

Groundwater Modeling for the Management of Agricultural Impacts on Surface Water Quality

A Dissertation

Presented to
the Faculty of the School of Engineering and Applied Sciences

University of Virginia

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Philosophy
in
Civil and Environmental Engineering

by
Wesley Owen Zell

August 2015

APPROVAL SHEET

This dissertation is submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Civil Engineering

Wesley Owen Zell
(Author)

This dissertation has been read and approved by the examining committee:

Teresa B. Culver, Ph.D. (Advisor)

Jonathan L. Goodall, Ph.D. (Committee Chairperson)

James A. Smith, Ph.D.

Todd M. Scanlon, Ph.D.

Ward E. Sanford, Ph.D.

Accepted for the School of Engineering and Applied Science:



Craig H. Benson, Dean, School of Engineering and Applied Science

August
2015

Abstract

While significant efforts have been made to reduce agricultural nitrogen export to the Chesapeake Bay, loads from many agriculturally-intensive areas continue to increase due to the discharge of legacy nitrates from surficial aquifers beneath agricultural systems. In order to improve understanding of the relationship between land surface management improvements and water quality response, the Chesapeake Bay Program (CBP) has endorsed the use of targeted small watershed studies as a means of better understanding the impact of agricultural best management practices (BMPs) on surface water quality. This dissertation supports the CBP objectives for small watershed studies by characterizing groundwater flow and nitrate transport and removal in one of the targeted watersheds. The dissertation uses an agriculturally-intensive catchment on the Maryland Eastern Shore to examine a series of questions related to understanding and managing nitrogen in agricultural systems, with particular focus on (i) modeling nitrogen transport subject to groundwater lag times and (ii) calibrating the simulation tools used for that modeling. The collective purpose of the four studies included in the dissertation is to improve our ability to calibrate and use numerical groundwater simulation tools, with the aim of better modeling the impact of agricultural intensification and de-intensification on water quality, and thus support management by improving our ability to interpret signs of decline and improvement in receiving streams.

The first study compares the range and character of different catchment-scale simulation models that result when an automated calibration routine is driven by various combinations of spatially- and temporally-distributed CFC, SF₆, ³H, and ³He information. While researchers commonly use groundwater-age information to calibrate subsurface flow

and transport models, rarely is there a rich enough age-tracer dataset to investigate the question of what number, spatial distribution, and measurement uncertainty of calibration targets are needed in order to adequately characterize a site. In this investigation 238 environmental-tracer measurements were used individually and in combination to estimate recharge rates, hydraulic conductivities, and effective porosity for a three-dimensional groundwater flow and transport model of our study site. The various combinations of spatially- and temporally-distributed CFC, SF₆, and ³H information used to drive the automated inverse modeling routine resulted in a range of catchment-scale simulation models and associated parameter uncertainty bounds. The study demonstrates that while tracer data can provide necessary supplemental information for the calibration of flow and transport models, the use of data from a single tracer or from a small tracer set may be insufficient to fully interpret the information content of the tracers.

In the second study we use the calibrated groundwater flow and transport model to resolve the key components of the nitrogen budget for the targeted watershed. While subsurface nitrate transport and catchment removal processes have been widely investigated, there have been few fully distributed, three-dimensional modeling studies of nitrate transport and removal in catchments with nitrogen removal rates that are highly spatially-variable, as is the case with our study site. We link the re-constructed time-variable land surface loadings to time-variable stream responses in two subcatchments that have similar land use histories but highly disparate nitrate export signatures, and we estimate the impact of soil denitrification and in-stream nitrogen removal as well as the potential influence of retarded nitrate transport. We show that in spite of spatial and temporal uncertainty in loading, multiple calibration scenarios agree that in-stream nitrate removal efficiencies

vary significantly between the two sub-catchments, with one stream removing 60-70% of incoming nitrogen loads and the contrasting stream removing only 15-30%.

While we use a steady state representation of the flow system in the descriptions of environmental tracer and nitrate transport in the first two studies, in some situations variability in base-flow age may impact in-stream solute concentrations. In the third study, we examine the impact of time-variable hydrologic forcing – such as that due to seasonal changes in precipitation and evapotranspiration – on the age of base-flow discharge. We develop a method for simulating the transient delivery of base-flow age from subsurface to receiving stream as a function of seasonal changes in hydrology and aquifer storage, and we apply the method to a variety of synthetic two-dimensional (2D) aquifers as well as to the study site. We found that the timing of maximum base-flow age relative to the timing of minimum base-flow discharge varied with both the hydraulic conductivity field and the annually averaged recharge, which determines the system mean age. The two assumptions of (i) an aquifer in which ages are vertically well-mixed and (ii) an aquifer in which ages are strongly stratified provide two end-members for estimation of how the base-flow age might respond to seasonal changes in recharge and base-flow, and the simulations in this study found that the change that occurs in real systems is somewhere in between. For the cases that we investigated, apparent ages inferred from SF₆ measurements while assuming piston-flow transport assumptions for the SF₆ were biased young, with biases especially pronounced with layered hydrogeology in which discharge consists of shallow surficial flow mixed with a contrasting regime of much older water. For one of the subcatchments in our Maryland study site we found that seasonal changes in recharge may only result in changes in base-flow age of 3 to 4 years, but that SF₆ apparent ages based on piston-flow

assumptions of tracer transport may underestimate the mean base-flow age by 60% and more closely resemble the median system age.

In the fourth and final study, we consider the impact of assumptions about age transport on the automated calibration of groundwater flow and transport models. As illustrated in the calibration of the flow and transport model used throughout this dissertation, parameters such as porosity and dispersivity must often be estimated through model calibration against data describing groundwater age. In such cases, the groundwater age observed at a point in the subsurface is often assumed to be a function of purely hydraulic processes. In this study we use the automated calibration of several synthetic aquifers to investigate the impact of that assumption on the resulting calibrated model under a variety of heterogeneity and dispersivity scenarios. We also consider the impact of applying advective-dispersive methods to the same range of scenarios. We show that as true system dispersivity increases, the capacity of kinematic simulations of age to translate the available system information into accurate parameter estimates decreases.

Table of Contents

Acknowledgements	vii
List of Figures	ix
List of Tables	xii
Chapter 1: Introduction	1
Chapter 2: Case Study Description	10
2.0 Introduction	10
2.1 Site Overview	10
2.2 Overview of Flow and Transport Model	13
2.3 References	17
Chapter 3: The Comparative Use of Environmental Tracer Data in Model Calibration to Examine Tracer Transport Processes and Resolve Hydrogeological Uncertainty	19
3.0 Introduction	19
3.1 Methods	25
3.2 Results	30
3.3 Discussion	39
3.4 Conclusions	48
3.5 References	49
Chapter 4: Simulating Nitrate Removal Mechanisms in an Agricultural Catchment with Contrasting Nitrate Base-flow Concentrations in Subcatchment Streams	53
4.0 Introduction	53

4.1 Methods	59
4.2 Results and Discussion	66
4.3 Conclusions and Future Work	83
4.4 References	84
Chapter 5: Numerical Simulation of Seasonal Changes in Base-flow Age	88
5.0 Introduction	88
5.1 Methods	94
5.2 Results and Discussion	102
5.3 Conclusions and Future Work	118
5.4 References	118
Chapter 6: The Impact of Kinematic Assumptions on Parameter Estimation for Advective-Dispersive Groundwater Systems	122
6.0 Introduction	122
6.1 Methods	130
6.2 Results and Discussion	138
6.3 Conclusions	155
6.4 References	156
Chapter 7: Conclusions and Future Work	159
Appendix A: Overview of Scripting Tools Developed for this Work	A-1
Appendix B: Subsurface Environmental Tracer Observations for the Upper Chester Study Site	B-1
Appendix C: Subsurface Nitrate Observations for the Upper Chester Study Site	C-1

Acknowledgements

Financial support for Chapters 3 and 4 of this dissertation was provided by the US Geological Survey Pathways Intern Program and by the National Science Foundation under award numbers CBET-0846244 and CCF-1451708. The William Walker Fellowship from the Virginia Water Resources Research Center funded my participation in an accelerated MODFLOW course at the Colorado School of Mines and consequently laid the foundation for the groundwater modeling integral to these studies.



One of my heroes, the Russian engineer/art historian/theologian Pavel Florensky, so hoped that his work in the world would be good, true, and beautiful that he created a new font solely for purpose of publishing his most important philosophical treatise. If I have any such hope for the work behind and beyond this dissertation, it is decidedly independent of font style or my formatting abilities, but very much the result of the generosity of many teachers and friends. Several persons - without whom my graduate studies would, literally, not have been possible - deserve thanks and recognition, and I will name a few of them here.

Thanks to the service of my dissertation committee, Drs. James Smith, Teresa Culver, Ward Sanford, Jonathan Goodall, and Todd Scanlon. Thanks to my advisor Teresa Culver, for her encouragement to apply to the PhD program, her patience with my slow acquisition of a new craft, her wisdom about the practice of environmental management, and her commitment to teaching the stewardship of water and soil. Thanks to Ward Sanford for his generosity of time and technical expertise, for teaching me so much about groundwater and hydrological modeling, and for his invitation to do research alongside the U.S. Geological Survey. Thanks to Jonathan Goodall and the 'help unlooked for' that he provided in the past year; his interest in and encouragement about my work came at a critical time. Thanks to Todd Scanlon for teaching me so much about catchment hydrology and involving me in his research. Thanks to Katherine Holcomb of the University of Virginia Advanced Computing Services and Engagement group for time-intensive assistance with code parallelization.

Thanks to the friendship of the Schlupfwinkelbruderschaft and their deep hope - often when I had none myself - that the work of our hands will be established.

Among the many advocates for our family and my vocational pursuits, special thanks to Susan Cunningham and Lynn Winn.

To Elias, Griffin, and Wiley - may the many constraints imposed by your dad's graduate studies soon be a distant memory, but may we always talk about water, maps, and the mending of things.

And thanks to Amy, for her commitment to me, for her kindness during these many distracted months, and for our life together.

This work is for Mom and Dad.

ברוך אדני יום | יום יעמס־לנו האל ישועתנו סלה:

List of Figures

Figure 1.1 Percent of Chesapeake Bay attaining water quality standards for dissolved oxygen, chlorophyll *a*, and water clarity.

Figure 1.2 Trend in nitrate concentrations for various flows in the Choptank River on the Maryland Eastern Shore.

Figure 2.1 Upper Chester study site and observation locations.

Figure 2.2 A cross-section of the hydrostratigraphic units for Upper Chester site.

Figure 3.1 North American atmospheric concentrations of CFCs, SF₆, and ³H.

Figure 3.2 Simulated vs. observed tracer concentrations and ³H/³He ages for the calibrated models.

Figure 3.3 Upper Chester parameter estimates using different tracer datasets as calibration targets.

Figure 3.4 Simulated steady state base-flow age distribution for Morgan Creek for the multiple calibrated models.

Figure 3.5 Scatter-matrix of apparent piston-flow ages interpreted from Upper Chester tracer measurements.

Figure 3.6 Tracer apparent ages and nitrate and dissolved oxygen concentrations at well nests in the Upper Chester.

Figure 3.7 Sensitivity of calibrated porosity to the value assumed for recharged excess air.

Figure 4.1 Observed stream nitrate concentrations at the Morgan Creek and Chesterville Branch gages.

Figure 4.2 Crop acreage, agricultural nitrogen inputs and exports, and estimated recharging nitrate concentrations for agricultural land in the Upper Chester.

Figure 4.3 Simulated and observed groundwater nitrate concentrations for observation wells in the Upper Chester.

Figure 4.4 Simulated and observed base-flow nitrate concentrations in Morgan Creek and Chesterville Branch.

Figure 4.5 Contributions of removal mechanisms to reducing exported loads in Morgan Creek and Chesterville Branch.

Figure 4.6 Recharge-dates predicted by the flow model for Upper Chester groundwater nitrate observations.

Figure 4.7 Age, nitrate, and dissolved oxygen observations for the well nest that includes KEBe167, KEBe166, and KEBe165.

Figure 4.8 Flow characteristics measured at the Morgan Creek and Chesterville Branch stream gages. Each marker represents a field measurement.

Figure 4.9 Base-flow stream nitrate concentrations from synoptic surface water sampling in Morgan Creek and Chesterville Branch.

Figure 5.1 Seasonal variations in stream nitrate concentrations and stream discharge for Chesterville Branch.

Figure 5.2 From McGuire and McDonnell (2006). Schematic illustrating the use of (i) a precipitation input signal plus (ii) the assumption that the streamflow preserves that signal to infer the TTD for a catchment.

Figure 5.3 Aquifer dimensions, hydraulic conductivity fields, and recharge spatial distributions for synthetic 2D scenarios.

Figure 5.4 Steady state base-flow age distributions for different aquifer scenarios subject to 40 cm/year recharge.

Figure 5.5 Steady state base-flow age distributions for the homogenous 1 m/day conductivity field subject to different recharge rates.

Figure 5.6 Comparison of base-flow age time series resulting from changes in the aquifer conductivity field.

Figure 5.7 Flow and age time series for the layered conductivity scenario with annually averaged recharge = 40 cm/year.

Figure 5.8 Effects of simulated grid size on the fit of simulated mean ages to a sinusoidal function for homogeneous aquifer with conductivity = 10 m/day and annual recharge = 40 cm/year.

Figure 5.9 Steady state age distribution at maximum base-flow conditions for homogeneous aquifer with hydraulic conductivity = 1 m/day.

Figure 5.10 Comparison of base-flow age time series resulting from changes in annually averaged recharge and aquifer mean age.

Figure 5.11 Comparison of base-flow age time series resulting from changes in the amplitude and spatial distribution of the seasonally-varying recharge.

Figure 5.12 Simulated seasonal base-flow age and SF₆ concentration for Morgan Creek.

Figure 6.1 Comparison of simulated advective flowpaths to a wellscreen and a registration volume surrounding the well screen.

Figure 6.2 Aquifer conductivity configurations and locations of calibration targets for synthetic cases.

Figure 6.3. WSSE progression for different heterogeneity scenarios and calibration methods.

Figure 6.4 Parameter estimates for Stratified heterogeneity scenarios.

Figure 6.5 Sequentially-calibrated parameter estimates for Upper Chester.

Figure 6.6 Simultaneously-calibrated parameter estimates for Upper Chester.

List of Tables

Table 3.1 Transport calibration methods.

Table 3.2 Mean and median base-flow age simulated at the Morgan Creek stream gage by the calibrated models.

Table 4.1 Calibrated nitrogen removal mechanisms and transport parameters.

Table 5.1 Multiples of steady state recharge used to decompose the annually averaged recharge into seasonally varied recharge.

Table 5.2 Comparison of base-flow SF₆ apparent ages and mean age of discharging particle distribution.

Table 6.1 Options for inclusion of groundwater age information in model calibration.

Table 6.2 Calibration methods and calibrated parameters for synthetic cases.

Table 6.3 Calibration targets for Upper Chester calibrations using steady state interpreted age.

Table 6.4 Calibrated system porosities for different calibration methods and Synthetic aquifer configurations.

Table 6.5 True and simulated mean base-flow age for different calibration methods and Synthetic aquifer configurations.

Table 6.6 True age as a function of dispersivity at select observation sites.

Chapter 1: Introduction

A defining characteristic of the Anthropocene (whether or not that term is used in a geologically-specific manner; cf. Autin and Holbrook, 2012, and Monastersky, 2015) is the human re-configuration of geochemical cycles (Steffen et al., 2011). The most widely discussed example of this is the transformation of fossil fuels to carbon dioxide. Carbon has, understandably, become the stuff of treaties and papal reflection (Francis I, 2015), and policy questions at the inter-governmental scale are focused on reducing the accelerated relocation of carbon and the associated impact on the climate. However, other geochemical interruptions in recent centuries have also had profound impacts on environmental systems, and for many aquatic systems no change has been more dramatic than alterations to the nitrogen cycle (Vitousek et al., 1997; Galloway et al., 2003; Galloway et al., 2008). By some estimates, anthropogenic conversion of atmospheric nitrogen to reactive forms through the Haber-Bosch process will soon surpass biological nitrogen fixation as the largest global producer of bioavailable nitrogen (Galloway et al., 2004). With curious symmetry, nitrogen has been pulled from the sky before ultimate deposition in estuaries and coastal bays at the same time that carbon has been re-routed to the atmosphere from prehistoric seas and swamps. Just as our capacity as humans to get and burn carbon has provided warmth, light, and innumerable resulting forms of life, community, and creativity (complicatedly interwoven with more tragic, carbon-energized forms of human agency), so also our capacity to make bioavailable nitrogen has dramatically increased food production through the application of inorganic nitrogen fertilizers to agricultural systems. However, this application of nitrogen fertilizers has, for much of the last century, exceeded the

demands of the crops that we grow, and it is largely – though not solely – a result of this agricultural use that more nitrogen than any other nutrient is entering the hydrosphere from the landscape (Schindler and Vallentyne, 2008).

Because of water quality problems associated with eutrophication, nitrogen has been a subject of policy concern in developed countries for several decades (e.g., Royal Commission, 1979; Burt et al., 2010). In the United States, the Chesapeake Bay is one of the most significant ecological and economic system subject to the adverse effects of nitrogen pollution. The Chesapeake Bay, like many estuarine systems, has chronic water quality problems that are driven by excessive nutrient loads from upstream sources (Kemp et al., 2005). Like most estuarine systems, eutrophication in the Chesapeake Bay is nitrogen limited most of the time (Science and Technical Advisory Committee, 2010; Vitousek et al., 1997; Schindler and Vallentyne, 2008). State and federal governments have made significant efforts since the early 1980's to improve water quality in the Chesapeake Bay. However, voluntary efforts by states have not impacted key indicators of the Bay's health (**Figure 1.1**), and as a result the Environmental Protection Agency (EPA) assumed oversight of the Bay remediation effort in 2010. Federal courts have recently confirmed that the EPA may require the governments of states in the Bay watershed to reduce the discharge of nitrogen, phosphorous, and sediment to Bay tidal waters and tributaries (Hicks, 2015). Each jurisdiction must submit a series of Watershed Implementation Plans (WIPs) that describe how they will make the reductions required to meet the intermediate goals of the Bay Total Maximum Daily Load (TMDL) by 2017 and the final goal by 2025.

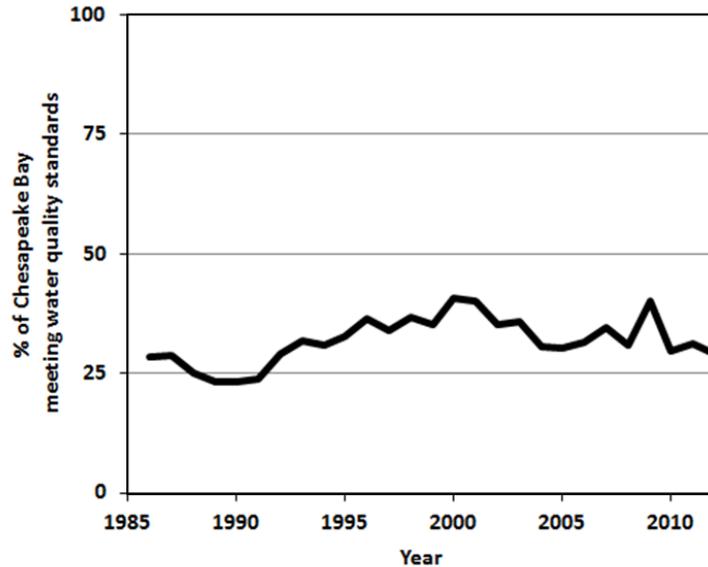


Figure 1.1 Percent of Chesapeake Bay attaining water quality standards for dissolved oxygen, chlorophyll *a*, and water clarity. Data from http://www.chesapeakebay.net/indicators/indicator/achievement_of_chesapeake_bay_water_quality_standards (downloaded 7/7/2015).

Agriculture is a primary driver of nitrogen to the Chesapeake Bay, and much of the agricultural nitrogen load to the Bay is transported via groundwater linkages. While significant efforts have been made to reduce agricultural nitrogen export to the Chesapeake Bay, loads to Bay tributaries and tidal waters from many agriculturally-intensive areas continue to increase (**Figure 1.2**). This temporal disconnect between land surface improvements and in-stream nitrogen loads is due to the accumulation of nitrates in surficial aquifers beneath agricultural systems during the past century (Puckett et al., 2011). At a given location the present in-stream nitrate concentration is an artifact of past land surface actions, as the base-flow nitrate concentration integrates agricultural nitrogen inputs that were distributed backwards in time and across the land surface in space. That is, any base-flow nitrate concentration is a complicated function of both transient surface phenomenon (such as precipitation and land use patterns) and subsurface characteristics

(such as aquifer travel times and dispersivities). It may consequently be difficult to extract the signal of management improvement from the noise that is associated with multiple, converging flow paths of varying age and nitrate concentration.

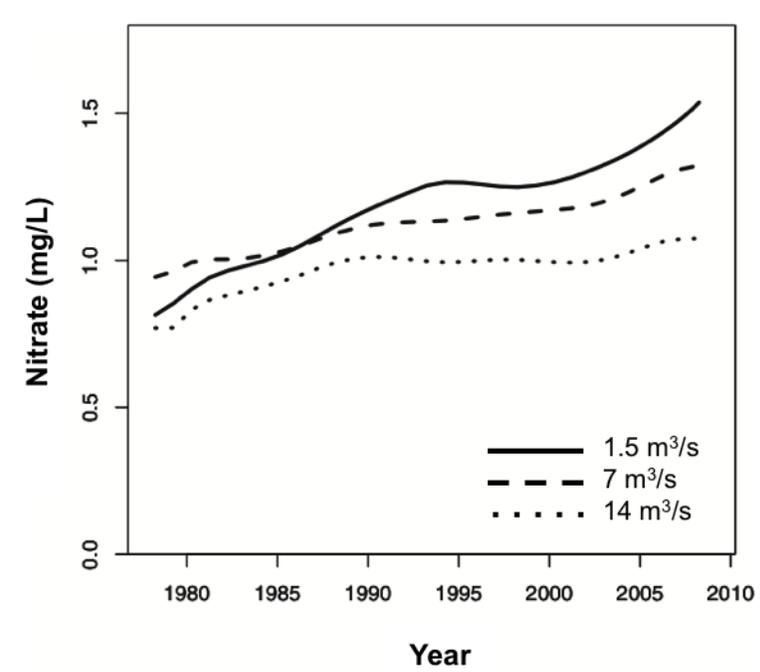


Figure modified from Hirsch et al. (2010)

Figure 1.2 Trend in nitrate concentrations for various flows in the Choptank River on the Maryland Eastern Shore. Nitrate estimated for April 1 of each year using the Weighted Regressions on Time, Discharges, and Seasons (WRTDS) model (Hirsch et al., 2010). The increasing concentration at low flow shows the impact of nitrates transported by groundwater flow and base-flow discharge.

