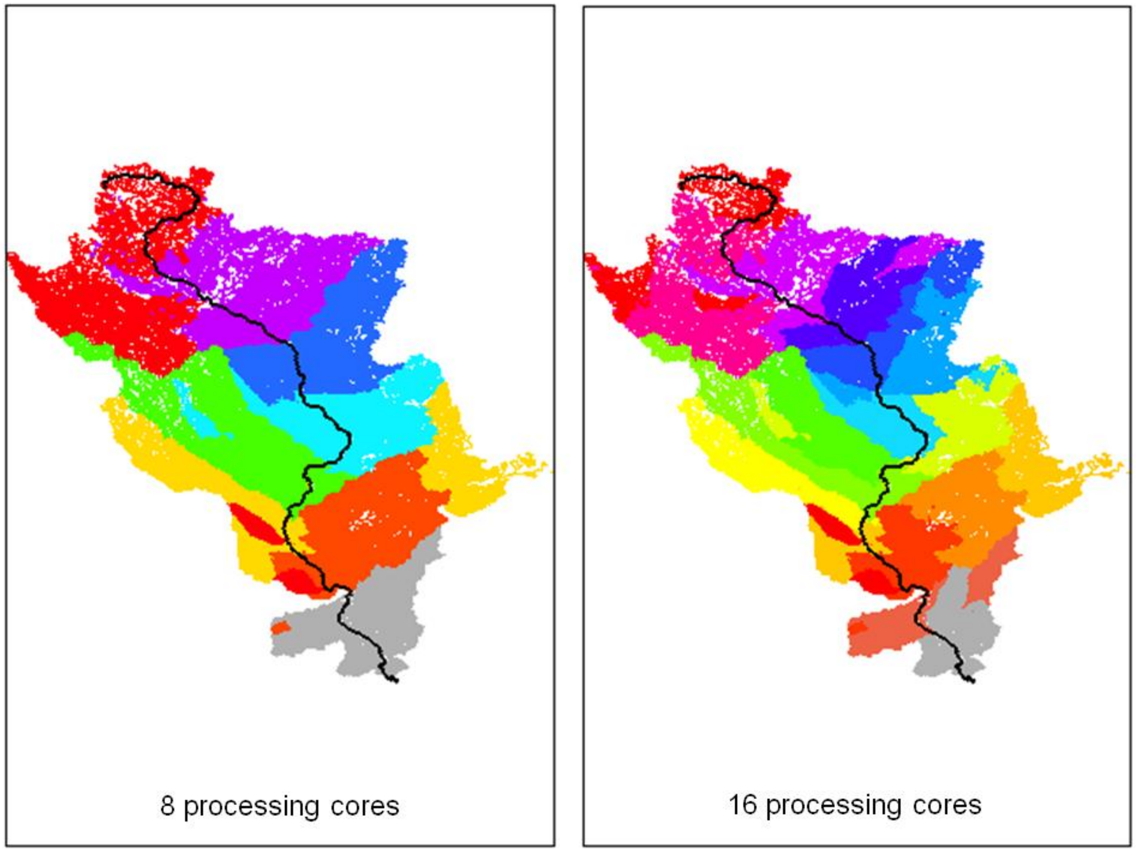

Figure 13 Longest path in the Upper Mississippi River Basin and location of sub-basins when RAPID is used in a parallel computing environment with 8 and 16 processing cores, different colors correspond to different cores.

