

Copyright

by

Ahmad Ali Tavakoly Zadeh

2014

The Dissertation Committee for Ahmad Ali Tavakoly Zadeh Certifies that this is the approved version of the following dissertation:

Flow and transport modeling in large river networks

Committee:

David R. Maidment, Supervisor

Zong-Liang Yang, Co-Supervisor

Spyridon A Kinnas

Danny Reible

Paola Passalacqua

Flow and transport modeling in large river networks

by

Ahmad A. Tavakoly Zadeh, B.E.; M.E.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2014

Dedication

This dissertation is dedicated to my family and my teachers

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my Ph.D. supervisors Dr. David Maidment and Dr. Zong-Liang Yang. I sincerely appreciate the freedom they have given me to explore and work on my research while supporting, guiding, and challenging my endeavors.

I would also like to thank Dr. Kinnas, Dr. Passalacqua, and Dr. Reible for being members of my dissertation committee. I thank them for their support and valuable comments that helped me write this dissertation.

Special thanks to Dr. Cédric H. David who encouraged me during this research and from whom I learned how to use the RAPID model. I will never forget his excellent advice and expertise he gave me during my journey at UT.

At Ecole des Mines de Paris, I am very grateful to Dr. Florence Habets and Dr. Nicolas Flipo for allowing me to visit their teams and for sharing their model source codes and data. I would also like to thank Dr. Emmanuel Ledoux and Dr. Gilles Billen.

Dr. Timothy L. Whiteaker, Dr. Lilit Yeghiazarian, Allen Teklitz, Dr. James McClelland, Claire G Griffin, Lisa Helper Meyer, Dr. Ernest To, and Brian A. Kiel were closely involved with the total nitrogen modeling in San Antonio and Guadalupe basins project. I am grateful to all of them.

My appreciation goes to members of the Land Environment and Atmospheric Dynamics research group, Dr. Hua Su, Dr. Long Zhao, Dr. Mingjie Shi, Xitian Cai, Maryia Halubok, Yonghwan Kwon, Qinjian Jin, Peirong Lin, Sagar Prasad Parajuli, Alex Resovsky, and Yongfei Zhang.

I would like to thank Patricia Bobeck and Bryan Barnett for their valuable time and constructive comments on the work contained herein.

Sharon Bernard, Marcy Betak, Susan Swanson-Cartwright, Scott Hammock, and Michael A Godwin at the CRWR Office and Danny Quiroz, Kris Powledge, Leslie McCroddan, and Dori Eubank at the Civil, Architectural and Environmental Engineering Department for helping with all my administrative paperwork and questions.

I am indebted to all my friends at the University of Texas and Mines Paristech, Rachel Chisolm, Dr. Georges Comair, Harish Sangireddy, Elisabeta Poci, Dr. Amanda Van Epps, Anne Mikelonis, Dr. Shane Walker, Dr. Lee Blaney, Dr. Clark Siler, Marcelo Somos, Andrea Ryan, Carlos Galdeano, Ali Morovat, Dr. Arash Motamed, Roxana Darvari, Dr. Wilfried Queyrel, Lauriane Vilmin, Dr. Firas Saleh, Dr. Ahmed Kayad Moussa, Dr. Cyril Bourgeois, Léna Abasq, Baptiste Labarthe, and Dr. Emad Jahangir.

This project was partially funded by the NASA Interdisciplinary Science Project NNX11AJ43G, by the French Mines Paristech, and by the French Programme Interdisciplinaire de Recherche sur l'Environnement de la Seine (PIREN-Seine) project.

Most importantly, I am deeply grateful to my parents Saideh and Mohammad, my sisters, Sara and Sepideh, and my brother Ali for their support and love throughout this journey since April 8th 2008 when I first came to the USA. I would not be at this point without their help. I thank My Aunt, Afsi Afshar for her support. My Uncles, Vahid Afshar and Mahmoud Afshar have been like fathers to me. Their supports and encouragement helped me to concentrate to my study. My cousin Eman Afshar, is like a brother to me. Thank you for believing in me and helping me reach my goals in life.

Flow and transport modeling in large river networks

Ahmad Ali Tavakoly Zadeh, Ph.D.

The University of Texas at Austin, 2014

Supervisor: David R. Maidment

Co-Supervisor: Zong-Liang Yang

The work presented in this dissertation discusses large scale flow and transport in river networks and investigates advantages and disadvantages of grid-based and vector-based river networks. This research uses the Mississippi River basin as a continental-case study and the Guadalupe and San Antonio rivers and Seine basin in France as regional-case studies. The first component of this research presents an extension of regional river flow modeling to the continental scale by using high resolution river data from *NHDPlus* dataset. This research discovers obstacles of flow computations for river a network with hundreds of thousands river segments in continental scales. An upscaling process is developed based on the vector-based river network to decrease the computational effort, and to reduce input file size. This research identifies drainage area as a key factor in the flow simulation, especially in a wetter climate. The second component of this research presents an enhanced GIS framework for a steady-state riverine nitrogen transport modeling in the San Antonio and Guadalupe river network. Results show that the GIS framework can be applied to represent a spatial distribution of flow and total nitrogen in a large river network with thousands of connected river segment. However, time features of the GIS environment limit its applicability to large scale time-varied modeling. The third component shows a modeling regional flow and transport with consideration of stream-

aquifer interactions at regional scale at high resolution. The STICS- Eau-Dyssée combined system is implemented for entire seine basin to compute daily nitrate flux in the Seine grid river network. Results show that river-aquifer exchange has a significant impact on river flow and transport modeling in larger river networks.

Table of Contents

LIST OF TABLES	XII
List of Figures	xiii
Chapter 1: Introduction	1
1.1. Background and motivation	1
1-2 Objectives	4
1.3. Outline of dissertation.....	5
Chapter 2: River Routing for the Mississippi River Basin using Grid and Vector Based River Networks	6
2.1. Abstract	6
2.2. Introduction.....	7
2.3. Methodology.....	9
2.3.1. The river routing model, RAPID	9
2.3.1.1. Optimization procedure in the RAPID model	11
2.3.2. Study area and datasets	12
2.3.3. River networks	16
2.3.3.1. Grid-based river network	16
2.3.3.2. Vector-based river network.....	16
2.3.3.3. Upscaling of catchment using NHDPlus dataset	17
2.3.4. Lateral inflow	18
2.3.4.1. Mosaic.....	20
2.3.4.2. VIC.....	20
2.3.5. Spatial variability of parameters	20
2.3.6. Criteria of performance.....	22
2.4. Results and discussion	24
2.4.1. Effect of topography on flow rate computation.....	24
2.4.2. The effect of cost functions and optimization results	25
2.4.3. Comparison of the grid-based and vector-based river networks using the Mosaic land surface model	25

2.4.4. Comparison of the grid-based and vector-based river networks using the VIC land surface model	29
2.4.5. Comparison of the simulated streamflow using the VIC and Mosaic land surface models for the vector river network	29
2.4.6. The effect of drainage area on the 9-year mean flow	34
2.4.7. Comparison of mean annual flow for grid-based and vector-based river networks using VIC and Mosaic land surface models	35
2.5. Conclusions.....	36
Chapter 3: A GIS framework for regional modeling of riverine nitrogen transport: Case study, San Antonio and Guadalupe basins.....	
39	39
3.1. Abstract.....	39
3.2. Introduction.....	40
3.3. Methods.....	42
3.3.1. The <i>NHDPlus</i> dataset	43
3.3.2. The river routing model, RAPID	45
3.3.3. Schematic Processor	46
3.3.4. Quantifying the effect of agriculture and urbanization on the nitrogen cycle across Texas	48
3.3.5. Nitrogen load removal in reservoirs and lakes	51
3.3.6. Modeling approach	51
3.4. Case Study: In-stream nitrogen load for the San Antonio and Guadalupe basins.....	53
3.4.1. Study area.....	53
3.4.2. Waste water treatment plants	55
3.4.3. Schematic network.....	56
3.5. Results.....	58
3.5.1. In-catchment TN	58
3.5.2. Calibration process.....	58
3.5.3. Comparison of TN load San Antonio and Guadalupe basins	60
3.5.4. Effect of reservoirs on the nitrogen removal	61
3.5.5. Effect of stream order on nitrogen delivery	62
3.5.6. ESRI Map service for water quality modeling	65

List of Tables

Table 1:	Spatial and temporal characteristics of study areas	4
Table 2:	Tributaries of the Mississippi Basin and their Characteristics.....	15
Table 3:	Statistics of the model outputs and observations for the grid river network using the Mosaic land surface model.....	27
Table 4:	Statistics of the model outputs and observations for the vector river network using the Mosaic land surface model.....	28
Table 5:	Statistics of the model outputs and observations for the grid river network using the VIC land surface model.....	30
Table 6:	Statistics of the model outputs and observations for the vector river network using the VIC land surface model.....	31
Table 7:	Summary of point source data	56
Table 8:	Percentage of point source and nonpoint source data in the San Antonio and Guadalupe basins	56
Table 9:	San Antonio and Guadalupe basins	58
Table 10:	Model calibration	60
Table 11:	Mean morphological characteristics of the Seine drainage network	70

List of Figures

Figure 1:	Framework of this research.....	3
Figure 2:	Study area and stations.....	14
Figure 3:	(a) schematic of data processing for the grid-based river network; (b) schematic of data processing for the vector-based river network; (c) snapping the gauge locations and the coupling process of LSM and river network and (d) the gauge locations and the coupling process of LSM and river network Mississippi river model	19
Figure 4:	(a) river parameters in the grid-based river networks and (b) river parameters in the vector-based river networks	23
Figure 5:	The observed and modeled streamflow at the Murray Dam near Little Rock station over three months: (a) model optimization using the first cost function and (b) model optimization using the second cost function	26
Figure 6:	The observed and modeled streamflow using grid-based and vector-based river networks over nine years: (a) Mississippi River at Vicksburg, MS; (b) Missouri River at Hermann, MO and (c) Mississippi River at Keokuk, IA.....	32
Figure 7:	The observed and modeled streamflow for VIC and Mosaic using vector river network over nine years: (a) Mississippi River at Vicksburg, MS; (b) Missouri River at Hermann, MO and (c) Mississippi River at Keokuk, IA.....	33
Figure 8:	Watershed delineation versus NHD <i>Plus</i> catchment at the Ohio River at Metropolis station	34

Figure 9:	Comparison of observed and calculated mean annual flow for all stations	36
Figure 10:	GIS-based modeling framework	44
Figure 11:	(a) An example of three elemental catchments in a schematic network including links and nodes; (b) A simple schematic network showing how to manipulate a function with a Schemalink.....	47
Figure 12:	Assigning nitrogen load to catchments on the basis of county estimates	50
Figure 13:	San Antonio and Guadalupe basins	54
Figure 14:	Schematic network for the San Antonio and Guadalupe basins	57
Figure 15:	Total delivered nitrogen load for the San Antonio and Guadalupe basins in 2008 and 2009.....	60
Figure 16:	Spatial variation of in-stream TN for the San Antonio and Guadalupe basins in 2008-2009	63
Figure 17:	Relation between Strahler order and (a) number of river reaches; (b) river reach channel width; (c) % of delivered TN; (b).....	64
Figure 18:	The Seine River Watershed	71
Figure 19:	Modeling framework	73
Figure 20:	General Simulation Units in the STICS model.....	74
Figure 21:	Surface and river grids for the Seine basin	76
Figure 22:	Schematic framework of spatial coupling of Eau-Dyssee and STICS79	
Figure 23:	Comparison of annual simulated nitrate flux with observation	81
Figure 24:	Simulated daily nitrate flux and inflow for the selected stations.....	82
Figure 25:	Simulated annual average nitrate flux and inflow for the selected stations	83

Figure 26: Mass balance of the nitrate flux and inflow for the Seine basin (a)
Cumulative annual mass balance; (b) daily mass balance during a year
.....85

Chapter 1: Introduction

1.1. BACKGROUND AND MOTIVATION

Hydrologic science lags behind atmospheric science in its ability to continuously assimilate observation data with simulation models. In contrast to the hydrologic modeling, meteorological modelers use standard techniques to adjust the state of models for numerical weather dynamics over the various space and time scales. Hydrologic science requires the integration of multidimensional, spatiotemporal data into an atmospheric model linked with a land surface model and a river model for long lead time forecasting.

The world's rivers are key components of the global hydrological cycle. They route fresh water from land within continental masses into the oceans and coastal areas. On the continental scale, water on land regulates heating and moistening of the atmosphere and directly affects Earth's climate. The river component of climate models is also very important because the runoff, and subsurface runoff produced by climate and land processes, can be validated by hydrograph data in the river routing models in addition to precipitation. Typically, river routing models reside within land surface models on the continental scale. This approach has modeling accuracy limitations and a loss of physical details from low resolution.

The next generation of river routing models should consider flow and transport on a continental scale, with high spatio-temporal resolution in terms of physical and biogeochemical processes.

Continental-scale River Transport Models (RTMs) traditionally use a grid-based approach in which water moves through flow networks made up of square grid cells. In contrast, vector-based river networks, or "blue lines" on maps, are increasingly becoming

available at global and continental scales in Geographic Information Systems datasets. For example, for the continental United States, the *NHDPlus* database describes the mapped streams and rivers as well as the catchments that surround them. This dataset provides river and catchment information for the continental United States using about three million connected river reaches, each with its associate reach catchment.

Over the past few years, the *NHDPlus* dataset has been used for regional scale river routing modeling. *David et al.* [2011b, 2013] applied the *NHDPlus* dataset to simulate flow in the San Antonio and Guadalupe basins and in the Texas Gulf Coast Hydrologic Region. The former study was conducted on a small domain (5,175 river reaches) and the latter study was implemented on the regional scale (68,143 river reaches). This study expands to continental scale river routing modeling of the Mississippi river basin, which encompasses 1/3 of continental United States.

Currently, the GIS environment is used as a framework to simulate hydrologic processes for small domains [*Whiteaker and Johnson, 2012; Johnson et al., 2013; Yang et al., 2014*]. There are challenges in developing a geospatial framework to link observations and measurements, river networks, large scale river flow, and transport modeling.

This work also explores how nitrogen transport modeling could be enhanced in by computing interactions of agriculture and groundwater nitrate contamination in regional flow and transport modeling. This proposed research is a contribution of collaboration between CRWR (Ahmad Tavakoly) and MINES ParisTech (Dr. Florence Habets and Dr. Flippo Nicolas) where the Eau-Dyssée hydrologic model has been developed. In this joint research project, the Eau-Dyssée model is coupled with an agronomic model (STICS) which is the source of nitrogen in the Seine basin in France.

This research explores the development of water flow and transport computation on large scale river networks and investigates advantages and disadvantages of grid-based and vector-based river networks. The framework of this research based on types of the river networks and study domains is shown in Figure 1. This research uses the Mississippi River basin as a continental-case study and the Guadalupe and San Antonio rivers and Seine basin in France as regional-case studies.

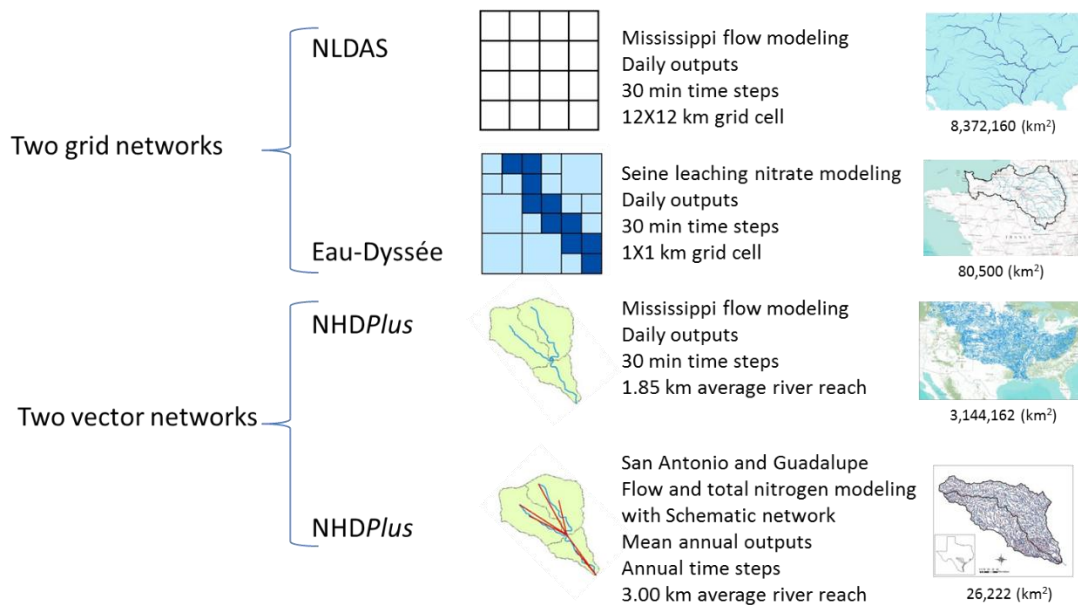


Figure 1: Framework of this research

The spatial and temporal extensions of studied domains are shown in Table 1. This table shows spatial extension of three basins, which are selected in this study, study periods, and river network characteristics in each study. Modeled variables and time steps of simulation also shown in this table.

Table 1: Spatial and temporal characteristics of study areas

Study area	River network	Number of stream reaches or grid cells	Average length of stream reaches or grid cells (km)	Area (km ²)	Time period	Time step	Variables
Mississippi	Vector	211,476	1.85	3,144,162	2000-08	30 min	Flow
Mississippi	Grid	58,140	12.13	8,372,160	2000-08	30 min	Flow
San Antonio and Guadalupe	Vector	5,196	3.00	26,222	2008-09	Mean annual	Total nitrogen and flow
Seine	Grid	6,481	1.00	80,500	1971-2010	30 min	Nitrate, flow

1-2 OBJECTIVES

This research shows the contribution in the flow and transport modeling of large scale river networks using two river networks: grid-based and vector-based river networks and focuses on continental scale river modeling and regional scale pollution modeling with river networks of thousands of connected stream reaches. This research builds upon the Routing Application for Parallel computation of Discharge (RAPID) model developed by *David* [2009].

The objective of this dissertation is to address the following three research questions:

1. What are the pros and cons of grid-based and vector-based river networks? Are there significant differences between river flows computed with these two approaches?

2. How can the GIS framework be developed to integrate a precomputed flow from RAPID and a GIS-based total nitrogen dataset for steady-state riverine nitrogen transport along a river network?

3. How does the interactions of agriculture and groundwater nitrate contaminants effect flow and transport modeling in large scale river networks? What are essential modifications for the dynamic coupling of Eau-Dyssée and STICS models?

1.3. OUTLINE OF DISSERTATION

This dissertation consists of a series of three related papers presented respectively in Chapter 2, Chapter 3 and Chapter 4. In the second chapter, river flow modeling is computed in 30 minute time step for the entire Mississippi river basin using both traditional grid-based river and vector-based river networks for a period of nine years (2000-2008). In particular, the technical challenge of input data preparation for entire domain with hundreds of thousands river reaches are described.

In the third chapter, a GIS framework is presented for steady-state flow and transport modeling in large river networks. The study is a 2-year case study (2008-09) of total nitrogen change in urban and rural regions (San Antonio and Guadalupe) along all river reaches in the San Antonio and Guadalupe River Basin.

In the fourth chapter, an enhanced technique for modeling nitrogen pollution from agriculture is presented. The STICS–Eau-Dyssée coupled models are applied to compute interactions of agriculture and groundwater nitrate contaminants over an entire Seine basin from 1971 to 2010. The total area of studied region is 80,500 km² and the river network includes 6481 river cells.

In the fifth chapter, the research results and findings are summarized and recommendations for future works are presented.

Chapter 2: River Routing for the Mississippi River Basin using Grid and Vector Based River Networks

2.1. ABSTRACT

Continental-scale river transport models (RTMs) traditionally use a grid-based approach in which water moves in flow networks made of square grid cells. Vector-based river networks – the “blue lines” from maps – are now increasingly becoming available at global and continental scales in Geographic Information Systems datasets. The vector-based approach may allow a better representation of river networks and a more accurate estimation of river model parameters than grid-based networks. The vector-based river network is also used for flood mapping. In this study, the Routing Application for Parallel computation of Discharge (RAPID) model was applied to the Mississippi River Basin for a period of nine years (2000-2008) using both grid-based and vector-based approaches for river networks and river parameters. Runoff data from the Mosaic and VIC land surface models available from the second phase of the North American Land Data Assimilation System (NLDAS2) were used as input to RAPID. The grid-based river network was derived from NLDAS2. The vector-based river network was extracted from the enhanced version of the National Hydrography Dataset (*NHDPlus*). It is shown that the vector-based approach allows for a more accurate representation of size and location of river basins than the grid-based approach which is particularly noticeable in wet climates (Ohio River). Analyses for the determination of river model parameters suggest that vector-based river networks are advantageous because the real length and slope of river segments can be obtained from the *NHDPlus* dataset. Finally, this study confirms that weighing the relative influence of each gauging station when optimizing river model parameters improves stream flow computations.

2.2. INTRODUCTION

Most climate system models include continental-scale land surface models (LSMs) along with river routing models as lower terrestrial boundaries in the global water cycle [*Larson et al.*, 2007].

Inclusion of continental-scale river models in climate systems have been in development since the work of *Miller et al.* [1994]. Large scale river routing models have been usually developed based on river networks derived from a Digital Elevation Model (DEM), which will be referred to as “grid-based river network” in this paper.

Some of these models include: LISFLOOD-FP [*Bates and De Roo*, 2000], RiTHM [*Ducharne et al.*, 2003], MODCOU [*Ledoux et al.*, 1989], and LEAF-Hydro-Flood [*Miguez-macho and Fan*, 2012]. In LISFLOOD-FP the channel length and slope are determined from a local drainage direction map that illustrates the line of the channel and each cell contains a marker representing the direction of the downstream cell. The latest version of LISFLOOD-FP solves a reduced complexity version of the Saint-Venant equations and calculates water depth and river flow between cells at each time step [*Bates et al.*, 2010]. In the RiTHM model (River-Transfer Hydrological Model), the grid slope is based on the meridian and zonal directions and the flow direction is limited to four direction classes: north, east, south, and west. Furthermore, the water travel time between two cells depends on the distance and the slope between the two grid-points. Modèle Couplé (MODCOU), which means “coupled model” in the French language, routes surface runoff using the linear reservoir scheme. The transfer time constants between cells are estimated based on the topography, the distance between cells and the drainage area. The LEAF-Hydro-Flood model is a grid-base model which solves the 1D

momentum equation. This model applies Manning equation to determine the mean flow depth.

Although grid-based stream networks have been prevalent within the past two decades, vector-based networks from mapped “blue lines” are increasingly becoming available at the continental and the global scale in geographic information system (GIS) datasets. For example, the Hydrological Data and maps based on Shuttle Elevation Derivatives at Multiple Scales (HydroSHEDS) provide vector-based river networks globally [Lehner *et al.*, 2006]. In the United States, a coherent description of topography and hydrographic features is available in the National Hydrography Dataset Plus (NHDPlus; Horizon Systems Corporation 2007) which includes vector-based river networks. The NHDPlus database describes the mapped streams and rivers as well as the catchments that surround them, including river and catchment information such as the river length and slope, and catchment area. Over the past few years, the NHDPlus dataset has been used in large scale river routing. The Routing Application for Parallel computation of Discharge (RAPID) is a river routing model that was first developed using NHDPlus river networks [David, 2009; David *et al.*, 2011b]. RAPID uses a matrix formulation of the Muskingum method to simultaneously compute discharge of water in many thousands of reaches of large river networks including at ungaged locations. RAPID has been applied to vector-based river networks to simulate flow in the San Antonio and Guadalupe basins and in the Texas Gulf Coast Hydrologic Region. The former study was conducted on a small domain (5175 river reaches) and the latter study was implemented on the regional scale (68,143 river reaches). [David *et al.*, 2011a, 2013]

The purpose of this paper is therefore to investigate the following research questions:

1. What are the pros and cons of grid-based and vector-based approaches; are there major differences between river flow computed with each approach? For this purpose, RAPID was applied to the Mississippi River Basin from 2000 to 2008 with two different river networks and with two land surface models.

2. To what extent does information such as slope and length of river reaches help improve flow simulations in large vector-based and large grid-based river networks? To answer this question, four types of spatial variability of the Muskingum K parameter were defined and tested.

3. How does the optimization cost function affect the corresponding calibrated river model results in large river networks? To answer this question, we used two different cost functions in the RAPID model to optimize the K and X parameters.

To answer the above questions, the river modeling is conducted in the Mississippi River Basin as described in the following sections. The river routing model, study area, datasets, river networks, lateral inflows, spatial variability of the river routing model, and criteria of performance are described in “Methodology”. Advantages and disadvantages of grid-based and vector-based approaches, influence of topography on river modeling, and effect of optimization functions in river model parameters are presented on “Results and discussion”. Finally, the implications of the results in large scale river modeling using grid-based and vector-based river networks are summarized in “Conclusions”.

2.3. METHODOLOGY

2.3.1. The river routing model, RAPID

The RAPID model is a river network routing model. Notable features of RAPID include the use of the “blue lines” on the map as well as grid network for river networks

and an automated parameter estimation procedure (<http://www.ucchm.org/david/index.html>). RAPID uses a matrix-based version of the Muskingum method to compute the flow and volume of water in river networks containing many thousands of reaches:

$$(\mathbf{I} - \mathbf{C}_1 \cdot \mathbf{N}) \cdot \mathbf{Q}(t + \Delta t) = \mathbf{C}_1 \cdot \mathbf{Q}^e(t) + \mathbf{C}_2 [\mathbf{N} \cdot \mathbf{Q}(t) + \mathbf{Q}^e(t)] + \mathbf{C}_3 \cdot \mathbf{Q}(t) \quad (1)$$

where: t is time and Δt is the river routing time step, \mathbf{I} is the identity matrix, \mathbf{N} is the river network matrix, \mathbf{Q} is a vector of outflows from each reach, \mathbf{Q}^e is a vector of lateral inflows for each reach, \mathbf{C}_1 , \mathbf{C}_2 and \mathbf{C}_3 are parameter diagonal matrices and for a given reach j , they can be represented as [McCarthy, 1938]:

$$\begin{aligned} C_{1j} &= \frac{K_j X_j - 0.5 \Delta t}{K_j (1 - X_j) + 0.5 \Delta t} \\ C_{2j} &= \frac{K_j X_j + 0.5 \Delta t}{K_j (1 - X_j) + 0.5 \Delta t} \\ C_{3j} &= \frac{K_j (1 - X_j) - 0.5 \Delta t}{K_j (1 - X_j) + 0.5 \Delta t} \end{aligned} \quad (2)$$

K_j is the storage constant (with a time dimension) and X_j is a dimensionless weighting factor. Lateral inflow in this model is calculated by the land surface model.

To avoid estimating (K_j and X_j) for all reaches, the RAPID model optimizes two multiplying factors λ_k and λ_x using the following equations:

$$K_j = \lambda_k \cdot \frac{L_j}{V_{wj}}, \quad X_j = \lambda_x \cdot 0.1 \quad (3)$$

where: L_j is a river reach length and V_{wj} is a wave celerity.

The optimization is executed uniformly over the study domain. Thus, the optimization gives unique values of λ_k and λ_x applied to all river reaches. At the end of the optimization procedure, the values of λ_k and λ_x coefficients which give the best statistical results are used for flow routing over the Mississippi basin. The optimization procedure is explained in the following subsection.

2.3.1.1. Optimization procedure in the RAPID model

RAPID uses the inverse method to estimate the K and X parameters that minimizes a cost function computed based on differences between computations and observations of stream flow at many gages located throughout river basins. The current version of RAPID includes two possible optimization cost functions that differ in how each gage is weighted depending on the average stream flow it measures. The best set of λ_k and λ_x coefficients have the minimum cost function. The first cost function (ϕ_1) is based on the sum of the error between daily observed and averaged flow over the optimization time period and along the number of selected stations (n):

$$\phi_1 = \sum_{t=t_0}^{t_f} \sum_{i=1}^n \left[\frac{\bar{Q}_i(t) - Q_i^{obs}(t)}{f} \right]^2 \quad (4)$$

where: $\bar{Q}_i(t)$ is the daily average of the computed flow, $Q_i^{obs}(t)$ is the daily observation, n is the number of the gauges used in the optimization, f is a scalar that applies to automate the optimization process, t_0 and t_f are the first and last day of the optimization period respectively.

