THE CATHOLIC UNIVERSITY OF AMERICA

Parameter and Uncertainty Estimation of Physical, Chemical, and Biological processes in Environmental Models using Stochastic Inverse Modeling

A DISSERTATION

Submitted to the Faculty of the

Department of Physics

School of Arts and Sciences

Of The Catholic University of America

In Partial Fulfillment of the Requirements

For the Degree

Doctor of Philosophy

By

Heather Ann Stewart

Washington, D.C.

2016

Parameter and Uncertainty Estimation of Physical, Chemical, and Biological processes in Environmental Models using Stochastic Inverse Modeling

Heather Ann Stewart, Ph.D.

Academic Director: Duília F. de Mello, Ph.D.

Abstract

Surface water is a vital resource for drinking water, food/fishing, recreation, transportation, industry, and agriculture. Human population growth and urbanization threaten this already precious resource via agricultural intensification and waste production. When existing vegetation is cleared for farmland, overland flow from rain is unimpeded and can more easily suspend and transport sediments. As topsoil is removed, farmers must increasingly rely upon fertilizers. Both agricultural runoff and municipal wastewater can contain contaminant like heavy metals and pathogens, as well as excess nutrients that pollute the water bodies they flow into. In small, balanced amounts, nutrients like nitrogen are necessary for healthy vegetation and wildlife. The nutrient loading from runoff and waste water, however, can cause potentially toxic algal blooms followed by heterotrophic bacteria that consume oxygen while feeding on the decaying algae. This results in hypoxia, a complete depletion of dissolved oxygen, and thus wildlife suffocation. Land management and wastewater treatment are used to prevent algal blooms, but the former has been difficult to evaluate quantitatively and the latter is very energy intensive.

This dissertation focuses on modeling these two systems: surface runoff and advanced nitrogen removal from wastewater. Environmental models are often complex combinations of physical, chemical, and biological process with many parameters, some of which cannot be directly measured or have a wide range of reported values in literature. As these models often

inform environmental policy and decision-making, it is also important that predictions incorporate uncertainty and confidence. I have used Bayesian inference with Markov-Chain Monte Carlo simulations to estimate parameters and uncertainty for these systems.

To quantify agricultural runoff, elemental fingerprints were collected from sediment samples of several land use types (agriculture, forest, unpaved roads, and streambanks) in the watershed of Laurel Hill Creek, PA. Fluvial samples to collect suspended sediments were taken following 6 rain events between 2010 and 2011. I used a Chemical Mass Balance (CMB) approach to assess the contribution of each source into the fluvial samples. The results confirm the influence of agricultural runoff as up to 80% of the sediments could be attributed to this land use. We also found that agriculture contribution increases with more intense rain events. This is important as climate change is predicted to increase the intensity of weather events.

Most modern treatment facilities perform nitrogen removal via activated sludge which convert ammonia to nitrogen gas via nitrite and nitrate. Some of the bacteria require oxygen (aerobic autotrophs) and others require an additional carbon substrate (heterotrophic) to thrive and accomplish these goals, but aeration and chemical addition are costly and contribute to the plant's carbon footprint. An experimental pilot reactor incorporating a new anaerobic ammonia oxidizing autotrophic bacteria process called anammox was operated from 2014-2015. Observations from this reactor and batch tests of sludge were used for model discrimination and parameter estimation using inverse modeling. When anammox bacteria are limited by both electron donors and acceptors, only the minimum of the two – rather than the product – was found to effect the growth rate. The probability density functions of estimated parameters can be used to optimize treatment operation by promoting anammox growth and reducing the need for aeration energy or chemical additions.

This dissertation by Heather Ann Stewart fulfills the dissertation requirement for the doctoral degree in Physics approved by Duília de Mello, Ph.D., as Academic Director, Arash Massoudieh, Ph.D., as Research Director, and Sen Nieh, Ph.D, as Reader.

_____

Duilia de Mello, Ph.D., Academic Director

_____

Arash Massoudieh, Ph.D., Research Director

_____

Sen Nieh, Ph.D., Reader

Table of Contents

List of Tables and List of Figures

Figures:

Tables:

Chapter 1: For theologian and mathematician Thomas Bayes,
who encouraged us to update our beliefs based on evidence.

Chapter 2: For marine biologist and author Rachel Carson,
who inspired a movement and new government agency by describing the silence of inaction.

Chapter 3: For Nobel Laureate biochemist and a French Resistance leader Jacques Monod,
who brought chemical reaction kinetics to microbial growth.

Chapter 4: For mathematician and first computer programmer Ada Lovelace,
whose gifts for math and science were nurtured by her mother's distain for poetry.

Acknowledgements

Each chapter is dedicated to a scientist who made this work possible. These are the types of giants Newton claimed helped him see further than others. The thesis as a whole, however, is dedicated to three giants in my life in terms of personality and inspiration: my parents Mary Jean and Mike, and big brother Michael. All three exemplify how to live life well: with an unquenchable thirst for knowledge, dedication to making the world a better place through hard work, and most importantly, a willingness to celebrate beauty and humor in daily life. Stewarts never turn down a glass of good wine, a good book, or a new kitten. Anyone who has met you, knows that one paragraph is inadequate to describe your love, support, and brilliance.

To my academic advisor, Duilia de Mello. Your scientific, academic, and career advice have been invaluable, but in 6 years I don't think I ever entered your office with good news. In the continuity equation for mental energy, you would be a major sink of negative thoughts and a fountain of confident solutions. Thank you for encouraging me to start this degree and also for advocating for my engineering master degree. Você foi o meu pilar e essa universidade tem muita sorte de ter você. Muito obrigada.

To my research advisor, Arash Massoudieh. You opened my eyes to the world of environmental engineering and brought computational physics to life. I have learned so much from you that I worry I may have overwritten prior information and knowledge. Thank you for your remarkable patience and support. دا شـ تم شما مـ ثل ا سـ تادی که بـ ودم شاذ س خوش خـ یـ لی من.

دندرك‌يم ی‌دوس‌ح نم هب هاگ‌ش‌يامزآ رد هگ‌يد ی‌اهوج‌ش‌نادـ مامت. زيچ همه زا منونمم.

To my co-authors and research managers at DC Water: Sudhir Murthy, Haydée De Clippelier, Imre Takács, Bernhard Wett, Tanush Wadhawan, Ahmed Al-Omari, and Allen Gellis.

Thank you for setting the bar so high in terms of research quality and dedication. Your enthusiasm for this work is truly admirable and my experimental results are a reflection of your tireless guidance. Receiving your stamp of approval on a presentation or abstract has been the highest honor. Go raibh maith agat, dank u, köszönöm, danke, ありがとうございますand شكرا .

There are so many other people I have leaned on over the past 6 years. They shaped my life; made me fall in love with Washington, DC; sparked my interest in green infrastructure; put up with my avant-garde cooking style; studied with me; taught me the 'bra'-'ket' quantum mechanics notation; taught me how to pipette and mop; found bugs in my code; gave me excuses to travel the world; helped me prepare for interviews and find a job I love; yoga-ed, kayaked, hiked, biked, and cocktailed with me. They made me Vietnamese sweet soup even though I wanted Phở. You have lifted me up like 100 unique balloons:

The Frye and Stewart clans especially my new sister-in-law Aubrey Merpi, AJ Cabrera, Betsy Milarcik, Bryce Carpenter, Charlotte Thurston, Chris Winnike, Christan Deegan, Emmaris Soto, Jamal Alikhani, Jordan Guerra, Karalyn Dennis, Meriem Alaoui, Norrell Griffiths, Oscar Abello, Pedro Shimizu, Ryan Norris, Sandra Blevins, Sean and Maria Luck, Stasha Orkwizewski, Tri Le, Xiaomeng Liang, the brave men and women of 1337 Newton Street, my hard-working lab mates at DC Water, the wonderful educators I had at Villanova and Catholic Universities, and Dr. Steven Kraemer for giving me an opportunity to escape middle school.

Thank you.

P.S. To my brother: I'm glad you're taller. It made it easier to track your giant footsteps.

Chapter 1 – Introduction

1.1 Motivation

Civilizations rise and fall with water. Planning a city away from a body of water is a modern invention; it is no coincidence that the ancient empires were known for their water infrastructure. Aside from drinking for survival, fresh surface water is used for transportation, food production, industry, recreation, religion, and identity. To borrow the District of Columbia Water and Sewer Authority's motto, Water is Life. It is the first thing we search for when observing other worlds, not because life isn't possible without, but because we can't imagine life without it.

The initial plan for this Ph.D. was to focus on just that: searching for evidence of water in the atmospheres of exoplanets. It is an exciting time in astronomy and astrobiology, but it is also an exciting time for earth. The effects of climate change threaten to disrupt many of the foundations of our global society from debilitating droughts and floods to city-swallowing sea-level rise (Paerl and Huisman, 2008). It is an exciting time to be in water treatment as it exists at the nexus of agriculture, public health, and energy. Waste water treatment facilities are often the largest consumers of electricity in a community. In Washington, D.C., the Potomac Electric Power Company's (Pepco) largest costumer is DC Water at $20.4 million and 27 Mega Watts per year, equivalent to powering over 25,000 homes (DC Water FY 2017 Budget, 2015). Pepco derives the majority of this electricity from non-renewable sources and is therefore contributing to green-house-gas emissions in addition to biological green-house-gas emissions from the waste water facility itself.

Trying to reverse trends in climate change and learning to deal with a changing climate is our generation's challenge. We are the scientists and engineers desperately trying to figure out

how to sustain the members of Apollo 13 using an odd array of tools, except we are both Houston and the astronauts.

1.2 Environmental Context

The reason that so many tax dollars are used to treat waste water is to protect the water quality of the surface water into which the treated water flows, which in turn is often used for drinking water or other types of human consumption (industry, recreation, transportation, agriculture, seafood production etc.). Waste water treatment facility effluent is an example of point source pollution, so named because the source is localized to a single observable spot with a single responsible entity. This effluent can contain a diverse array of constituents from pathogens to heavy metals and pharmaceuticals. The primary concern of waste water treatment, however, is excess nutrient loading (carbon, nitrogen, and/or phosphorus) into receiving water. High concentrations of nitrogen as nitrate (NO_3) in water is toxic at high concentrations and has been linked to methemoglobinemia in infants and toxic effects on livestock (Carpenter et al., 1998). As point-source polluters are easier to hold accountable, this area of research is more developed.

Non-point-source pollution, on the other hand, is diffuse or of unidentifiable origin. The main source of concern from non-point source pollution is sediment transport. When existing vegetation is cleared for farmland, overland flow from rain is unimpeded and can more easily suspend and transport sediments. As topsoil is removed, farmers (or urban landscapers and backyard gardeners) must increasingly rely upon fertilizers, either derived synthetically from the energy intensive Haber-Bosch process or from livestock. Suspended sediments act as a vehicle for transport of contaminants including heavy metals, organics and nutrients into water bodies. The scope of this research is limited to source identification for non-point-source pollution and

modeling advanced nitrogen treatment for point-source pollution. Both research areas are vital for regulators charged with protecting resources like the Chesapeake Bay.

Table 1-1 Break down of surface water pollution sources. Modified from Carpenter et al., 1998.

Point source	**Non-point source**
Wastewater effluent (municipal and industrial)	Runoff from agriculture (including return flow from irrigated agriculture)
Runoff and infiltration from animal feedlots	Runoff from pasture and range
Runoff and leachate from waste disposal	Urban runoff from unsewered areas
Runoff from mines, oil fields, unsewered industrial sites	Septic tank leachate and runoff from failed septic systems
Storm sewer outfalls from cities	Runoff from construction sites <2 hectares
Overflows of combined storm and sanitary sewers	Runoff from abandoned mines
Runoff from construction sites >2 hectares	Atmospheric deposition over a water surface
	Activities on land that generate contaminants, such as logging, wetland conversion, construction, and development of land or waterways

If plants and animals need naturally occurring nutrients to live, why are we concerned with allowing too much to enter surface water or aquifers? A small amount of nutrient loading is required for a healthy river or estuary. However, the amount of nutrients that humans introduce can wreak havoc on the ecological balance and cause eutrophication (Monballiu et al., 2013). The most extreme example of such an imbalance is an algal bloom (also known as a red or brown tide) and the resultant oxygen dead zones. Algal blooms consist of a widespread explosion in algae populations when excess nitrogen and/or phosphorus enter a water body. Sunlight and warmer temperatures exacerbate the problem. Heterotrophic bacteria feed on the decaying algae and consume dissolved oxygen until the waters become anoxic. At this point all of the aquatic life that requires oxygen suffer as well as wildlife (or humans) that rely on the aquatic life. Some algal blooms can even be toxic for humans or other animals, causing long-term liver, digestive, and neurological damage from contact (Paerl and Huisman, 2008). Figure 1-1 shows a photo of an algal bloom in nearby Virginia; Figure 1-2 shows a true color satellite image of the Chesapeake

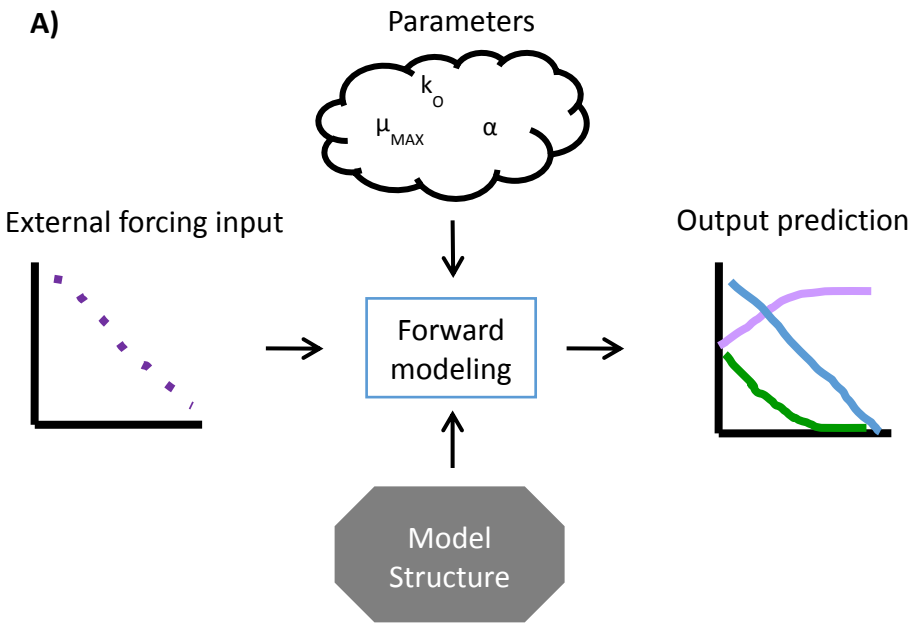
Bay and the pervasiveness of algal blooms. The Bay watershed contains both agriculture (light green and beige regions) and some of our country's largest urban areas (white and grey regions).

Figure 1-1. Close up of an algal bloom in the Chesapeake Bay at Norfolk, Virginia, on April 20, 2009 (courtesy of The Virginian-Pilot newspaper).

Figure 1-2. Landsat image of the Chesapeake Bay with Washington, D.C., highlighted.

Soil erosion and excessive nutrient loading are two major threats to water quality and environmental health. Over the North Atlantic Ocean coastline, non-point sources of nitrogen are up to nine times greater than from waste water treatment facilities (Carpenter et al., 1998). Best management practices to reduce soil erosion require a good understanding of where these sediments originate. Erosion rates from developed watersheds can be 50 and 500 times greater

than those for agriculture and for undisturbed vegetation, respectively (Novotny and Olem 1994). Typical methods involve unmixing models assuming chemical mass balances of elemental profiles in downstream suspended-sediments from samples taken from different land use/land cover types (Davis and Fox, 2009). Most modern treatment facilities perform nitrogen removal via activated sludge bacteria which convert ammonia to nitrogen gas. These reactions are modeled using Activated Sludge Models (Henze et al., 1987). Environmental systems are represented by combinations of physical, biological, and chemical models with many parameters, some of which cannot be directly measured. However, relevant data can be used to estimate these parameter values and their uncertainties via inverse modeling. The demand for realistic models is proportional to the importance of the system. The goal of this dissertation was to improve or validate models that quantify sources of non-point source pollution and estimate parameters and uncertainty of point-source pollution treatment using Bayesian inference and Markov Chain Monte Carlo (MCMC) simulations.



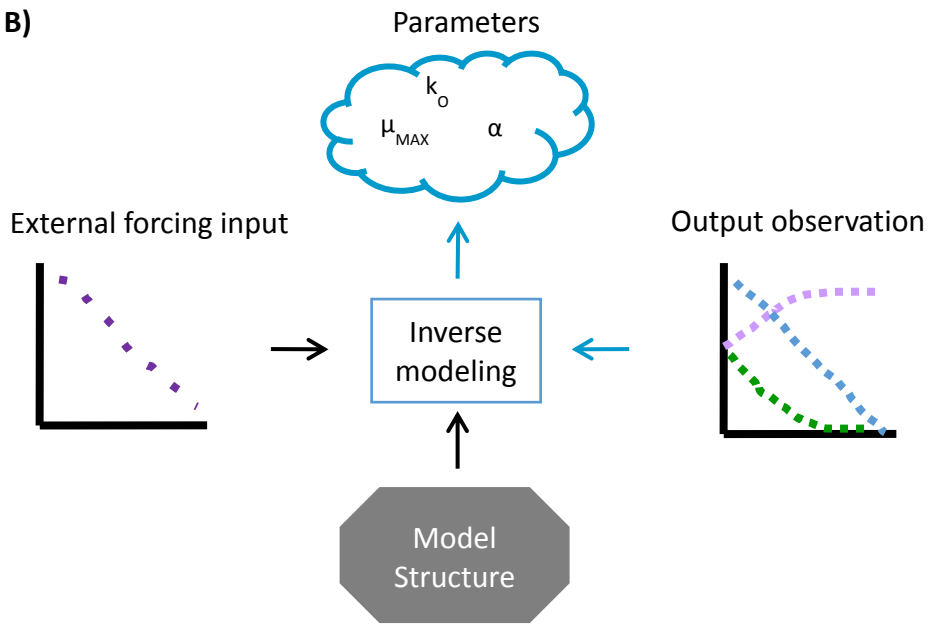


Figure 1-3. Schematic of A) Forward modeling and B) Inverse Modeling. Parameters cloud represents distributions rather than point values.

1.3 Bayesian Inverse Modeling

Current methods of sediment source apportionment and activated sludge parameter estimation produce single value results and (sometimes) an uncertainty estimate. Bayesian inference with MCMC provides more information regarding the credibility of its results via probability distribution functions for each parameter and incorporates sources of uncertainty and prior knowledge (Sharifi et al., 2014). The problem with manual model calibration when modeling complex systems like activated sludge reactors, is that it cannot be guaranteed that the obtained parameter set is the only set capable of reproducing the observations. Uniqueness and identifiability are significant issues in these models (Arnaldos et al., 2015). This research will contribute by producing reliable evidence to promote evidenced-based best land management practices. Likewise, more realistic and plant-specific sludge parameters will produce more reliable models which can then predict optimal full-scale operating conditions.

Although the methodology of each chapter was slightly different in its application of Bayesian parameter estimation, they all follow this format: observations were gathered from both inputs and outputs of a system, one or more model structures and error structures were selected, prior knowledge about parameters was used to form prior probability distributions of values, the parameter space was sampled using MCMC simulations such that new information could be used to update the prior estimates to generate posterior probability distributions with the associated correlations between parameters, sensitivities between parameters and observations, and most importantly, the uncertainty of a parameter and the model predictions that rely upon it.

1.4 Chapters Overview

The topics are put in order of increasing system complexity as well as the order in which the research was completed.

Chapter 2 Laurel Hill Creek Sediment Source Apportionment

A Bayesian Chemical Mass Balance (CMB) approach was used to assess the contribution of potential sources for fluvial samples from Laurel Hill Creek in southwest Pennsylvania. The Bayesian approach provides the joint probability density functions of the sources' contributions considering the uncertainties due to source and fluvial sample heterogeneity and measurement error. Both elemental profiles of sources and fluvial samples and ^{13}C and ^{15}N isotopes were used for the source apportionment. The sources considered include stream bank erosion, forest, roads and agriculture (pasture and cropland). Agriculture was found to have the largest contribution, followed by stream bank erosion. Also, road erosion was found to have a significant contribution in three of the samples collected during lower intensity rain events. The source apportionment was performed with and without isotopes. The results were largely consistent, however, the use of

isotopes was found to slightly increase the uncertainty in most of the cases. The correlation analysis between the contributions of sources shows strong correlations between stream bank and agriculture whereas roads and forest seem to be less correlated to other sources. Thus the method was better able to estimate roads and forest contributions independently. The hypothesis that the contributions of sources are not seasonally changing was tested by assuming that all ten fluvial samples had the same source contributions. This hypothesis was rejected, demonstrating a significant seasonal variation in the sources of sediments in the stream.

Chapter 3 Dual Limitation Modeling

Substrate limitation occurs frequently in wastewater treatment and knowledge about microbial behavior at limiting conditions is essential for the use of biokinetic models in system design and optimization. Monod kinetics are well-accepted for modelling growth rates when a single substrate is limiting, but several models exist for treating two or more limiting substrates simultaneously. In this study three dual limitation models (*multiplicative*, *minimum*, and *Bertolazzi*) were compared based on experiments using nitrite-oxidizing bacteria (limited by dissolved oxygen and nitrite) and Anammox bacteria (limited by ammonium and nitrite). A deterministic likelihood-based parameter estimation followed by Bayesian inference was used to estimate model-specific parameters. The *minimum* model outperformed the other two by a slight margin in three ways. 1) Parameters estimated using this model were closest to parameters estimated from single limitation batch tests. 2) Simulations based on model-specific parameters best described the experimental observations. 3) Simulations based on single limitation parameters best described the experimental observations. Full-scale simulations of mainstream deammonification reactor indicate that dual substrate limitation model selection can have as much as a 75% difference in predicted performance.

Chapter 4 Mainstream Deammonification

This chapter deals with modeling an advanced nitrogen removal pilot that incorporates the anammox process. The populations of bacteria that perform anammox are slow growing and can easily be dominated by other groups (aerobic ammonia-oxidizing, aerobic nitrite-oxidizing, and heterotrophic denitrifying bacteria) so that operational conditions must be carefully monitored (Monballiu et al, 2013). Current procedures assume default values for parameters such as maximum reaction rates and half-saturation coefficients for each of the bacteria populations. Each of these populations has a diverse composition of strains that vary between treatment plants and with time due to physical and chemical environmental conditions (temperature, substrate, dissolved oxygen concentrations etc). Batch tests are used whenever possible to isolate some populations and to determine their properties, but some experiments have too many unknowns for a deterministic solution. Data from these tests as well as pilot reactor simulations were used to generate probability density functions (PDFs) of difficult-to-obtain reaction parameters via Bayesian inverse modeling and MCMC simulations.

Chapter 2 - Sediment Source Apportionment using Bayesian Chemical Mass Balance and Isotope Fingerprinting

2.1 Introduction

Excessive soil erosion in watersheds can have a wide range of adverse environmental and economical effects including negative impacts on water quality, aquatic life habitat, crop productivity, reservoir sedimentation, restriction of light penetration, reduction of suitable spawning grounds for fish, reduction of primary production and causing undesired changes in channel morphology (Stutter et al., 2007; Davis and Fox, 2009; Collins et al., 2010; Owens and Xu, 2011; Mukundan et al., 2012). In addition, suspended sediments act as a vehicle for transport of contaminants including heavy metals, organics, and nutrients into water bodies (Sibbesen and Sharpley, 1997; Palmer and Douglas, 2008; Ballantine et al., 2009; Massoudieh et al., 2010; Massoudieh et al., 2013). The potential damage caused by excess nutrients entering surface water has been discussed in the introduction of this thesis.

In order to base watershed management strategies aimed at controlling and minimizing soil erosion and therefore nutrient loading on sound scientific knowledge, the contribution of various sources or land uses to the sediments of receiving water bodies must be quantified (Owens and Xu, 2011; Mukundan et al., 2012). When the potential sources of sediments are assumed to be known, various physical or chemical properties of the sediments including their elemental profiles, the ratios of various rare stable and unstable isotopes to the dominant isotope (or heavy to light isotope; also known as isotope ratios), magnetic properties, particle size distributions, density and color have been used to infer their contribution in receiving water bodies. (Davis and Fox, 2009; Martínez-Carreras et al., 2010). These methods are collectively referred to as sediment fingerprinting. Various tracer-based approaches in sediment erosion and sediment fingerprinting

studies have been extensively reviewed (Walling, 2005; Davis and Fox, 2009; Guzman et al., 2013). These reviews and recent literature indicate that among the methods that have been used to apportion sources in sediment fingerprinting in watersheds, Chemical Mass Balance Modeling (CMB) (Kelley and Nater, 2000; Minella et al., 2008; Collins et al., 2010; Collins et al., 2012; Massoudieh et al., 2013; Massoudieh and Kayhanian, 2013) and stable isotopes (^{15}N, ^{13}C, ^{18}O) (Bellanger et al., 2004; Fox and Papanicolaou, 2007; Fox, 2009) have become most common.

CMB is based on inferring the contribution of a number of possible sediment sources from the elemental composition of a select group of elements in the sources' (land in watershed) and fluvial (suspended in stream or streambed) sediments. The underlying assumption in all sediment fingerprinting methods including CMB is that the elements used as fingerprints are preserved as the sediments are transported from the source to the sampling location (Walling et al., 1999). When stable isotopes such as ^{15}N and ^{13}C are used for sediment source apportionment, it is often assumed that the isotope ratios (i.e. the ratio of the rare isotope to the dominant isotope of an element, e.g. $^{13}C/^{12}C$) are preserved and also that the isotopic composition of displaced organic carbon and nitrogen does not vary with time (Bellanger et al., 2004; Fox, 2009). This means that the rates of solid-water exchange or transformation of substances containing isotopes are the same for the rare stable isotopes and the corresponding abundant isotope. Although the preferential biological uptake and solids-water mass exchange due to the difference in molecular diffusion of the light and heavy isotopes may result in the alteration of the isotope ratios (Jaisi et al., 2010; Jaisi et al., 2011) this effect may be considered small if the transport time of the sediments from the source to the sampling location is short. In some studies the isotope ratio of the mixture of the fluvial samples is considered to be the weighted average of the isotope ratios of the mixing sources weighted based on the contribution of each source (Fox and Papanicolaou, 2007). From a mass

balance perspective, this implies that all the sources have a similar concentration of the light (abundant) elements for which the heavy (rare) isotopes are used for fingerprinting. In other words when the isotope ratio or "δ" value of the fluvial samples is assumed to be a linear combination of the ratios of the contributing sources without taking into account the content of the corresponding light (abundant) isotopes, the inherent assumption is that the content of the light isotopes in the sources are similar. Mukundan et al., (2010) used CMB and isotopes $\delta^{15}N$ and $\delta^{137}Cs$ separately to apportion sources of suspended sediments and got comparable results. Other authors have found large variations in Total Nitrogen (TN) and $\delta^{15}N$ of streambed sediments and have attributed this to the possibility of biogeochemical factors affecting the $\delta^{15}N$ in temporary storage (Fox et al., 2010). Mulholland et al. (2009) showed that the uptake of NO_3 by denitrifying bacteria in the sediments can significantly affect the $\delta^{15}N$ ratio of dissolved nitrate. However the inorganic nitrogen associated with sediment particles probably undergoes a slower transformation as a result of hydrolysis and nitrification which produce NO_3.

When CMB is used for sediment source apportionment, it is important to recognize and quantify the effect of various sources of uncertainty including

- heterogeneity in sources' and fluvial samples' elemental composition
- errors in analytical methods used to measure concentration of elements
- model structural error due to the violation of the mass balance assumption as a result of sediment-water mass exchange, omission of possible contributing sources, and preferential transport of sediment particles based on their size.

To quantify the effect of these uncertainties on the estimated contribution of sources, stochastic methods based on Bayesian inference or other Monte Carlo-based methods have been incorporated

into mixing models (Franks and Rowan, 2000; Fox and Papanicolaou, 2008; Palmer and Douglas, 2008; Keats et al., 2009; Massoudieh et al., 2013; Massoudieh and Kayhanian, 2013).

A sediment source study was conducted in Laurel Hill Creek, Pennsylvania, from 2010-11 using the sediment-fingerprinting approach (Harris et al., 2001). Sources were identified as roads, stream banks, agriculture (cropland and pasture), and forest. Tracers used to apportion sediment between the sources included elemental and stable isotope analysis. A deterministic multivariate statistical analysis using an unmixing model was used to discriminate sediment sources (Collins et al., 2010; Gellis and Walling, 2011). Results using 10 samples collected from 6 storm events indicated that agriculture was the major source of sediment (53%) followed by stream banks (30%), unpaved roads (17%), and forest (<1%).

In this study, this dataset collected by Sloto et al., (2012) (elemental profiles and the isotope ratios of ^{13}C and ^{15}N) was used with a Bayesian Markov Chain Monte Carlo approach to infer the contribution of sources in fluvial sediments in Laurel Hill Creek. The Bayesian CMB model used is based on Massoudieh et al., (2012) with some modification in order to include isotope ratios in the inference of contribution of sources. The information from the isotope ratios is based on the assumption of mass balance of the rare and abundant isotopes as the sediments are transported. The analysis was performed with and without the isotopes for comparison. One objective of the study was to contrast sediment-source results using elemental profiles and stable isotopes as tracers. Temporal variability of the sources' contributions and their dependence on environmental factors relevant to mobilizing sediments from different sources (rain intensity and flow hydrograph) was also analyzed by applying the model collectively to the data from all fluvial samples (in addition to independently) and then comparing the ability of the model to reproduce the elemental profiles.

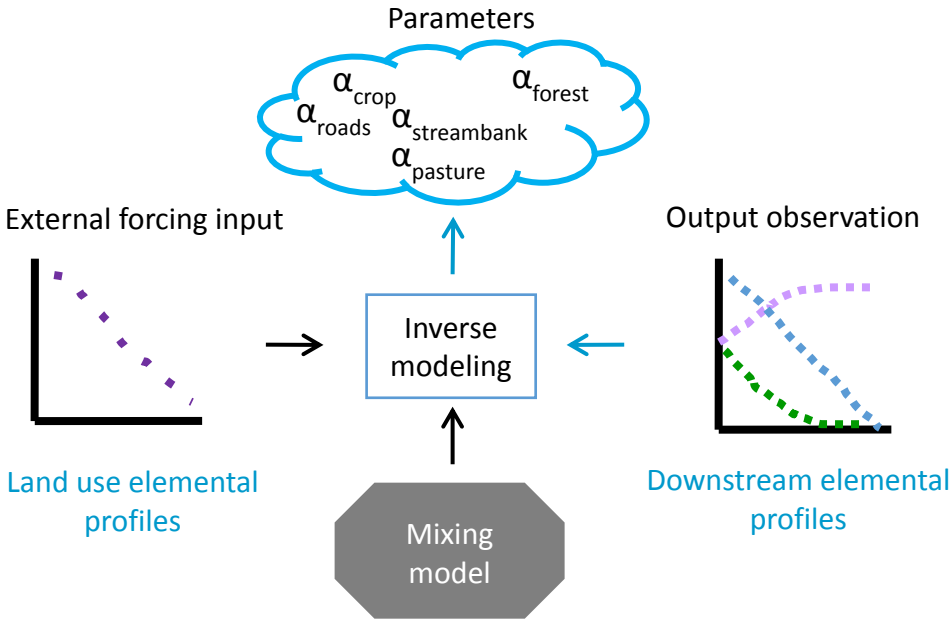


Figure 2-1. Inverse modeling schematic for sediment source apportionment.

2.1.1 Description of the study area

Laurel Hill Creek is a 324 km^2 watershed mostly in Somerset County, PA, with small areas extending into Fayette and Westmoreland Counties (Figure 2-1). The Laurel Hill Creek watershed lies within the Allegheny Mountain Section of the Appalachian Plateaus Physiographic Province. The valleys are broad, undulating surfaces with shallow to deep stream incision (Mulholland et al., 2009). Elevation in the Laurel Hill Creek ranges from 396m at the mouth of the watershed at Confluence, PA, to approximately 884 m along the ridge tops. Underlying geology in the watershed consists of sedimentary rocks of Pennsylvanian and Mississippian ages.

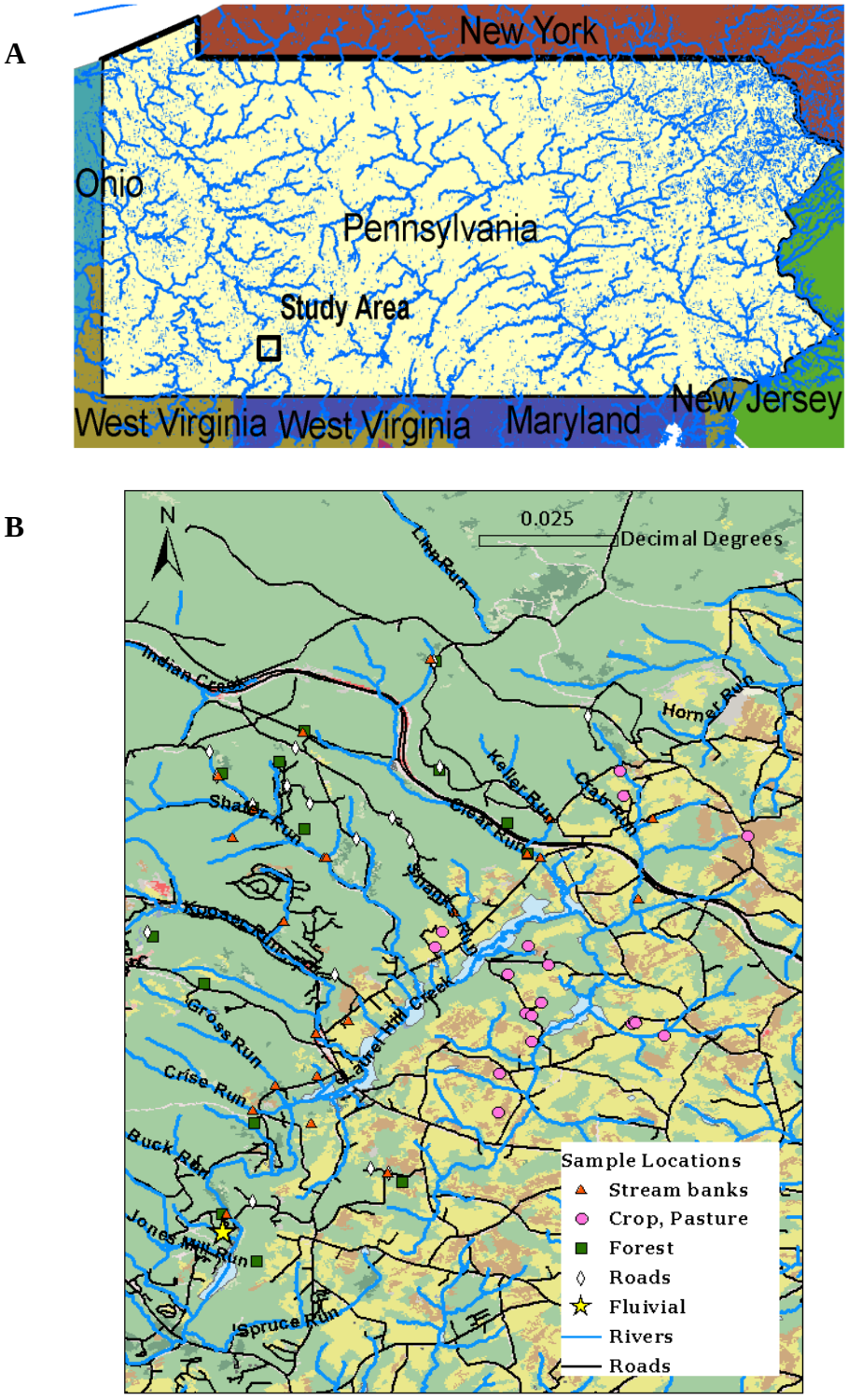


Figure 2-2. A) Location of the study site **B)** Laurel hill creek and the locations of source and fluvial samples.

The climate of the Laurel Hill Creek watershed is continental (Sloto et al., 2012). The daily maximum temperature ranges from about 24 to 29 degrees Celsius in July, and the daily minimum temperature ranges from about -9 to -7 degrees Celsius in January (Fox et al., 2010). The annual precipitation (1971-2000) in the watershed varies spatially with elevation from 1331 mm at Laurel Mountain to 1076 mm at Somerset (NOAA, 2001). The average annual snowfall was 1394 mm for Confluence during 1971–2000.

Land use in Laurel Hill Creek watershed at the time of study was 63.4 percent forest, 27.2 percent agricultural, 4.9 percent residential, 3.0 percent wetlands and open water, and 1.6 percent commercial/industrial and mining (Sloto et al., 2012). Together forested and agricultural areas account for more than 90 percent of the watershed. The upper one-third of the watershed is mainly agricultural (cropland and pasture). Agricultural land use predominantly coincides with the low relief areas. The watershed also has great recreational value, including state parks and ski resorts.

2.1.2 Sample collection and analysis

Sample collection and laboratory analysis was performed by Sloto et al. (2012). Potential sediment sources in Laurel Hill Creek were identified as: (1) stream banks, (2) forest, (3) unpaved roads, (4) cropland, and (5) pasture. Site selection for sampling sediment in areas with forest and agriculture was based on landowner permission and the ability to obtain a spatially representative data set. Unpaved roads were identified on topographic maps (7.5 minute) and distinguished in the field as ‘semi-paved sand or gravel’, ‘maintained sand and gravel’, and ‘non-maintained sand and gravel’. Samples from unpaved roads were obtained by sweeping the surface sediment with a small broom into a plastic pan. Approximately 50 m of road were sampled at each site. Soil samples from agricultural and forested areas were collected from the soil surface with a plastic

hand shovel. To account for variability in the tracer properties at agricultural and forested sites, sediment was collected across transects and composited into one sample. Transects were typically 100 by 30 m.

Site selection for sampling stream banks was based on a spatial analysis of streams in the Laurel Hill Creek watershed (Sloto et al., 2012). The streams were classified by Strahler order into first, second, third, and fourth order, where a first order branch has no further bifurcations, second order has one split, and so on. The lengths of streams in each order were summed and samples collected for each stream order were weighted to the total length of each stream order. Eroding stream banks were sampled from the bottom to the top of the bank face. Samples were collected at three to five transects spaced 10 m apart and composited into one sample. Fluvial (suspended-sediment) samples were collected during 6 rain events between March 2010 and April 2011 at the U.S. Geological Survey (USGS) streamflow station (ID 03079600) during high (storm) flows using an automatic sampler (Tables 2-1 through 2-3).

2.1.3 Laboratory Analyses for Sediment Fingerprinting

Agriculture (cropland and pasture), forest, unpaved road, and stream bank samples were taken to the laboratory, dried at 60°C, disaggregated using a pestle and mortar, and wet-sieved through a 63-micron polyester sieve to remove sand. Sample weights before and after sieving were recorded to determine the percentage of sand in the samples.

The silt and clay fractions (less than 63 microns) of suspended sediment, upland soil, and channel corridor samples were sent to the USGS Geology Discipline research laboratory in Denver, Colorado, for elemental analyses. The elements and isotopes that were measured are listed in Table 2-1. At the USGS laboratory, samples were analyzed using inductively coupled plasma

combined with mass spectrometry (ICP-MS) after multi-acid decomposition (a mixture of hydrochloric, nitric, perchloric, and hydrofluoric acids). Specific details regarding this method can be found in Taggart (2002).

Table 2-1. The starting time and sampling duration for each of the 10 fluvial samples.

Sample	Start	Sampling duration (hr)
1	3/12/10 8:00	1.68
2	3/12/10 9:42	1.38
3	5/17/10 17:00	11
4	5/18/10 5:00	11
5	9/30/10 12:30	6
6	9/30/10 19:30	17
7	11/16/10 20:30	16
8	4/5/11 5:00	7
9	4/5/11 12:15	7.75
10	4/19/11 20:00	13

Samples were analyzed for carbon and nitrogen stable isotopes ($^{13}C/^{12}C$ and $^{15}N/^{14}N$), total carbon (C), and total nitrogen (N) at the University of California Stable Isotope Laboratory (Table 2-2 and 2-3) using an Elementar Vario EL Cube (ELEMENTAR Analysensysteme, GmbH, Hanau, Germany) elemental analyzer interfaced to a Sercon 20-20 isotope ratio mass spectrometer (Sercon Ltd., Cheshire, UK). Samples were combusted at 1,000°C in a reactor packed with cerium dioxide, copper oxide, and lead chromate. Following combustion, oxides were removed in a reduction reactor (reduced copper at 650°C). Water was removed with magnesium perchlorate. Carbon dioxide was removed from the carrier stream by an adsorption trap allowing nitrogen to be analyzed. Following the completion of the nitrogen analysis, the adsorption trap was heated releasing the trapped carbon dioxide for analysis.

Carbon and nitrogen isotope values ($\delta^{13}C$ and $\delta^{15}N$) are reported in per mil (‰) notation with respect to Vienna Pee Dee Belemnite (PDB) and atmospheric N_2, respectively. During analysis, samples were interspersed with laboratory standards, which were previously calibrated against National Institute of Standards and Technology (NIST) Standard Reference Materials. The preliminary isotope ratio for each sample was measured relative to reference gases analyzed with the sample. The long-term standard deviation is 0.2‰ for $\delta^{13}C$ and 0.3‰ for $\delta^{15}N$. The samples analyzed for carbon were acid fumed according to the procedures of Harris et al., (2001)

Two sources of error were determined for the fingerprint dataset - one from field sampling and the other from analytical errors. Field sampling error was determined by collecting replicates of the source samples at selected locations. Analytical errors were discerned by taking a split of the field sample. Results of the field and analytical errors were presented in Sloto et al., (2012). A bracketing technique was used to exclude the elements with their fluvial concentrations significantly above or below the maximum and minimum of source samples respectively. Only elements whose maximum and minimum value in the fluvial samples was within the maximum and minimum values in the source samples were considered. The elements that satisfied this requirement are highlighted in Tables 2-2 and 2-3 and Figure 2-5. The Kruskal–Wallis *H*-test and Mann-Whitney *U*-test have been used to verify the ability of individual tracers to discriminate source categories in previous studies (Collins et al., 1997; Collins et al., 2013). These methods are based on the statistical significance of the difference between the mean of a certain fingerprinting property for one source class verses others and not based on the physical or chemical conservativeness of the tracers. In a Bayesian framework, if a piece of information does not provide any knowledge about the model parameters its inclusion or lack thereof will not affect the posterior distribution. So although inclusion of those tracers may not bear any benefits, it does not

adversely affect the outcome of the analysis. On the other hand, if inclusion of a particular tracer can impact the posterior distributions of contributing sources, it means that the tracer has some level of ability to discriminate between the sources.