Thus, while numerous studies have connected the increase in groundwater nitrates to the intensification of agricultural practices during the last century, the magnitude and timing of the impact of agricultural practices that are intended to lower groundwater nitrate levels is not well understood (Meals et al., 2010). The uncertainty associated with predicting the water quality benefits of landscape action can have serious implications for the management of environmental systems like the Chesapeake Bay. The most basic questions

for such a system are: (1) Who is responsible for the degraded water quality and adverse ecological/economic impacts in the downstream receiving waters? (2) How much cumulative management action (for example, changes in farming practice) must be taken in order to improve the water quality and reverse the ecological/economic impacts? and (3) When will the degraded system respond to management improvements? Where the time scale of ecological response is much longer than the time scale of public opinion and legislative action (as is true of the Chesapeake Bay), analyses are required that will enable stakeholders and decision-makers to discern the impacts - or explain the lack of impacts - of mandated conservation measures. For example, the full engagement of the Bay agricultural community in the Phase III Watershed Implementation Plans (WIPs) will be supported by analytical tools that can distinguish between the failure and pending success of hardship incurred due to the Phase II WIPs.

Groundwater transport processes and timescales are not well represented in the present suite of Chesapeake Bay Program (CBP) models. The process-based Phase 5.3.2 Watershed Model (WSM) provides a steady state estimate of the delivery of nitrogen, phosphorous, and sediment to the Chesapeake Bay as a function of land use, point sources, population, and atmospheric deposition (USEPA, 2010). The WSM conversion of land use to nutrient delivery rates is supported by estimates from the empirical Spatially Referenced Regression on Watershed Attributes (SPARROW) model. SPARROW is similarly steady state, relating estimates of long-term averaged land surface loading to downstream observations in order to calibrate parameters that describe delivery and transport (Ator et al., 2011). These limitations thus require supplementary studies to qualify the modeling results of the WSM, to better estimate the timescales of groundwater improvement, and to

continue improving its parameterization and calibration procedures. A critical but uncertain parameter in the WSM and the regulatory application of its results is the magnitude of load reductions due to agricultural best management practices (BMPs) and the timing of water quality benefits from those load reductions. In order to improve understanding of the relationship between land surface management improvements and water quality response, the Chesapeake Bay Program (CBP) has endorsed the use of targeted small watershed studies as a means of: (i) focusing resources for the accelerated implementation of agricultural best management practices (BMPs); (ii) documenting that implementation and making higher resolution land use data available to researchers and key federal partners such as the US Geological Survey; and (iii) explaining the magnitude and time-scale of water quality responses to agricultural BMPs (Science and Technical Advisory Committee, 2015). This dissertation supports the CBP objectives for small watershed studies by characterizing groundwater flow and nitrate transport and removal in one of the targeted watersheds. In so doing, this dissertation examines a series of questions related to understanding and managing nitrogen in agricultural systems, with particular focus on (i) modeling nitrogen transport subject to groundwater lag times and (ii) calibrating the simulation tools used for that modeling.

The order of the dissertation is as follows. In Chapter 2 we introduce the Upper Chester study site, an agricultural watershed on the Maryland Eastern Shore, and we describe the modeling framework that is common to the analysis of the subsequent chapters. The Upper Chester site is one of three agriculturally-intensive watersheds in the Chesapeake Bay drainage that has been targeted by the U.S. Department of Agriculture (USDA) for further study; as a result, it is an important site for the sort of small watershed study

described above. In Chapter 3 we apply a uniquely rich dataset of environmental tracer observations to the calibration of the flow and transport model, and we compare the range and character of different catchment-scale simulation models that result when an automated calibration routine is driven by various combinations of spatially- and temporally-distributed CFC, SF₆, ³H, and ³He information. In Chapter 4 we use the calibrated model to link estimated historic land surface inputs to subsurface and base-flow nitrate concentrations in order to characterize the controls on nitrate removal in the Upper Chester. While subsurface nitrate transport and catchment removal processes have been widely investigated, there have been few fully distributed, three-dimensional modeling studies of nitrate transport and removal in catchments with nitrogen removal rates that are highly spatially-variable, as is the case in the Upper Chester. Chapters 3 and 4 use the steady state flow model to describe environmental tracer and nitrate transport; under these steady state conditions, the age of base-flow discharging to a stream is, by extension, constant. In some situations, variability in base-flow age may impact in-stream solute concentrations, and in Chapter 5 we use several synthetic aquifers as well as the Upper Chester model to examine the impact of time-variable hydrologic forcing – such as that due to seasonal changes in precipitation and evapotranspiration – on the age of base-flow discharge. Finally, in Chapter 6, we return to the topic of the use of groundwater age information in parameter estimation and consider the impact of assumptions about age transport on the automated calibration of groundwater flow and transport models. In sum, the purpose of the studies described in the following chapters is to improve our ability to calibrate and use numerical groundwater simulation tools, with the immediate aim of better modeling the impact of agricultural intensification and de-intensification on water

quality, and thus improve our ability to interpret signs of decline and improvement in receiving streams.

REFERENCES

Ator, S.W., Brakebill, J.W., and Blomquist, J.D. (2011) *Sources, fate, and transport of nitrogen and phosphorus in the Chesapeake Bay watershed*. U.S. Geological Survey Scientific Investigations Report 2011-5167, 27 p.

Autin, W.J., and Holbrook, J.M. (2012). Is the Anthropocene an issue of stratigraphy or pop culture? *GSA Today*, 22(7), 60-61.

Burt, T. P., Howden, N. J. K., Worrall, F., Whelan, M. J., & Bierzoza, M. (2010). Nitrate in United Kingdom Rivers: Policy and Its Outcomes Since 1970. *Environmental Science and Technology*, 45(1), 175-181.

Francis I (2015). *On Care for Our Common Home – Laudato Si'* [Encyclical Letter]. Retrieved from <http://w2.vatican.va/content/francesco/en/encyclicals.index.html#encyclicals>.

Galloway, J. N., Aber, J. D., Erisman, J. W., Seitzinger, S. P., Howarth, R. W., Cowling, E. B., & Cosby, B. J. (2003). The nitrogen cascade. *Bioscience*, 53(4), 341-356.

Galloway, J.N., Dentener, F.J., Capone, D.G., Boyer, E.W., Howarth, R.W., Seitzinger, S.P., Asner, G.P., Cleveland, C.C., Green, P.A., Holland, E.A., Karl, D.M., Michaels, A.F., Porter, J.H., Townsend, A.R., Vorosmarty, C.J. (2004). Nitrogen cycles: past, present, and future. *Biogeochemistry*, 70(2), 153-226.

Galloway, J.N., Townsend A.R., Erisman, J., Bekunda, M., Cai, Z., Freney, J., Martinelli, L., Seitzinger, S., Sutton, M. (2008). Transformations of the nitrogen cycle: recent trends, questions, and potential solutions. *Science*, 320(5878), 889-892.

Hicks, Josh (2015, July 6). US appeals court upholds Chesapeake Bay clean-up plan. *The Washington Post*. Retrieved from <http://www.washingtonpost.com>.

Hirsch, R. M., Moyer, D. L., & Archfield, S. A. (2010). Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs1. *JAWRA Journal of the American Water Resources Association*, 46(5), 857-880.

Kemp, W. M., Boynton, W. R., Adolf, J. E., Boesch, D. F., Boicourt, W. C., Brush, G., Cornwell, J.C., Fisher, T.R., Glibert, P.M., Hagy, J.D., Harding, L.W., Houde, E.D., Kimmel, D.G., Miller, W.D., Newell, R.I.E., Roman, M.R., Smith, E.M., & Stevenson, J. C. (2005). Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Marine Ecology Progress Series*, 303(21), 1-29.

Meals, D. W., Dressing, S. A., & Davenport, T. E. (2010). Lag time in water quality response to best management practices: A review. *Journal of Environmental Quality*, 39(1), 85-96.

Monastersky, R. (2015). Anthropocene: the human age. *Nature*, 519(7542), 144-147.

Puckett, L. J., Tesoriero, A. J., & Dubrovsky, N. M. (2011). Nitrogen contamination of surficial aquifers: A growing legacy. *Environmental Science and Technology*, 45(3), 839.

Royal Commission on Environmental Pollution (1979). Seventh Report: Agriculture and Pollution. Her Majesty's Stationary Office, Command Paper 7644.

Schindler, D. W., & Vallentyne, J. R. (2008). *The Algal Bowl*. University of Alberta Press.

Science and Technical Advisory Committee (2010). Small Watershed Monitoring Designs: A Report from the Chesapeake Bay Program Scientific and Technical Advisory Committee. Science and Technical Advisory Committee Publication 10-004. Downloaded from www.chesapeake.org/stac/stac_pubs.php.

Science and Technical Advisory Committee (2015). Estimating Land Management Effects on Water Quality. Science and Technical Advisory Committee Publication 15-002. Downloaded from www.chesapeake.org/stac/stac_pubs.php.

Steffen, W., Grinevald, J., Crutzen, P., McNeill, J. (2011). The Anthropocene: conceptual and historical perspectives. *Philosophical Transactions of the Royal Society*, 369(1938), 842-867.

United States Environmental Protection Agency (USEPA). (2010). Chesapeake Bay Phase 5.3 Community Watershed Model. EPA 903S10002 - CBP/TRS-303-10. U.S. Environmental Protection Agency, Chesapeake Bay Program Office, Annapolis MD.

Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., Schlesinger, W.H., and Tilman, D. G. (1997). Human alteration of the global nitrogen cycle: sources and consequences. *Ecological applications*, 7(3), 737-750.

Chapter 2: Case Study Description

2.0 INTRODUCTION

This chapter introduces the study site and the numerical model that was used to simulate groundwater flow and the transport of environmental tracers, age, and nitrates as described in chapters 3-6 of this dissertation. Model details common to all simulation studies, such as the model framework, the basic conceptualization of flow and transport processes, and the head and discharge data used to calibrate the flow model, are described here. Further details regarding the simulation of environmental tracer and nitrate transport, including the data and parameter estimation methods specific to that transport, are described in subsequent chapters.

2.1 SITE OVERVIEW

The study site is located on the Maryland Delmarva Peninsula and combines the USDA-designated Upper Chester Showcase Watershed (Nelson and Spies, 2013) with the adjacent Morgan Creek subcatchment (**Figure 2.1**). For purposes of this study, 'Upper Chester' hereafter refers collectively to the Showcase Watershed plus the Morgan Creek watershed. The 125-km² study site is a low-relief agricultural watershed underlain by variably permeable unconsolidated sediments that dip to the southeast (Böhlke and Denver, 1995). The topmost fluvial deposits of the Pensauken formation and the underlying marine deposits of the Aquia formation together form an unconfined aquifer that is responsible for the majority of base-flow discharge to both Morgan Creek and Chesterville Branch (Böhlke and Denver, 1995). The Aquia Formation is underlain by the Aquia Confining Unit, which

likewise dips to the southeast, outcropping at Morgan Creek and forming its bed sediments (Puckett et al., 2008). The identification of the confining unit conductivity is one focal purpose of this study. The Aquia Confining Unit separates the surficial unconfined aquifer from the Hornerstown Aquifer; the conductivity of the confining unit – and the associated proportion of older, pre-agricultural water that discharges to Morgan Creek – may be an important driver of stream water chemistry in Morgan Creek (Böhlke and Denver, 1995; Bachman et al., 2002). In order to account for the effects of weathering on Confining Unit permeability, we assumed that Confining Unit hydraulic conductivity was twice its areal average in locations where the Confining Unit outcropped at the land surface.

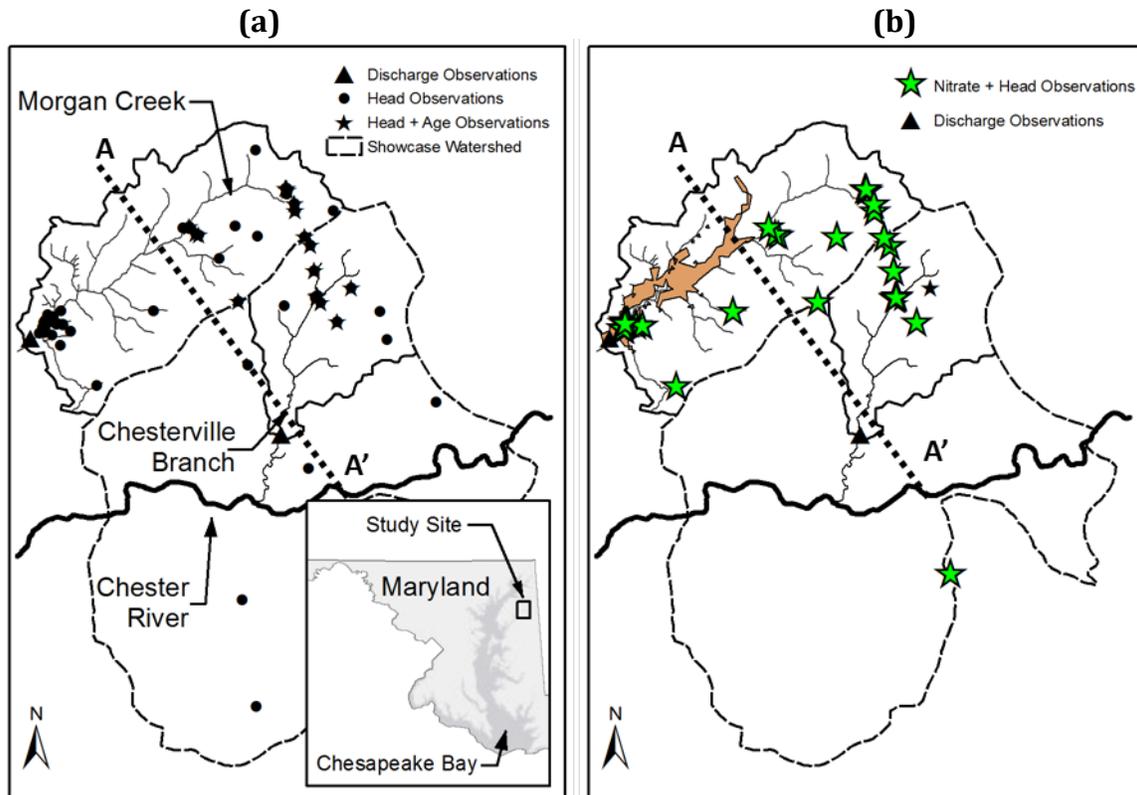


Figure 2.1 Upper Chester study site and observation locations. Note that some subsurface observation locations indicate nested wells with observations at multiple elevations. Shaded area in panel (b) shows the simulated location of Aquia Confining Unit in the lower reaches of Morgan Creek (see discussion in text). See Figure 2.2 for cross-section A-A'.

Results from preliminary calibration runs showed that hydraulic conductivity estimates for the Pensauken and Aquia units are highly correlated given the available data. As a result, in the numerical model the Pensauken and Aquia Aquifers were combined into a single hydrogeologic unit that we refer to as the ‘Surficial Aquifer’ (**Figure 2.2**). For ease of reading we hereafter refer to the Aquia Confining Unit as simply the ‘Confining Unit’ and to the Hornerstown Aquifer as the ‘Confined Aquifer’ (**Figure 2.2**).

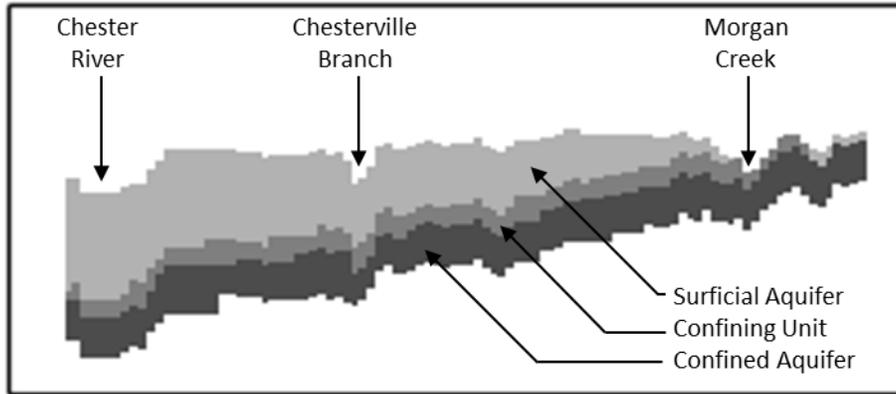


Figure 2.2 A cross-section of the hydrostratigraphic units for Upper Chester site. See **Figure 2.1** for cross-section location (A-A').

2.2 Overview of Flow and Transport Model

2.2.1 – Model Framework and Simulation of Groundwater Flow

Groundwater flow in the Upper Chester was modeled using the U.S. Geological Survey (USGS) numerical simulation code MODFLOW (Harbaugh, 2005). MODFLOW-2005 was used for steady state solutions, which were subsequently used to simulate the transport of environmental tracers and nitrates as described in chapters 3 and 4 of this dissertation. A version of MODFLOW that was modified by S.S. Papadopoulos and Associates for purposes of better handling water table fluctuations was used for the transient head solution used in Chapter 5 of this dissertation and is described in more detail there.

For all simulations of flow and transport in the Upper Chester the finite-difference grid consisted of 220 rows, 180 columns, and 27 layers with square model cells 76 m across. Layer thicknesses were 1.5 m for the top 16 layers and 3.0 m or 7.6 m for the lower layers.

A deterministic hydrogeologic framework for the study site was constructed using LIDAR elevation data from the USGS National Map (<http://nationalmap.gov>) and regional geologic unit elevation contours digitized by the Maryland Geological Survey (Andreasen et al., 2007) (**Figure 2.2**). Hydrogeologic units that fall within the model domain were sampled onto the finite difference grid and adjusted to conform to local hydrogeological data that is available from previous studies in the area (e.g., Drummond, 1998; Böhlke and Denver, 1995; Bachman et al., 2002). No head or tracer data exists for the hydrogeologic units below the confined Hornerstown Aquifer; consequently, for our simulations the bottom of the active simulation volume corresponds to the bottom of the confined aquifer (**Figure 2.2**).

MODFLOW calculates the head at each cell in the model grid by solving a finite-difference discretization of the general equation for groundwater flow through porous media,

$$S \frac{\partial h}{\partial t} = \frac{\partial}{\partial x} \left(K_x \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_y \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_z \frac{\partial h}{\partial z} \right) + W, \quad (2.1)$$

where S is specific storage, h is head at location and time (x,y,z,t) , K_i for $i = x, y, z$ is the hydraulic conductivity along the appropriate axis, and W is the volumetric flux due to inputs and withdrawals at location and time (x,y,z,t) . There is little evidence of irrigation (i.e., pumping) or other relevant stresses in the study area (Hancock and Brayton, 2006), and we consequently assumed that the single input and single output to the aquifer system were recharge and base-flow discharge, respectively. We assumed that a groundwater divide and associated zero-flux boundary condition coincided with the watershed divide that delineates the Upper Chester study site (**Figure 2.1**).

Base-flow discharge to receiving streams was simulated using MODFLOW's drain package. Under this conceptualization, each cell in the top layer of the model is defined as a head-dependent (i.e., Cauchy type) boundary condition with the boundary elevation equal to the elevation of the land surface at that point. Simulated base-flow discharge consequently occurs at a model location when the water table elevation in a cell exceeds the land surface elevation for that cell. A combination of the USGS code Zonebudget (Harbaugh, 1990) and Python scripting was used to post-process the drain outflow calculated by MODFLOW at each time step; this post-processing calculated the total simulated base-flow in Morgan Creek and in Chesterville Branch by aggregating the discharge from all model cells located upstream of the gage in each reach. The Chester River (**Figure 2.1**) was simulated as a general head boundary at constant stage of 0.76 m, which is the mean high water elevation at Chestertown (email communication, Buck Nickerson), approximately 5 km downstream of the active model area.

Our calibration targets for simulated head were long-term average groundwater levels from observation wells that had ten or more groundwater level measurements. Discharge calibration targets were average discharge measurements in Morgan Creek and Chesterville Branch (**Figure 2.1**). Base-flow was assumed to be 85% of measured streamflow per the results of Sanford et al. (2011), who examined base-flow components of streams in other similar low-relief, high-infiltration catchments on the Delmarva Peninsula. Base-flow residuals from preliminary calibration runs suggested that recharge rates per unit area are higher in the Chesterville Branch drainage than the Morgan Creek drainage. Reilly et al. (1994) found for their 2D transect simulation that reducing recharge in areas of more poorly drained soils in the catchment uplands improved the model fit. Applying this

insight to the entire Upper Chester improved the base-flow residuals for our 3D model since soils in the Morgan Creek catchment are on average more poorly drained than in the Chesterville Branch catchment. Consequently, two recharge zones, distinguishing well-drained soil types from other soil types, were derived from USDA SSURGO data for Kent County and Queen Anne's County (Soil Survey Staff).

2.2.1 – Simulation of Subsurface Transport

For all transport simulations, travel times through the unsaturated zone were assumed to be negligible due to the high water table and thin unsaturated zone on the Delmarva Peninsula. Initial tests compared the performance of two different approaches to the simulation of subsurface transport: an advection-only simulation using MODPATH (Pollock, 2012) and a fully advective-dispersive model (ADM) using MT3DMS (Zheng and Wang, 1999). However, the parameter estimates resulting from ADM simulation of both age and tracer concentration (not shown) appeared to be impacted by numerical dispersion introduced by the finite difference solution to the ADM. Several authors have described the challenges of using lower order finite difference ADM schemes to simulate transport in advection-dominated systems (e.g., Zheng and Bennett, 2002). From their simulation of a 2D cross-section in the Chesterville Branch subcatchment, Reilly et al. (1994) concluded that hydrodynamic dispersion was very small. They estimated longitudinal dispersivity to be less than 0.3 m and inferred from the sharp vertical fronts between nested wells vertical that transverse dispersivity was effectively zero. Chapter 6 of this dissertation shows that the use of a Eulerian ADM conceptual model to simulate age in a low-dispersive system may have unpredictable effects on the estimation of flow and transport parameters. The use of

an ADM for the Upper Chester confirms those results, and also suggests that the numerical dispersion associated with an ADM method may confound automated parameter estimation whether tracer concentrations or age are simulated. Consequently, hydrodynamic dispersion in the Upper Chester was assumed to be minimal, and advection-only transport was simulated solely in the remainder of this study. For all transport described in subsequent chapters, which include studies describing the transport of environmental tracers (Chapters 3 and 5), nitrate (Chapter 4), and age (Chapters 5 and 6), the particle-tracking code MODPATH (Pollock, 2012) and a series of Python scripts was used to (i) convert the simulated head field calculated by MODFLOW into a simulated velocity field and (ii) generate a distribution of advective travel times for subsurface and surface water observation locations. These travel time distributions (TTD) were convolved with land surface solute input functions in order to simulate the solute concentrations at the given location. More details on the use of this method to calculate environmental tracer and nitrate concentrations are provided in the next two chapters.