In Eq. (4), the magnitude of the flow affects the cost function. A station with larger flow influences ϕ_1 more than a station with smaller flow for the contributed fractional error. Consequently, in this function, gaging stations do not have the same weight. As a result, the ϕ_1 cost function could constrain the Nash efficiency. To resolve this issue, a second cost function ϕ_2 is developed in RAPID, which gives the same weight to all stations regardless of the magnitude of the flow:

$$\phi_2 = \sum_{t=t_0}^{t_f} \sum_{i=1}^n \left[\frac{\overline{Q_i}(t) - Q_i^{obs}(t)}{Q_i^{obs}} \right]^2 \quad (5)$$

where: $\overline{Q_i^{obs}}$ is the daily average of the observed flow over the time period.

Based on the cost function, the optimized model coefficients (K and X) are introduced into the RAPID model to compute discharge.

2.3.2. Study area and datasets

Based on Hydrologic Unit Code (HUC-2 codes), the United States is divided into 21 major regions, six of which define the Mississippi river basin [Seaber *et al.*, 1987]. These six regions are 5, 6, 7, 8, 10 (divided into 10L and 10U), and 11 (Figure 2). Latitudes and longitudes vary from 27.9375N to 51.5625N and from 115.0625W to 76.9375W, respectively. The total surface area is approximately 2,981,076 km².

Simulated flows are compared to observations for seven stations. Criteria for selecting gaging stations are observed-daily flow availability for the period studied, assessment of the model performance in the wet (region 5, 6, most part of region 7, and 8) and dry regions (region 10, and most part of region 11) and consistency with the literature [Maurer *et al.*, 2001; Lohmann *et al.*, 2004; Troy *et al.*, 2008]. In addition, the

selected gages cover the main basins and the Mississippi River at Vicksburg, MS gauge was selected as the Mississippi basin mouth. Table 2 shows the location of the stream gages, drainage area, and *NHDPlus* region corresponding to each station.

The NLDAS, *NHDPlus* and USGS datasets provide the required input data for river modeling. NLDAS Data processing includes two steps: 1) process land surface model (LSM) output to define lateral flow and 2) create grid-based river network. A grid-based river network was determined using an eight-direction method applied to a 0.125-degree digital elevation model (DEM).

Runoff data from the Mosaic and VIC land surface models available from the North American Land Data Assimilation System (NLDAS2) were used as input to RAPID. The *NHDPlus* dataset was used to create vector-based river networks. Furthermore, daily stream flow data from USGS; National Water Information System (NWIS) were obtained as observations to optimize and compare the RAPID model outputs. Daily stream flow from USGS are easily accessible, however; we acknowledge that naturalized flow in a daily basis is the best scenario for optimization since anthropogenic effects of both management and use (e.g. reservoirs, diversions) are considered.

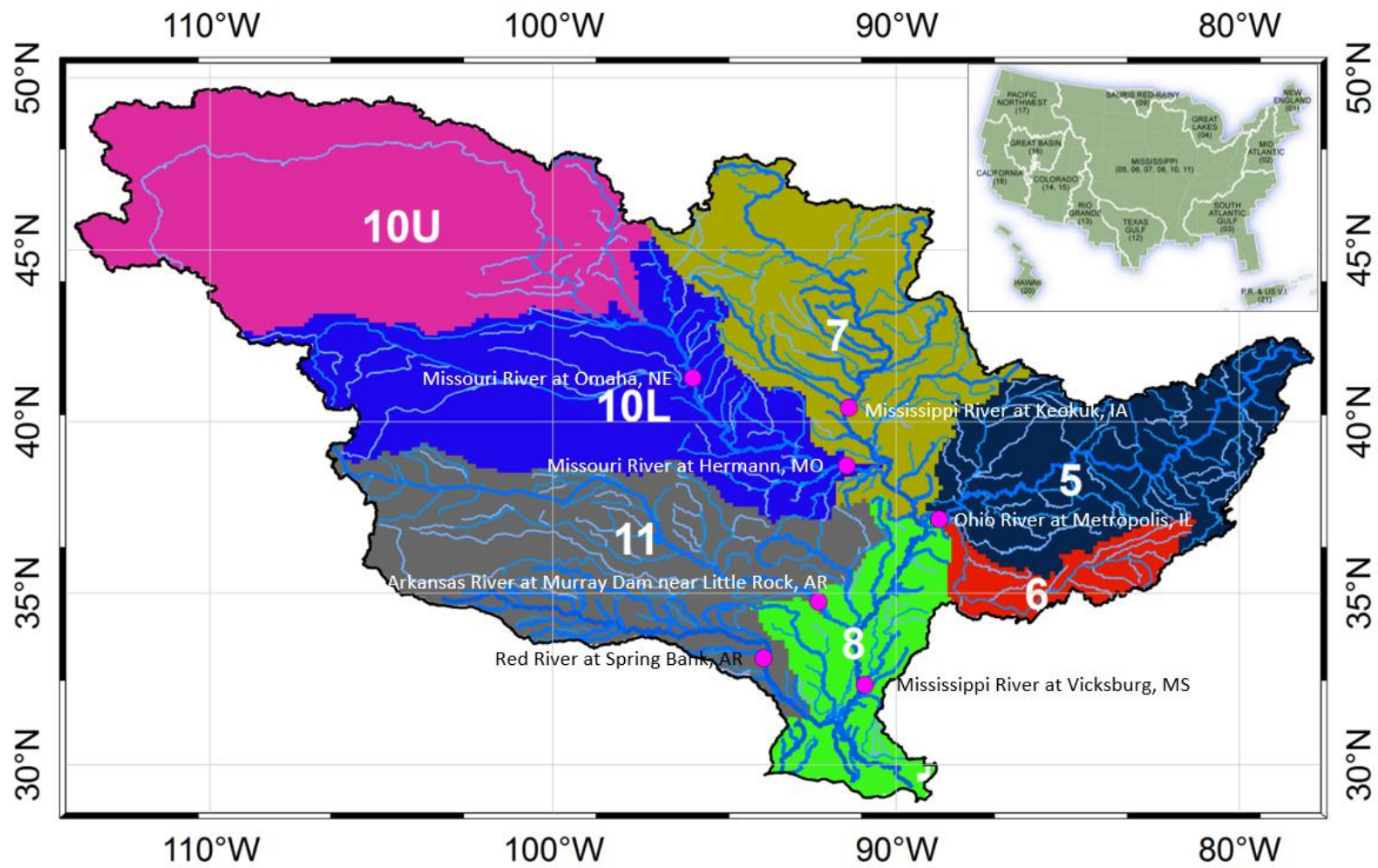


Figure 2: Study area and stations

Table 2: Tributaries of the Mississippi Basin and their Characteristics.

Station Name	LON	LAT	Drainage area(km ²)			9-year average flow	NHD <i>Plus</i> Region
			USGS	Grid	Vector	Observed flow (m ³ s ⁻¹)	
Ohio River at Metropolis, IL	-88.74	37.15	525,768	401,383	523,498	8,051	5
Mississippi River at Keokuk, IA	-91.37	40.39	308,209	281,641	298,719	2,147	7
Mississippi River at Vicksburg, MS	-90.91	32.32	2,953,881	2,733,922	2,913,317	16,660	8
Missouri River at Hermann, MO	-91.44	38.71	1,353,269	1,324,115	1,310,467	1,983	10L
Missouri River at Omaha, NE	-95.92	41.26	846,302	805,372	817,083	747	10L
Arkansas River at Murray Dam near Little Rock, AR	-92.36	34.79	409,296	330,460	391,738	1,232	11
Red River at Spring Bank, AR	-93.86	33.08	-	270,822	146,785	544	11

* Datum is above NGVD29

2.3.3. River networks

2.3.3.1. Grid-based river network

The 1/8th-degree resolution over the NLDAS domain which is based on the GTOPO30 Global 30 Arc Second (~1-km) Elevation Dataset is used to create a grid-based river network map (<http://ldas.gsfc.nasa.gov/nldas/NLDASelevation.php>). Since the runoff data were also obtained from the NLDAS data set, coupling the LSM output with a grid-based river network is straight forward. Nevertheless, to determine the station locations, a snapping process was performed over this type of river network. A Snap Pour Point tool is used to determine the cell of highest flow accumulation for the delineated catchment. Figure 3 shows the coupling process for a grid-based river network. To create the grid-based river network, flow direction grid cells (FDR) and flow accumulation grid cells (FAC) were created using the eight-direction pour point model (D8). The D8 method assumes that the steepest decent of eight neighboring cells determines the flow direction [Olivera *et al.*, 2002]. Once the flow direction is determined, the connection between grid cells can be calculated; consequently, for each grid the upstream grid cells and downstream grid cell can be determined. The rectangular domain in the grid-based river network, which covers the Mississippi basin, includes 58,140 grid cells.

2.3.3.2. Vector-based river network

The NHD*Plus* database provides the mapped streams and rivers as well as the catchments that surround them in the United States (<http://www.horizon-systems.com/nhdplus/data.php>) and used to create vector-based river network. The NHD*Plus* dataset is an integration of the National Hydrography Dataset (NHD), the National Elevation Dataset (NED), and the National Watershed Boundary Dataset

(WBD). The river network in *NHDPlus* is not from a DEM and is based on the USGS Quad maps which were digitized into a vector dataset called a Digital Line Graph. Finally, the streams (i. e. “bluelines”) and hydrologic features from the maps were transformed to the NHD dataset. Unlike the grid-based river network, in the vector-based river network the real location of a gauging station can be easily determined. However; coupling the LSM and vector-river network requires GIS processing with the assumption that the contributing catchment area for each river reach is approximately equal to the area of one grid cell.

The *NHDPlus* dataset includes a functional feature for upstream and downstream navigation called “value-added attributes” (VAA). According to the “VAA” feature, each river reach in the national network and contributing local catchment has the same unique integer identifier, COMID. Besides COMID, the FROMNODE, TONODE, and DIVERGENC attributes are applied to produce river connectivity. More details on how to use this dataset to create river connectivity are given in *David et al.* [2011a]. The attribute table of the NHD flowline feature shows that the Mississippi basin has a total of 1,197,396 reaches, with an average length of 1.82 km and average catchment area of 2.75 km².

2.3.3.3. Upscaling of catchment using NHDPlus dataset

Flow simulation for 1,197,396 river reaches is computationally expensive. The runoff file, which is the main input file in the RAPID model, for 1,197,396 river reaches is about 118 gigabytes (GB). The RAPID model in this study uses one processor to read input files and reading a file with 118 GB size is very time consuming. To reduce the simulation time, a process was developed to decrease number of river reaches One of the Value Added Attributes (VAAs), “ThinnerCod”, is an ordinal number that displays the

density of the river network with values ranging from 0 to 6, denoting an increasing density of the river network to a maximum which reflects the entire river network. ThinnerCod=0 shows all non-network rivers and the ThinnerCod = 1 denotes the main river network of the 8-digit Hydrologic Unit Codes, which is the least dense river network. In the *NHDPlus* dataset, rivers with higher ThinnerCod flow to river reaches with a lower ThinnerCod. For instance, if a river reach has ThinnerCod=2, the upstreams have $\text{ThinnerCod} \geq 2$.

The upscaling process includes: First, determine rivers with only ThinnerCod=1-5 and merge cumulative upstream rivers with ThinnerCod=6 for each river reach. Second, dissolve catchments corresponding to the ThinnerCod=6 and upstream rivers to create an equilibrium catchment of the downstream river with ThinnerCod=1-5. Details on the upscaling process are given in [Tavakoly *et al.*, 2012]. The upscaling process decreases the number of reaches from 1,197,396 river reaches to 211,476 river reaches with average length of 1.86 km and average catchment area of 15 km². The size of the runoff input file is reduced to 20 GB using 211,476 river reaches which is five times smaller than the file size for Mississippi river network with 1,197,396 river reaches.

2.3.4. Lateral inflow

In the NLDAS dataset, the hourly Mosaic and VIC output data are available in the grb format (<ftp://ldas3.ncep.noaa.gov/pub/raid0/nldas/>). From the LSM outputs, the surface and subsurface runoff data for the entire NLDAS domain were obtained and summed to compute the lateral flow from 01 January 2000 to 31 December 2008.

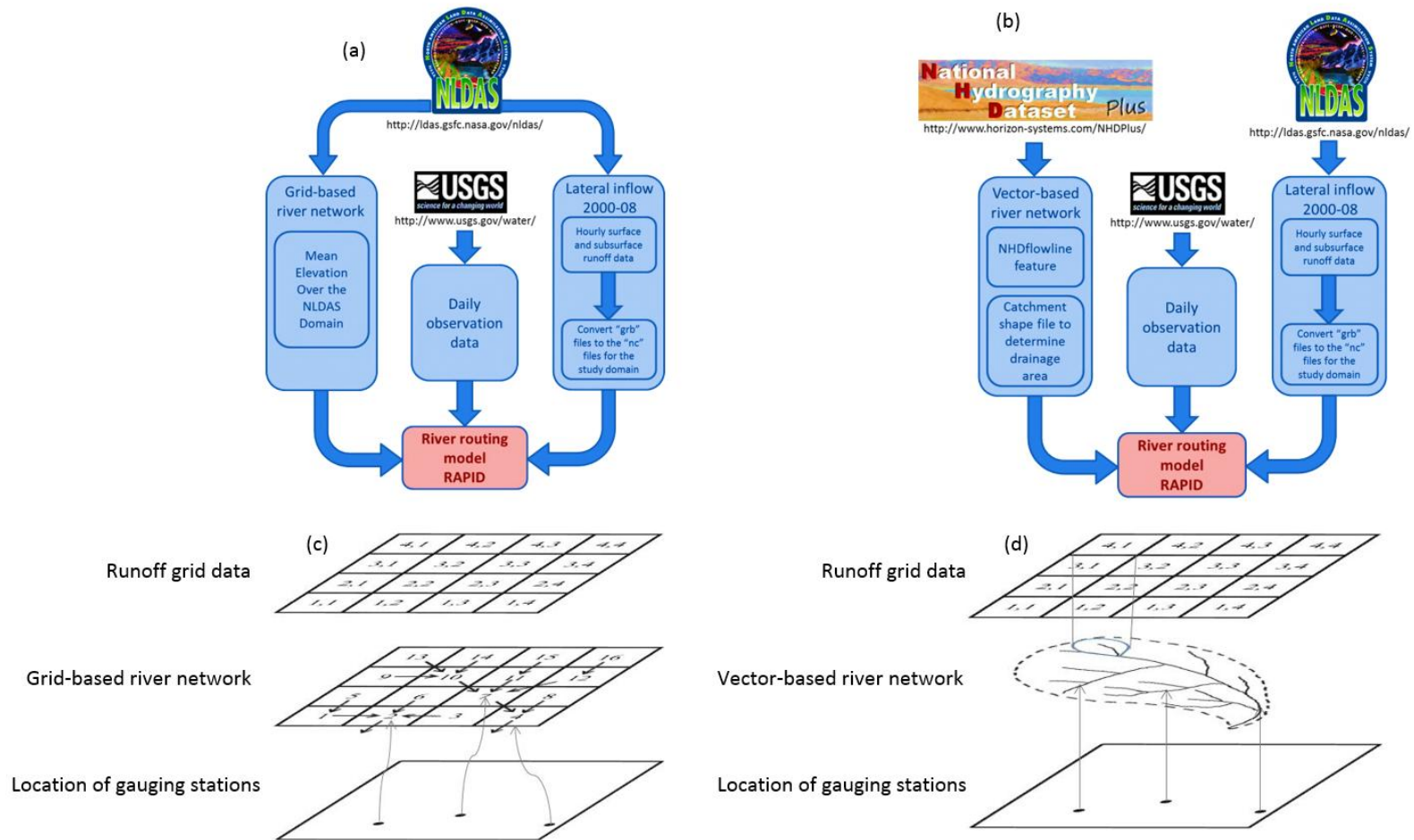


Figure 3: (a) schematic of data processing for the grid-based river network; (b) schematic of data processing for the vector-based river network; (c) snapping the gauge locations and the coupling process of LSM and river network and (d) the gauge locations and the coupling process of LSM and river network Mississippi river model

2.3.4.1. Mosaic

The Mosaic model was developed to use with an atmospheric general circulation model (GCM) [Suarez and Koster, 1996]. A grid in this model divided into several homogeneous subgrids. Each of sub-regions contains a single vegetation or bare soil. The Mosaic model simulates surface and subsurface runoff based on the through fall on the saturated fraction and on the free drainage, which depends on the slope, respectively [Lohmann et al., 1998].

2.3.4.2. VIC

The variable infiltration capacity (VIC) model is a semi-distributed macroscale hydrological model [Liang et al., 1994]. The most unique capabilities of the VIC model include: variability of infiltration capacity curve and land surface vegetation classes in subgrids [Gao et al., 2010]. The model has been extensively applied nationally in the United States and globally such as: Mississippi river basins [Berbery et al., 2003; Lohmann et al., 2004], North America [Lucas-Picher et al., 2003], Texas and Maryland [Meng and Quiring, 2008], Rhine river basin [Hurkmans et al., 2008], and China [Xie et al., 2007].

2.3.5. Spatial variability of parameters

K and X parameters need to be determined in the RAPID model. Fread [1993] estimated X to be between 0.1 and 0.3 in most streams. The Muskingum K parameter is estimated by Eq. (6) as follows [Tewolde and Smithers, 2007]:

$$K_j = \frac{L_j}{V_{wj}} \quad (6)$$

where: K_j = storage constant, L_j = reach length, V_{wj} = wave celerity.

For the optimization procedure, the initial value of K (K_{ini}) has to be defined. In this study four experiments were defined for K_{ini} . In the first and second experiments, all river reaches were assumed to have the constant wave celerity. The wave celerity of water in the large basins ranges from 0.5 to 5 m/s [Lohmann *et al.*, 2004], hence the reference wave celerity was defined as 1.35m/s=4.86 km/hr into the model. In the first two experiments the wave celerity is assumed to be independent of the topography. Nevertheless, the travel time of flow wave is affected by topography, particularly for the high topography areas. Therefore, in the last two experiments the K_{ini} are defined based on the river slope. To avoid the influence of the extreme values on the celerity, based on the cumulative probability function the 5% and 95% thresholds corresponding to $\sqrt{S_i}$ were computed and all values of $\sqrt{S_i}$ were modified in the interval [x0.5, x0.95] in the last case. Four scenarios of the K_{ini} are:

$$\text{Case (1)} \quad K_{ini}^1 = \frac{\bar{L}}{C_0} \quad (7)$$

$$\text{Case (2)} \quad K_{ini}^2 = \frac{L_i}{C_0} \quad (8)$$

$$\text{Case (3)} \quad K_{ini}^3 = \alpha \frac{L_i}{\sqrt{S_i}} \quad (9)$$

$$\text{Case (4)} \quad K_{ini}^4 = \alpha \frac{L_i}{(\sqrt{S_i})'}, \quad (\sqrt{S_i})' \in P[0.05, 0.95] \quad (10)$$

where: \bar{L} is the mean of the river length, which equals to 12.13 and 1.86 km for grid-based and vector-based river networks. C_0 is the reference water wave celerity, S_i is

the river slope, $\alpha = \frac{K^2}{L_i \sqrt{S_i}}$ is the inverse of a velocity and L_i is the river length which ranges $\forall_i \in [1, 211476]$ and $\forall_i \in [1, 58140]$ in vector-based and grid-based networks respectively.

For the vector-based river network, the river length and slope were obtained from *NHDPlus* dataset. Whereas for the grid-based river network, the river length was assumed to be the distance between the center of a grid cell and its downstream grid cell (according to the D8 method) and the slope was calculated based on elevation difference between the two grid cells (Figure 4).

2.3.6. Criteria of performance

In this study the following statistical criteria were used to assess the model simulations on both vector and gridded river networks with observations [*Wu et al.*, 2011]:

Root mean square error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Q_i^{obs} - Q_i^{sim})^2}{n}} \quad (11)$$

The RMSE is used to measure the accuracy of the model prediction and has a minimum value of 0.0.

Nash-Sutcliffe Efficiency (NSE)

$$NSE = 1 - \left[\frac{\sum_{i=1}^n (Q_i^{obs} - Q_i^{sim})^2}{\sum_{i=1}^n (Q_i^{obs} - \overline{Q_i^{obs}})^2} \right] \quad (12)$$

The NSE ranges from $-\infty$ to 1. A NSE equal to 1 means that the model prediction is perfectly matched with the observations. An efficiency lower than zero indicates that the mean observation predicts better than the model [Nash and Sutcliffe, 1970].

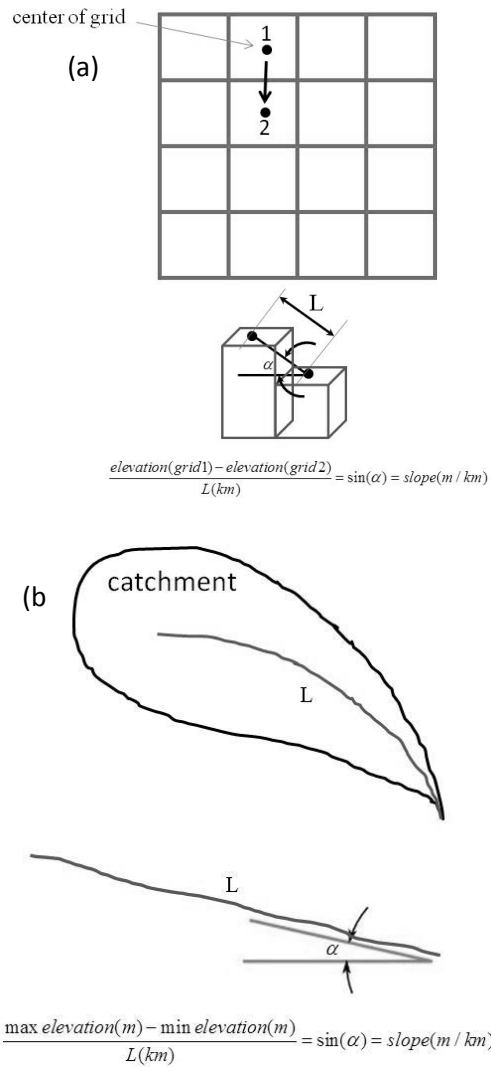


Figure 4: (a) river parameters in the grid-based river networks and (b) river parameters in the vector-based river networks

2.4. RESULTS AND DISCUSSION

This section discusses the effect of topography on the RAPID optimization for two optimization cost functions, for two river networks, and for two land-surface models. Subsequently, the differences between the grid-based with vector-based river networks using two land surface models are discussed. The RAPID model is run using 3-hourly lateral inflow volumes and a 30-minute time step. In the optimization process, one-year of data (2000) is selected to reduce computational time. Furthermore, *David et al.* [2011b] showed that to optimize RAPID, one year or less is sufficient. Due to the variation in topography among subbasins, processing optimization independently for each main subbasin may provide better results, but it has only a slight effect on the flow simulation. Statistical results of two river networks using two land surface models for selected stations are shown in table 3-6. Results show that using large scale river routing model improves river flow simulation in comparison with lumped model in all cases. Correlation (ρ) is higher in all simulation cases compare to the lumped model results. NSE also improves by coupling the RAPID model and land surface models.

2.4.1. Effect of topography on flow rate computation

The RAPID model was run for both types of river networks: grid-based (58,140 grid cells) and vector-based (211,476 reaches) networks and for both land surface models with different sets of parameters (λ_k and λ_x) and two optimization functions.

The statistical terms for topography experiments, river networks, and land surface models are listed in Tables 3-6. The mean values of statistical criteria for all stations show that ρ and NSE are highest using the second cost function and last two topography experiments. RMSE also represents the lowest value for the second cost function.

Results also shows that eliminating the extremes of the $\sqrt{S_i}$ (case four, K_{ini}^4) slightly changes the model performance. The model results using four scenarios of topographies for both const functions are shown in Figure 5. The K_{in}^3 and K_{in}^4 experiments are improved the model performance for both cost functions and track the observed flow variation relatively better than the first and second experiments.

2.4.2. The effect of cost functions and optimization results

A comparison of predicted hydrographs using ϕ_1 and ϕ_2 ; and observed steam flow for the Murray Dam near Little Rock gauging station over three months is also shown in Figure 5. The summary statistics and Figure 5 also indicate that regardless of the initial value of the K (K_{ini}), using the second cost function (ϕ_2) to optimize river routing model gives a better performance. The second cost function (ϕ_2) reduces the oscillations and the model was capable of reaching the peak flow in the same time span (Figure 5b).

2.4.3. Comparison of the grid-based and vector-based river networks using the Mosaic land surface model

The simulation results for the grid and vector river networks using the Mosaic land surface model are shown in Table 3 and Table 4. Using the Mosaic land surface model for the vector-based river network, NSE and RMSE, for both cost functions and both K_{ini}^3 and K_{ini}^4 generally outperform the grid-based river network. The model results also show that the simulated mean annual flow using the Mosaic land surface is underestimated for both river networks. The mean annual flow is also comparable for all stations except at the Ohio River at Metropolis station. At this station the vector river network shows better mean annual flow simulation than the grid river network compare to the observation. The reason can be explained by comparison of the drainage area for both river networks. Comparisons of the drainage area based on the USGS

measurements, *NHDPlus* dataset and DEM delineation (Gird river network) indicate that, drainage area provided by the vector river network (based on the *NHDPlus* dataset) is more accurate (Table 2). Hence the vector river network shows better model performances. More detail regarding the effect of drainage area on the flow simulation for both river networks is discussed in the section 2.4.6.