2.1.4 Bayesian Source apportionment using chemical mass balance and rare stable isotopes

The chemical mass balance model developed by Massoudieh et al., (2013) was modified to incorporate the information from rare isotopes in conjunction with the elemental profiles in order to infer the contribution of sources. The mass balance method for the elemental composition of the suspended sediments can be written in matrix form as:

$$\mathbf{C}=\mathbf{Y}\bullet\mathbf{X}$$

$$\begin{bmatrix} Al \\ As \\ Bi \\ \vdots \end{bmatrix}_{Fluvial} = \begin{bmatrix} Al & Al & Al & Al \\ As & As & As & As \\ Bi & Bi & Bi & Bi \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}_{Sources} \cdot \begin{bmatrix} X \\ X \\ X \\ X \end{bmatrix}_{Contributions} \qquad \textit{(Eq. 2-1)}$$

where $\mathbf{C}=[c_i]_{I\times 1}$ is a vector representing the true elemental composition of the fluvial sample normalized by the sum of the measured elemental concentrations in the recipient elemental profile, henceforth referred to as the fluvial sample elemental composition vector. $\mathbf{Y}=\left[y_{ij}\right]_{I\times J}=\left[\varsigma_{ij}\Big/\sum_{j=1}^{J}\varsigma_{ij}\right]_{I\times J}$ is the *true* source elemental composition matrix, each of its columns representing the normalized elemental composition of sources by the sum of the measured elements' concentrations in each source and ς_{ij} is the actual elemental content (mg/kg) of element *i* in source *j*. $\mathbf{X}=\left[x_j\right]_{J\times 1}$ is the source contribution vector, containing the true values of fractional contributions of each source (as expressed by the sum of the measured elements) into the fluvial

sample. *I* is the number of elements measured and *J* is the number of sources considered. The contribution of each sediment source, ξ_j then can be calculated using the following equation:

$$\xi_j = \frac{x_j \sum_{i=1}^{J} \varsigma_{ij}}{\sum_{j=1}^{I} \left(x_j \sum_{i=1}^{J} \varsigma_{ij} \right)} \qquad \textit{(Eq. 2-2)}$$

Since X is defined as the fractional contribution of sources the sum of its elements should be unity:

$$\sum_{j=1}^{J} x_j = 1 \qquad \textit{(Eq. 2-3)}$$

The sediment used throughout this analysis was sieved to less than 63 microns and reflects the sediment size class where most tracer concentration and activity would be found. The size and organic matter corrections were not applied due to lack of information about the exact particle size distribution of sources and fluvial sediments and the inadequacy of the knowledge about how size and organic matter content affect each element and this may result in some error. Applying size and organic correction factors in the sediment fingerprinting approach is often a 'black box' approach (Koiter et al., 2013) where the sign of the slope of between size and tracer concentrations can be positive or negative with different tracers (Gellis and Noe, 2013). Further research is needed to determine how important these correction factors are in the sediment-fingerprinting approach. It should be noted that due to measurement errors and heterogeneities in the elemental composition of both the fluvial and source samples, one can never know the exact values of the true fluvial sample elemental composition vector, $\mathbf{C}$, or the true source elemental composition matrix $\mathbf{Y}$; and therefore one can never know the true source contribution vector $\mathbf{X}$. To incorporate the rare isotopes, the isotope ratios as defined by the concentration of rare isotopes normalized by

the concentration of the corresponding abundant isotopes is used. The true isotope ratio of rare isotopes in the fluvial samples can be found using the following relationship:

$$\gamma_i = \frac{\sum_{j=1}^{J} x_j y_{ij} \varphi_{ij}}{\sum_{j=1}^{J} x_j y_{ij}} \qquad \textit{(Eq. 2-4)}$$

where γ_i is the isotope ratio of rare isotope of element *i* in the fluvial sample, and φ_{ij} is the isotope ratio of the rare isotope of element *i* in source *j*. y_{ij} in Eq. (2-4) is the fraction of abundant element corresponding to the rare isotope *i* in source *j*.

The goal here is to infer the posterior probability distribution of $\mathbf{X}$ based on observed elemental compositions of sources and fluvial samples henceforth referred to as $\tilde{\mathbf{Y}}$ and $\tilde{\mathbf{C}}$ using Bayesian inference. Considering that the prior distributions of $\mathbf{X}$ and $\mathbf{Y}$ are independent, based on Bayes' theorem, the posterior distribution of $\mathbf{X}$ and $\mathbf{Y}$, given the observed source and fluvial sample elemental compositions can be expressed as:

$$\pi(\mathbf{Y},\mathbf{X},\boldsymbol{\Lambda},\boldsymbol{\Lambda}_{is} \mid \tilde{\mathbf{C}},\tilde{\mathbf{Y}},\tilde{\boldsymbol{\Gamma}},\tilde{\boldsymbol{\Phi}}) \propto \pi(\tilde{\mathbf{C}} \mid \mathbf{Y},\mathbf{X}) \cdot \pi\left(\tilde{\boldsymbol{\Gamma}} \mid \mathbf{Y},\mathbf{X},\boldsymbol{\Phi}\right) \cdot \pi\left(\tilde{\boldsymbol{\Phi}} \mid \boldsymbol{\Phi}\right) \cdot \pi(\tilde{\mathbf{Y}} \mid \mathbf{Y}) \pi(\mathbf{Y}) \pi(\mathbf{X}) \pi\left(\boldsymbol{\Phi}\right)$$

(Eq. 2-5)

where $\boldsymbol{\Lambda}$ is the variance-covariance matrix for the observation error for fluvial samples' elemental composition, $\boldsymbol{\Lambda}_{is}$ is the variance-covariance matrix for the observed error of the isotope ratios in the fluvial samples, $\tilde{\mathbf{C}} = \left[\tilde{c}_{il}\right]_{I \times L}$ is the matrix containing the observed samples' elemental profiles of the fluvial samples where $\tilde{c}_{il}$ is the normalized observed concentration of element *i* in fluvial sample j, $\tilde{\mathbf{Y}}$ is the observed elemental profile of sources, $\tilde{\boldsymbol{\Phi}} = \left[\tilde{\varphi}_{ij}\right]_{I' \times J}$ is the matrix containing the isotope ratios in the sources, I' is the number of rare isotopes used in the study, and $\tilde{\boldsymbol{\Gamma}} = \left[\tilde{\gamma}_{il}\right]_{I' \times L}$ is the matrix containing the isotope ratios in the fluvial samples. In Eq. (2-5)

$\pi(\tilde{\mathbf{C}}|\mathbf{Y},\mathbf{X})\cdot\pi(\tilde{\boldsymbol{\Gamma}}|\mathbf{Y},\mathbf{X},\boldsymbol{\Phi})\cdot\pi(\tilde{\boldsymbol{\Phi}}|\boldsymbol{\Phi})\cdot\pi(\tilde{\mathbf{Y}}|\mathbf{Y})$ is the likelihood function, and $\pi(\mathbf{Y}),\pi(\mathbf{X})$ and $\pi(\boldsymbol{\Phi})$ are the prior distributions for $\mathbf{Y}$, $\mathbf{X}$, and $\boldsymbol{\Phi}$ respectively. After inspecting the histograms of normalized elemental contents of sources, it was found that the c/(1-c) transformation of the normalized elemental contents can be reasonably assumed to follow a log-normal distribution. This transformation also ensures that the normalized elemental contents vary between 0 and 1. Other mathematical distributions may better describe the histograms of some of the elements in some of the sources, but for practical reasons we sought to use a single form for all element contents in all the sources and the fluvial samples. Assuming the c/(1-c) transformation of the recipient elemental profiles $\tilde{\mathbf{C}}$ is log-normally distributed (Massoudieh et al., 2013) with known variances with its elements independent of each other conditional to true concentrations (i.e. independent observed error), and also assuming that the standard deviation of the log-transformed elemental fractions is the same for all the elements (i.e. $\boldsymbol{\Delta} = diag(\delta_c)$), the first component of the likelihood function can be expressed as:

$$\pi(\tilde{\mathbf{C}}\,|\,\mathbf{Y},\mathbf{X}) \propto \frac{1}{\sigma_c^{I\cdot L}\prod_{i=1}^{I}\prod_{l=1}^{L}\tilde{c}_{il}(1-\tilde{c}_{il})}\, e^{-\sum_{l=1}^{L}\sum_{i=1}^{I}\frac{\left[\ln\left(\frac{\tilde{c}_{il}}{1-\tilde{c}_{il}}\right)-\ln\left(\frac{\sum y_{ij}\cdot x_j}{1-\sum y_{ij}\cdot x_j}\right)\right]^2}{2\sigma_c^2}} \qquad \textit{(Eq. 2-6)}$$

where σ_c is the standard deviation of the logarithm of elemental compositions and $\sum y_{ij}\cdot x_j$ is the "true" recipient elemental composition for element *j*. Making the same assumptions about the error structure of observed isotope ratios while assuming that the error associated with the observed isotope ratios are log-normally distributed, the second term in the likelihood function can be expressed as:

$$\pi\left(\tilde{\mathbf{\Gamma}} \mid \mathbf{Y}, \mathbf{X}, \mathbf{\Phi}\right) \propto \frac{1}{\sigma_{is}^{I' \cdot L} \prod_{i=1}^{I'} \prod_{l=1}^{L} \tilde{\gamma}_{il}} e^{-\sum_{l=1}^{L} \sum_{i=1}^{I'} \frac{\left[\ln\left(\tilde{\gamma}_{il}\right) - \ln\left(\sum \frac{x_j y_{ij} \varphi_{ij}}{x_j y_{ij}}\right)\right]^2}{2\sigma_{is}^2}} \qquad \textit{(Eq. 2-7)}$$

where σ_{is} is the standard deviation of the log transformed observed isotope ratios in the fluvial samples. Assuming that the isotope ratio of rare isotopes in the sources are log-normally distributed, the third term in the likelihood function can be expressed as:

$$\pi\left(\tilde{\mathbf{\Phi}} \mid \mathbf{\Phi}\right) \propto \frac{1}{\prod_{j=1}^{J} \prod_{i=1}^{I'} \tilde{\varphi}_{ij} \sigma_{\varphi,ij}} e^{-\sum_{j=1}^{n} \sum_{i=1}^{m} \frac{\left[\ln \tilde{\varphi}_{ij} - \ln \varphi_{ij}\right]^2}{2\sigma_{\varphi,ij}^2}} \qquad \textit{(Eq. 2-8)}$$

where $\sigma_{\varphi,ij}$ is the standard deviation of the isotope ratio on rare isotope i in source j that can be estimated from multiple observations of the isotope ratios in samples of each source:

$$\sigma_{\varphi,ij} = \overline{\left[\left(\ln \tilde{\varphi}_{ij} - \overline{\ln \tilde{\varphi}_{ij}}\right)^2\right]}^{1/2} \qquad \textit{(Eq. 2-9)}$$

In order to make sure all source elemental fractions $\tilde{y}_{ij}$ vary between zero and one, similar to the elemental composition of fluvial samples the transformation $\tilde{y}_{ij} / \left(1 - \tilde{y}_{ij}\right)$ of observed source elemental compositions are assumed to be log-normally distributed and therefore we can express the forth component of the likelihood function as:

$$\pi(\tilde{\mathbf{Y}} \mid \mathbf{Y}) \propto \frac{1}{\prod_{j=1}^{n} \prod_{i=1}^{m} \tilde{y}_{ij}\left(1 - \tilde{y}_{ij}\right) \sigma_{y,ij}} e^{-\sum_{j=1}^{n} \sum_{i=1}^{m} \frac{\left[\ln\left(\frac{\tilde{y}_{ij}}{1-\tilde{y}_{ij}}\right) - \ln\left(\frac{y_{ij}}{1-y_{ij}}\right)\right]^2}{2\sigma_{y,ij}^2}} \qquad \textit{(Eq. 2-10)}$$

where $\sigma_{y,ij}$ is the standard deviation of element *j* measured in source *i* calculated from multiple observations of elemental fractions for each source*:*

$$\sigma_{y,ij} = \left[\overline{\left(\ln \frac{\tilde{y}_{ij}}{1-\tilde{y}_{ij}} - \overline{\ln \frac{\tilde{y}_{ij}}{1-\tilde{y}_{ij}}} \right)^2} \right]^{1/2} \qquad \textit{(Eq. 2-11)}$$

The prior distribution for X was considered a Dirichlet distribution with parameters all equal to one and therefore satisfying constraint in Eq. (2-2):

$$\pi(\mathbf{X}) = \begin{cases} \pi(x_1,...x_n) = n! & \textit{for } 0 \le \sum_{j=1}^{n-1} x_j \le 1 \\ \pi(x_1,...x_n) = 0 & \textit{otherwise} \\ x_n = 1 - \sum_{j=1}^{n-1} x_j \end{cases} \qquad \textit{(Eq. 2-12)}$$

Since no additional information about the source elemental profiles y_{ij} and rare isotope ratios φ_{ij} other than the measured values are available, a uniform PDF between 0 and 1 was considered for the prior distribution of Y and a flat distribution (non-informative prior) is considered as the prior distribution of φ_{ij}. Assuming a flat distribution for φ_{ij} allows the term $\pi(\mathbf{\Phi})$ to be removed from Eq. (2-5). Substituting Eqs. (2-6 through 2-12) into Eq. (2-4), the following relationship for the posterior probability is obtained:

$$\pi(\mathbf{Y},\mathbf{X},\mathbf{\Lambda},\mathbf{\Delta}_{is} \mid \tilde{\mathbf{C}},\tilde{\mathbf{Y}},\tilde{\mathbf{\Gamma}},\tilde{\mathbf{\Phi}}) \propto \frac{1}{\sigma_c^{I\cdot L}\prod_{i=1}^{I}\prod_{l=1}^{L}\tilde{c}_{il}} e^{-\sum_{l=1}^{L}\sum_{i=1}^{I}\frac{\left[\ln(\tilde{c}_{il})-\ln\left(\sum y_{ij}\cdot x_j\right)\right]^2}{2\sigma_{c,i}^2}} \cdot \frac{1}{\sigma_{is}^{I'\cdot L}\prod_{i=1}^{I'}\prod_{l=1}^{L}\tilde{\gamma}_{il}} e^{-\sum_{l=1}^{L}\sum_{i=1}^{I'}\frac{\left[\ln(\tilde{\gamma}_{il})-\ln\left(\sum\frac{x_j y_{ij}\varphi_{ij}}{x_j y_{ij}}\right)\right]^2}{2\sigma_{is}^2}}$$
$$\cdot \frac{1}{\prod_{j=1}^{J}\prod_{i=1}^{I'}\tilde{\varphi}_{ij}\sigma_{\varphi,ij}} e^{-\sum_{j=1}^{n}\sum_{i=1}^{m}\frac{\left[\ln\tilde{\varphi}_{ij}-\ln\varphi_{ij}\right]^2}{2\sigma_{\varphi,ij}^2}} \cdot \frac{1}{\prod_{j=1}^{n}\prod_{i=1}^{m}\tilde{y}_{ij}\left(1-\tilde{y}_{ij}\right)\sigma_{y,ij}} \cdot e^{-\sum_{j=1}^{n}\sum_{i=1}^{m}\frac{\left[\ln\left(\frac{\tilde{y}_{ij}}{1-\tilde{y}_{ij}}\right)-\ln\left(\frac{y_{ij}}{1-y_{ij}}\right)\right]^2}{2\sigma_{y,ij}^2}} U(\mathbf{Y};0,1)\pi(\mathbf{X})$$

(Eq. 2-13)

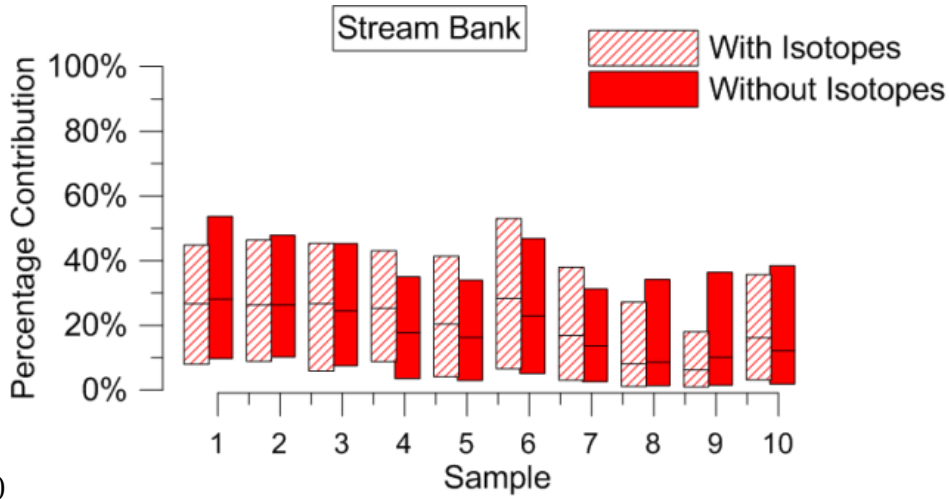
Eq. (2-12) can be used to calculate various moments of the posterior probability distributions of **X** and **Y** through integration. Due to the large number of dimensions, evaluating the integral in Eq. (2-12) using conventional methods is prohibitive. Therefore a Markov Chain

Monte Carlo (MCMC) approach (Gamerman and Hedibert, 2006) was used to generate random samples according to the posterior distribution of X and Y. Specifically in this research the Metropolis-Hasting Algorithm (Metropolis et al., 1953) was used to obtain a sequence of random numbers from the posterior probability distribution presented in Eq. (2-12). A program using the C++ programming language was used to perform the MCMC Bayesian inference. The posterior moment approach suggested by Geweke (1992) and Geweke and Tanizaki (2001) was used to evaluate the convergence of the MCMC algorithm. An adaptive perturbation factor was used starting from a large perturbation and reducing it until a minimum acceptance rate by all the chains are achieved (Massoudieh et al., 2013; Massoudieh and Kayhanian, 2013). For this research, the program generated a total of 5,000,000 samples with 8 chains for analyzing each of the fluvial samples and an average acceptance rate of 0.0976 was achieved.

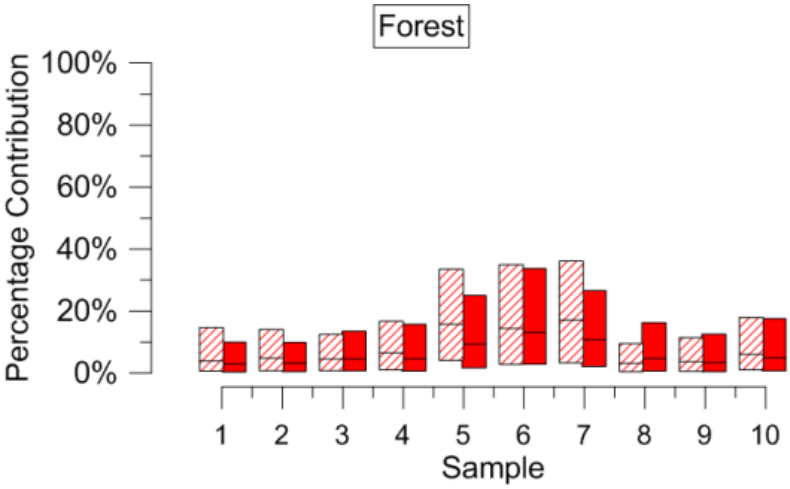
2.2 Results and Discussion

The ten fluvial samples collected were analyzed using the method outlined in section 2.1.2. Table 2-2 shows elemental profiles of all the potential sources considered for this study. Table 2-3 contains the elemental profiles of the ten fluvial samples. The standard deviations of elemental profiles $\sigma_{y,ij}$ and isotope concentrations $\sigma_{\varphi,ij}$ were calculated via Eqs. (2-8 and 2-10) for each source sample. The analyses were performed once with and once without ^{13}C and ^{15}N isotopes. Figure 2-2 shows the result of analyzing each sample individually with and without isotopes. The floating bars in this figure show the 95% credible intervals (CI) for the contribution of each source into each of the ten samples. Overall the contributions of the sources into the ten fluvial samples were consistent and (with the exception of events 5-7) no definitive conclusion about a significant temporal variation in the contribution among the events could be made (i.e. 95% CI.s have significant overlaps). Also, possibly because there were only two isotopes versus 27 elements, it

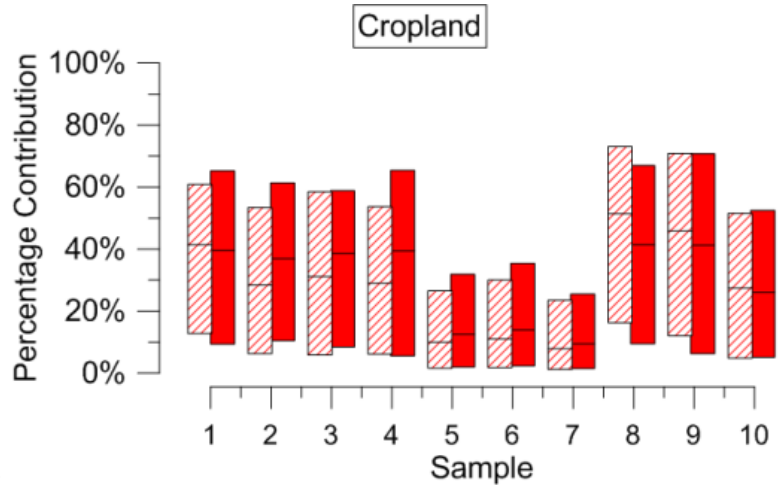
can be seen that in general the inclusion of isotopes did not substantially narrow down the credible intervals. The results are to a large extent similar between the cases where isotopes were used and the cases where isotopes were not used. This indicates that in this study the isotopes do not contain a significant amount of additional information regarding the sources with respect to the metallic elements; though they do provide an additional confirmation to the results obtained from the metallic elements. It should be noted that the conclusion about the usefulness of isotopes versus metallic elements should not necessarily be generalized to other datasets obtained under different settings. Although arguably the assumed conservation of C and N isotopes in the fluvial sediments as they are transported from the source can be questioned, the observed variability of isotope ratios among the source samples of each source may to some degree capture the alteration of the isotope ratios during sediment mobilization and transport. This conjecture needs to be tested through further research. Cropland, pasture and stream-bank erosion were found to have the largest contribution into the fluvial sediments with an expected value of contribution around 10%-70%, 10%-70% and 5%-50% respectively (Figure 2-2).



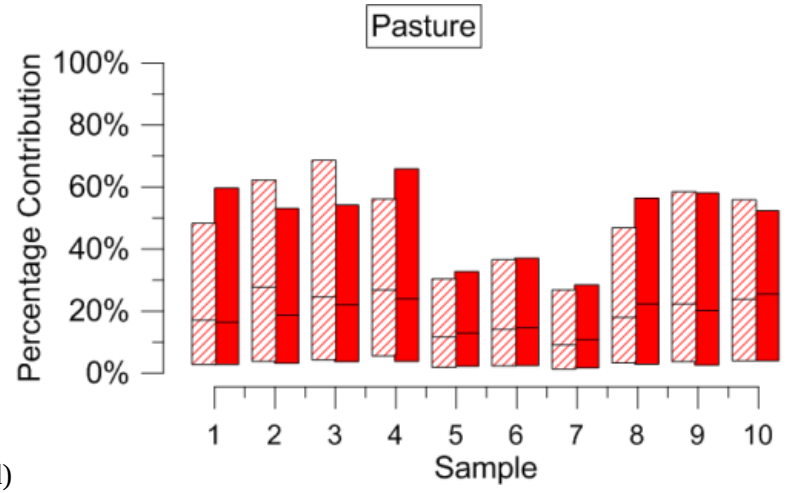
a)


Forest
Percentage Contribution
100%
80%
60%
40%
20%
0%
1
2
3
4
5
6
7
8
9
10
Sample

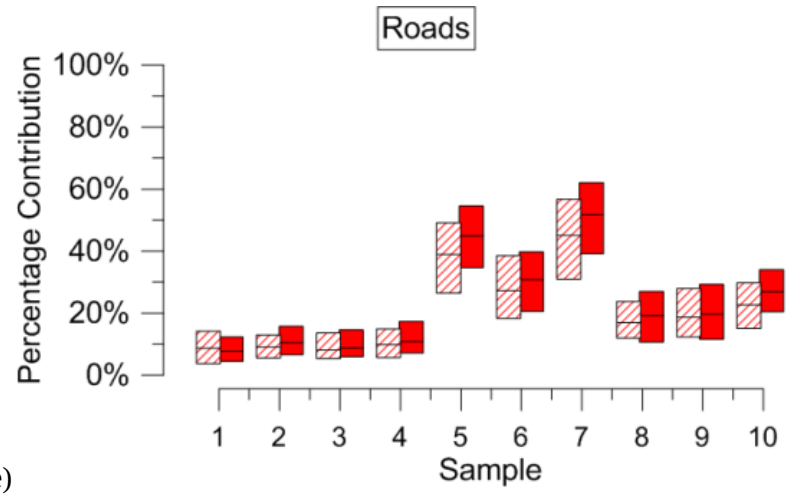
b)


Cropland
Percentage Contribution
100%
80%
60%
40%
20%
0%
1
2
3
4
5
6
7
8
9
10
Sample

c)


Pasture
Percentage Contribution
100%
80%
60%
40%
20%
0%
1
2
3
4
5
6
7
8
9
10
Sample

d)



e)

Figure 2-3. CI source contribution into the ten individual fluvial samples for each source (a-e).

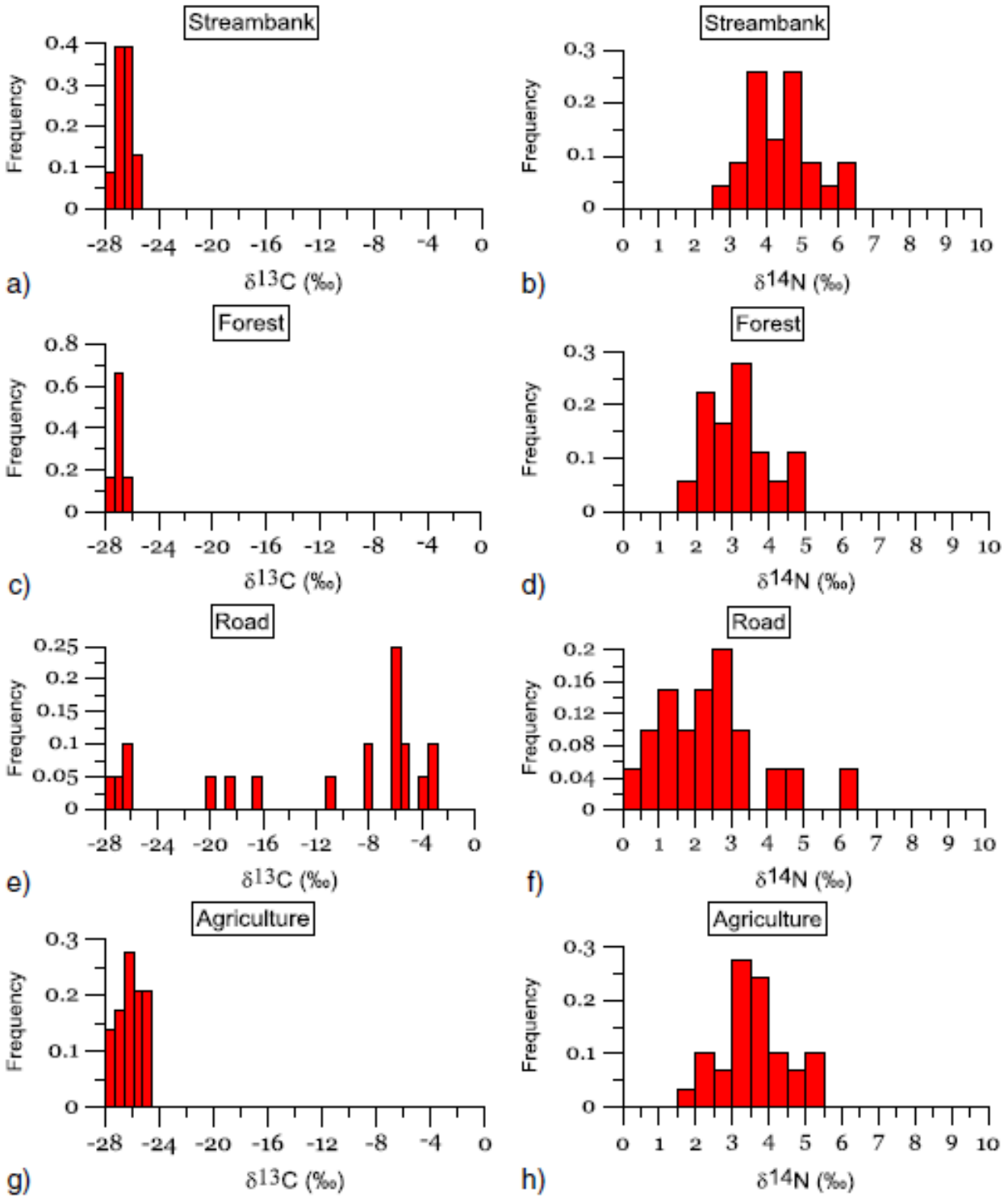


Figure 2-4. Distribution of δ13C and δ15N in the potential sediment sources.

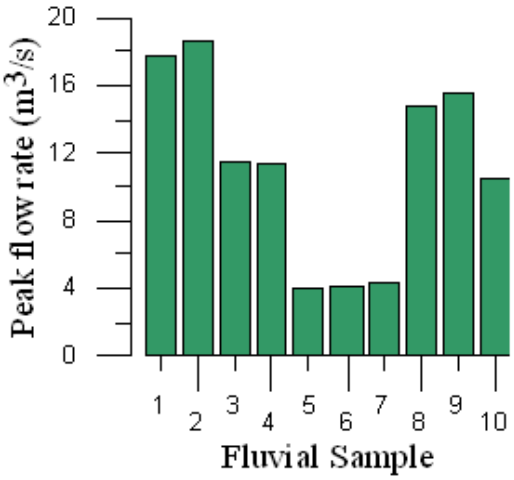
Figure 2-3 shows the distribution of isotope ratios $\delta^{13}C$ and $\delta^{15}N$ in all the potential sources. The figures show a results when pasture and cropland were combined into agriculture for reasons discussed in the following section. Results from split pasture and cropland are not shown because they were very similar to agriculture. As it can be seen there is no significant difference between the distributions of the two isotopes among all of the sources except for ^{13}C in some of the road samples. The similarity between the isotope ratios of all the sources can explain the fact that including isotopes did not result in narrowing the credible intervals. In the sediment fingerprint analysis by Sloto et al., (2012) ^{13}C was also not found to be a significant tracer distinguishing the sediment sources. Aside from unpaved roads, the Laurel Hill Creek sediment sources' $\delta^{13}C$ values indicate mainly biogenic sources of carbon (i.e. between -33‰ to -24‰). In the case of roads the range of $\delta^{13}C$ indicates the presence of both biogenic and non-biogenic (mantle or mineral carbon) (i.e. $\delta^{13}C > -10$). The difference between the sources in terms of $\delta^{15}N$ is less statistically significant although a similar pattern of $\delta^{13}C$ can be observed in the sense that roads seem to have a different isotope ratio range. This means that the two isotopes can potentially contain information about the road source. However the high variability of ^{13}C among the road sources can weaken its ability to reduce the uncertainty in the estimation of sources. Also, the fact that isotope ^{15}N is the only constituent for which the observed value was outside the predicted 95% C.I. (Figure 2-5) can be attributed to the differential transformation or solid-water exchange of heavy and light N isotopes.

The 95% CI range for the contribution of forest sources vary between close to zero to up to 20% except for three events (5, 6, and 7) where the forest contribution was found to have an upper limit of roughly 40%. For these three events the contribution of cropland and pasture was roughly half of that for other events. Similarly, roads were found to contribute between 2-20% in

all events except for the three events 5, 6, and 7 where the contribution was significantly higher (between 20 and 60%).

Figure 2-4a shows the peak flow rates during fluvial sampling and Figure 2-4b shows the flow hydrograph measured at the same location by a USGS gauge at station 03079600. It appears that the highest contribution of sediments from roads occurs during low stream flow and during the months of September and November. For these samples (5, 6, and 7) the contribution of cropland and pasture was inferred to be smaller than the other events. It appears that the contribution of cropland and pasture is significantly correlated to the peak hydrograph indicating that the overland flow as a result of less significant rain events results in significantly smaller mobilization of sediments from these sources. The relative contribution of roads and forest therefore increases during low intensity rains indicating that their contribution is less affected by the event intensity compared to pasture and cropland.

a)



b)

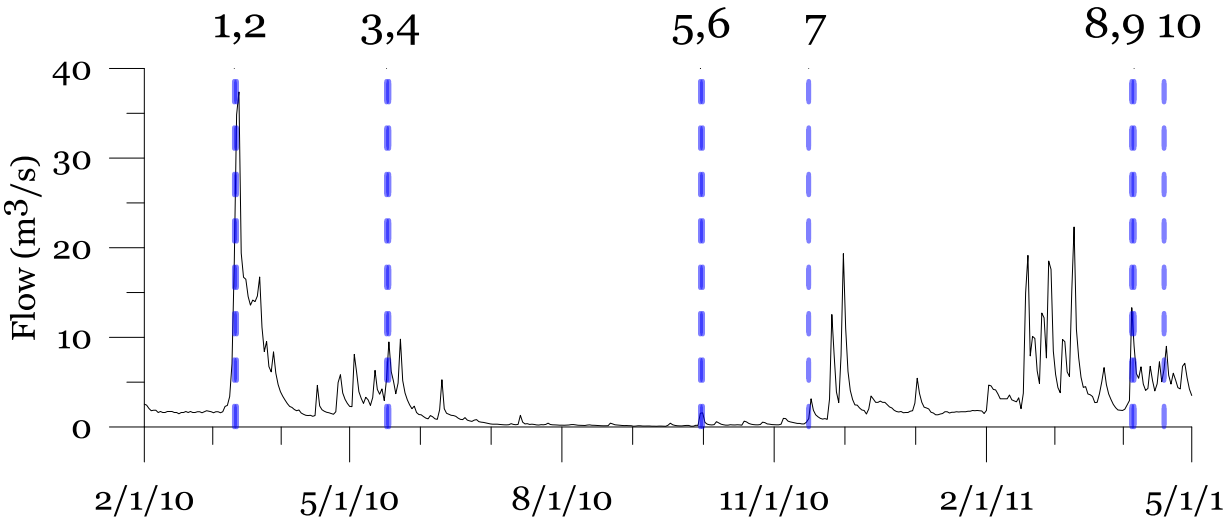


Figure 2-5. a) Peak flow during sampling and b) stream hydrograph during the sampling period in Laurel Hill Creek. The hydrograph is obtained from USGS gauge # 03079600 near Bakersville, PA.

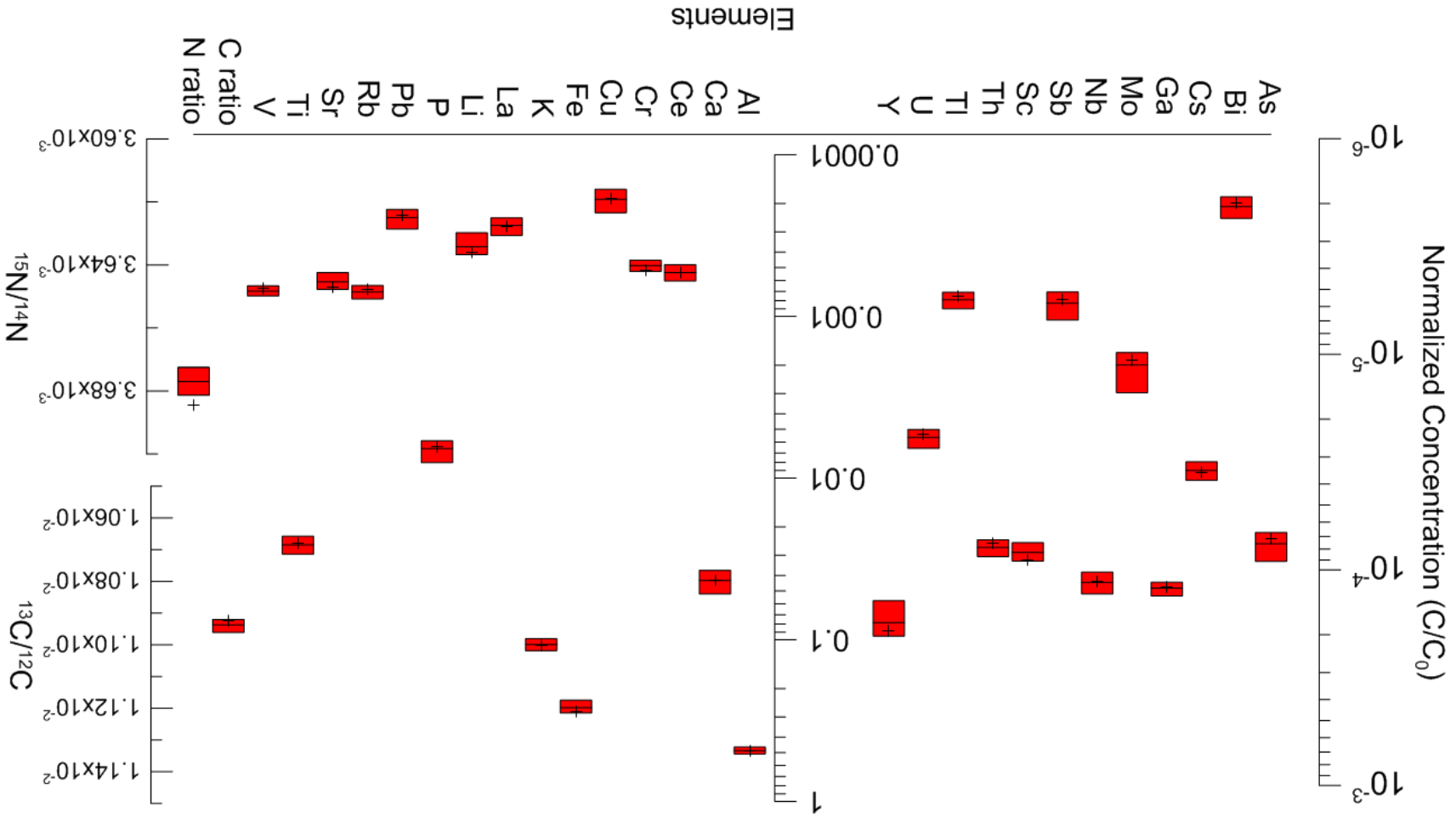


Figure 2-6. Modeled 95% CI and observed elemental profile and isotope ratios for fluvial sample no. 1- March 12, 2010. The plus signs (+) show the observed normalized elemental content.

Figure 2-5 shows the 95% C.I. for the predicted and observed elemental profiles of the elements and the isotopes for fluvial sample number 1. Overall the mixing model captured the observed elemental profiles well. Isotope ^{15}N is the only tracer for which the observed value was outside the predicted 95% C.I. This can be attributed to the differential transformation or differential solid-water exchange of heavy and light N isotopes. The modeled vs. observed elemental profiles and isotope ratios for the rest of the samples are similar to sample no. 1, with ^{15}N underestimated for most cases. These results are not presented for the sake of brevity.

In order to evaluate the ability of the method to discriminate between different sources, the scatter plots representing the probability space of contribution of sources as well as the correlation matrix of the posterior source contributions are shown in Figure 2-6. A high correlation between the posterior probabilities of two sources would indicate collinearity and therefore a lack of ability by the method to discriminate between them. Particularly a large negative correlation between two sources shows that one can replace the other and still result in the same observed elemental profile or the same overall agreement between measured and modeled profiles.

The highest correlation is between pasture and cropland. Moderate negative correlations between other sources are expected due to the fact that the sum of contributions should be one. However the significant negative correlation between pasture and cropland is probably due to the similarity between their elemental fingerprints and the lack of ability of the CMB method to discriminate between them. This is expected as the lands are sometimes alternatively used as pasture and cropland. Similar results were obtained by Sloto et al., (2012). Therefore the fingerprinting analysis was repeated with cropland and pasture data combined under the term agriculture.

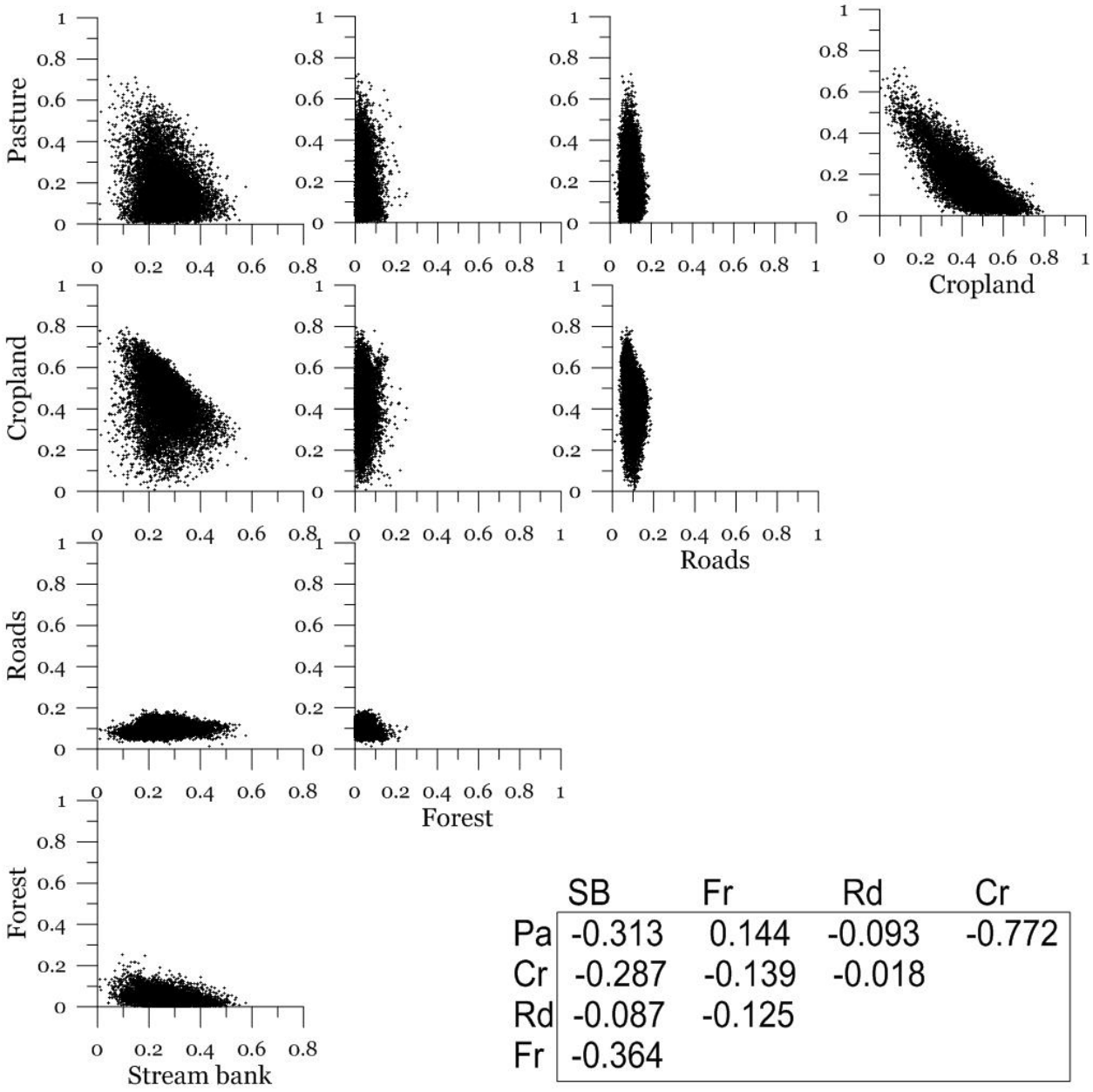


Figure 2-7. Scatter plots representing the probability space of joint contribution of sources and the correlation matrix for the posterior source contribution for sample no. 1 on March 12, 2010.