2.3 REFERENCES

- Andreasen, D. C., Achmad, G., Staley, A. W., & Hodo, R. M. (2007). *Hydrogeologic framework of the Maryland coastal plain*. Maryland Geological Survey Report, Baltimore, MD.
- Bachman, L. J., Krantz, D. E., & Bohlke, J. (2002). *Hydrogeologic Framework, Ground-water Geochemistry, and Assessment of Nitrogen Yield from Base Flow in Two Agricultural Watersheds, Kent County Maryland*. Environmental Protection Agency.
- Böhlke, J. K., & Denver, J. M. (1995). Combined use of groundwater dating, chemical, and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, Atlantic coastal plain, Maryland. *Water Resources Research*, 31(9), 2319-2339.
- Drummond, D. D. (1998). Hydrogeology, simulation of ground-water flow, and ground-water quality of the Upper Coastal Plain aquifers in Kent County. *Maryland Geological Survey Report of Investigations*, (68), 76.

Harbaugh, A. W. (2005). *MODFLOW-2005, the US geological survey modular ground-water model: The ground-water flow process*. US Department of the Interior, US Geological Survey Reston, VA, USA.

Harbaugh, A.W., (1990). *A computer program for calculating subregional water budgets using results from the U.S. Geological Survey modular three-dimensional ground-water flow model*. U.S. Geological Survey Open-File Report 90-392, 46 p.

Hancock, T. C., & Brayton, M. J. (2006). *Environmental setting of the Morgan Creek Basin, Maryland, 2002-04*. US Geological Survey Open File Report 2006-1151.

Nelson, J., & Spies, P. (2013). The Upper Chester river watershed: Lessons learned from a focused, highly partnered, voluntary approach to conservation. *Journal of Soil and Water Conservation*, 68(2), 41A-44A.

Pollock, D. W. (2012). *User guide for MODPATH version 6: A particle tracking model for MODFLOW*. US Department of the Interior, US Geological Survey.

Puckett, L. J., Zamora, C., Essaid, H., Wilson, J. T., Johnson, H. M., Brayton, M. J., & Vogel, J. R. (2008). Transport and fate of nitrate at the ground-water/surface-water interface. *Journal of Environmental Quality*, 37(3), 1034-50.

Reilly, T. E., Plummer, L. N., Phillips, P. J., & Busenberg, E. (1994). The use of simulation and multiple environmental tracers to quantify groundwater flow in a shallow aquifer. *Water Resources Research*, 30(2), 421-433.

Sanford, W. E., Nelms, D. L., Pope, J. P., & Selnick, D. L. (2011). *Quantifying components of the hydrologic cycle in Virginia using chemical hydrograph separation and multiple regression analysis*. US Department of the Interior, US Geological Survey.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database. Available online at <http://sdmdataaccess.nrcs.usda.gov/>.

Zheng, C., & Bennett, G. D. (2002). *Applied contaminant transport modeling* (2nd Edition). Wiley-Interscience New York.

Zheng, C., & Wang, P. P. (1999). *MT3DMS: A modular three-dimensional multispecies transport model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems; documentation and user's guide*.

Chapter 3: The Comparative Use of Environmental Tracer Data in Model Calibration to Examine Tracer Transport Processes and Resolve Hydrogeological Uncertainty¹

3.0 INTRODUCTION

Numerical simulation of groundwater solute transport is a critical tool for predicting the behavior of contaminants, including the identification of the probable location and timing of their introduction to the subsurface and the management of any risks they pose to human and environmental health. Field observations of groundwater age can be an important means of establishing rates of recharge and subsurface velocities; consequently, they may also be used to constrain estimates of the model parameters critical to the prediction of contaminant transport and management impact (Sanford, 2011).

Groundwater age is the time between a water molecule's entry into a groundwater system and the observation of that water at some subsequent point in space and time. Data describing this transit time is available from analyses of environmental tracer concentrations in subsurface and base-flow samples. A number of different tracers exist, with the applicability of any individual tracer to a particular groundwater question dependent upon the timescales and potential confounding environmental factors involved. See Cook and Herczeg (2000), Kazemi et al. (2006), or Turnadge and Smerdon (2014) for comprehensive reviews of different tracer sources and some methodological considerations for their use in groundwater studies.

¹ A version of this chapter is currently under review with *Journal of Hydrology*.

Chlorofluorocarbons (CFCs), sulfur hexafluoride (SF_6), and tritium (^3H) are particularly useful for the study of shallow groundwater systems that have been (and still are) subject to the various types of anthropogenic contamination over the past century. CFCs (i.e., CFC-11, CFC-12, and CFC-113) and SF_6 are gaseous atmospheric constituents that dissolve into recharging waters (Plummer and Busenberg, 2000), while ^3H is an isotopic variation on the water molecule itself (Solomon and Cook, 2000). The timescale of intensification for the atmospheric concentrations of these tracers roughly corresponds to the timescale of intensification of chemical fertilization of agricultural land (**Figure 3.1**).

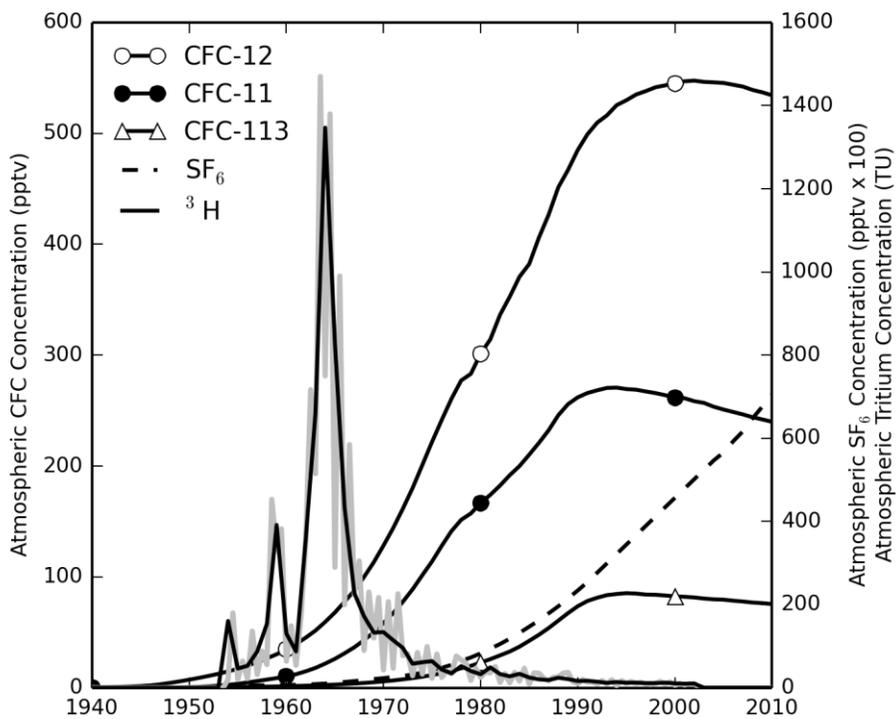


Figure 3.1 North American atmospheric concentrations of CFCs, SF_6 , and ^3H . All lines show the annually-averaged concentration except for the grey line, which shows all available data for atmospheric ^3H . Data downloaded from http://water.usgs.gov/lab/software/air_curve/index.html.

For a problem in which the timing and location of a contaminant source is adequately known, a time series of data that describes the plume development for the contaminant of interest may itself be used to calibrate a model that can then be applied for predictive and management purposes. However, problems of nonpoint source subsurface contaminants, such as those associated with agricultural inputs, may include large source-term uncertainties due to the lack of data describing the spatial and temporal variability of land surface loading; this uncertainty is compounded by a similar lack of data describing contaminant transformations (e.g., denitrification) that may occur between the land surface and the water table, or along a flow path through the saturated zone. These uncertainties make it difficult to estimate transport parameters through an inverse modeling exercise constrained only by data pertaining to the contaminant of interest, as a result making the model less useful for management. In such cases, sampled subsurface tracer data, combined with our relative confidence in the spatial and temporal distribution of inputs from the atmosphere, may enable improved estimates of system-wide transport parameters such as effective porosity.

A number of studies have used CFCs, SF₆, ³H and/or the tritium/tritiogenic helium (³H/³He) ratio to calibrate the flow or transport parameters of numerical groundwater models. These studies have typically involved the use of tracers to calibrate a single realization of a two-dimensional (2D) model of a transect or a single piezometer nest. For example, Reilly et al. (1994), Cook et al. (1995), Portniaguine and Solomon (1998), and Bauer et al. (2001) used CFC concentrations and/or ³H/³He ages to estimate flow and transport parameters for two-dimensional (2D) groundwater models. Mattle et al. (2001) used a single tracer species (³H/³He ages) with a three-dimensional (3D) model to estimate aquifer recharge

due to river leakance. Zuber et al. (2005) used SF₆ observations from several wells to calibrate a transport model for an aquifer in Poland. In addition to the numerical simulation models employed in these studies, other studies have used these same tracers to calibrate lumped-parameter models of flow and transport. For example, Long and Putnam (2009) paired a convolution model with transient CFC and ³H data at a single well to estimate the time-varying relative contributions of quick-flow and slow-flow reservoirs for a karst system in South Dakota.

In spite of these previous works, few if any studies have examined the effect of tracer species selection or tracer dataset size on the automated calibration of a fully 3D numerical simulation model. The present paper examines the impact of different tracers on parameter estimation for a coastal aquifer system by applying a uniquely rich dataset of subsurface atmospheric tracer observations to the calibration of a 3D flow and transport model. The study compares the range and character of different catchment-scale simulation models that result when an automated calibration routine is driven by various combinations of spatially- and temporally-distributed CFC, SF₆, ³H, and ³He information.

Model calibration may be sensitive to tracer species selection for a variety of reasons, including those factors that impact the recharge of atmospheric tracers to the water table or their subsequent transport through the subsurface. For example, the recharging concentration of a dissolved tracer is a function of the atmospheric concentration, the variables that govern partitioning between the gas and dissolved phases (i.e., recharge temperature and elevation), and the amount of excess air introduced at the time of recharge (Plummer et al., 2006; Goody et al., 2006). When using inverse modeling to estimate

system transport parameters, inaccurate representation of the tracer mass boundary condition may result in inaccurate interpretation of system velocities and, therefore, system effective porosity. However, the different tracers – and therefore the models that result from using the different tracers as calibration targets – vary in their sensitivity to representation of these factors. For example, the relative solubilities of SF₆ and CFCs make SF₆ much more sensitive to errors in estimating the excess air content of recharging waters (Goody et al., 2006). In cases where the assumed amount of excess air cannot be well-constrained (e.g., by sampled ratios of atmospheric nitrogen to argon), parameters that are estimated against SF₆ data may be subject to uncertainty that can be interrogated by comparison with estimations made against other tracers (Busenberg and Plummer, 2000).

Inaccurate estimates of system transport parameters may similarly result from inaccurate conceptualization of tracer transport processes. For example, most studies that have used atmospheric tracer data to calibrate groundwater models have assumed conservative tracer transport (e.g., Szabo et al., 1996; Hunt et al., 2006). However, several field and laboratory studies have observed the non-conservative behavior of CFCs due to degradation or sorption in a variety of hydrogeological settings. Cook et al. (1995) found that CFC-11 was degraded in both the highly organic unsaturated zone as well as in anaerobic portions of a Canadian aquifer, while CFC-113 appeared to be retarded by sorption in the unsaturated zone. Bauer et al. (2001) found evidence of delayed CFC-113 transport, which they modeled with linear sorption ($R = 1.5$). Results from Happell et al. (2003) indicated the removal of CFCs in the anoxic sediments of the groundwater-surface water interface. Both Sebol et al. (2007) and Hinsby et al. (2007) observed the degradation of CFCs in anaerobic zones due to pyrite oxidation in unconfined aquifers.

For CFCs that recharged when the atmospheric concentration was monotonically increasing (**Figure 3.1**), either degradation or retardation of CFCs in the subsurface would result in an under-estimation of system velocities (i.e., an exaggeration of the apparent age) and an associated over-estimation of system porosity if that mass loss or delay is not accounted for in the transport model.

In sum, the dataset and methods described below further the investigation of important uncertainties associated with tracer transport. The study demonstrates the comparative use of inverse modeling to highlight consistency or detect variability in transport processes between different tracer species. As such, this is a tracer-specific demonstration of the more general practice of using model calibration as a tool to corroborate a conceptual model or identify its shortcomings (Hill and Tiedeman, 2006; Hill, 2006).

In the course of using parameter estimation as a tool for comparing tracer transport, the study also demonstrates the utility of atmospheric-derived tracers to characterize important hydrogeological features. For example, argillaceous leaky confining units, such as the clay and silt Aquia Confining Unit located in the Upper Chester, are often important at controlling groundwater age distribution (Sanford, 2011) but their vertical leakance is notoriously difficult to quantify (Neuzil, 1994). In this regard, previous studies at this site have suggested that the complicated hydrology and nutrient transport in the agriculturally intensive catchment may be further clarified by better characterizing the hydraulic connectivity between the surficial and lower aquifers (Böhlke and Denver, 1995; Bachman et al., 2002).

3.1 METHODS

Multiple environmental tracer species were used individually and in combination to estimate recharge rates, conductivities, and effective porosity for the Upper Chester model. For each method, flow and transport parameters were estimated both sequentially and simultaneously. For the sequential cases, flow parameters (i.e., recharge rates and hydraulic conductivities) were first calibrated against head and discharge data; the resulting estimated values of the flow parameters were then assigned while the effective porosity was calibrated against the tracer concentrations or interpreted ages designated by each transport calibration method (**Table 3.1**). (In the following, 'porosity' refers to effective porosity.) For the simultaneous cases, recharge rates, conductivities, and porosity were calibrated simultaneously.

Table 3.1 Transport Calibration Methods.

<u>Method Name</u>	<u>Transport Calibration Targets</u>	<u>Number of Observations</u>
CFC11	CFC-11 concentrations	68
CFC12	CFC-12 concentrations	75
CFC113	CFC-113 concentrations	52
SF6	SF ₆ concentrations	35
Composite	CFC-11, CFC-12, CFC-113, and SF ₆ concentrations	230
TritHel_Ages	Steady state ages derived from the ³ H/ ³ He ratio	8

Groundwater levels and stream discharges used to constrain calibrated hydraulic conductivities and steady state recharge are described in **Section 2.2** of this dissertation.

Porosity was estimated in this study using CFC, SF₆, ³H, and ³He data collected during previous studies (Busenberg and Plummer, 2000; Dunkle et al., 1993; Ekwurzel et al., 1994). This entire tracer dataset consisted of 238 measurements taken between 1990 and 1998 at 28 observation wells in the active model area (**Figure 2.2; Table 3.1**). The number of observations available and the type of tracer sampled varied for each well; some wells were sampled once for a single tracer, while other wells were sampled multiple times for multiple tracers. No data exists to describe spatially varying porosity, and preliminary calibration experiments using depth-dependent variable porosity did not appreciably impact the resulting parameter estimates. Consequently, a homogenous porosity field was used in all simulations, such that each calibration method estimated a single, spatially-averaged effective porosity.

3.1.1 - Calibration methods using solute concentrations of environmental tracers

The concentration $C_{k,j}(t)$ of a conservative tracer species k at location j and time t in a steady state system may be calculated as the convolution integral

$$C_{k,j}(t) = \int_0^{\infty} C_{k,atm}(t - \tau)g_j(\tau)d\tau \quad (3.1)$$

where $C_{k,atm}$ is the tracer input function (i.e., the time series of recharge tracer concentration of species k) and the system functional $g_j(\tau)$ is equal to the travel time distribution (TTD) of groundwater sampled at location j . As alluded to above, a variety of parametric and non-parametric distributions have been derived for $g(\tau)$ subject to various aquifer conditions (Cook and Böhlke, 2000; Maloszewski and Zuber, 1982; McCallum et al.,

2014). In contrast, we used the numerical model to generate the TTD for each observation well by backward tracking particles from the well screen to the recharge location in order to determine the advective travel time and associated recharge rate for each particle. Well-screens were assumed to be 1-m in length (Reilly et al., 1994) and particles were released at 1 cm intervals. The resulting recharge-weighted histogram of travel times was then convolved with the annually averaged atmospheric tracer time series. Note that this is more precisely expressed as the discrete time form

$$C_{k,j}[t] = \sum_{\tau=0}^{\infty} C_{k,atm}[t - \tau]g_j[\tau] \quad . \quad (3.2)$$

The input function $C_{k,atm}$ for each species k was derived from the tracer atmospheric mixing ratios for North America (**Figure 3.1**). Concentrations in air were converted to dissolved concentrations using Henry's law and assuming constant values of 2 cm³/L recharged excess air and 10° C recharge temperature. Busenberg and Plummer (2000) used N₂-Ar ratios in groundwater to estimate the recharge temperature in the Upper Chester to be 10.5° ± 1.3° C, and they found excess air in nearly all groundwater samples to be between 0 and 3 cm³/L. Dunkle et al. (1993) estimated the recharge temperature on the Delmarva to be 9° ± 2° C.

In addition to calibrations involving single tracer species, a composite data set that included all CFCs and SF₆ measurements was used as transport calibration targets in order to examine the impact of additional information on the parameter estimation process. For the Composite method, both the observations at wells and the input signal for each tracer species were normalized (and therefore located on a common dimensionless scale) by

dividing all observed concentrations and simulated input concentrations by the maximum dissolved input concentration for that species.

3.1.2 - Calibration using ages inferred from environmental tracers

The methods described in the previous section for simulating the subsurface transport of conservative tracers may also be applied to tritium, using the atmospheric history of ^3H as the recharging boundary condition and the combined concentration of both the parent species and the decay product ($^3\text{H} + ^3\text{He}$) as the calibration target at observation locations. However, the extreme non-linearity of the tritium input signal (**Figure 3.1**), and the associated sensitivity of simulated $^3\text{H} + ^3\text{He}$ to small changes in porosity, make the simulation of $^3\text{H} + ^3\text{He}$ transport poorly suited for the regression methods used in this study (see description of those methods below). While use of an annually-averaged time series reduces the volatility of the ^3H input signal (**Figure 3.1**), the use of this approximation for estimating Upper Chester parameters resulted in physically unrealistic calibrated porosities with very large uncertainty bounds (results not shown).

In contrast, with simultaneous measurement of ^3H and ^3He , the decay relationship between the two species may be used to calculate a $^3\text{H}/^3\text{He}$ age that is independent of uncertainties associated with the input signal (Solomon and Cook, 2000). For the Trit_Hel_Ages method (**Table 3.1**), $^3\text{H}/^3\text{He}$ ages reported by Ekwurzel, et al. (1994) were used as calibration targets, and the simulated age at each observation location was equal to the mean of the TTD generated with backward particle tracking.

3.1.3 - Automated parameter estimation algorithm

For all calibration methods, automated parameter estimation was performed with the inverse modeling software UCODE (Poeter et al., 2005), which uses a modified Gauss-Newton nonlinear method to minimize a least squares objective function. The standard deviations of tracer measurements were used to weight CFC and SF₆ observations during calibration. Weights for the ³H/³He age observations were set equal to 10% of the age. In order to test the sensitivity of the calibration results to the weight assignments we repeated the calibration methods without weighting observations (i.e., with all observations weighted equally). Calibration used a maximum of 50 iterations, which in most cases allowed stabilization of estimated parameter sets even if (in the case of some simultaneous methods) the calibration did not converge according to the designated criteria of less than 1% change in all parameter values. The optimized parameter set was the identified set resulting in the lowest weighted sum of square errors (i.e., not necessarily the set at the 50th iteration).

3.1.4 - Simulation of steady state base-flow age in receiving streams

For each calibrated model (i.e., using the parameter sets estimated by the various calibration methods), the steady state TTD of base-flow in receiving streams was simulated using travel times calculated by MODPATH (Pollock, 2012). Particles were uniformly distributed across the water table and tracked forward to their discharge location. The travel times of discharging particles were then aggregated for each stream (i.e., Morgan Creek and Chesterville Branch) in order to construct the empirical cumulative distribution function (ECDF) and the associated mean and median base-flow age for each stream.

3.2 RESULTS

This section compares the parameter sets estimated by the various calibration methods with particular attention given to the following model characteristics:

1. *Estimated system porosity.* The effective porosity, particularly of the Surficial Aquifer, is a key control on the rate of agricultural contaminant delivery to local streams and is therefore of focal interest to the research questions for which the calibrated model will be used. Note that for the sequential transport calibration methods, the porosity was the only calibrated parameter (i.e., it was estimated after previously calibrating recharge and hydraulic conductivities against head and discharge information). The range of total porosity for sand is typically reported as 0.25 – 0.50 (Freeze and Cherry, 1977). Sanford et al. (2009) examined several cores from further south on the Delmarva Peninsula and found that the total porosity for unconsolidated sands in both surficial and confined aquifers ranged from approximately 0.35 – 0.55. The effective porosity may be as high as 94 percent of the total porosity in unconsolidated sands (Hudak, 1994).
2. *Estimated conductivity of the Confining Unit.* As mentioned above, the conductivity of the confining unit between the Surficial Aquifer and the Confined Aquifer may also be important for explaining agricultural contaminant concentrations in Morgan Creek.
3. *Estimated mean age of base-flow discharge to Morgan Creek.* In surface water systems where contaminant transport is dominated by base-flow (as is the case for nitrate delivery in the Upper Chester), the travel time distribution of base-flow ages discharging to a receiving stream provides an important integrated hydrological

metric for system susceptibility to contamination and the timescale of system recovery from contamination.