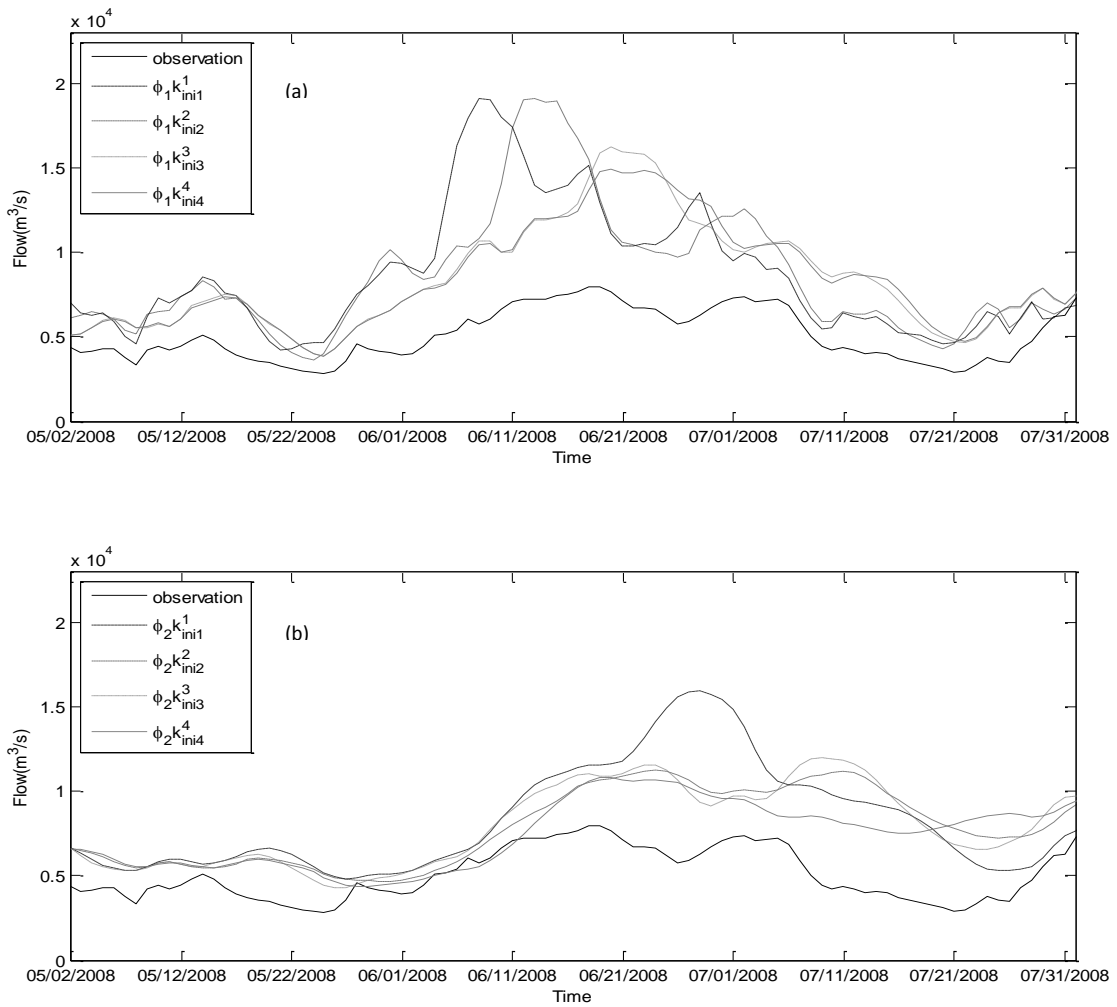


Figure 5: The observed and modeled streamflow at the Murray Dam near Little Rock station over three months: (a) model optimization using the first cost function and (b) model optimization using the second cost function

Table 3: Statistics of the model outputs and observations for the grid river network using the Mosaic land surface model

Gauging station		Observation	Lumped Model	Grid-Based River Network							
				ϕ_1				ϕ_2			
			Mosaic	K_{ini}^1	K_{ini}^2	K_{ini}^3	K_{ini}^4	K_{ini}^1	K_{ini}^2	K_{ini}^3	K_{ini}^4
Ohio River at Metropolis, IL	Average	8051	3072	3071	3071	3071	3071	3069	3069	3068	3068
	ρ	-	0.51	0.68	0.68	0.67	0.69	0.77	0.77	0.77	0.78
	NSE	-	-0.44	-0.21	-0.21	-0.22	-0.20	-0.13	-0.13	-0.13	-0.13
	RMSE	-	6966	6407	6388	6425	6358	6182	6180	6171	6167
Mississippi River at Keokuk, IA	Average	2147	718	718	718	718	718	718	718	718	718
	ρ	-	0.35	0.48	0.47	0.44	0.43	0.59	0.62	0.60	0.59
	NSE	-	-1.14	-0.75	-0.77	-0.84	-0.87	-0.53	-0.48	-0.51	-0.53
	RMSE	-	2134	1931	1942	1980	1994	1802	1775	1790	1804
Mississippi River at Vicksburg, MS	Average	16660	7809	7786	7793	7792	7790	7788	7784	7783	7783
	ρ	-	0.10	0.67	0.67	0.71	0.72	0.67	0.82	0.85	0.85
	NSE	-	-1.06	-0.65	-0.62	-0.57	-0.55	-0.44	-0.39	-0.35	-0.35
	RMSE	-	11757	10542	10428	10258	10190	9827	9640	9494	9494
Missouri River at Hermann, MO	Average	1983	1222	1221	1221	1221	1199	1221	1220	1220	1220
	ρ	-	0.45	0.65	0.65	0.62	0.66	0.78	0.80	0.81	0.79
	NSE	-	-1.07	-0.34	-0.35	-0.41	-0.31	0.11	0.17	0.21	0.17
	RMSE	-	1868	1504	1510	1545	1489	1229	1185	1154	1187
Missouri River at Omaha, NE	Average	747	486	486	486	486	486	486	486	486	485
	ρ	-	0.40	0.53	0.52	0.52	0.49	0.56	0.56	0.58	0.55
	NSE	-	-10.16	-7.93	-8.22	-7.73	-8.88	-6.88	-6.82	-6.27	-7.12
	RMSE	-	802	717	728	709	754	673	671	647	683
Arkansas River at Murray Dam near Little Rock, AR	Average	1232	733	733	733	733	733	732	732	732	732
	ρ	-	0.41	0.61	0.61	0.60	0.67	0.73	0.75	0.77	0.78
	NSE	-	-0.35	0.16	0.16	0.17	0.28	0.40	0.43	0.46	0.46
	RMSE	-	1518	1195	1196	1193	1106	1015	989	960	959
Red River at Spring Bank, AR	Average	544	455	454	456	455	455	453	455	455	454
	ρ	-	0.36	0.45	0.46	0.47	0.53	0.60	0.61	0.66	0.74
	NSE	-	-0.95	-0.36	-0.33	-0.28	-0.01	0.19	0.21	0.35	0.51
	MSE	-	891	744	736	720	642	642	566	515	446
Mean	Average	4481	2071	2067	2068	2068	2065	2067	2066	2066	2066
	ρ	-	0.37	0.58	0.58	0.58	0.60	0.67	0.70	0.72	0.72
	NSE	-	-2.17	-1.44	-1.48	-1.41	-1.51	-1.04	-1.00	-0.89	-1.00
	MSE	-	3705	3291	3276	3261	3219	3053	3001	2962	2963

Table 4: Statistics of the model outputs and observations for the vector river network using the Mosaic land surface model

Gauging station	Observation	Lumped Model	Vector-Based River Network								
			ϕ_1				ϕ_2				
			Mosaic	K_{ini}^1	K_{ini}^2	K_{ini}^3	K_{ini}^4	K_{ini}^1	K_{ini}^2	K_{ini}^3	K_{ini}^4
Ohio River at Metropolis, IL	Average	8051	3891	3881	3886	3886	3887	3879	3879	3883	3883
	ρ	-	0.51	0.73	0.74	0.76	0.76	0.64	0.87	0.77	0.77
	NSE	-	-0.30	0.05	0.05	0.10	0.10	0.11	0.11	0.19	0.20
	RMSE	-	6977	6311	6293	5502	5508	5978	5761	5477	5492
Mississippi River at Keokuk, IA	Average	2147	820	818	818	819	819	817	818	818	818
	ρ	-	0.38	0.53	0.55	0.63	0.63	0.63	0.65	0.69	0.76
	NSE	-	-1.06	-0.60	-0.55	-0.37	-0.20	-0.37	-0.34	-0.28	-0.18
	RMSE	-	2208	1845	1817	1801	1801	1915	1707	1670	1689
Mississippi River at Vicksburg, MS	Average	16660	8390	8334	8335	8339	8341	8329	8324	8335	8337
	ρ	-	0.52	0.71	0.74	0.83	0.84	0.71	0.73	0.80	0.83
	NSE	-	-0.96	-0.24	-0.24	-0.40	-0.46	-0.40	-0.30	-0.28	-0.24
	RMSE	-	12073	10416	10204	9882	9681	9618	9592	9323	9121
Missouri River at Hermann, MO	Average	1983	1229	1197	1199	1199	1199	1198	1197	1198	1198
	ρ	-	0.44	0.62	0.64	0.66	0.76	0.63	0.66	0.77	0.76
	NSE	-	-1.11	-0.32	-0.31	0.06	0.03	-0.31	-0.11	0.10	0.11
	RMSE	-	1989	1490	1489	1302	1294	1489	1367	1281	1261
Missouri River at Omaha, NE	Average	747	507	507	507	507	507	507	507	507	507
	ρ	-	0.38	0.49	0.54	0.54	0.55	0.49	0.51	0.53	0.52
	NSE	-	-11.93	-9.48	-8.30	-8.30	-7.25	-8.74	-8.39	-7.46	-6.95
	RMSE	-	909	777	777	771	771	749	736	735	713
Arkansas River at Murray Dam near Little Rock, AR	Average	1232	748	746	746	747	747	746	746	746	747
	ρ	-	0.40	0.62	0.63	0.73	0.75	0.70	0.70	0.76	0.77
	NSE	-	-0.38	0.18	0.21	0.41	0.44	0.35	0.36	0.44	0.44
	RMSE	-	1377	1180	1163	1060	1032	1052	1041	1031	1027
Red River at Spring Bank, AR	Average	544	360	360	360	360	360	360	360	360	360
	ρ	-	0.38	0.49	0.51	0.68	0.68	0.58	0.62	0.76	0.76
	NSE	-	-0.65	-0.14	-0.07	0.36	0.36	0.15	0.24	0.51	0.57
	MSE	-	862	682	660	539	538	588	555	472	441
Mean	Average	4481		2263	2265	2265	2266	2262	2261	2264	2264
	ρ	-		0.60	0.62	0.69	0.71	0.62	0.68	0.73	0.74
	NSE	-		-1.51	-1.32	-1.16	-1.00	-1.32	-1.20	-0.97	-0.86
	MSE	-		3243	3200	2980	2946	3056	2965	2856	2821

2.4.4. Comparison of the grid-based and vector-based river networks using the VIC land surface model

The statistical results of the RAPID simulations using the VIC land surface model for both river networks are shown in Table 5 and Table 6. The calculated average flow for both river networks show that using the VIC land surface model improves the river routing results. Table 5 and Table 6 show that using the VIC land surface model, the grid-based and vector-based river networks show the comparable statistical results. However, the average flow calculated by the vector river network at the Ohio River at Metropolis station is significantly better using vector river network. The reason is explained in the previous section. Figure 6 plots daily simulated flow by RAPID using two different river networks, with using VIC, (K_{ini}^3 and ϕ_2) and observed flow for 9 years. In this figure, comparisons are shown for three gauging stations: (a) Mississippi River at Vicksburg, (b) Missouri river at Hermann, MO, and (c) Mississippi river at Keokuk, IA. For all stations the model results are slightly different compare to the observations.

2.4.5. Comparison of the simulated streamflow using the VIC and Mosaic land surface models for the vector river network

This section compares the RAPID outputs using VIC and Mosaic land surfaces for the vector river network. Figure 7 depicts the observed and modeled streamflow for VIC and Mosaic using vector river network at three stations: Mississippi River at Vicksburg, MS, Missouri River at Hermann, MO, and Mississippi River at Keokuk, IA. For all stations simulated flow by RAPID using Mosaic is underestimated and flow simulated using VIC is capable of reaching peak flows. The correlation coefficient (ρ) is higher using the VIC land surface model.

Table 5: Statistics of the model outputs and observations for the grid river network using the VIC land surface model

Gauging station	Observation	Lumped Model	Grid-Based River Network								
			ϕ_1				ϕ_2				
			VIC	K_{ini}^1	K_{ini}^2	K_{ini}^3	K_{ini}^4	K_{ini}^1	K_{ini}^2	K_{ini}^3	K_{ini}^4
Ohio River at Metropolis, IL	Average	8051	6413	6375	6374	6376	6376	6392	6385	6397	6397
	ρ	-	0.58	0.86	0.86	0.87	0.89	0.89	0.90	0.91	0.92
	NSE	-	0.02	0.64	0.69	0.65	0.64	0.70	0.71	0.75	0.76
	RMSE	-	5752	3487	3484	3419	3252	2926	2874	2858	2760
Mississippi River at Keokuk, IA	Average	2147	2245	2243	2242	2244	2244	2240	2241	2241	2242
	ρ	-	0.46	0.73	0.76	0.86	0.86	0.86	0.86	0.90	0.91
	NSE	-	-0.66	0.44	0.51	0.73	0.72	0.76	0.80	0.81	0.82
	RMSE	-	1983	1093	1021	769	753	709	659	642	640
Mississippi River at Vicksburg, MS	Average	16660	18081	17987	17951	17989	17993	17929	17926	17906	17923
	ρ	-	0.08	0.93	0.94	0.93	0.92	0.76	0.77	0.87	0.90
	NSE	-	-0.99	0.83	0.83	0.81	0.78	0.52	0.58	0.72	0.78
	RMSE	-	12174	3415	3380	3584	3802	5678	5294	4298	3802
Missouri River at Hermann, MO	Average	1983	3139	3130	3125	3133	3133	3121	3122	3124	3127
	ρ	-	0.49	0.70	0.71	0.75	0.74	0.70	0.80	0.83	0.84
	NSE	-	-3.61	-0.89	-0.71	-0.62	-0.61	-0.67	-0.79	-0.64	-0.63
	RMSE	-	2941	1786	1701	1652	1646	1663	1659	1680	1739
Missouri River at Omaha, NE	Average	747	1362	1359	1358	1360	1360	1355	1356	1357	1359
	ρ	-	0.41	0.55	0.58	0.58	0.59	0.49	0.53	0.53	0.57
	NSE	-	-37.09	-25.72	-21.31	-20.55	-20.17	-20.76	-18.28	-17.62	-17
	RMSE	-	1481	1240	1133	1114	1104	1119	1053	1035	1029
Arkansas River at Murray Dam near Little Rock, AR	Average	1232	1663	1659	1655	1660	1660	1658	1656	1658	1657
	ρ	-	0.52	0.83	0.83	0.83	0.82	0.82	0.80	0.84	0.83
	NSE	-	-0.80	0.46	0.47	0.49	0.49	0.53	0.53	0.58	0.51
	RMSE	-	1754	964	957	931	930	915	895	850	850
Red River at Spring Bank, AR	Average	544	945	945	942	945	945	944	943	944	944
	ρ	-	0.54	0.82	0.83	0.80	0.84	0.80	0.81	0.88	0.88
	NSE	-	-1.18	0.11	0.17	0.19	0.23	0.36	0.34	0.36	0.37
	MSE	-	943	603	582	575	560	560	518	510	508
Mean	Average	4481	4835	4814	4807	4815	4816	4806	4804	4804	4807
	ρ	-	0.44	0.78	0.79	0.80	0.81	0.76	0.78	0.82	0.84
	NSE	-	-6.33	-3.45	-2.76	-2.61	-2.56	-2.65	-2.30	-2.15	-2.11
	MSE	-	3861	1798	1751	1721	1721	1939	1850	1696	1618

Table 6: Statistics of the model outputs and observations for the vector river network using the VIC land surface model

Gauging station	Observation		Lumped Model	Vector-Based River Network							
				ϕ_1				ϕ_2			
				VIC	K_{ini}^1	K_{ini}^2	K_{ini}^3	K_{ini}^4	K_{ini}^1	K_{ini}^2	K_{ini}^3
Ohio River at Metropolis, IL	Average	8051	8379	8331	8346	8344	8338	8314	8327	8331	8319
	ρ	-	0.56	0.77	0.79	0.82	0.88	0.85	0.88	0.90	0.91
	NSE	-	-0.32	0.72	0.77	0.81	0.87	0.61	0.67	0.81	0.87
	RMSE	-	7034	3932	3625	3540	2807	3241	2807	2524	2524
Mississippi River at Keokuk, IA	Average	2147	2462	2452	2455	2456	2455	2450	2453	2455	2452
	ρ	-	0.49	0.79	0.84	0.88	0.88	0.83	0.88	0.88	0.90
	NSE	-	-0.74	0.57	0.61	0.66	0.72	0.73	0.76	0.77	0.75
	RMSE	-	2028	1013	959	887	785	813	760	755	751
Mississippi River at Vicksburg, MS	Average	16660	19684	19526	19534	19531	19534	19490	19487	19491	19493
	ρ	-	0.07	0.88	0.90	0.91	0.90	0.89	0.90	0.92	0.93
	NSE	-	-1.28	0.67	0.67	0.68	0.70	0.68	0.70	0.71	0.71
	RMSE	-	13036	5298	4938	4718	4661	4689	4660	4624	4649
Missouri River at Hermann, MO	Average	1983	3172	3104	3001	3002	3001	3104	3110	3110	3115
	ρ	-	0.49	0.68	0.75	0.76	0.77	0.68	0.75	0.76	0.83
	NSE	-	-3.66	-0.71	-0.66	-0.67	-0.65	-0.69	-0.61	-0.62	-0.45
	RMSE	-	2957	1830	1824	1792	1798	1783	1798	1759	1760
Missouri River at Omaha, NE	Average	747	1419	1414	1416	1416	1416	1411	1414	1416	1414
	ρ	-	0.14	0.48	0.51	0.57	0.51	0.54	0.57	0.57	0.56
	NSE	-	-39.29	-24.54	-24.67	-24.62	-21.25	-22.00	-21.88	-21.97	-19.44
	RMSE	-	1605	1193	1281	1214	1212	1143	1210	1150	1150
Arkansas River at Murray Dam near Little Rock, AR	Average	1232	1717	1713	1714	1714	1714	1712	1713	1715	1713
	ρ	-	0.51	0.80	0.81	0.78	0.82	0.80	0.82	0.82	0.85
	NSE	-	-0.94	0.30	0.30	0.36	0.47	0.43	0.44	0.47	0.49
	RMSE	-	1920	1094	1077	1045	1003	1001	1004	993	983
Red River at Spring Bank, AR	Average	544	804	803	803	803	803	803	802	803	803
	ρ	-	0.57	0.82	0.82	0.84	0.85	0.88	0.89	0.87	0.88
	NSE	-	-0.57	0.43	0.44	0.47	0.52	0.57	0.57	0.58	0.63
	MSE	-	842	489	468	481	477	436	411	413	412
Mean	Average	4481		5335	5324	5324	5323	5326	5329	5332	5330
	ρ	-		0.75	0.77	0.80	0.80	0.78	0.81	0.82	0.84
	NSE	-		-3.22	-3.22	-3.19	-2.66	-2.81	-2.76	-2.75	-2.35
	MSE	-		2121	2025	1954	1821	1872	1807	1745	1747

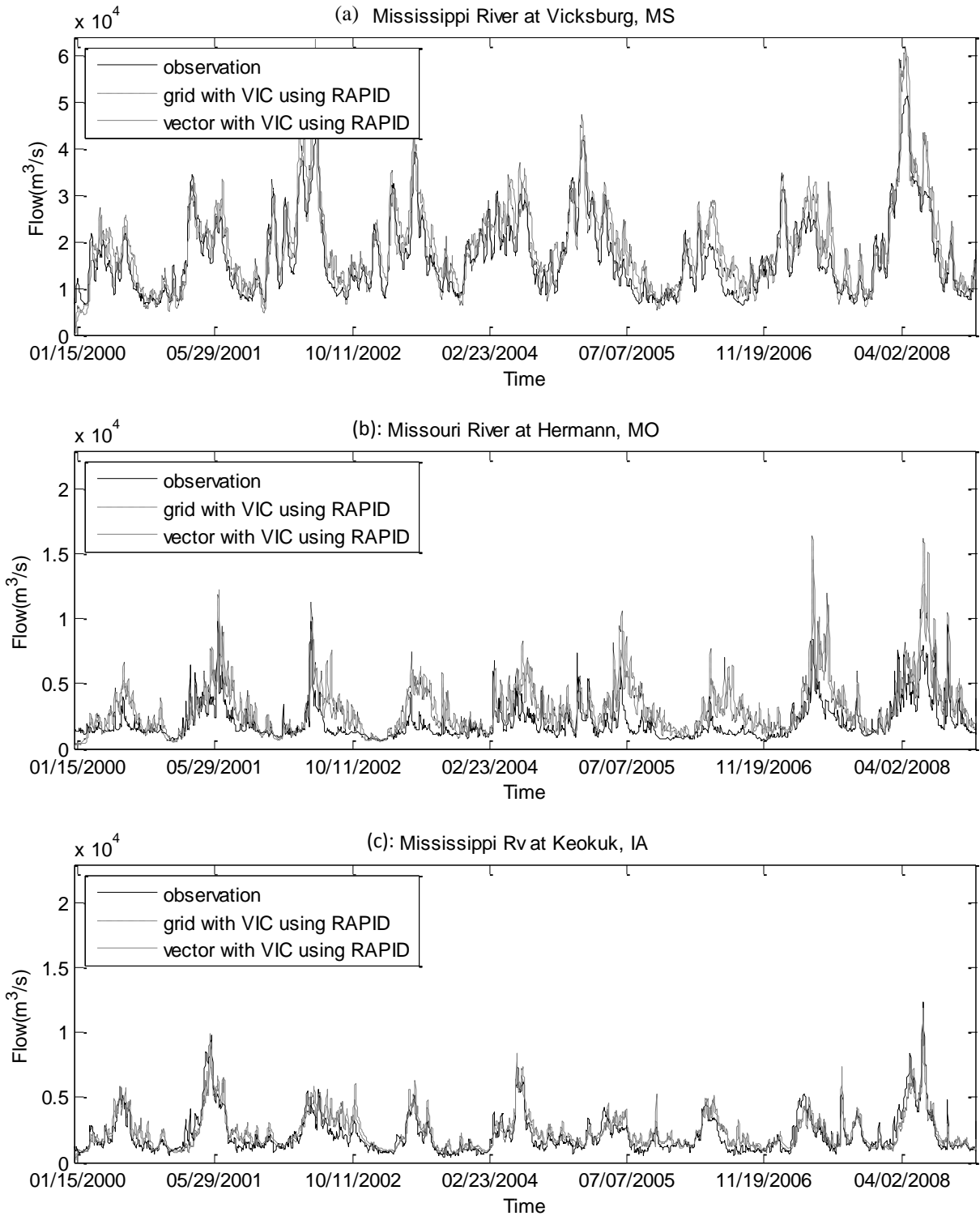


Figure 6: The observed and modeled streamflow using grid-based and vector-based river networks over nine years: (a) Mississippi River at Vicksburg, MS; (b) Missouri River at Hermann, MO and (c) Mississippi River at Keokuk, IA

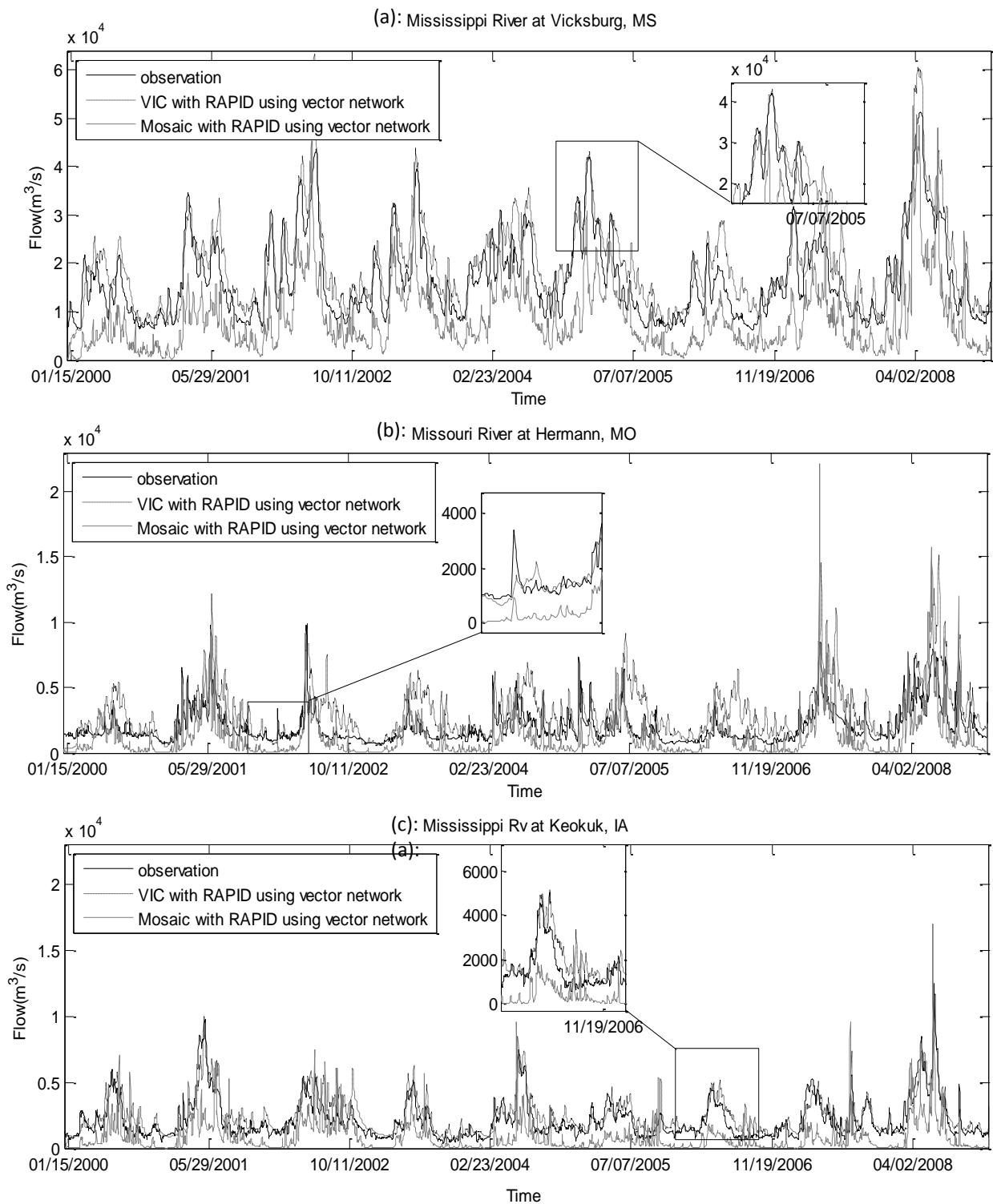


Figure 7: The observed and modeled streamflow for VIC and Mosaic using vector river network over nine years: (a) Mississippi River at Vicksburg, MS; (b) Missouri River at Hermann, MO and (c) Mississippi River at Keokuk, IA

2.4.6. The effect of drainage area on the 9-year mean flow

The 9-year mean flow, using the same land surface model for both river networks is almost constant in all stations except at the Ohia River at Metropolis station. Comparisons of grid-based and vector-based river networks show that the average flow is significantly different at this station. To compare the 9-year flow average, the delineated watershed and respective drainage area for each station is analyzed using DEM and NHDPlus data (Table 2 and Figure 8).

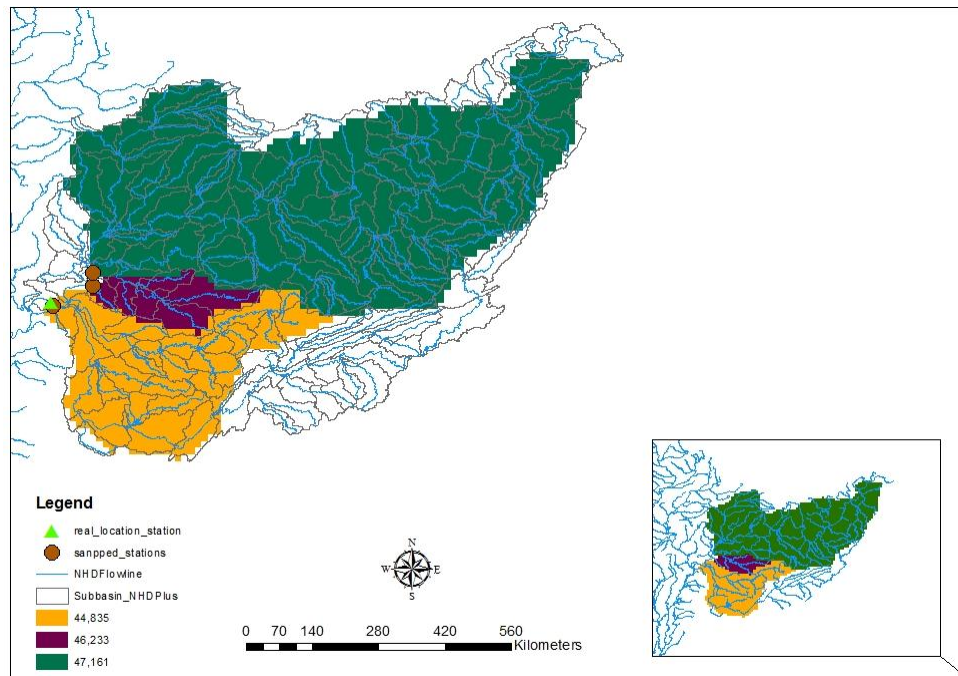


Figure 8: Watershed delineation versus NHDPlus catchment at the Ohio River at Metropolis station

At the Metropolis station, the vector-based river network produced 26% and 30% larger mean flow than the grid-based river network using the Mosaic and VIC land surface models respectively. This is likely the result of the NHDPlus dataset dedicating a

30% larger basin (523,498 km²) to this station, than that which the DEM dedicated (401,383 km²). The actual basin, based on the USGS measurement (Table 2) suggests the *NHDPlus* dataset is more accurate. In addition, to get a delineated watershed using a DEM at this station, three points should be snapped. Figure 8 also shows the real and snapped locations of Metropolis station. For other stations the 9-year mean flow calculated for both networks are similar to each other because the drainage area calculated for grid and vector-based networks are very comparable. However; using *NHDPlus* dataset one can determine drainage area with higher accuracy.