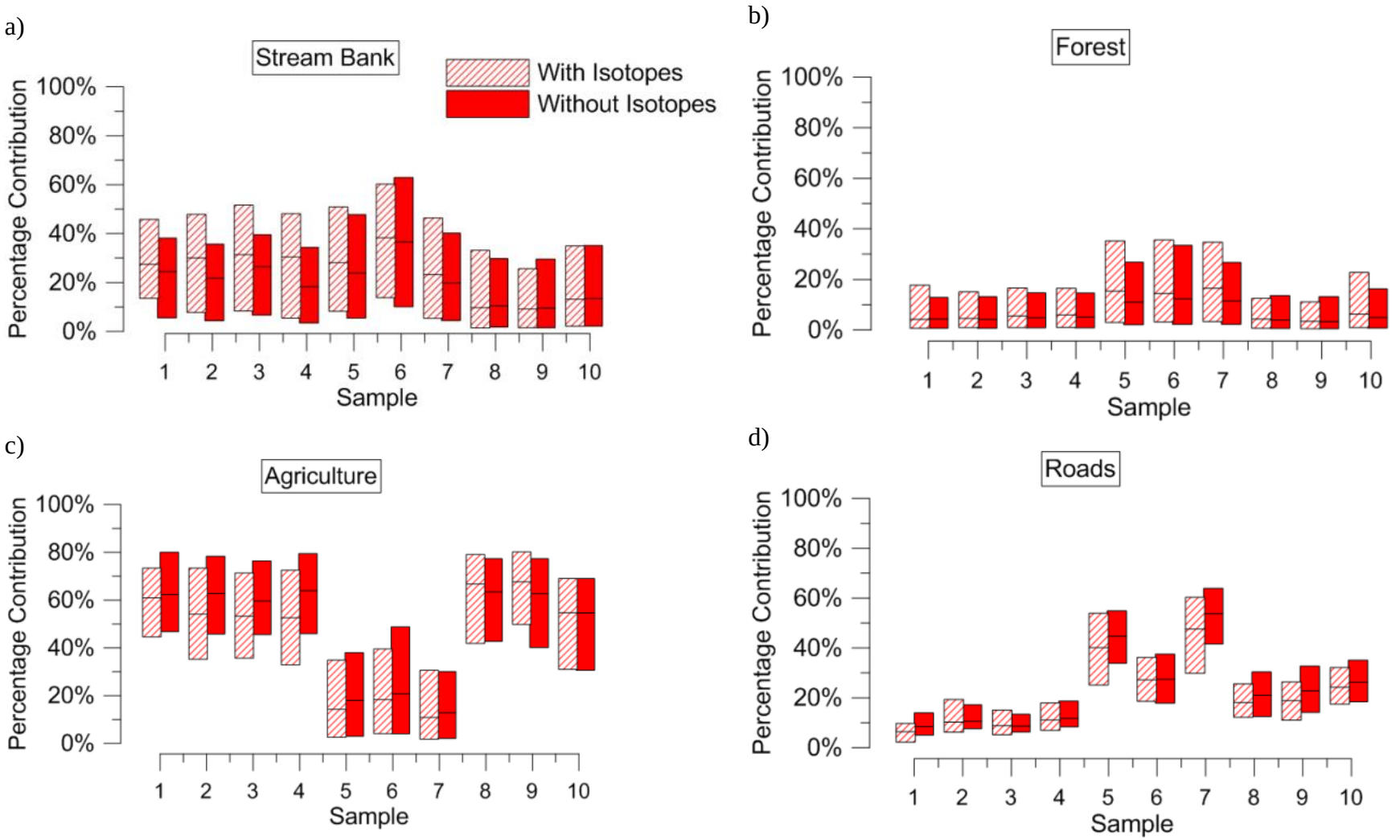


Figure 2-8. 95% CI source Contribution with Pasture and Cropland combined as Agriculture.

Figure 2-7 shows the 95% brackets representing the contribution of each source when cropland and pasture are considered a single source type. It seems that the contribution of other sources is not affected significantly and the contribution of the new Agriculture source is almost the same as the sums of the contributions of cropland and pasture in the original analysis. Also, the influence of smaller rain events on agricultural contribution is preserved. This confirms the conclusion that the method is not able to discriminate between cropland and pasture but can estimate their collective contribution. Interestingly, there is no major correlation between road and other sources pointing to its distinct signature. After roads, forest has the smallest correlation with other sources.

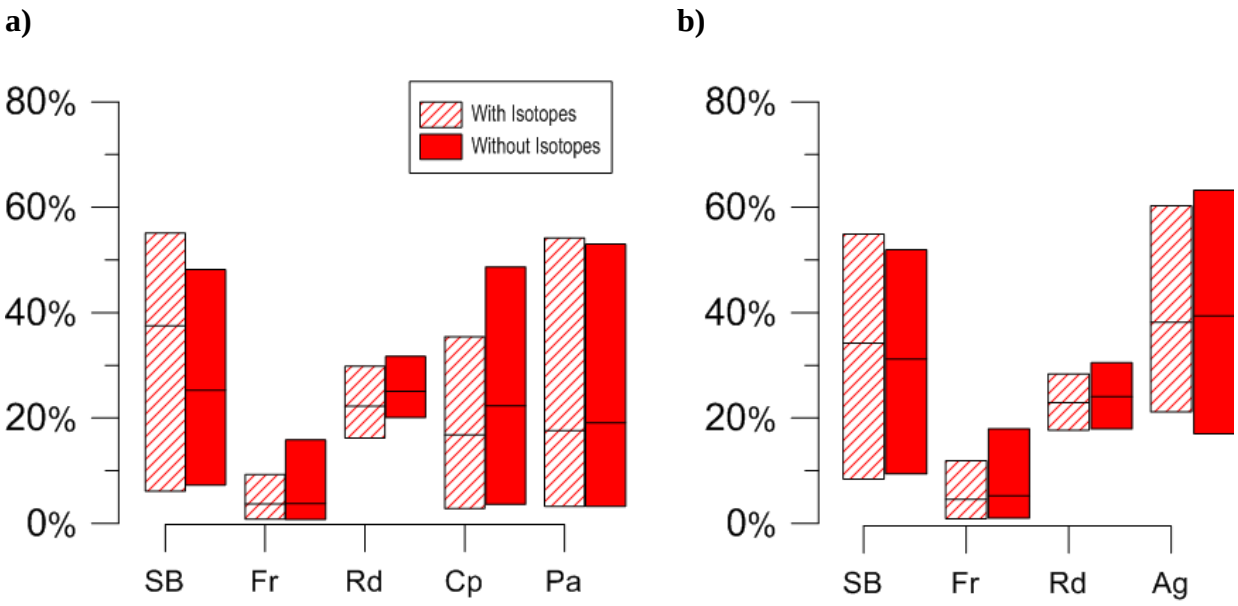


Figure 2-9. Source contribution into one lumped fluvial sample as a result of analysis with and without isotope ratios for a) five sources: Stream Bank, Forest, Roads, Cropland, and Pasture and (b) when Cropland and Pasture were combined as Agriculture

As it was mentioned before, the contribution of sources toward the ten fluvial samples was consistent to a large degree except for the three events with less intense precipitation as can be noticed from the large overlaps between the 95% CI brackets (Figure 2-2 and 2-7). One hypothesis to describe this consistency is that the sources' contribution for all of the samples are in fact statistically similar and the resulting variation in source contributions for different samples is a result of the random heterogeneities in the elemental compositions of the fluvial samples and the contributing sources. To test this hypothesis, the method was applied to all of the fluvial samples collectively, thus assuming that the source contribution came from the same probability distributions and the differences in their elemental composition were solely due to heterogeneity and observation error (i.e. the likelihood function was evaluated using all the fluvial samples collectively). As can be seen from Eq. (2-6), the method was capable of considering multiple recipient samples, thus the variations between elemental compositions of samples are attributed to their heterogeneities rather than to the difference between their contributing sources. Figure 2-8 shows the 95% CIs for the contribution of the sources with pasture and cropland split (a) as well as combined as agriculture (b). When all ten fluvial samples were considered together the CIs were much wider, particularly for stream bank and agriculture (or cropland and pasture), indicating the inability of the method to infer the contributing sources with high confidence. The overall pattern was to a large degree consistent with the case when the method was applied to individual fluvial samples; and similar results were obtained whether with or without isotopes. Figure 2-9 shows the modeled vs. observed elemental composition of elements for all ten fluvial samples when pasture and cropland were combined as agriculture and isotopes included (the other cases had similar results – not shown). The fact that the observed elemental compositions in most cases are outside the inferred 95% C.I. brackets indicates that the unmixing model in this case is not

capable of capturing the observed elemental profiles well. In other words no single combination of the sources considering an added uncertainty due to random heterogeneities can reproduce the elemental compositions of all of the fluvial samples, which leads to the conclusion that there is a statistically significant temporal variation in the contribution of sources.

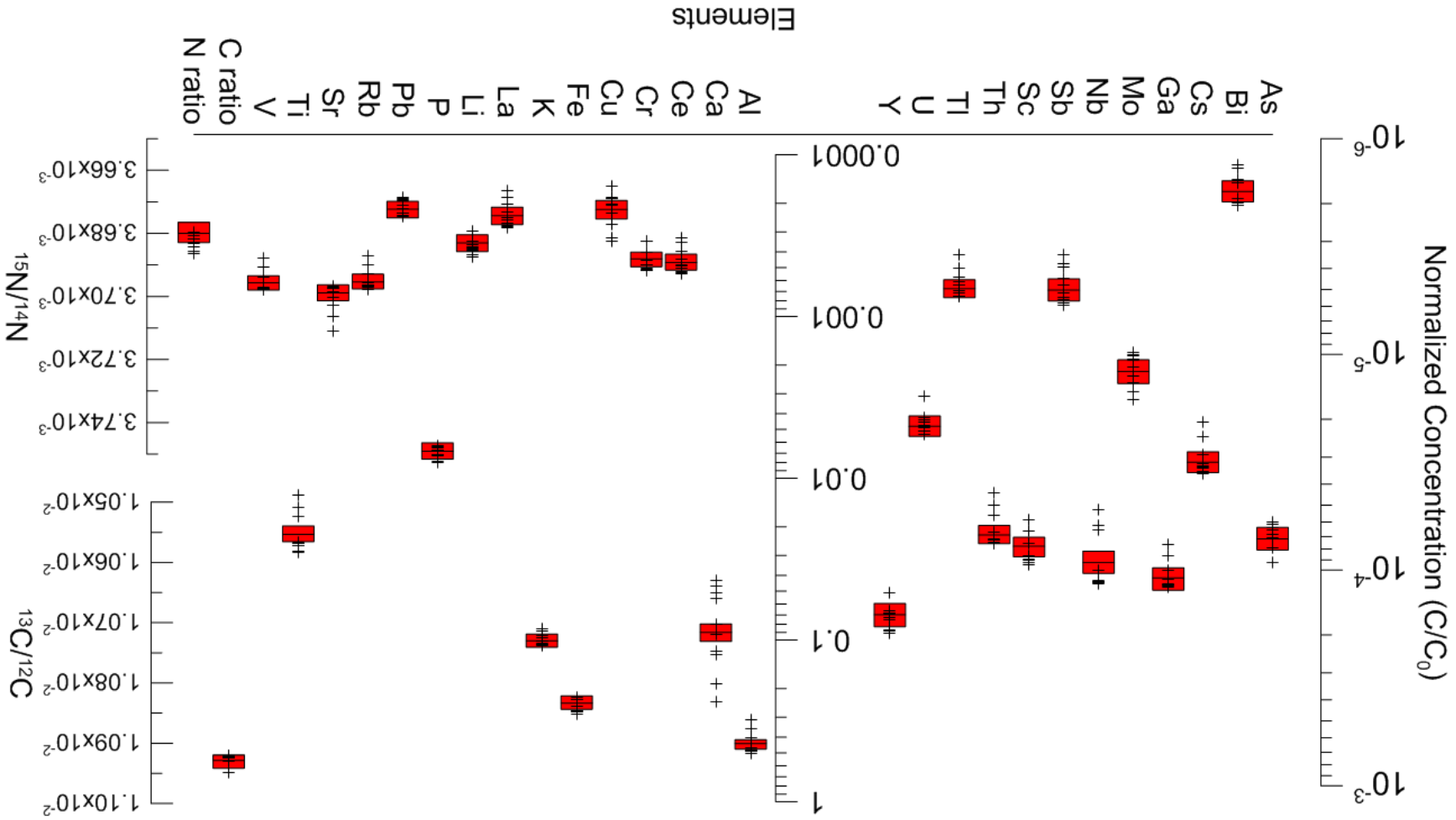


Figure 2-10. Modeled 95% CI and observed elemental profile and isotope ratios for lumped samples. The plus signs (+) show the observed normalized elemental content.

2.3 Summary and Conclusions

A Bayesian Chemical Mass Balance approach was used to assess the contribution of five potential sources (stream bank, roads, forest, cropland and pasture) in fluvial samples in Laurel Hill Creek, PA. Ten fluvial samples and 91 samples representing the sources were available. Multiple samples of each source were used to assess the variability of elemental profiles among each source. In the analysis, the elemental profiles of the sources and fluvial samples and the isotope ratios of ^{15}N and ^{13}C were used. The analyses were performed once using only the elemental profiles and once using a combination of elemental profiles and isotope ratios; it was found that in most cases inclusion of the isotope ratios as fingerprinting agents does not significantly reduce the uncertainty. This can be due to the fact that the isotope signatures are not generally different among the sources. It should be noted that lack of consideration for the preferential delivery of the finer particles and the possible organic matter alteration due to biotransformation and its effects on the tracers during the transport of sediments can add some uncertainty to the results. Cropland and pasture were found to be the most significant sources of sediments in all of the samples, followed by stream bank. In three of the fluvial samples collected during the fall, roads were found to have a significant contribution while the contribution of cropland and pasture was found to be less than during the other events. It was found that for these three events the peak runoff was substantially lower than the rest of the events. It seems that the reason for a lower contribution of cropland and pasture during these three events can be attributed to the inability of low intensity precipitation events to mobilize sediments from these two sources whereas the other sources are less affected by the intensity of the event. When cropland and pasture were treated as separate sources, the uncertainty of their contributions was large enough such that no definitive conclusion about their significance could be made (i.e. the lower limit of their

contribution CIs was small). However, the upper bound of the credible intervals of these two sources was determined to be significant. Correlation analysis on the posterior distributions of the source contributions show a large correlation between the contributions of these sources which indicated the inability of the method to confidently discriminate between them. This may be due to the rotation of field use between cropland and pasture. So analysis was also performed where cropland and pasture were combined and collectively referred to as agriculture. The result of the analysis with the combined sources had higher confidence.

Also, it was shown that the differences between the elemental compositions of the fluvial samples represented variability due to differences in actual source contribution rather than simply heterogeneity by comparing model predictions on individual samples vs all ten collectively. The model predictions had much higher uncertainty and were unable to capture the observations well when all dates were included, suggesting a temporal variability in source contributions. Despite this variability, agriculture was a significant contributor of sediment to the Laurel Hill Creek while only representing a quarter of the land in the watershed. These findings provide an uncertainty estimate that was missing from the previous deterministic multivariate statistical analysis by Sloto et al., (2012) which found the same order of contribution (agriculture (53%) followed by stream banks (30%), unpaved roads (17%), and forest (<1%)). This quantitative information can be used to justify and enforce best management practices to protect the water qualities of these streams, such as restoration of wetlands, creation of riparian buffers, and appropriate use of fertilizers regarding quantity and timing with precipitation. Figure 2-10 shows the unfortunate abundance of impaired streams in the Laurel Hill watershed; the impaired designation is given for streams with high sedimentation and/or nutrient loads like nitrogen and low dissolved oxygen concentrations.

Storm runoff from agricultural fertilizer is a prime example of non-point source pollution. The following chapters focus on the other type, point-source nutrient pollution.

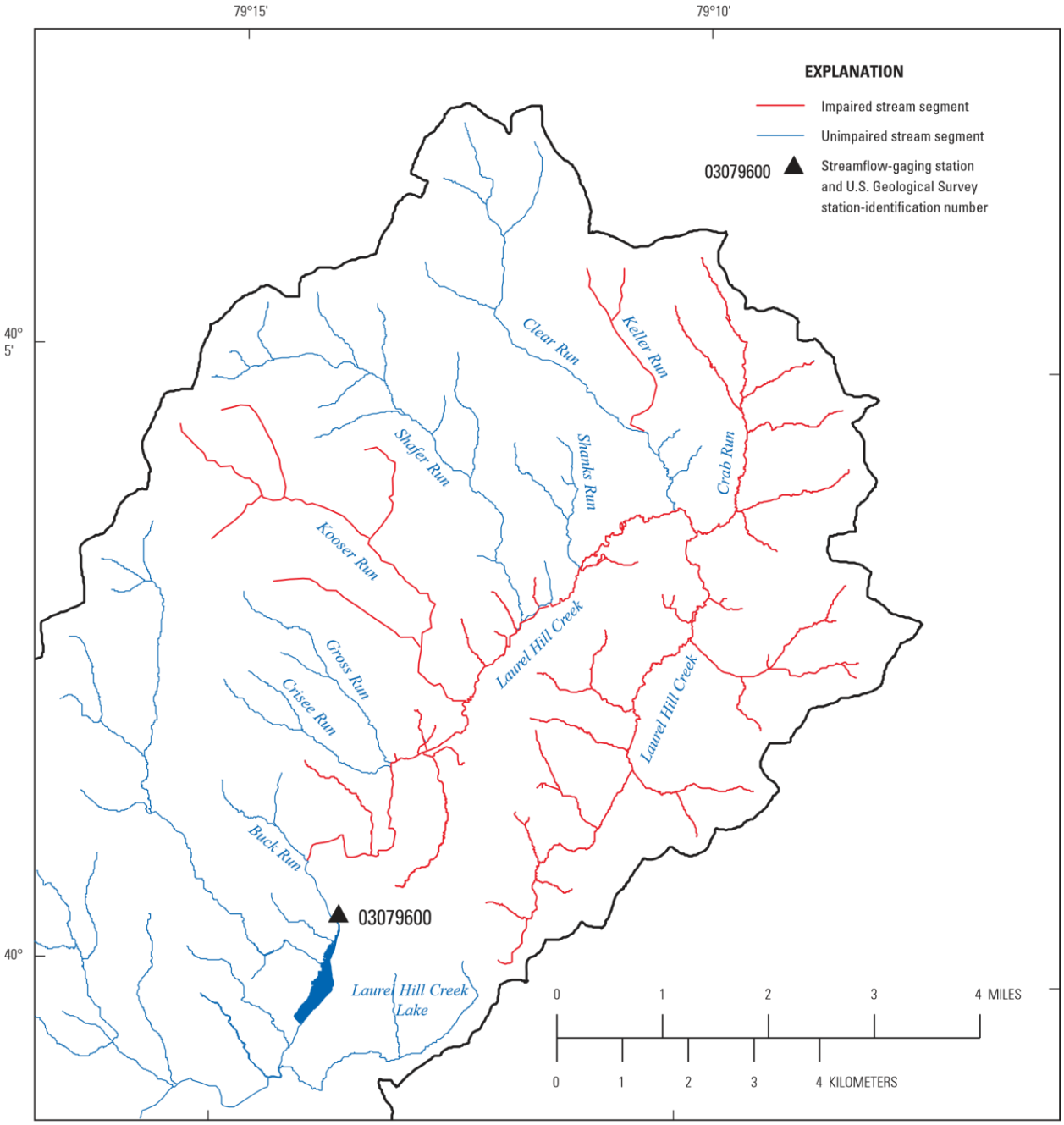


Figure 2-11. Impaired stream segments in the Laurel Hill Creek watershed, Somerset County, Pennsylvania (Sloto et al., 2013).

Table 2-2. The elemental profiles and isotope ratios of the 10 fluvial samples used in the study. The rows shown using italics letters indicate the elements that were excluded from the analysis after bracketing.

	FS-1	FS-2	FS-3	FS-4	FS-5	FS-6	FS-7	FS-8	FS-9	FS-10
Ag (mg/kg)	*0.106*	*0.103*	*0.138*	*0.120*	*0.093*	*0.12*	*0.138*	*0.103*	*0.086*	*0.129*
Al (mg/kg)	72900	73100	75900	75100	56800	68200	47000	73900	69200	69000
As (mg/kg)	10.8	10.4	11.3	11.1	12.7	15.5	9.6	8.8	9.3	9.2
Ba (mg/kg)	*517*	*513*	*564*	*538*	*651*	*662*	*578*	*483*	*459*	*493*
Be (mg/kg)	*3.0*	*3.8*	*3.5*	*3.3*	*3.9*	*3.9*	*3*	*3.4*	*3.1*	*3.5*
Bi (mg/kg)	0.30	0.29	0.32	0.32	0.22	0.27	0.2	0.23	0.22	0.24
Ca (mg/kg)	6680	8160	7590	9060	31600	21000	38900	12400	13800	19300
Cd (mg/kg)	*2.6*	*2.5*	*2.8*	*2.6*	*3.8*	*3.1*	*3.8*	*1.5*	*1.5*	*2.1*
Ce (mg/kg)	80.9	81.0	85.7	79.4	56.3	66.7	49.2	71.6	68.3	65.8
Co (mg/kg)	33.8	33.4	36.9	29.6	49.3	41.9	48.9	22.3	20.5	25.4
Cr (mg/kg)	77.5	76.7	81.2	81.4	55.4	68.3	51.4	70.4	68.1	67.1
Cs (mg/kg)	5.3	5.1	5.5	5.6	3.9	4.9	3.1	4.9	4.7	4.8
Cu (mg/kg)	28.0	27.8	32.4	36.3	25.2	31.3	40.5	29.7	46.1	51.5
Fe (mg/kg)	42100	41700	42900	43500	44200	48200	38800	34400	33300	33900
Ga (mg/kg)	18.0	17.8	18.5	18.5	13.9	16.9	11.4	17.1	16.6	16.5
K (mg/kg)	16200	16200	15900	15200	14100	15900	12800	15500	15100	15300
La (mg/kg)	41.9	42.8	43.0	41.6	29.7	34.2	25.1	37	34.7	33.7
Li (mg/kg)	60.0	54.4	60.3	58.9	55.6	63.8	44.8	61.5	61.4	58
Mg (mg/kg)	*4650*	*4850*	*4910*	*5240*	*7660*	*7390*	*9510*	*5830*	*5920*	*6910*
Mo (mg/kg)	1.6	1.6	1.8	1.6	2.6	2.5	1.9	2	1.4	1.5
Nb (mg/kg)	17	17	18	18	10	11	7.9	17	16	15
Ni (mg/kg)	56.1	53.3	57.4	54.0	70.4	73.4	70.1	42.1	40.6	49.3
P (mg/kg)	957	1020	1120	1250	1170	1330	1070	934	919	1010
Pb (mg/kg)	*35.9*	*35.2*	*37.9*	*33.8*	*29.6*	*38.5*	*30.9*	*27.8*	*26.4*	*29*
Rb (mg/kg)	102	98.9	103	103	77	92.1	63.2	93.9	91	92
Sb (mg/kg)	*0.84*	*0.88*	*0.93*	*0.86*	*0.61*	*0.8*	*0.67*	*0.58*	*0.49*	*0.77*
Sc (mg/kg)	13.5	13.6	13.4	13.6	10.6	12.7	8.8	13.9	13.1	13
Sr (mg/kg)	98.8	102	102	105	161	144	186	104	109	127
Th (mg/kg)	11.3	11.5	11.7	11.3	8.1	9.44	6.63	10.7	10.3	10
Ti (mg/kg)	3790	3950	4110	3990	2430	2900	1910	4250	4020	3740
Tl (mg/kg)	0.81	0.79	0.84	0.80	0.64	0.74	0.52	0.7	0.68	0.69
U (mg/kg)	3.53	3.47	3.57	3.38	3.18	3.48	2.34	3.24	3.07	2.99
V (mg/kg)	101	103	107	107	80.3	96.9	65.8	97.2	96.1	92.6
Y (mg/kg)	28.8	29.9	29.7	26.9	24.9	26.7	19.1	24.5	22.9	23.7
Zn (mg/kg)	*346*	*333*	*402*	*349*	*403*	*338*	*396*	*264*	*252*	*355*
%C	4.45	4.53	5.04	4.94	5.84	5.53	5.53	3.28	3.25	3.84
%N	0.431	0.455	0.480	0.533	0.604	0.578	0.774	0.409	0.446	0.555
$\delta^{13}C$	-28.03	-27.8	-28.18	-27.85	-27.30	-28.25	-25.73	-27.91	-27.78	-27.99
$\delta^{15}N$	5.83	5.076	5.37	6.47	5.51	6.21	5.14	4.54	4.76	4.66

Table 2-3. Mean concentration of elements in source samples. The rows shown using italics letters indicate the elements that were excluded from the analysis after bracketing.

	Stream bank	Forest	Roads	Cropland	Pasture
No. of samples	24	18	20	20	9
Ag (mg/kg)	*0.09*	*0.06*	*0.09*	*0.11*	*0.10*
Al (mg/kg)	67772.92	58013.89	37720.00	71367.50	64977.78
As (mg/kg)	12.97	11.09	5.40	12.77	11.20
Ba (mg/kg)	*413.42*	*331.72*	*309.08*	*465.20*	*442.67*
Be (mg/kg)	*2.49*	*1.31*	*1.04*	*1.43*	*1.28*
Bi (mg/kg)	0.34	0.39	0.12	0.28	0.26
Ca (mg/kg)	2865.71	1155.50	103200.90	3027.00	2954.44
Cd (mg/kg)	*1.02*	*0.39*	*0.27*	*0.33*	*0.39*
Ce (mg/kg)	72.79	50.51	52.75	80.12	71.34
Co (mg/kg)	*20.45*	*10.17*	*6.83*	*13.03*	*12.38*
Cr (mg/kg)	68.07	55.88	41.11	72.00	62.93
Cs (mg/kg)	5.04	4.34	2.74	5.15	4.53
Cu (mg/kg)	32.38	22.53	17.58	36.25	33.76
Fe (mg/kg)	37520.83	26580.56	17766.50	36052.50	31422.22
Ga (mg/kg)	17.09	14.91	8.49	18.16	16.12
K (mg/kg)	14582.08	11932.50	13959.00	15805.00	14311.11
La (mg/kg)	36.76	25.48	26.64	40.55	35.56
Li (mg/kg)	50.08	36.41	23.96	45.79	40.80
Mg (mg/kg)	*3829.58*	*3085.00*	*3936.50*	*4454.25*	*4066.67*
Mo (mg/kg)	2.38	3.51	0.89	1.84	1.78
Nb (mg/kg)	11.00	11.12	7.26	20.00	17.56
Ni (mg/kg)	31.54	16.41	11.46	22.46	20.91
P (mg/kg)	946.31	1090.08	406.15	981.28	1142.33
Pb (mg/kg)	*57.56*	*55.24*	*15.30*	*27.48*	*27.46*
Rb* (mg/kg)	102.16	89.24	64.68	104.92	95.87
Sb (mg/kg)	*1.00*	*1.07*	*0.53*	*0.82*	*0.75*
Sc (mg/kg)	11.29	8.64	6.43	12.11	10.69
Sr (mg/kg)	73.85	73.76	138.85	78.21	71.69
Th (mg/kg)	10.60	8.34	8.37	12.11	10.96
Ti (mg/kg)	2623.75	2639.83	1939.75	4516.25	4024.44
Tl (mg/kg)	0.91	0.82	0.38	0.80	0.72
U (mg/kg)	3.54	2.56	2.16	3.69	3.41
V (mg/kg)	95.41	84.30	47.21	106.52	93.99
Y (mg/kg)	23.33	12.17	17.07	18.28	17.04
Zn (mg/kg)	*168.94*	*91.03*	*54.08*	*111.87*	*117.00*
C%	6.87	12.48	4.71	2.80	4.21
N%	0.53	0.86	0.13	0.31	0.42
$\delta^{13}C$	-26.59	-26.99	-11.92	-25.90	-26.62
$\delta^{15}N$	4.44	3.10	2.41	3.76	3.29

Chapter 3 - Dual Substrate Limitation Modeling and Implications for Mainstream Deammonification

3.1 Introduction

3.1.1 Activated sludge Modeling

Modern waste water treatment facilities harness the natural response to excess nitrogen by hosting and fostering the growth of nitrifying and denitrifying organisms. These organisms are called activated sludge and form the usually brown, fluffy biomass in water treatment reactors. In order to maintain effluent concentrations below regulation limits, extensive physical, chemical, and biological models have been proposed to describe these systems, called Activated Sludge Models (ASM). The basic state variables for these models are the biomass groups, X_i, and the inorganic substrates they consume or produce, S_i.

$$\frac{dX_i}{dt} = X_i \mu_i(S) \quad \text{and} \quad \frac{dS}{dt} = \frac{X_i}{Y_i} \mu_i(S) \qquad \textit{(Eq. 3-1)}$$

where μ_i is the doubling time for a given biomass group and Y_i is a stoichiometric parameter relating the yield of biomass produced for a given unit of substrate. The overwhelming majority of these models describe the growth of organisms with respect to substrate concentration with the Monod equation (Monod, 1949; Henze et al., 2000)

$$\mu_i(S) = \mu_{i,\max} \frac{S}{K_S + S} \qquad \textit{(Eq. 3-2)}$$

where $\mu_{i,max}$ is the unlimited growth rate and K_s is the substrate concentration at which the modeled biomass growth rate is half of the maximum growth rate as illustrated in Figure 3-1. The name of his term, K_s, is debated within the ASM community: half-saturation index, constant, coefficient or

affinity constant (Arnaldos et al., 2015). I try to use the generic ‘half-saturation parameter’. This model is empirical, but analogous to Michaelis-Menten enzyme kinetics.

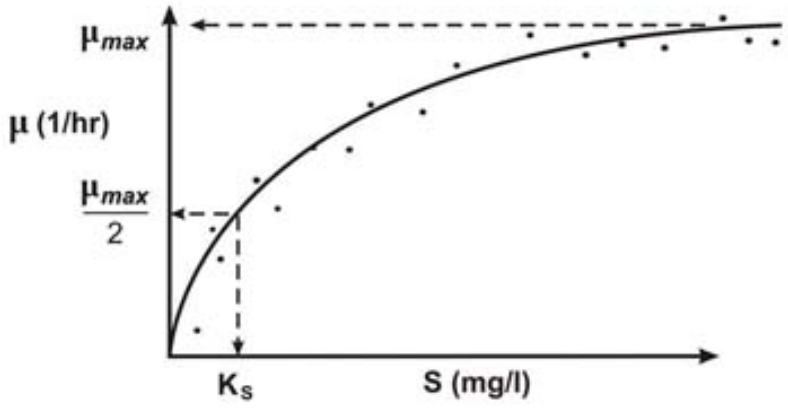


Figure 3-1. Monod kinetics. A typical relationship between growth rate and substrate concentration.

Traditional or conventional nitrogen removal is carried out in two steps. In the first step NH_3 is oxidized by aerobic and autotrophic Ammonia-Oxidizing Biomass (AOB) to NO_2, which is then oxidized to NO_3 by aerobic and autotrophic Nitrite Oxidizing Biomass (NOB). In the second step Ordinary Heterotrophic Organisms (OHO) or denitrifiers convert NO_3 to NO_2 and again to N_2 in an anoxic environment with a supplemental carbon source such as methanol. This process can be improved by skipping the NO_3 formation (called Nitrite Shunt) or by fostering Anaerobic Ammonia-Oxisizing Biomass (AnAOB) to achieve the Anammox process (Mainstream Deammonification). The details of implementing the Anammox process are further discussed in section 4.1.1.

3.1.2 Advanced Nitrogen Removal

Nitrogen shortcut technologies like mainstream deammonification offer significant cost savings over traditional biological nitrogen removal. Nitrite (NO_2) plays a decisive role in the success of each of these technologies because it lies at a critical juncture of three processes described in Figure 3-2 (Al-Omari et al., 2015):

(1) Full nitrification by nitrite-oxidizing bacteria (NOB) *(red)*

(2) Denitrification by ordinary heterotrophic organisms (OHO) *(orange)*

(3) Deammonification by anaerobic ammonia-oxidizing bacteria (AnAOB) *(black)*

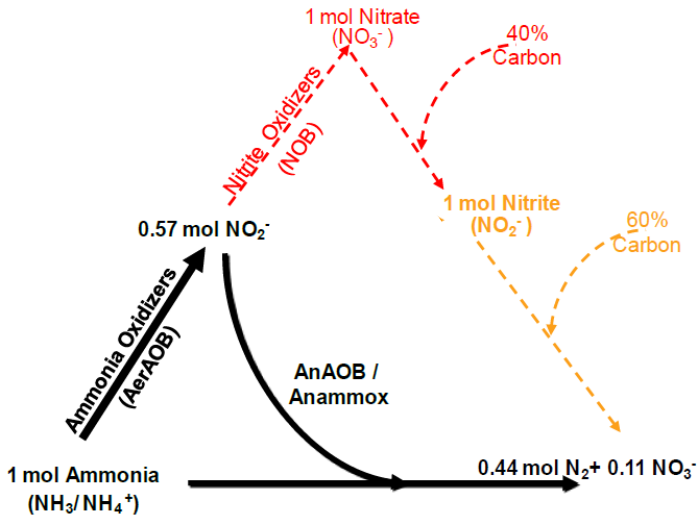


Figure 3-2. The potential pathways of nitrogen removal in waste water treatment. (Modified from a presentation by Haydée De Clippeleir).

The competition between these players depends on their environmental fitness traits, i.e. their maximum growth rates (μ) and affinity or half-saturation indices (K) for different substrates (Maynard et al., 2015). For the sake of simplicity, the term 'substrate' shall refer to both electron donors and electron acceptors. The ability to model this competition therefore depends on accurate and reliable methods to predict how limitation from two or more substrates will affect the overall

rate of a process. The particular environment of mainstream deammonification often involves more than one limiting substrate: NH_3 and NO_2 are near limiting conditions for AnAOB, and NO_2 and dissolved oxygen (DO) are near limiting conditions for NOB (Neethling and De Clippeleir, 2015). Monod kinetics are well-accepted for modelling reaction rates as a function of a single non-inhibitory substrate (Kovarova-Kovar and Egli, 1998; Henze et al., 2000; Gonzo et al., 2014), but several models exist for treating two or more substrates simultaneously (MeGee et al., 1972; Mankad and Bungay, 1988; Zinn et al., 2004; Okpokwasili and Nweke, 2005).

Currently, the *multiplicative* model (*Eq. 3-3*), which simply multiplies single Monod terms, is most commonly used in Activated Sludge Models as it is continuous, smooth, and easy to handle in numerical simulations particularly when the numerical methods rely on computation of Jacobians of mass balance equations (MeGee et al., 1972; Makinia, 2010). This expression, however, may underestimate growth rates when more than one substrate concentration is limiting. Underestimation of growth rates particularly worsens when the number of limiting substrates increases, due to the nature of Monod type equations: seven limiting components at 10 times the half saturation concentration for each substrate results in a growth rate approximately 50% of maximum. This effect is visualized in Figure 3-3 along with other alternative models discussed in later sections.

Bungay (1994) put it bluntly: "It makes no sense that the growth rate becomes vanishingly small as the number of potentially limiting nutrients is increased when each is present in reasonable concentrations." As a result, the maximum growth rate is often artificially inflated to compensate for this under-estimation of the growth rate which in turn can result in unrealistic estimates of growth when none or only one of the substrates are actually limiting. Another frequently proposed dual substrate model is the *minimum* formula, which is not smooth (i.e. not 2nd order

differentiable), but theoretically relates to a rate-limiting reactant in chemical kinetics and Liebig's law of the minimum (Von Liebig, 1840; Bader, 1978).

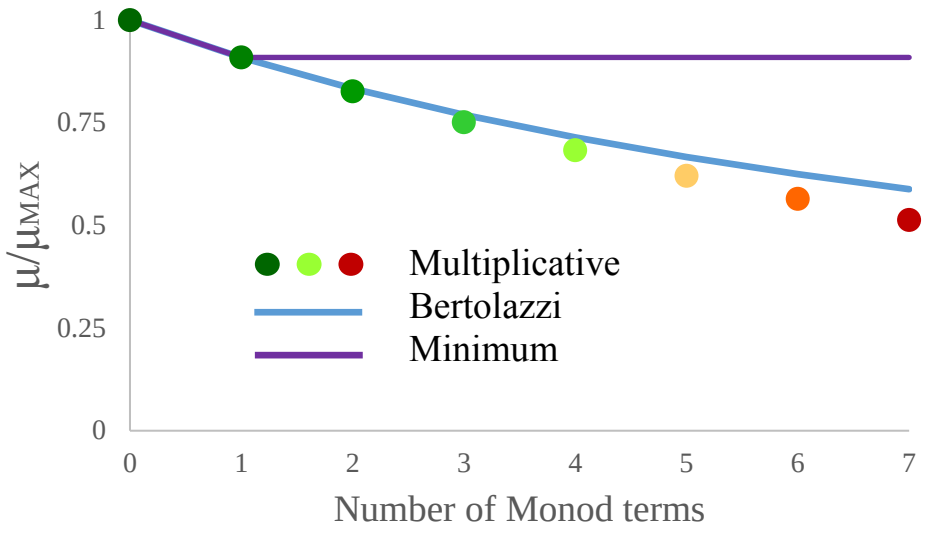


Figure 3-3. Conceptual comparison of multiple substrate limiting models where each substrate is at 10 times that half saturation coefficient concentration.

Previous studies on this topic include batch and chemostat experiments of pure cultures as well as computer simulations. Chemostat experiments are able to directly measure yield indices by measuring the effluent biomass and substrate concentrations at different dilution rates. However, it is difficult to maintain low substrate conditions, and can lead to inconclusive results regarding the appropriate dual substrate model (Panikov, 1979). Each steady state experiment results in a single observational point in the dual substrate space which makes the amount of work needed to generate adequate amount of data tedious. Batch tests can explore this space more thoroughly and effectively by resulting in time-series of data from each single experiment, but must rely upon standard, constant yield and stoichiometry values.

Chemostat and batch test studies found that the *multiplicative* model did not fit experimental data for *Pseudomonas aeruginosa;* but instead of exploring other dual limitation models, they used the *multiplicative* model with alternatives to the Monod function: Tessier and Contois models (Beyenal et al., 2003; Tanase et al., 2011). Maynard et al. (2015) successfully used the *minimum* model to predict competition outcomes between two pure culture strains of *Saccharomyces cerevisiae* in replicated chemostats under various substrate limiting conditions. However, the multiplicative model would have predicted the same outcomes; the experiment was not designed to compare the fidelity of model predictions. Several simulation studies have predicted large differences between model predictions of different models. Odencrantz (1992) found a six-fold deviation between models when diffusion was incorporated. The deviation would likely increase when biofilm heterogeneity is incorporated (Gonzo et al., 2014). These simulations highlight the importance of experimentally verifying the best dual substrate model.

Dual and multiple substrate biokinetic models have been categorized as either interactive (two or more limiting substrates have a synergistic effect) or non-interactive (only one limitation effects growth at a time) (Chapra, 1997). In this study, three models were compared: one fully non-interactive dual substrate limitation model - *minimum* (Bader, 1978)

$$\mu = \mu_{max} \cdot MIN\left[\left(\frac{S_1}{S_1+K_1}\right),\left(\frac{S_2}{S_2+K_2}\right)\right] \qquad \textit{(Eq. 3-3)}$$

and two interactive models - *multiplicative* (MeGee et al., 1972)

$$\mu = \mu_{max}\left(\frac{S_1}{S_1+K_1}\right)\left(\frac{S_2}{S_2+K_2}\right) \qquad \textit{(Eq. 3-4)}$$

and *Bertolazzi* (2005)

$$\mu = \mu_{max}\frac{1}{\frac{K_1}{S_1}+\frac{K_2}{S_2}+1} \qquad \textit{(Eq. 3-5)}$$

where S_x are the substrate concentrations and K_x are the associated half-saturation indices. Only two limiting substrates were considered for simplicity. These models were selected for their ability to satisfy the following criteria:

1) When both substrates are in excess, the formula reduces to the maximum growth rate.
2) When one substrate is in excess and only one is limiting, the formula should reduce to a single Monod term.
3) The formula uses established half-saturation indices and does not over-parameterize the system – does not require excessive, non-measurable parameters in order to achieve a good model fit.