3.2.1 - Comparison of calibrated parameter sets

Figure 3.2 shows the normalized simulated versus observed tracer concentrations for each of the calibrated models. **Figure 3.3** shows the parameter estimates and the associated 95% confidence intervals that result from the various sequential and simultaneous calibration methods. The range of calibrated recharge values was similar to estimates from previous studies on the Upper Eastern Shore (Bohlke and Denver, 1995; Green et al., 2008). The range of calibrated hydraulic conductivities for the Surficial and Confined Aquifers are also consistent with field studies and model calibration from elsewhere on the Delmarva Peninsula (Andreasen, 2007; Sanford et al., 2012).

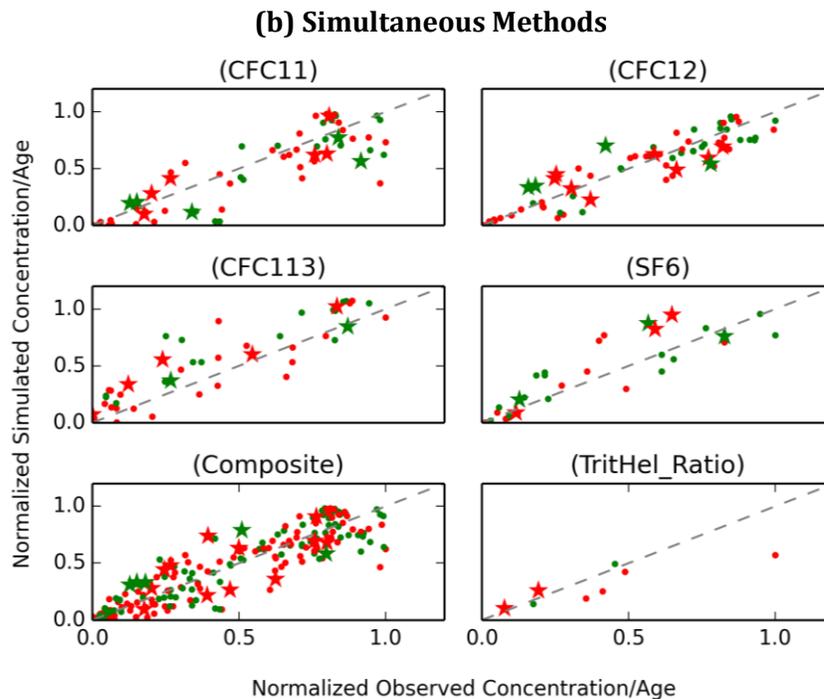
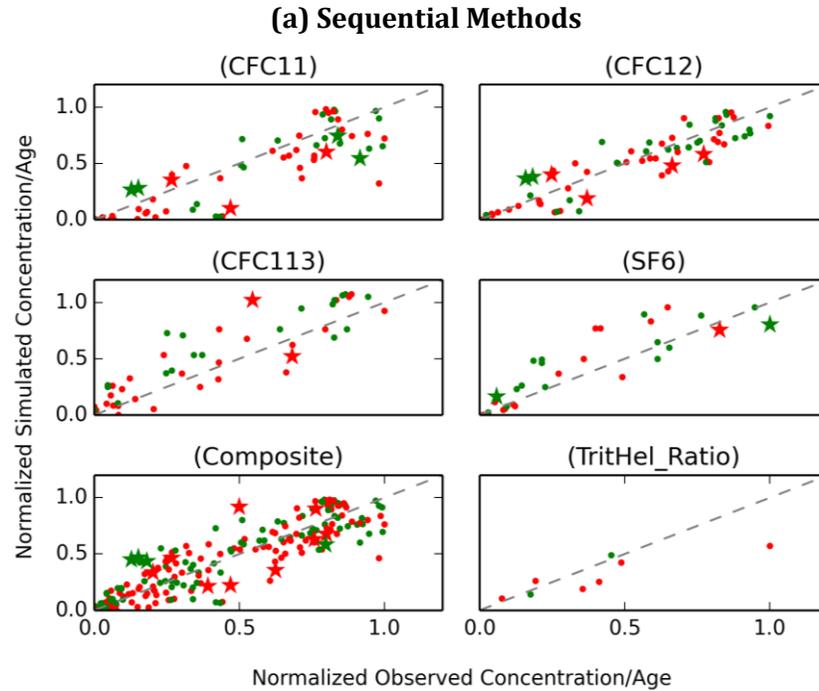


Figure 3.2 Simulated vs. observed tracer concentrations and $^3\text{H}/^3\text{He}$ ages for the calibrated models. Note that each subplot shows the results of a different calibrated model (e.g., the CFC11 plot shows the simulated vs. observed CFC-11 concentrations that are simulated by the model calibrated against CFC-11 concentrations). Red and green markers show observations in Chesterville Branch and Morgan Creek, respectively. The stars show high-leverage observations, with leverage evaluated by Cook's D values reported by UCODE.

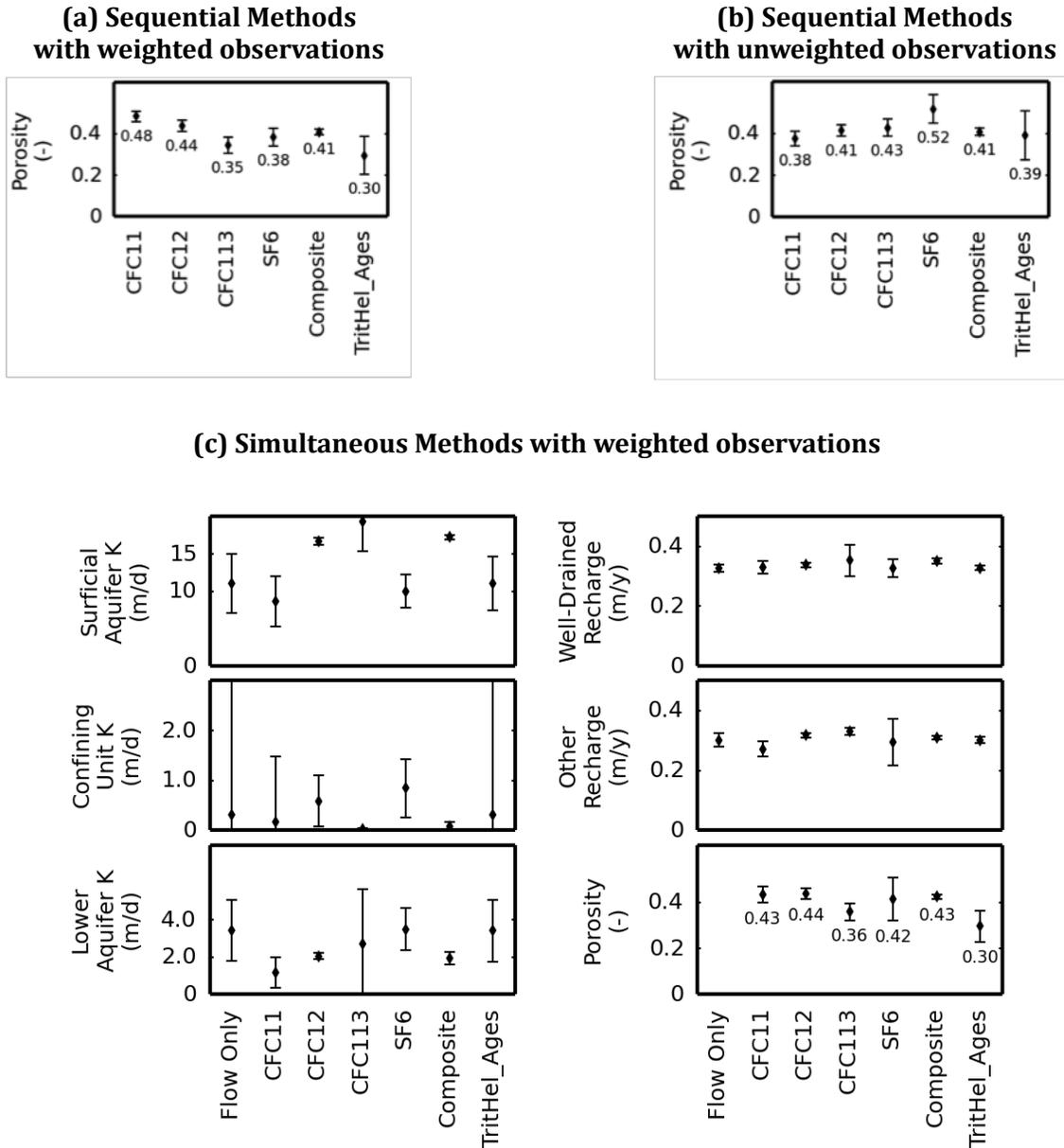


Figure 3.3 Upper Chester parameter estimates using different tracer datasets as calibration targets. Error bars show 95% confidence intervals calculated by UCODE. The individually labeled porosity values specify the calibrated values (i.e., rather than the confidence interval). See **Table 3.1** for method descriptions. ‘Flow Only’ refers to hydraulic conductivity and recharge parameters calibrated against head and discharge data. The sequential methods were each calibrated using a flow field described by the Flow Only parameters.

The estimated porosities are within the range of values expected from Sanford et al. (2009).

The calibrated porosities are more highly variable for the sequential methods than the simultaneous methods. For the weighted sequential methods (i.e., for the case in which

each method used an identical flow field and only estimated the porosity – **Figure 3.3a**) it is notable that the porosities estimated with CFC-11 and CFC-12 data are the highest (0.48 and 0.44, respectively), and that the porosities estimated with CFC-113 and SF₆ are both (i) lower than the CFC-11 and CFC-12 estimates and (ii) similar to one another. The porosity estimated from the Composite tracer dataset (0.41) was equal to the mean of the porosity estimates from the datasets of individual tracer species (**Figure 3.3a**).

The porosity estimate using the Composite dataset did not vary between the weighted and unweighted sequential methods. In contrast, the calibration methods that used individual tracers showed sensitivity to the weighting scheme (compare **Figures 3.2a** and **3.2b**).

Interestingly, when compared to the weighted sequential methods, the unweighted sequential methods resulted in porosity estimates that moved towards (and in some cases past) the Composite porosity estimate. The sequential use of the unweighted SF₆ dataset resulted in the highest porosity estimate of all methods examined in this study.

For both CFC-11 and SF₆, simultaneously calibrating all parameters resulted in a porosity estimate closer to that of the Composite method than did the sequential use of those tracers (compare **Figures 3.2a** and **3.2c**). The inclusion of multiple tracer species (i.e., the 230 transport observations of the Composite dataset) resulted in very small confidence intervals for all parameters and consistent porosity estimates between the weighted sequential, unweighted sequential, and simultaneous cases (0.41, 0.41, and 0.43, respectively). For CFC-11, the lower porosity estimate that resulted from the simultaneous procedure was accompanied by similar reductions in both the calibrated hydraulic conductivities as well as the recharge calibrated for the more poorly drained recharge

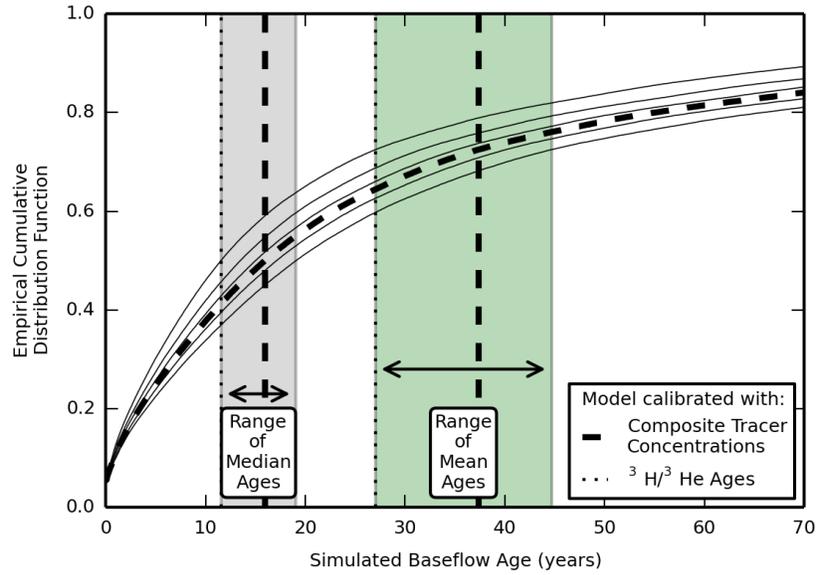
areas. In other words, when compared to the sequential CFC-11 method, the simultaneous use of the CFC-11 transport observations resulted in a model that reduced flow rates and offset those reductions by a reduction in effective porosity (which raises advective velocities). For both the CFC-113 and the Composite methods, the estimated porosity was consistent between the sequential and simultaneous procedures, but the simultaneous procedure resulted in increased recharge rates as well as a significantly reduced calibrated hydraulic conductivity for the Confining Unit. When compared to the weighted sequential method, the simultaneous use of the $^3\text{H}/^3\text{He}$ dataset did not affect the flow parameter estimates, the porosity estimate, or the associated uncertainty bounds (**Figures 3.2a** and **3.2c**). For both of those cases, the use of the $^3\text{H}/^3\text{He}$ ages resulted in the lowest porosity estimate of all methods examined in this study.

The Confining Unit hydraulic conductivity estimate resulting from individual tracer datasets were within an order of magnitude, with the exception of the CFC-113 estimate, which was an order of magnitude less than the estimates made against other datasets (0.03 m/day for CFC-113, versus 0.16, 0.60, and 0.81 m/day, respectively, for CFC-11, CFC-12, and SF₆; see **Figure 3.3c**). Of the methods using individual tracer datasets, the CFC-113 data – with more consistent tracer measurements at the four wells located in the Confined Aquifer – also resulted in the smallest uncertainty bounds. No $^3\text{H}/^3\text{He}$ age data exists for wells located below the Confining Unit; as a result the calibration using $^3\text{H}/^3\text{He}$ ages could not constrain the Confining Unit parameter estimate. The use of all data resulted in an estimate of 0.08 m/day with very small uncertainty bounds (**Figure 3.3c**). In contrast, calibration without tracer information (Flow Only) was unable to constrain its estimate of the confining unit conductivity.

3.2.2 - Comparison of the resulting simulated distributions of base-flow age

The management purposes of the Upper Chester simulation model include predicting the residual catchment response to historic nitrate loading and the future catchment response to nitrate mitigation strategies. As stated above, the estimated distribution of base-flow ages, including descriptors such as the mean and median ages, serve as important indicators of this response. **Figure 3.4** shows the simulated age distributions of base-flow discharging to Morgan Creek for the models calibrated by different tracer datasets. **Table 3.2** specifies the mean age for each of the calibrated models.

(a) Sequential Methods



(b) Simultaneous Methods

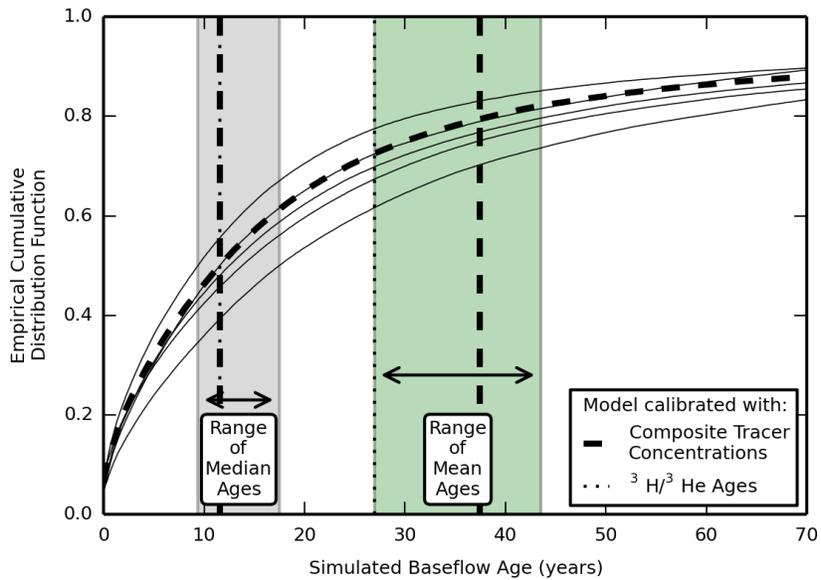


Figure 3.4 Simulated steady state base-flow age distribution for Morgan Creek for the multiple calibrated models. Age distribution simulated at the stream gage (see **Figure 2.1** for gage location). In-stream travel times were assumed to be negligible.

Table 2. Mean and median base-flow age simulated at the Morgan Creek stream gage by the calibrated models. See **Figure 2** for gage locations. ‘Leakance to Confined Aquifer’ calculated as the inflow to the Confined Aquifer divided by the total recharge to the model.

Method Name	<u>Sequential Methods</u>		<u>Simultaneous Methods</u>		
	Mean age	Median age	Mean age	Median age	Leakance to Confined Aquifer
CFC11	44.7	19.1	43.5	13.8	0.20
CFC12	40.1	17.1	38.2	12.5	0.21
CFC113	31.6	13.5	29.8	9.4	0.19
SF6	35.2	15.0	38.6	17.5	0.31
Composite	37.4	15.9	37.4	11.6	0.19
Trit_Hel_Ages	27.0	11.5	27.0	11.5	0.29

The mean base-flow age predicted by the models calibrated with individual tracers are within 20-30% of the mean age predicted by the model calibrated with the Composite dataset. The weighted sequential and simultaneous sets of methods indicate very similar ranges of mean base-flow age (simulated base-flow age for the unweighted sequential methods are not shown). In other words, while the simultaneous calibration methods resulted in different parameter values than the sequential methods (**Figure 3.3**), those differences had little aggregated effect on the predicted mean base-flow age. In contrast, when compared to the sequential methods, the simultaneous use of flow and tracer information generally reduced the predicted median ages by 25-40% (as a result of reducing the Confining Unit conductivity and thereby increasing the contribution of shallower, faster flow paths). All methods indicate that 80-90% of base-flow is younger than 70 years and has therefore recharged since the onset of post-1940 agricultural

intensification (**Figure 3.4**). As expected, larger calibrated porosities result in slower system velocities and therefore generally correlate with older base-flow ages; e.g., for both the sequential and simultaneous sets of methods, the CFC-11 method resulted in the largest estimated porosity (**Figure 3.3**) and the oldest simulated mean base-flow ages (**Figure 3.4, Table 3.2**). When compared to the sequentially-calibrated model, the simultaneous use of the Composite dataset resulted in a lower median age due to the adjustments to the recharge rate and the Confining Unit hydraulic conductivities (**Figure 3.3c**). The SF6 method is the only method for which the mean and median ages were higher using sequential calibration than when using simultaneous calibration.

3.3 DISCUSSION

In this section we examine the factors that contribute to the variability in the parameter sets that result from calibration using different tracer datasets. In the absence of knowledge about the true parameter values, we consider the Composite porosity estimates to be the most reliable due to their derivation from the largest dataset and the smaller confidence intervals that result from this dataset when compared to the confidence intervals associated with other species (**Figure 3.3**). However, the Composite dataset is subject to the same potential impacts of any errors in excess air estimates or transport model formulation as its constituent species. (It should also be noted that because the confidence intervals shown in **Figure 3.3** do not integrate these sources of uncertainty they should consequently be considered to be minimum confidence intervals). In addition, for the weighted calibration methods, the Composite parameter estimates are markedly different than the Trit_Hel_Ages estimates (**Figure 3.3**). As previously described, the

$^3\text{H}/^3\text{He}$ ages are independent of the atmospheric history and therefore less subject to uncertainty in the input signal. As such, the parameter estimates that result from the Trit_Hel_Ages method provide an important provisional benchmark for the accuracy of the estimates made by other methods. However, due to the relatively small number of observations (**Table 2.1**), the estimates made with the Trit_Hel_Ages method are characterized by large confidence intervals and as such provide limited evidence.

The sequential methods each used the same flow field but generated a range of porosity estimates. Furthermore, while the simultaneous calibration methods resulted in more consistent porosity estimates, the various methods did not result in consistent adjustments to other model parameters (**Figure 3.3**). The variability of parameter estimates and the non-overlapping confidence intervals between methods suggest the following potential explanations:

1. The CFC-11 and CFC-12 datasets are impacted by some form of mass loss or transport delay and are therefore not well-described by the conservative advective transport assumed by our simulation model.
2. The SF₆ dataset is a function of a higher excess air amount than the value assumed by our simulation model (2 cm³/L).
3. The spatial distribution of measurement uncertainty and the associated calibration weights is not consistent between species, such that the different calibration methods were leveraged by large residuals at different observation locations.

These interpretations are considered in more detail below.

3.3.1 - Investigating the potential impact of tracer retardation or degradation

As discussed earlier, because the atmospheric CFC input signal was increasing for most of the recharge times sampled by this dataset, subsurface degradation of tracer mass or retardation of tracer transport would make apparent ages artificially old (**Figure 3.1**). For a model that assumes conservative, advection-only transport, calibration against artificially old age information would result in elevated porosity estimates.

It is not expected that CFC retardation would affect calibration results at the Upper Chester study site. CFC sorption is commonly associated with the fraction of organic carbon (f_{oc}) in groundwater systems (Plummer and Busenberg, 2000), but the Upper Chester has generally low organic content, except localized surface areas of higher organic content in small depressional wetlands that have been converted to agricultural use (Denver et al., 2014). Furthermore, in previous studies CFC-113 was more affected by sorption than CFC-11 and CFC-12 (Choung and Allen-King, 2010; cf. Cook et al., 1995, and Bauer, 2001), whereas in the present study CFC-113 appears to be less impacted by mass loss or delay than CFC-11 and CFC-12.

CFC-11 and CFC-12 apparent ages are older than the CFC-113 and SF₆ apparent ages for a few deeper measurements (**Figure 3.5**), which may suggest some preferential degradation of CFC-11 and CFC-12. The reducing conditions required for CFC degradation have been observed in portions of the Upper Chester (Böhlke and Denver, 1995). For example, **Figure 3.6a** shows a well nest in which, for the deepest well in the Chesterville Branch catchment (Be159), low O₂ and low nitrate indicate the potential impact of reducing conditions on contributing transport pathways, and the CFC-12 apparent ages are notably older than the CFC-113 apparent age. However, the lack of a CFC-113 age gradient with depth, and the

dissimilarity of CFC-113 ages to both SF₆ and CFC-11 ages, may indicate CFC-113 contamination. Further, previous investigators have observed that CFC-12 is more resistant to degradation than CFC-113 (Sebol et al., 2007; Oster et al., 1996). Thus, if CFC degradation is an important component of tracer transport in the Upper Chester we would expect that the porosity estimate using CFC-12 would be lower than the estimate using CFC-113, which was not the case (**Figures 3.2a** and **3.2b**). Finally, analysis of observation influence using Cook's D as a metric of leverage does not suggest that measurement locations at which CFC-11 and CFC-12 ages are older than CFC-113 or SF₆ ages (**Figure 3.5**) have more influence than other observations on the porosity estimate (results not shown). In sum, there is little evidence from the porosity estimate that CFC degradation or retardation impacted the calibration results or should be included in the transport model.