2.4.7. Comparison of mean annual flow for grid-based and vector-based river networks using VIC and Mosaic land surface models

Mean annual flow for all stations (except Red river at Spring Bank, AR) for two river networks and land surface models are compared with observed mean annual flow (Figure 9). For all stations simulated-mean annual flow using VIC shows a better performance of the RAPID model. Considering river networks, the vector-based river network produces mean annual flow with higher accuracy than the does the grid-based river network at the Ohio River at Metropolis station. For other stations the simulated mean annual flow for both river networks are slightly different.

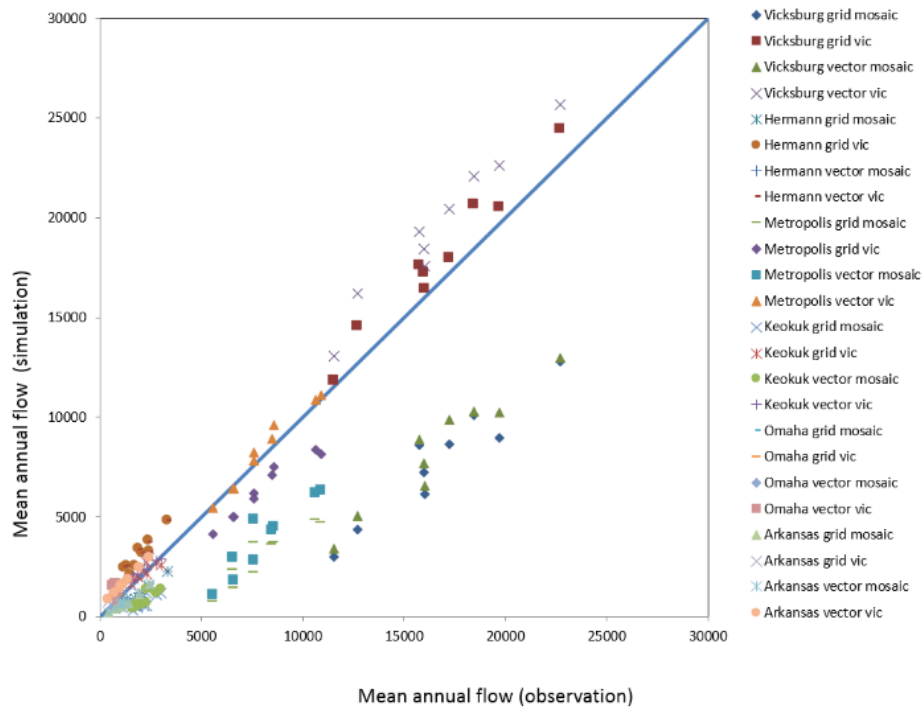


Figure 9: Comparison of observed and calculated mean annual flow for all stations

2.5. CONCLUSIONS

We simulated and compared flow rates over the entire Mississippi river basin using gridded and vector river networks. For this purpose a river routing model called RAPID was run with runoff from Mosaic and VIC for a 9-year period from 2000 to 2008. The gridded river network was generated from the 1/8 degree topography data and the total number of grid cells is 58,140 with an average grid length of 13 km. The vector river network was derived from the *NHDPlus* dataset. The vector river network has 211,476 blue lines with an average river length of 1.86 km. The model results were compared against observed daily flow retrieved from the USGS.

The summary statistics in Tables 3 and 5 indicate that, coupling a land surface model and a river routing model significantly improves large scale river modeling

compare to a land surface model only. Furthermore, various experiments to optimize the “K” factor showed that taking channel geometry into account is a key parameter in the model optimization. The results for the spatial variation of the wave celerity (K_{ini}^3 and K_{ini}^4) reveal an improved correlation with observations compared to constant wave celerity (K_{ini}^3 and K_{ini}^4). The results also display a decrease in RMSE for spatially varied wave celerity for all river reaches. Two cost functions were used in the optimization procedure. The second cost function, which assigns the same weight to all gauges, subtly increases the efficiency and decreases the RMSE in both river networks and land surface models. This finding is in agreement with a previous study by *David et al.*, [2011a].

Both river networks mimic observed hydrographs in the selected gauges and are able to capture peak flow. However; the location of river gauges is more easily determined within the vector river network (e.g. no need for snapping) hence, one can use as many gauges as needed. The 9-year averaged flow and visual comparison of simulated and observed hydrographs show the model underestimation for both river networks compared to observations using Mosaic. Whereas, application of the VIC land surface model significantly improve the RAPID performance.

Comparison of the 9-year mean flow also indicates that a drainage area is a key factor, especially in a wetter climate. For example in a wetter climate, the vector river network produced 26% and 30% more 9-year averaged flow than the grid-based river network at the Ohio River, Metropolis Station using Mosaic and VIC. For other stations, where watershed areas are comparable the differences between the gridded and vector approaches are small.

In this study, the gages were selected based on the literature and on the data available for the study period. Therefore, using more gages for comparison would help improve the results.

Chapter 3: A GIS framework for regional modeling of riverine nitrogen transport: Case study, San Antonio and Guadalupe basins

3.1. ABSTRACT

While anthropogenic nitrogen inputs are recognized as the primary driver of coastal eutrophication world-wide, modeling these inputs remains challenging. This paper presents a GIS framework for integrating a regional GIS-based nitrogen dataset (Texas Anthropogenic Nitrogen Dataset, TX-ANB) and a GIS-based river routing model (Routing Application for Parallel computation of Discharge, RAPID) to simulate steady-state riverine total nitrogen (TN) transport in river networks containing thousands of river reaches. A two-year case study was conducted in the San Antonio and Guadalupe basins with thousands of river reaches during wet and dry years (2008-09). This paper investigates total nitrogen export in more urbanized and rural drainage basins (San Antonio and Guadalupe) and considers the effect of reservoirs on TN transport.

Results show that in both years the San Antonio basin contributed a larger quantity of delivered TN than the Guadalupe basin. The reason is that the San Antonio basin is affected by urban activities including point sources (PS), in addition to the agricultural activities. Both basins delivered more TN to the Gulf of Mexico in 2009 than in 2008. Furthermore, TN removal in the San Antonio and Guadalupe basins is inversely related to stream orders: the higher the order the less TN is removed.

3.2. INTRODUCTION

Nutrient transport from the atmosphere, across the land, and to the coast is tightly coupled to climate and land cover/use characteristics. Human modifications of the environment such as reservoir building also have a major effect on this tightly-coupled process. Human impacts on nitrogen transport from watersheds to coastal waters are of particular interest, because productivity of estuarine ecosystems is often limited by availability of this nutrient. Unexpected interactions between land use and climate change can alter the amplitude, frequency, and duration of nutrient pulses in streams and rivers [Kaushal *et al.*, 2010]. Human activity including the cultivation of N-fixing crops, fossil-fuel burning, the discharge of industrial and domestic effluents, and extensive use of fertilizers accelerate the increase of nitrogen load to watersheds. These activities lead to increased riverine fluxes to coastal areas and lakes [Zhang *et al.*, 2010]. Food and energy production has increased nitrogen abundance by more than a factor of ten since the late-19th century [Galloway *et al.*, 2004].

Excessive nutrients generated by natural and anthropogenic activity lead to eutrophication and adverse environmental consequences such as hypoxia and reduced fish and shellfish production [Schaefer and Alber, 2007; Boesch *et al.*, 2009]. Eutrophication has adversely affected 40% of the estuarine resources in the U.S. [Howarth *et al.*, 2006]. While hypoxia along the inner continental shelf of the Gulf of Mexico has become nationally important, localized hypoxia and degradation of aquatic resources in bays and estuaries along the Texas coast is becoming more significant [Rebich *et al.*, 2011]. The consequences of overabundant nutrients on riverine export emphasize the necessity to improve our understanding of nitrogen transport in large coastal watersheds [Alexander *et al.*, 2002].

Existing water quality models range from sophisticated deterministic methods to simple export coefficient approaches. Deterministic models describe nitrogen transport, attenuation, and transformation in detail. As a result, they capture spatial and temporal variations in sources and sinks present in watersheds. Growing computational resources encourage researchers to develop and to apply such models (e.g. SWAT, [Arnold *et al.*, 1998]; HSPF, [Bicknell *et al.*, 2005]; RIVERSTRAHLER, [Billen *et al.*, 1994]; INCA, [Whitehead *et al.*, 1998]; AGNPS, [Young *et al.*, 1989]). However, the complexity of deterministic models creates large amounts of data that requires sophisticated calibration processes. Furthermore, computational demands and uncertainty in model coefficients also pose potential problems for modeling large-scale river networks [Alexander *et al.*, 2002; Liu *et al.*, 2008; Huang *et al.*, 2009; Marcé and Armengol, 2009].

Lack of water quality data and the weakness of uncertainty measurements prevent the implementation of complex deterministic models [Marcé and Armengol, 2009]. Under these conditions, a simple water quality modeling approach such as an export coefficient method can be applied to estimate the total annual load from a river basin. Export coefficient approaches predict total annual nutrient loads primarily as a function of land-use types [Johnes, 1996]. The export coefficients, which represent the percentage of total load delivered to drainage networks, are obtained from literature. Furthermore, the flow regime in the export coefficient method is estimated using regression methods and databases such as the National Hydrography Dataset Plus (NHDPlus) (e.g. [Yang *et al.*, 2014], [Johnson *et al.*, 2013]) to represent weather changes from year to year; The NHDPlus dataset provides mean annual velocity and mean annual flow for all river reaches averaged over multiple years. The regional river routing model simulation has not been previously applied to the export coefficient method. Therefore, the accuracy of

the export coefficients approach to consider climatologic and physiographic characteristics of a specific watershed with thousands of river reaches needs to be improved. This study uses the river routing model called RAPID [David *et al.*, 2011b] to calculate mean annual flow and velocity for 2008 and 2009 using daily simulated data. The RAPID model uses the *NHDPlus* dataset to represent the river network and computes water flow and volume for all *NHDPlus* flowlines.

In this study we introduced a GIS framework to integrate a regional GIS-based nitrogen dataset (Texas Anthropogenic Nitrogen Dataset, TX-ANB) and a GIS-based river routing model (Routing Application for Parallel computation of Discharge, RAPID) to simulate riverine nitrogen load in river networks containing thousands of river reaches. This project is a two-year study (2008-2009) of total nitrogen (TN) load variation in urban and rural regions on all river reaches in the San Antonio and Guadalupe basins; 2008 was a drier year and 2009 was a wetter year.

3.3. METHODS

Given the stated goal, a GIS modeling framework has been established to estimate nitrogen load for the San Antonio and Guadalupe drainage basins. Figure 10 shows the framework, which uses an in-state dataset and models. This research builds upon the river routing model RAPID that was developed by David [2009] and incorporates nitrogen inputs from the TX-ANB dataset. The core of the GIS modeling framework is the schematic processor [Whiteaker *et al.*, 2006; Johnson, 2009], which works with a schematic network. The schematic network is a system of links and nodes that represent hydrologic features and their connectivity in a watershed. The schematic processor takes nitrogen inputs from the TX-ANB nitrogen budget and receives flow from the RAPID model to route the nitrogen transport through a river network. Wastewater treatment plant

and reservoir information are also added to the schematic network. The calibration process is performed within the schematic processor by comparing modeling outputs with measured total nitrogen data. In the final step, results are shared as an Esri web map service and published in ArcGIS online. Details regarding the modeling framework and dataset are described below.

3.3.1. The NHD*Plus* dataset

The major dataset used in this project is NHD*Plus* version 2, which provides the vector river network for the RAPID model and schematic network. NHD*Plus* coherently describes topography and hydrographic features for 21 regions that cover the entire USA [NHD*Plus*; Horizon Systems Corporation 2007]. The NHD*Plus* database contains mapped streams and river reaches as well as the catchments that surround them. A river reach is a segment of a river and a catchment is the incremental drainage area for a given reach. NHD*Plus* provides significant spatial detail and a realistic representation of stream networks based on the NHD (1:100,000 scale), improved networking, feature naming, and “Value-Added Attributes” (VAA). Value-Added Attributes (VAA) include capabilities for modeling and upstream and downstream navigation. Applications include the ability to determine all upstream flowlines and to calculate cumulative drainage areas for a given flowline. Moreover, NHD*Plus* is an application-ready set of geospatial data products, integrating the medium-resolution NHD, the National Elevation Dataset (NED), and the National Watershed Boundary Dataset (WBD) [Simley and Carswell, 2009]. NHD*Plus* version 2, released to the public in June 2012, includes several improvements over version 1.

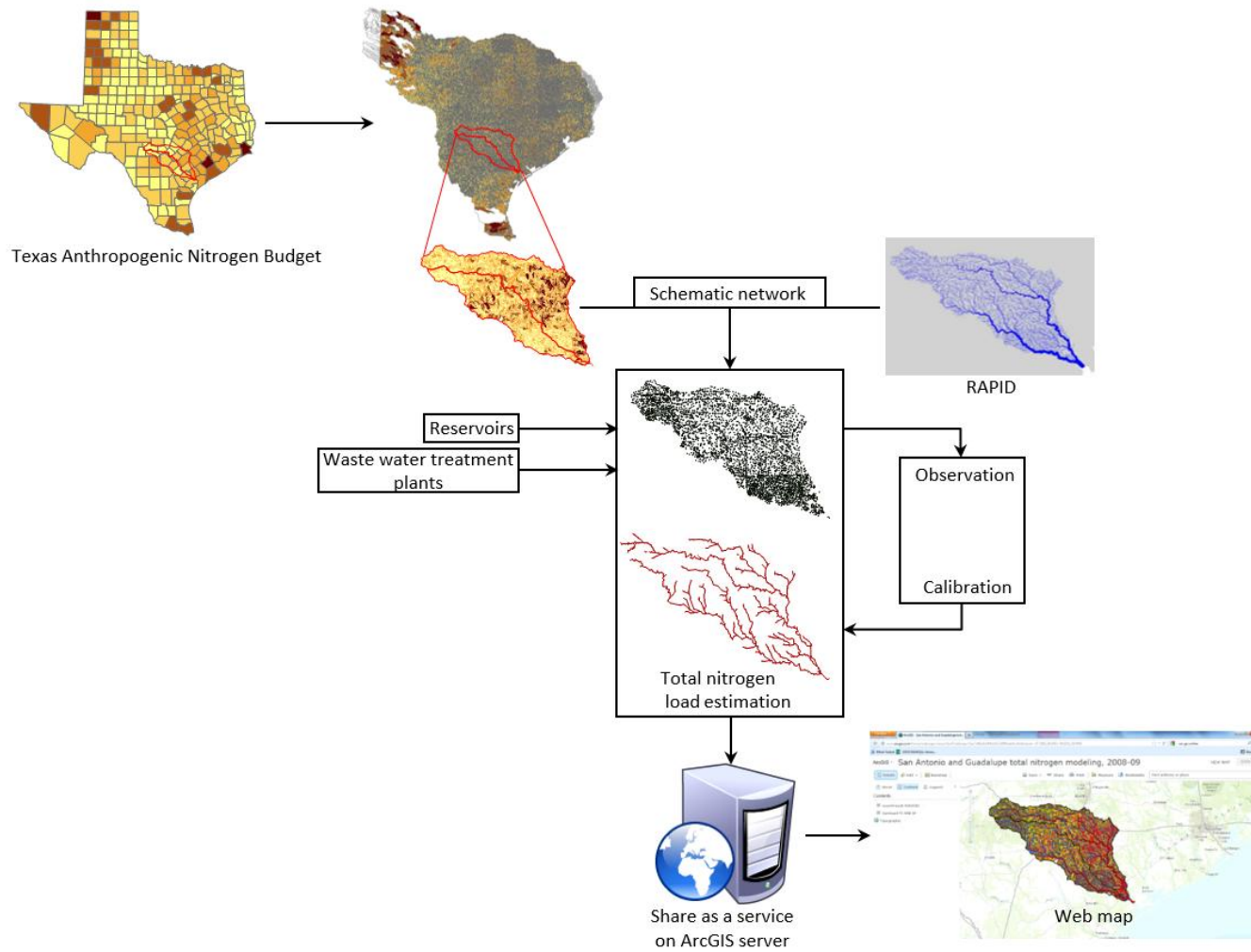


Figure 10: GIS-based modeling framework

In the second version isolated networks are connected. For example, the San Antonio and Guadalupe basins contain 5,193 river reaches in *NHDPlus* v.2 and 5,175 river reaches in *NHDPlus* v.1. The names of water body and lake features have been added to version 2. More details about improvements can be found in *McKay*, [2012]. In this study we make use of improvements in the second version of *NHDPlus* dataset to create the vector river network.

3.3.2. The river routing model, RAPID

River flow in the San Antonio and Guadalupe basins is simulated by a river routing model called (RAPID) [*David*, 2009; *David et al.*, 2011b]. Advantages of RAPID include the use of “blue lines” to represent the river network and the ability to run it on a parallel computing machine (<http://www.ucchm.org/david/rapid.htm>). The RAPID model was first developed using *NHDPlus* river networks which is available for the continental United States. This model couples with land surface models to obtain surface and subsurface runoff and uses a matrix-based version of the Muskingum method to compute water flow and volume everywhere throughout river networks

In addition to the San Antonio and Guadalupe basins, RAPID has been applied to the Texas Gulf Coast Hydrologic region [*David et al.*, 2013], to the river network of SIM-France [*David et al.*, 2011a], and to the Mississippi basin for both grid and vector river networks. For the San Antonio and Guadalupe basins, the RAPID model uses the real-time water decision support system developed by the University of Illinois at Urbana Champaign. The real-time water decision support system is used in this study to run RAPID. The resulting daily outputs are used to calculate an annual average for all river reaches in the San Antonio and Guadalupe basins. The real-time modeling system downloads surface and subsurface runoff from land surface models and runs RAPID to

predict discharge of water. Visualization of the RAPID simulation as a web application enables decision makers to define different climate scenarios and understand the impact of extreme events such as flooding and drought on river networks (Accessed: <http://rapid.ncsa.illinois.edu:8080/rapid/>).

3.3.3. Schematic Processor

The schematic processor developed by *Whiteaker et al.* [2006] broadens the implementation of Arc Hydro from a pre-processing basis to a framework for performing hydrologic modeling. The schematic processor operates by associating a behavior type (RECEIVE or PASS) to a schematic network to pass information through a watershed and to move water or pollutants downstream. The schematic network, which is created by the Arc Hydro toolset, represents hydrologic features as a network of links and nodes [*Maidment, 2002*]. SchemaNodes show hydrologic features such as catchments or stream junctions and SchemaLinks describe the connections between nodes. RECEIVE behavior processes values received from upstream features plus incremental values. PASS behavior processes a total value for a given feature to generate the value that is passed to the next downstream feature of the network. In addition, the schematic processor has a capability to apply additional programming to pass and/or receive values [*Whiteaker and Johnson, 2012*].

The schematic processor has been used to calculate bacterial loads in Copano Bay, Texas by *Johnson et al.*, [2013]. They developed a GIS-based steady-state total maximum daily load (TMDL) balance model to address the majority of Texas's bacteria TMDLs. The merit of the modeling approach used in their study is to simplify hydrologic and water quality processes and to reduce the data and time resources required for model development. Following their approach, we apply the schematic processor to compute

total annual nitrogen loads produced by nonpoint sources (NPS) plus wastewater treatment plants (WWTPs), which are point sources (PS). The accumulation of upstream nitrogen load and decay of loads through networks are behaviors illustrated by schema nodes and links. Figure 11 shows the implementation of a schematic processor for a simple river-catchment network. The red and green nodes represent the starting points of every stream and sub-watershed centroid, respectively. Bold blue and dashed green links exhibit the connection between stream and sub-watersheds (Figure 11a). Decaying nitrogen loads associated with links prior to passing to the next downstream features are considered a PASS behavior, which produces a ‘passed’ value. The ‘passed’ value is then equal to the ‘total’ value received from upstream nodes minus the amount of decay (Figure 11b).

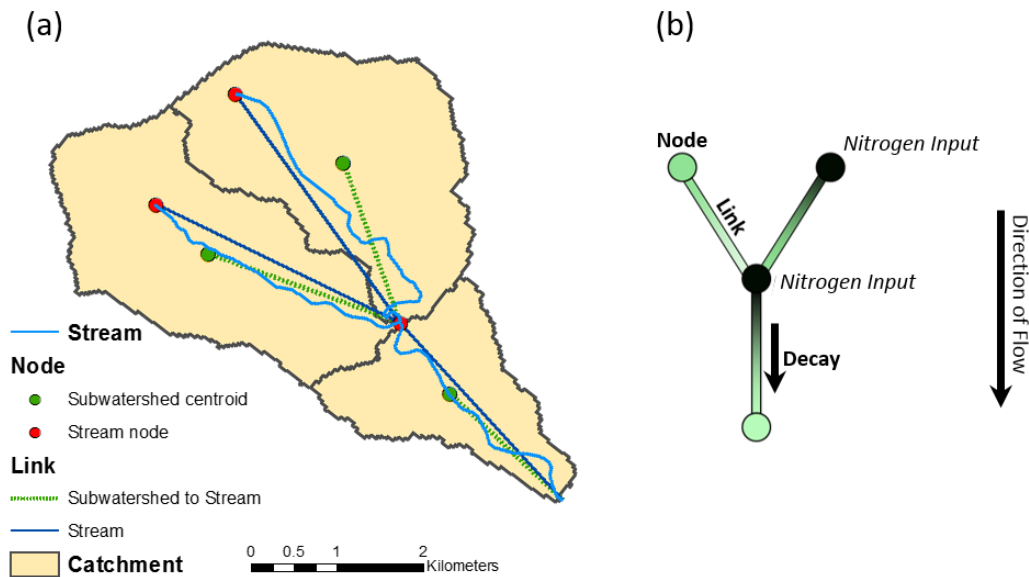


Figure 11: (a) An example of three elemental catchments in a schematic network including links and nodes; (b) A simple schematic network showing how to manipulate a function with a Schemalink

3.3.4. Quantifying the effect of agriculture and urbanization on the nitrogen cycle across Texas

Anthropogenic land use increases nutrient loading from the land to coastal areas by way of rivers [Atasoy *et al.*, 2006; Mayorga *et al.*, 2010]. Hong *et al.*, [2011] developed a Net Anthropogenic Nitrogen Input toolbox (NANI toolbox) that quantifies all anthropogenic nitrogen sources into a national N budget. This budget uses methods developed by Boyer *et al.*, [2002]. The national budget is useful as a large-scale look at N inputs from human influences [Hong *et al.*, 2011]. However, concerns arise when applying this method to some smaller scale river basins. One major problem with the NANI toolbox in Texas basins results from the Net Food and Feed calculation, which accounts for livestock and human consumption processes as a net input. This net source leads to large negative inputs in some basins, which is not likely to be accurate. The extensive livestock industry in Texas creates the need to ensure that the livestock population does not cancel out N inputs from other sources. In addition, enhancing a national dataset (NANI) with regional input ensures highly accurate land use and land coverage and legume fixation rates within Texas. Texas has the third-largest urban area in the United States, and is home to one of the ten largest agriculture industries in the nation. Quantification of all Texas nitrogen sources was recently described [Meyer, 2012]. This nitrogen budget for Texas is called “Texas Anthropogenic Nitrogen Budget” or “TX-ANB”. Due to the limitation of data availability, TX-ANB was constructed at county-level resolution for two years (2008 and 2009). Meyer [2012] quantifies sources and inputs of N within Texas based on datasets of fertilizer input, crop cultivation area, livestock population, and atmospheric deposition.

The TX-ANB does not account for nitrogen contributions by feed imports; it assumes the N from food imported for agriculture is described by accounting for animal

manure N estimates. Imported food for human consumption is considered as leaving the system. Human waste is sent to wastewater treatment facilities where it is treated, thus leaving the system. However, leakage from septic tanks is not considered; it is a point source that is difficult to estimate on a state scale.

To integrate TX-ANB with the *NHDPlus* dataset, the nitrogen budget for *NHDPlus* catchments needs to be determined. GIS is utilized to scale the TX-ANB dataset to the catchment scale. ArcGIS “Intersect” and “Dissolve” tools are applied to estimate the nitrogen budget at the catchment scale. The “Intersect” tool computes the area of the portion of *NHDPlus* catchment feature that overlaps with the Texas county feature. The total nitrogen load for every portion is computed by multiplying the area and total nitrogen load per unit area of county obtained from the TX-ANB dataset. Once the total nitrogen load is computed for portions of catchments, the “Dissolve” tool is run to aggregate catchment areas and total nitrogen. Figure 12 shows the GIS processor of computing total nitrogen load for the *NHDPlus* catchment feature. This process was applied to the entire state of Texas; however, this paper focuses on the 5,195 catchments of the San Antonio and Guadalupe basins.

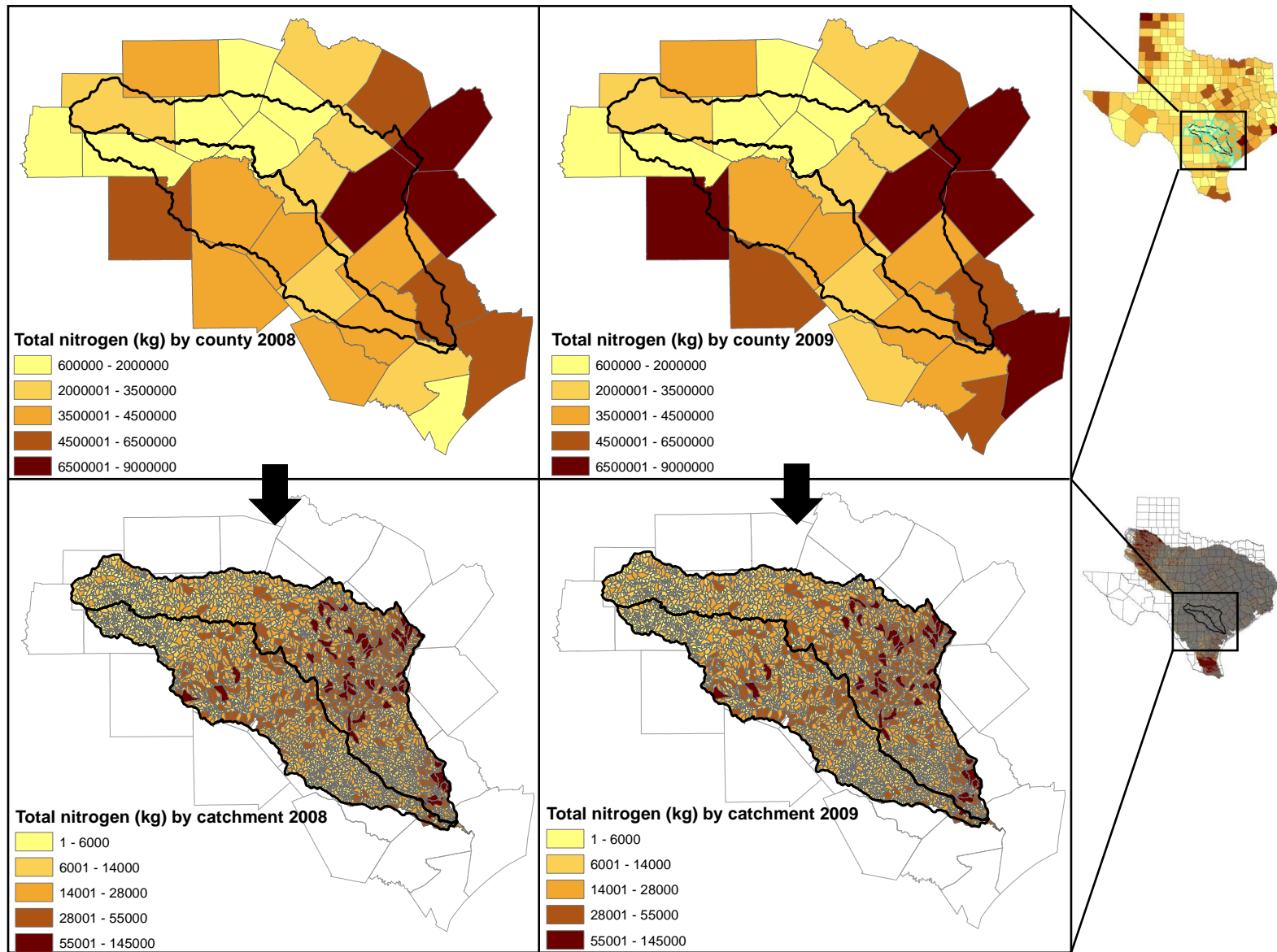


Figure 12: Assigning nitrogen load to catchments on the basis of county estimates

3.3.5. Nitrogen load removal in reservoirs and lakes

Reservoirs play a significant role as nitrogen sinks, as compared to other in-stream nitrogen retention processes [Garnier *et al.*, 2000]. The amount of N removed in lakes and reservoirs (TN_{rem} : $kg\ yr^{-1}$) can be estimated as follows:

$$TN_{rem} = TN_{in} \times R \quad (13)$$

where TN_{in} is total nitrogen load inflow to lake and reservoir surface waters and R is the fraction of total nitrogen retained within lakes and reservoirs. The removal fraction (R) is calculated as [Harrison *et al.*, 2009]:

$$R = 1 - \exp\left(\frac{-V_f \times A}{Q \times 1000}\right) \quad (14)$$

where V_f is the apparent settling velocity for TN ($m\ year^{-1}$) in lake or reservoir sediments determined by Harrison *et al.*, [2009] to be $8.91\ (m\ year^{-1})$, Q is inflow to lakes and reservoirs ($km^3\ year^{-1}$), and $A\ (km^2)$ is the surface area of individual lakes (for large lake analysis).