The third formula evaluated in this study, *Bertolazzi*, is an attempt to bridge the gap between *multiplicative* and *minimum*. It was developed based on the criteria 1-3 listed above to generalize the Monod equation for multiple substrates and can be theoretically derived from enzyme kinetics (Bertolazzi, 2005). Dual limitation models that do not fit the criteria above by including inhibition (Liu et al., 1992; Haas et al., 1994; Ben-Youssef and Vazquez-Rodriguez, 2011), by predicting reduced growth rate with increasing substrate (Mankad and Bungay, 1998) or by requiring new parameters (Legovic and Cruzado, 1997; van den Berg, 1998; Ramkrishna et al., 2012) were not considered.

Double Monod surfaces for growth rate as a function of DO and NO_2 concentrations, based on each formula are plotted in Figure 3-5 for NOB as an example biomass. At non-limiting concentrations (10 times the half-saturation index, DO = 2.7 mg O_2/L and NO_2= 3.3 mg NO_2-N/L) all three models predict similar values for μ / μ_{max}: 88%, 88%, and 92% by *multiplicative*, *Bertolazzi*, and *minimum*, respectively. However, at dual substrate limiting conditions (half of the

half-saturation indices, DO = 0.14 mg O_2/L and NO_2 = 0.17 mg N O_2-N/L) the models substantially differ in terms of the predicted μ / μ_{max} being 11%, 20%, and 33%, for *multiplicative*, *Bertolazzi*, and *minimum*, respectively. From these values and the conceptual double Monod surfaces in Figure 3-4, it is clear that *Bertolazzi* is an intermediary between the other two models.

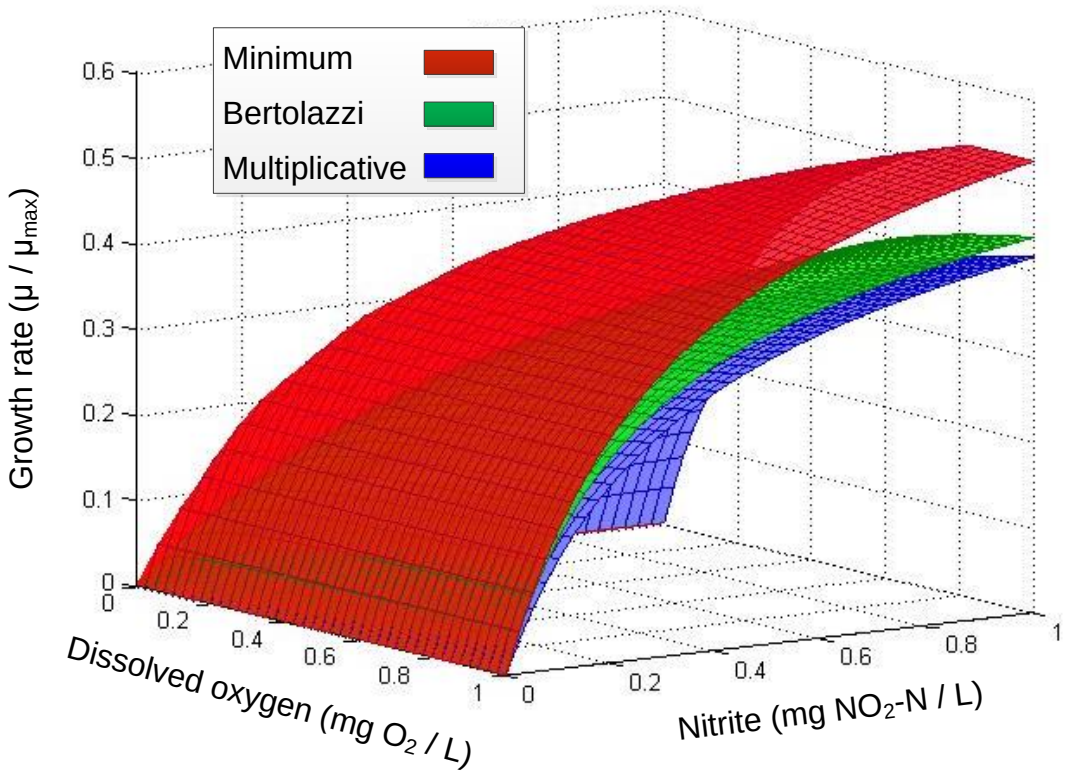


Figure 3-4. Conceptual double Monod surfaces for nitrite oxidizing bacteria with oxygen and nitrite as substrates. The 3 models used the following indices and maximum growth rates: $\mu_{max,NOB}$ = 0.7 [1/day], $K_{NO2,\ NOB}$ = 0.33 [mg NO_2-N / L], and $K_{O,\ NOB}$ = 0.27 [mg O_2 / L] in this schematic.

Modern wastewater treatment facilities can operate under dual constraint of lower nutrient effluent limits while minimizing energy consumption. One method to achieve sustainable nitrogen removal is via mainstream deammonification, which reduces oxygen demand by 63% and organic carbon demand by up to 90% compared to conventional full nitrification-denitrification (Siegrist

et al., 1998). A major challenge in applying this process is preventing full nitrification. Several pilot studies have studied different strategies for preventing nitrite oxidation by aerobic organisms, with promising results (Lotti et al., 2014; Lemaire et al., 2014; Han et al., 2016, and De Clippeleir et al', 2014). The initial motivation for this research was to understand why different and seemingly opposing strategies have had similar results in NOB out-selection through the lens of multiple limitation models. This was achieved by estimating the likelihood of each of the three alternative models accurately predicting NOB and AnAOB reaction rates during dual substrate limiting conditions using experimental data and rigorous probabilistic inference. Experimental observations were obtained using sludge from the DC Water's Blue Plains Advanced Wastewater Treatment Facility (AWWTF) under single, double, and non-limiting substrate conditions. These observations were used to estimate maximum growth rate and substrate half-saturation parameters. Models were then evaluated based on their ability to predict substrate utilization rates under various substrate conditions using parameters estimated from single substrate limitation.

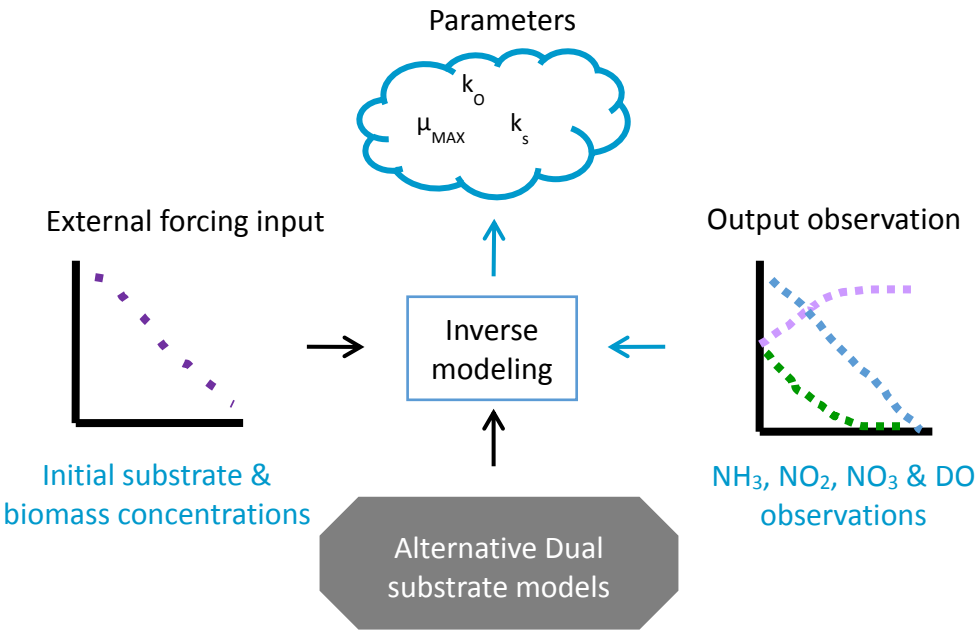


Figure 3-5. Schematic for inverse modeling of dual substrate limitation experiments.

3.2 Materials and Methods

3.2.1 Batch experiment I: NOB

The methodology for observing aerobic NO_2 consumption combined aspects of respirometry tests and activity tests (Smith et al., 1998). The mixed liquor sample was taken from the mainstream deammonification pilot at the Blue Plains AWWTF. The sample was centrifuged at 630 *g* for 15 min to remove as much residual substrate as possible. The sample was then resuspended with dechlorinated tap water. The sample was divided into three batches for three different substrate (NO_2) spike regimes: low, medium, and high (1.5, 3, and 6 mg NO_2-N / L respectively) (Figure 3-5). Total suspended solids and volatile suspended solids (TSS, VSS) samples were taken from each batch prior to its test. Tests were brief enough (less than 90 minutes) to permit ignoring the change in biomass concentration. $NaHCO_3$ and H_2SO_4 were added to each batch to maintain an alkalinity greater than 300 mg $CaCO_3$/L and a pH between 7 and 7.5. Mixed liquor was agitated and aerated above 7 mg O_2/L using air stones. An identical spike of KNO_2 (potassium nitrite) was administered to 15 0.35 L bottles containing magnetic stirrers before they were filled with mixed liquor, immediately sealed, and placed on a multi-positional stir plate (Figure 3-6). One more bottle was filled with mixed liquor without a NO_2 spike in order to estimate OHO oxygen consumption. This bottle along with the first filled bottle were fitted with YSI 5010 BOD probes (Xylem Inc, New York, USA). These bottles were essentially respirometry tests where the probe measured and logged the DO concentration at one second intervals using HyperTerminal version 7.0 software (observations were averaged to the time step of the inverse modeling algorithm: 1.5 minutes). The time at which each bottle was filled was noted so that the delay created by filling bottles sequentially could be accounted for in the analysis. In this way 16

reactors were created with the same initial conditions. One by one, bottles were sacrificed (unplugged and sampled) to measure substrate concentrations. Nitrogen, phosphorus, and COD concentrations were measured using Hach vials [HACH GmbH]. TSS and VSS were measured according to standard methods (APHA, 2005).

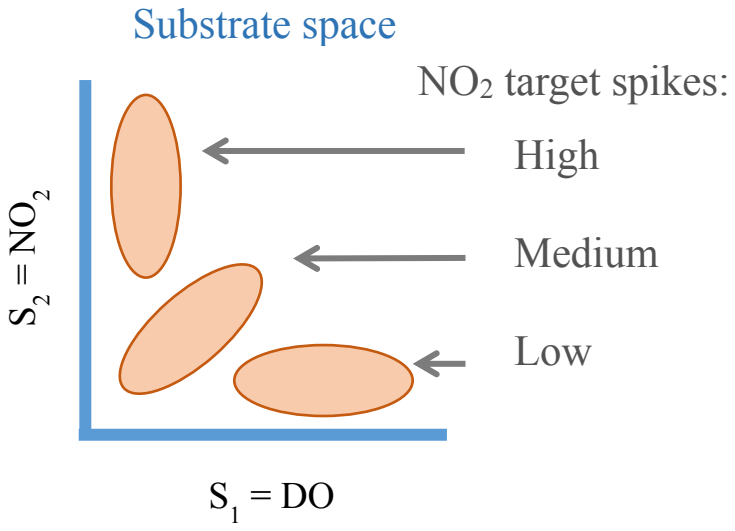


Figure 3-6 Experimental design of Batch test I for NOB.

Figure 3-7 Experimental setup for Batch test I for NOB.

3.2.2 Batch experiment II: AnAOB

The second experimental design consisted of multiple batch tests of AnAOB activity with varying initial conditions. These initial conditions were chosen to amplify the deviations between dual substrate models. The mixed liquor sample was taken from the side stream deammonification pilot at the Blue Plains AWWTF (Zhang et al., 2016). The side stream is a source of AnAOB seed for many mainstream deammonification systems and is therefore a reasonable sample source for this experiment. In order to isolate the mostly granular AnAOB fraction from the flocculent sludge, the sample was sieved at 212 microns. It was then centrifuged at 630 *g* for 15 min and resuspended with dechlorinated tap water. Despite centrifuging, some denitrifying OHO activity was still observed. The sample was divided into five batches for five different substrate regimes: both NH_3

and NO_2 in excess, limiting NO_2, limiting NH_3, double limitation, and no NH_3. The last batch test allowed the modeling software to separate OHO activity from AnAOB activity. TSS and VSS samples were taken at the start and end of each batch test and found not to increase/decrease more than 4%. Nitrogen gas was continuously diffused into the reactor to maintain a DO below 0.01 mg O_2 / L. Alkalinity and pH were maintained and substrates were measured as in the NOB Batch experiment.

3.2.3 Inverse modeling

The BIOEST program (Sharifi et al., 2014) was used to estimate the biokinetic parameters from the experimental results based on biokinetic reactions for NOB and aerobic heterotrophs (NOB batch experiment)

$$\frac{dDO}{dt} = -\frac{1.14 - Y_{NOB}}{Y_{NOB}} X_{NOB} \mu_{NOB,MAX} f(NO_2, DO) - \frac{1 - Y_{OHO,Aer}}{Y_{OHO,Aer}} X_{OHO} \mu_{OHO,MAX} \frac{DO}{DO + k_{O,OHO}}$$

(Eq. 3-6)

$$\frac{dNO_2}{dt} = -\frac{1}{Y_{NOB}} X_{NOB} \mu_{NOB,MAX} f(NO_2, DO) \qquad \textit{(Eq.3-7)}$$

$$\frac{dNO_3}{dt} = +\frac{1}{Y_{NOB}} X_{NOB} \mu_{NOB,MAX} f(NO_2, DO) \qquad \textit{(Eq.3- 8)}$$

and AnAOB and anoxic heterotrophs (AnAOB batch experiment)

$$\frac{dNH_3}{dt} = -\frac{1}{Y_{AnAOB}} X_{AnAOB} \mu_{AnAOB,MAX} f(NH_3, NO_2) \qquad \textit{(Eq.3- 9)}$$

$$\frac{dNO_2}{dt} = -\left(1.52 + \frac{1}{Y_{AnAOB}}\right) X_{AnAOB} \mu_{AnAOB,MAX} f(NH_3, NO_2) - \frac{1}{Y_{OHO,An}} X_{OHO} \mu_{OHO,MAX} \frac{NO_2}{NO_2 + k_{NO2,OHO}}$$

(Eq. 3-10)

using the stoichiometric values (Y_{NOB} = 0.09, $Y_{OHO,Aer}$ = 0.67, and $Y_{OHO,An}$ = 0.54) given by Al-Omari (2015), (Y_{AnAOB} = 0.159) from Dapena-Mora et al. (2004) and the three different rate models, *f(S_1,S_2) (Eq. 3-3 to 3-5).* As the biomass concentration of a specific organism group is not easily estimated, the reported growth rates are formulated as

$$R_{MAX} = \frac{X\mu_{MAX}}{VSS} \qquad \textit{(Eq.3-11)}$$

with units of *[mg biomass COD / mg VSS / day]*. As a consequence, values for R_{max} are lower than expected μ_{MAX} by as much as an order of magnitude.

Parameter ranges given in Table 3-1 based on previous experiments and literature review determined the *prior* distributions (shaded regions in Figure 3-7) (Dapena-Mora et al., 2014; Al-Omari et al., 2015). Using Bayesian inverse modeling a selected number of parameters in the model were estimated. The Bayesian method provided joint probability distribution of model parameters that can better be used to compare the fidelity of different models. First, a deterministic parameter estimation was performed using a genetic algorithm; the deterministic values were then used as initial estimates for a stochastic parameter estimation using MCMC simulations in order to reduce the burn-in period. The MCMC samples formed the *posterior* joint probability distribution of parameters.

Table 3-1. Prior estimate range and median parameter values derived from MCMC simulations for each model as well as from single limitation only for both NOB and AnAOB experiments. The model fit comparison is also included. A better model has lower DIC (Deviance Information Criterion), higher Bayes factor and lower forward fit error.

	Prior Range	**Single Limitation**	***multiplicative***	***Bertolazzi***	***minimum***
$\mathbf{R_{max,\ NOB}}$ [mg COD/mg VSS/day]	0.015-0.16	0.040	0.044	0.044	0.042
$\mathbf{K_{O,\ NOB}}$ [mg O_2/L]	0.1-1	0.27	0.27	0.29	0.27
$\mathbf{K_{NO2,NOB}}$ [mg N/L]	0.1-1	0.33	0.33	0.34	0.34
$\mathbf{R_{max,\ OHO}}$ [mg COD/mg VSS/day]	0.609-1.22	0.689	0.678	0.679	0.676
$\mathbf{K_{O,\ OHO}}$ [mg N/L]	0.01-0.3	0.05	0.05	0.05	0.05
Forward fit error	-	-	35.5	34.2	30.1
DIC	-	-	55.2	54.9	55.1
Bayes Factor	-	-	0.333	0.333	0.333
$\mathbf{R_{max,\ AnAOB}}$ [mg COD/mg VSS/day]	0.0026-0.01	0.0061	0.0056	0.0056	0.0057
$\mathbf{K_{NH3,AnAOB}}$ [mg N/L]	0.1-10	0.77	0.5	0.55	0.92
$\mathbf{K_{NO2,AnAOB}}$ [mg N/L]	0.05-1.5	0.33	0.14	0.17	0.26
$\mathbf{R_{max,\ OHO}}$ [mg COD/mg VSS/day]	0.0033-0.033	0.0034	0.0032	0.0032	0.0034
$\mathbf{K_{NO2,\ OHO}}$ [mg N/L]	0.1-0.8	0.59	0.34	0.4	0.48
Forward fit error	-	-	196.5	194.5	194.3
DIC	-	-	6.85	5.57	3.73
Bayes Factor	-	-	0.3333	0.3328	0.3338

In Bayesian inference the posterior distributions of parameters given some observed data are computed based on the likelihood of observing the measured data and the prior distribution of parameters based on Bayes' Theorem:

$$p(\boldsymbol{\theta}\,|\,\mathbf{y}) = p(\mathbf{y}\,|\,\boldsymbol{\theta})\,p(\boldsymbol{\theta}) \qquad \textit{(Eq. 3-12)}$$

where $\boldsymbol{\theta}$ represents the vector of model parameters $\mathbf{y}$ is the vector of all observed data $p(\boldsymbol{\theta}\,|\,\mathbf{y})$ is referred to as posterior distribution which is the probability distribution of the parameters after considering the observed data, $p(\mathbf{y}\,|\,\boldsymbol{\theta})$ is the likelihood of observing the measured data given some parameter values that is calculated based on an assumption of error structure using the presumed model, and $p(\boldsymbol{\theta})$ represent the prior information about the parameters that can represent range of values for given parameters known based on previous studies. For more details about the Bayesian inference applied to biological wastewater treatment model see (Sharifi et al., 2014).

In this study a normal and additive error structure was assumed. Also the prior distribution of the parameters was considered lognormal with their mean and standard deviation extracted from the literature. A global sensitivity analysis was also performed to determine the sensitivity of each of the outputs with respect to the parameters. In Bayesian inverse modeling the Deviance information criteria (DIC) for each model can be used to quantify the effectiveness or goodness-of-fit of models. DIC is defined as

$$DIC = 2\overline{D(\theta)} - D(\bar{\theta}) \qquad \textit{(Eq. 3-13)}$$

where $\overline{D(\theta)}$ is the mean deviance and $D(\bar{\theta})$ is the deviance of the mean. These are defined as

$$\overline{D(\theta)} = -2E[\ln p(\boldsymbol{y}|\theta)] - 2\log f(y) \qquad \textit{(Eq. 3-14)}$$

$$D(\bar{\theta}) = -\ln p(\boldsymbol{y}|\bar{\theta}) - 2\log f(y) \qquad \textit{(Eq. 3-15)}$$

where f is a standardizing term that is only a function of the observed data; and $\bar{\theta}$ is the expected value of the parameters. Credible intervals for uncertainty analysis were obtained from the posterior distributions of parameters and model predictions. This statistical step should improve model discrimination over previous studies that struggled with large uncertainties (Kovarova-Kovar and Egli, 1998).

3.2.4 Full-scale simulation

It was not immediately apparent how different model structures would impact predictions of performance for a full-scale activated sludge system. For this reason the SUMO modeling software (Dynamita, Nyons, France) was used to simulate various nitrogen removal processes under defined operational conditions under which substrate limitation and competition for NO_2 varied. The baseline for the process configurations mimicked the mainstream nitrogen removal pilot at Blue Plains but was scaled up to represent the size of one full-scale reactor. The system consisted of ten consecutive 2000 m^3 reactors with a step feed into tanks 1, 3, and 6. The influent flow rate was 120000 m^3/d and the characteristics were based on the actual Secondary Effluent from Blue Plains. Biokinetic parameters were selected from a BioWin 3.1 model calibration of the Blue Plains pilot (Al-Omari et al., 2015). The Sumo2 two-step nitrification denitrification framework was used to analyze this system because it incorporates the relevant processes in nitrogen shortcut systems. The following growth rate equations were modified to reflect the three different substrate models: AOB (NH_3 and DO), NOB (NO_2 and DO), AnAOB (NO_2 and NH_3), OHO (readily biodegradable substrate, S_B and DO), OHO (S_B and NO_2), and OHO (S_B and NO_3).

All trials began with the same initial conditions and were run until the simulation reached steady-state. The operational conditions (Table 3-2) were chosen to evaluate the impact of one or two substrate limitations and two or three competitors for NO_2 for the biokinetic characteristics of sludge at Blue Plains. Different kinetic parameters would require different conditions for successful operation.

Table 3-2. Selected conditions for each operational strategy of the SUMO modeling study. The influent nitrogen was fixed at 20 NH_3 – N g / m^3 and 25 degrees C for all simulations. The influent C / N ratio is given in *(mg rbCOD / mg NH3-N)*

Operational strategy	Aerobic fraction	DO set point (mg O_2 / L)	Dynamic SRT (d)	Influent C / N	Seed type
(A) Conventional Full Nit-Denit	100 %	0.1	18	14	None
(B) Nitrogen Shunt	25 %	2.0	15	8	None
(C) Deammonification: high DO	25 %	2.0	15	4	AnAOB
(D) Deammonification: low DO	100 %	0.1	6	4	AnAOB, AOB

3.3 Results

3.3.1 Parameter estimates from Single limitation

When measured in batch tests, kinetic parameters are estimated using a single substrate limitation model, focusing on the limiting substrate. For the purpose of comparison and validation, the parameter estimation was performed using subsets of experimental data that exhibited only

single limitation: DO limitation to estimate $K_{O,NOB}$ and $K_{O,OHO}$; NO_2 limitation to estimate $K_{NO2,NOB}$, $K_{NO2,AnAOB}$, and $K_{NO2,OHO}$; and NH_3 limitation to estimate $K_{NH3,AnAOB}$. The median and 95% credible intervals for parameters from single estimation are represented by grey bars in Figure 3-7 and median values are given in Table 3-1. The bar height corresponds to the uncertainty of a given parameter. The resulting distributions are well within the prior distributions. The VSS of the NOB experiment was 523 ± 26 mg/L. NH_3 was negligible (< 0.01 mg / L) ensuring no NO_2 was produced by AOB. The VSS of the AnAOB experiment was 3660 ± 120 mg/L.

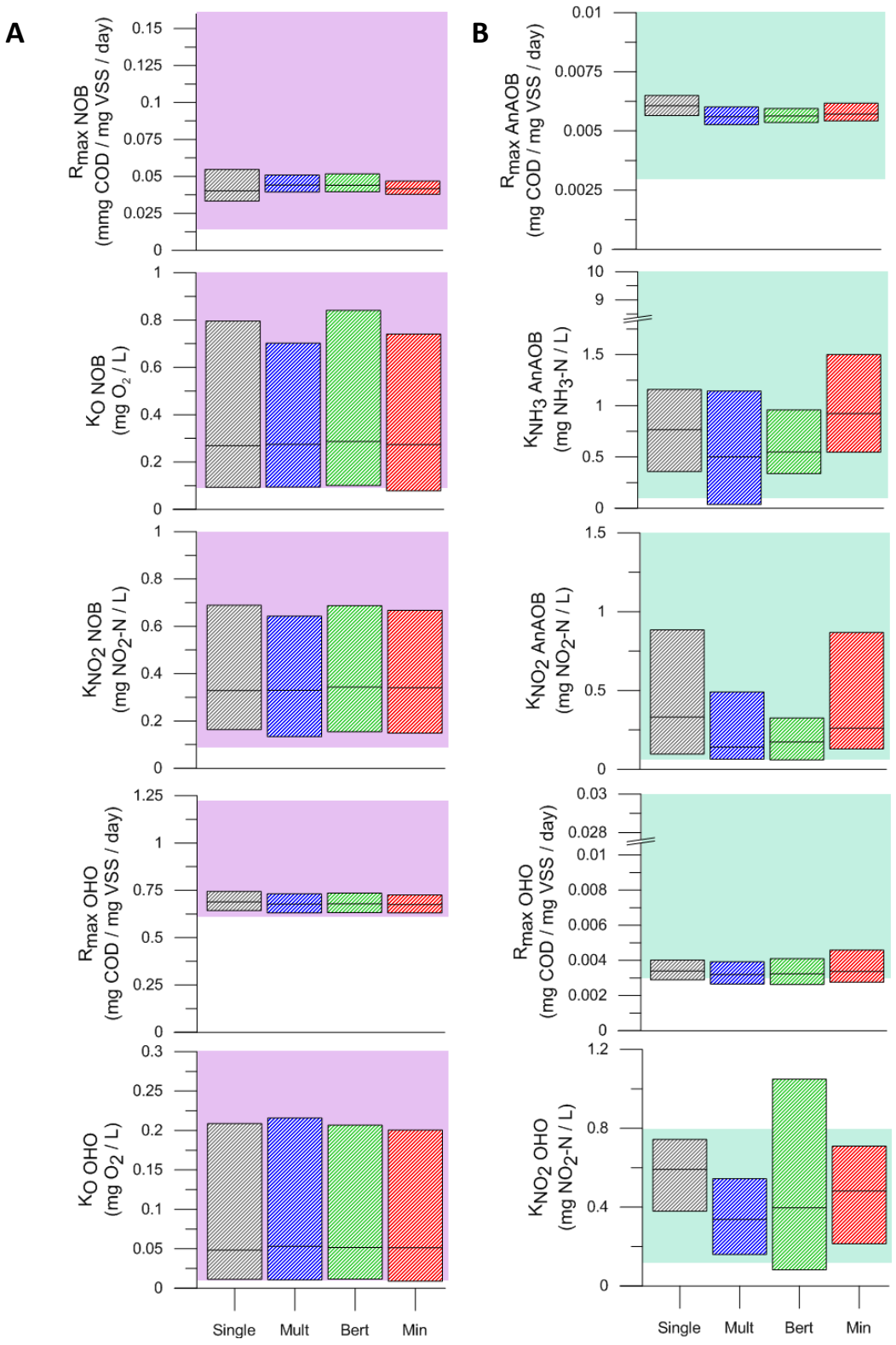


Figure 3-8. The 95% credible interval and median of parameters for single limitation and each dual limitation model for NOB (A) and AnAOB (B) experiments. The prior range for each parameter is shaded lavender (NOB) or teal (AnAOB) in the background.

3.3.2 Model Evaluation using parameters from single limitation

A dual substrate model should be able to predict reaction rates given the parameters derived from singly-limited batch tests and the initial conditions. The results of such forward predictions are given in Figure 3-8. The median single parameters from Table 3-1 were used in each model and compared to the observations for both the NOB and AnAOB experiments. The three models had nearly identical predictions for experiments where only one substrate was limiting as expected (Figure 3-8 NOB-A,B and AnAOB-A,B,D). However, when both substrates were not in excess the models deviated, particularly in the AnAOB experiment. For both experiments, the *minimum* model best approximated the observed data (lowest Forward fit value in Table 3-1). The *Bertolazzi* model predictions were in between those of the *minimum* and *multiplicative*, both in terms of predicted concentrations and error fit.

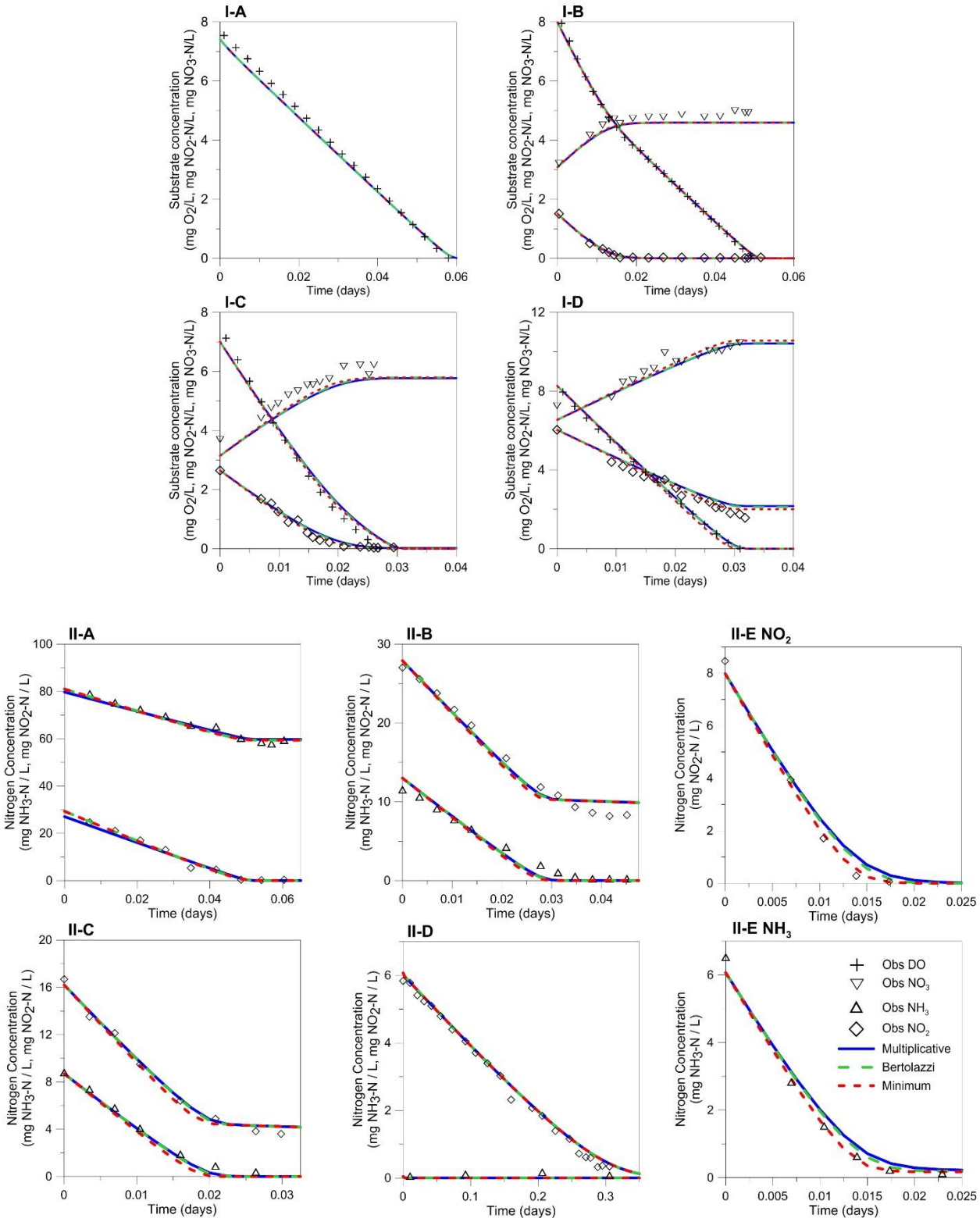


Figure 3-9. Forward model predictions vs. observations for NOB (A-D) and AnAOB (A-F).The parameters estimated from single limitation batch tests were used in each dual limitation model.

3.3.3 Model Evaluation using Inverse modeling

The Bayesian Inverse modeling described in section 2.2 was used to estimate maximum growth rate and half saturation indices based on the experimental data shown in Figure 3-8. Parameters estimated with a given dual substrate model should match those from single limitation experiments. For the NOB experiment the *minimum* model's estimate of the maximum growth rate for NOB is closest to the single limitation value. The rest of the parameters are similar between models and match the single limitation value (95% credible intervals are given in Figure 3-7 and median values in Table 3-1). For the AnAOB experiment the *minimum* model parameter estimates are closest to the single limitation overall. The *minimum* model also had the largest Bayes' Factor which is an index quantifying model appropriateness when a limited number of models are compared for the AnAOB experiment (Bayes' Factors were indistinguishable for the NOB experiment) (Kass and Raftery, 1995).

The higher the sensitivity value associated between a parameter (half-saturation indices, maximum rates) and an observable (DO, NH_3, NO_2, or NO_3) the more influence the parameter has on that predicted observable. A strong positive or negative correlation between two parameters is indicative of the lack of identifiability of model parameters given the observed data. The sensitivity of each parameter and correlation between parameters are similar between models for the NOB experiment whereas there is some deviation between models in the AnAOB experiment (Table 3-3).

Table 3-3. Parameter correlations and sensitivity analysis of parameters on observed constituents (DO, NO2, and NO3) for each dual substrate model for both a) NOB and b) AnAOB experiments.

a)

Model	Parameter	$R_{MAX,NOB}$	$K_{O,NOB}$	$K_{NO2,NOB}$	$R_{MAX,OHO}$	DO	NO_2	NO_3
Multiplicative	$R_{MAX,NOB}$					3.70	14.80	0.30
	$K_{O,NOB}$	0.62				0.59	1.52	0.03
	$K_{NO2,NOB}$	0.65	0.04			0.58	7.28	0.05
	$R_{MAX,OHO}$	-0.01	0.05	-0.05		25.14	7.29	0.11
	$K_{O,OHO}$	-0.04	-0.03	-0.01	0.47	1.87	0.38	0.00
Bertolazzi	$R_{MAX,NOB}$					3.80	15.90	0.29
	$K_{O,NOB}$	0.69				0.53	0.81	0.02
	$K_{NO2,NOB}$	0.66	0.16			0.58	8.53	0.05
	$R_{MAX,OHO}$	-0.05	0.02	-0.08		25.50	8.87	0.10
	$K_{O,OHO}$	-0.10	-0.08	-0.05	0.53	1.79	0.76	0.00
Minimum	$R_{MAX,NOB}$					3.85	14.99	0.29
	$K_{O,NOB}$	0.56				0.43	0.10	0.01
	$K_{NO2,NOB}$	0.71	0.28			0.49	8.41	0.04
	$R_{MAX,OHO}$	-0.03	0.00	-0.03		25.39	7.52	0.09
	$K_{O,OHO}$	-0.07	-0.07	-0.03	0.44	1.72	0.73	0.00

b)

Model	Parameter	$R_{MAX\ AnAOB}$	$K_{NH3,AnAOB}$	$K_{NO2,AnAOB}$	$R_{MAX,OHO}$	NH_3	NO_2
Multiplicative	$R_{MAXAnAOB}$					5.73	3.37
	$K_{NH3,AnAOB}$	0.30				1.93	0.41
	$K_{NO2,AnAOB}$	-0.01	-0.47			0.74	0.37
	$R_{MAX,OHO}$	-0.13	0.24	0.35		2.33	14.62
	$K_{NO2,OHO}$	0.16	-0.02	0.55	0.35	0.60	3.61
Bertolazzi	$R_{MAXAnAOB}$					5.45	2.92
	$K_{NH3,AnAOB}$	0.62				1.80	0.32
	$K_{NO2,AnAOB}$	0.23	-0.01			0.24	0.27
	$R_{MAX,OHO}$	0.09	0.26	0.48		2.52	13.64
	$K_{NO2,OHO}$	-0.01	-0.02	0.12	0.39	0.63	3.85
Minimum	$R_{MAXAnAOB}$					6.15	3.56
	$K_{NH3,AnAOB}$	0.71				2.84	0.62
	$K_{NO2,AnAOB}$	0.77	0.58			0.26	0.39
	$R_{MAX,OHO}$	0.31	0.55	0.42		2.27	13.19
	$K_{NO2,OHO}$	0.07	0.03	0.11	0.30	0.41	3.89

3.3.4 Full-scale simulation

The largest deviation between the simulated nitrogen removal pathways for each dual substrate model occurs in the Low DO Deammonification strategy (Figure 3-9). For conventional full nitrification-denitrification, the *multiplicative* model predicted the best performance in terms of nitrogen removal, followed by *Bertolazzi* and *minimum*. However, *multiplicative* and *Bertolazzi* both exhibited NO_2 cycling where NOB re-nitrify NO_2 after denitratation – the first step of denitrification – resulting in much lower C/N removal efficiency.

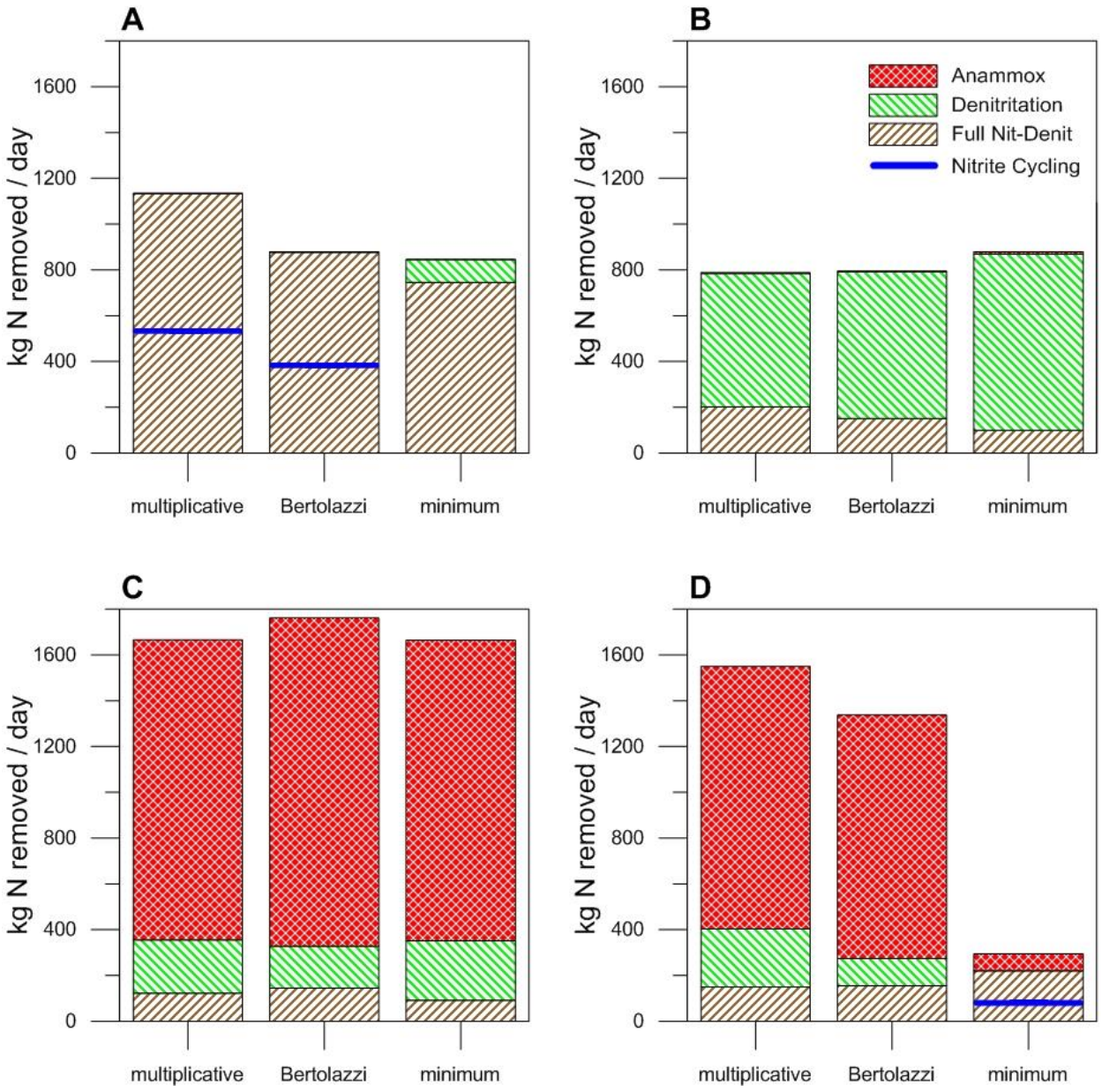


Figure 3-10. Comparison of the nitrogen removal pathway predictions (full denitrification, denitritation only, and Anammox) from the three different dual substrate models (*multiplicative*, *Bertolazzi*, and *minimum*) for four operational strategies (A - Conventional Full Nit-Denit, B – Nitrogen Shunt, C – Deammonification-high DO, D – Deammonification-low DO) resulting from SUMO simulation.

3.4 Discussion

3.4.1 Single Limitation Parameter Estimation

The half-saturation indices for NOB for DO and NO_2-N are nearly equivalent (ca 0.3 mg / L); this leads to maximum deviation between dual substrate models. When half-saturation indices are very dissimilar, the dual limitation model selection has less of an impact. The parameter estimate for $K_{O,NOB}$ is consistent with reported values for suspended growth systems (de Kruek et al., 2007; Arnaldos et al., 2015). The $K_{NO2,NOB}$ estimate from this study is just below the reported range of (0.8 - 7.5 mg NO_2-N / L), but this range includes attached growth systems, which have higher apparent saturation indices due to diffusion effects (Brockmann et al., 2008). These lower half-saturation indices may indicate the dominance of *Nitrospira sp*. (high affinity) over *Nitrobacter sp* (low affinity) (Arnaldos et al., 2015).

Despite coming from the NH_3-rich environment of a sidestream pilot reactor (1000 mg NH_3-N / L) and possible diffusion effects of granule diameters greater than 212 um, the AnAOB half-saturation index for NH_3 was low, less than 1.5 mg NH_3-N/L. This is larger than found by Strous et al. (1999) or Oshiki et al. (2011), but these systems had floc sizes < 100 μm thus removing mass transfer effects. Also, the strains of AnAOB present in the sludge from this study may have had different physiological characteristics than *Candidatus* Brocadia anammoxidans or sinica studied by Strous et al. (1999) or Oshiki et al. (2011), respectively. The K_{NO2} value was similar to that used by Dapena-Mora et al. (2004). Both values were within an order of magnitude of concentrations present in mainstream deammonification, and thus influential as sidestream granules are a potential seed for mainstream systems (Wett et al., 2015).

OHO activity was incorporated in the reactions network for single limitation parameter estimation specifically to remove their influence. The combined consumption of DO by OHO and NOB is observed in batch B until NO_2 is depleted, after which DO is consumed only by OHO. The relative OHO activity was much lower in the AnAOB experiment than in NOB experiment, however, continued NO_2 consumption by OHO can be seen in AnAOB batch tests B and C after the NH_3 is depleted.