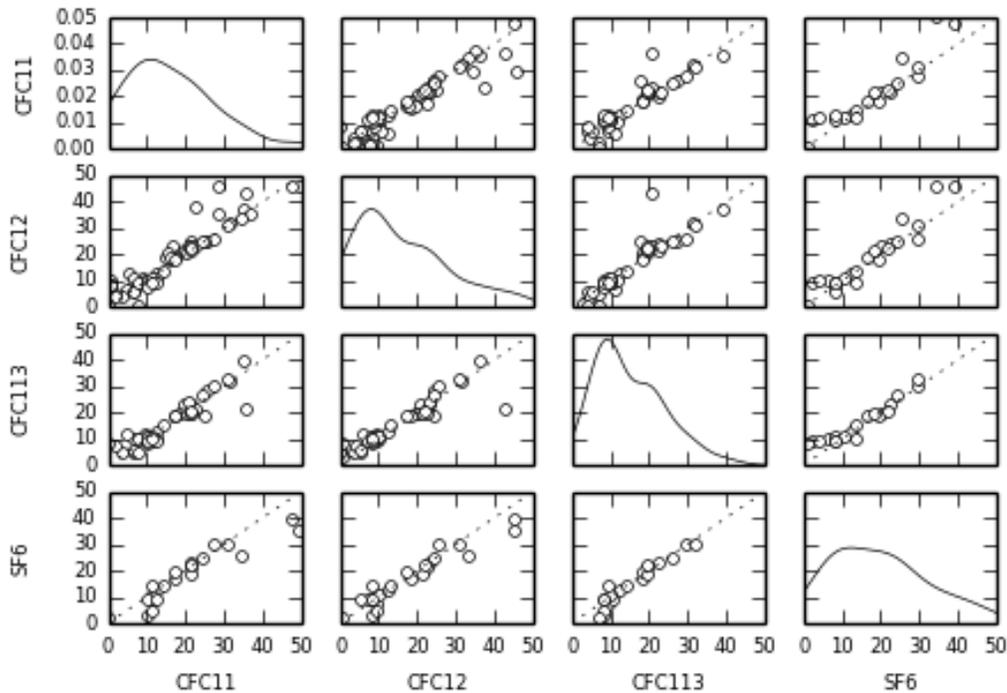


Figure 3.5 Scatter-matrix of apparent piston-flow ages interpreted from Upper Chester tracer measurements. Integer values on axes are apparent age in years. Plots on diagonal show distribution of apparent ages for each tracer (represented by kernel density estimate).

The variations in the conductivity estimate for the Confining Unit were due to small variations in the tracer concentrations sampled below the Confining Unit. For the four wells located in the Confined Aquifer, the sampled waters were anoxic (data from Bachman et al., 2002) and the tracer concentrations were either zero or near the limits of the measureable tracer amounts (e.g., apparent recharge dates between 1945 and 1965; see **Figure 3.1**). Each of these four wells contained a small amount of SF₆ (0.02 to 0.07 fmol/L; see data for wells Be189, Be200, Be210, and Be211 in Busenberg and Plummer, 2000) while, in contrast, only one of the wells in the Confined Aquifer contained CFC-113. In the absence of any subsurface sources of background SF₆ this may suggest conservative SF₆ transport but the preferential degradation of CFC-113 at some sites due to reducing

conditions in the silt and clay deposit (cf. Hinsby, 2007). However, given the uncertainty associated with sampling very small tracer concentrations (Busenberg and Plummer, 2000) and the mixed evidence of the CFC data (e.g., the CFC-11 apparent ages are younger than the SF₆ apparent ages at Be210) it is not clear that there is a difference in transport mechanisms between the different tracer species. More clear is the evidence of hydraulic connectivity between the Surficial and Confined Aquifer since each of the four wells contained a measurable amount of at least one tracer species. The small confidence intervals on the Confining Unit hydraulic conductivity estimated by the Composite method (**Figure 3.3c**) particularly shows that tracer concentrations in the Confined Aquifer can be explained by transport from the Surficial Aquifer through the silt and clay deposit (i.e., rather than from some more distant recharge location).

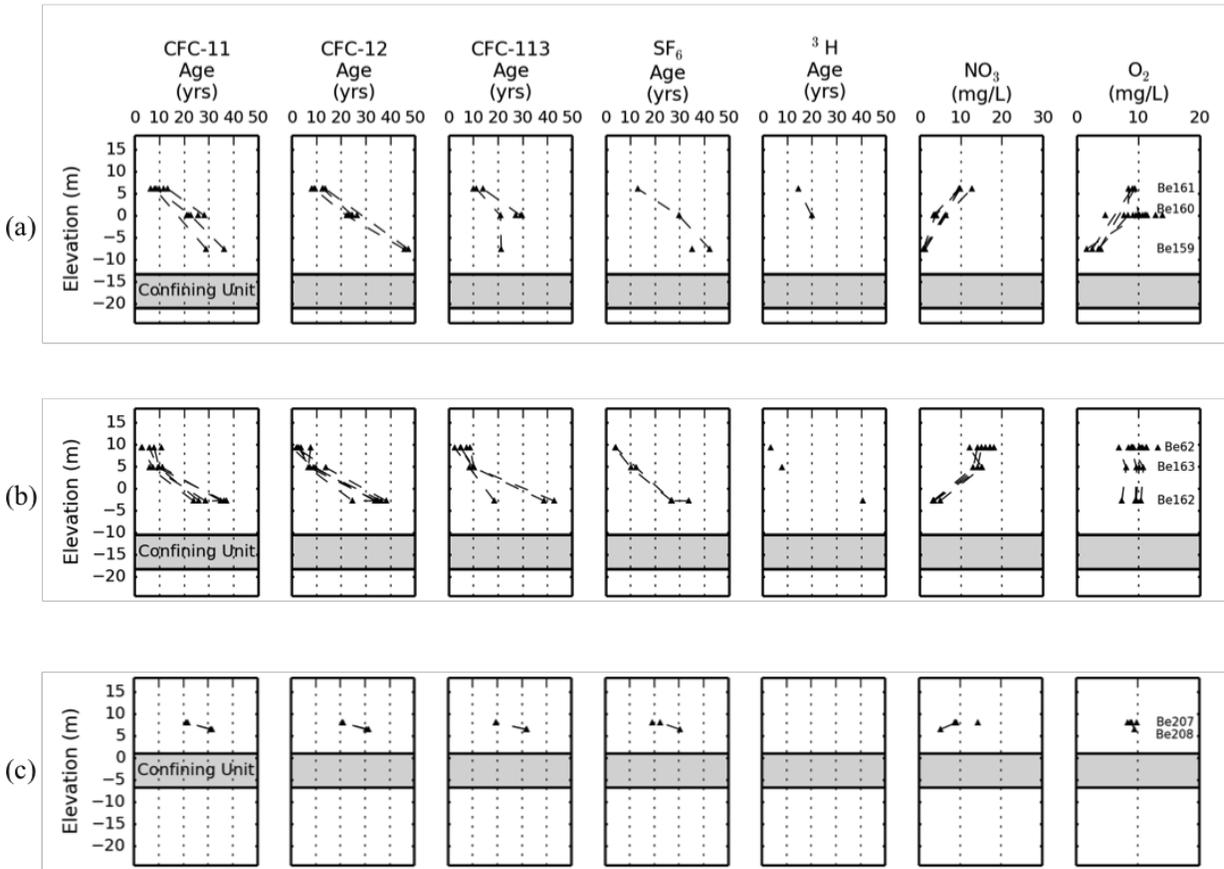


Figure 3.6 Tracer apparent ages and nitrate and dissolved oxygen concentrations at two well nests in the Chesterville Branch subcatchment (**a** and **b**) and one well nest in the Morgan Creek subcatchment (**c**). Identifiers for the individual wells in each nest are listed in the far right panel. The dashed lines connect measurements made on the same date.

3.3.2 - Investigating the potential impact of errors in model boundary conditions

As with CFC-113, sequential calibration using SF₆ resulted in a lower porosity estimate than the estimate from the CFC-11 or CFC-12 method (**Figure 3.3a**). The porosity estimate using the simultaneous SF₆ method was marginally lower than CFC-11 or CFC-12 but with much larger uncertainty (**Figure 3.3b**). SF₆ is not known to degrade; it would therefore be expected that, in the event of CFC degradation, calibration against SF₆ data would result in lower porosities (i.e., faster advective velocities and younger apparent ages) than

calibration against CFC data. However, while SF₆ observations are less susceptible to mass removal than CFC observations, SF₆ simulation is much more sensitive to excess air assumptions than is CFC simulation because of their relative solubilities (Goddy et al., 2006). In order to examine the sensitivity of the porosity estimate to the excess air assignment we performed additional sequential calibrations using SF₆ and CFC-113 and assigning 1, 3 and 4 cm³/L excess air as inputs to the Henry's Law conversions (noting that Busenberg and Plummer, 2000, found evidence of excess air values as high as 3 cm³/L on the Delmarva Peninsula). The SF₆ porosity estimate was highly sensitive and the CFC-113 estimate was not (**Figure 3.7**). In other words, it could be the case that our general assumption (2 cm³/L for all recharging tracers) underestimates the actual excess air, such that our porosity calibrated with SF₆ is in erroneously low. In this case, the SF₆ porosity estimates would be more consistent with the CFC-11 and CFC-12 estimates; however, this would not explain the lower CFC-113 porosity estimate.

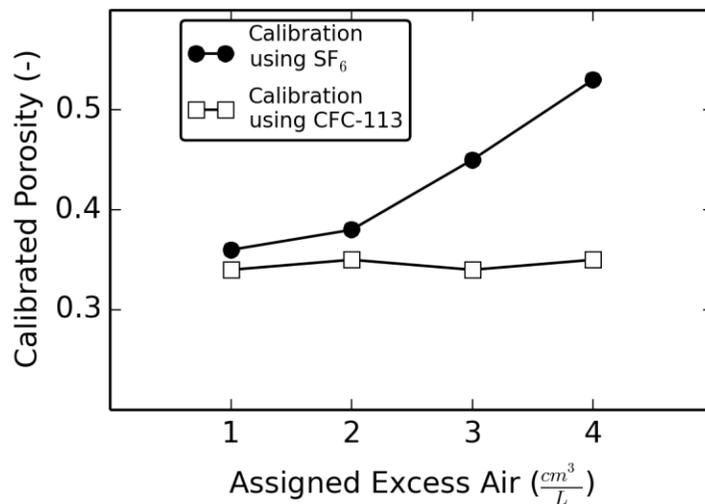


Figure 3.7 Sensitivity of calibrated porosity to the value assumed for recharged excess air.

3.3.3 - Investigating the potential impact of measurement uncertainty and calibration weights

Much of the distinction between the sequentially-estimated parameter sets may be explained by examining the most influential observations in each tracer dataset and comparing the weighted calibration results to the unweighted calibration results for each method (**Figure 3.3**). For example, the CFC-11 porosity estimate is strongly influenced by the observation at Be208, a well located in the Morgan Creek headwaters. The reported measurement standard deviation for Be208 is very low (1%, from Busenberg and Plummer, 2000), which results in this observation being heavily weighted during the regression.

Figure 3.6c shows the steep age and nitrate gradients at this near-stream location; these gradients reflect the convergence of disparate flow paths near their discharge in Morgan Creek. We tested the sensitivity of the CFC-11 calibration results to the confidence assigned at well Be208 and found that by simply increasing the Be208 measurement uncertainty from 1% to 2%, the porosity estimate using CFC-11 as calibration targets dropped from 0.48 to 0.44. The porosity estimate using the unweighted CFC-11 dataset reflects this same sensitivity to calibration weights (**Figure 3.3b**).

Similarly, the low porosity value estimated using CFC-113 reflects the influence of Be163, an observation location in the Chesterville Branch headwaters with a low (< 1%) CFC-113 measurement uncertainty (Busenberg and Plummer, 2000). The lack of age gradient and nitrate gradient at Be62 and Be163 (**Figure 3.6b**) indicates locally-complex hydrology, including ponding and focused recharge from a swale at the land surface (oral communication, J.K. Böhlke). The CFC-113 apparent age is consistent with the ages derived from other tracers (**Figure 3.6b**); however, unlike CFC-113, the larger sample standard deviations for measurements associated with other tracer species reflect the uncertainty in

this area of the model. By relaxing the relative importance of this observation, the porosity estimate using CFC-113 data increases from 0.35 to 0.43 (**Figures 3.2a** and **3.2b**), closer to the estimate derived from the Composite dataset.

In sum, the spatial distribution of measurement uncertainty is not consistent across species, and this variation in confidence explains much of the variation in calibration results. It may be further noted that neither the measurement uncertainty nor the simulated residuals are spatially biased (**Figure 3.2**). For CFC-11, CFC-12, and CFC-113, damping the influences of individual measurements by using an unweighted regression resulted in porosity estimates similar to the estimate derived from the collective evidence of the weighted Composite dataset.

3.4 CONCLUSIONS

CFC, SF₆, and ³H/³He datasets were used individually and in combination to calibrate a groundwater flow and transport model for an agricultural catchment on the Delmarva Peninsula. Sequential calibrations, in which individual tracer species were used with an identical flow model to calibrate the effective porosity, resulted in a wide range of porosity estimates. Biases due to un-modeled transport processes (e.g., degradation or retardation) are not evident in the datasets and do not appear to be responsible for the differences in calibration results. Instead, the differences between calibrated models can be largely explained by the variation in the spatial distribution of measurement uncertainty for the different tracer species. The use of a composite dataset of multiple tracer species was judged to provide the most accurate parameter estimates; for the calibrations performed with individual tracer species, small adjustments to the calibration weights at a few

locations resulted in parameter estimates consistent with those derived from the weighted composite dataset and suggested that the different species are consistently described by conservative advective transport. Sensitivity analysis did show that errors in the assignment of recharging excess air could also explain the inconsistency between the porosity estimate using SF₆ and the porosity estimates using CFC-11 and CFC-12. Simulated base-flow age using models calibrated by individual tracer species varied by +/- 20% from the age simulated by a composite dataset of multiple species. The use of the tracer dataset allowed estimation of the hydraulic conductivity for a semi-confining unit that was not previously well understood, showing that the confined aquifer does receive recharge from the surficial aquifer.

This work demonstrates that while tracer data can provide necessary supplemental information for the calibration of flow and transport models, the use of data from a single tracer or from a small tracer set may be insufficient to fully interpret the information content of the tracers. While use of multiple weighting schemes with datasets of individual tracers may be important for suggesting the range of possible models, the combined use of multiple tracers is less sensitive to the weighting scheme and results in more confident parameter estimates.

3.5 REFERENCES

Andreasen, D. C., Achmad, G., Staley, A. W., & Hodo, R. M. (2007). Hydrogeologic framework of the Maryland coastal plain. *Maryland Geological Survey Report, Baltimore, MD.*

Bachman, L. J., Krantz, D. E., & Bohlke, J. (2002). *Hydrogeologic Framework, Ground-water Geochemistry, and Assessment of Nitrogen Yield from Base Flow in Two Agricultural Watersheds, Kent County Maryland.* Environmental Protection Agency.

- Bauer, S., Fulda, C., & Schafer, W. (2001). A multi-tracer study in a shallow aquifer using age dating tracers ^3H , ^{85}Kr , CFC-113 and SF_6 - indication for retarded transport of CFC-113. *Journal of Hydrology*, 248(14), 34.
- Böhlke, J. K., & Denver, J. M. (1995). Combined use of groundwater dating, chemical, and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, Atlantic coastal plain, Maryland. *Water Resources Research*, 31(9), 2319-2339.
- Busenberg, E., & Plummer, L. N. (2000). Dating young groundwater with sulfur hexafluoride: Natural and anthropogenic sources of sulfur hexafluoride. *Water Resources Research*, 36(10), 3011-3030.
- Choung, S., & Allen-King, R. M. (2010). Can chlorofluorocarbon sorption to black carbon (char) affect groundwater age determinations? *Environmental Science and Technology*, 44(12), 4459-4464.
- Cook, P.G., Solomon, D.K., Plummer, L.N., Busenberg, E., & Schiff, S.L. (1995). Chlorofluorocarbons as tracers of groundwater transport processes in a shallow, silty sand aquifer. *Water Resources Research*, 31(3), 425-434.
- Cook, P. G., & Böhlke, J.K. (2000). Determining timescales for groundwater flow and solute transport. In *Environmental tracers in subsurface hydrology* (pp. 1-30). Springer.
- Cook, P. G., & Herczeg, A. L. (2000). *Environmental tracers in subsurface hydrology*. Springer.
- Denver, J.M., Ator, S.W., Lang, M.W., Fisher, T.R., Gustafson, A.B., Fox, R. XXX (2014). Nitrate fate and transport through current and former depressional wetlands in an agricultural landscape, Choptank Watershed, Maryland, United States. *Journal of Soil and Water Conservation*, 69(1), 1-16.
- Dunkle, S. A., Plummer, L. N., Busenberg, E., Phillips, P. J., Denver, J. M., Hamilton, P. A., Michel, R.L., & Coplen, T. B. (1993). Chlorofluorocarbons (CCl_3F and CCl_2F_2) as dating tools and hydrologic tracers in shallow groundwater of the Delmarva Peninsula, Atlantic Coastal Plain, United States. *Water Resources Research*, 29(12), 3837-3860.
- Ekwurzel, B., Schlosser, P., Smethie, W. M., Plummer, L. N., Busenberg, E., Michel, R. L., Weppernig, R., & Stute, M. (1994). Dating of shallow groundwater: Comparison of the transient tracers $^3\text{H}/^3\text{He}$, chlorofluorocarbons, and ^{85}Kr . *Water Resources Research*, 30(6), 1693-1708.
- Freeze, R. A., & Cherry, J. A. (1977). *Groundwater*. Prentice-Hall.
- Goody, D. C., Darling, W. G., Abesser, C., & Lapworth, D. J. (2006). Using chlorofluorocarbons (CFCs) and sulphur hexafluoride (SF_6) to characterise groundwater movement and residence time in a lowland chalk catchment. *Journal of Hydrology*, 330(1-2), 44-52.
- Happell, J. D., Price, R. M., Top, Z., & Swart, P. K. (2003). Evidence for the removal of CFC-11, CFC-12, and CFC-113 at the groundwater--surface water interface in the everglades. *Journal of Hydrology*, 279(1), 94-105.
- Hill, M. C. (2006). The practical use of simplicity in developing ground water models. *Groundwater*, 44(6), 775-781.

- Hill, M. C., & Tiedeman, C. R. (2006). *Effective groundwater model calibration: with analysis of data, sensitivities, predictions, and uncertainty*. John Wiley & Sons.
- Hinsby, K., Højberg, A. L., Engesgaard, P., Jensen, K. H., Larsen, F., Plummer, L. N., & Busenberg, E. (2007). Transport and degradation of chlorofluorocarbons (CFCs) in the pyritic Rabis Creek aquifer, Denmark. *Water Resources Research*, 43(10).
- Hudak, P. F. (1994). Effective porosity of unconsolidated sand; estimation and impact on capture zone geometry. *Environmental Geology*, 24(2), 140-143.
- Hunt, R. J., Feinstein, D. T., Pint, C. D., & Anderson, M. P. (2006). The importance of diverse data types to calibrate a watershed model of the Trout Lake Basin, Northern Wisconsin, USA. *Journal of Hydrology*, 321(1), 286-296.
- Kazemi, G. A., Lehr, J. H., & Perrochet, P. (2006). *Groundwater age*. John Wiley & Sons.
- Long, A. J., & Putnam, L. D. (2009). Age-distribution estimation for karst groundwater: Issues of parameterization and complexity in inverse modeling by convolution. *Journal of Hydrology*, 376(3), 579-588.
- Maloszewski, P., & Zuber, A. (1982). Determining the turnover time of groundwater systems with the aid of environmental tracers: 1. Models and their applicability. *Journal of Hydrology*, 57(3), 207-231.
- Mattle, N., Kinzelbach, W., Beyerle, U., Huggenberger, P., & Loosli, H. H. (2001). Exploring an aquifer system by integrating hydraulic, hydrogeologic and environmental tracer data in a three-dimensional hydrodynamic transport model. *Journal of Hydrology*, 242(3), 183-196.
- McCallum, J. L., Engdahl, N. B., Ginn, T. R., & Cook, P. (2014). Nonparametric estimation of groundwater residence time distributions: What can environmental tracer data tell us about groundwater residence time? *Water Resources Research*, 50(3), 2022-2038.
- Neuzil, C. E. (1994). How permeable are clays and shales? *Water Resources Research*, 30(2), 145-150.
- Oster, H., Sonntag, C., & Münnich, K. O. (1996). Groundwater age dating with chlorofluorocarbons. *Water Resources Research*, 32(10), 2989-3001.
- Plummer, L. N., & Busenberg, E. (2000). Chlorofluorocarbons. In *Environmental tracers in subsurface hydrology* (pp. 441-478). Springer.
- Plummer, L. N., Busenberg, E., & Cook, P. G. (2006). Use of chlorofluorocarbons in hydrology. *International Atomic Energy Agency*.
- Poeter, E.P., Hill, M.C., Banta, E.R., Mehl, Steffen, Christensen, Steen. (2005). *UCODE_2005 and six other computer codes for universal sensitivity analysis, calibration, and uncertainty evaluation*. US Department of the Interior, US Geological Survey.