Reservoirs and lakes are added to the schematic network as nodes associated with equation (14) and the nitrogen load delivered to the downstream river reach (TN_{out}) is determined as follows:

$$TN_{out} = TN_{in} - TN_{rem} \quad (15)$$

3.3.6. Modeling approach

The model framework presented in this work uses a steady-state mass balance that relies on the first-order decay rate to transport pollution within river reaches and accounts

for nonpoint and point sources. In the modeling approach, stream discharge is assumed to be uniform, uptake occurs primarily in the main channel, dispersion is negligible, and solutes are transported by advection. Based on these assumptions, the equation describing pollutant travel through a plug flow is expressed as [Chapra, 1997; Runkel, 2007]:

$$V \frac{dC}{dx} = -kC \quad (16)$$

The analytical solution is:

$$C_x = C_0 e^{-\frac{k}{V}x} \quad (17)$$

The load is calculated by multiplying both sides by a discharge:

$$L_i = L_0 e^{-\frac{k}{V}x} \quad (18)$$

where V is an average velocity over a cross section (m/yr), C_x is the constituent concentration (amount/m³), x is distance (m), C_0 is the initial concentration (amount/m³), C_x , is the concentration at distance x (amount/m³), k is the decay rate (yr⁻¹), L_0 is the initial load (kg yr⁻¹), and L_i is the load at a distance x (kg yr⁻¹).

The initial load includes nonpoint and point source nitrogen loads. The TX-ANB dataset provides the nonpoint source for each *NHDPlus* catchment. Point source pollutants come from wastewater treatment plants and industrial facilities. The following equation shows the initial load input to streams:

$$L_0 = \theta * (NB)_i * A_i + PS \quad (19)$$

where θ is the ratio of TN loading to rivers; NB is the nitrogen budget per area; A_i is the catchment area; and PS is the point source.

Once the L_0 is computed for river reaches, it is plugged into equation (18) to transport the TN load along the schematic network.

3.4. CASE STUDY: IN-STREAM NITROGEN LOAD FOR THE SAN ANTONIO AND GUADALUPE BASINS

The river routing model is coupled with the nitrogen dataset for the San Antonio and Guadalupe basins. Output flow from the RAPID model and the TX-ANB nitrogen budget are applied to annual nitrogen modeling in the San Antonio and Guadalupe basins.

3.4.1. Study area

The San Antonio and Guadalupe basins cover an area of 26,222 km² in south-central Texas. The basins extend from Kerr and Gillespie counties toward the Gulf of Mexico. The attribute table of the *NHDPlus* flowline feature shows that the San Antonio and Guadalupe basins have a total of 5,193 river reaches with an average length of 3.00 km and an average catchment area of 5.08 km². The study area is shown in Figure 13. The Guadalupe basin, which can be considered a relatively rural region, is 45% agricultural, 52% forest, 2% water, and 2% urban. The San Antonio basin, the “more urban” case, is 41% agricultural, 48% forest, 2% water, 8% urban, and 1% barren. Seven reservoirs and lakes are present within the San Antonio and Guadalupe basins: three lakes in the San Antonio basin and four lakes and reservoirs in the Guadalupe basin. The removal fraction within reservoirs ranges from 95% for Braunig Lake to less than 5% for Lake Gonzales; this range is due to the ratio of inflow to lakes and the surface area of

lakes. Lake and reservoir data are provided by the Texas Water Development Board, 2012.

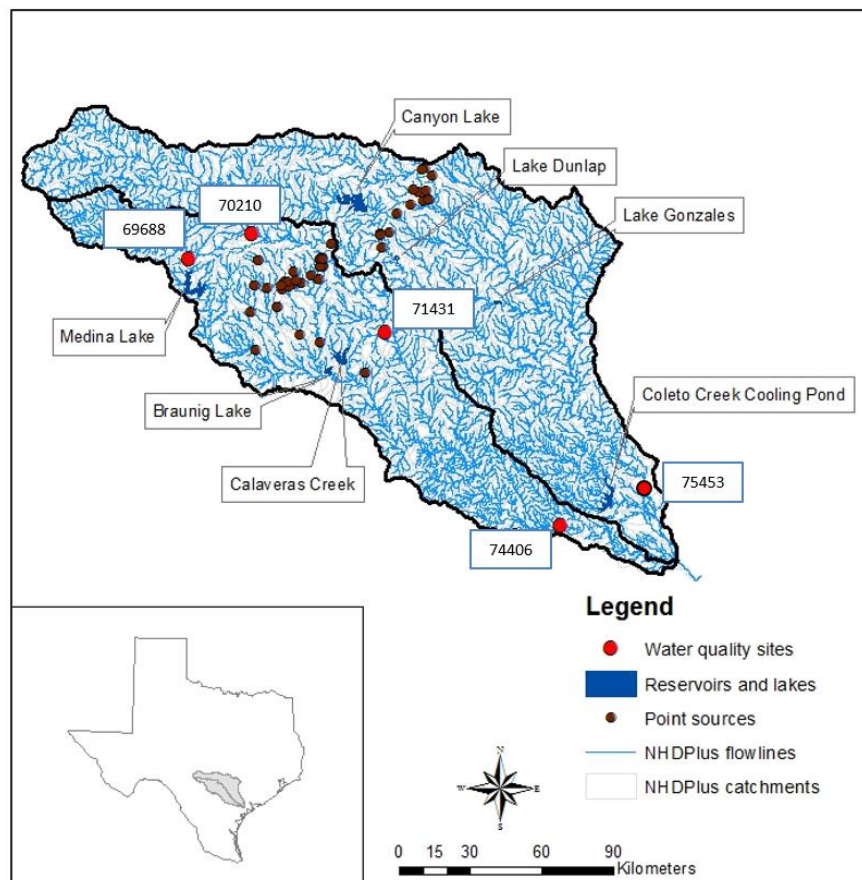


Figure 13: San Antonio and Guadalupe basins

The modeling framework for this case study is unique in that the RAPID output is integrated with the TX-ANB dataset based on the NHDPlus river network. Mean annual flow is produced by RAPID for the San Antonio and Guadalupe basins for 2008-09. Modeled stream flows were calibrated by comparison to flows measured at 36 gauging stations within San Antonio and Guadalupe basins [David *et al.*, 2011b]. The simulated flows were then used to calculate mean annual flow and velocity for NHDPlus flowlines

in the study area. In this study the mean annual velocity is based on the measured river width [Kiel and Cardenas, 2014] and was computed using the Leopold law [Leopold and Langbein, 1962].

3.4.2. Waste water treatment plants

Waste water treatment plants (WWTPs) and industrial facilities are considered point sources (PSs). Location and data regarding annual TN load from these facilities are obtained from the Discharge Monitoring Report (DMR) Pollutant Loading Tool (Accessed at: <http://cfpub.epa.gov/dmr/>). Point sources contribute substantially less to the TN load in the Guadalupe basin due to the smaller percentage urban area in this basin. Among cities in the study area, San Antonio, San Marcos, New Braunfels, and Kyle have the most PSs, respectively. Point sources are assigned to the appropriate SchemaNodes of catchments based on their spatial locations in the watershed. The 2-digit Standard Industrial classification (SIC) is used to categorize point sources (Accessed: <https://www.osha.gov/pls/imis/sicsearch.html>) as shown in Table 7. Sanitary services including sewage systems contribute the most TN: the average annual load from sanitary services was 4.4 tons in 2008 and 3.8 tons in 2009 for the two rivers combined. The percentage of non-point source (based on the TX-ANB dataset) and point source loads (based on the Waste water treatment plants data) for the San Antonio and Guadalupe basins are shown in Table 8. Point source loads are 14% of total load in the San Antonio basin and about 5 percent of the total nitrogen load in the Guadalupe basin.

Table 7: Summary of point source data

SIC Description	2008		2009	
	Average Concentration	Average Pollutant Load	Average Concentration	Average Pollutant Load
	(mg/L)	(kg/yr)	(mg/L)	(kg/yr)
Sanitary Services	1.54	4434.35	1.44	3859.98
Fishing, hunting and trapping	0.19	834.04	0.04	87.86
Heavy construction except building	0.34	11.09	0.33	17.96
Educational services	1.43	80.20	1.43	38.28
Service industrial machinery	2.38	0.07	2.37	0.40

Table 8: Percentage of point source and nonpoint source data in the San Antonio and Guadalupe basins

Basin	Non-point sources (%)	Point sources (%)
San Antonio	86	14
Guadalupe	95	5

3.4.3. Schematic network

A first-order decay rate equation is assigned to the schematic network. The schematic network is a framework in which the schematic processor applies equations 18 and 19 to simulate nitrogen transport through the San Antonio and Guadalupe basins. Arc Hydro tools including attribute and network tools [Maidment, 2002] are used to build a schematic network for the San Antonio and Guadalupe basins as shown in Figure 14. The following attribute tools are used: *Assign HydroID*, *Generate From/To Node* for Lines, *Find Next Downstream Junction*, and *Store Area Outlets*. The final step involves the use of the Network tools, *Node/Link Schema Generation* to generate schema node and schema link types 1 and 2. Point sources and Lakes/reservoirs are added to the network

by manually adding types 3 and 4 separately. SchemaNode types 1 and 2 indicate the centroid of catchments and endpoints of NHDPlus river reaches. Catchments are connected to rivers by SchemaLink type 1. River segments are connected by SchemaLink type 2. SchemaLink type 3 connects point sources to the nearest downstream junction point. The fraction of nitrogen retained in lakes/reservoirs is determined by applying equation 14 to the outlet river reach of the reservoir. The remaining nitrogen load is transferred to the nearest downstream junction, via SchemaLink type 4.

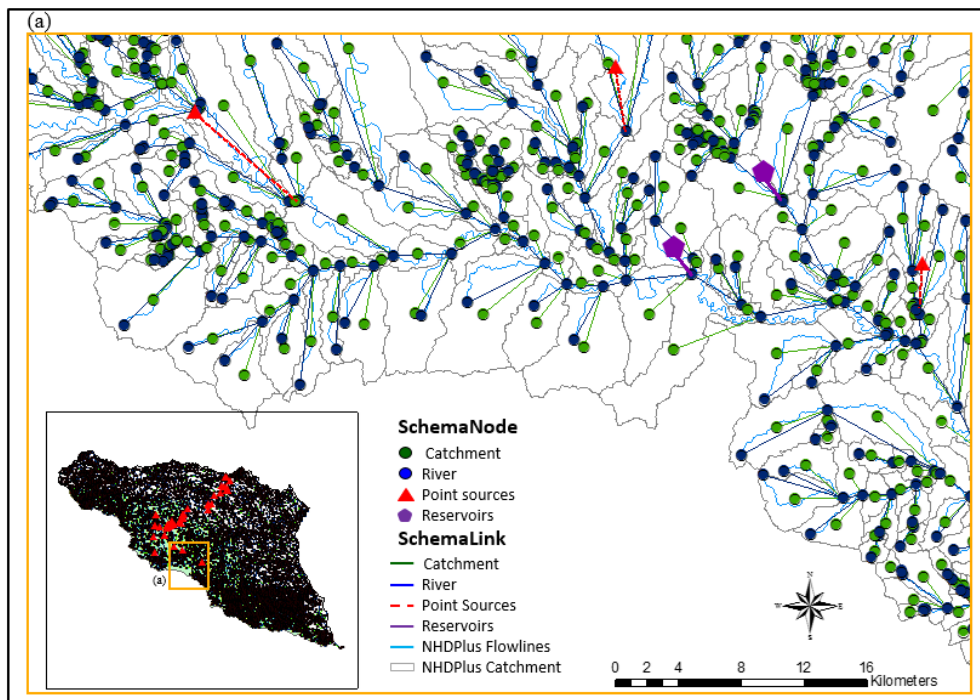


Figure 14: Schematic network for the San Antonio and Guadalupe basins

3.5. RESULTS

3.5.1. In-catchment TN

Total nitrogen accounted for by TX-ANB within the San Antonio and Guadalupe basins is represented in Table 9. The nitrogen budget for the San Antonio basin increased by two percent from 2008 to 2009, whereas in the Guadalupe basin the nitrogen budget declined slightly (0.5%) between the two study years.

Furthermore, based on the TX-ANB dataset, the in-catchment total nitrogen per square kilometer is ~20% higher in the Guadalupe basin than the San Antonio basin because the percentage agricultural land use is greater for the Guadalupe basin.

Table 9: San Antonio and Guadalupe basins

Basin	Area (km ²)	TX-ANB (kg/ km ²)		RAPID mean annual flow (m ³ /s)		Reservoirs and lakes	Reservoir surface area (km ²)
		2008	2009	2008	2009		
San Antonio	10,866	1,361	1,390	0.08	0.32	Braunig Lake	5.16
						Calaveras Lake	12.92
						Medina Lake	21.96
Guadalupe	15,480	1,741	1,732	0.10	0.41	Canyon Lake	32.60
						Coletto Creek reservoir	11.29
						Lake Dunlap	1.05
						Lake Gonzales	1.17

3.5.2. Calibration process

Five monitoring stations are used to calibrate the total nitrogen load (Figure 13). Monitored water quality data are obtained from University of Texas Marine Science Institute [Mooney and McClelland, 2012] and the Texas Commission on Environmental Quality (Accessed: <http://www.tceq.texas.gov/waterquality/clean-rivers/data/samplequery.html>). Land-to-water delivery ratio and reach decay coefficients

are considered in model calibration. The decay coefficient can vary by stream size [Alexander et al., 2000; Rebich et al., 2011]. Therefore the in-stream loss coefficient is estimated for two stream classifications: the first category includes streams with flows $\leq 1.4 \text{ m}^3/\text{s}$ (91% of streams in the region) and the second category is made up of streams with flow $> 1.4 \text{ m}^3/\text{s}$ and $\leq 28.3 \text{ m}^3/\text{s}$ (the rest of streams in the region). Total estimated loads are compared with measured values at selected stations and the TN load difference is calculated (Table 10). The TN load difference between estimated and observed values is less than ten percent for all stations which shows a good fit. The decay coefficient of the first category has a major role in model calibration due to the dominant percentage of streams. The resulting decay rate for the first-order stream class is “0.38” day^{-1} ; for the second order stream class it is “0.12” day^{-1} . Furthermore, total loads estimated in the present study are compared with published values [Dunn, 1996; Rebich et al., 2011]. Dunn [1996] computed trends in nutrient inflows to the Gulf of Mexico for the period of 1972-1993 reported an estimated mean annual TN yield of $343 \text{ kg}/\text{km}^2/\text{yr}$ for the San Antonio and Guadalupe basins.

A steady-state statistical model *SPAtially Referenced Regression On Watershed attributes* or SPARROW improves on previous regression models by using a mechanistic model structure to relate in-stream nitrogen transport, including a first-order decay rate with spatial data on nitrogen sources (for example atmospheric deposition and soil permeability) and stream properties (streamflow and water time travel) [Smith et al., 1997; Alexander et al., 2007]. Rebich et al., [2011] developed the SPARROW model to estimate total nutrient loads for the Lower Mississippi Texas-Gulf (LMTG) region. They estimated a total nitrogen load of $311 \text{ kg}/\text{km}^2/\text{yr}$ for the study region. In our study the

annual average load is 301 kg km²/yr, which is in general agreement with published values.

Table 10: Model calibration

HydroID of station	Monitored load (kg yr ⁻¹)	Modeled load (kg yr ⁻¹)	Difference (%)
69688	562021	612210	8.93
70210	238720	255143	6.88
74406	2346934	2460291	4.83
71431	420561	448275	6.59
75453	1081124	1121774	3.76

3.5.3. Comparison of TN load San Antonio and Guadalupe basins

In this study, total nitrogen loads (including TX-ANB and point-source contributions) delivered to the coastal zone in 2008 and 2009 by the San Antonio and Guadalupe basins are compared (Figure 15).

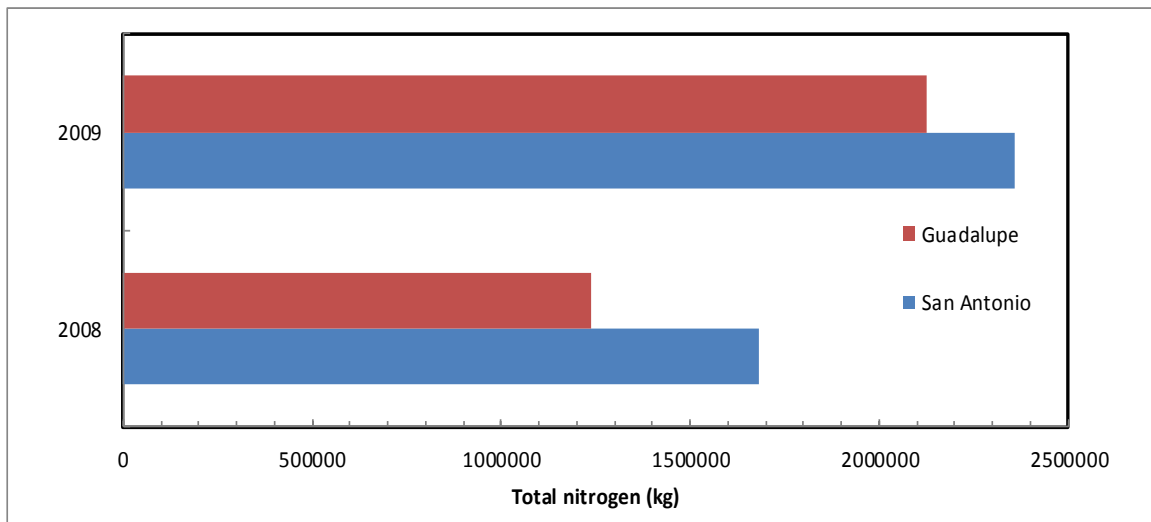


Figure 15: Total delivered nitrogen load for the San Antonio and Guadalupe basins in 2008 and 2009

The model output indicates that the San Antonio basin delivers 26% more TN load than the Guadalupe basin in a drier year (2008) and 10% more than the Guadalupe basin's contribution in a wetter year (2009). The difference can be explained by the fact that the San Antonio basin is affected more substantially by urban activities including point sources, in addition to the agricultural activities. Hence, the San Antonio basin has more impact on the downstream environment. The San Antonio basin delivered 57% of the TN load delivered by the San Antonio and Guadalupe basins in 2008. In 2009, the San Antonio basin contributed 52%. Moreover, because of higher flow, both basins delivered more nitrogen load to the Gulf of Mexico in a wetter year (2009) than in a drier year (2008). Figure 16 shows the spatial variation of in-stream TN load in 2008 and 2009. This figure provides TN load for all river reaches, which may be helpful for watershed-scale management. This figure also demonstrates the increase in delivered TN from 2008 to 2009.

3.5.4. Effect of reservoirs on the nitrogen removal

We also investigate the effect of reservoirs on total nitrogen loads in the San Antonio and Guadalupe basins. The TN removal in reservoirs is calculated based on the equations (14) and (15). Three reservoirs in the San Antonio basin were considered: Medina Lake, Calaveras Lake, and Braunig Lake. Medina Lake, located 64 km northwest of San Antonio, was constructed on the Medina River in 1913 for water supply purposes; it has a surface area of 22 km² and, reduces TN load by 77%. Calaveras Lake, located 32 km southeast of San Antonio, and Braunig Lake, 27 km south of San Antonio, removes almost 100% of the TN load entering these lakes. Four reservoirs in the Guadalupe basin, Canyon Lake, Coletto Creek Reservoir, Lake Dunlap, and Lake Gonzales, have less impact on TN load than do the San Antonio basin reservoirs. Lake Dunlap and Lake

Gonzales, the smallest (about 1.1 km²), reduce the TN load by four percent. Canyon Lake, located 25.7 km northwest of New Braunfels on the Guadalupe River, is the largest reservoir in this study (about 33 km²); it reduces TN load by 74%. The Coletto Creek Reservoir, located 24 km southwest of Victoria on Coletto Creek, removes 52% of the TN load.

3.5.5. Effect of stream order on nitrogen delivery

Alexander et al. [2000] showed that water travel time in stream channels of different sizes influences the nitrogen loss rate. Their results also show that the nitrogen loss rate in streams decreases rapidly with increasing channel size [*Alexander et al.*, 2000]. We calculate the percentage of delivered TN for all Strahler stream orders; that is, delivered percentage of TN input to a river reach. Figure 17 shows the average percentage of TN delivery, the contribution of each stream order to total river reaches, and mean river channel width versus stream order for both San Antonio and Guadalupe basins. The results show that smaller streams make larger contributions to the total number of streams (Figure 17a) and the increased TN delivery is consistent with increasing channel width, as shown in Figure 17b. The TN delivery increases with increasing stream order (Figure 17c). In other words, Figure 17c shows that reach-specific N removal (removed proportion of N input to a reach) is greater in first-order reaches (25-30%) than in higher order reaches (2-5% in 5th and 6th orders in the Strahler order classification). Estimation of higher TN delivery in larger streams (as compared to smaller streams) in the San Antonio and Guadalupe basins is consistent with results reported in the literature [*Seitzinger et al.*, 2002; *Alexander et al.*, 2007, 2009].

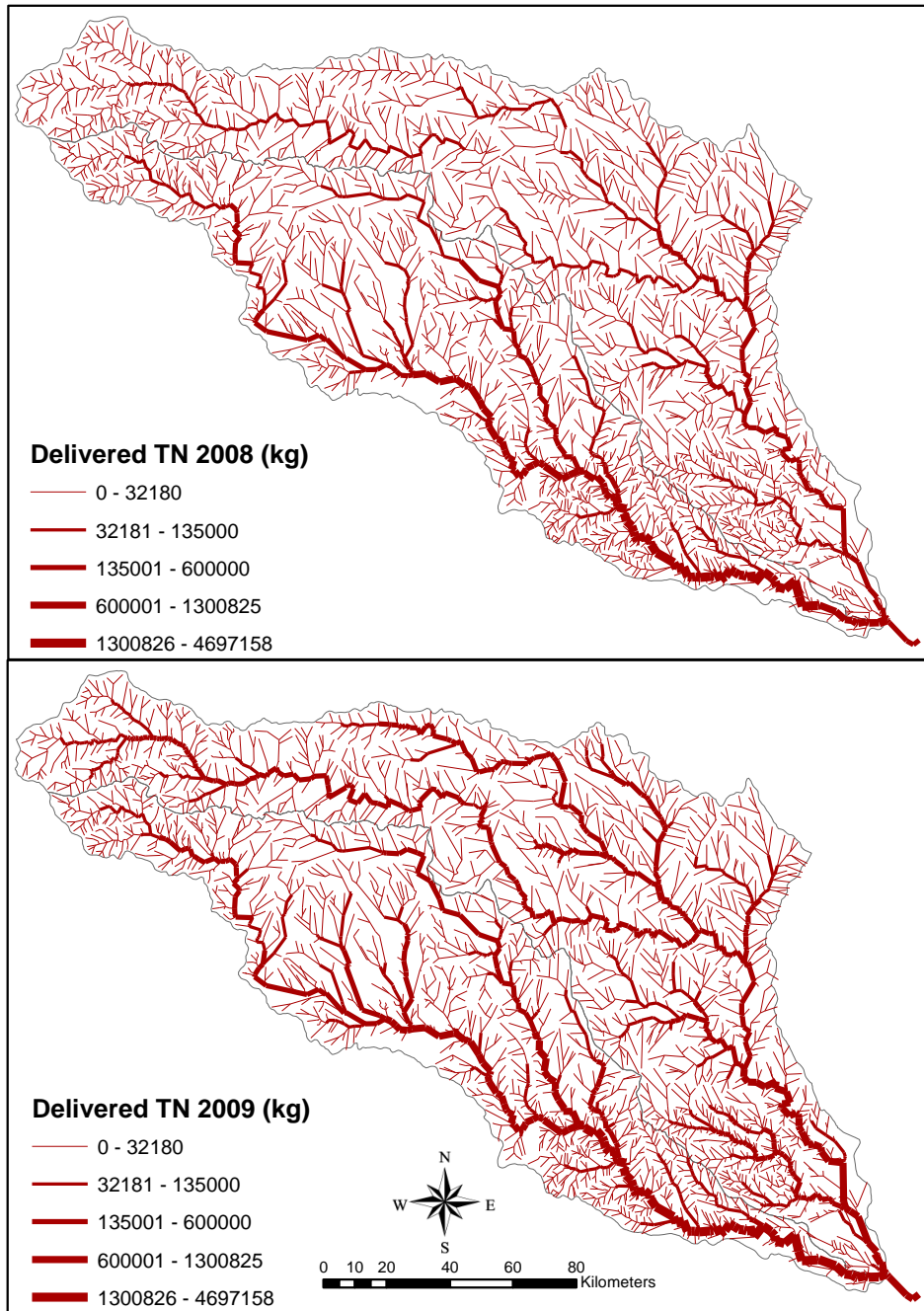


Figure 16: Spatial variation of in-stream TN for the San Antonio and Guadalupe basins in 2008-2009

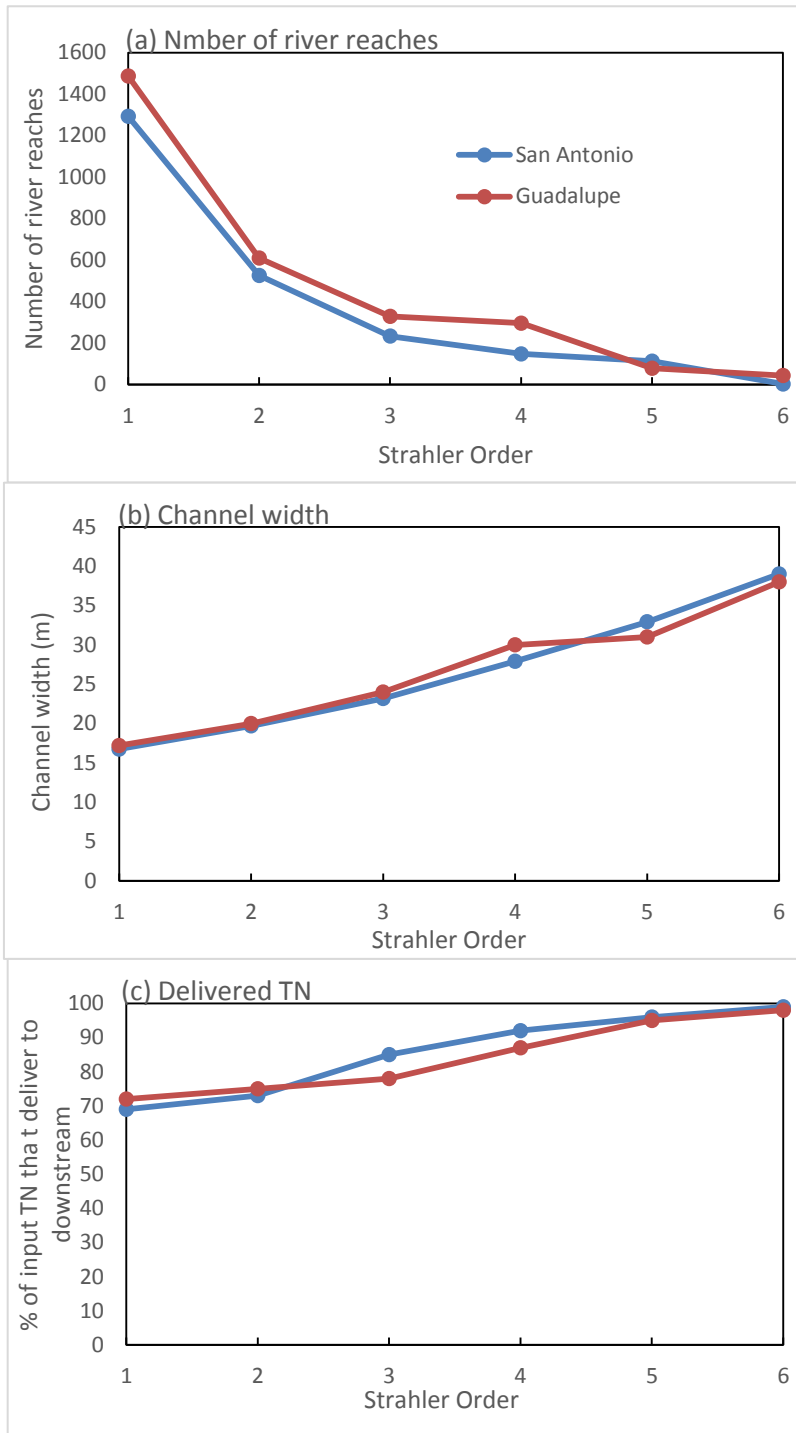


Figure 17: Relation between Strahler order and (a) number of river reaches; (b) river reach channel width; (c) % of delivered TN; (b)

3.5.6. ESRI Map service for water quality modeling

An ESRI map service consolidates model output data, river network and reservoir maps, and model input data into a single view for reporting. The ESRI map service in this project is used to support the GIS modeling framework for the San Antonio and Guadalupe River basins. The following steps explain how this process is conducted. A map document is opened in the ArcMap application. From the File menu in ArcMap, “Share As Service” is selected, then “Publish a service” is selected from the “Share As Service” menu. In the “Publish a Service” dialog box, “Connect to ArcGIS Server” is chosen to create a new connection to the server. The “Add ArcGIS server” is opened as the next window and “Publish GIS Services” is selected. On the next page, the URL of the ArcGIS Server site is required. In this project, the URL of “ArcGIS Server in the Center for Research in Water Resources in the University of Texas at Austin” was used.