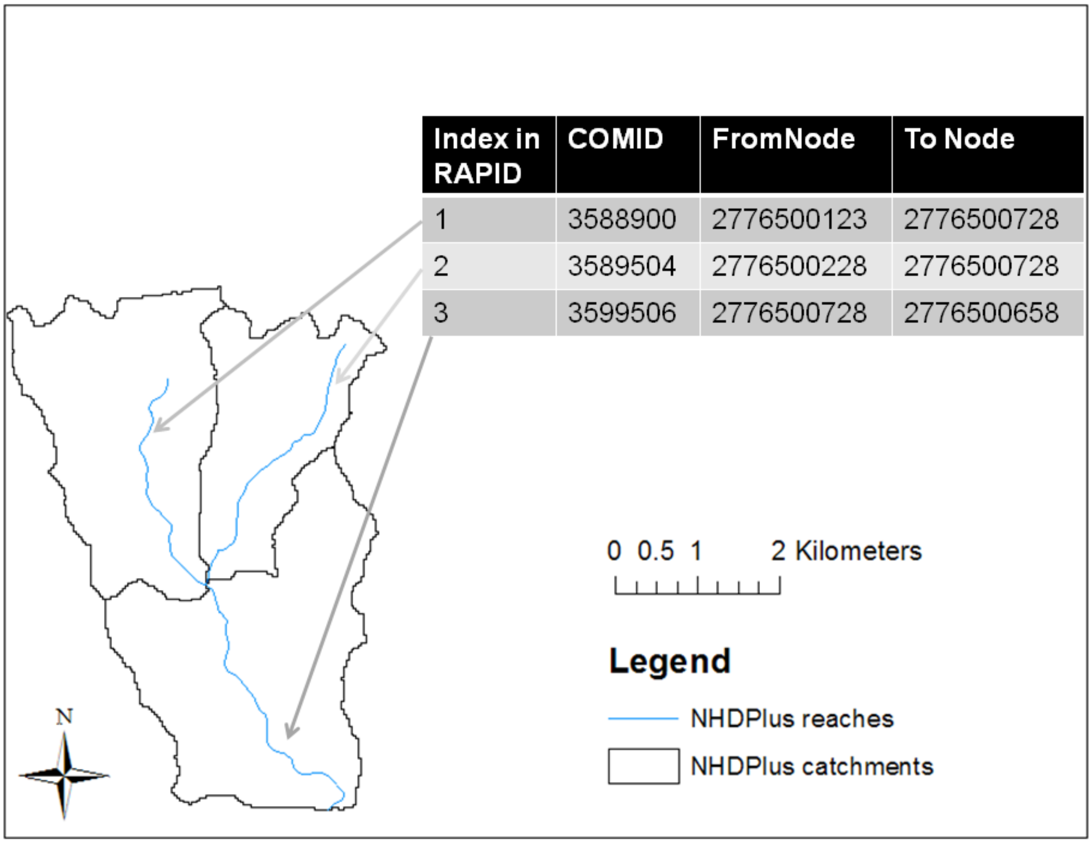

Figure 14 NHDPlus connectivity between reaches, nodes and catchments

Figure captions

Figure 1 Guadalupe and San Antonio Basins

Figure 2 NHDPlus river network and catchments for the Guadalupe and San Antonio Basins

Figure 3 Upper Mississippi River Basin

Figure 4 River network

Figure 5 Principle of flux coupler between Noah and RAPID

Figure 6 Lagged cross-correlation as a function of lag time

Figure 7 Wave celerities are estimated for eleven different sub-basins within the Guadalupe and San Antonio river basins. Location of 36 gaging stations used for optimization and names of the 15 gaging stations used for estimation of wave celerities. The same sub-basins are used for distributed parameters in RAPID

Figure 8 Statistics of river reach lengths in Guadalupe and San Antonio River Basins

Figure 9 Ratio between observed and modeled stream flow at 16 gages, location of the Edwards Aquifer. Location of the two largest springs in Texas.

Figure 10 Hydrograph of observed, lumped and routed flows for the Guadalupe River near Victoria, using (k^{γ},x^{γ})

Figure 11 Comparison of computing time between the traditional Muskingum method and matrix methods

Figure 12 Total computing time for matrix method with an iterative solver as a function of the number of processing cores, number of iterations needed, total computing time for the traditional Muskingum method.

Figure 13 Longest path in the Upper Mississippi River Basin and location of sub-basins when RAPID is used in a parallel computing environment with 8 and 16 processing cores, different colors correspond to different cores.

Figure 14 NHDPlus connectivity between reaches, nodes and catchments