- Pollock, D. W. (2012). *User guide for MODPATH version 6: A particle tracking model for MODFLOW*. US Department of the Interior, US Geological Survey.
- Portniaguine, O., & Solomon, D. K. (1998). Parameter estimation using groundwater age and head data, Cape Cod, Massachusetts. *Water Resources Research*, 34(4), 637-645.
- Reilly, T. E., Plummer, L. N., Phillips, P. J., & Busenberg, E. (1994). The use of simulation and multiple environmental tracers to quantify groundwater flow in a shallow aquifer. *Water Resources Research*, 30(2), 421-433.
- Sanford, W., Pope, J., & Nelms, D. (2009). *Simulation of groundwater-level and salinity changes in the Eastern Shore, Virginia*. US Department of the Interior, US Geological Survey.
- Sanford, W. (2011). Calibration of models using groundwater age. *Hydrogeology Journal*, 19(1), 13-16.
- Sanford, W. E., Nelms, D. L., Pope, J. P., & Selnick, D. L. (2011). *Quantifying components of the hydrologic cycle in virginia using chemical hydrograph separation and multiple regression analysis*. US Department of the Interior, US Geological Survey.
- Sanford, W.E., Pope, J.P., Selnick, D.L., & Stumvoll, R.F. (2012). *Simulation of groundwater flow in the shallow aquifer system of the Delmarva Peninsula, Maryland and Delaware*. US Department of the Interior, US Geological Survey.
- Sebol, L. A., Robertson, W. D., Busenberg, E., Plummer, L. N., Ryan, M. C., & Schiff, S. L. (2007). Evidence of CFC degradation in groundwater under pyrite-oxidizing conditions. *Journal of Hydrology*, 347(1), 1-12.
- Solomon, D. K., & Cook, P. G. (2000). ^3H and ^3He . In *Environmental tracers in subsurface hydrology* (pp. 397-424). Springer.
- Szabo, Z., Rice, D. E., Plummer, L. N., Busenberg, E., Drenkard, S., & Schlosser, P. (1996). Age dating of shallow groundwater with chlorofluorocarbons, tritium/helium, and flow path analysis, Southern New Jersey Coastal Plain. *Water Resources Research*, 32(4), 1023-1038.
- Turnadge, C., & Smerdon, B. D. (2014). A review of methods for modeling environmental tracers in groundwater: Advantages of tracer concentration simulation. *Journal of Hydrology*, 519, 3674-3689.
- Zuber, A., Witczak, S., Rózański, K., Śliwka, I., Opoka, M., Mochalski, P., Kuc, T., Karlikowska, J., Kania, J., Jackowicz-Korczynski, M., & Duliński, M. (2005). Groundwater dating with ^3H and SF_6 in relation to mixing patterns, transport modelling and hydrochemistry. *Hydrological Processes*, 19(11), 2247-2275.

Chapter 4: Simulating Nitrate Removal Mechanisms in an Agricultural Catchment with Contrasting Nitrate Base-flow Concentrations in Subcatchment Streams²

4.0 INTRODUCTION

As point source controls on nutrient discharges to receiving waters have improved in recent decades, non-point source loading has become an increasingly important contributor to surface water quality problems. In coastal estuaries and bays, where primary productivity is typically nitrogen limited (Vitousek et al., 1997; Schindler and Vallentyne, 2008), algal blooms and dissolved oxygen deficits may persist due to the ongoing discharge of groundwater nitrates that have accumulated in surficial aquifers during the past century (Puckett et al., 2011). For example, studies have estimated that roughly half of nitrogen inputs to the Chesapeake Bay are transported via recharge to the subsurface and subsequent base-flow discharge to Bay tributaries or direct groundwater discharge to the Bay itself (Lindsey et al., 2003). On the agriculturally-intensive Eastern Shore of the Chesapeake Bay, the fraction of nitrate transported via groundwater may be as high as 70% (Ator and Denver, 2012).

While numerous studies have documented the linkages between agricultural nitrogen loads and surface water degradation (e.g., Böhlke and Denver, 1995; Rupert, 2008), the magnitude and timing of load reductions due to agricultural best management practices (BMPs) that are intended to reduce groundwater nitrates and improve surface water

² This chapter is being prepared for submission to *Environmental Science and Technology*.

quality at the catchment scale are not well understood (Meals et al., 2010). In gaining streams, in-stream nitrate concentrations are an aggregate measure of the transport due to many converging groundwater flow paths, such that the in-stream nitrate measurement at any point in time is a composite response to spatially- and temporally-distributed land use management and landscape processes. As a result, while loading reductions due to BMPs have been documented at laboratory and field scales (e.g., Staver and Brinsfield, 1998), the anticipated effects of these practices are often difficult to detect at the outlets of agricultural watersheds in which they have been widely implemented (Osmond et al., 2012). For example, Sutton et al. (2009) concluded from 20 years of base-flow nitrate measurements in a Maryland agricultural catchment that nitrate discharge did not decrease in spite of catchment-wide implementation of agricultural BMPs. They suggested that, while BMP implementation may have slowed the rate of increase of in-stream nitrate concentrations, the potential impact of BMPs could not be separated from other spatially- and temporally-distributed processes affecting in-stream loads. Gitau et al. (2010) similarly compared trends in stream nitrate concentrations with a detailed land use and BMP-implementation history for an agricultural catchment in Arkansas. They concluded that the impact of BMP implementation could not be disentangled from the contributing effects of other land use changes.

These studies illustrate two fundamental challenges to detecting water quality impacts from agricultural BMPs such as cover crop implementation or nutrient management plans. First, the failure to detect improvement may be because those improvements are not yet evident given the length of groundwater travel times and the associated lag between land

surface action and surface water response (Sanford and Pope, 2013). In this case, the use of groundwater flow and transport modeling or lumped parameter methods to characterize the travel time distribution (TTD) of base-flow discharge may be needed in order to estimate the likely timescale of system response and the concomitant timescale of monitoring programs required to discern that response. Second, the identification of water quality responses due to management improvements – or the explanation of disparate responses across agricultural catchments subject to the same regime of improvements – requires accounting for a variety of confounding environmental factors. These confounding factors, which may vary widely between catchments, include background rates of nitrogen removal at the land surface, in the aquifer, or in the receiving stream. In this case, analysis or prediction of the effects of land surface conservation practices may require prior characterization of nitrate transport processes and removal mechanisms.

In the current Chesapeake Bay Watershed Model (WSM) of the Chesapeake Bay Program, assigned BMP efficiencies are derived from small-scale experiments and perturbed in order to account for the uncertainty due to differences in setting (e.g., topography, geology), maintenance, and operation between the experimental site and the application site (USEPA, 2010). It has been recognized that the lack of field-scale and catchment-scale measures for BMP nitrate-reduction performance is a significant source of uncertainty for the Bay WSM predictions and the regulatory decisions that are based on those predictions (Scientific and Technical Advisory Committee, 2005). In order to better assess the catchment-scale water quality benefit of agricultural management improvements in the Chesapeake Bay basin, the U.S. Department of Agriculture (USDA) has targeted three agriculturally-intensive

watersheds (i.e., characterized by high rates of agricultural nutrient export) for increased BMP implementation and associated water quality monitoring. BMP implementation was initially funded through allocations from the Chesapeake Bay Watershed Initiative, which was included in the 2008 Farm Bill (Science and Technical Advisory Committee, 2010). The U.S. Geological Survey (USGS) is coordinating monitoring and analysis of these targeted watersheds [Science and Technical Advisory Committee, 2010; Nelson and Spies, 2013]. The current phase of these watershed studies effectively began in 2010 and includes: (i) land use data collection (including reconstruction of recent nutrient inputs and conservation efforts); (ii) accelerated, state- and federally-funded implementation of BMPs; and (iii) more intensive watershed monitoring (Nelson and Spies, 2013). In some cases, including the Upper Chester watershed that is examined in this study and described in more detail below, the focus watersheds have a prior history of surface water and groundwater data that was collected in the course of individual studies or at longer term monitoring sites (e.g., USGS surface water gages).

The objective of the study described in this chapter is to use groundwater modeling to link estimated historic land surface inputs to subsurface and base-flow nitrate concentrations in order to characterize the controls on nitrate removal in the Upper Chester (MD) targeted watershed (**Figure 2.1**). The study makes an important contribution to the catchment-scale examination of subsurface nitrate transport and catchment removal processes. Though these phenomena have been widely investigated, there have been few fully distributed, three-dimensional modeling studies of nitrate transport and removal in catchments with nitrogen removal rates that are highly spatially-variable, as is the case in

the Upper Chester. The Morgan Creek and Chesterville Branch subcatchments (**Figure 2.1**) have similar land-use histories, soil types, and stream discharge rates but widely different in-stream nitrate levels. Concentrations at the Morgan Creek USGS gage hovered between 2 and 3 mg NO₃-N/L for the duration of its sampling history; in contrast, concentrations at the Chesterville Branch gage have increased from 4-6 mg/L in the early 1990's and currently persist near 10 mg/L (**Figure 4.1**). Field data suggest several potential reasons for this variability though no catchment scale studies have integrated the available data in order to test or quantify their relative contributions. For example, Bohlke and Denver (1995) reported elevated nitrate $\delta^{15}\text{N}$ levels, excess dissolved N₂, and indicators of pyrite reduction at the glauconitic confining unit that outcrops at the lower reaches of Morgan Creek (**Figure 2.1b**). Microbial reduction of nitrate to N₂ prefers $\delta^{14}\text{N}$, leaving an enriched nitrate $\delta^{15}\text{N}$ value in the remaining nitrate and excess N₂ in the water sample; therefore, this evidence from lower Morgan Creek indicates that some fraction of potential nitrate loads from the surficial aquifer may be removed by denitrification due to contact with the Confining Unit before discharge as base-flow. The Confining Unit is tens of meters deeper beneath Chesterville Branch (**Figure 2.2**); as a result, fewer discharging base-flow paths encounter the Confining Unit and base-flow concentrations consequently show no evidence of denitrification.

Additionally, Bachman et al. (2002) observed increasing silica concentrations in a downstream direction on Morgan Creek. Groundwater silica concentrations elsewhere on the Delmarva Peninsula have been shown to positively correlate with tritium-derived groundwater ages (Clune and Denver, 2012); therefore, increased silica in the lower

reaches of Morgan Creek may indicate the dilution of agricultural nitrates with older, higher silica, nitrate-free water that reaches the stream from the confined aquifer. Sediment cores in lower Morgan Creek show an abrupt change in the elevation of the Confining Unit and thus suggest a discontinuity that could allow influx of older groundwater from the Confined Aquifer (Bachman et al., 2002).

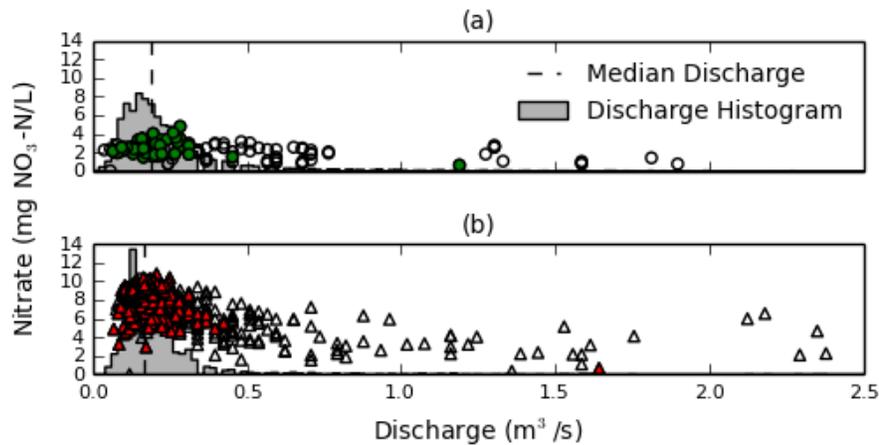


Figure 4.1 Observed stream nitrate concentrations at the (a) Morgan Creek and (b) Chesterville Branch gages (see **Figure 2.1** for gage locations). Hollow markers show all nitrate observations; filled markers show those observations for which the associated base-flow component was greater than 90% of the total discharge (see Section 4.1.3).

Finally, because the Morgan Creek stream channel is downcut into the outcropping Aquia Confining Unit, direct groundwater discharge through the streambed is limited. Groundwater instead emerges through seeps at the edge of a near-stream floodplain and travels to the main channel via small rivulets and sheetflow. In a study of the lower reach, Duff et al. (2008) found nitrate concentrations in the groundwater, rivulets, and stream to be 10.2 mg-N/L (median value), 5 mg-N/L (mean value), and 2.8 mg-N/L (mean value), respectively. This suggests that riparian nitrogen removal may account for a significant

portion of load reductions to Morgan Creek. There is no equivalent field data describing nitrate processing at or near the Chesterville Branch stream channel.

This study uses the steady state flow model calibrated in Chapter 3 of this dissertation to derive travel time distributions (TTDs) for each observation location in a large dataset of subsurface and surface water nitrate measurements. Identifying these controls on nitrate transport and removal will allow subsequent disentanglement of the background nitrate removal from water quality improvements that may be due to management actions. As such, this study provides an important baseline assessment of the Upper Chester targeted watershed in support of future efforts to detect any water quality benefits from BMP implementation.

4.1 METHODS

4.1.1 - Simulation of nitrate transport

The time-variable nitrate concentrations at observation wells and at the Morgan Creek and Chesterville Branch stream gages were simulated by convolving the steady-state travel time distribution (TTD) at each location with the time-variable nitrate input function (described in Section 4.1.2) minus reductions due to nitrate removal mechanisms. As described in Chapter 3 of this dissertation, previous studies have found solute transport in the study area to be strongly advection-dominated, and TTDs for this study were generated without dispersion using MODPATH (Pollock, 2012). TTDs at observation wells were generated by backward tracking particles from each well location to the land surface and then aggregating the travel times of all pathlines into a distribution. In order to account for

the uncertainty associated with the particle interpolation scheme and the spatial distribution of loading, as well as small scale mixing near the observation wells, backward particles were tracked from a cylindrical registration volume (radius = 5 meters, height = 10 meters) centered on the reported location of the observation well (cf. **Figure 6.1** and the associated discussion in Section 6.0.2 of this dissertation). TTDs at the stream gages were generated by forward tracking particles from the land surface to discharge locations and then aggregating the travel times of all pathlines that discharged to either Morgan Creek or Chesterville Branch into a TTD for that stream.

Several factors govern the delivery of excess nitrates to the water table and their transport through the subsurface to discharge locations. These factors include not only nitrogen inputs (e.g., the rate and timing of fertilizer applications) and plant uptake, but also climate variables, soil drainage type, and soil organic matter content (Vinten et al, 1993). For example, multiple researchers have shown the particular sensitivity of leachate concentrations to precipitation patterns, as rainfall deficits during the growing season reduces crop uptake efficiencies and increases pools of excess nitrate (Burt et al., 2008), while large rainfall amounts post-harvest accelerate nitrate flushing from the root zone to the water table (Staver and Brinsfield, 1998). In addition, there is some evidence of nitrate sorption in agricultural soils, as Clay et al. (2004) estimated a retardation factor of 1.37 from column studies of clay-loam soils.

We estimated three parameters describing nitrate transport and removal in the Upper Chester: (i) nitrate removal at the land surface due to soil denitrification; (ii) retardation of

nitrate transport relative to the advective velocities described by the calibrated flow model; and (iii) nitrate removal at the stream due to the aggregated impact of in-stream and near-stream processes (e.g., denitrification at the groundwater-streambed interface, denitrification due to hyporheic exchange, and biotic uptake). The Surficial Aquifer in the Upper Chester has high dissolved oxygen concentrations and low organic carbon content, and previous researchers have found little evidence of denitrification in the aquifer itself (Green et al., 2008); as a result, we assumed conservative transport between the root zone (i.e., after removal due to soil denitrification) and the discharge location, with the important exception of flow paths that contacted the Aquia Confining Unit. Based upon the evidence of Bohlke and Denver (1995), we assumed without calibration, complete removal of nitrogen from these flow paths due to pyrite reduction and denitrification. Approximately 35% of simulated discharge to Morgan Creek contacts the Confining Unit, and a high fraction of those flow paths recharge from agriculturally-loaded areas.

4.1.2 Annual estimates of excess nitrogen

Loading rates of nitrate recharging from agricultural land to the water table were derived from county-level nitrogen budgets. These budgets were based on agricultural inputs, agricultural outputs, and atmospheric deposition. County-level budgets were estimated and converted to areal loading rates as follows. For the years 1930-2006, the total mass of agricultural nitrogen available for recharge to groundwater in Kent County during year i was calculated as

$$N_{tot,in,i} = N_{ag,in,i} - N_{ag,out,i} \quad (4.1)$$

where $N_{ag,in}$ is the agricultural nitrogen inputs, and $N_{ag,out}$ is the agricultural nitrogen exports. Historical county-level agricultural nitrogen inputs were derived from estimated and reported inorganic fertilizer sales (Alexander and Smith, 1990; Gronberg and Spahr, 2012) and estimates of poultry manure production (Sanford and Pope, 2013) (**Figure 4.2**). Historical county-level agricultural nitrogen exports were derived from the annual production of corn, soybeans, and wheat as published by the National Agricultural Statistics Service (NASS). The amount produced of each crop was converted to mass nitrogen by assuming the nitrogen content of harvested crops to be 0.9, 1.1, and 1.5 pounds nitrogen per bushel for corn, soybeans, and wheat, respectively (Murrell, 2008). The nitrogen within reported harvested silage (which is not reported for the full period of record) was assumed to remain in the catchment and thus be available for leaching.

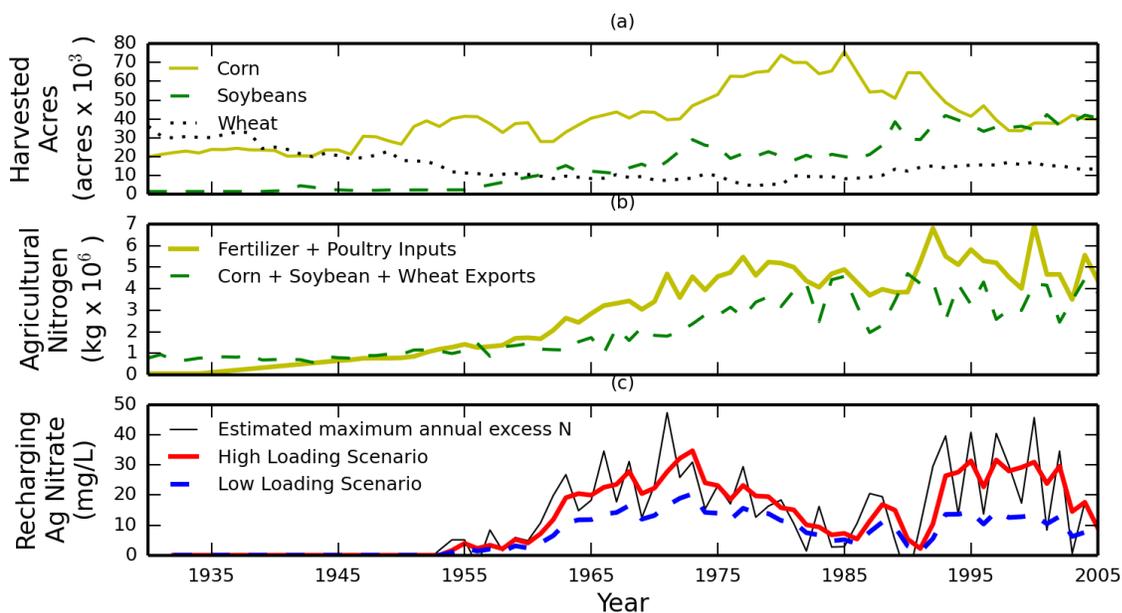


Figure 4.2 (a) Crop acreage, (b) agricultural nitrogen inputs and exports, and (c) estimated recharging nitrate concentrations for agricultural land in the Upper Chester. The High Loading and Low Loading scenarios are calculated using a three-year window; see discussion in text.

For each year, county-level estimates of the residual nitrogen available after crop uptake (Eq. 4.1) were converted to an areal loading rate that was applied within the model domain. With the exception of a few years for which a land use snapshot was available, the spatial and temporal distribution of crops – and, therefore, loading – is uncertain. That is, corn receives a much higher fraction of total fertilizer inputs for a given year than other crops (Hancock and Brayton, 2006), but in recent decades most agricultural land in the model area has been in rotation between corn, soybeans, and wheat (Hancock and Brayton, 2006). Given these data limitations, we assumed a constant spatial extent of agricultural land, and for all simulated years nitrogen loads were uniformly distributed across model cells that were categorized as row crop or pasture land by the 2008 Cropland Data Layer. In order to address the loading uncertainty, we estimated upper and lower bounds on the loading rates that were applied to an agricultural model cell in a given year. The upper estimation (referred to as the High Loading scenario; **Figure 4.2c**) assumed that fertilizers were applied to only corn and therefore calculated the areal loading rate (mass N · area⁻¹ · year⁻¹) for year i as

$$Rate_{Ag,High,i} = \frac{N_{tot,in,i}}{Area\ Corn_i} \quad (4.2)$$

where $N_{tot,in,i}$ is the county-level mass of nitrogen remaining after crop export for year i (Eq. 4.1) and $Area\ Corn_i$ is the county-level area of harvested corn for the year i . In contrast, the Low Loading scenario assumed that the excess nitrogen was uniformly available to the cumulative harvested area for corn, soybeans, and wheat as reported by NASS and therefore calculated the areal loading rate as

$$Rate_{Ag,Low,i} = \frac{N_{tot,in,i}}{Area (Corn + Soybeans + Wheat)_i} \quad (4.3)$$

For both scenarios, additional nitrogen from atmospheric deposition was uniformly applied to the entire model domain. Rates of nitrate wet deposition were obtained from the National Atmospheric Deposition Program monitoring site in Wye, Maryland, approximately 30 miles southeast of the study site (data downloaded from <http://nadp.sws.uiuc.edu> on 6/4/2015). Wet deposition data was available from 1983-2006. We assumed zero wet deposition for years prior to 1935; for years between 1935 and 1983 we used a linear interpolation to estimate annual wet deposition rates. No data on nitrate dry atmospheric deposition for the Maryland Eastern Shore was available, and so dry deposition was not included in the analysis.

We further accounted for the uncertainties associated with (i) crop rotation and (ii) variability in the timing of leaching driven by variability in precipitation inputs by transforming the loading to the water table in year i to the three-year moving average of the rates calculated with equations 4.1, 4.2, and 4.3 (**Figure 4.2c**). In summary, the High Loading scenario, which we simulated as the worst-case scenario, overestimates the total nitrogen inputs to the subsurface because it derives the fertilizer intensity from only corn acreage and applies that intensity to all agricultural acreage in the study area. In contrast, the Low Loading scenario is more closely mass conservative, but cannot capture localized loading variability. Spatial variability in loading may have important consequences, such as high nitrate leaching from cornfields.

4.1.3 Parameter estimation method

Parameters characterizing nitrate transport and removal were estimated in two phases. For both phases we calibrated parameters using the automated parameter estimation code UCODE (Poeter et al., 2005). In the first phase, the soil removal fraction and the transport retardation were calibrated against groundwater nitrate observations. Subsurface nitrate observations were aggregated into annually-averaged concentrations at each observation well. Groundwater nitrate measurements included observations from three closely-spaced transects of 3-4 piezometers each that sampled shallow groundwater in the lower reach of Morgan Creek. For reasons that are further discussed in the Results section, these observations were aggregated into a single calibration target for each transect. In sum, 233 total nitrate measurements were aggregated through annually-averaging and aggregation of transects into 96 subsurface calibration targets that date between 1988 and 2004.