Once the map service is published on the ArcGIS Server, it can be used by the web application. “ArcGIS.com map viewer” was applied to view and share the map service over the web. To have access to the server a new ESRI Global Account must be created. To do this, open a new web browser and navigate to the “ArcGIS Server Services Directory”. From the services list, the previously published map service is selected. The web map can then be saved and shared with the public. A web service of this project can be accessed at: (<http://bit.ly/1uplqrd>). The web service provides the following information for the San Antonio and Guadalupe basins: TX-ANB dataset at the catchments scales, simulated flow by the RAPID model for all NHDPlus river reaches, TN point sources, and computed in-stream TN represented by SchemaLink. Information for the simulated time period (2008-09) can be viewed by clicking on each layer. The ESRI map service in this project helps users to assess how much TN is present in every

catchment, how it changed from 2008 to 2009, how flow regimes changed during the study years, and how in-stream TN varied based on in-catchment TN and river flow in the vector based NHDPlus dataset.

3.6. CONCLUSIONS

A GIS modeling framework was developed to link a GIS-based river routing model (RAPID) and a GIS-based total nitrogen dataset (TX-ANB). The modeling framework was designed to estimate TN load in the San Antonio and Guadalupe basins in 2008 and 2009. Model results compare well with previous studies; the findings show that in both years the San Antonio basin contributed more delivered TN than the Guadalupe basin. Both basins delivered more TN to the Gulf of Mexico in 2009 than in 2008.

The effect of reservoirs on TN load was considered. In the San Antonio basin, Calaveras Lake and Braunig Lake remove almost 100% of the TN load that enters these lakes; Medina Lake reduces the TN load by 77%. In contrast, reservoirs in the Guadalupe basin have less impact on TN load; of the Guadalupe reservoirs, Canyon Lake provides the maximum reduction (74%). We also showed that TN removal in the San Antonio and Guadalupe basins is inversely related to stream orders: the higher the order the less TN removal.

On a regional scale, the goal of this project was to represent the application of a GIS modeling framework to the San Antonio and Guadalupe basins to provide a web-based support system that, can be used to evaluate in-catchment TN, simulated river flow, and in-stream TN.

Chapter 4: Modeling interactions of agriculture and groundwater nitrate contaminant: The STICS–Eau-Dyssée coupled models applied over the Seine basin

4.1. ABSTRACT

Nutrient enrichment is one of the major environmental phenomena in the French coastal zone. The Seine river system represents a sophisticated hydrosystem with several aquifers. This basin is well-known as one of the most productive agricultural areas in France, and includes the megalopolis of Paris. The present study demonstrates the coupling of the agronomic STICS model, and Eau-Dyssée, a distributed hydrologic modeling system. The Eau-Dyssée modeling system is modified in this paper to simulate nitrate transport from the surface to rivers and aquifers, and also from aquifers to the surface and rivers. The STICS- Eau-Dyssée combined system is implemented for entire seine basin to compute daily nitrate flux in the Seine grid river network. The simulation covers 39 years (1971-2010). The simulated results were similar to the measured data on an annual basis, without any calibration performed. Model results showed that simulated nitrate highly depends on the inflow produced by surface and subsurface waters.

4.2. INTRODUCTION

Two anthropogenic activities are food and energy productions which have greatly increased nitrogen creation by over a factor of ten compared to the late-19th century [Galloway *et al.*, 2004]. Human activity such as the cultivation of N-fixing crops, burning of fossil fuel, discharging of industrial and domestic effluents, and extensive usage of fertilizer have recently accelerated the increase of nitrogen load to watersheds. These activities lead to increase riverine flux to coastal areas and lakes [Zhang *et al.*, 2010]. Nitrogen load in the form of nitrate is a key industrial fertilizer which is the major cause of eutrophication in the water system [Vitousek *et al.*, 1997]. Increasing nitrate concentration in surface water and groundwater is a major concern in watersheds with extensive agricultural activities [Foster, 2000]. The European Union and World Health Organization defined the threshold of $50 \text{ mg-NO}_3\text{L}^{-1}$ for the potable water in Europe. Furthermore, the Water Framework Directive set a goal that all groundwater bodies must reach this limitation by 2015 [Billen *et al.*, 2007b; Even *et al.*, 2007; Flipo *et al.*, 2007a; Rivett *et al.*, 2008; Thieu *et al.*, 2009; Philippe *et al.*, 2011].

Nutrient enrichment and eutrophication are major environmental phenomena in the French coastal zone. The Seine river system represents a very important case study of river nutrient chemistry in a regional ecosystem. The Seine basin includes the megalopolis of Paris with about ten million inhabitants. Its watershed has an intensive agriculture and three major aquifer systems [Billen *et al.*, 2007a].

Deterministic or physically distributed models are strong tools to compute flow and transport in a river system with aquifers and to understand the behavior of complex large river networks such as the Seine river system [Flipo *et al.*, 2007a]. Within the framework of the PIREN-Seine research program, several numerical models have been

developed to tackle different objectives of the Seine Hydrosystem. Each model is adapted for a specific temporal and spatial scale and has a capability to share data with other models in the concept of a modeling tool-box. Examples of models developed and implemented for the Seine Hydrosystem are the Riverstrahler model [Billen *et al.*, 1994, 2001; Garnier, 1995; Garnier *et al.*, 2005; Thieu *et al.*, 2009] and the CAWAQS model (CAttachment Water Quality Simulator) [Flipo *et al.*, 2007a, 2007b].

The Eau-Dyssée modeling platform is developed by MINES ParisTech and is based on the existing models and databases of the PIREN-Seine program. The hydrogeological section of the model uses the same principles as the MODCOU model [Ledoux *et al.*, 1989; Habets *et al.*, 2008]. The Eau-Dyssée platform has three features which are: the composition of different modules represents a process and each module can communicate externally with other modules; an improved river-aquifer interaction with a water level fluctuation module called QtoZ [Saleh *et al.*, 2011]; and finally the incorporation of a river network model called RAPID in the river network component [David *et al.*, 2011a]. RAPID directly computes water flow and volume for each cell of the river network and allows flexibility in the number and location of river gages. This version of Eau-Dyssée did not model concentration of contaminants such as nitrates.

The aim of the present study is to simulate nitrate flux leaching into the river network for the entire Seine basin, including several layers of aquifers. The infiltrating nitrate flux leaving the root zones is obtained from the agronomic model STICS (Simulateur mulTidisciplinaire pour les Cultures Standard) [Brisson *et al.*, 1998, 2002]. For this purpose, the Eau-Dyssée platform is modified to transport nitrate from land to rivers and to aquifer waters due to the importance of long residence time in aquifer systems. This model also simulates nitrate traveling from aquifers to surface and rivers.

4.3. METHODOLOGY

4.3.1. Study area

The Seine, the second largest river in France after the Rhine, is 776 km long. It rises at an altitude of 446 m at Source-Seine in the department of Côte-d'Or in Burgundy and discharges into the English Channel near the city of Le Havre [Massei *et al.*, 2010]. The Seine river network contains more than 5,000 river reaches and 25,000 km of tributaries (Table 11). Table 11 also shows that the stream gradients are not extremely varied; although altitude ranges from 0 to 856m above sea level, 90% of the basin is below 300m. The Seine basin covers an area of 80,506 km² in northern France (Figure 18). The hydrological regime of the Seine River network is considered to be pluvial oceanic, with varied seasonal flows (high flows in winter and low flows in summer) that reflect rainfall distribution throughout the year. The Seine river network also has several aquifers that play an important role in sustaining base flow [Ducharne, 2008].

Table 11: Mean morphological characteristics of the Seine drainage network

Order	Number	Length (km)	Mean		Catchment area (km ²)
			Width (m)	Slope (m/m)	
1	2887	11714	2.5	0.0151	42234
2	1440	5591	5.9	0.0062	17190
3	848	3801	12.6	0.0034	12493
4	354	1932	23.0	0.0022	6982
5	186	1102	50.2	0.0014	4497
6	115	601	88.9	0.0013	2416
7	36	383	200.9	0.0011	2113
Seine	5866	25124	10.5	0.0097	87926

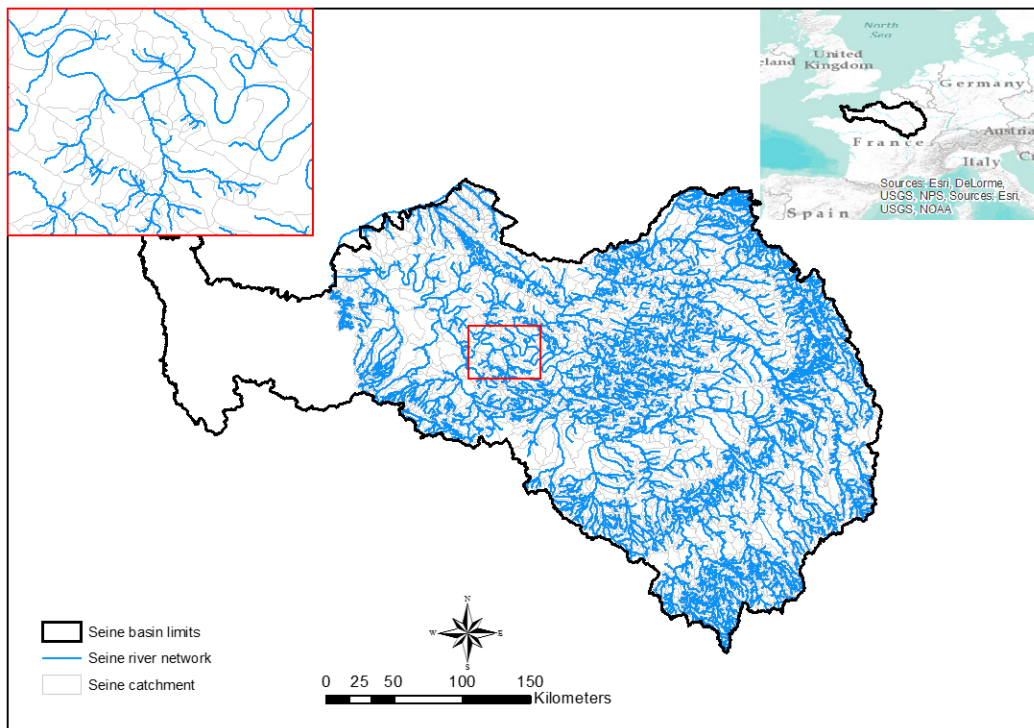


Figure 18: The Seine River Watershed

The Seine basin is predominantly covered by an intensive agricultural industry. The Seine basin is also home to approximately 30% of the total population of France (20 million inhabitants), including 10 million people in Paris. Besides agricultural activity, pollution sources including industrial sites and population density lead to water quality degradation and nutrient enrichment of the Seine River. This nutrient enrichment causes eutrophication problems in most of the main branches of the river system. Nutrient enrichment of the water is largely a function of nitrate that comes from fertilizer runoff and industrial pollutants. The excess nitrate cause eutrophication and adverse environmental effects such as hypoxia (oxygen depletion), harmful algal blooms (growth of phytoplankton in a water body), and reduced fish and shellfish production.

4.3.2. Models description

This research is based on the coupling of an agronomic model, STICS, and a regional hydrological model, Eau-Dyssée (Figure 19). The study period covers 39 years (1971-2010) and the models are coupled over the entire Seine basin.

4.3.2.1. The STICS crop model

STICS is a daily time step crop model that simulates (a) crop yields in terms of quantity and quality and (b) the environment in terms of drainage and nitrate leaching. STICS has the following features: adaptability to various crops using the same set of equations and specific parameters, ability to simulate various climate and soil conditions, possibility of adding new modules, and external communication with other models and developers. The input variables are related to climate, soil, and the crop system [Chnebel *et al.*, 2004].

The STICS model accounts for crop growth and development and nitrogen and water balances. The upper boundary of the system corresponds to climatic variables including radiation, solar radiation, daily minimum and maximum temperature, precipitation, and reference evapotranspiration. The lower boundary corresponds to the soil/sub-soil interface. The STICS model characterizes the overhead biomass (carbon and nitrogen), leaf area index (LAI), biomass of harvested organs, and root length profile of crops. Therefore, vegetative plant parts (leaves, branches, or tillers) are considered as a whole [Brisson *et al.*, 2003]. Soil is described as horizontal consequent layers; every layer is characterized by its water, mineral and organic nitrogen content. Roots, which are defined by their density distribution in the soil, represent the interaction between soil and crop.

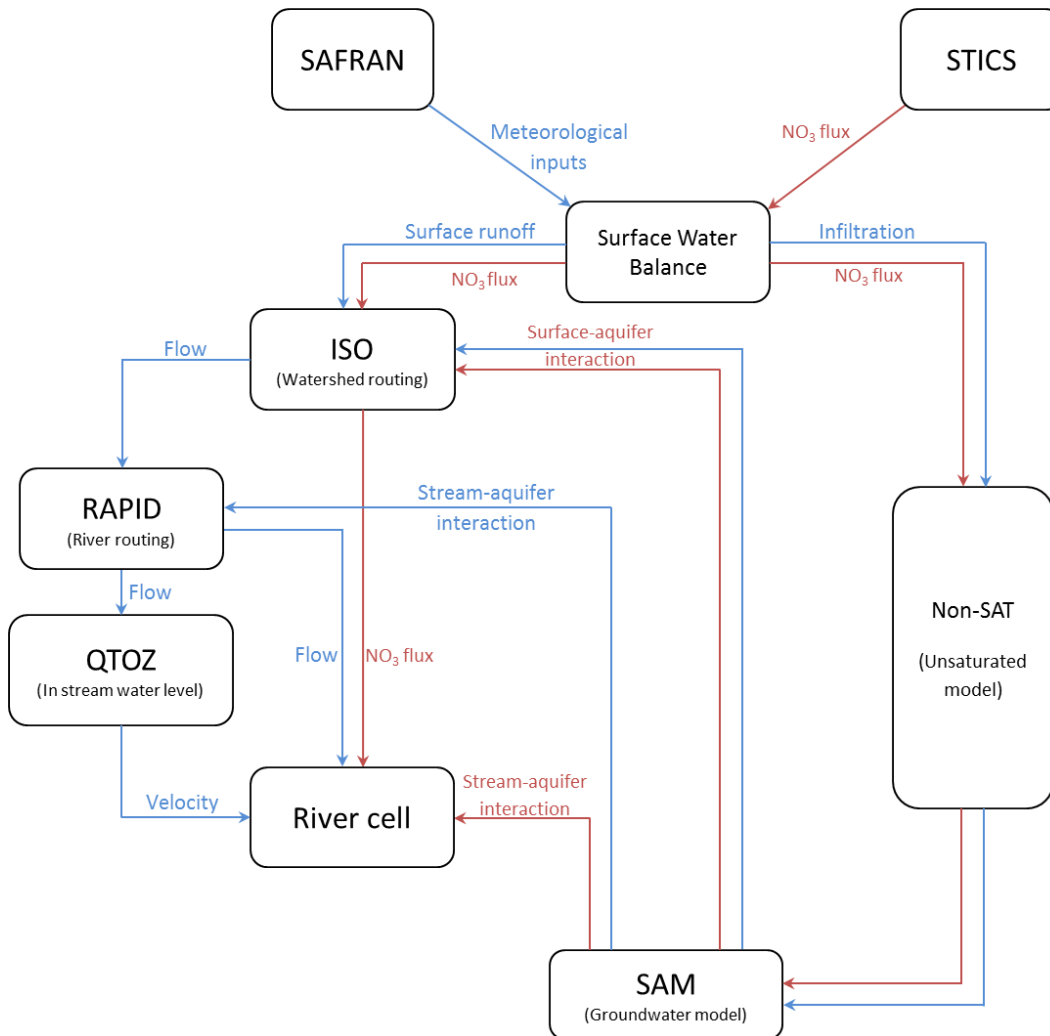


Figure 19: Modeling framework

The core of the STICS model includes four primary sets of modules. The first set of modules includes phenology, shoot growth, and yield formation. This set considers the ecophysiology of aerial plant parts. The second set contains four modules (root growth, water balance, nitrogen balance, and soil transfer) that simulate the interaction between underground plant parts and soil functions. The third module is crop management, which accounts for water transfer through the canopy, the status of water and heat balances in

the soil-crop system, and fertilizers. The fourth module is the microclimate, which calculates temperature and air humidity through the canopy.

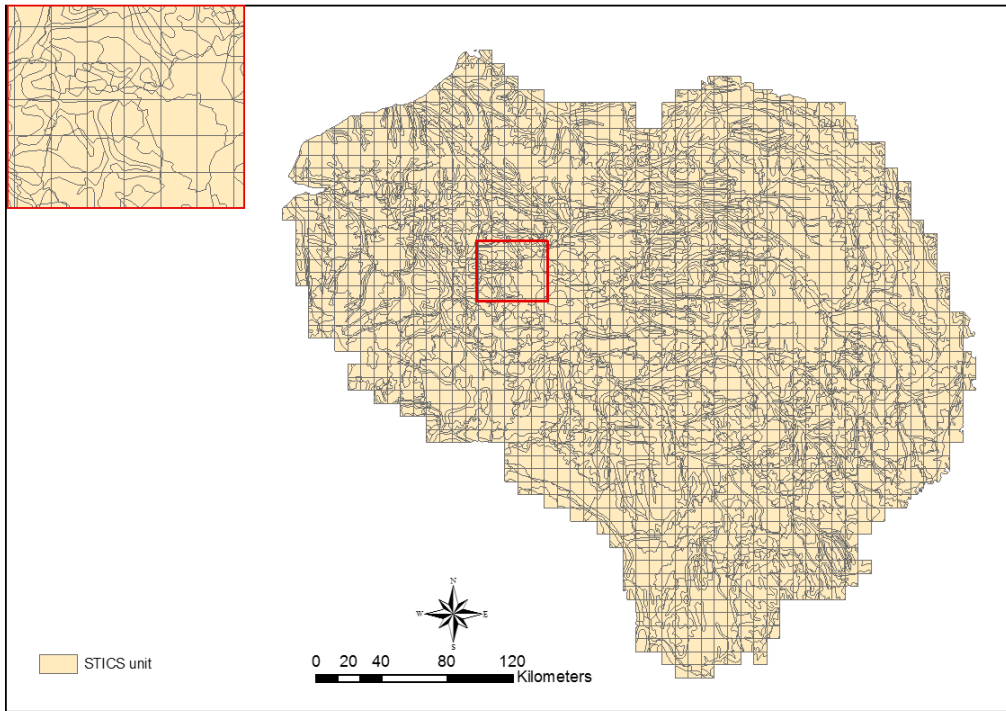


Figure 20: General Simulation Units in the STICS model

Three principal databases are used in the STICS model to characterize the pedology, meteorology, and agriculture of the Seine basin. The three databases are: soil, agriculture, and meteorological; each has a different spatial resolution [Ledoux *et al.*, 2007]. The intersection of these databases generates spatial units called General Simulation Units (GSU) that share the same spatial, pedological, agricultural, and meteorological characteristics. STICS, which is run on each GSU, is made up of 9596 squares with an average unit area of 12 km² over the Seine basin (Figure 20).

4.3.2.2. The Eau-Dyssée platform

The Eau-Dyssée platform is a hydrometeorological and biogeochemical model that couples existing specialized modules to simulate water resources (quantity and quality) at regional scales. The Eau-Dyssée model consists of modules that compute elements of the water cycle (Figure 19): surface and unsaturated components, the groundwater dynamic or saturated component, and water routing in the river network component [Saleh *et al.*, 2011].

The surface component uses a seven-parameter conceptual parameter model to compute water mass balance at a daily time step for each cell of the surface mesh [Deschesnes *et al.*, 1985]. In this module, the domain is divided into units called production functions that are extracted from the interaction of land-use and geological units [Flipo *et al.*, 2012]. Inputs to the surface component include precipitation and potential evapotranspiration with a daily time step and a regular 8-km grid produced by a mesoscale atmospheric analysis system, SAFRAN, [Durand *et al.*, 1999; Quintana-Seguí *et al.*, 2008]. The outputs are actual evapotranspiration (AET), surface runoff, infiltration, and soil storage. The total number of surface cells covering the Seine basin is 35,698; the average resolution is 1x1 km² (Figure 21).

The ISO module is used to route surface runoff to the river network. A number of isochronal zones are defined in the ISO module to determine the delay between runoff generation and the time that runoff reaches the nearest river cell.

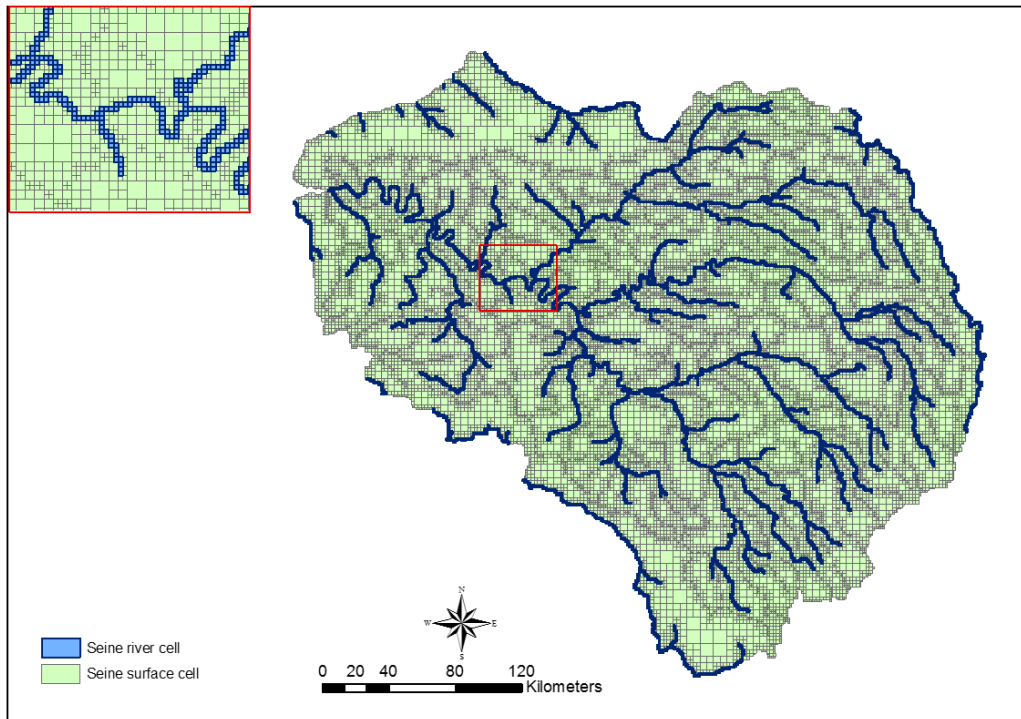


Figure 21: Surface and river grids for the Seine basin

In-stream-water routing is handled by the RAPID model [David, 2009]. RAPID solves the matrix-based Muskingum equation and has the ability to run in a parallel computing environment with actual speedup. Inputs to RAPID are surface and subsurface runoff; outputs are water flow and volume in each cell of the grid river network (Figure 21). The Seine River network is composed of 6,481 river grid-cells of which 3,519 interact with aquifers. RAPID was successfully implemented in basins of varying scales: the San Antonio and Guadalupe basins in Texas [David *et al.*, 2011b]; France [David *et al.*, 2011a]; the Texas Gulf coast region [David *et al.*, 2013]; and the Mississippi basin. The QtoZ module calculates water levels at a given river cell in the Eau-Dyssée platform [Saleh *et al.*, 2011]. At each time step QtoZ receives a discharge computed by RAPID; it computes a water level that is sent to the groundwater model (SAM) to simulate stream-

aquifer interactions. QtoZ has three options for calculating water levels: (a) fixed water level, (b) rating curve, and (c) the Manning equation. The Manning equation option is modified in this project to estimate flow velocity and cross-sectional area for each river cell at every time step.

Infiltration is vertically portioned by the production function transferred to groundwater within the unsaturated zone. The unsaturated zone component NONSAT [Ledoux *et al.*, 1989; Gomez *et al.*, 2003; Flipo *et al.*, 2007c] consists of a sequence of reservoirs; the number of reservoirs depends on the distance between soil horizons and the saturated zone, which is initially calculated on the basis of hydraulic head distribution.