3.4.2 Model Comparison

In general, when parameters are calibrated to a given model structure, the model is able to recreate the observed data well – even if this is achieved by selecting unrealistic parameter values. Thus, the model simulations based on the median parameter estimation for each model are similar to each other for both Experiments and capture the observed data equally well. The 95% credible interval for the MCMC simulations of each model overlap very closely with each other for the NOB experiment and capture the observations well (Figure 3-10); they overlap for all but batch test E for the AnAOB experiment (Figure 3-11). Although none of the dual substrate models selected unreasonable values, they do deviate from the values estimated from single-limitation. As predicted, the *multiplicative* and *Bertolazzi* models estimated higher maximum growth rate (the NOB experiment) and lower half-saturation indices (the AnAOB experiment) in order to compensate for lower calculated rates.

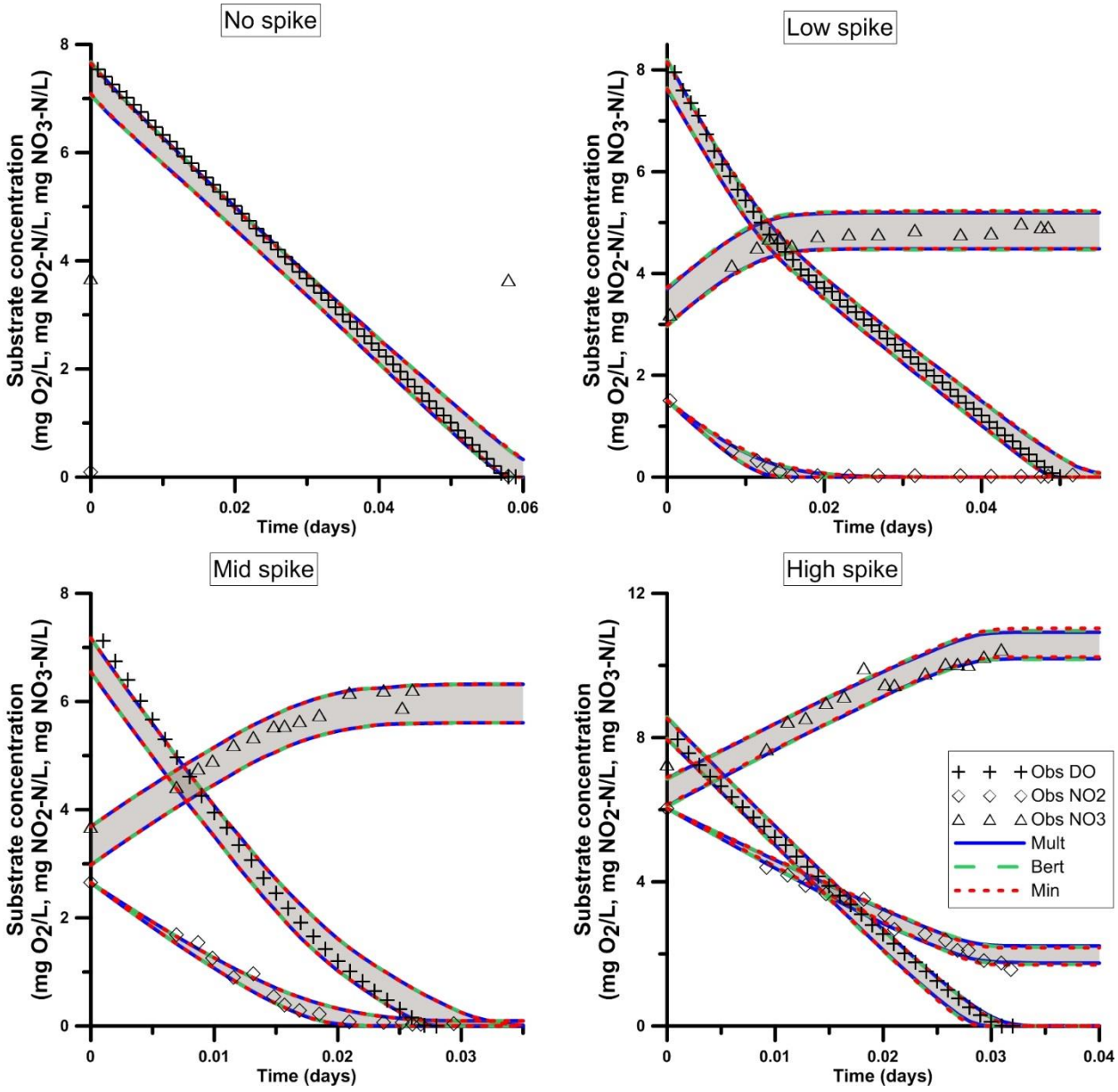
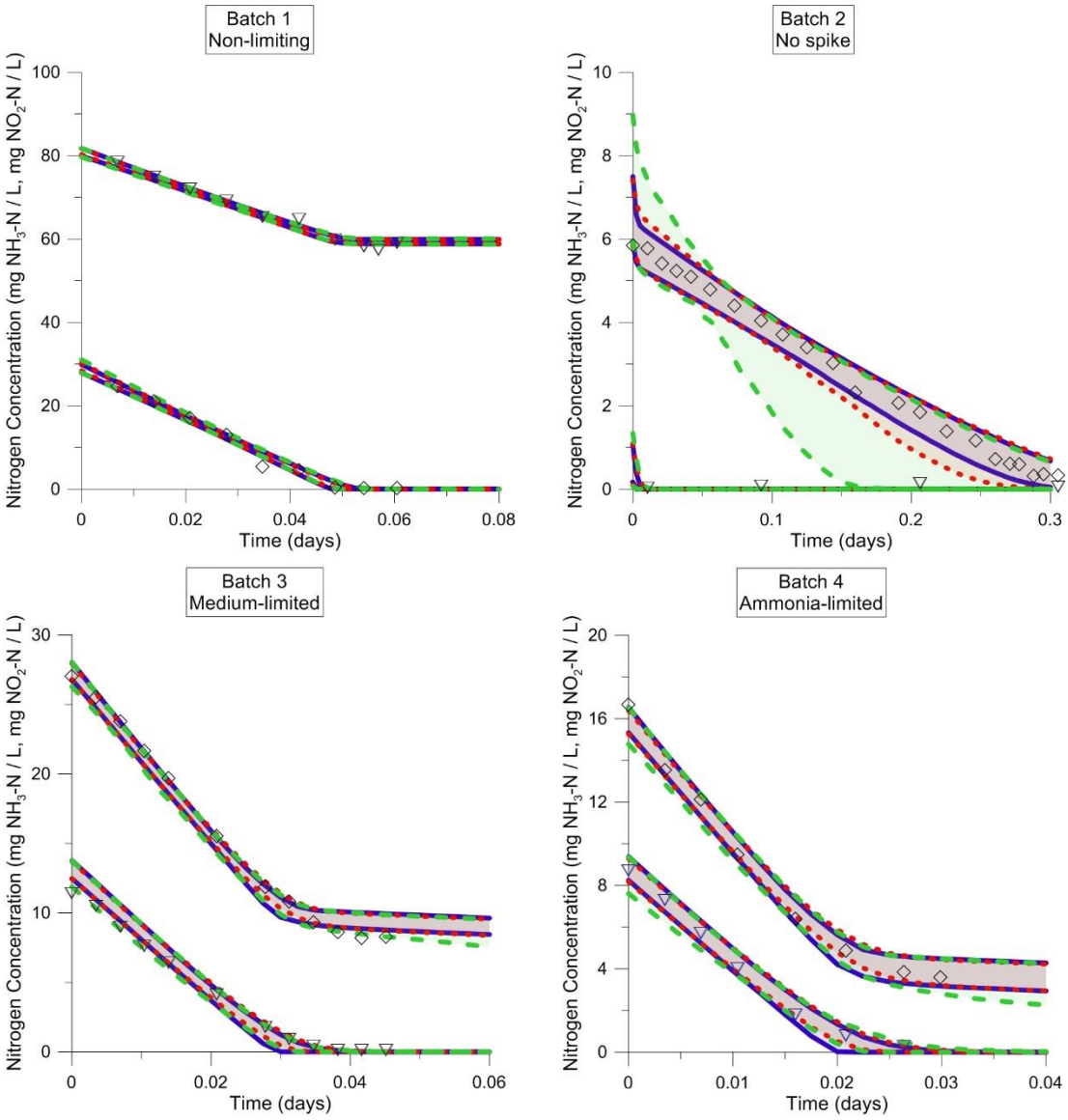


Figure 3-11. Model predictions vs observations for NOB experiment. The 95% credible interval for MCMC output for each dual limitation model is plotted against the observations of DO, NO_2, and NO_3 for NOB batch test C.


Batch 1
Non-limiting
Batch 2
No spike
Batch 3
Medium-limited
Batch 4
Ammonia-limited
Nitrogen Concentration (mg NH3-N / L, mg NO2-N / L)
Time (days)

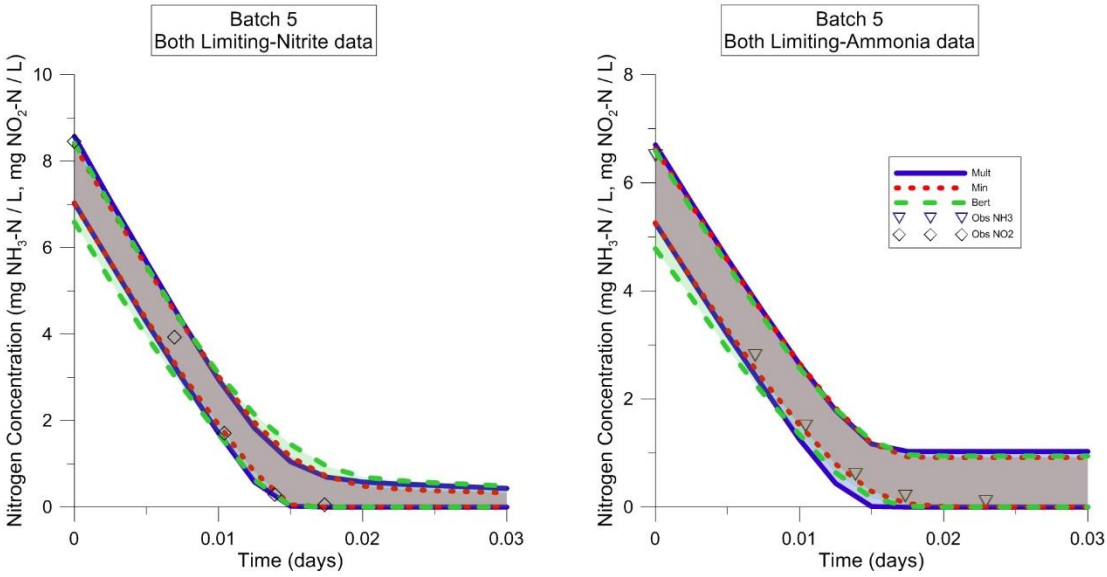


Figure 3-12. Model predictions vs observations for AnAOB experiment. The 95% credible interval for MCMC output for each dual limitation model is plotted against the observations of NH_3 and NO_2 for all 5 batch tests.

The maximum growth rate for OHO had the highest sensitivity, i.e. it had the most influence on the observations for both experiments, followed by the maximum growth rate for NOB and AnAOB for those respective experiments (Table 3-3). The sensitivity of NOB half-saturation indices was similar, which may be a result of similar values. Moreover for the AnAOB experiment the higher half-saturation index ($K_{NH3,AnAOB}$) had a higher impact on results. Overall, the half-saturation indices had a strong positive correlation to maximum reactions rates, meaning a higher estimate for R_{max} tended to accompany a high estimate of K_s for given observations. The *minimum* model, and to a lesser extent *Bertolazzi*, also showed a weak positive correlation between the half-saturation indices for NOB.

Of the models considered, the *minimum* model performed best both by selecting parameters close to those from single limitation tests and by best representing the observations using the single

limitation parameters. The other two models (especially *multiplicative*) either overestimated maximum growth rates (as in the NOB experiment) or underestimated the half-saturation indices (as in the AnAOB experiment). These results suggest that limiting biokinetic reactions behave as limiting chemical reactions. When using parameters based upon single limitation, the *multiplicative* and *Bertolazzi* models underestimated reaction rates at substrate concentrations similar to those found in mainstream deammonification systems (both experiments). When this effect is extended to more than two terms – commercial software can easily include five Monod terms for a given reaction – systems are undoubtedly miscalibrated by artificially overestimating maximum rates. It is preferable for system operators to be able to use parameter estimates from single limitation batch tests in a multi-substrate model. Based on the result of this study it is deemed that the *minimum* model provides a more realistic estimate of the parameter compared to the other two models. Further work could explore dual substrate limitation for different microbial groups such as AOB and OHO and whether the *minimum* model applies for other factors that affect growth rates such as pH and inhibition.

3.4.3 Model approximations

It is worth noting that, when implementing dual substrate limitation models in numerical Activated Sludge Models it is sometime beneficial to avoid formulations that contain discontinuous first-order derivatives such as in the minimum function. For this purpose, the *minimum* model can be approximated using the following analytic expression:

$$\mu \approx \mu_{max} \frac{-1}{n} \ln \left[e^{-n\left(\frac{S_1}{S_1+K_1}\right)} + e^{-n\left(\frac{S_2}{S_2+K_2}\right)} \right] \qquad \textit{(Eq. 3-16)}$$

where *n* is a smoothness factor such that the maximum deviation between *Eq. (3-16)* and *Eq. (3-4)* is equal to $\ln(2)/n$ (Cook, 2011). Other analytic alternatives exist that are everywhere smooth and differentiable, but this formula lends itself well to the topic at hand and may be used if a given software cannot handle the minimum function.

3.4.4 Full-scale Simulations

The impact of model selection in small-scale batch tests may seem small, but the impact is magnified in the full-scale simulations, with highly variable operational conditions. In general, the *minimum* model will predict higher reaction rates than the *Bertolazzi* model, which will in turn predict higher rates than the commonly-used *multiplicative* model at low substrate concentrations. Although there is some deviation between models for each strategy, the largest deviation occurs in strategies with more limitations and competition (Figure 3-9 A,D). The Nitrite-cycling in full nitrification-denitrification is likely due to a worse NOB out-selection indicated by high NOB/AOB ratio as AOB were more limited due to lower affinity for both DO and NH_3 than NOB for DO and NO_2. The *minimum* model predicted higher aerobic OHO rates; increased competition for readily biodegradable substrate (S_B) reduced the system's denitrifying potential and therefore the total nitrogen removal. Deammonification with continuous low aeration exhibited the most model deviation stemming from different degrees of NOB-washout. The substrate limitations on NOB were not enough for effective washout in the *minimum* model as it generally predicts faster growth rates. There was also a slight increase in model deviation between the nitrite shunt and deammonification with high DO strategy, suggesting that growth rates are sensitive to the number of competitors in a system; this is most prominent when competition contributes to substrate limitation.

As expected, the predicted operational performance depends strongly on which dual substrate limitation model is used. Model selection is crucial for design optimization. For the mainstream nitrogen system based on the pilot at DC Water AWWTP, this is particularly relevant for operational strategies most likely to exhibit multiple limitations for NO_2 competitors, such as simultaneous nitrification-denitrification or deammonification at a low DO set point. Model selection is less important when there are fewer competitors and only single limitation in a system such as a nitrite shunt scenario with a high DO set point and intermittent aeration. As we move toward more advanced nutrient treatment strategies we must be cognizant of the impact of model selection on operational control success.

Although the *minimum* model had the best experimental fit, it may not be the most appropriate choice to model an arbitrary nitrifying system. Obviously the model used for calibration should be used for prediction. However, if parameters are directly measured independently, the *minimum* model is recommended. If maximum rate and affinity parameters are estimated together (model-fitting) the *multiplicative* model (or one of the other proposed interactive models (Lee et al., 1984) may be used, as the maximum rates will already be artificially raised. Further study should examine which model is most appropriate for other microbial nitrifying environments, such as biodegradation of environmental contaminants in soil (Okpokwasili and Nweke, 2005; Massucci et al., 2015; Maynard et al., 2015). More accurate models will play an even larger role in heterogeneous systems, such as biofilms (Odencrantz, 1992; Gonzo et al., 2014)

3.4.5 Theoretical impact of diffusion on dual limitation

The strong correlation between the maximum growth rate and a half-saturation coefficient for an organism group is not a new discovery. It is also known that these parameter values seem

to vary over time for a given system. Many studies have sought to analyze the identifiability of these two parameters and design experiments to maximize information (Button, 1993; Dochain et al., 1995; Vanrolleghem et al., 1995). Recent studies have proposed both empirical and physical models that attempt to understand the relationship between them (Shaw et al., 2013; Arnaldos et al., 2015). The physical models tend to focus on diffusion of substrate through the floccular (irregular and loose) or granular (spherical and compact) structure that these organism colonies form or through the cell wall of the organisms themselves. It therefore warrants discussion on how diffusion would impact a system with two substrates. Given the continuity equation

$$\frac{d\rho}{dt} + \oiint \vec{J} \bullet d\vec{s} = -\mu(\rho) \qquad \textit{(Eq. 3-17)}$$

and Fick's Law

$$\vec{J} \approx -D\vec{\nabla}\rho = -D\frac{d\rho}{dr} \qquad \textit{(Eq. 3-18)}$$

where ρ is the substrate concentration or density and $\vec{J}$ is the mass flux we can define the diffusion equation for this substrate diffusing from surface boundary *b* into a spherical, homogenous biofilm of small radius r_o

$$\frac{d\rho}{dt} - D4\pi r_o^2 \frac{(\rho_b - \rho)}{(r_o - 0)} = -\mu(\rho) \qquad \textit{(Eq. 3-19)}$$

If we now consider two substrates S_1 and S_2, the enzyme *E* that facilitates the uptake of these substrates and rate constants k_1 *(reaction of E with* S_1*)*, k_2 *(reaction of E and* S_1 *with* S_2*)*, and k_3 *(dissociation of E from final product)* we have the following equations at the center

$$\frac{dS_1}{dt} = D4\pi r_o\left(S_{1,b} - S_1\right) - k_1 S_1 E$$
$$\frac{dE_{S1}}{dt} = +k_1 S_1 E - k_2 E_{S1} S_2$$
$$\frac{dS_2}{dt} = D4\pi r_o\left(S_{2,b} - S_2\right) - k_2 E_{S1} S_2 \qquad \textit{(Eq. 3-20)}$$
$$\frac{dE_{S1,S2}}{dt} = +k_2 E_{S1} S_2 - k_3 E_{S1,S2}$$
$$\frac{dE}{dt} = -k_1 S_1 A + k_3 E_{S1,S2}$$

If we assume steady state conditions and that the total amount of the enzyme E_o is conserved we find after some algebraic manipulations that the difference between the substrates $S_1 - S_2 = S_{1,b} - S_{2,b}$ is constant and the overall substrate uptake rate is given as

$$\mu(S_1, S_2) = \frac{3D}{r_o}(S_{1,b} - S_1) = \frac{3D}{r_o}(S_{2,b} - S_2) = \frac{E_0 S_1 S_2}{\frac{S_2}{k_1} + \frac{S_1}{k_2} + \frac{S_1 S_2}{k_3}} \qquad \textit{(Eq. 3-21)}$$

which is symmetrical with respect to both substrates. Unless the substrates are at an equal concentration at the boundary, the diffusion effect will exacerbate the limitation of the Monod term of the substrate at lower concentration. To then use the *multiplicative* model of dual limitation, would over incorporate the diffusion effect.

Similarly, if we take a step back and examine the use of the extant half-saturation parameter put forth by Shaw et al. (2013),

$$K_s = K_o + \frac{V_{\max}}{(4\pi r_o D)} \qquad \textit{(Eq. 3-22)}$$

where K_o is the intrinsic parameter without the effect of diffusion, and V_{max} is the maximum growth rate per cell (proportional to μ_{max}) we can see that there is an order of magnitude difference between the *minimum* and *multiplicative* models with respect to diffusion or boundary layer thickness:

$$\mu = \mu_{\max} \frac{S_1}{K_{0,1} + \frac{V_{\max}}{(4\pi r_o D)} + S_1} \frac{S_2}{K_{0,2} + \frac{V_{\max}\alpha}{(4\pi r_o D)} + S_2}$$

$$\mu = \frac{\mu_{\max} S_1 S_2}{K_{0,1}K_{0,2} + \frac{V_{\max}^2 \alpha}{(4\pi r_o D)^2} + \frac{V_{\max}}{(4\pi r_o D)}(\alpha K_{0,1} + K_{0,2} + \alpha S_1 + S_2) + K_{0,1}S_1 + K_{0,2}S_2 + S_1 S_2} \quad \textit{(Eq. 3-23)}$$

$$\mu = \frac{\mu_{\max} S_1 S_2}{\frac{V_{\max}^2 \alpha}{(4\pi r_o D)^2} + \frac{V_{\max}}{(4\pi r_o D)}(\alpha S_1 + S_2) + S_1 S_2}$$

where α is a place holder for stoichiometric factors and the intrinsic parameters K_0 vanish to zero. Thus it would be an enlightening experiment if someone were to set up steady state chemostats for a given biofilm of different thicknesses. If the growth response is linearly or quadratically proportional to the diffusion, this will be more evidence for the *minimum* or *multiplicative* model, respectively.

3.5 Summary

This study has reported the results of dual substrate model discrimination based on experiments using NOB limited by dissolved oxygen and NO_2 and AnAOB limited by NH_3 and NO_2. Both a deterministic genetic algorithm and Bayesian inference using Markov Chain Monte Carlo simulations were used to estimate model-specific parameters and determine which model (*multiplicative*, *Bertolazzi*, and *minimum*) best describes the experimental observations. A full-scale simulation was performed using the SUMO commercial software to observe the potential impact of model selection on system performance and control. The main messages are the following:

- For both microbial groups studied, the *minimum* model best fit the experimental data using parameters derived from single substrate limiting batch tests.

- For both microbial groups studied, the *minimum* model selected for parameter values closer to that of single limitation estimates than the *multiplicative* model – which overestimated maximum rates or underestimated half-saturation indices.
- For advanced nitrogen removal systems with substrate competition and low concentrations for two or more substrates – like mainstream deammonification – model selection has a large impact on performance prediction and therefore control strategies.

Chapter 4 - In situ Parameter and Uncertainty Estimation for Mainstream Deammonification

4.1 Introduction

4.1.1 Mainstream Deammonification

As global populations continue to urbanize, nutrient loading to receiving urban water bodies increases and the importance of nitrogen treatment also grows. The current practice for most modern wastewater treatment facilities to remove nitrogen is to artificially enhance what would happen in nature; that is to foster complete biological nitrification and denitrification before the effluent enters the open water. As described in Chapter 1, the nitrification step traditionally relies on aerobic autotrophic bacteria that require dissolved oxygen (DO) as an electron acceptor to oxidize ammonia to nitrate. The process of aerating reactor tanks is very energy intensive and is easily the largest draw on electricity for a facility (which is usually derived via fossil fuels). The second step, denitrification, traditionally uses heterotrophic organisms that use readily biodegradable carbon as an electron donor and nitrate as the acceptor. Due to the historical progression of wastewater treatment, where advancements are often tagged onto the end of the line, readily biodegradable carbon has often already been removed by the time the wastewater stream enters the nitrogen treatment reactors. In order to compensate for this, carbon sources like methanol or acetate are added to the denitrification reactors. As we move toward a framework of viewing nutrients in wastewater as a resource worthy of recovery, the idea of adding more carbon nutrients after removing them in a previous step seems highly inefficient - especially when carbon can be redirected to digestion reactors that create electricity producing – methane gas.

Fortunately, organisms have been found that enable more efficient biological nitrogen removal. A cutting edge process that takes advantage of one of these recently discovered organisms

is via anaerobic ammonium oxidation or the Anammox process, where some ammonia is partially oxidized to nitrite and then the two are converted directly to dinitrogen gas according to the reaction below (Strous, 1998):

$$NH_4^+ + 1.32\ NO_2^- + 0.066\ HCO_3^- + 0.13\ H^+ \rightarrow$$

$$1.02\ N_2 + 0.256\ NO_3^- + 0.066\ CH_2O_{0.5}N_{0.15} + 2.03\ H_2O \qquad \textit{(Eq. 4-1)}$$

When this 'shortcut' nitrogen removal process is used to treat mainstream wastewater (main incoming flow) it is called mainstream deammonification. It is a very active area of research because it can have substantial financial and environmental benefits due its potential to reduce aeration and external carbon demand, reduce greenhouse gas nitrous oxide (N_2O) emissions, and reduce sludge production (Al-Omari et al., 2015; Turk and Mavinic, 1986; De Clippeleir et al., 2013; Siegrist et al., 2008; Monballiu et al., 2013). Although there has been clear success in using partial nitrification/anammox in side-stream (flows emanating from diverted processes like sludge dewatering before transport) treatment (Peng et al., 2012; Gao et al., 2009; Winkler et al., 2011; Vlaeminck et al., 2012), there still exist several challenges in effectively using it in mainstream systems due to the slow growth rate of anaerobic ammonium oxidizing bacteria (AnAOB) compared to the other participating organism groups (ammonia oxidizing bacteria - AOB, nitrite oxidizing bacteria - NOB, and ordinary heterotrophic organisms - OHO), and lack of success in effective out-selection of NOB; side stream processes have the benefit of higher temperatures and ammonia concentrations which increase free ammonia concentrations that are inhibitory to NOB (Al-Omari et al., 2015; Ali et al., 2014). Several experimental and modeling studies have attempted to find the optimal conditions resulting in effective deammonification in mainstream and side-stream treatment (Fernandez et al., 2011; Fux et al., 2002; Al-Omari et al., 2015; Hao et al., 2002; Hubaux et al., 2015). Compared to classical nutrient removal systems, Mainstream

Deammonification is rather complex with nearly a dozen state variables and their accompanying nonlinear differential equations.

4.1.2 Model Framework Development

In 1987 the International Water Association established the first Activated Sludge Model (ASM) that provided the framework for future models. Full nitrification-denitrification, bacteria growth and decay, alkalinity consumption, and hydrolysis of entrapped organics were the basis of the original ASM. The model framework for this study was based on ASM1 with the additions from literature that incorporate separate nitritation from nitratation reactions (Wyffels et al, 2004) and Anammox reactions (Dapena-Mora, 2004). Extensive developments in ASM have been made in the areas of phosphorus precipitation, carbon digestion, and sulfur sorption/desorption but this study is limited to nitrogen modeling. The main components or constituents of this model are listed in Table 4-1. Other compounds were part of the original framework from ASM1, namely dissolved N_2 and particulate biodegradable carbon substrate. These components were removed from the model framework for this work as they were not needed for rate expressions, were not consumed by any process, and did not relate to any observations.

Table 4-1. Compounds or state variables modeled in this study.

Symbol	*Name*	*Units*
X_I	*Inert organic particulates*	*mg COD / L*
X_S	*Slowly biodegradable substrate*	*mg COD / L*
X_{OHO}	*Ordinary heterotrophic organism biomass*	*mg COD / L*
X_{AOB}	*Aerobic ammonia oxidizing organism biomass*	*mg COD / L*
X_{NOB}	*Aerobic nitrite oxidizing organism biomass*	*mg COD / L*
X_{AnAOB}	*Anaerobic ammonia oxidizing organism biomass*	*mg COD / L*
S_O	*Dissolved oxygen*	*mg O_2 / L*
S_S	*Readily biodegradable substrate*	*mg COD / L*
S_{NO2}	*Soluble inorganic nitrite*	*mg NO_2-N / L*
S_{NO3}	*Soluble inorganic nitrate*	*mg NO_3-N / L*
S_{NH3}	*Soluble inorganic ammonia*	*mg NH_3-N / L*

The Gujer or Petersen matrix for this model is given in Table 4-2(a-c) and the relevant parameters are listed in Table 4-3. This format conveniently arranges how each process (rows) either produces (positive) or consumes (negative) each of the components (columns). The stoichiometric factors convert each component into units of 'chemical oxygen demand' or COD. All of the constituents in activated sludge models are reported in units of COD in order to simplify the Gujer matrix. Each process is in units of mg COD / (L * d). The time derivative for a given compound C_j is given by

$$\frac{dC_j}{dt} = \sum_i r_i s_{i,j} \qquad \textit{(Eq. 4-2)}$$

where r_i is the i^{th} rate expression and $s_{i,j}$ is the stoichiometric factor for the j^{th} compound in regards to the i^{th} process. Unfortunately only soluble COD can easily and reliably be measured in our lab, so the biomass is estimated by the volatile suspended solids (VSS) concentration. VSS are usually 80% of the Total suspended solids (TSS). TSS is reported as the concentration of dry particulates and colloidal particles greater than 0.45 microns. Biomass concentrations are empirically related to VSS by the following equation

$$\sum_n X_n = 1.42(VSS) \qquad \textit{(Eq. 4-3)}$$

where n are the four biomass groups: AOB, NOB, OHO, and AnAOB (Henze et al., 1987). Although VSS was measured in this study, this relationship was only used as a rough quality check on simulation results.

The component called readily biodegradable COD (S_S) consists of soluble compounds with low molecular weights that can be immediately metabolized by organisms. As the influent to a municipal treatment facility has a highly diverse chemical makeup, it is not practical to quantify and measure the many compounds that could potentially fall under the term S_S. Instead, it is simply calculated as the difference between influent and effluent soluble COD, where the effluent soluble COD is assumed to be inert. (Makinia et al., 2010).

The experiments of the previous chapter were all batch tests performed at room temperature. The mainstream deammonification pilot, on the other hand, was operated using real plant influent that caused the temperature in the reactors to range from 18 to 32 C. Like other chemical or biochemical reactions, activated sludge processes are sensitive to temperature and are

generally modeled according to the modified Arrhenius equation which relates rate constants with the activation energy of a given reaction

$$\mu(T) = Ae^{\frac{E_a}{k_B T}} \qquad \textit{(Eq. 4-4)}$$

where A is a collision factor, E_a is the activation energy $[J / mol]$ and k_B is the Boltzmann constant. Due to the heterogeneities of these systems, rather than obtain the activation energy, a reference term is determined, typically at a temperature of 293 K and an Arrhenius coefficient θ is estimated according to the following simplification

$$\frac{\mu(T)}{\mu(T_{ref})} = e^{\frac{E_a (T - T_{ref})}{k_B T T_{ref}}} \doteq \theta^{(T - T_{ref})} \qquad \textit{(Eq. 4-5)}$$

In the temperature ranges expected in water bodies (273 – 313 K), *Eq. (4-5)* is valid as θ for a given process is relatively constant (~ 1% deviation) (Chapra, 1997).

Table 4-2. Reaction network for Mainstream deammonification pilot modeling. Due to size constraints the model is split into three tables where Table 4-2 **a)** gives stoichiometric factors for particulate constituents (X_j), Table 4-2 **b)** gives the stoichiometric factors for soluble constituents (S_j), and Table 4-2 **c)** gives the rate expressions for each process.

a)

Process	X_I	X_S	X_{OHO}	X_{AOB}	X_{NOB}	X_{AnAOB}
Aerobic growth of heterotrophs			1			
Anoxic growth of heterotrophs on nitrite			1			
Anoxic growth of heterotrophs on nitrate			1			
Decay of heterotrophs	f_i	1 - f_i	-1			
Aerobic growth of autotrophic ammonia oxidizers				1		
Decay of aerobic ammonia oxidizers	f_i	1 - f_i		-1		
Aerobic growth of autotrophs nitrite oxidizers					1	
Decay of aerobic nitrite oxidizers	f_i	1 - f_i			-1	
Anoxic growth of ammonia oxidizers						1
Decay of anoxic ammonia oxidizers	f_i	1 - f_i				-1
Hydrolysis of entrapped organics		-1				

b)

Process	S_O	S_S	S_{NO2}	S_{NO3}	S_{NH3}
Aerobic growth of heterotrophs	$-\frac{\left(1-Y_{OHO,O}\right)}{Y_{OHO,O}}$	$-\frac{1}{Y_{OHO,O}}$			$-i_{nbm}$
Anoxic growth of heterotrophs on nitrite		$-\frac{1}{Y_{OHO,NO2}}$	$-\frac{\left(1-Y_{OHO,NO2}\right)}{1.71\ Y_{OHO,NO2}}$		$-i_{nbm}$
Anoxic growth of heterotrophs on nitrate		$-\frac{1}{Y_{OHO,NO3}}$	$\frac{\left(1-Y_{OHO,NO3}\right)}{1.14\ Y_{OHO,NO3}}$	$-\frac{\left(1-Y_{OHO,NO2}\right)}{1.14\ Y_{OHO,NO3}}$	$-i_{nbm}$
Decay of heterotrophs					$i_{nbm} - f_p\, i_{np}$
Aerobic growth of autotrophic ammonia oxidizers	$-\frac{\left(3.43-Y_{AOB}\right)}{Y_{AOB}}$		$\frac{1}{Y_{AOB}}$		$\frac{1}{Y_{AOB}} - i_{nbm}$
Decay of aerobic ammonia oxidizers					$i_{nbm} - f_p\, i_{np}$
Aerobic growth of autotrophs nitrite oxidizers	$-\frac{\left(1.14-Y_{NOB}\right)}{Y_{NOB}}$		$-\frac{1}{Y_{NOB}}$	$\frac{1}{Y_{NOB}}$	$-i_{nbm}$
Decay of aerobic nitrite oxidizers					$i_{nbm} - f_p\, i_{np}$
Anoxic growth of ammonia oxidizers			$-1.52 - \frac{1}{Y_{AnAOB}}$	*1.52*	$-\frac{1}{Y_{AnAOB}} - i_{nbm}$
Decay of anoxic ammonia oxidizers					$i_{nbm} - f_p\, i_{np}$
Hydrolysis of entrapped organics		1			

c)

Process	*Rate expression*
Aerobic growth of heterotrophs	$\mu_{OHO} X_{OHO} \frac{S_O}{K_{O,OHO} + S_O} \frac{S_S}{K_{S,OHO} + S_S}$
Anoxic growth of heterotrophs on nitrite	$\mu_{OHO} \eta_{NO2} X_{OHO} \frac{K_{O,OHO}}{K_{O,OHO} + S_O} \frac{S_{NO2}}{K_{NO2,OHO} + S_{NO2}} \frac{S_S}{K_{S,OHO} + S_S}$
Anoxic growth of heterotrophs on nitrate	$\mu_{OHO} \eta_{NO3} X_{OHO} \frac{K_{O,OHO}}{K_{O,OHO} + S_O} \frac{S_{NO3}}{K_{NO3,OHO} + S_{NO3}} \frac{S_S}{K_{S,OHO} + S_S}$
Decay of heterotrophs	$b_{OHO} X_{OHO}$
Aerobic growth of autotrophic ammonia oxidizers	$\mu_{AOB} X_{AOB} \frac{S_O}{K_{O,AOB} + S_O} \frac{S_{NH3}}{K_{NH3,AOB} + S_{NH3}}$
Decay of aerobic ammonia oxidizers	$b_{AOB} X_{AOB}$
Aerobic growth of autotrophs nitrite oxidizers	$\mu_{NOB} X_{NOB} \frac{S_O}{K_{O,NOB} + S_O} \frac{S_{NO2}}{K_{NO2,NOB} + S_{NO2}}$
Decay of aerobic nitrite oxidizers	$b_{NOB} X_{NOB}$
Anoxic growth of ammonia oxidizers	$\mu_{AnAOB} X_{AnAOB} \frac{K_{O,AnAOB}}{K_{O,AnAOB} + S_O} \frac{S_{NO2}}{K_{NO2,AnAOB} + S_{NO2}} \frac{S_{NH3}}{K_{NH3,AnAOB} + S_{NH3}}$
Decay of anoxic ammonia oxidizers	$b_{AnAOB} X_{AnAOB}$
Hydrolysis of entrapped organics	$\mu_{Hyd} X_{OHO} \frac{X_S}{K_{S,Hyd} X_{OHO} + X_S}$

Table 4-3. Mainstream deammonification model parameters. Parameters with a range (high and low) were estimated while fixed parameters

Parameter type	*Symbol*	*low*	*high*	*θ*
Maximum growth rates (1 / d)	μ_{OHO}	*0.1*	*10*	*1.029*
	μ_{AOB}	*0.8*	*2*	*1.062*
	μ_{NOB}	*0.7*	*1.5*	*1.07*
	μ_{AnAOB}	*0.08*	*0.15*	*1.1*
Oxygen affinity or inhibition (mg O_2 / L)	$K_{O,OHO}$	*0.8*		
	$K_{O,AOB}$	*0.1*	*1.5*	
	$K_{O,NOB}$	*0.1*	*1.5*	
	$K_{O,AnAOB}$	*0.05*		
Substrate affinity (mg S or N / L)	$K_{S,OHO}$	*2*		
	$K_{NO2,OHO}$	*0.1*	*1.5*	
	$K_{NO2,NOB}$	*0.05*	*1.5*	
	$K_{NO2,AnAOB}$	*0.1*	*1.5*	
	$K_{NO3,OHO}$	*0.1*	*1.5*	
	$K_{NH3,AOB}$	*0.3*	*1.5*	
	$K_{NH3,AnAOB}$	*0.1*	*2*	
Correction factors (-)	η_{NO2}	*0.6*		
	η_{NO3}	*0.6*		
Decay rate (1 / d)	b_{OHO}	*0.62*		
	b_{AOB}	*0.17*		
	b_{NOB}	*0.17*		
	b_{AnAOB}	*0.019*		
Hydrolysis of entrapped organics	μ_{hyd}	*3*		
	$K_{s,hyd}$	*0.03*		
Biomass yield (mg COD biomass produced / mg COD consumed)	$Y_{OHO,O}$	*0.666*		
	$Y_{OHO,NO2}$	*0.54*		
	$Y_{OHO,NO3}$	*0.54*		
	Y_{AOB}	*0.15*		
	Y_{NOB}	*0.09*		
	Y_{AnAOB}	*0.114*		
Fractions	i_{nbm}	*0.086*		
	i_{np}	*0.06*		
	f_p	*0.08*		
	f_i	*0.08*		

4.1.3 Motivation for Inverse modeling and Parameter estimation

Mathematical process modeling is necessary to predict system responses and sensitivities to environmental changes. It can also discover optimal operating conditions or control strategies resulting in effective utilization of the anammox process as well as help implement model predictive control (MPC) algorithms aiming at maintaining optimal conditions under dynamically changing conditions. The predictions of mathematical process models of activated sludge systems in general are affected by the large number of biokinetic and stoichiometric parameters representing various processes. For example, the effectiveness of controlling the anammox process by controlling DO level highly depends on the difference between DO affinity constants for ammonia oxidation and nitrite oxidation (Al-Omari et al., 2015; Hanaki et al., 1990; Bernet et al., 2001). Thus more certain and accurate parameter estimates and more realistic physical models can lead to improved effluent quality and reduced operation cost. However, there is a great deal of uncertainty associated with the parameters involved in Mainstream Deammonification. A wide range of values have been reported for many of these parameters; the oxygen affinity coefficient for AOB ($K_{O,AOB}$) ranging from 0.1 to 1.45 mg DO/L (Blackburne et al., 2008; Wett et al., 2013; Sin et al., 2008). Empirical and mechanistic models have been proposed to relate system characteristics like floc size distribution on oxygen mass transfer rates and therefore the affinity coefficients; such effects are likely at least partly responsible for the variability of the values estimated in different studies either by direct measurement or model-fitting (Beccari et al., 1992).

Although batch and continuous experiments have been used to determine the values of parameters controlling deammonification (Blackburne et al., 2008), the values obtained may not be transferable to full-scale systems due to the inherent differences between batch and full-scale reactor environments and the interactive effects of various environmental factors influencing the

outcome in full-scale reactors (Sharifi et al., 2014). It is also possible that parameter values evolve in response to certain reactor conditions as different strains within an organism group capitalize on intrinsic advantages and dominate (Wett et al., 2011). Estimating the values of the parameters based on full-scale data or pilot studies representing the full-scale systems can potentially provide additional information about effective values of the parameters in real systems.

The goal of this study was to evaluate the ability of long-term pilot studies to provide information about the values of deammonification parameters. For this purpose a pilot representing a full-scale nitrification-denitrification plant was operated and monitored over a period of 380 days. Then an automatic probabilistic parameter estimation method was used to estimate the joint probability distributions of biokinetic parameters. These resultant distributions are a needed improvement over the many conflicting single values existing in the literature to date because the PDFs also provide uncertainty analysis that can improve decision-making (Sharifi et al, 2014). The Bayesian parameter estimation algorithm allows quantifying the identifiably of the parameters based on observed data and the collinearity between the parameters. A sensitivity analysis was performed to understand the effect of each of the parameters on the deammonification rates.

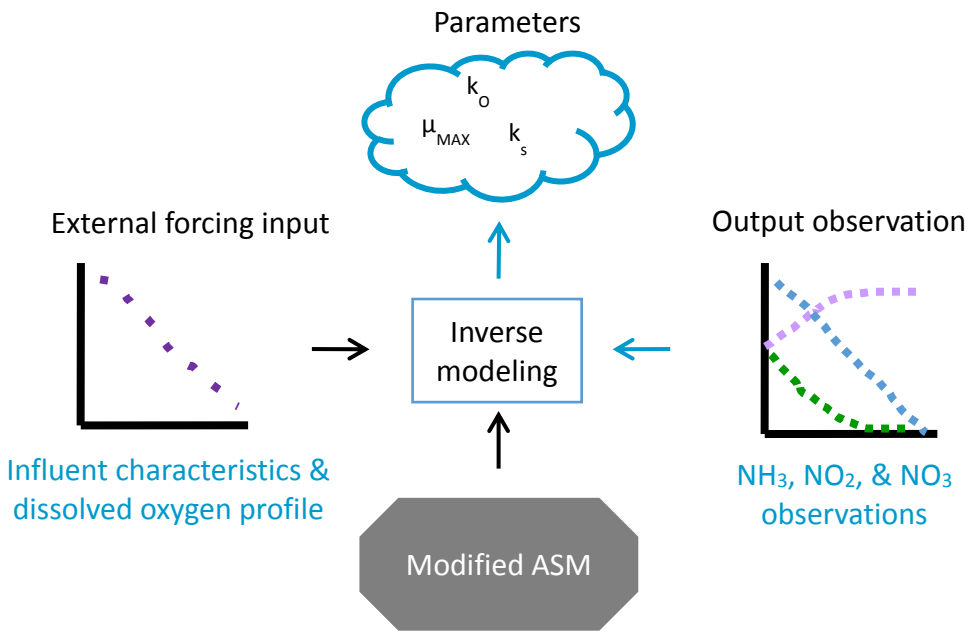


Figure 4-1. Schematic for inverse modeling of mainstream deammonification reactor.