Soil denitrification was assumed to be a function of soil drainage type and calculated with the expression

$$\left(\frac{S}{5}\right)^m \quad (4.1)$$

where **S** is a rank of soil drainage type on a scale of 1 to 5 and **m** is a calibrated parameter (cf. Sanford and Pope, 2013). *S* for each land surface model cell was derived from the spatial distribution of soil drainage type as identified by the 'drclasswt' attribute in the Soil Survey Geographic Database (SSURGO) dataset for Kent County MD (Soil Survey Staff).

Most soils in the study site are well-drained or moderately well-drained (corresponding to S=1 and S=2, respectively).

In the second phase of parameter estimation, parameters describing nitrate mass removal due to in-stream and near-stream processes for both Morgan Creek and Chesterville Branch were calibrated against base-flow nitrate observations. Base-flow nitrate concentrations were extracted from the complete record of in-stream nitrate concentrations by (i) separating the stream discharge record into base-flow and event components and (ii) selecting those nitrate concentrations for which the associated discharge measurement had a base-flow component that was evaluated to be 90% or higher (**Figure 4.1**). The base-flow component of total discharge was calculated using the Web-based Hydrograph Analysis Tool (WHAT) (www.engineering.purdue.edu/~what/; cf. Lim et al., 2005).

4.2 RESULTS AND DISCUSSION

4.2.1 - Characterization of Nitrate Transport and Removal Mechanisms

Table 4.1 shows the (i) calibrated values of the parameters describing nitrogen transport and removal and (ii) the relative model performances under the different loading and parameterization scenarios. **Figure 4.3** and **Figure 4.4** show the simulated and observed nitrate concentrations for the Low Loading + No Retardation calibration scenario. Note that for each loading/retardation scenario, calibration was performed in two phases (see Section 4.1.3) and thus two sum of squared error (SOSE) terms apply. The SOSEs for the High Loading scenario is much higher than the SOSEs for the Low Loading scenario because

the higher loading assumptions increase the maximum possible error for an individual observation.

The similarity across the scenarios of the terms describing near-stream/in-stream removal at Chesterville Branch and Morgan Creek indicates that these values accurately describe the role of these systems in the nitrogen budget for each subcatchment. For Morgan Creek, stream processes remove 60-70% of incoming loads, while for Chesterville Branch stream process only remove 15-30% of incoming loads. Note that the stream removal term for Morgan Creek is in addition to removal due to denitrification at the Confining Unit, which is already accounted for in the model and is assumed to completely remove nitrate from flow paths that intersect it. This means that if the actual removal percentage due to the Confining Unit is lower than modeled, then the removal rates due to stream processes at Morgan Creek are higher than estimated here. Conversely, if the spatial extent and influence of the Confining Unit is greater in reality than is represented by the model, then the Morgan Creek removal rates would be lower than shown in **Table 4.1**.

For the High Loading Scenario, the addition of the retardation parameter reduced the SOSE by 10%, and for the Low Loading Scenario the addition of the retardation parameter reduced the SOSE by 40%. The estimated value of the retardation parameter was similar under both loading scenarios (1.39 for the High Loading scenario and 1.45 for the Low Loading scenario), and the confidence intervals associated with both loading scenarios overlapped. Notably, the parameter estimates under both loading scenarios are close to the value of 1.37 estimated by Clay et al. (2004) for clay-loam column studies.

Table 4.1. Calibrated nitrogen removal mechanisms and transport parameters. Values in parentheses under each estimated parameter are the lower and upper 95% confidence intervals calculated by UCODE. No confidence intervals could be calculated for the soil removal terms under the Low Loading scenarios.

			High Loading		Low Loading	
			Parameter Estimate	SOSE	Parameter Estimate	SOSE
Fraction of Incoming Loads Removed in Stream	With Retardation	Morgan Creek	0.61 (0.56/0.66)	154	0.67 (0.63/0.71)	153
		Chesterville Branch	0.14 (0.10/0.17)		0.18 (0.15/0.22)	
	Without Retardation	Morgan Creek	0.62 (0.57/0.67)	155	0.70 (0.66/0.73)	154
		Chesterville Branch	0.25 (0.23/0.28)		0.33 (0.30/0.36)	
Fraction of Incoming Loads Removed in Soils	With Retardation	Retardation	1.39 (1.32/1.47)	8009	1.45 (1.38/1.54)	6455
		Well drained	0.36 (0.26/0.49)		0 (--)	
		Moderately well drained	0.56 (0.46/0.67)		0 (--)	
	Without Retardation	Poorly drained	0.87 (0.83/0.91)	8113	0.03 (--)	6733
		Well drained	0.39 (0.29/0.51)		0 (--)	
		Moderately well drained	0.58 (0.50/0.69)		0 (--)	
		Poorly drained	0.88 (0.84/0.91)		0.12 (--)	

Note that the soil removal term is determined by a single calibrated parameter, m (see Section 4.1.3), but described in **Table 4.1** by its resulting removal estimate for specific soil

types. The estimated soil removal terms vary significantly between loading scenarios but are not sensitive to the presence or absence of the retardation parameter. For the High Loading scenarios, the majority of soils in the model domain (i.e., well drained and moderately drained soils) are estimated to have removed approximately 40-60% of excess nitrogen, while under the Low Loading scenarios no soil removal occurred. These estimates of soil removal percentages should be considered bounds associated with the range of possible loading scenarios rather than estimates of the values themselves.

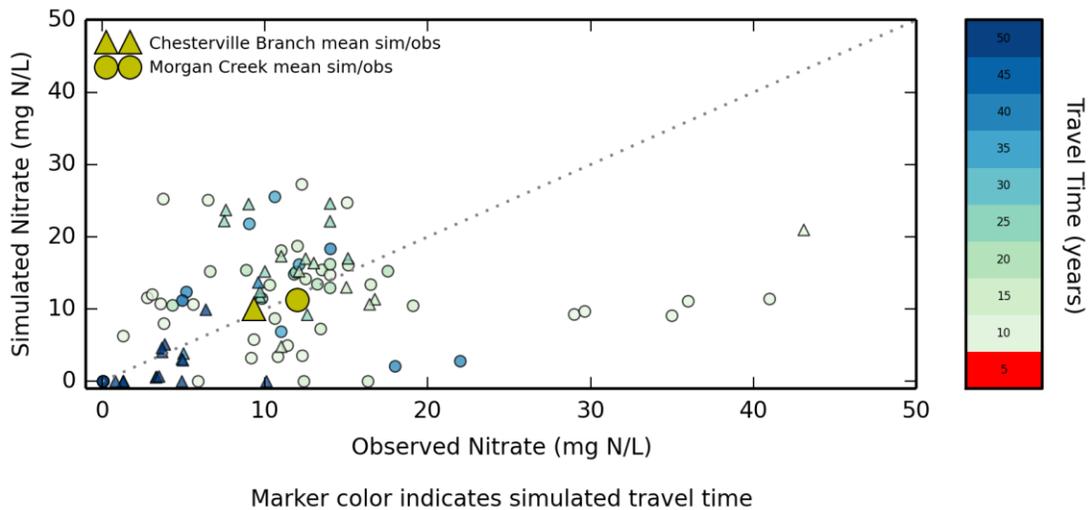


Figure 4.3 Simulated and observed groundwater nitrate concentrations for observation wells in the Upper Chester. The mean simulated/observed value is the mean subsurface measurement for the specified catchment.

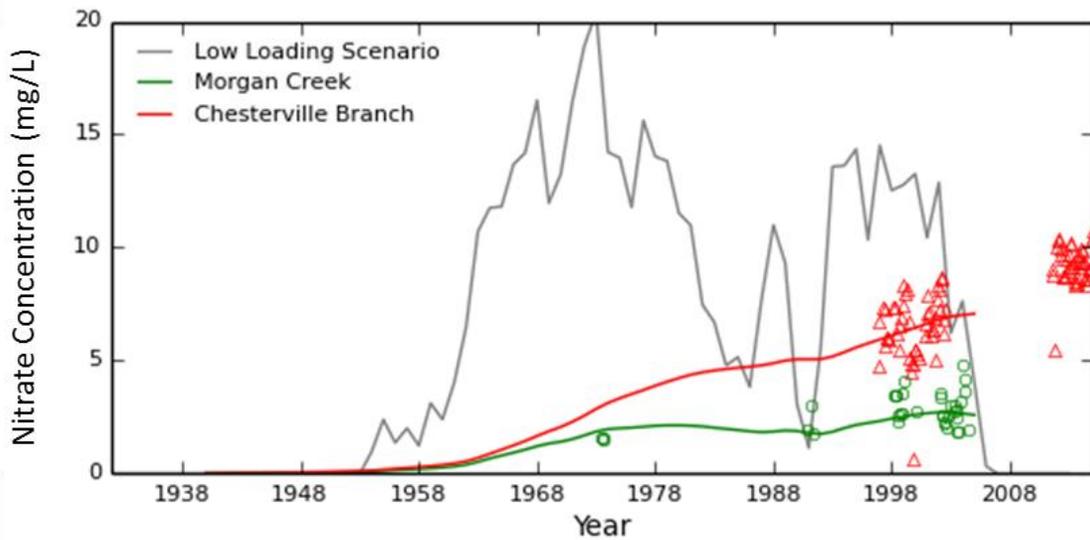


Figure 4.4 Simulated and observed base-flow nitrate concentrations in Morgan Creek (green circles and green line) and Chesterville Branch (red triangles and red line). The gray line shows the nitrate concentration recharging to the water table for simulations using the Low Loading Scenario. Concentrations simulated until 2006, which is the most recent year for published fertilizer sales.

Morgan Creek has lower median age because the Confining Unit creates a shallower surficial system with a higher fraction of faster delivery. It is therefore more sensitive to changes in land use. Note, for example, the small decline in the Morgan Creek simulated base-flow concentrations during the late 1980's in response to the decline in recharging nitrate during that same time period (**Figure 4.4**). Chesterville Branch, in contrast, has older median and mean ages than Morgan Creek. Therefore, for the Chesterville Branch simulated base-flow nitrate, the rate of increase slowed in response to the decline in recharging nitrate during the late 1980's but the concentrations did not decrease.

The effect of the different removal mechanisms on the total nitrogen budget in each catchment can be seen in **Figure 4.5**. The dotted line in each panel shows the loads that would be exported by base-flow discharge in a given year if no nitrogen were removed by

any of the calibrated processes (i.e., soil denitrification, Confining Unit denitrification, or in-stream removal). The dashed line in each panel shows the level to which loads were reduced by soil denitrification and Confining Unit removal. The solid line in each panel shows the actual loads leaving each catchment (i.e., the level to which loads are reduced by in-stream removal processes).

For the Low Loading scenario shown in **Figure 4.5**, soil denitrification removed approximately 1% of annual loads to both catchments. The Confining Unit removal partially accounts for the difference in stream nitrate concentrations. In the Chesterville Branch catchment, no flowpaths that recharged since 1940 and carry agricultural nitrate contacted the Confining Unit. Therefore, for Chesterville Branch there was no removal due to Confining Unit contact, and the dotted and dashed lines coincide (**Figure 4.5b**). In contrast, the Confining Unit removed approximately 15% of total nitrate loads to Morgan Creek in a given year. However, the differences between Morgan Creek and Chesterville Branch are more a function of in-stream removal processes than denitrification at the Confining Unit. Morgan Creek drains a larger catchment than Chesterville Branch and consequently has more loads recharging to the water table and potentially reaching the stream (dotted lines in **Figure 4.5**). Even after removal due to Confining Unit contact, the loads to Morgan Creek are larger than the loads to Chesterville Branch (dashed lines in **Figure 4.5**). However, after in-stream removal processes, the loads leaving the catchment are lower in Morgan Creek than in Chesterville Branch (solid lines in **Figure 4.5**).

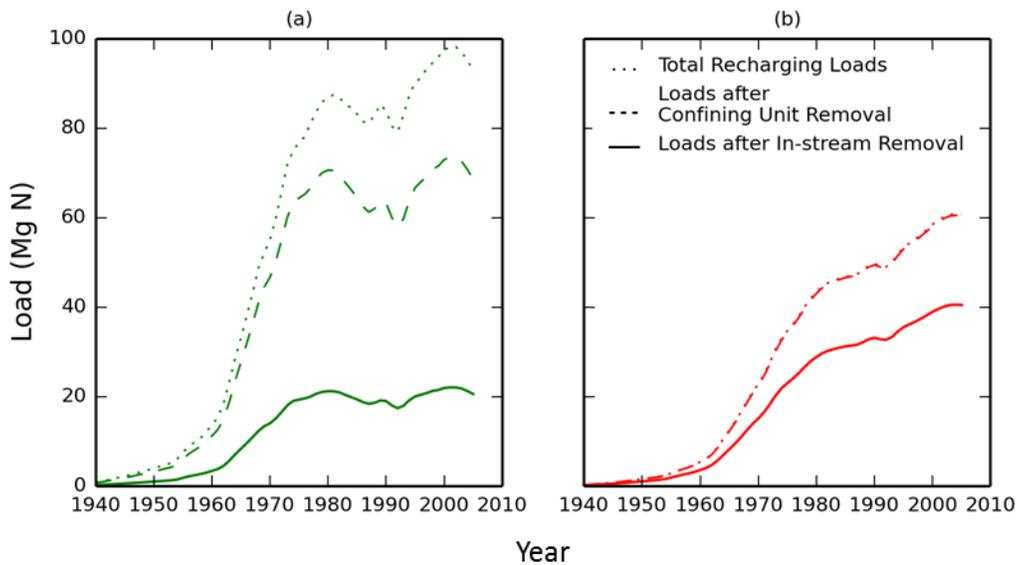


Figure 4.5 Contributions of removal mechanisms to reducing exported loads in (a) Morgan Creek and (b) Chesterville Branch. Time series shown for calibrated parameters from Low Loading/No Retardation scenario (see Table 4.1). Line styles identified in panel (b) legend apply to both panels.

4.2.2 – Model performance, and the use of the calibrated flow model to corroborate the nitrogen loading history

Figure 4.6 shows the model-predicted recharge date for subsurface nitrate observations in the Upper Chester and compares those to the estimated loading time series (described in Methods, above). The simulated recharge date shown on the horizontal axis refers to the recharge date that is predicted by the calibrated flow model given the date and location of the subsurface nitrate observations shown on the vertical axis.

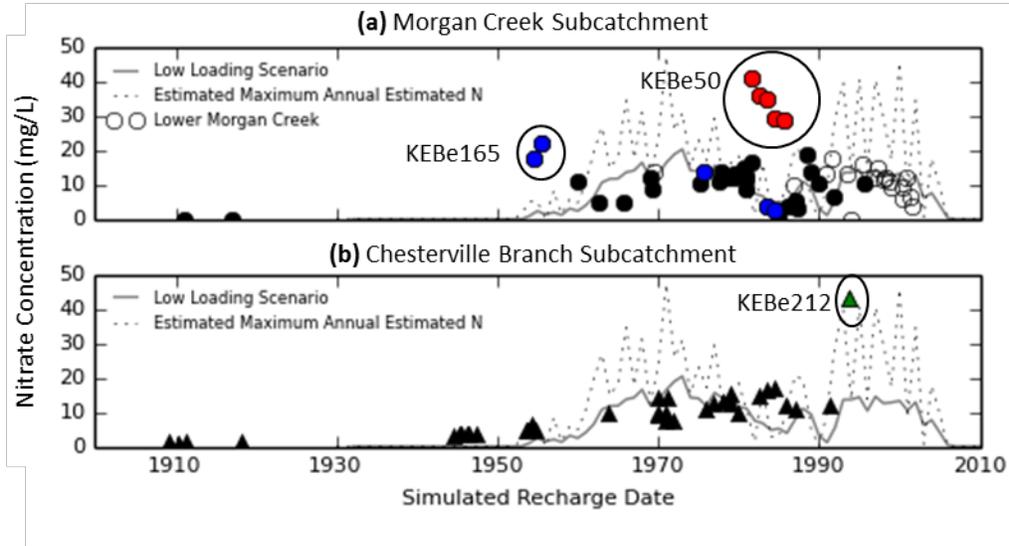


Figure 4.6 Recharge-dates predicted by the flow model for Upper Chester groundwater nitrate observations. The Simulated Recharge Date (horizontal axis) is the recharge date predicted by the calibrated flow model given the date and location of the nitrate observation.

Given the uncertainty associated with estimating historic nitrate loads from county-level crop data and distributing according to limited spatial information, the general consistency between the assumed loading time-series and the observed subsurface nitrate concentrations provides an important corroboration of our loading methodology. **Figure 4.6** simultaneously suggests: (i) the general capacity of the flow model to relate land surface inputs to subsurface observations; (ii) the uncertainties in the transport simulation that may be introduced by the flow model's simplifications of local heterogeneities; and (iii) the uncertainties in the transport simulation that may be introduced by both highly heterogeneous land surface loading and sub-surface removal mechanisms. That is, the recharge dates simulated by the flow model map the trend of the majority of nitrate observations (shown by black triangles and black circles) to the trend in recharging nitrate

for those portions of the aquifer in which oxic conditions prevail and transport is expected to be conservative. Subsurface concentrations for observations that recharged prior to the late 1980's increase with time. In the late 1980's and early 1990's, increasing energy costs drove increasing fertilizer costs, resulting in the conversion of cropland from corn to soybeans and associated decreases in fertilizer input (written communication with Ann Baldwin, Natural Resource Conservation Service liaison for Kent County, MD). During this same time period, intermittent dry conditions may have contributed to the volatile crop productivity evident in **Figure 4.2b**. This combination of lower fertilizer nitrogen inputs and lower nitrogen uptake rates may be responsible for the variability in subsurface observations that recharged around 1990. Green et al. (2008) used $\delta^{15}\text{N}$ and age information to reconstruct the loading history for a well transect in Morgan Creek and similarly found evidence of some transition in loading rates around 1990. The early 1990's began a period of increased implementation of nutrient management plans leading up to their mandatory implementation by all Maryland farmers beginning in 1998 (written communication with Ann Baldwin); however, it appears from the nitrogen balance calculated in this study that loading again increased in the 1990's, and that loading reductions observed by Green et al. (2008) may have been due to a market-driven or hydrologically-driven pause rather than an improvement in land use management.

Exceptions to the general trends just described may be briefly considered as follows. For well KEBe165 (**Figure 4.6a**), there is evidence that spatial heterogeneities in recharge conditions and temporal heterogeneities in land use are responsible for the apparent disconnect between the modeled recharge date and the nitrate observation. Nitrate

observations for the well nest including KEBE167, KEBE166, and KEBE165 are shown by blue circles on **Figure 4.6a**. **Figure 4.7** additionally shows simulated and observed ages for this nest. Focused recharge from an unsimulated, edge-of-field drainage pond that is located near this well nest would account for higher transport velocities and younger observed ages at wells KEBE165 and KEBE166 than is predicted by the flow model. At the same time, time-variable land use data from the National Land Cover Dataset and the Cropland Data Layer shows that the plots adjacent to the well nest have been used for both field crops and hay pasture during the last two decades. Noting that the low nitrate concentrations at KEBE167 (**Figure 4.6**) were sampled in successive years (1990, 1991), these low concentrations could be explained by recharge from a non-fertilized period. In other words, supporting information for this location collectively suggests that (a) the recharge date for well KEBE165 is approximately 1970 (rather than 1955, as predicted by the flow model; **Figure 4.6a**) and (b) the recharge concentrations for well KEBE167 may not be coordinated with the agricultural loading trend.

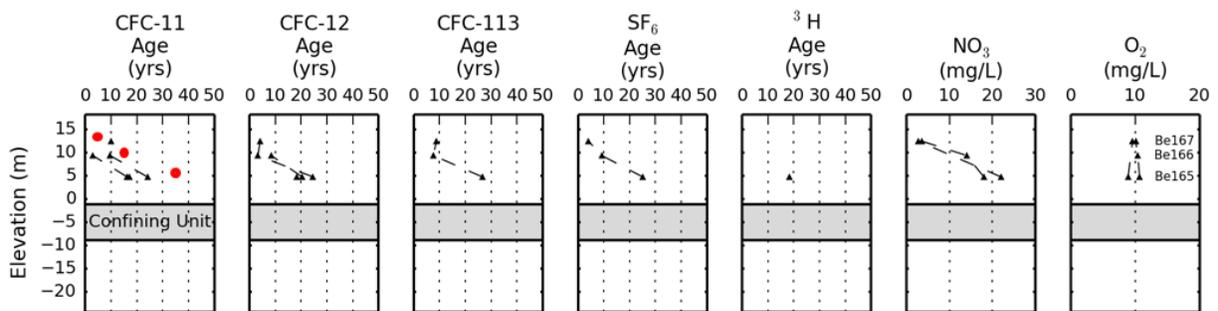


Figure 4.7 Age, nitrate, and dissolved oxygen observations for the well nest that includes KEBE167, KEBE166, and KEBE165. Triangles show observations; the dashed lines connect observations made on the same day. The red circles (•) on the leftmost panel show model simulated ages for the three wells.

The single observation at well KEBE212 (shown by a green triangle on **Figure 4.6b**) may be similarly independent of the general agricultural loading history because it is located near a nursery located in the headwaters of Chesterville Branch. No data on loading rates at the nursery are available. However, it may be noted that other observations from wells that likely sample groundwater flow paths recharging from the nursery exhibit much lower nitrate concentrations, on the order of 10-15 mg-N/L.

The five observations at well KEBE50 are shown by red circles on **Figure 4.6a**. These samples were collected between 1988 and 1992 and similarly appear to be the consequence of a loading hotspot at the recharge location. Residuals from both simulated heads and simulated atmospheric tracer concentrations are very small (results not shown); we are consequently confident in the model-predicted recharge dates. There is no evidence of land use other than row crop agriculture for the recharge location identified by the flow model for this well. However, unlike the likely causal link between the nursery and the high concentrations at KEBE212, no data exists that would explain why this location more than others would preserve such a high recharging nitrate signal.