The saturated zone component SAM (Simulation des Aquifères Multicouches) is a regional spatially-distributed model that applies the diffusivity equation to compute both the temporal distribution of hydraulic heads and the flow in multilayer aquifer units [Ledoux *et al.*, 1989]. Flow is 2-D in each aquifer and vertically 1-D in the aquitard between two horizontal layers [Flipo *et al.*, 2012]. An aquitard is a low-permeability bed parallel to an aquifer. SAM also calculates the water flux exchanged between aquifer grid-cells and stream grid-cells using water levels calculated by the QtoZ module. SAM also uses water levels calculated by the QtoZ module to calculate the water flux exchanged between aquifer grid-cells and stream grid-cells.

4.3.2.3. Modification to the Eau-Dyssée platform to transfer nitrate

In this research, solute transport is added to watershed routing within the surface component of the Eau-Dyssée platform. The ISO module (Figure 19) routes runoff to the river network. The watershed is divided into a number of isochronal zones representing the number of travel time steps to flow into the nearest river cell. The same methodology

is applied to route solute transport through the watershed. With the new modification, the ISO module is able to accumulate the mass of solute constituents and define the total mass transferred to river cells. The accumulated solute transport is the sum of solute from the water balance, surface runoff, unsaturated zone, and aquifer overflow. The solute can be in flux or in concentration units. A new module called ISO solute is added to the Eau-Dyssée platform to model watershed solute transport and to provide riverine nitrogen inputs. This module has the advantage of including the solute flux received from aquifer overflow and adding it to solute transport in rivers. Nitrate flows to rivers and to the surface through aquifer overflow. Nitrate also flows to rivers through stream-aquifer interactions. The Eau-Dyssée platform is also modified to consider the nitrate flux from aquifer overflows to river reaches. As a result, the Eau-Dyssée platform can be applied to regional watershed flow and solute modeling with stream and aquifer interactions.

4.3.3. Model Applied to Seine River

The Eau-Dyssée and STICS models are coupled for nitrogen transport simulation in Seine basin. The simulation period is from August 1, 1971, to July 31, 2010. STICS provides nitrogen flux in General Simulation Units (GSU). The schematic framework of spatial coupling of Eau-Dyssée and STICS is presented in Figure 22. The first step in superimposing Eau-Dyssée and nitrate concentrations is to calculate nitrate flux for each surface cell. GIS is utilized to determine the spatial contribution of nitrate flux to surface cells. The intersect tool in ArcGIS is applied to correlate corresponding GSUs to the surface cells and to compute the areal proportion of the GSU that overlaps with the surface cells.

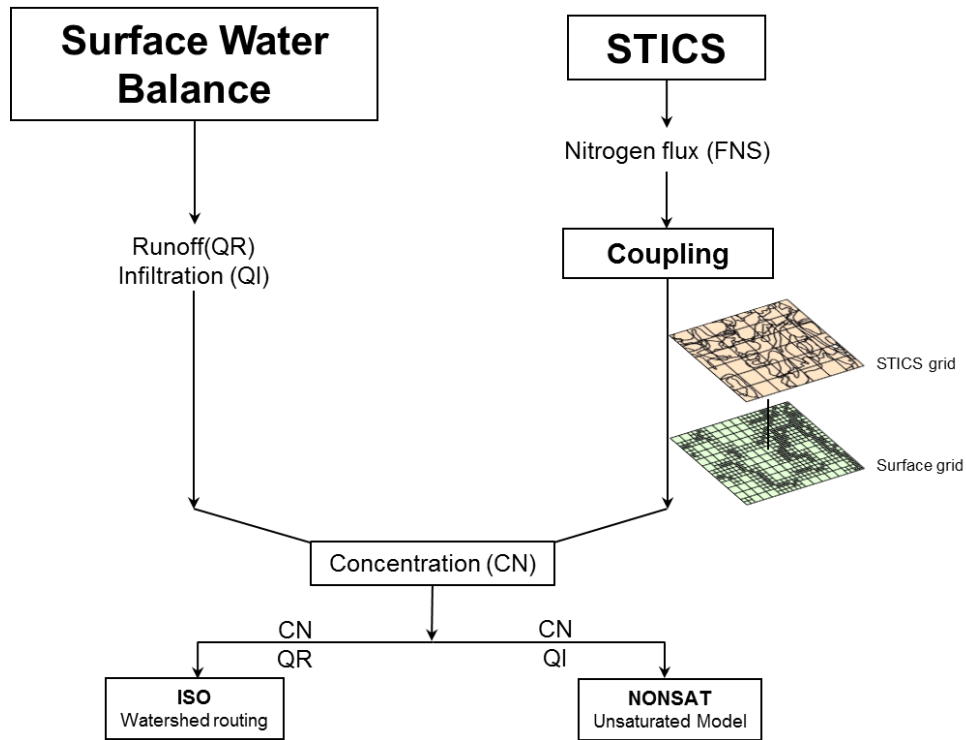


Figure 22: Schematic framework of spatial coupling of Eau-Dyssee and STICS

The second step is the dilution of the nitrate flux (STICS output) for each surface cell by runoff and infiltration. Runoff and infiltration are calculated by the surface water balance component of Eau-Dyssée:

$$\bar{C} = \frac{\varphi}{Q_R + Q_I} \quad (20)$$

where φ is nitrate flux (kg); Q_R and Q_I are daily runoff and infiltration (m^3).

The Eau-Dyssée model is run for 39 years (1971-2010) and computes a daily nitrate flux entering the river cells.

4.4. RESULTS AND DISCUSSION

The Eau-Dyssée platform is run using a 30-minute time step to simulate daily nitrate flux leaching to the river cells for 39 years (1971-2010). Three stations are selected to evaluate the model outputs (Figure 23). Annual results are compared at (1) the Poses station, (2) the L'oise a Jaux station, and (3) the Montereau-Fault-Yonne station. The Poses station is selected as a station downstream of Paris and also downstream of the main streams of the Seine river network. The L'oise a Jaux station is located in the north part of Seine basin with high agricultural activity. The L'oise a Jaux station is downstream of the Compiègne city on the Oise River. The Montereau-Fault-Yonne station is on the Seine River and is located downstream of Troyes, Nogent-sur-Seine, and Sens cities.

4.4.1. Evaluation of simulate nitrate flux

One should note that, the focus of this research is simulation of the leaching of nitrate to the river cells. The calibration of in-stream nitrate transport is not a topic of this paper. The annual leaching nitrate flux to river cells are compared with observations at the selected stations. Observations are available from 1985 to 2010. For most of stations observations are measured once a month during a year. Daily measured data is available for some of stations.

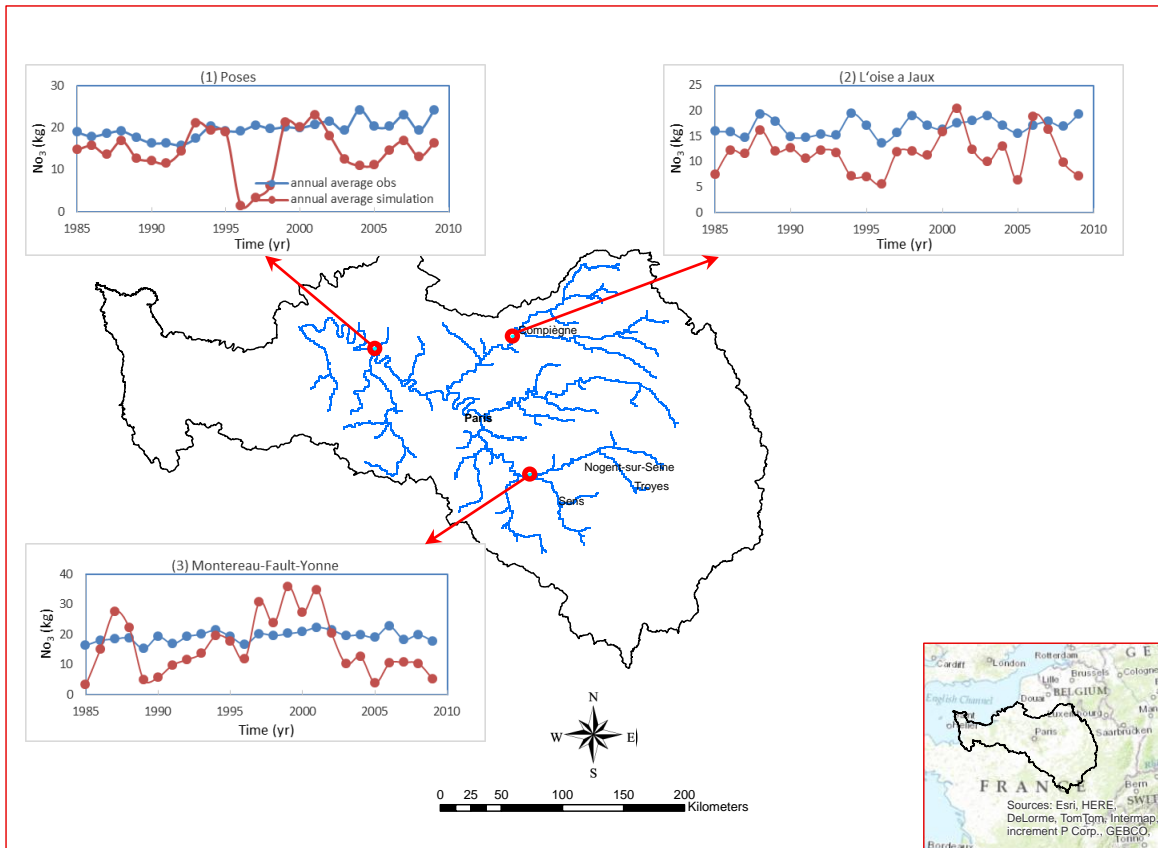


Figure 23: Comparison of annual simulated nitrate flux with observation

The Results of simulation of the annual variations generally agree with the measured observations averaged annually. At the Pose and L'oise a Jaux stations, the simulated nitrate flux between 1996 and 1999 does not trend with the observations. These years were dry years and nitrate flux leaching to rivers was transferred by surface and subsurface inflow produced by the Eau-Dyssée. The model result is also slightly underestimated at the Pose station, such an underestimation is likely due to point source loading to the Seine River in the Paris area. Comparison of between leaching nitrate flux and inflow is discussed in the next section.

4.4.2. Nitrate delivery through surface and groundwater flow

Nitrate flux leaching to rivers is carried by surface runoff and groundwater. The coupling of the Eau-Dyssée and STICS models in this study enables the study of nitrate loads in the Seine basin over the long term with the different temporal resolutions. Figure 24 shows the daily nitrate load and inflow simulated by the Eau-Dyssée platform for the selected stations. The high variation of nitrate load can be seen from this figure.

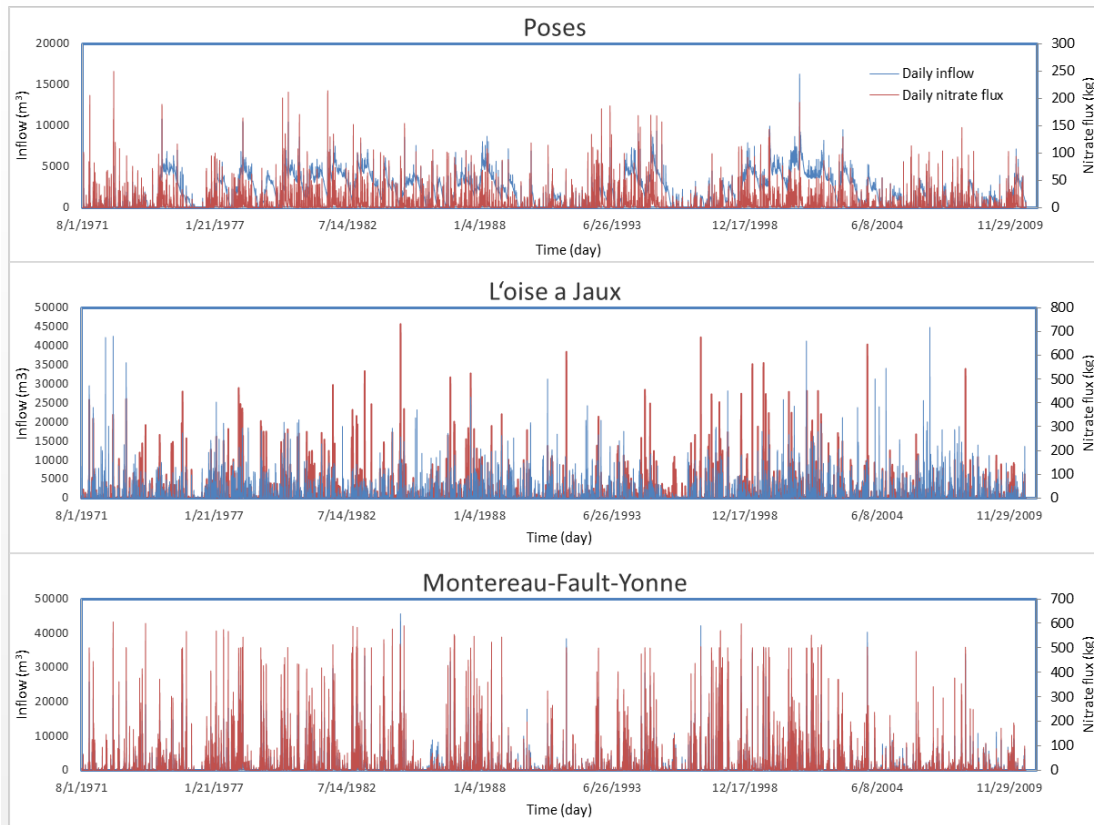


Figure 24: Simulated daily nitrate flux and inflow for the selected stations

To better understand nitrate delivery, the annual average nitrate flux is compared with annual average inflow in Figure 25. Results show that nitrate delivered to river cells highly depend on inflow. The amount of leaching nitrate is low during the dry years and it increases during the wet years. For example, the lowest amount of nitrate was delivered in 1996 and 1997 at the poses station, the annual average of inflow at these years were 173 and 382 cubic meters. In contrast, in wet years the annual average is significantly increased.

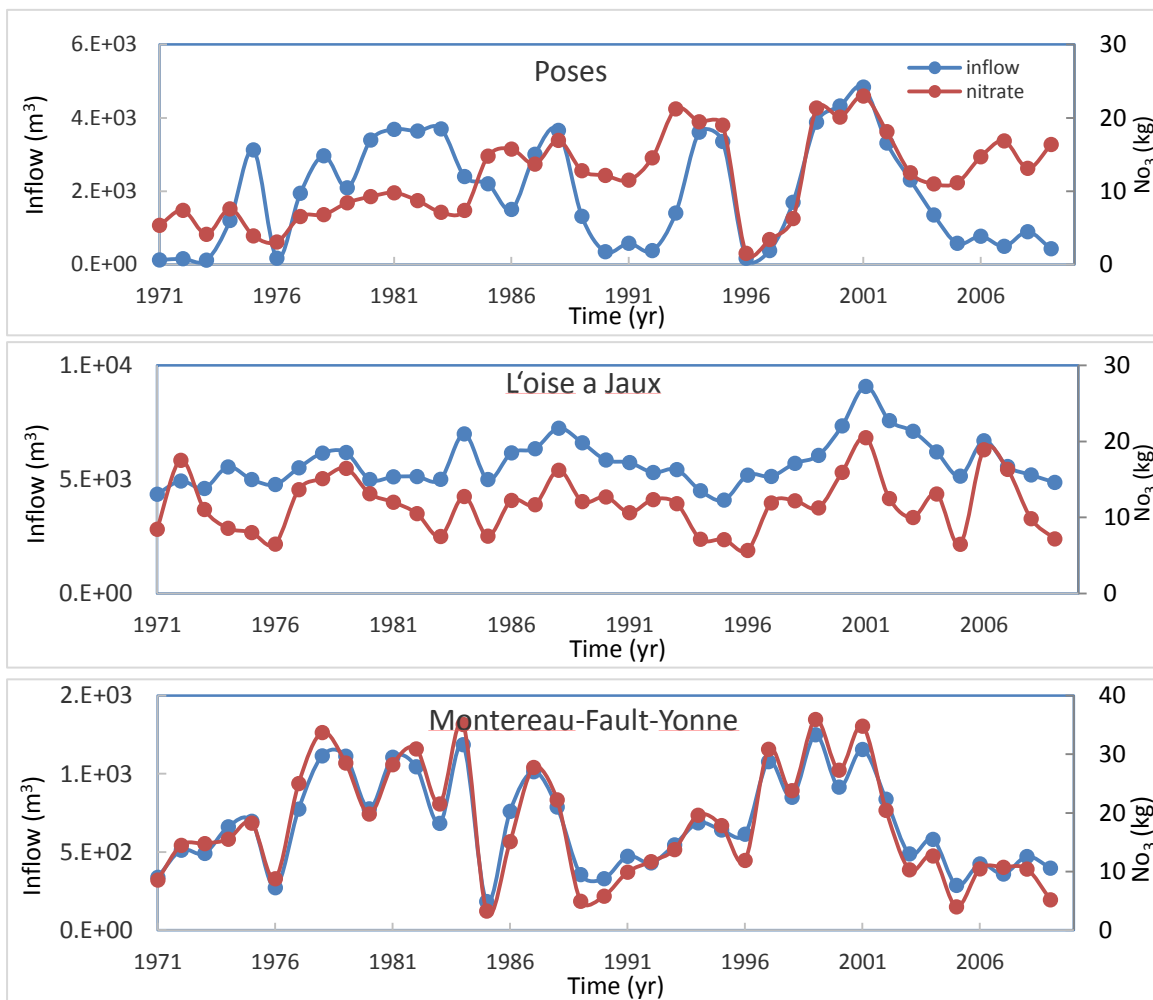


Figure 25: Simulated annual average nitrate flux and inflow for the selected stations

The model outputs show that 2000 and 2001 are characterized by very high discharge. The delivered nitrate is also increased for these years. In general, low nitrate load is delivered to rivers in 1976, 1985, 1990-91, 1996, and 2005-06, which are characterized as dry years. On the other hand, high amount of nitrate delivered to rivers in 1988, 1993, 2000, and 2001. These years have a high discharge and categorized as wet years.

4.4.3. Annual nitrate mass balance at the basin scale

The Eau-Dyssée model produces mass balance of flow that enters the river network for the simulation time period. This model is modified to generate mass balance for nitrates as part of this research. Figure 26 depicts the annual mass for inflow (m^3) and nitrate flux (kg). The cumulative mass balance for nitrate and inflow is shown in Figure 26(a) and the daily mass balance is shown in Figure 26(b). The trend of nitrate flux is agrees with the inflow mass balance. Furthermore, based on the model result, the Seine basin delivers $7.4E8$ (kg) nitrate to the Seine river network in one year. This amount of nitrate is delivered by $2.3E10$ (m^3) of inflow including runoff and groundwater overflow, which represents the surface and stream interaction with aquifer. The daily balance shows the gradual increasing of leaching nitrate starting in spring and during fall semester and decreasing starting from the end of the fall semester (Figure 26b).

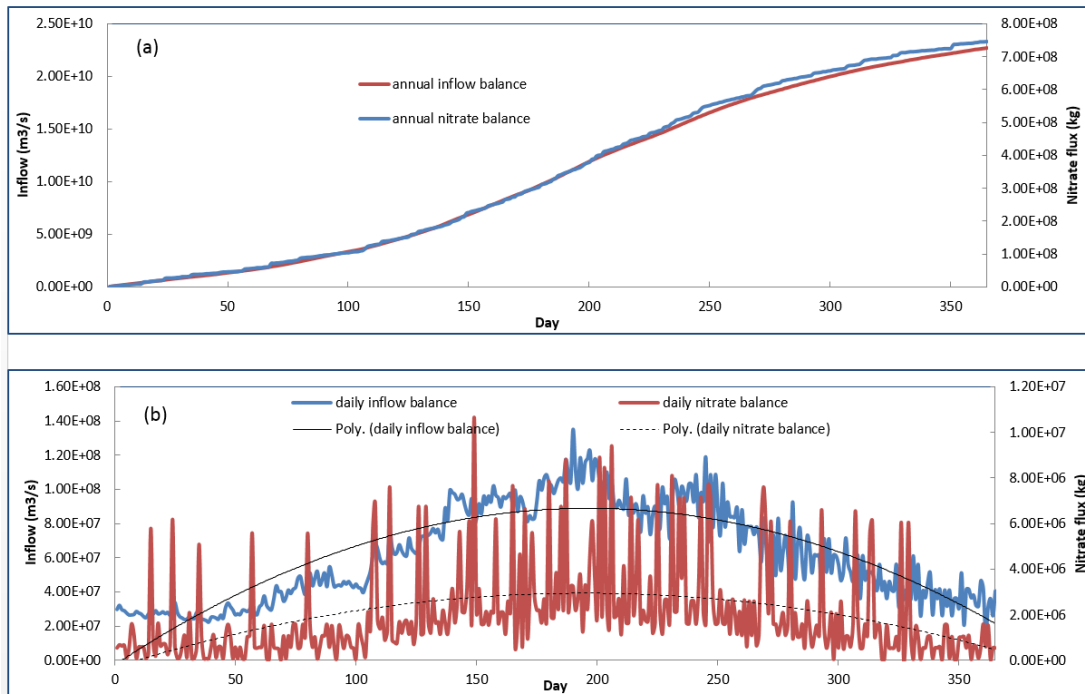


Figure 26: Mass balance of the nitrate flux and inflow for the Seine basin (a) Cumulative annual mass balance; (b) daily mass balance during a year

4.5. CONCLUSIONS

This paper described the implementation of coupled Eau-Dyssée and STICS models to the entire Seine basin. The Eau-Dyssée modeling platform was modified to also simulate nitrate transport in addition to water transport from surface to rivers and aquifers. Furthermore, the surface and river interactions with aquifers were considered for the nitrate transport. The Eau-Dyssée platform computes a daily nitrate flux leaching to rivers for 39 years (1971-2010). The leaching nitrate flux was simulated with the river grid network, which is the backbone of the river routing module of Eau-Dyssée.

The simulated results compared favorably to the measured data on an annual basis. Model results showed that simulated nitrate highly depends on the inflow produced by surface and subsurface waters. Variation of mean annual nitrate from year to year can

be explained by the hydrologic regime of that year. Larger amounts of nitrate flux transports during wet years and less transport during dry years. The daily nitrate mass balance showed the maximum amount of nitrate was delivered to rivers in the late summer and beginning of the fall semester.

The results of this research can be used to simulate in-stream nitrate transport. Point source nitrate pollution should be considered separately for the riverine transport modeling. As part of this research the hydraulic variables of river cells such river flow velocity is computed for the entire basin for the studied period (1971-2010).

Chapter 5: Conclusions

5.1. SUMMARY AND CONCLUSIONS

The objective of this study is to identify optimum modeling techniques or tools with existing datasets. This work covers two spatial scales and two temporal scales. There are three components in this research. The first component advances continental river modeling by developing a high resolution, distributed continental scale river routing model. The second component advances hydrological, steady-state models by developing a high resolution GIS-based regional model for Texas Rivers. The third component advances continental scale research by modeling regional flow and transport while considering stream-aquifer interactions at regional scale at high resolution. Beyond this study, work will need to be done to refine models to reflect river-aquifer interaction and the absorption of pollutants. Below are conclusions from each chapter followed by recommendations for future research.

5.1.1. Question one

What are the pros and cons of grid-based and vector-based river networks? What significant differences may exist between river flows computed using these two approaches? How can a dynamic regional river routing model be extended to continental scales using GIS-based datasets?

Chapter 2 presents an extension of regional river flow modeling to the continental scale by using high resolution river data from *NHDPlus* dataset. Previously, this dataset has been used for river flow simulation on a regional scale in Texas and California. In this study, an unsteady-state modeling framework was developed for the entire Mississippi river network, which is the largest drainage system in North America. An upscaling process is developed on top of the vector-based river network to decrease the

computational effort, and to reduce input file size. This study shows a direct comparison between grid-based and vector-based river networks in the continental river routing model.

The RAPID model was run with runoff obtained from Mosaic and VIC land surface models from 2000 to 2008. The grid-based river network was generated using the 1/8 degree topographical data obtained from NLDAS dataset. The vector river network was derived from the *NHDPlus* dataset.

The benefit of vector-based river network is that the locations of river gauges are more easily determined (e.g. no need for snapping), therefore one can use as many gauges as needed. Another advantage of vector river network is that, physical visualization of river segments is superior to the grid river network. Vector-based river networks are advantageous because the real length and slope of river segments can be obtained from the *NHDPlus* dataset. On the other hand, an advantage the grid-based river network provides is that, the coupling process of the LSM with a grid-based river network is straight forward, since the runoff data of LSMs are also obtained from the NLDAS dataset.

This research identifies drainage area as a key factor in the flow simulation, especially in a wetter climate. Comparison of simulated river flow showed that, the grid and vector river networks have very comparable results, when the drainage area is determined accurately. Both river networks mimic observed hydrographs in the selected gauges, and are able to capture peak flows. However, the vector-based approach allows for a more accurate representation of size and location of river basins, especially in wet climates. The application of the VIC land surface model significantly improves the RAPID performance.

5.1.2. Question two

How can a GIS framework be developed for a steady-state riverine nitrogen transport modeling along a large river network? How can the GIS environment be used to integrate a precomputed flow from RAPID and GIS-based total nitrogen (TN) dataset?

In chapter 3, an enhanced GIS-based modeling framework for steady-state modeling of total nitrogen is presented. This study explores challenges in developing a geospatial framework, such as the GIS environment, that links measurements, river networks, large scale river flow and transport modeling together. The regional GIS modeling framework developed in this study, can be characterized as a steady-state process-based approach. This modeling framework extends GIS application from small domain with hundreds of river segments to a large river network with thousands of connected river segments. The schematic network was built upon the *NHDPlus* river network for the San Antonio and Guadalupe basins. This GIS framework integrates the GIS-based river routing model (RAPID) and GIS-based total nitrogen dataset (TX-ANB) for 2008 and 2009. This framework also includes the effect of reservoirs on in-stream total nitrogen transport.

Findings show that the GIS framework can be applied to represent a spatial distribution of flow and water quality factors in a large river network with thousands of river segment. However, time features of the GIS environment limit its applicability to large scale unsteady-state modeling.

Regarding the modeling results of the San Antonio and Guadalupe basins, the San Antonio basin contributed more delivered TN in both years (2008-09). Furthermore, both basins delivered more TN to the Gulf of Mexico in 2009 than in 2008. In the San Antonio basin reservoirs have more impact on TN load than they do in the Guadalupe basin. This

research also shows that TN removal in the San Antonio and Guadalupe basins is inversely related to stream orders: the higher the order the less TN removal.

5.1.3. Question three

How do the interactions of agriculture and groundwater nitrate contaminants effect flow and transport modeling in large scale-river networks? What are essential modifications for the dynamic coupling of Eau-Dyssée and STICS models?

Chapter 4 presents, an enhanced technique for flow and transport modeling in large river networks with river-aquifer interfaces. The Seine basin was selected for this section because it is a very important hydrosystem that includes the megalopolis of Paris, a city of approximately ten million inhabitants. The Seine basin has an intensive agricultural industry and three major aquifer systems. An unsteady modeling framework is developed by coupling the agronomic model (STICS) and the hydrological platform (Eau-Dyssée). Eau-Dyssée was extensively modified to add nitrate concentration to the volumetric flow variable. This required the modification of the model's components. The STICS–Eau-Dyssée models were applied to compute leaching nitrate contaminant to rivers from 1971 to 2010. Total area of the studied region is 80,500 km²; the river network includes 6481 river cells with an average of 1 km² per cell. The coupling process includes modeling nitrate transport in saturated and unsaturated zones as well as stream-aquifer interactions.

The river network in the Eau-Dyssée platform is represented by a grid-based river network, because the vector network is not available for other modules such as groundwater and surface modules. It would be advantageous to have vector-based modules in Eau-Dyssée for higher resolutions. The model simulates the transport of nitrate variable from surface to aquifers to rivers, and also the transport from surface to

ivers. The modeling framework in this study has the flexibility to add or remove nitrate constituent without excessive programming effort.