4.2 Materials and Methods

4.2.1 Pilot operation and data collection

The mainstream pilot reactor used in this study was a scaled down version of the current full-scale denitrification reactor at Blue Plains AWWTP. It consisted of ten consecutive, 20-liter, cells with diffusion stones and mixers, followed by a 160 L conical settling tank with a 1 rotation per minute scraper. The pilot had a step feed with returned activated sludge (RAS) into the first cell, one third of the feed entered into the first (anoxic) cell, and the remaining two thirds of feed into the next two anoxic cells (Figure 4-1). The feed rate was 1.5 m^3/day and pulled from a 5,300 L storage tank of the previous day's full-scale plant effluent from either Chemically Enhanced Primary Treatment clarifiers, effluent from secondary treatment (high rate activated sludge system for removal of organics), or a combination of the two. The feed was cycled through UV light disinfection before entering the pilot. Ammonium bicarbonate and sodium bicarbonate were added

to the feed tank to maintain the NH_3 loading rate (~ 20 mg N / L in the feed) and alkalinity (~ 200 mg $CaCO_3$ / L) of the pilot. No other chemicals (like methanol) were added. Wasting was performed by collecting mixed liquor from the last reactor cell daily and pouring it over a vibrating sieve, such that granular AnAOB could be returned to the system. NH_3, NO_2, NO_3, and COD (soluble and total) were measured in duplicate from the influent and effluent from filtered samples to monitor daily performance using Hach vials (HACH GmbH, Düsseldorf, Germany). TSS and VSS were measured from the last cell and clarifier daily using standard methods (APHA, 2005). Once per week samples were measured from influent, effluent, and each reactor cell to get a snapshot of the nutrient profile across the pilot. The more comprehensive weekly profile data formed the basis of observations fed to the model described in the next sections. During the course of the study the following operations were varied: number and arrangement of aerobic cells, dissolved oxygen setpoint, dynamic sludge retention time (SRT) (via wasting rate), source of influent, and AnAOB seeding rate. Dynamic SRT is an estimate of the average time that sludge has resided in the system and depends on rate at which biomass (M) is changing. Biomass changes due to growth, decay, effluent solids escape, as well as the rate at which biomass is wasted (removed) from the system (Takacs and Patry, 2002).

$$\frac{d(SRT)}{dt} = 1 - \frac{SRT}{M}\frac{dM}{dt} \qquad \textit{(Eq. 4-6)}$$

The hydraulic retention time (HRT) on the other hand is only a function of flow rates and reactor volumes. The airflow into the aerobic cells is controlled by LabView program (National Instrument Corporation, Austin, USA) connected to online LDO probes (HACH GmbH, Düsseldorf, Germany). Two of these probes can be seen in cells five and eight of Figure 4-1 below.

These changes reflected different control strategies and/or responses to weather effects. Of the more than eighty weekly profile data sets collected on the pilot, the following are all from a

period of one year with the same feeding and wasting procedure, good measurement replication in the duplicates, reasonable nitrogen mass balances, sufficiently low dissolved oxygen levels in anoxic cells to prevent aerobic activity, and are not within a week of a system failure, such as a pump or mixer failure, so that dynamic SRT calculation is reliable. Table 4-4 below summarizes the main operational conditions for each profile date. The temperature reported is from the mixed liquor in the last cell.

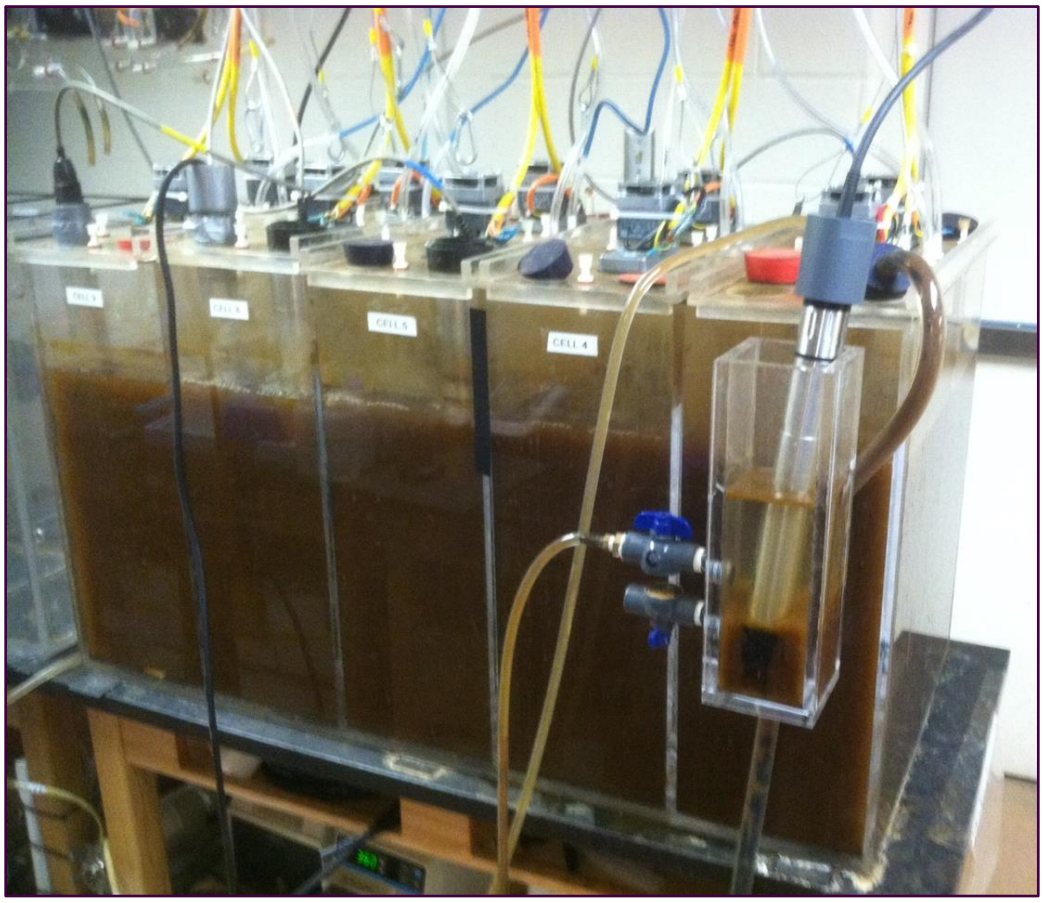


Figure 4-2. Pilot reactor in operation. Probes, mixing motors, and sampling ports can be seen on the lids of cells 1, 4, 5, 8, and 9. The other cells are behind. The activated sludge (brown) obscures the aeration stones and mixing blades, as well as the alternating high and low ports between cells to promote a serpentine flow that prevents short circuiting. The influent sampling box shows how quickly the suspended sludge will settle if not constantly mixed. The settling tank is to the left of the reactor (not shown).

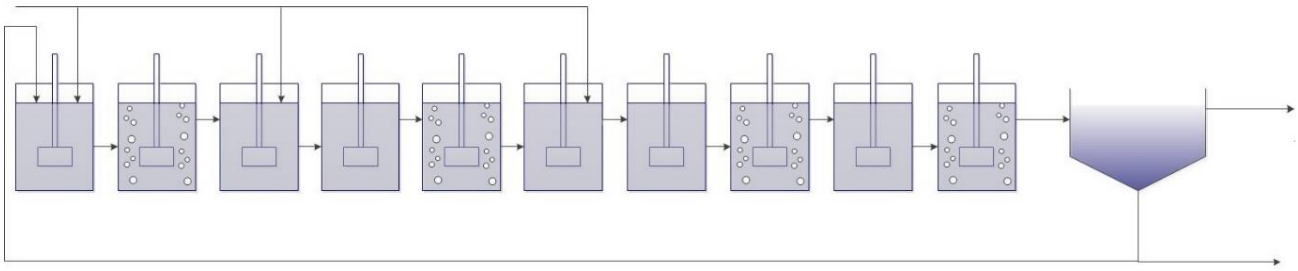


Figure 4-3. Schematic diagram of how the mainstream deammonification pilot was modeled, with feed, Return Activated Sludge (RAS), Wasted Activated Sludge (WAS), and effluent flows labeled. The case shown has cells 2, 5, 8, and10 aerated (bubbles) while other cells were flushed with nitrogen gas

Table 4-4. Operational conditions of pilot. **a)** gives the physical operation inputs while **b)** gives the biological conditions and **c)** gives the dissolved oxygen profiles in mg O_2 / L. The results from the AnAOB activity test near 2015/02/04 were inconclusive.

a)

	Date	*Feed NH_3*	*Feed NO_2*	*Feed NO_3*	*S_S*	*Temp*	*DO setpt*	*aeration pattern*	*Feed pattern*
		mg/L	mg/L	mg/L	mg/l	°C	mg/L	Cell #	Cell #
1	2014-06-24	22.96	2.9	1.06	12.45	30.4	1.5	2,5,8	1,3,6
**2*	2014-07-01	22.61	0.78	0.43	4.1	31.8	1.8	2,5,8	1,3,6
3	2014-08-05	23.82	5.73	1.84	15.6	30.7	1.3	2,5,8	1,3,6
4	2014-10-28	28.72	5.66	2.26	13.8	24.8	3	2,5,8,10	1,3,6
**5*	2015-02-04	23.27	0.77	0.55	14.5	20.6	1	2,5,8,10	1,3,6
**6*	2015-03-24	27.02	0.14	0.43	21.7	21.3	3.5	2,4,6,8,10	1,3
†*7*	2015-05-05	31.50	0.14	0.39	23.5	24.1	3	2,4,6,8,10	1,3,5
**8*	2015-07-08	38.61	4.56	1.4	11.5	25.3	2.8	2,4,6,8	1,3,5

b)

	Date	*VSS in cell 10*	*SRTdyn*	*Performance*	*AnAOB activity*	*AOB activity*	*NOB activity*	*DIC*
		mg/L	d	%	mg N/g vss/d	mg N/g vss/d	mg N/g vss/d	
1	2014-06-24	555	18.8	61	102	141	47	-79
**2*	2014-07-01	750	14.1	79	95	165	74	-163
3	2014-08-05	763	15.7	44	185	176	227	-190
4	2014-10-28	800	9.9	26	19	235	221	127
**5*	2015-02-04	1098	10.4	13	-	216	216	120
**6*	2015-03-24	642	4.3	19	6	276	249	-238
†*7*	2015-05-05	1105	4.7	20	92	226	229	
**8*	2015-07-08	585	4.4	25	61	348	344	-381

c)

	Date	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	2014-06-24	0.01	1.01	0.02	0.02	1.33	0.03	0.01	1.84	0.05	0.01
**2*	2014-07-01	0.02	1.51	0.03	0.03	1.78	0.03	0.03	1.98	0.03	0.03
3	2014-08-05	0.02	1.34	0.03	0.03	1.52	0.03	0.03	1.25	0.06	0.03
4	2014-10-28	0.01	1.03	0.03	0.01	1.68	0.03	0	2.59	0.08	3.74
**5*	2015-02-04	0	0.92	0.01	0.01	1.53	0.01	0.01	0.97	0.02	0.44
**6*	2015-03-24	0.03	3.39	0.08	3.25	0.09	3.51	0.06	0.51	0.1	4.37
†7	2015-05-05	0.02	2.29	0.03	2.4	0.05	3.64	0.08	0.74	0.04	2.09
**8*	2015-07-08	0.06	3.21	0.08	2.92	0.12	2.43	0.1	2.26	0.09	0.05

4.2.2 BIOEST Inverse modeling

A Bayesian parameter estimation algorithm was used for the estimation of activated sludge kinetics parameters. This approach provided joint probability distribution of the parameters and therefore higher-moment quantification of parameter uncertainty (Massoudieh et al., 2013). The inputs for the application are the system flow configuration and operational conditions, the reaction model, prior estimate ranges for parameters, and measured observations of performance. The pilot was simulated as ten consecutive completely stirred tank reactors (CSTRs), a clarifier, and with fixed oxygen concentrations (see Figure 4-2). A clarifier model is needed to calculate the values of $i,r\ C$ for settleable (particulate, X) components. In this study the clarifier was simplified by assuming zero particulate concentration in the effluent and a quasi-steady state approximation according to Takacs et al. (1991). Assuming a conversation of mass, the change in a given component is given by

$$\frac{d(V_k C_{i,k})}{dt} = \omega_{r,k} Q_r C_{i,r} + \omega_k Q C_{i,in} + \left(\sum_{k'=1}^{k-1} (\omega_{r,k'} Q_r + \omega_{k'} Q) \right) C_{i,k-1}$$

$$- \left(\sum_{k'=1}^{k} (\omega_{r,k'} Q_r + \omega_{k'} Q) \right) C_{i,k} + V_k \sum_{l=1}^{N} \varphi_{l,i} R_l + k_{La} V_k (C_{i,sat} - C_{i,k}) + \dot{m}_{i,k}$$

Eq. (4-7)

where V_k is the volume of cell k in L, $C_{i,k}$ is the concentration of component i in cell k in mg / L, Q is the inflow rate to the entire system in L / d, ω_k is the fraction of influent entering cell k, $\omega_{r,k}$ is the fraction of return flow entering cell k (equal to one for the first cell and zero for all others), $C_{i,r}$ is the concentration of component i in the return flow, $C_{i,in}$ is the concentration of component i in the influent, Q_r is the flow rate of the return flow, N is the total number of reactions in the model framework, R_l is the reaction rate of reaction number l, $\varphi_{l,i}$ is the stoichiometric coefficient for reaction l affecting component i, k_{La} is the rate constant for the mass transfer of component i across the liquid-gas interface of bubbles or the atmosphere in units of (1/d) with a saturation concentration $C_{i,sat}$ (this is only used for oxygen in this study) according to Henry's Law and $\dot{m}_{i,k}$ is the mass transfer rate of the influx of component i into cell k (equal to zero for this system as there were no chemical additions). Henry's law describes the mass transfer equilibrium for a dissolvable compound at gaseous and liquid boundary. For a given partial pressure for a compound, p, the corresponding equilibrium or saturation concentration in water is given by

$$C_{sat} = \frac{pRT}{H_e} \qquad \textit{Eq. (4-8)}$$

where H_e is the unitless Henry's constant for that compound, R is the universal gas constant in (atm m^3) / (K mole), and T is the air temperature in K. In *Eq. (4-6)*, the term on the left side is the rate of change in the total mass of the components i in cell k; the first term on the right side is the

mass inflow due to return flow; the second term is the mass inflow of the influent; the third term is the inflow from the preceding cell; the fourth term is the outflow of the components; the fifth term is the production or disappearance of the components due to the reactions; the sixth term is the effect of rate-limited mass transfer (e.g., aeration); and the last term is the direct addition of the components (e.g., addition of carbon source for denitrification).

Previous attempts to model oxygen transfer into the reactors from the diffusors were unsuccessful due to imperfectly sealed cell lids such that oxygen entering/exiting the headspace could not be measured. Thus rather than model the consumption and replenishment of dissolved oxygen across the pilot, the observed values were set as fixed values in the model, and k_La was set to 0. Although the feed into the reactor varies day-to-day due to full fluctuations in the full plant, each profile date was simulated as steady state with constant influent, wasting, and return flow rates. A steady-state assumption was made in order to estimate biomass fractions as the methodology for measuring the concentration of an organism group like AOB is intensive and targets specific bacteria species. Thus unknown strains could be unaccounted for and the estimate would be incomplete. The reaction network used for this study extended the basic ASM1 with two-step nitrification and denitrification (Wyffels et al, 2004) in order to monitor NO_2 consumption via AnAOB. The keen reader will note that the reaction rate expressions in Table 4-3(c) contain many instances of dual substrate limitation as well a third Monod-like term for inhibition. Despite the results of the previous chapter which promoted the use of the minimum function for these expressions, the multiplicative model was used. This is because 1) the pilot study began before the dual limitation experiments were performed and 2) the resultant parameter PDFs are meant to be used by other systems which will likely be modeled using multiplicative rates. Prior estimates of parameters were based upon previous batch tests of the pilot, model calibration of the pilot (Al-

Omari et al., 2015), and ranges reported in literature (Wyffels et al, 2004; Henze et al., 1987). The observations included soluble NH_3, NO_2, and NO_3 measured in each cell and readily biodegradable COD in the influent.

Inverse modeling was used to estimate parameter values that are most likely to reproduce observed pilot data. Simulations were run in three parts. In the first part, several dates were run independently using fixed maximum growth rates from previous pilot model calibration (Al-Omari et al., 2015), generating a set of half-saturation parameter estimates for each date. In part two, the same dates were run independently without maximum growth rates fixed. In the last part, several dates were simulated holistically to see how well one set of parameters fit observations from different dates with different operational conditions. Profiles dates were chosen to represent a variety of operational conditions. Sensitivity analysis and correlation were performed according to (Sharifi et al, 2014).

4.2.3 Model validation and optimization

Validation of the model and parameter estimations were performed by simulating two weeks of pilot performance. The observed daily effluent quality was compared to the predicted effluent quality of the pilot using the median parameter estimates from a steady-state inverse modeling of a profile within a month of the dynamic simulation. These parameters were then used to assess the impact of the following operational conditions on performance: dissolved oxygen concentration and arrangement of aerobic zones, influent composition, solids retention time, and temperature. The configurations that result in the best utilization of the Anammox process and maximized nitrogen removal are optimal. A separate validation for each inverse modeling study

was performed by summing the estimated particulate components (X_{Total}) according to *Eq. (4-3)* and comparing this to the observed VSS measurements.

4.3 Results

The observations collected for this study were analyzed several different ways, producing a rather large body of results. For the reader's benefit the analysis categories have been organized in Table 4-5 according to the number of unfixed (estimated) parameters, the number of time series included, and whether the simulation was steady-state or dynamic. Each simulation was evaluated in 4 ways.

1) How well the MCMC 95% credible interval models fit the observed nitrogen concentrations. These figures are called Modeled vs Measured. Observations outside the shaded model region suggest unreliable parameter values or model structure.
2) Whether the estimated parameter distributions are reasonable compared to previous studies and values reported in literature and have a smooth curve. These figures are called joint probability distributions and are summarized in parameter 95% credible interval floating bar charts. Estimate ranges that far exceed the prior estimate (shaded region) suggest that the model was not able to gain new information on the value of that parameter while jagged distributions suggest the simulation was not able to converge on in the parameter space.
3) Whether the parameters were highly correlated. The correlograms quantify the identifiability of a parameter set.
4) The sensitivity of each parameter with respect to each observation which quantifies relative importance of each parameter.

Table 4-5. Organization of presented results. The 12 estimated parameters include maximum growth rates.

Section	*Name*	*Type of analysis*	*Data considered*	*Number of parameters estimated*
4.3.1	Independent dates	Steady-state inverse	7 single profiles	8
	Independent dates	Steady-state inverse	7 single profiles	12
4.3.3	Holistic analysis	Steady-state inverse	4 profiles simultaneously	8
	Holistic analysis	Steady-state inverse	4 profiles simultaneously	12
4.3.4	Transient analysis	Dynamic inverse	2 weeks of daily observations	12
4.3.5	Optimization	Forward	Simulation – no data	12

4.3.1 Independent Steady state analysis

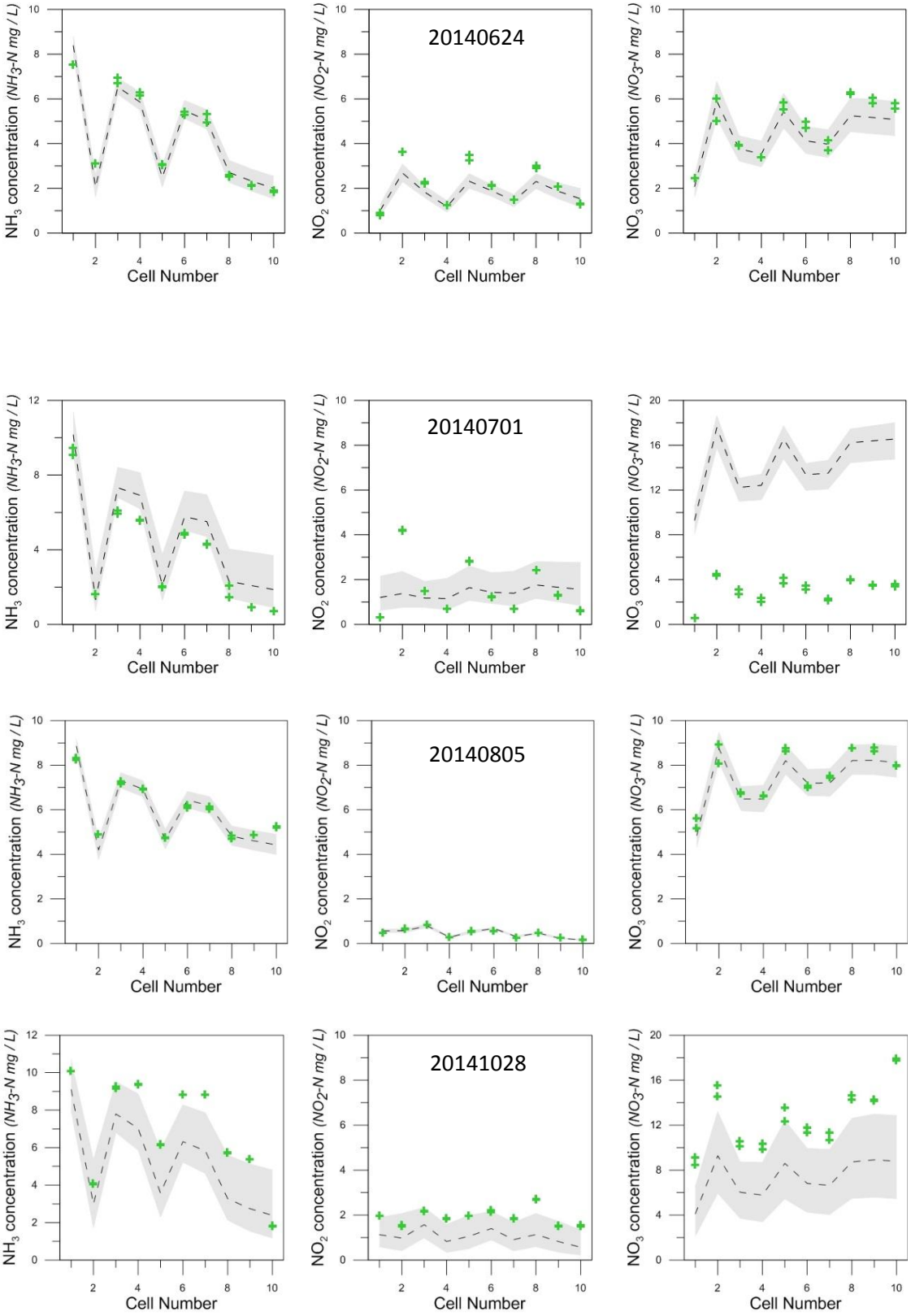
4.3.1.1 Modeled vs measured constituent profiles

For the first set of simulations only eight half saturation parameters (K_S) were estimated (Table 4-3). The remaining (24) were set to values from literature or previous studies, under the assumption that they were not as susceptible to fluctuation or evolution. These include maximum growth rates for each of the four main organism groups. The solid grey bands in Figure 4-3 below represent the 95% credible interval of MCMC realizations for the steady state simulations with the dashed line showing the median value. 1000 of the 500,000 MCMC samples were selected to visualize the range in performance predicted by models using those parameter sets. Two nitrogen measurements for each of the ten cells of the reactor are shown in green for each date.

Every profile presented has influent entering the first cell and the second cell aerated (see operation Table 4-4). This can be visually confirmed by the signs of nitrification occurring between the first and second cell: nitritation where ammonia is consumed and nitratation where nitrate is produced. The biomass and therefore reactions rates decrease in later cells due to dilution by the

subsequent influent feed lines. Thus, the first aerobic cell always has the steepest drop. As the influent had higher concentrations of ammonia and lower concentrations of nitrite and nitrate, one can confirm that feed entered cells one, three, and six on June 24, 2014, because of the increase in ammonia and dilution in nitrate. Whether the nitrite concentration rises or falls between any two cells depends upon the difference between the AOB and NOB reaction rates.

The first measure of simulation success is how well the model fits the measured data, that is to say how narrow the 95% credible interval is and how much of the data is captured within this band. After reasonable model to measurement fits are confirmed, parameter posterior distributions, correlations and sensitivity should be analyzed. Although the predicted models using fixed maximum growth rates and estimated half saturations indices captures the observations well for the first, third, and seventh dates, this is not always the case. Specifically the fits for the second and sixth dates do a poor job of capturing the value or trend of the observations. The predictions for NO_2 were particularly difficult to capture with these parameter sets. NO_2 is a crucial constituent as it is a substrate for NOB, ANAOB, and OHO and plays a big part in control strategies (De Clippeleir et al., 2012; Regmi et al., 2014). These discrepancies between modeled and measured concentrations suggest that inverse modeling was not able to find parameter values for the half-saturation indices that complement the fixed maximum growth rate parameters. Thus the observed maximum rates in this system seem to differ from literature values. To test this, the simulations were performed again allowing the maximum rates to be estimated as well. As will be seen in the next section, estimated maximum growth rates significantly improved the model predictions. For this reason, although all stages of analysis were performed with both fixed and estimated maximum growth rates (parameter estimates, correlation and sensitivity) for both steady state and dynamic, only results from estimated growth rates are presented.


20140624
20140701
20140805
20141028
NH3 concentration (NH3-N mg / L)
NO2 concentration (NO2-N mg / L)
NO3 concentration (NO3-N mg / L)
Cell Number

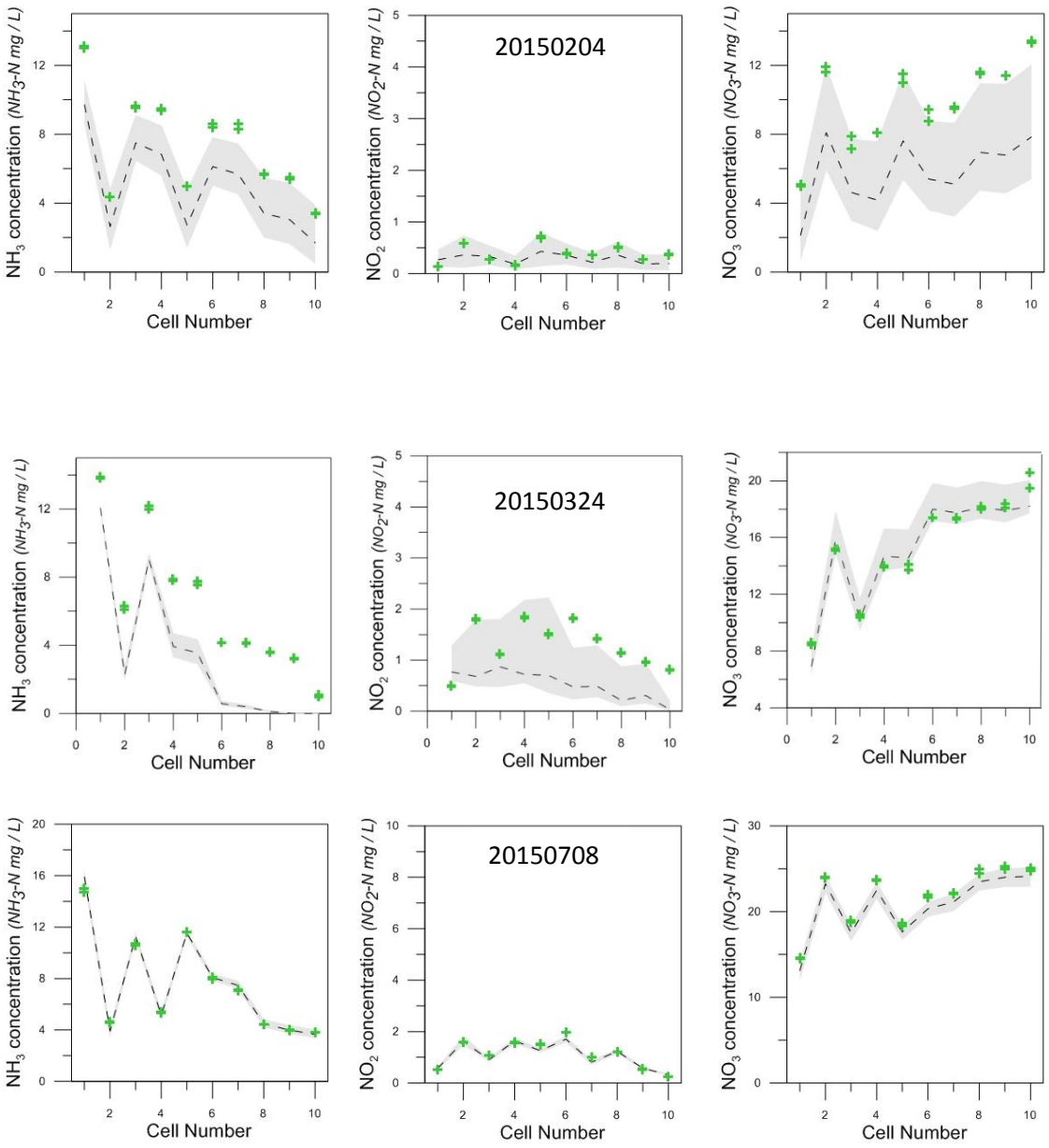
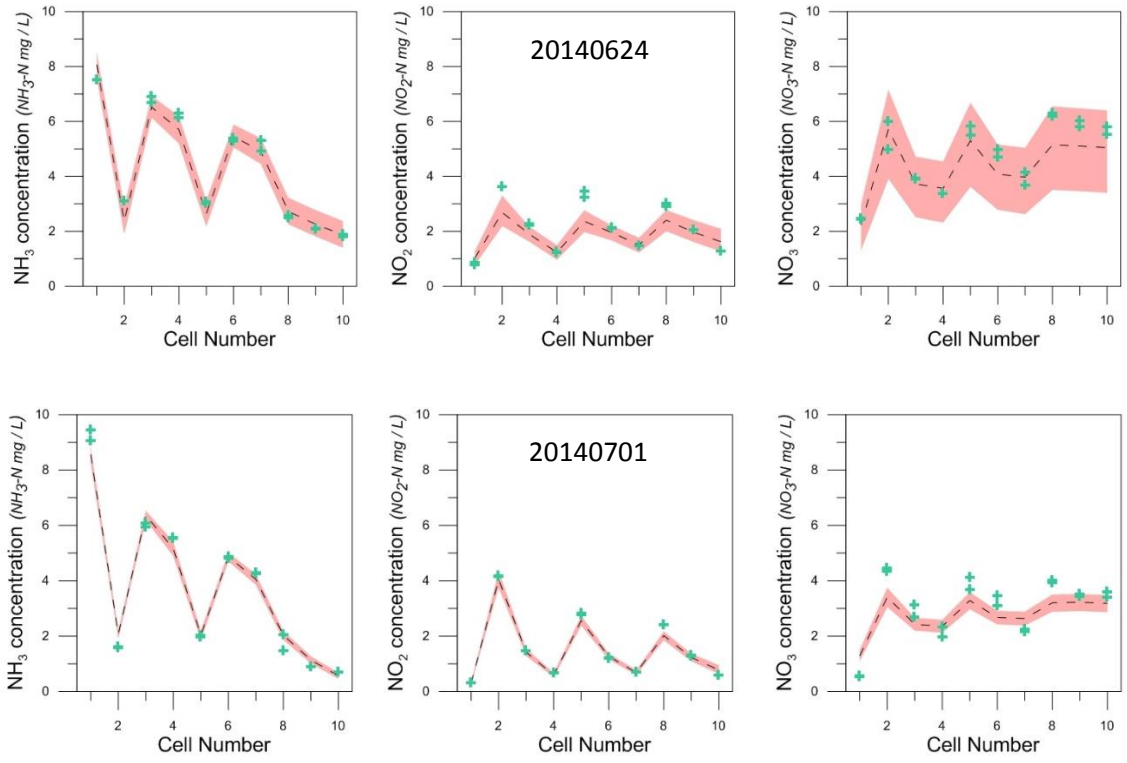
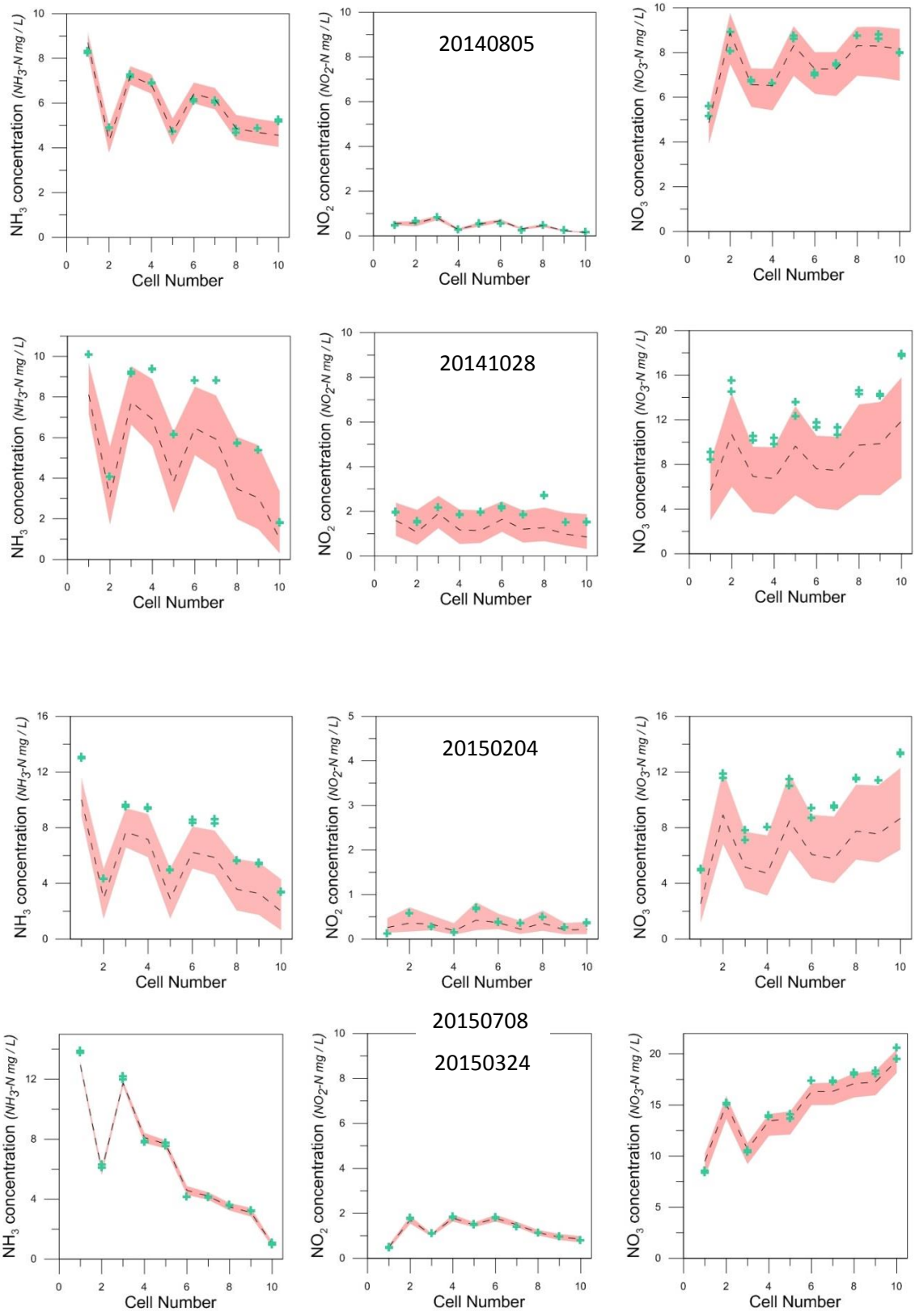


Figure 4-4. Modeled vs measured comparison for nitrogen concentrations in each cell of the mainstream deammonification pilot for 7 independent steady state analyses where the maximum growth rates were ***fixed***. The green crosses represent the observed nitrogen concentrations in each cell on for a given date. Grey regions and dashed lines represent the 95% credible interval and median predictions from inverse modeling.

The predicted models using both estimated maximum growth rates and half saturation indices captured nearly all the observations very well. For these simulations the MCMC prediction

bands are salmon (Figure 4-4) in order to differentiate from the similar-looking grey plots in Figure 4-3. The nitrate predictions are consistently less certain than those for ammonia and nitrite. This is most likely a result of higher measurement error. The nitrate measurement is less accurate due to interference at nitrite concentrations as low as 1 mg / L. Thus samples must be diluted with deionized water to avoid this interference. In response to this, the weighting of the nitrate observations was set lower that nitrite and ammonia in the inverse modeling program by a factor of two.




20140805
20141028
20150204
20150708
20150324
NH3 concentration (NH3-N mg / L)
NO2 concentration (NO2-N mg / L)
NO3 concentration (NO3-N mg / L)
Cell Number

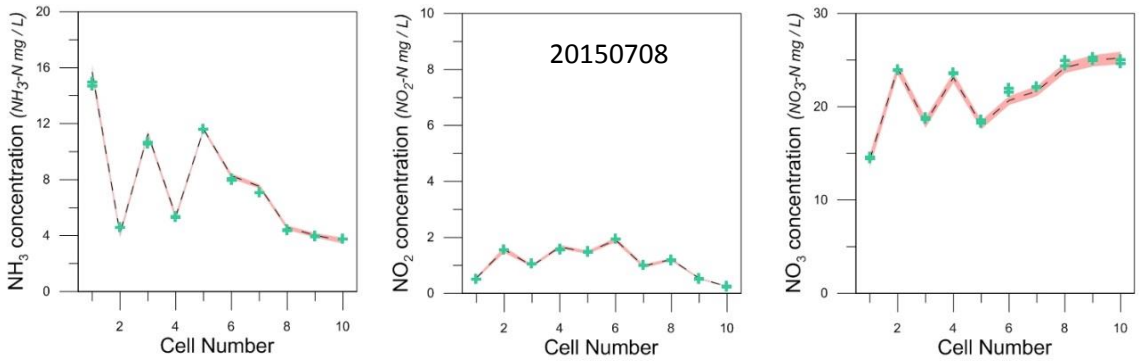


Figure 4-5. Modeled vs measured comparison for nitrogen concentrations in each cell of the mainstream deammonification pilot for 7 independent steady state analyses where the maximum growth rates were ***estimated***. The green crosses represent the observed nitrogen concentrations in each cell on for a given date. Grey regions and dashed lines represent the 95% credible interval and median predictions from inverse modeling.

As a rough validation test, the mean estimated individual biomass group concentrations in the tenth cell were compared with the measured VSS concentrations for each date in cell ten. As not all sources of biomass are included in the model framework, it is simply enough that the estimated total biomass is less than the observed VSS / 1.42 according to *Eq. (4-3)*. Figure 4-5 shows that this requirement was indeed met. It is worth mentioning that inactive or dead biomass is present in VSS, while the biomass estimated by inverse modeling is only the viable, active fraction. Thus a strong correlation between VSS and X_{total} is not expected. Table 4-4 lists the observed maximum activity measured for AOB, NOB, and AnAOB from batch tests of pilot mixed liquor performed in the same week as the profile observations. The specific biomass estimates for NOB, AOB, and AnAOB do positively correlate to the observed activities for each date.

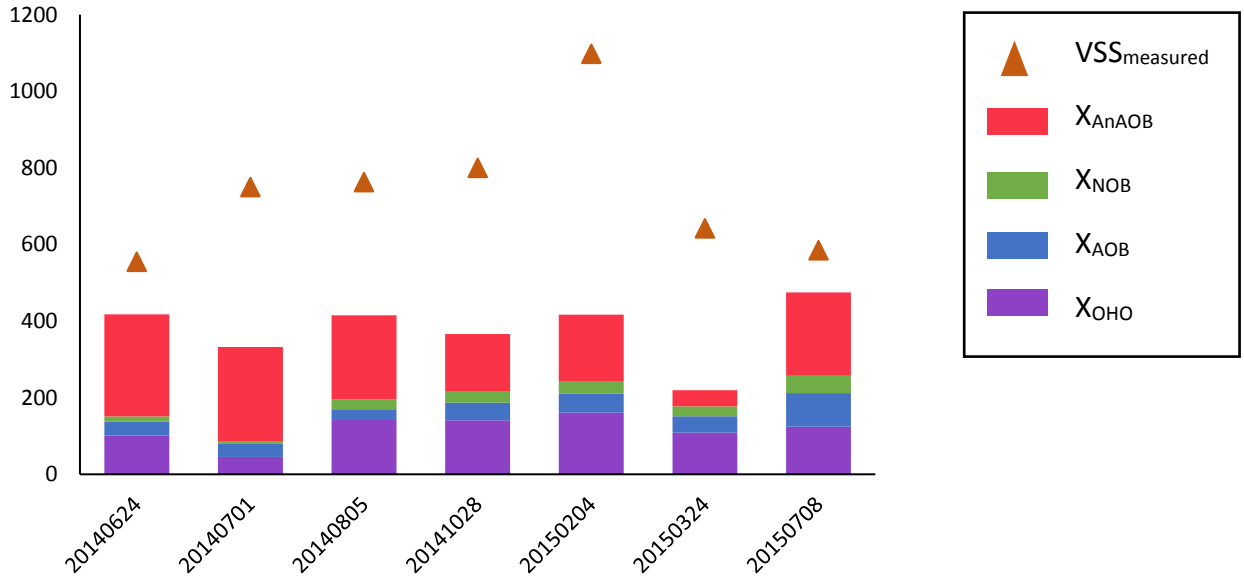


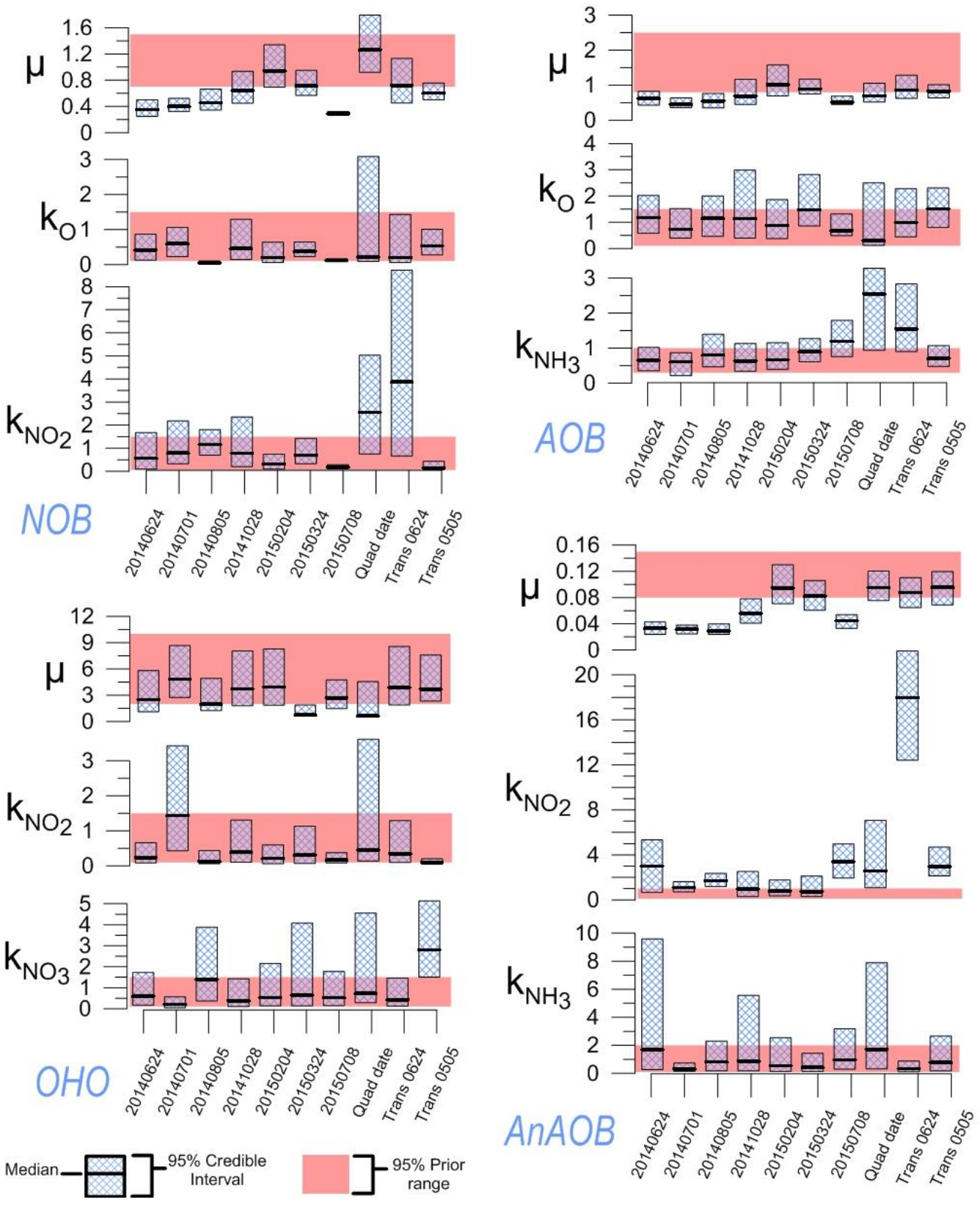
Figure 4-6. VSS validation. Biomass estimates for each group are given for each independent steady state analysis profile with the corresponding measured VSS (divided by 1.42). The vertical axis is in mg COD / L.

4.3.2.2 Parameter Estimation

It is important to note that although this was a mainstream deammonification pilot, there was not always mainstream deammonification achieved in the reactor. This can be seen in 2014/10/28 and 2015/02/04 where the ammonia does not decrease, is not consumed, between cells three and four or six and seven, that is there is no anoxic ammonium removal. One reason for this is poor nitrogen shunt where NOB are consuming nitrite in the aerobic zone before it reaches the anoxic cell. Another reason is a lack of AnAOB biomass. A vibrating sieve is used to separate the AnAOB granular sludge from floccular sludge comprised of the other biomass groups; however, granules may break apart and not be retained on top of the sieve or the seed granules being fed into the system could have been inactive (dead). At least two dates had very low measured AnAOB activity, where activity is reported as the amount of substrate in mg consumed per day per gram

of volatile suspended solids under ideal batch conditions. As in Chapter 3 for the oxygen half saturation index for NOB, if a given process is not observed, the parameters that describe that process cannot be expected to have informative estimates. This is one explanation for why several of the parameter estimates for AnAOB are either well above reasonable ranges or have very large uncertainty compared to literature values and batch experiments. Due to this large range of results in the AnAOB estimates in the floating bar chart in Figure 4-6 A, Figure 4-6 B zooms in on the prior credible range to improve visibility. The NOB parameter estimates for the first two dates should similarly be qualified due to the near washout or removal of NOB via nitrite shunt.

Overall K_o is in the range of 0.5 to 1.5 mg DO / L, which is a bit higher than the prior range. This confirms the message from Chapter 3 on K_o estimation. K_{NH3} values are consistently lower than K_o values except for 2015/07/08 which is probably because this date had a high NH_3 residual; there were no observations at low NH_3 concentrations with which to estimate K_{NH3}. The large uncertainty for $K_{NH3,AnAOB}$ for date 2014/06/24 is likely due to the relatively large deviation in the NH_3 measurement in cell 7, although the lower end of the range is still within the reasonable values. This measurement deviation of only 0.39 mg NH_3-N / L creates noticeable uncertainty in the parameter estimates because each data point is so important. Systems with online nitrogen sensors with high temporal resolution and sensitivity would greatly improve parameter estimates. OHO parameter estimates look fine taking into account that we don't have *Ss* (soluble readily biodegradable substrate) observations in each cell of the pilot, only in the influent. Estimates from holistic (Quad date) and transient simulations will be discussed in sections 4.3.3 and 4.3.4.


μ
k_O
k_{NO2}
NOB
k_{NH3}
AOB
k_{NO3}
OHO
AnAOB
20140624
20140701
20140805
20141028
20150204
20150324
20150708
Quad date
Trans 0624
Trans 0505
Median
95% Credible Interval
95% Prior range

A

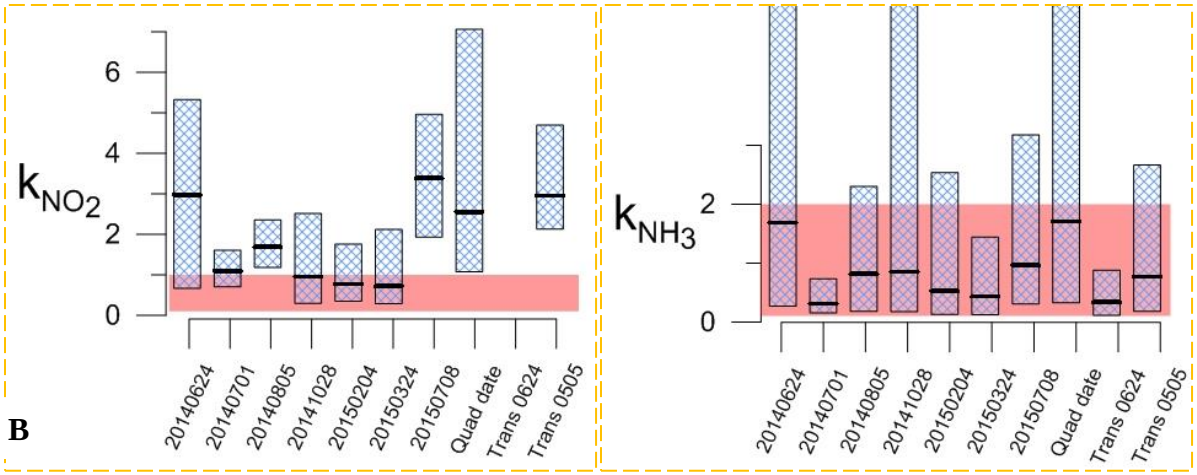


Figure 4-7. A) Parameter estimate comparisons for analyses where maximum growth rates were ***estimated***. Parameters are grouped by their associated biomasses. **B)** Zoom in on floating bar chart for AnAOB parameters. Salmon bands present the prior 95% credible interval behind the blue hatched floating bars which represent the posterior 95% credible interval and median for each parameter for each of seven independent steady state dates, the holistic Quad date, and the two dynamic analyses.

4.3.2.3 Correlation and Sensitivity

The correlation coefficients between parameters quantify how much an increase in one parameter leads to an increase or decrease in the estimate of another parameter. The area of the circles in Figure 4-7 represents the correlations for each independent steady state, the holistic steady state, and the two transient simulations. The larger the purple (white) circle, the stronger the positive (negative) correlation. Values greater than 0.6 are labeled to highlight which simulations had more or fewer highly correlated parameters. The higher the correlation between two parameters, the harder it is to independently identify each parameter from the data set provided.

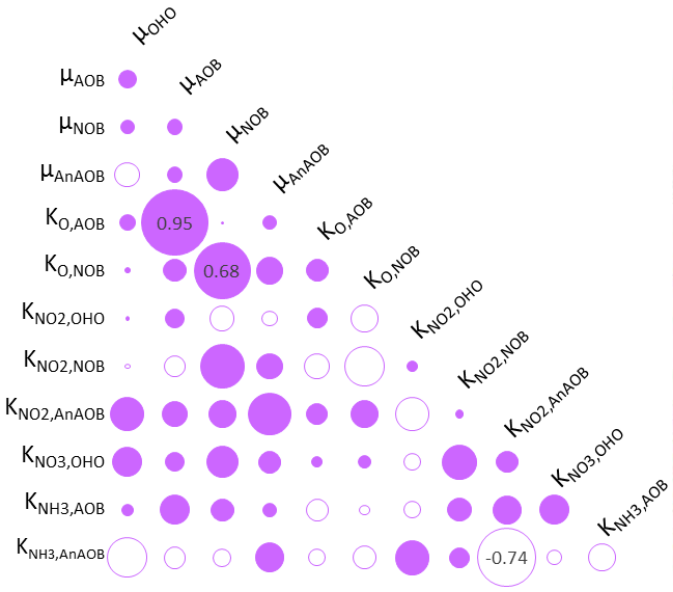
μ_{AOB} and K_{oAOB} are highly correlated across the board (> 0.90) except for transient 2014/06/24. This highlights the importance of estimating these parameters. Special orientations could be configured that have more variety in oxygen concentrations in different cells or these values could be set to batch test values as such (respirometry) tests are relatively inexpensive.

Although as discussed in Chapter 3, these respirometry tests also have identifiability issues. Thus it is not surprising that these same parameter pairs have identifiability issues in situ. It is interesting to note that K_{NO2} and K_{NO3} for OHO do not seem correlated, despite how intertwined the denitrification process rate expressions are.

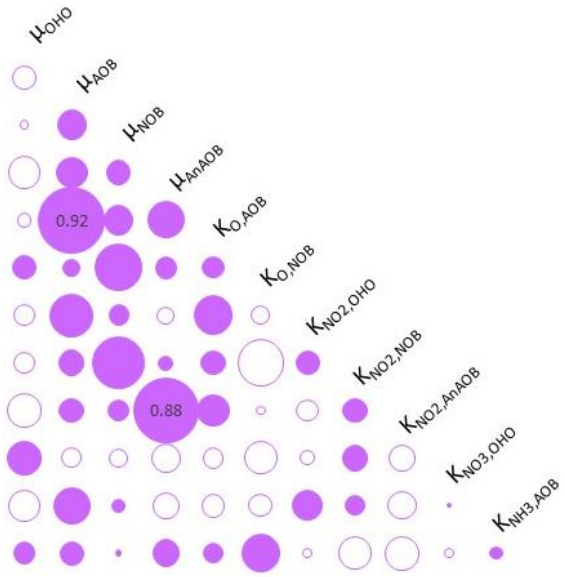
Although the model captures the observations very well for date 2015/07/08, the parameter estimations have very different 95% credible intervals for NOB compared to other dates. The parameter correlations for this simulation are much stronger than any of the other dates, almost twice as many correlation pairs above 0.6. This suggests that the system does not have a unique solution. On the other hand, the credible interval projections for dates 2014/10/28 and 2015/02/04 have much wider uncertainties (while still capturing data values and trends) compared to 2015/07/08. These dates also only have one pair of significantly correlated parameters. The operational conditions of these dates may induced more limiting conditions, thus making the model more sensitive to parameters related to those substrate limitations.

The global sensitivity of each of the twelve estimated parameters on each nitrogen constituent within each cell of the pilot is given in the Figure 4-8 below. It is quite clear that AOB parameters dominate. μ_{AOB} had the highest sensitivity for NH3 for every simulation, closely followed by $K_{O,AOB}$. This is understandable, as aerobic oxidation of ammonia is a prerequisite for any of the other main processes to occur. As can be seen, aeration pattern has a large impact on sensitivity of NH_3 observations while NO_3 observations are rather flat. This is most likely because NH_3 is involved in both aerobic and anoxic processes, while NO_3 is only involved in anoxic and therefore the same parameters in every cell. It is interesting to note that the half saturation indices for NOB often had higher sensitivities than the maximum growth rate for NOB, particularly in 2014/06/24 and 2014/07/01 steady state simulations. This suggests that the operational strategies

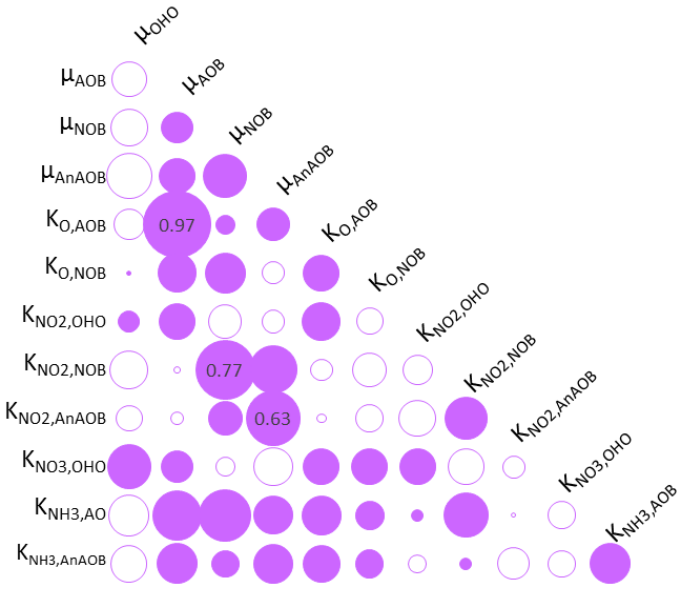
have indeed been successful in suppressing full nitrification; if the rates are influenced by half saturation parameters, the process must be substrate limited.



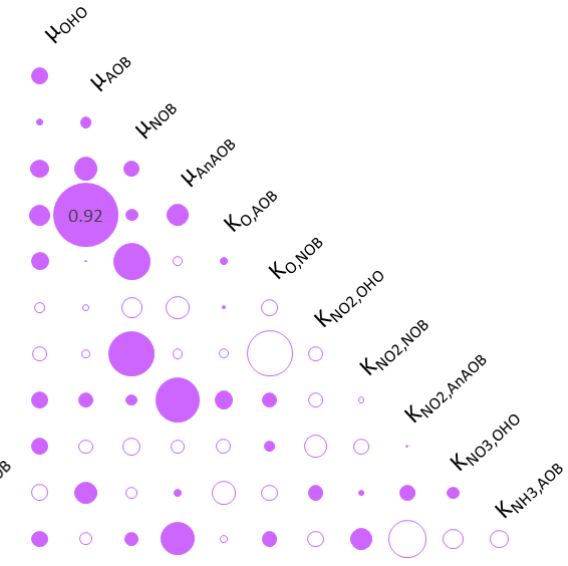
20140624



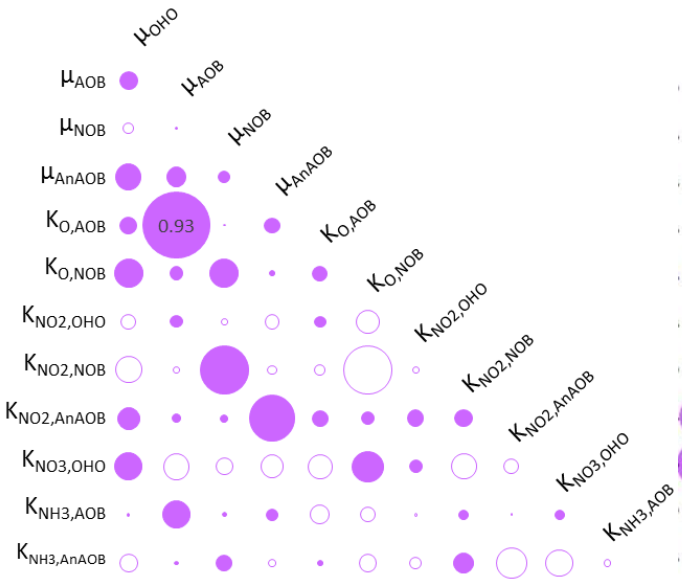
20140701



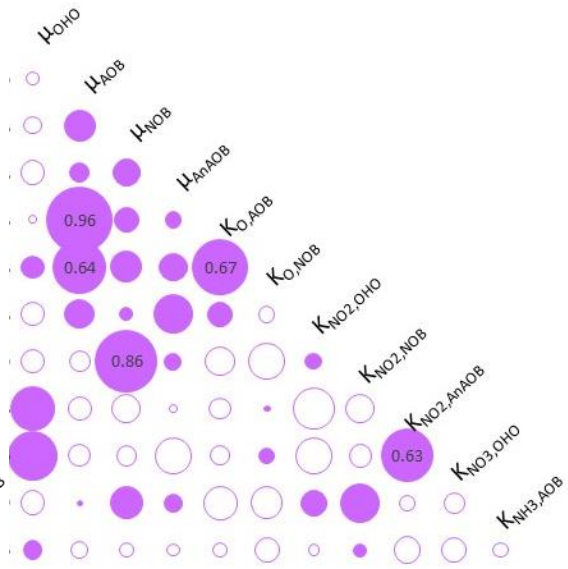
20140805



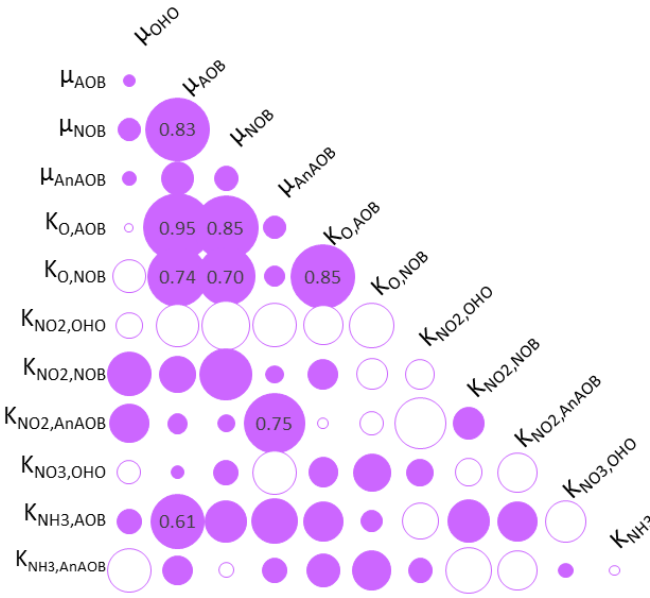
20141028



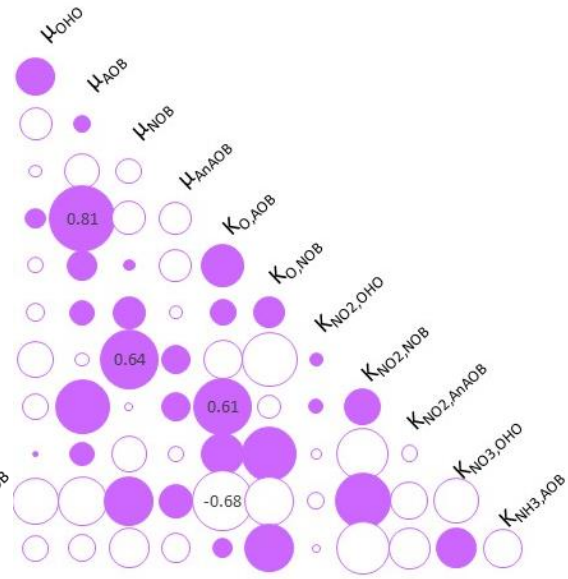
20150204



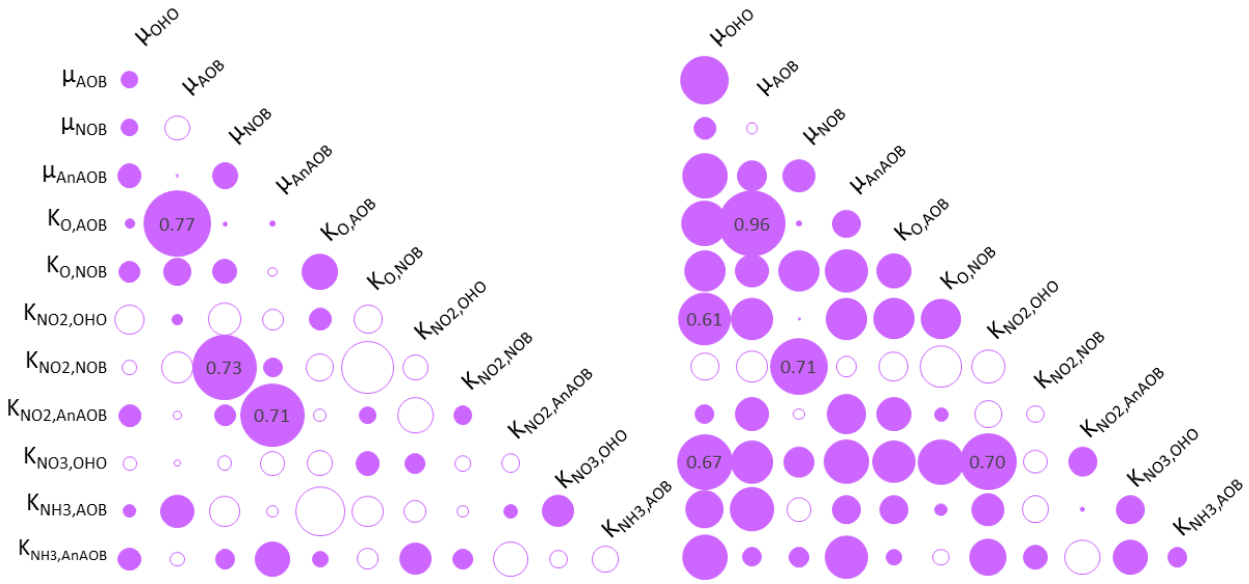
20150324



20150708

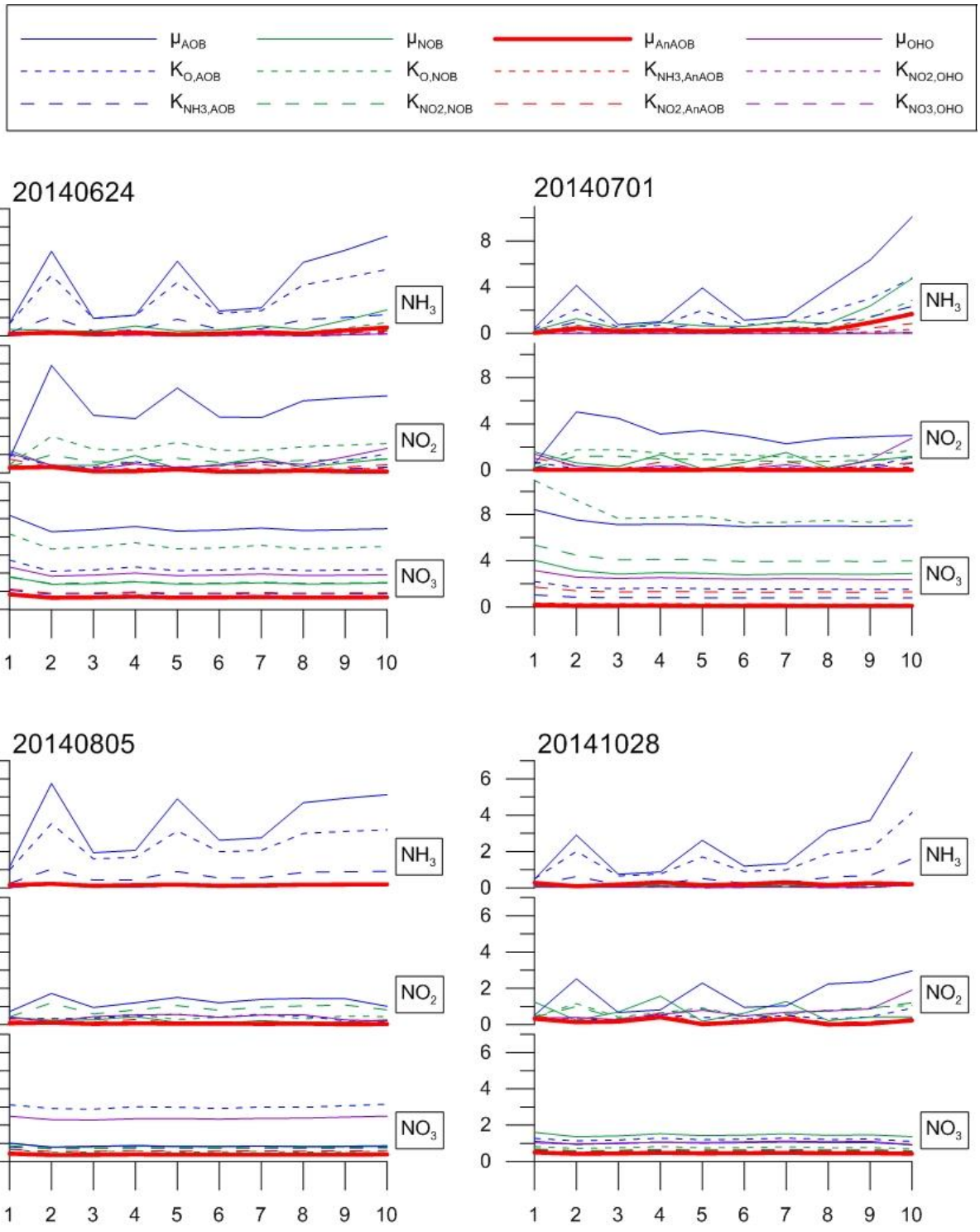


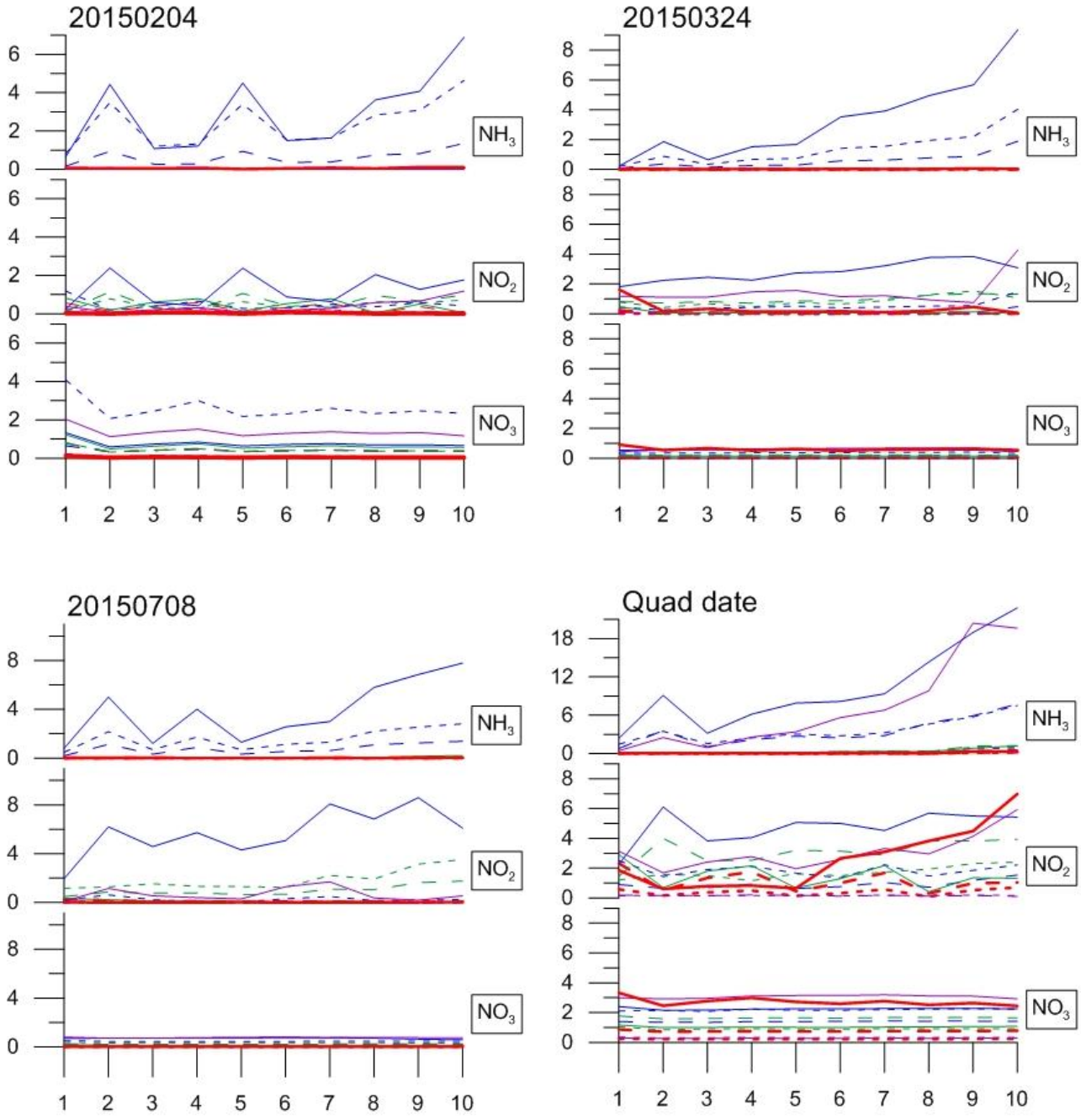
Quad date



Transient 20140624 Transient 20150505

Figure 4-8. Correlograms for each inverse modeling analysis where maximum growth rates were ***estimated***.


μAOB
μNOB
μAnAOB
μOHO
KO,AOB
KO,NOB
KNH3,AnAOB
KNO2,OHO
KNH3,AOB
KNO2,NOB
KNO2,AnAOB
KNO3,OHO
20140624
20140701
20140805
20141028
NH3
NO2
NO3
1 2 3 4 5 6 7 8 9 10


20150204
20150324
20150708
Quad date
NH3
NO2
NO3

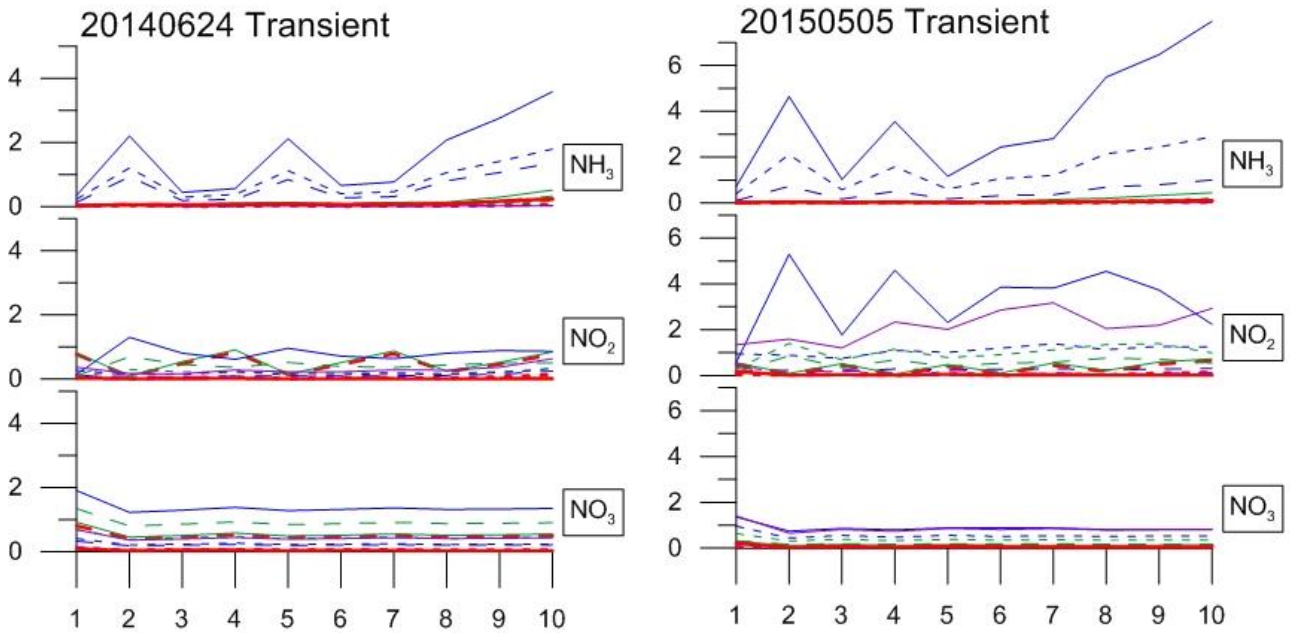


Figure 4-9. Sensitivity plots for each inverse modeling analyses where maximum growth rates were ***estimated***. Parameters are grouped by biomass group by color (blue, purple, red, green) for (AOB, NOB, AnAOB, OHO). Solid and dashed lines represent maximum growth rates and half-saturations indices. The horizontal axes represent the cells of the pilot.

4.3.3 Holistic Steady state analysis

The parameter estimation method for this section considered multiple steady state profile dates simultaneously. Four dates were selected with different operational conditions, at different seasons and spanning the year of data in order test the hypothesis that the parameters under study are constant. Analysis was first performed with fixed maximum growth rates. As expected from the poor fit between predicted concentrations and observations for independent steady state dates, the holistic results were also poor when maximum growth rates were fixed (Figure 4-9). Examination of the resulting posterior PDFs for each parameter suggest that the model struggled to converge, as they do not form smooth bell-shaped-like curves (Figure 4-10).

The analysis for the four dates analyzed simultaneously with estimated maximum growth rates, on the other hand, does seem to have converged and do a somewhat better job of capturing the observations, but the trends are not closely followed (Figures 4-11 and 4-12). Floating bar charts based on the 95% credible interval of the parameters' PDFs are included in Figure 4-6 under the label 'Quad date'. The uncertainty of estimates for all of the biomass groups for Quad date are much higher than for the independent dates. This, along with the fact that the model predictions of ammonia are much wider than for the independent dates, suggest that a single set of parameter PDFs are unable to describe observations from different operational conditions and seasons apart.

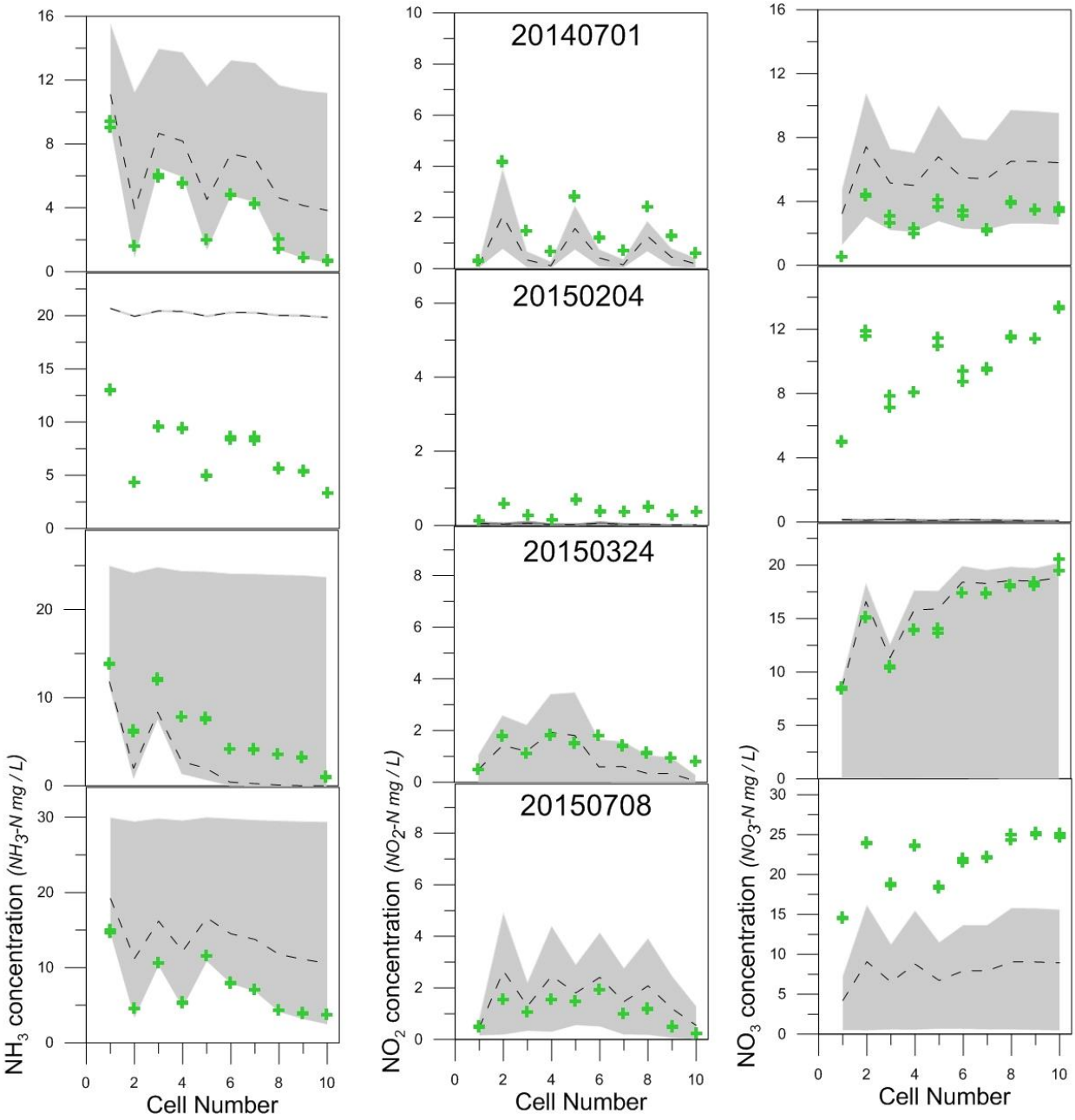


Figure 4-10. Holistic modeled vs measured for four dates analyzed simultaneously with maximum growth rates ***fixed***. The grey regions and dashed lines represent the 95% credible interval and median predictions. Green crosses represent observed nitrogen concentrations.

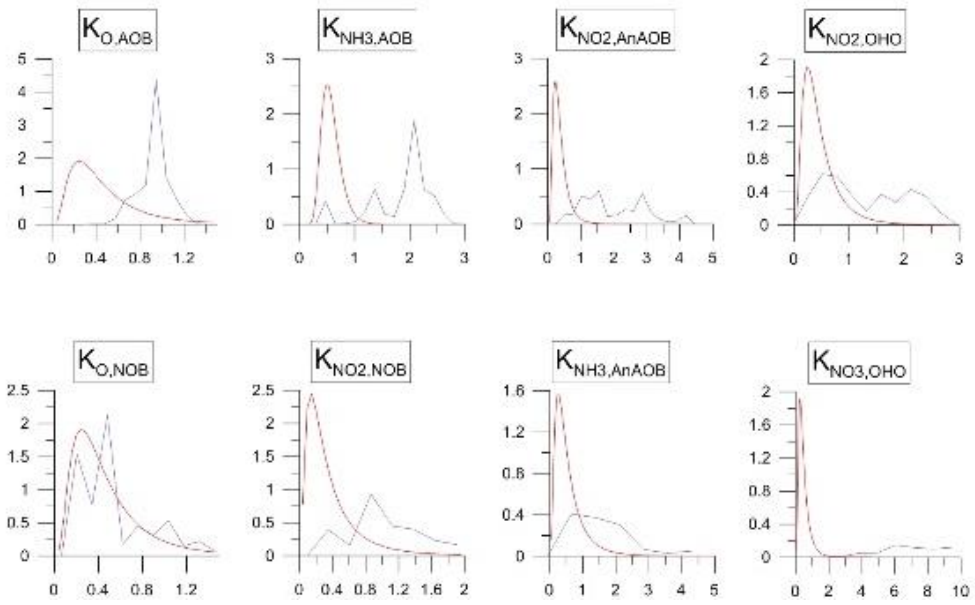


Figure 4-11. Probability distribution functions for estimated kinetics parameters for the holistic, four date steady-state analysis with maximum growth rates ***fixed***. Red and blue lines represent the prior and posterior distributions respectively. Horizontal axes are in the units of the estimated parameter given in Table 4-3. The vertical axes represent the normalized probability of a given parameter value.

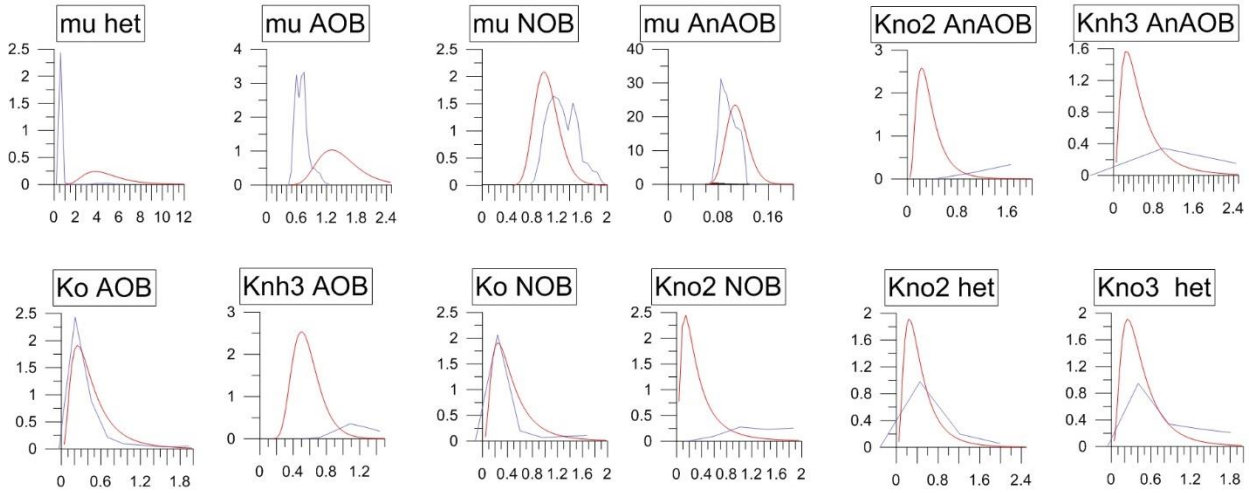


Figure 4-12. Probability distribution functions for estimated kinetics parameters for the holistic, four date steady-state analysis with maximum growth rates ***estimated***. Red and blue lines represent the prior and posterior distributions respectively. Horizontal axes are in the units of the estimated parameter given in Table 4-3. The vertical axes represent the normalized probability of a given parameter value.