The simulated results compared favorably to the measured data on an annual basis. Model results showed that simulated nitrate flux highly depends on the inflow produced by surface and subsurface waters. Variation of mean annual nitrate from year to year can be explained by the hydrologic regime of that year. Larger amounts of nitrate flux are transported during wet years.

In summary, this research demonstrated that it is possible to couple climate models with vector-based river networks on a continental scale. However, many obstacles are encountered which make the process more difficult for novices. Furthermore, a variety of datasets from various geographies complicate modeling because there is no standard for hydrologic modeling like there is for climate modeling. Steady-state pollution modeling can also be coupled with GIS-based river routing model and nitrogen dataset in a regional scale. However, the current schematic processor is not adequate for unsteady pollution transport in large river networks. Finally, an integration of multiple models, such as the Eau-Dyssée platform requires a standard format of data and communication protocols between the models. Establishing standards for environmental modeling will unlock the ultimate potential of each model.

5.2. RECOMMENDATION FOR FUTURE WORK

This research shows continental scale modeling can be greatly improved. Currently it is tedious and difficult to produce surface and subsurface runoff data for RAPID modeling.

One practical solution is to retrieve runoff data produced by the National Weather Service in catchment scales. Currently, grid-based land surface models (e.g. VIC, Noah-

MP, and Mosaic), provide runoff data for river routing models. More importantly, the river-aquifer exchange requires more research in the horizontal water movement in continental river flow scales. Also, the RAPID model is “Input/Output (I/O) bound” for the continental river flow modeling scale data. In this model, the NetCDF file is used to store runoff data, which is the largest file. However, major modifications are required in the I/O operation step of the RAPID model in order to simulate river flow for the entire United States.

This study improves river-aquifer exchange modeling in regional scales. Therefore, the present work could serve as basis for in-stream transport modeling in the Seine basin. Calibration of transport in large river networks, using hundreds of gauges for period of time should be considered because this would significantly enhance flow transport modeling. Hence, a web service for water quality data on a scale of the continental United States for surface and subsurface pollution data is beneficial. Currently the *NHDPlus* dataset includes the USGS gauges feature for the entire United States; this feature could be extended to include water quality parameters.

References

- Alexander, R. B., R. A. Smith, and G. E. Schwarz (2000), Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico, *Nature*, 403(6771), 758–61, doi:10.1038/35001562.
- Alexander, R. B., P. J. Johnes, E. W. Boyer, and R. A. Smith (2002), A comparison of models for estimating the riverine export of nitrogen from large watersheds, *Biogeochemistry*, 57/58, 295–339.
- Alexander, R. B., E. W. Boyer, R. a. Smith, G. E. Schwarz, and R. B. Moore (2007), The Role of Headwater Streams in Downstream Water Quality, *JAWRA*, 43(1), 41–59, doi:10.1111/j.1752-1688.2007.00005.x.
- Alexander, R. B., J. K. Böhlke, E. W. Boyer, M. B. David, J. W. Harvey, P. J. Mulholland, S. P. Seitzinger, C. R. Tobias, C. Tonitto, and W. M. Wollheim (2009), Dynamic modeling of nitrogen losses in river networks unravels the coupled effects of hydrological and biogeochemical processes, *Biogeochemistry*, 93(1-2), 91–116, doi:10.1007/s10533-008-9274-8.
- Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams (1998), LARGE AREA HYDROLOGIC MODELING AND ASSESSMENT PART I: MODEL DEVELOPMENT, *JAWRA*, 34(1), 73–89.
- Atasoy, M., R. B. Palmquist, and D. J. Phaneuf (2006), Estimating the effects of urban residential development on water quality using microdata, *J. Environ. Manage.*, 79(4), 399–408, doi:10.1016/j.jenvman.2005.07.012.
- Bates, P. ., and A. P. . De Roo (2000), A simple raster-based model for flood inundation simulation, *J. Hydrol.*, 236(1-2), 54–77, doi:10.1016/S0022-1694(00)00278-X.
- Bates, P. D., M. S. Horritt, and T. J. Fewtrell (2010), A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling, *J. Hydrol.*, 387(1-2), 33–45, doi:10.1016/j.jhydrol.2010.03.027.
- Berbery, E. H., Y. Luo, K. E. Mitchell, and A. K. Betts (2003), Eta model estimated land surface processes and the hydrologic cycle of the Mississippi basin, *J. Geophys. Res.*, 108(D22), 1–20, doi:10.1029/2002JD003192.
- Bicknell, B. R., J. C. Imhoff, J. L. Kittle, Jr., T. H. Jobes, and A. S. Donigian, Jr. (2005), *Hydrological Simulation Program-Fortran: HSPF Version 12.2 User's Manual*, AQUA TERRA Consultants Mountain View, California.
- Billen, G., J. Garnier, and P. Hanset (1994), Modelling phytoplankton development in whole drainage networks : the RIVERSTRAHLER Model applied to the Seine river system, , 1991, 119–137.

- Billen, G., J. Garnier, A. Ficht, and C. Cun (2001), Modeling the response of water quality in the Seine river estuary to human activity in its watershed over the last 50 years, *Estuaries*, 24(6), 977–993.
- Billen, G., J. Garnier, J. Némery, M. Sebilo, a Sferratore, S. Barles, P. Benoit, and M. Benoît (2007a), A long-term view of nutrient transfers through the Seine river continuum., *Sci. Total Environ.*, 375(1-3), 80–97, doi:10.1016/j.scitotenv.2006.12.005.
- Billen, G., J. Garnier, J.-M. Mouchel, and M. Silvestre (2007b), The Seine system: introduction to a multidisciplinary approach of the functioning of a regional river system., *Sci. Total Environ.*, 375(1-3), 1–12, doi:10.1016/j.scitotenv.2006.12.001.
- Boesch, D. F. et al. (2009), Nutrient enrichment drives Gulf of Mexico hypoxia, *EOS, Trans. Am. Geophys. Union*, 90(14), 117, doi:10.1029/2009EO140001.
- Boyer, E. W., C.L. Goodale, N.A. Jaworski, R.W. Howarth (2002), Anthropogenic nitrogen sources and relationships to riverine nitrogen export in the northeastern U.S.A., *Biogeochemistry*, 57(58), 137-169.
- Brisson, N., B. Mary, and D. Ripoche (1998), STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn, *Agronomie*, 18(5-6), 311–346.
- Brisson, N. et al. (2002), Original article STICS : a generic model for simulating crops and their water and nitrogen balances . II . Model validation for wheat and maize, *Agronomie*, 22, 69–92, doi:10.1051/agro.
- Brisson, N., C. Gary, E. Justes, and R. Roche (2003), An overview of the crop model STICS, *Eur. J. of Agronomy*, 18(3-4).
- Chapra, S. C. (1997), *Surface water-quality modeling*, McGraw-Hill New York, USA.
- Chnebelen, N. S., B. N. Icoullaud, H. B. Ourennane, and A. C. Outurier (2004), Original article The STICS model to predict nitrate leaching following agricultural practices, *Agronomie*, 24, 423–435, doi:10.1051/agro.
- David, C. (2009), Towards river flow computation at the continental scale, 132 pp., The University of Texas at Austin.
- David, C. H., F. Habets, D. R. Maidment, and Z.-L. Yang (2011a), RAPID applied to the SIM-France model, *Hydrol. Process.*, n/a–n/a, doi:10.1002/hyp.8070.
- David, C. H., D. R. Maidment, G.-Y. Niu, Z.-L. Yang, F. Habets, and V. Eijkhout (2011b), River Network Routing on the NHDPlus Dataset, *J. Hydrometeorol.*, 12(5), 913–934, doi:10.1175/2011JHM1345.1.
- David, C. H., Z.-L. Yang, and S. Hong (2013), Regional-scale river flow modeling using off-the-shelf runoff products, thousands of mapped rivers and hundreds of stream

- flow gauges, *Environ. Model. Softw.*, 42, 116–132, doi:10.1016/j.envsoft.2012.12.011.
- Deschesnes, J., J.-P. Villeneuve, E. Ledoux, and G. Girard (1985), Modeling the hydrologic cycle: the MC model. Part I - principles and description, *Nord. Hydrol.*, 16(5), 257–272.
- Ducharne, A. (2008), Importance of stream temperature to climate change impact on water quality, *Hydrol. Earth Syst. Sci.*, 12(3), 797–810, doi:10.5194/hess-12-797-2008.
- Ducharne, A., C. Golaz, E. Leblois, K. Laval, J. Polcher, E. Ledoux, and G. de Marsily (2003), Development of a high resolution runoff routing model, calibration and application to assess runoff from the LMD GCM, *J. Hydrol.*, 280(1-4), 207–228, doi:10.1016/S0022-1694(03)00230-0.
- Dunn, D. (1996), Trends in nutrient inflows to the Gulf of Mexico from streams draining the conterminous United States, 1972-93, *U.S. Geol. Surv. Water-resources Investig. Rep. 96-4113*.
- Durand, Y., G. Giraud, E. Brun, L. Merindol, and E. Martin (1999), A computer-based system simulating snowpack structures as a tool for regional avalanche forecasting, *J. Glaciol.*, 45(151), 469–484.
- Even, S. et al. (2007), New tools for modelling water quality of hydrosystems: an application in the Seine River basin in the frame of the Water Framework Directive., *Sci. Total Environ.*, 375(1-3), 274–91, doi:10.1016/j.scitotenv.2006.12.019.
- Flipo, N., N. Jeannée, M. Poulin, S. Even, and E. Ledoux (2007a), Assessment of nitrate pollution in the Grand Morin aquifers (France): combined use of geostatistics and physically based modeling., *Environ. Pollut.*, 146(1), 241–56, doi:10.1016/j.envpol.2006.03.056.
- Flipo, N., S. Even, M. Poulin, S. Théry, and E. Ledoux (2007b), Modeling nitrate fluxes at the catchment scale using the integrated tool CAWAQS., *Sci. Total Environ.*, 375(1-3), 69–79, doi:10.1016/j.scitotenv.2006.12.016.
- Flipo, N., C. Rabouille, M. Poulin, S. Even, M. Tusseau-vuillemin, and M. Lalande (2007c), Primary production in headwater streams of the Seine basin : The Grand Morin river case study, , 375, 98–109, doi:10.1016/j.scitotenv.2006.12.015.
- Flipo, N., C. Monteil, M. Poulin, C. de Fouquet, and M. Krimissa (2012), Hybrid fitting of a hydrosystem model: Long-term insight into the Beauce aquifer functioning (France), *Water Resour. Res.*, 48(5), doi:10.1029/2011WR011092.

- Foster, S. S. D. (2000), Assessing and Controlling the Impacts of Agriculture on Groundwater—from Barley Barons to Beef Bans, *Q. J. Eng. Geol. Hydrogeol.*, 33, 263–280.
- Fread, D. L., 1993: Flow routing. *Handbook of Hydrology*, D. r. Maidment, Ed., McGraw-Hill, 10.17-10.18.
- Galloway, J. N. et al. (2004), Nitrogen Cycles: Past, Present, and Future, *Biogeochemistry*, 70(2), 153–226, doi:10.1007/s10533-004-0370-0.
- Gao, H., Q. Tang, X. Shi, C. Zhu, and T. Bohn (2010), Water budget record from Variable Infiltration Capacity (VIC) model, in ... *Water Cycle Data Records*, ..., pp. 120–173.
- Garnier, J. (1995), Seasonal succession of diatoms and Chlorophyceae in the drainage network of the Seine River: Observations and modeling, *Limnol. Oceanogr.*, 40(4), 750–765.
- Garnier, J., G. Billen, N. Sanchez, and B. Leporcq (2000), Ecological functioning of the Marne reservoir (upper Seine basin, France), *Regul. Rivers Res. Manag.*, 71(16), 51–71.
- Garnier, J., J. Némery, G. Billen, and S. Théry (2005), Nutrient dynamics and control of eutrophication in the Marne River system: modelling the role of exchangeable phosphorus, *J. Hydrol.*, 304(1-4), 397–412, doi:10.1016/j.jhydrol.2004.07.040.
- Gomez, E., E. Ledoux, J. Monget, and G. De Marsily (2003), Distributed surface-groundwater coupled model applied to climate or long term water management impacts at basin scale, *European Water Resources Association*, (1), 3–8.
- Habets, F. et al. (2008), The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France, *J. Geophys. Res.*, 113(D6), 1–18, doi:10.1029/2007JD008548.
- Harrison, J. a., R. J. Maranger, R. B. Alexander, A. E. Giblin, P.-A. Jacinthe, E. Mayorga, S. P. Seitzinger, D. J. Sobota, and W. M. Wollheim (2009), The regional and global significance of nitrogen removal in lakes and reservoirs, *Biogeochemistry*, 93(1-2), 143–157, doi:10.1007/s10533-008-9272-x.
- Hong, B., D. P. Swaney, and R. W. Howarth (2011), A toolbox for calculating net anthropogenic nitrogen inputs (NANI), *Environ. Model. Softw.*, 26(5), 623–633, doi:10.1016/j.envsoft.2010.11.012.
- Horizon Systems Corporation, cited 2007: National Hydrography Dataset Plus: Documentation. [Available online at <http://www.horizon-systems.com/nhdplus/documentation.php>].
- Howarth, R. W., D. P. Swaney, E. W. Boyer, R. Marino, N. Jaworski, and C. Goodale (2006), The influence of climate on average nitrogen export from large

- watersheds in the Northeastern United States, *Biogeochemistry*, 79(1-2), 163–186, doi:10.1007/s10533-006-9010-1.
- Huang, S., C. Hesse, V. Krysanova, and F. Hattermann (2009), From meso- to macro-scale dynamic water quality modelling for the assessment of land use change scenarios, *Ecol. Modell.*, 220(19), 2543–2558, doi:10.1016/j.ecolmodel.2009.06.043.
- Hurkmans, R. T. W. L., H. de Moel, J. C. J. H. Aerts, and P. a. Troch (2008), Water balance versus land surface model in the simulation of Rhine river discharges, *Water Resour. Res.*, 44(1), 1–14, doi:10.1029/2007WR006168.
- Johnes, P. J. (1996), Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: the export coefficient modelling approach, *J. Hydrol.*, 183(3-4), 323–349, doi:10.1016/0022-1694(95)02951-6.
- Johnson, S. L. (2009), A general method for modeling coastal water pollutant loadings, 338 pp., University of Texas.
- Johnson, S. L., D. R. Maidment, and M. J. Kirisits (2013), TMDL Balance: A Model for Coastal Water Pollutant Loadings, *JAWRA J. Am. Water Resour. Assoc.*, 49(4), 838–850, doi:10.1111/jawr.12044.
- Kaushal, S. S., M. L. Pace, P. M. Groffman, L. E. Band, K. T. Belt, P. M. Mayer, and C. Welty (2010), Land use and climate variability amplify contaminant pulses, *Eos (Washington. DC).*, 91(25), 221–222.
- Kiel, B., and M. Cardenas (2014), Lateral hyporheic exchange throughout the Mississippi River network, *Nat. Geosci.*, 7(May), 413–417, doi:10.1038/NGEO2157.
- Larson, J., A. Craig, J. Drake, D. Erickson III, M. Branstetter, and M. Ham (2007), A massively parallel dynamical core for continental-to global-scale river transport, in *Proceedings of the International Congress on Modelling and Simulation (ModSim 2007)*, pp. 532–538.
- Ledoux, E., G. Girard, and G. De Marsily (1989), Spatially distributed modeling: conceptual approach, coupling surface water and groundwater, *Morel-Seytoux HJ, Ed. Unsaturated flow Hydrol. Model. Pract.*, 435–454.
- Ledoux, E., E. Gomez, J. M. Monget, C. Viavattene, P. Viennot, a Ducharne, M. Benoit, C. Mignolet, C. Schott, and B. Mary (2007), Agriculture and groundwater nitrate contamination in the Seine basin. The STICS-MODCOU modelling chain., *Sci. Total Environ.*, 375(1-3), 33–47, doi:10.1016/j.scitotenv.2006.12.002.
- Lehner, B., K. Verdin, and A. Jarvis (2006), HydroSHEDS TechDoc v10,
- Leopold, L., and W. Langbein (1962), The concept of entropy in landscape evolution, *U.S. Geol. Surv. Prof.*, (Paper 500-A), A1–A20.

- Liang, X., D. P. Lettenmaier, E. F. Wood, and J. Burges (1994), based model of land surface water and energy fluxes for general circulation models, , 99.
- Liu, Z.-J., D. E. Weller, T. E. Jordan, D. L. Correll, and K. B. Boomer (2008), Integrated Modular Modeling of Water and Nutrients From Point and Nonpoint Sources in the Patuxent River Watershed, *J. Am. Water Resour. Assoc.*, 44(3), 700–723, doi:10.1111/j.1752-1688.2008.00200.x.
- Lohmann, D. et al. (1998), The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) phase 2(c) Red–Arkansas River basin experiment:: 3. Spatial and temporal analysis of water fluxes, *Global and Planetary Change*, 161–179.
- Lohmann, D. et al. (2004), Streamflow and water balance intercomparisons of four land surface models in the North American Land Data Assimilation System project, *J. Geophys. Res.*, 109(D7), 1–22, doi:10.1029/2003JD003517.
- Lucas-Picher, P., V. Arora, D. R. Cayan, and R. Laprise (2003), Implementation of a large-scale variable velocity river flow routing algorithm in the Canadian Regional Climate Model (CRCM), *Atmosphere-ocean*, 139–153.
- Maidment, D., 2002. Arc Hydro: GIS for Water Resources. ESRI Press, Redlands, California, ISBN-13: 978-1589480346.
- Marcé, R., and J. Armengol (2009), Modeling nutrient in-stream processes at the watershed scale using Nutrient Spiralling metrics, *Hydrol. Earth Syst. Sci. Discuss.*, 6(1), 501–533, doi:10.5194/hessd-6-501-2009.
- Massei, N., B. Laignel, J. Deloffre, J. Mesquita, A. Motelay, R. Lafite, and A. Durand (2010), Long-term hydrological changes of the Seine River flow (France) and their relation to the North Atlantic Oscillation over the period 1950-2008, *Int. J. Climatol.*, 30(14), 2146–2154, doi:10.1002/joc.2022.
- 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, , 106(2000).
- Mayorga, E., S. P. Seitzinger, J. a. Harrison, E. Dumont, A. H. W. Beusen, A. F. Bouwman, B. M. Fekete, C. Kroeze, and G. Van Drecht (2010), Global Nutrient Export from WaterSheds 2 (NEWS 2): Model development and implementation, *Environ. Model. Softw.*, 25(7), 837–853, doi:10.1016/j.envsoft.2010.01.007.
- McCarthy, G. T., 1938: The unit hydrograph and flood routing. Proc. Conf. of the North Atlantic Division, New London, CT, U.S. Engineer Department, 1-19.
- McKay, L., Bondelid, T., Dewald, T., Rea, A., Moore, R., and Johnston, J. “NHDPlus Version 2: User Guide”, 2012

- Meng, L., and S. M. Quiring (2008), A Comparison of Soil Moisture Models Using Soil Climate Analysis Network Observations, *J. Hydrometeorol.*, 9(4), 641–659, doi:10.1175/2008JHM916.1.
- Meyer, L. H. (2012), Quantifying the Role of Agriculture and Urbanization in the Nitrogen Cycle across Texas, 90 pp., The University of Texas at Austin.
- Miguez-macho, G., and Y. Fan (2012), The role of groundwater in the Amazon water cycle : 1 . Influence on seasonal streamflow , flooding and wetlands, *J. Geophys. Res.*, (January), doi:10.1029/2012JD017539.
- Miller, J. R., G. L. Russell, and G. Caliri, 1994: Continental-scale river flow in climate models, *J. Climate*. 7 (6), 914-928.
- Mooney, R. F., and J. W. McClelland (2012), Watershed Export Events and Ecosystem Responses in the Mission–Aransas National Estuarine Research Reserve, South Texas, *Estuaries and Coasts*, 35(6), 1468–1485, doi:10.1007/s12237-012-9537-4.
- Nash, J. E. and J. V. Sutcliffe, 1970: River flow forecasting through conceptual models part I: A discussion of principles, *J. Hydrol.*, 10 (3), 282–290.
- Olivera, F., M. S. Lear, James S. Famiglietti, and K. Asante (2002), Extracting low-resolution river networks from high-resolution digital elevation models, *Water Resour. Res.*, 38(11), doi:10.1029/2001WR000726.
- Philippe, E., F. Habets, E. Ledoux, P. Goblet, P. Viennot, and B. Mary (2011), Improvement of the solute transfer in a conceptual unsaturated zone scheme: a case study of the Seine River basin, *Hydrol. Process.*, 25(5), 752–765, doi:10.1002/hyp.7865.
- Quintana-Seguí, P., P. Le Moigne, Y. Durand, E. Martin, F. Habets, M. Baillon, C. Canellas, L. Franchisteguy, and S. Morel (2008), Analysis of Near-Surface Atmospheric Variables: Validation of the SAFRAN Analysis over France, *J. Appl. Meteorol. Climatol.*, 47(1), 92–107, doi:10.1175/2007JAMC1636.1.
- Rebich, R. a., N. a. Houston, S. V. Mize, D. K. Pearson, P. B. Ging, and C. Evan Hornig (2011), Sources and Delivery of Nutrients to the Northwestern Gulf of Mexico from Streams in the South-Central United States, *JAWRA*, 47(5), 1061–1086, doi:10.1111/j.1752-1688.2011.00583.x.
- Rivett, M. O., S. R. Buss, P. Morgan, J. W. N. Smith, and C. D. Bemment (2008), Nitrate attenuation in groundwater: a review of biogeochemical controlling processes., *Water Res.*, 42(16), 4215–32, doi:10.1016/j.watres.2008.07.020.
- Runkel, R. L. (2007), Toward a transport-based analysis of nutrient spiraling and uptake in streams, *Limnol. Oceanogr. Methods* 5, 50–62.
- Saleh, F., N. Flipo, F. Habets, A. Ducharne, L. Oudin, P. Viennot, and E. Ledoux (2011), Modeling the impact of in-stream water level fluctuations on stream-aquifer

- interactions at the regional scale, *J. Hydrol.*, 400(3-4), 490–500, doi:10.1016/j.jhydrol.2011.02.001.
- Schaefer, S. C., and M. Alber (2007), Temperature controls a latitudinal gradient in the proportion of watershed nitrogen exported to coastal ecosystems, *Biogeochemistry*, 85(3), 333–346, doi:10.1007/s10533-007-9144-9.
- Seaber, P.R., F.P. Kapinos, and G.L. Knapp, 1987: Hydrologic Unit Maps: U.S. Geological Survey Water-Supply Paper 2294, 63 p.
- Seitzinger, S., R. Styles, E. Boyer, R. Alexander, G. Billen, R. W. Howarth, B. Mayer, and N. Van Breemen (2002), Nitrogen retention in rivers: model development and application to watersheds in the northeastern USA, *Biogeochemistry*, 57(1), 199–237.
- Simley, J., and W. Carswell (2009), The National Map–Hydrography, *US Geol. Surv. Fact Sheet 2009–3054*, (March).
- Smith, R., G. Schwarz, and R. B. Alexander (1997), Regional interpretation of water-quality monitoring data, *Water Resour. Res.*, 33(12), 2781–2798.
- Suarez, M. J., and R. D. Koster (1996), Technical Report Series on Global Modeling and Data Assimilation Volume 9 Energy and Water Balance Calculations in the Mosaic LSM, , 9.
- Tavakoly, A. A., C. H. David, D. R. Maidment, Z.-L. Yang, and X. Cai (2012), An Upscaling Process for Large-Scale Vector Based River Network Using the NHDPlus Dataset, in *AWRA 2012 SPRING SPECIALTY CONFERENCE*, pp. 1–6.
- Tewelde, M., and J. Smithers (2007), Flood routing in ungauged catchments using Muskingum methods, *Water SA*, 32(3), 379–388.
- Thieu, V., G. Billen, and J. Garnier (2009), Nutrient transfer in three contrasting NW European watersheds: the Seine, Somme, and Scheldt Rivers. A comparative application of the Seneque/Riverstrahler model., *Water Res.*, 43(6), 1740–54, doi:10.1016/j.watres.2009.01.014.
- Troy, T. J., E. F. Wood, and J. Sheffield (2008), An efficient calibration method for continental-scale land surface modeling, *Water Resour. Res.*, 44(9), 1–13, doi:10.1029/2007WR006513.
- Vitousek, P., J. Aber, R. Howarth, and G. Likens (1997), Human alteration of the global nitrogen cycle: sources and consequences, *Ecol. Appl.*, 7(November 1996), 737–750.
- Whiteaker, T. L., and S. Johnson (2012), A GIS Tool for Simulating Processes Along Schematic Networks, in *AWRA 2012 Spring Specialty Conference on Geographic Information Systems and Water Resources*, pp. 1–6.

- Whiteaker, T. L., D. R. Maidment, J. L. Goodall, and M. Takamatsu (2006), Integrating Arc Hydro Features with a Schematic Network, *Trans. GIS*, 10(2), 219–237, doi:10.1111/j.1467-9671.2006.00254.x.
- Whitehead, P., E. Wilson, D. Butterfield, and K. Seed (1998), A semi-distributed integrated flow and nitrogen model for multiple source assessment in catchments (INCA): Part II–application to large river basins in south Wales and eastern England, *Sci. Total Environ.*, 210, 559–583.
- Wu, H., J. S. Kimball, N. Mantua, and J. Stanford (2011), Automated upscaling of river networks for macroscale hydrological modeling, *Water Resour. Res.*, 47(3), 1–18, doi:10.1029/2009WR008871.
- Xie, Z., F. Yuan, Q. Duan, J. Zheng, M. Liang, and F. Chen (2007), Regional Parameter Estimation of the VIC Land Surface Model: Methodology and Application to River Basins in China, *J. Hydrometeorol.*, 8(3), 447–468, doi:10.1175/JHM568.1.
- Yang, G., E. P. H. Best, T. Whiteaker, A. Teklitz, and L. Yeghiazarian (2014), A screening-level modeling approach to estimate nitrogen loading and standard exceedance risk , with application to the Tippecanoe River watershed , Indiana, *J. Environ. Manage.*, 135, 1–10, doi:10.1016/j.jenvman.2014.01.003.
- Young, R., C. Onstad, D. D. Bosch, and W. P. Anderson (1989), AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds, *J. soil water ...*, 44(2), 168–173.
- Zhang, Z., F. Tao, J. Du, P. Shi, D. Yu, Y. Meng, and Y. Sun (2010), Surface water quality and its control in a river with intensive human impacts--a case study of the Xiangjiang River, China., *J. Environ. Manage.*, 91(12), 2483–90, doi:10.1016/j.jenvman.2010.07.002.

Vita

In May 1981, Ahmad was born in Ahvaz, Iran, home of the largest river in Iran, the “Karūn”. He is the son of Mohammad Tavakoly and Saideh Afshar. After growing up in Ahvaz and graduating high school in 1999, Ahmad enrolled in the Agriculture Engineering department at Ahvaz University, which is the top school in the field of hydrology and hydraulic engineering in Iran. He received a bachelor’s degree in 2003. Unlike the rest of his family who went into medicine, Ahmad was concerned about the effects of pollution in the Karūn River. He pursued a master’s degree at Ahvaz University and developed his master thesis on the topic of river hydrodynamic modeling and water quality. In 2006 he obtained a Masters of Science in Hydraulic Structure and nominated as a distinguished researcher in the graduate school. From 2006 to 2008 he worked on various projects from hydrodynamic and water quality modeling of rivers to designing green houses. On April 2008 he became a US resident and decided to continue his academic career at the University of Texas at Austin. Ahmad enrolled in the doctoral program in Environmental and Water Resources Engineering at the University of Texas at Austin August 2008. He was awarded the University of Texas At Austin Walter L. and Reta Mae Moore Graduate Fellowship in Water Resources for the 2008-2009 academic years. Since then, he is researching on development of geospatial framework for flow and transport modeling in large scale river networks under the direction of Dr. David R. Maidment and Dr Zong-Liang Yang.

Permanent address: tavakoly@utexas.edu

This dissertation was typed by the author.