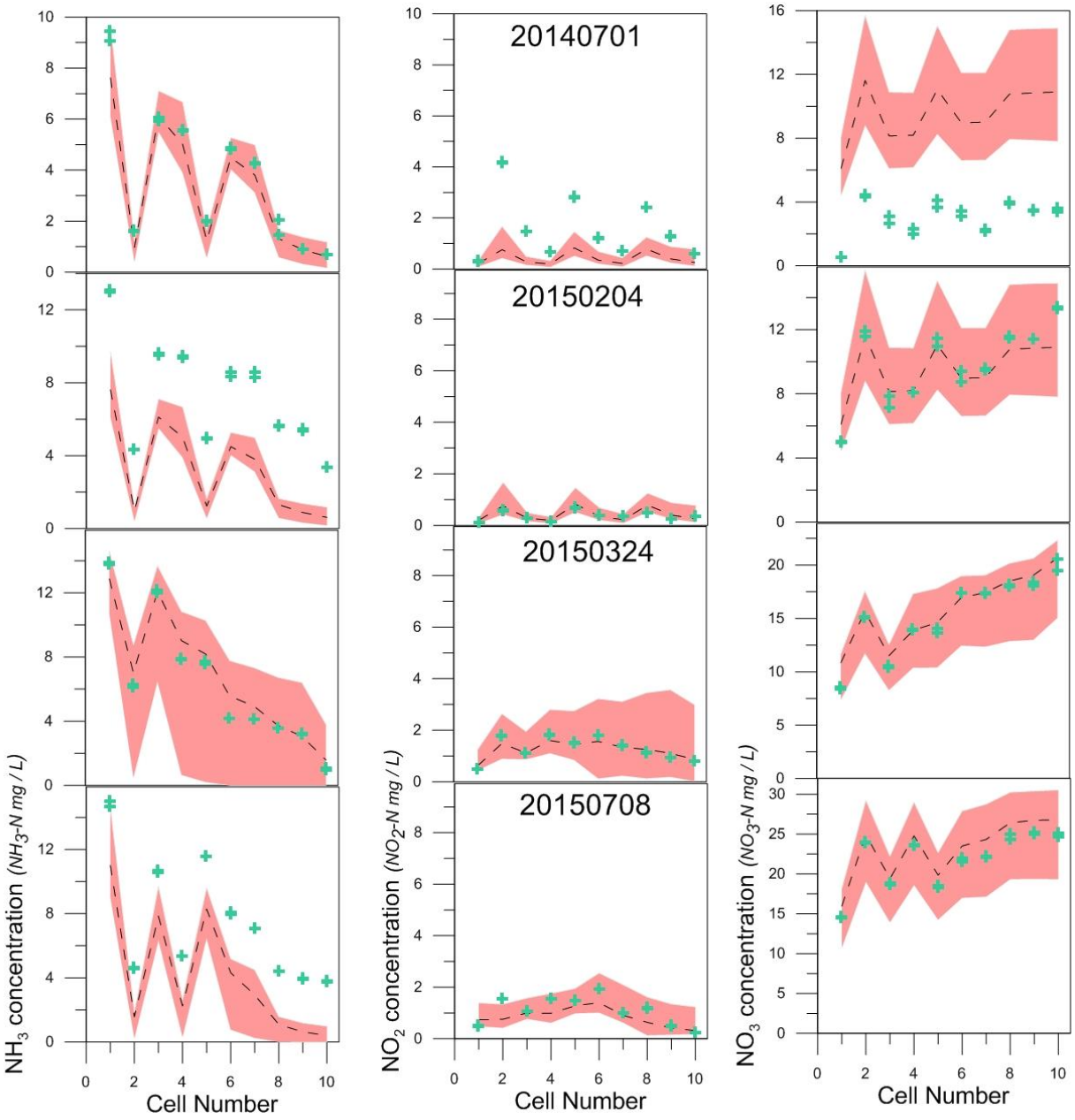


Figure 4-13. Holistic modeled vs measured for four dates analyzed simultaneously with maximum growth rates ***estimated***. The salmon regions and dashed lines represent the 95% credible interval and median predictions. Green crosses represent observed nitrogen concentrations.

4.3.4 Transient Inverse Modeling

Two transient inverse modeling simulations were performed for roughly two weeks each at both the beginning and end of the study period. The first date starts on 2014/06/24, when the pilot was achieving very high total nitrogen removal and exhibiting NOB suppression and high AnAOB activity. The other starts on 2015/05/05, when the reactor was not performing well. The date 2015/05/05 was dropped from the steady-state analysis when it was found to have greater than average mass balance errors for nitrogen, possibly due to an unstable influent. However, as there are many more data points in the transient simulation it is worth including in the analysis.

All things considered, the modeled vs measured plots in Figure 4-13 and 4-14 look rather good for both data sets both for the effluent and across the cells at weekly intervals. The 95% range of MCMC outputs are represented by a green band while the nitrogen observations from cell 10 (the effluent) are shown as purple crosses. Profile comparisons for dates within the data set are shown as well. For the first time series, starting on 2014/06/24, the model was able to capture weekly trends in each for each nitrogen species, but the model would certainly do a better job if there had been observations during the weekends. This impact of poor temporal resolution emphasizes the need for and potential benefits of accurate online nitrogen measurements. The nitrate data is not as well captured.

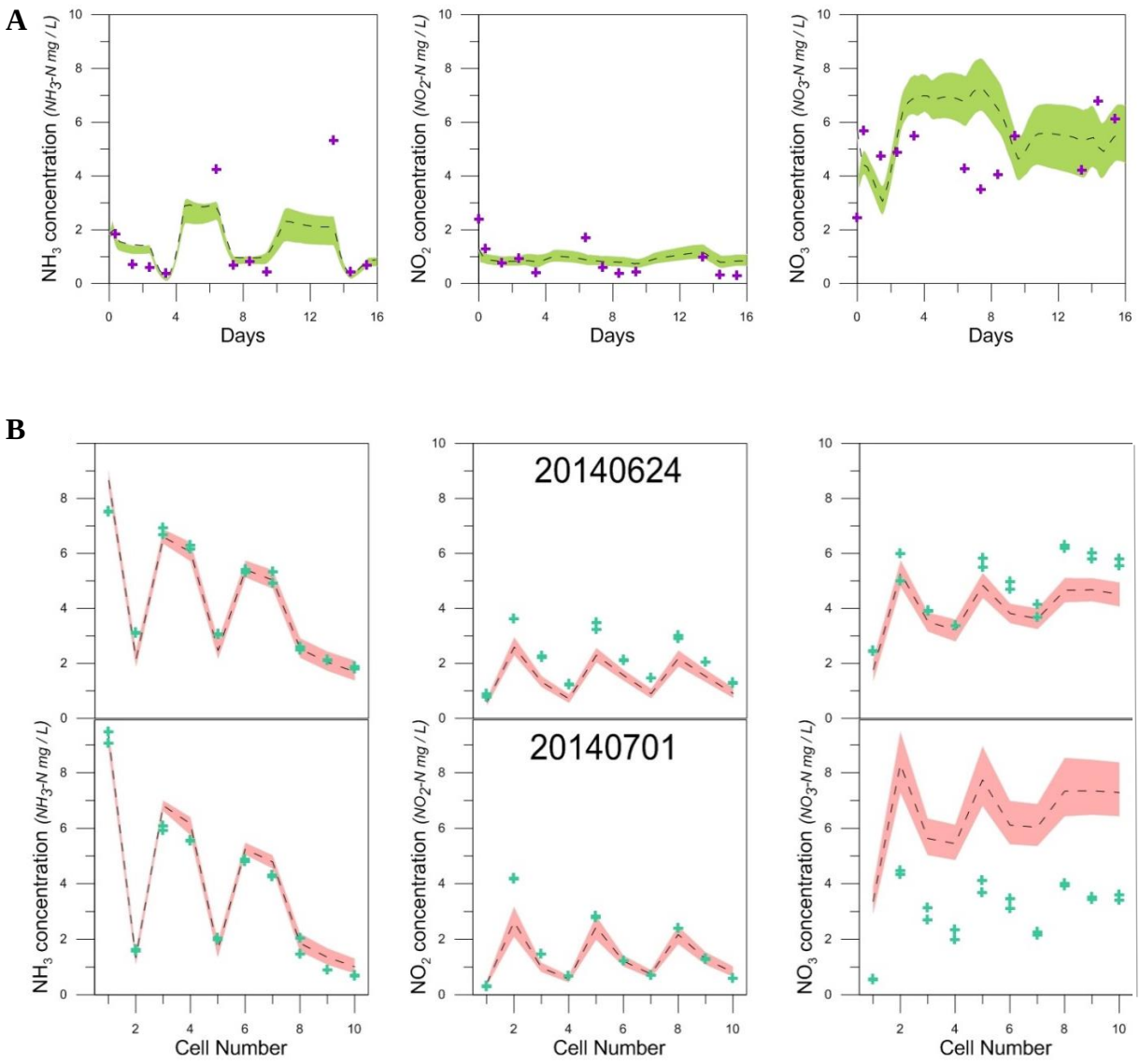


Figure 4-14. Dynamic inverse modeling vs measured for two weeks in summer 2014. **A)** Green regions and dashed lines represent the 95% credible interval and median model predictions for the effluent (cell 10) while purple crosses represent the daily observations. **B)** Salmon regions and dashed lines represent the 95% credible interval and median model predictions of weekly profiles while purple crosses represent the observation from each cell.

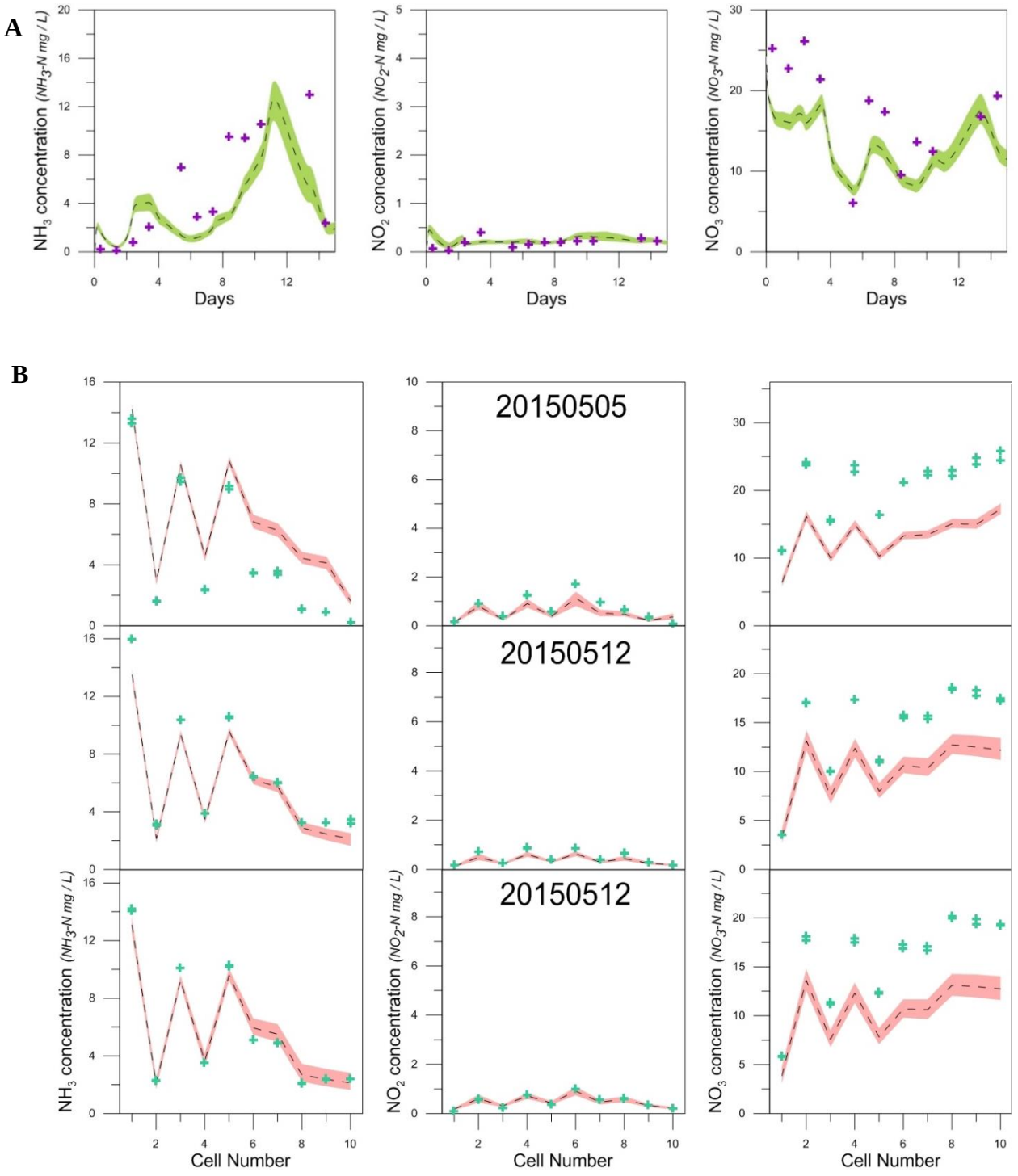


Figure 4-15. Dynamic inverse modeling vs measured for two weeks in spring 2015. **A)** Green regions and dashed lines represent the 95% credible interval and median model predictions for the effluent (cell 10) while purple crosses represent the daily observations. **B)** Salmon regions and dashed lines represent the 95% credible interval and median model predictions of weekly profiles while purple crosses represent the observation from each cell.

4.4 Conclusion

Overall takeaways from this study are that the maximum growth rates for AOB, NOB, OHO, and AnAOB were found to be slightly higher than reported in literature and that parameters for OHO were sensitive to the influent source (primary or secondary effluent). As expected, half-saturation coefficients for AnAOB varied with different sources of seed which had varying granule size and potential diffusion effects. Steady state simulations provided the most reliable estimates for this system that experienced disruptive operational changes daily. For a meso-scale or full scale system, with more stable operation, transient and holistic analysis will likely be more effective.

Further work could include modeling a reactive clarifier. Biomass is still active in the most likely anoxic conditions of the clarifier and several studies have shown the potential impact of these reactions on plant performance (Koch et al., 1999). Currently denitrifying activity that occurs in the clarifier is attributed to the first cell, as it is assumed that the constituent concentrations leaving cell 10 are the same as entering cell 1 with the RAS stream. This is difficult to correct, however, as the clarifier is not a completely mixed system. Further work should also include the minimum dual substrate limitation model in the reaction network and compare the resultant PDFs from such simulations as in Chapter 3.

This research is also the foundation for future work in optimizing control strategies. The future of controls will find optimal operations based on updated estimated parameters, called model predictive control (MPC) algorithms. The model, methodology, and the mainstream deammonification pilot are currently being used for this purpose by the next generation of graduate researchers with the goal of reducing the carbon footprint of wastewater treatment.

5 – Summary

5.1 Main findings

5.1.1 Laurel Hill

This dissertation has validated certain water quality modeling assumptions and questioned others. The Laurel Hill watershed source apportionment study validated the conservation of mass assumption for most of the elements studied (not nitrogen isotopes) based on the strong agreement between modeled and observed downstream profile for individual dates. The main finding of the study was that agricultural runoff accounted for the majority of downstream sediments (up to 80%) and that particles from unpaved roads and forests only contributed significantly during light rain events (< 3 m^3/ s resultant stream flow). Thus as climate change threatens more frequent and more intense rain events, managing agricultural runoff will only become more important.

5.1.2 Batch experiments

This study questioned the assumption that multiple limiting substrate factors should be multiplied together when modeling activated sludge performance. The main finding was that the minimum model best fit the observations for AnAOB at varying limiting conditions for substrates NH_3 and NO_3, and also parameter estimates using the minimum model most closely resembled estimates from single limitation tests. However, the differences in predictions between the three models studied (multiplicative, minimum, and Bertolazzi) were not very large particularly for the NOB experiment. The deviation becomes much more significant as more terms are incorporated, when more biomass groups are incorporated, and/or when multiple substrates are limiting. The simulated operational performance was found to depend strongly on which multiple substrate

limitation model is used. For a mainstream nitrogen system, this is particularly relevant for operational strategies most likely to exhibit multiple limitations for nitrite competitors, such as simultaneous nitrification-denitrification or deammonification at a low DO set point. As we move toward more advanced nutrient treatment strategies we must be cognizant of the impact of model selection on operational control success.

5.1.3 Mainstream pilot

This study validated the method of Bayesian inverse modeling to estimate in situ activated sludge rate kinetics parameters. It also questioned the assumption that these parameters based on biodiverse groups are constant. The main finding was that parameter uncertainty and agreement between model and observations were much worse when multiple dates at different operations and seasons were analyzed holistically rather than individually. Estimated parameters were found to vary with operation changes, although direct correlations between operation and parameter changes were not determined. Transient inverse modeling should promise for parameter estimation but higher temporal resolution than five observations per week are advised. Other important findings are that the parameters with the highest correlation (μ_{max} and K_o for AOB) and therefore worse identifiability are also the parameters for which the model is most sensitive. Thus observations at varying NH_3 and DO concentrations are especially important for estimating AOB parameters in situ and therefore for improving overall model predictions.

5.2 Research implications and looking forward

5.2.1 Source apportionment

The importance of this research lies in its ability to inform policy-makers (Figure 5-1) of best management practices in watershed management. It provides evidence with a certain level of confidence that agricultural runoff is not only occurring, but contributes more than its proportional share of surface area. This is crucial as non-point sources of pollution are difficult to quantify and regulate. The proposed methodology of Bayesian Chemical Mass balance could be used to evaluate the effects of different strategies to manage runoff, either before-and-after a change, or comparing a statistically similar number of watersheds either with or without a given strategy. This methodology was also the first to examine the usefulness of including rare and abundant isotope ratios. It may be informative to apply this methodology with other isotope ratios, such as Cesium of Oxygen. Potential study sites are unfortunately abundant as a large proportion of waterways are impaired in the United States.

5.2.2 Dual Limitation Modeling

The results of this model discrimination are important because of the sheer number of water treatment facilities that use activated sludge models. By switching to the minimum model, facilities will not have to use subjective, plant-specific correction factors to maximum growths rates. Although the models may have similar predictions for only two substrates, the multiplicative model becomes unrealistic for the number of terms often found in full models. The model minimum will also allow operators to identify limiting factors, assuming these results hold for more than two limiting substrates. It would be beneficial to develop and perform experiments for

activated sludge under three or more limiting substrates. There are also compounds that have been found to slow biological growths rates, called inhibitors. Further experiments could examine the interaction between multiple limiting substrates and one or more inhibitors.

Figure 5-1. The National Mall and Tidal Basin just south of the White House (not shown) and leading to the Capitol.

5.2.3 In-Situ Parameter Estimation

The results of the mainstream deammonification pilot were in many ways hindered by observations frequency and artifacts caused by the small scale of the reactor. Despite this, steady-

state analyses for individual dates and short-term dynamic analyses were able to estimate up to twelve key kinetics parameters for a complex nitrogen shortcut strategy. A larger scale system is already in plans with online (automatic and nearly continuous) nitrogen probes. Increased observations could potentially tease out relationships between operation and parameter values as well as be used in advanced control systems that use updated parameter estimates to optimize operational conditions of the system. These controls will enable treatment facilities to minimize energy usage and green-house gas emissions while still maintaining effluent quality (Figure 5-2).

Figure 5-2. View of DC Water effluent into the Potomac River just south of the Tidal Basin.

Bibliography

Al-Omari A., Wett B., Nopens I., De Clippeleir H., Han M., Regmi P., Bott C., and Murthy S. (2015). Model-based evaluation of mechanisms and benefits of mainstream shortcut nitrogen removal processes. *Water Science & Technology*, **71.6**:840-847.

APHA. (2005). Standard Methods for the Examination of Water and Wastewater. APH Association, Washington, DC, USA. American Water Works Association (AWWA) & Wat. Env. Fed. (WEF).

Arnaldos M., Amerlinck Y., Rehman U., Maere T., Van Hoey S., Nuessens W., and Nopens I. (2015). From the affinity constant to the half-saturation index: Understanding conventional modeling concepts in novel wastewater treatment processes. *Water Research*, **70:** 458-470.

Bader F.G. (1978). Analysis of Double-Substrate Limited Growth. *Biotechnology and Bioengineering*, **20**:183-202.

Bae W. and Rittmann B. (1995). A Structured Model of Dual-Limitation Kinetics. *Biotechnology and Bioengineering*, **49**:683-689.

Ballantine D.J., Walling D.E., Collins A.L., Leeks G.J.L. (2009). The content and storage of phosphorus in fine-grained channel bed sediment in contrasting lowland agricultural catchments in the UK. Geoderma, **151**: 141-149.

Bellanger B., Huon S., Velasquez F., Valles V., Girardin C., Mariotti A. (2004). Monitoring soil organic carbon erosion with delta(13)C and delta(15)N on experimental field plots in the Venezuelan Andes. Catena, **58**: 125-150.

Ben-Youssef C. and Vazquez-Rodriguez. (2011). Model-based design of different fedbatch strategies for phenol degradation in acclimatized activated sludge cultures. *Bioresource Technology*, **102:** 3740-3747.

Bertolazzi E.(2005). A Combination Formula of Michaelis-Menten-Monod Type. *Computers and Mathematics with Applications* **50**:201-215.

Beyenal H., Chen S., Lewandowski Z. (2003). The double substrate growth kinetics of *Pseudonomas aeruginosa. Enzyme and Microbial Technology,* **32**: 92-98.

Brockmann D., Rosenwinkel K., Morgenroth E. (2008). Practical Identifiability of Biokinetic Parameters of a Model Describing Two-Step Nitrification in Biofilms. *Biotechnology and Bioengineering,* **101**(3): 497-514.

Bungay H. (1994). Growth rate expressions for two substrates one of which is inhibitory. *Journal of Biotechnology* **34:** 97-100.

Button D.K. (1993). Nutrient-limited microbial growth kinetics: overview and recent advances. *Antonie vuan Leeuwenhoek,* **63:** 225-235.

Carpenter S.R., Caraco N.F., Correll D.L., Howarth R.W., Sharpley A.N., Smith V.H. (1998). Nonpoint Pollution of Surface Waters with Phosphorus and Nitrogen. *Ecological Applications* **8**(3): 559-568.

Chapra S. (1997). *Surface Water-Quality Modeling.* Boston, Mass: McGraw-Hill.

Collins A.L., Walling D.E., Leeks G.J.L. (1997). Source type ascription for fluvial suspended sediment based on a quantitative composite fingerprinting technique. Catena, **29**: 1-27.

Collins A.L., Walling D.E., Webb L., King P. (2010). Apportioning catchment scale sediment sources using a modified composite fingerprinting technique incorporating property weightings and prior information. Geoderma, **155**: 249-261.

Collins A.L., Zhang Y., Walling D.E., Grenfell S.E., Smith P., Grischeff J., Locke A., Sweetapple A., Brogden D. (2012). Quantifying fine-grained sediment sources in the

River Axe catchment, southwest England: application of a Monte Carlo numerical modelling framework incorporating local and genetic algorithm optimisation. Hydrological Processes, **26**: 1962-1983.

Collins A.L., Zhang Y.S., Duethmann D., Walling D.E., Black K.S. (2013). Using a novel tracing-tracking framework to source fine-grained sediment loss to watercourses at sub-catchment scale. Hydrological Processes, **27**: 959-974.

Cook J. (2011). Basic properties of the soft maximum. UT MD Anderson Cancer Center Department of Biostatistics Working Paper Series. Working Paper 70.

Dapena-Mora A., Van Hulle S., Campos J.L., Mendez R., Vanrolleghem P., Jetten M. (2004). Enrichment of Anammox biomass from municipal activated sludge: experimental and modelling results. *J Chem Technol Biotechnol*, **79:**1421-1428.

Davis C.M., Fox J.F. (2009). Sediment Fingerprinting: Review of the Method and Future Improvements for Allocating Nonpoint Source Pollution. Journal of Environmental Engineering-Asce, **135**: 490-504.

DC Water Approved FY 2017 Budget. (2015). https://www.dcwater.com/investor_relations/budget_information.cfm

De Clippeleir H., Han M., Al-Omari A., Bott C., Wett B., and Murthy S. (2014). From nitrite shunt to mainstream deammonification strategy: pilot-scale demonstration. *Proceedings of the Water Environment Federation*, **6**: 4244-4248.

De Clippeleir H., Vlaeminck S., Wilde F., Daeninck K., Mosquera M., Boeckx P., Verstraete W., and Boon N. (2013). One-stage partial nitritation/anammoc at 15 °C on pretreated sewage: feasibility demonstration at lab-scale. *Appl Microbiol Biotechnol*, **97**(23): 10199-10210.

De Kruek M., Picioreanu C., Hosseimi M., Xavier J., van Loosdrecht M. (2006). *Biotechnology and Bioengineering*, **97**(4):801-815.

Dochain, D., Vanrolleghem, P., and Van Daele, M. (1995). Structural identifiability of biokinetic models of activated sludge respiration. *Water Research*, **29**(11): 2571-2578.

Egli T. and Zinn M. (2003). The concept of multiple-nutrient-limited growth of microorganisms and its application in biotechnological processes. *Biotechnology Advances*, **22:**35-43.

Fox J.F., Davis C.M., Martin D.K. (2010). Sediment Source Assessment in a Lowland Watershed Using Nitrogen Stable Isotopes1. JAWRA Journal of the American Water Resources Association, **46**: 1192-1204.

Fox J.F., Papanicolaou A.N.(2007). The use of carbon and nitrogen isotopes to study watershed erosion processes. Journal of the American Water Resources Association, **43**: 1047-1064.

Fox J.F., Papanicolaou AN. (2008). An un-mixing model to study watershed erosion processes. Advances in Water Resources, **31**: 96-108.

Fox J.F. (2009). Identification of Sediment Sources in Forested Watersheds With Surface Coal Mining Disturbance Using Carbon and Nitrogen Isotopes1. Journal of the American Water Resources Association, **45**: 1273-1289.

Franks S.W., Rowan J.S. (2000). Multi-parameter fingerprinting of sediment sources: Uncertainty estimation and tracer selection.

Gamerman D., Hedibert F.L. (2006). Markov Chain Monte Carlo - Stochastic Simulation for Bayesian Inference. Second Edition Edn., Chapman & Hall/CRC.

Gellis A., Noe G. (2013). Sediment source analysis in the Linganore Creek watershed, Maryland, USA, using the sediment fingerprinting approach: 2008 to 2010. Journal of Soils and Sediments, **13**: 1735-1753.

Gellis A.C., Walling D.E. (2011). Sediment-source fingerprinting (tracing) and sediment budgets as tools in targeting river and watershed restoration programs: Simon A., Bennett S.., Castro J.M., eds., Stream Restoration in Dynamic Fluvial Systems: Scientific Approaches, Analyses, and Tools. In: American Geophysical Union Monograph Series 194, pp: 263-291.

Geweke J., Tanizaki H. (2001). Bayesian estimation of state-space models using the Metropolis–Hastings algorithm within Gibbs sampling. Computational Statistics & Data Analysis, **37**: 151-170.

Geweke J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bayesian Statistics 4, Bernardo J.M., Berger J., Dawid A.P., Smith J.F.M.(eds.) Oxford University Press, pp: 169-193.

Gonzo E., Wuertz S., Rajal V. (2014). The Continuum Heterogeneous Biofilm Model with Multiple Limiting Substrate Monod Kinetics. *Biotechnology and Bioengineering*, **111** (11): 2252-2264.

Guzman G., Quinton J.N., Nearing M.A., Mabit L., Gomez J.A. (2013). Sediment tracers in water erosion studies: current approaches and challenges. Journal of Soils and Sediments, **13**: 816-833.

Haas C. (1994). Unified Kinetic Treatment for Growth on Dual Nutrients. *Biotechnology and Bioengineering* **44:**154-164.

Han M., De Clippeleir H., Al-Omari A., Wett B., Vlaeminck S., Bott C., and Murthy S. (2016). Impact of carbon to nitrogen ratio and aeration regime on mainstream deammonification. Water Science and Technology, **74**(2): 375-384.

Harris D., Horwáth W.R., van Kessel C. (2001). Acid fumigation of soils to remove carbonates prior to total organic carbon or CARBON-13 isotopic analysis. Soil Sci. Soc. Am. J., **65**: 1853-1856.

Henze M., Gujer W., Mino T., van Loosdrecht M., (2000). Activated sludge models ASM1, ASM2, ASM2d and ASM3. IWA Publishing: London.

Jaisi D.P., Blake R.E., Kukkadapu R.K. (2010). Fractionation of oxygen isotopes in phosphate during its interactions with iron oxides. Geochimica Et Cosmochimica Acta, **74**: 1309-1319.

Jaisi D.P., Kukkadapu R.K., Stout L.M., Varga T., Blake R.E. (2011). Biotic and Abiotic Pathways of Phosphorus Cycling in Minerals and Sediments: Insights from Oxygen Isotope Ratios in Phosphate. Environmental Science & Technology, **45**: 6254-6261.

Kass R., Raftery A. (1995). Bayes factors. *Journal of the American Statistical Association*, **90**(430): 773-795.

Keats A., Cheng M.T., Yee E., Lien F.S. (2009). Bayesian treatment of a chemical mass balance receptor model with multiplicative error structure. Atmos. Environ., **43**: 510-519.

Kelley D.W., Nater E.A. (2000). Source apportionment of lake bed sediments to watersheds in an Upper Mississippi basin using a chemical mass balance method. Catena, **41**: 277-292.

Koch G., Pianta R., Krebs P., and Siegrist H. (1999). Potential of the denitrification and solids removal in the rectangular clarifier. *Water research*, **33**(2):309-318.

Koiter A.J., Owens P.N., Petticrew E.L., Lobb D.A. (2013). The behavioural characteristics of sediment properties and their implications for sediment fingerprinting as an approach for identifying sediment sources in river basins. Earth-Science Reviews, **125**: 24-42.

Kovarova-Kovar K., Egli T. (1998). Growth Kinetics of Suspended Microbial Cells: From Single-Substrate-Controlled Growth to Mixed-Substrate Kinetics. *Microbiology and Molecular Biology Reviews*, **62**:646-666.

Lee A.L., Ataal M.M., Schuler M.L. (1984). Double-Substrate-Limited Growth of *Escherichia coli*. *Biotechnology and Bioengineering*, **26**:1398-1401.

Legovic T., Cruzado A. (1997). A model of phytoplankton growth on multiple nutrients based on the Michaelis-Menten-Monod uptake, Droop's growth and Liebig's Law. *Ecological Modelling*, **99:**19-31.

Lemaire R., Zhao H., Thomson C., Christensson M., Piveteau S., Hemmingsen S., Veuillet F., Zozor P., Ochoa J. (2014). Mainstream Deammonification with ANITA^{TM}Mox Process. *Proceedings of the Water Environment Federation*.

Liu H., Wang S., Hsu H. (1992). Double-substrate interactive model. *Chemical Engineering Science,* **48**(11):2169-2172.

Lotti T., Kleerebezem R., Hu Z., Kartal B., Jetten M.S.M., van Loosdrecht M.C.M. (2014). Simultaneous partial nitritation and anammox at low temperature with granular sludge. *Water Research* **66**:111-121.

Makinia J. (2010). *Mathematical Modelling and Computer Simulations of Activated Sludge Systems.*London: IWA Publishing.

Mankad T., Bungay H.R. (1988). Model for microbial growth with more than one limiting nutrient. *Journal of Biotechnology,* **7:**161-166.

Martínez-Carreras N., Udelhoven T., Krein A., Gallart F., Iffly J.F., Ziebel J., Hoffmann L., Pfister L., Walling D.E. (2010). The use of sediment colour measured by diffuse reflectance spectrometry to determine sediment sources: Application to the Attert River catchment (Luxembourg). Journal of Hydrology, **382**: 49-63.

Massoudieh A., Bombardelli F.A., Ginn T.R. (2010). A biogeochemical model of contaminant fate and transport in river waters and sediments. Journal of Contaminant Hydrology, **112**: 103-117.

Massoudieh A., Gellis A., Banks W.S., Wieczorek M.E. (2013). Suspended sediment source apportionment in Chesapeake Bay watershed using Bayesian chemical mass balance receptor modeling. Hydrological Processes, **27**: 3363-3374.

Massoudieh A., Kayhanian M. (2013). A Bayesian Chemical Mass Balance Method for Surface Water Contaminant Source Apportionment. Journal of Environmental Engineering, **139**: 250-260.

Massoudieh A., Lu N, Liang X, Nguyen T., Ginn T. (2013). Bayesian Process-indentification in bacteria transport in porous media. *Journal of Contaminant Hydrology,* **153**: 78-91.

Massucci A., Guimera R., Amaral L., Sales-Pardo M. (2015). Scaling and optimal strategy: Two principles determining microbial growth in complex media. *Physical Review E,* **91:** 062703.

Maynad D., Leonard K., Drake J., Hall D., Crother T., Bradford M. (2015). Modelling the multidimensional niche by linking functional traits to competitive performance. *Proc. R. Soc. B.* **282:** 20150516.

MeGee R.D. III, Drake J.F., Fredrickson A.G., Tsuchiya H.M., (1972). Studies in intermicrobial symbiosis. Saccharomyces cerevisiae and Lactobacillus casei, Can. J. Microbiol. **18:** 1733.

Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H., Teller E. (1953). Equations of State Calculations by Fast Computing Machines. Journal of Chemical Physics, **21**: 1087-1092.

Minella J.P.G., Walling D.E., Merten G.H. (2008). Combining sediment source tracing techniques with traditional monitoring to assess the impact of improved land management on catchment sediment yields. Journal of Hydrology, **348**: 546-563.

Monballiu A., Desmidt E., Ghyselbrecht K., De Clippeleir H., van Hulle S.W.H., Verstraete W., and Meesschaert B. (2013). Enrichment of anaerobic ammonium oxidizing (Anammox) bacteria from OLAND and conventional sludge: features and limitations. Separation and Purification Technology, 104: 130–137.

Monod J. (1949). The growth of bacterial cultures. *Annual Review of Microbiology*, **3:** 371-394.

Mukundan R., Radcliffe D.E., Ritchie J.C., Risse L.M., McKinley R.A.(2010). Sediment Fingerprinting to Determine the Source of Suspended Sediment in a Southern Piedmont Stream. Journal of Environmental Quality, **39**: 1328-1337.

Mukundan R., Walling D.E., Gellis A.C., Slattery M.C., Radcliffe D.E. (2012). Sediment Source Fingerprinting: Transforming From a Research Tool to a Management Tool1. JAWRA Journal of the American Water Resources Association, **48**: 1241-1257.

Mulholland P.J., Hall R.O., Sobota D.J., Findlay S.E.G., Burgin A.J. (2009). Nitrate removal in stream ecosystems measured by N-15 addition experiments: Denitrification. Limnol. Oceanogr., **54**.

Neethling, J. B., De Clippeleir, H. (Eds.). (2015). *Shortcut Nitrogen Removal – Nitrite Shunt and Deammonification.* Alexandria, VA: Water Environment Federation.

Novotny V., and H. Olem. (1994). Water quality: prevention, identification and management of diffuse pollution. *Van Nostrand Reinhold*, New York, New York, USA.

Odencrantz J. (1992). Comparison of minimum-rate and multiplicative monod biodegradation kinetic models applied to *in situ* bioremediation. *Proceedings of the Fifth International Conference on Solving Groundwater Problems with Models*. Dallas, Texas, 479-493.

Okpokwasili G., Nweke C. (2005). Microbial growth and substrate utilization kinetics. *African Journal of Biotechnology*, **5**(4): 305-317.

Oshiki M., Shimokawa M., Fujii N., Satoh H., Okabe S. (2011). Physiological characteristics of the anaerobic ammonium-oxidizing bacterium, *Candidatus* Brocasia sinica. *Microbiology*, **157**: 1706-1713.

Owens P., Xu Z. (2011). Recent advances and future directions in soils and sediments research. J Soils Sediments, **11**: 875-888.

Paerl H and Huisman J. (2008). Blooms like it Hot. *Science*, **320:** 57-58.

Palmer M.J., Douglas G.B. (2008). A Bayesian statistical model for end member analysis of sediment geochemistry, incorporating spatial dependences. Journal of the Royal Statistical Society: Series C (Applied Statistics), **57**: 313-327.

Panikov N. (1979). Steady State Growth Kinetics of *Chlorella vulgaris* under Double Substrate (Urea and Phosphate) Limitation. *J. Chem. Tech. Biotechnol.*, **29:** 442-450.

Pavlou S., Fredrickson A. (1989). Growth of Microbial Populations in Nonminimal Media: Some Considerations for Modeling. *Biotechnology and Bioengineering*, **34**(7): 971-989.

Pusker R., Miller M., Holgate B., Bunce R., Park H., Chandran K., Wett B., Murthy S., and Bott C. (2014). Control of aeration, aerobic SRT and COD input for mainstream deammonification. *Water Research*, **57**: 162-171.

Sharifi S., Murthy S., Takacs I., Massoudieh A. (2014). Probabilistic Parameter Estimation of Activated Sludge Processes using Markov Chain Monte Carlo. *Water Research* **50**: 254-266.

Shaw A., Takacs I., Pagilla K.R., Murthy S. (2013). A new approach to assess the dependency of extant half-saturation coefficients on the maximum process rates and estimate intrinsic coefficients. *Water Research*, **47**: 5986-5994.

Sibbesen E., Sharpley A.N. (1997). Setting and justifying upper critical limits for phosphorus in soils.

Siegrist H., Reithaar S., Koch G., and Lais P. (1998). Nitrogen loss in a nitrifying contractor treating ammonium-rich wastewater without organic carbon. *Water Science and Technology* **38**(8): 241-248.

Sloto R.A., Gellis A.C., Galeone D.G. (2012). Total Nitrogen and Suspended-Sediment Loads and Identification of Suspended-Sediment Sources in the Laurel Hill Creek Watershed, Somerset County, Pennsylvania, Water Years 2010–11. USGS.

Smith L., McCarty P., Kitanidis P. (1998). Spreadsheet Method for Evaluation of Biochemical Reaction Rate Coefficients and Their Uncertainties by Weighted Nonlinear Least-Squares Analysis of the Integrated Monod Equation. *Appl Env Microbio.* **64-**6: 2044-250.

Steininger G., Dosso S., Dettmer J., Holland C. (2014). Model selection using Bayesian samples: An introduction to the deviance information criterion. *J. Acoust. Soc. Am*. **136**: 2085.

Stinson B., Murthy S., Bott C., Wett B., Al-Omari A., Bowden G., Mokhyerie Y., De Clippeleir H. (2013). Roadmap Toward Energy & Chemical Optimization Through The Use of Mainstream Deammonification at Enhanced Nutrient Removal Facilities. *Proceedings of the Water Environment Federation, Nutrient*, **30**:702-731.

Strous M., Kuenen J., Jetten M. (1999). Key Physiology of Anaerobic Ammonium Oxidation. *Applied and Environmental Microbiology*, **65** (7): 3248-3250.

Strous M., Heijnen J.J., Kuene J.G., Jetten M.S.M. (1998). The Sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microoganisms. *Appl. Microbiol. Biotechnol,* **50**: 589-596.

Stutter M.I., Langan S.J., Demars B.O.L. (2007). River sediments provide a link between catchment pressures and ecological status in a mixed land use Scottish River system. Water Res., **41**: 2803-2815.

Taggart J.E. (2002). Analytical methods for chemical analysis of geologic and other materials. In: U.S. Geological Survey: U.S. Geological Survey Open-File Report 02-0223.

Takacs, I.; Patry, G. G. (2004). The Dynamic Solids Residence Time. Proceedings of the 3rd World Water Congress, Melbourne, Australia, April 7–12; IWA: London, United Kingdom.

Tanase C., Chirvase A., Ungureanu C., Caramihai M., Muntean O. (2011). Study of Double-substrate limited growth of *Pseudomonas aeruginosa* in aerobic bioprocess. *Revue Roumaine de Chimie*, **56** (12): 1147-1153.

Van den Berg H. (1998). Multiple nutrient limitation in unicellulars: reconstructing Liebig's Law. *Mathematical Biosciences*, **149:**1-22.

Vanrolleghem P., Van Daele M., and Dochain D. (1995). Practical identifiability of a biokinetic model of activated sludge respiration. *Water research*, **29**(11): 2561-2570.

Von Liebig J. (1840). Die organische Chemie in ihrer Anwendung auf Agricultur und Physiologie (Organic Chemistry in its Application to Agriculture and Physiology). Braunschweig, Germany: Vieweg.

Walling D.E., Owens P.N., Leeks G.J.L. (1999). Fingerprinting suspended sediment sources in the catchment of the River Ouse, Yorkshire, UK. Hydrological Processes, **13**: 955-975.

Walling D.E.(2005). Tracing suspended sediment sources in catchments and river systems. Science of the Total Environment, **344**: 159-184.

Wett B., Jimenez A., Takacs I., Murthy S., Bratby J.B., Holm N.C., Ronner-Holm S.G.E. (2011). Models for nitrification process design: one or two AOB populations? *Water Science & Technology* **64** (3): 568-578.

Wett B., Podmirseg S., Gomez-Brandon M., Hell M., Nyhuis G., Bott C., Murthy S. (2015). Expanding DEMON Sidestream Deammonification Technology towards Mainstream Application. *Water Environment Research*, **87** (12): 2084-2089.

Wyffels S., Van Hulle S.W.H., Boeckx P., Volcke E.I.P., Van Cleemput O., Vanrolleghem P.A., and Verstraete W. (2004). Modeling and simulation of oxygen-limited partial nitritation in a membrane-assisted bioreactor. Biotechnology and Bioengineering, 86 (5): 532–542.

Zhang Q., De Clippeleir H., Su C, Al-Omari A., Wett B., Vlaeminck S., Murthy S. (2016). Deammonification for Digester supernatant pretreated with thermal hydrolysis: overcoming inhibition through process optimization. *Appl. Microbiol. Biotechnol.* Epub ahead of print.

Zinn M., Witholt B., Egli T. (2004). Dual nutrient limited growth: models, experimental observations, and applications. *Journal of Biotechnology*, **113**:263-279.