Finally, several nitrate observations are from closely-spaced transects of piezometers that sample groundwater at depths of 0-3 m beneath the stream bed on the lower reaches of Morgan Creek (shown by hollow circles on **Figure 4.6a**). This very small area (i.e., all observations separated by less than 300 meters) is subject to very steep nitrate gradients. Chemical and isotopic evidence indicates that these gradients are due to a combination of

at least three factors: (a) converging flow paths of widely disparate ages from the Surficial Aquifer and Confined Aquifer; (b) nitrate removal due to denitrification in the Confining Unit sediments; and (c) concentrated near-stream loading that possibly originates at a dairy operation waste retention pond near the lower right bank of Morgan Creek (Puckett et al., 2008; Bohlke and Denver, 1995; Bachman et al., 2002). As described in Section 4.1.3, these multiple measurements were aggregated into a single measurement for each transect at a given point in time. **Figure 4.6a** shows the aggregated measurements, which as aggregates are well represented by model, rather than all individual measurements from lower Morgan Creek.

In summary, while it is not expected that a catchment-scale model can reproduce all local complexity, the flow model successfully maps the general trend of the observed subsurface nitrate to the likely nitrate loading history. This corroborates its utility as a simulation tool linking the land surface to surface water and identifying the controlling removal mechanisms.

4.2.3 - Discussion of contrasting stream removal efficiencies

Given the available data, the current analysis cannot conclusively identify why the in-stream/near-stream removal rates vary so much between the two catchments. However, two potential factors may be considered. Recent studies have used regression analysis of nitrogen removal across a range of stream types (Mulholland et al., 2008; Alexander et al., 2009; Böhlke et al., 2009) as well as modeling of nitrate transport through individual stream networks (Wollheim et al., 2006) to show that nitrate removal efficiency in smaller

order streams is a function of both stream depth (or velocity) and stream nitrate concentration. At shallow depths and low velocities, a greater fraction of the stream nitrates has longer exposure to (i) denitrifying microbes in streambed sediments and (ii) nitrogen uptake services from in-channel biota (Alexander et al., 2009). Less well understood is the finding that nitrogen removal efficiency declines with increasing stream nitrate concentration. For example, Mulholland et al. (2008) found across a range of smaller order streams that increasing the stream nitrate concentration from 1.5 to 15 mg/L may reduce the nitrate removal fraction by more than half (from approximately 25% to 10%; see Figure 4 in Mulholland et al., 2008). Scanlon et al. (2010) similarly observed that the background in-stream nitrate removal rates in a small forested stream network were approximately 80% but dropped to less than 5% following a widespread tree defoliation event that quadrupled the in-stream nitrate concentration. While the disruption of increased nitrate concentrations on benthic and aquatic community nitrate processing is not well described, we may note that declining nitrogen removal functions with increasing nitrate concentrations have also been observed in higher organisms such as freshwater mussels (Spooner et al., 2013).

The potential for differences in (i) stream channel and flow characteristics or (ii) stream water quality to explain the differences in stream removal rates between Morgan Creek and Chesterville Branch is briefly considered here. The Morgan Creek riparian zone is thickly wooded, with tree debris common in the stream channel (Duff et al., 2008). As described above, the Confining Unit which outcrops at the lower reaches may not only account for substantial nitrogen removal through denitrification, but also controls the

manner in which discharge enters the main channel (i.e., through rivulets and sheetflow in the near-stream floodplain, rather than through the bed sediments). While the Chesterville Branch stream network has not been characterized with the same detail, it is expected that base-flow discharge to Chesterville Branch is via upwelling through the sandy bed sediments (part of the Aquia Aquifer formation), bypassing the riparian zone processing that is an important control in Morgan Creek. The organic content of the Chesterville Branch bed sediments, and the associated denitrification potential of those sediments (cf. Gu et al., 2008) is not known.

A coarse comparison of the stream channels may be made using (i) velocities and associated cross-sectional flow areas sampled at the catchment outlets (**Figure 4.8**) and (ii) National Hydrography Dataset (NHD) representation of the stream networks (shown in **Figure 2.1**). It is important to note that it is not known how representative outlet characteristics are of velocities and cross-sectional flow areas throughout the respective stream networks. Nevertheless, we may tentatively conclude that the Chesterville Branch catchment has shorter in-stream residence times due to a shorter network (**Figure 2.1**) and higher velocities (**Figure 4.8b**). These variable residence times may account for some of the variation in stream nitrate removal.

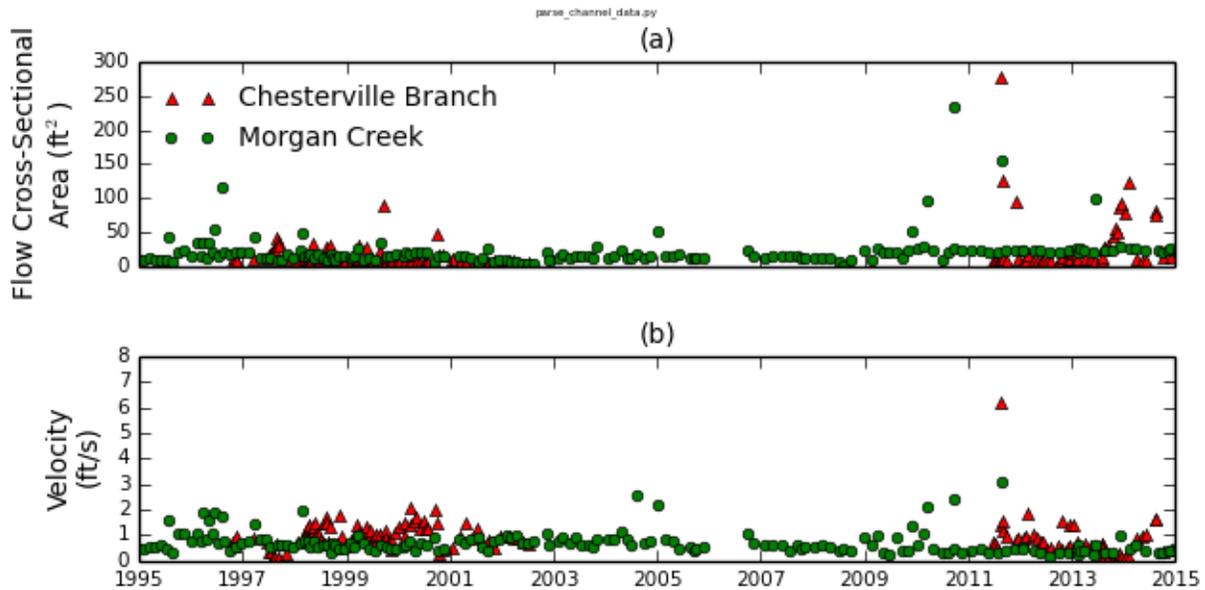


Figure 4.8 Flow characteristics measured at the Morgan Creek and Chesterville Branch stream gages. Each marker represents a field measurement. See **Figure 2.1** for locations.

In addition, evidence from a small set of synoptic studies suggests that Chesterville Branch headwater concentrations have historically been much higher than headwater concentrations in Morgan Creek (**Figure 4.9**). These conclusions are likewise tentative because of the few spatially-distributed snapshots that include both Morgan Creek and Chesterville Branch (Morgan Creek is not recently sampled, while multiple sites on Chesterville Branch have been sampled twice yearly since 2011). In the early 1990's (i.e., at the time at which the stream networks were simultaneously sampled) surficial aquifer nitrate concentrations in each catchment had nitrate concentrations of 10-20 mg NO₃-N/L for observation wells near the upstream-most site in both catchments. However, Morgan Creek headwater concentrations were significantly lower than aquifer concentrations, while Chesterville Branch headwater concentrations were not.

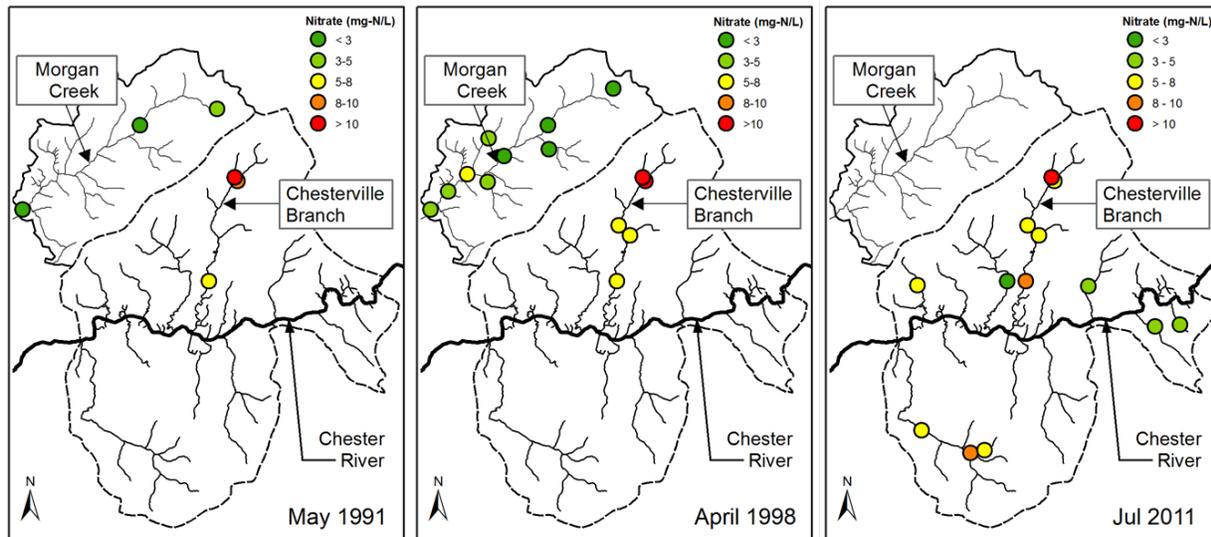


Figure 4.9 Base-flow stream nitrate concentrations from synoptic surface water sampling in Morgan Creek and Chesterville Branch. Sites shown in the rightmost panel (July 2011) are currently sampled bi-annually.

4.2.4 – Management implications

These results raise an additional question that is of potential significance for management of nitrogen export from lower order streams. Most basically stated for Chesterville Branch: are the low removal rates in Chesterville Branch (i) a characteristic of the natural system or (ii) a legacy of stream degradation? That is, given the evidence of suppressed stream processing efficiency with elevated nitrate concentrations (Mulholland et al., 2008; Alexander et al., 2009; Scanlon et al., 2010), it is important to consider whether nitrate removal services in Chesterville Branch may be improved (i.e., restored) by reducing the headwater loads. Further study is required to evaluate the relative importance of headwater loads (versus loads from tributaries or base-flow discharge further

downstream) in the high concentrations at the subcatchment outlet, and whether these loads are responsible for degrading the in-stream processing capacity.

The same essential question may be applied in different terms to Morgan Creek: does the nitrate load reduction provided by Confining Unit denitrification help maintain the in-stream processing potential of Morgan Creek? That is, in the absence of the damping that is due to the Confining Unit, would the in-stream/near-stream removal rates in Morgan Creek be as high as they are observed to be in this study? This requires further specification of the spatial extent of the Confining Unit and its potential impact on not only base-flow loads to the lower reaches but also to the Morgan Creek headwaters. Alternatively, are there other removal mechanisms in the Morgan Creek headwaters that filter higher loads and thus preserve the relatively high rates of nitrate removal along the reach? If these alternative mechanisms can be identified, their maintenance may be a priority for management of loads from the Morgan Creek subcatchment.

These results also illustrate the complexity of predicting and managing agricultural nitrate export at a regional scale. For example, highly variable nitrate export due to highly variable removal rates from lower-order streams could have important implications for the economics of nutrient reduction in the Chesapeake Bay watershed. Under some proposed systems, BMP implementation in both subcatchments would qualify for the same nutrient reduction credits even though management action in Morgan Creek may be redundant with natural processes. That is, the natural denitrification potential in the aquifer may establish a lower limit on nitrate concentrations for the management timescale of interest. This

would therefore mean, for example, that a municipality that offsets nitrate loads from future development through the purchase of credits generated in the Morgan Creek catchment will not, in fact, be offsetting those future loads, since no further load reductions are to be expected as a result of the agricultural BMPs.

4.3 CONCLUSIONS AND FUTURE WORK

The assessment and management of agricultural nitrogen loadings to receiving waters requires both (i) characterizing groundwater nitrate transport to ecological systems and (ii) characterizing the background capacity of those systems to process excess nitrate. This is particularly true in areas of agricultural land-use, where the relative contribution of base-flow to nitrate loads is higher due to the ongoing discharge of nitrates that have accumulated in surficial aquifers during the past century. We used the travel time distributions generated by a well-calibrated flow model in order to relate land surface loadings to stream responses and resolve the key components of the catchment nitrogen budget. We examined adjacent agricultural subcatchments with similar land use histories but disparate nitrate export signatures in order to quantify the removal fractions of various removal mechanisms. We showed that in spite of spatial and temporal uncertainty in loading, multiple calibration scenarios agreed that in-stream nitrate removal efficiencies vary significantly between the two streams.

Future research should integrate improved land use information and more spatially distributed watershed monitoring. Under Section 1619 of the Farm Bill, the USDA is developing a detailed, privacy-protected database of land use and agricultural practices

associated with programs administered by the Farm Service Agency (FSA) and Natural Resources Conservation Service (NRCS). This emerging dataset will provide spatially- and historically-distributed data on BMP implementation and conservation measures (Hively et al., 2013). Watershed monitoring since 2011 now includes: (i) bi-annual monitoring of base-flow discharge and several water quality parameters at multiple sites within the watershed; and (ii) high-frequency (e.g., 15-minute or daily) monitoring of both total discharge and nitrate concentration at the Chesterville Branch outlet (Science and Technical Advisory Committee, 2010). Given the lag times between land surface action and stream response, several years of water quality monitoring will be required to fully leverage the higher resolution land use data now being collected. However, more immediate further investigations will consider whether the emerging land use dataset provides any information upon which reconstructions of historical loading may be conditioned, thus reducing the loading uncertainty described in this chapter.

4.4 REFERENCES

Alexander, R. B., & Smith, R. A. (1990). *County-Level estimates of nitrogen and phosphorus fertilizer use in the United States, 1945 to 1985*. USGS open file report 90--130. US Department of Interior, US Geological Survey.

Alexander, R. B., Böhlke, J. K., Boyer, E. W., David, M. B., Harvey, J. W., Mulholland, P. J., Wollheim, W. M. (2009). Dynamic modeling of nitrogen losses in river networks unravels the coupled effects of hydrological and biogeochemical processes. *Biogeochemistry*, 93(1-2), 91-116.

Ator, S. W., & Denver, J. M. (2012). Estimating contributions of nitrate and herbicides from groundwater to headwater streams, Northern Atlantic Coastal Plain, United States. *JAWRA Journal of the American Water Resources Association*, 48(6), 1075-1090.

Bachman, L. J., Krantz, D. E., Böhlke, J., & Hantush, M. M. (2002). Hydrogeologic framework, groundwater geochemistry, and assessment of nitrogen yield from base flow in two agricultural watersheds, Kent County, Maryland. Environmental Protection Agency.

- Böhlke, J. K., & Denver, J. M. (1995). Combined use of groundwater dating, chemical, and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, Atlantic Coastal Plain, Maryland. *Water Resources Research*, 31(9), 2319-2339.
- Böhlke, J. K., Antweiler, R. C., Harvey, J. W., Laursen, A. E., Smith, L. K., Smith, R. L., & Voytek, M. A. (2009). Multi-scale measurements and modeling of denitrification in streams with varying flow and nitrate concentration in the Upper Mississippi River Basin, USA. *Biogeochemistry*, 93(1-2), 117-141.
- Clay, D. E., Zheng, Z., Liu, Z., Clay, S. A., & Trooien, T. P. (2004). Bromide and nitrate movement through undisturbed soil columns. *Journal of Environmental Quality*, 33(1), 338-342.
- Clune, J. W., & Denver, J. M. (2012). *Residence time, chemical and isotopic analysis of nitrate in the groundwater and surface water of a small agricultural watershed in the coastal plain, Bucks Branch, Sussex County, Delaware*. US Geological Survey Scientific Investigations Report 2012-5235.
- Duff, J. H., Tesoriero, A. J., Richardson, W. B., Strauss, E. A., & Munn, M. D. (2008). Whole-stream response to nitrate loading in three streams draining agricultural landscapes. *Journal of Environmental Quality*, 37(3), 1133-44.
- Gitau, M. W., Chaubey, I., Gbur, E., Pennington, J. H., & Gorham, B. (2010). Impacts of land-use change and best management practice implementation in a conservation effects assessment project watershed: Northwest Arkansas. *Journal of Soil and Water Conservation*, 65(6), 353-368.
- Green, C. T., Puckett, L. J., Böhlke, J. K., Bekins, B. A., Phillips, S. P., Kauffman, L. J., Denver, J.M., Johnson, H. M. (2008). Limited occurrence of denitrification in four shallow aquifers in agricultural areas of the United States. *Journal of Environmental Quality*, 37(3), 994-1009.
- Gronberg, J. M., & Spahr, N. E. (2012). *County-level estimates of nitrogen and phosphorus from commercial fertilizer for the conterminous United States, 1987-2006*. US Department of the Interior, US Geological Survey.
- Gu, C., Hornberger, G. M., Herman, J. S., & Mills, A. L. (2008). Influence of stream-groundwater interactions in the streambed sediments on NO₃⁻ flux to a low-relief coastal stream. *Water Resources Research*, 44(11).
- Hancock, T. C., & Brayton, M. J. (2006). *Environmental setting of the Morgan Creek Basin, Maryland, 2002-04*. US Geological Survey Open File Report 2006-1151.
- Hively, W.D., Devereux, O.H., and Claggett, Peter (2013). Integrating federal and state data records to report progress in establishing agricultural conservation practices on Chesapeake Bay farms. US Geological Survey Open-File Report 2013-1287.
- Lim, K. J., Engel, B. A., Tang, Z., Choi, J., Kim, K. S., Muthukrishnan, S., & Tripathy, D. (2005). Automated web GIS based hydrograph analysis tool, WHAT1. *JAWRA Journal of the American Water Resources Association*, 41(6), 1407-1416.
- Lindsey, B. D., Phillips, S. W., Donnelly, C. A., Speiran, G. K., Plummer, L. N., Böhlke, J.K., Busenberg, E. (2003). *Residence times and nitrate transport in ground water discharging to streams in the Chesapeake Bay Watershed*. US Department of the Interior, US Geological Survey.

Meals, D. W., Dressing, S. A., & Davenport, T. E. (2010). Lag time in water quality response to best management practices: A review. *Journal of Environmental Quality*, 39(1), 85-96.

Mulholland, P. J., Helton, A. M., Poole, G. C., Hall, R. O., Hamilton, S. K., Peterson, B. J., Dahm, C. N. (2008). Stream denitrification across biomes and its response to anthropogenic nitrate loading. *Nature*, 452(7184), 202-205.

Murrell, T.S. (2008). Measuring Nutrient Removal, Calculating Nutrient Budgets. *Soil Science: Step-by-step Field Analysis*. 259-181.

Nelson, J., & Spies, P. (2013). The Upper Chester river watershed: Lessons learned from a focused, highly partnered, voluntary approach to conservation. *Journal of Soil and Water Conservation*, 68(2), 41A-44A.

Osmond, D., Meals, D., Hoag, D., Arabi, M., Luloff, A., Jennings, G., Line, D. (2012). Improving conservation practices programming to protect water quality in agricultural watersheds: Lessons learned from the National Institute of Food and Agriculture-Conservation Effects Assessment Project. *Journal of Soil and Water Conservation*, 67(5), 122A-127A.

Poeter, E.P., Hill, M.C., Banta, E.R., Mehl, Steffen, Christensen, Steen. (2005). *UCODE_2005 and six other computer codes for universal sensitivity analysis, calibration, and uncertainty evaluation*. US Department of the Interior, US Geological Survey Reston, VA, USA.

Puckett, L. J., Tesoriero, A. J., & Dubrovsky, N. M. (2011). Nitrogen contamination of surficial aquifers: A growing legacy. *Environmental Science and Technology*, 45(3), 839.

Puckett, L. J., Zamora, C., Essaid, H., Wilson, J. T., Johnson, H. M., Brayton, M. J., & Vogel, J. R. (2008). Transport and fate of nitrate at the ground-water/surface-water interface. *Journal of Environmental Quality*, 37(3), 1034-50.

Rupert, M. G. (2008). Decadal-scale changes of nitrate in ground water of the united states, 1988--2004. *Journal of Environmental Quality*, 37(5 - Supplement), S240-S248.

Sanford, W. E., & Pope, J. P. (2013). Quantifying groundwater's role in delaying improvements to Chesapeake Bay water quality. *Environmental Science and Technology*, 47(23), 13330-13338.

Scanlon, T. M., Ingram, S. M., & Riscassi, A. L. (2010). Terrestrial and in-stream influences on the spatial variability of nitrate in a forested headwater catchment. *Journal of Geophysical Research: Biogeosciences*, 115(G2).

Schindler, D. W., & Vallentyne, J. R. (2008). *The Algal Bowl*. University of Alberta Press.

Science and Technical Advisory Committee (2005). *Understanding Lag Times Affecting the Improvement of Water Quality in the Chesapeake Bay: A Report from the Chesapeake Bay Program Scientific and Technical Advisory Committee*. Science and Technical Advisory Committee Publication 5-001. Downloaded from www.chesapeake.org/stac/stac_pubs.php.

Science and Technical Advisory Committee (2010). *Small Watershed Monitoring Designs: A Report from the Chesapeake Bay Program Scientific and Technical Advisory Committee*. Science and

Technical Advisory Committee Publication 10-004. Downloaded from www.chesapeake.org/stac/stac_pubs.php.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database. Available online at <http://sdmdataaccess.nrcs.usda.gov/>.

Spooner, D. E., Frost, P. C., Hillebrand, H., Arts, M. T., Puckrin, O., & Xenopoulos, M. A. (2013). Nutrient loading associated with agriculture land use dampens the importance of consumer-mediated niche construction. *Ecology Letters*, 16(9), 1115-1125.

Staver, K. W., & Brinsfield, R. B. (1998). Using cereal grain winter cover crops to reduce groundwater nitrate contamination in the Mid-Atlantic Coastal Plain. *Journal of Soil and Water Conservation*, 53(3), 230-240.

Sutton, A. J., Fisher, T. R., & Gustafson, A. B. (2009). Historical changes in water quality at German Branch in the Choptank River Basin. *Water, Air, & Soil Pollution*, 199(1), 353-369.

Wollheim, W. M., Vörösmarty, C. J., Peterson, B. J., Seitzinger, S. P., & Hopkinson, C. S. (2006). Relationship between river size and nutrient removal. *Geophysical Research Letters*, 33(6).

Vinten, A. J. A., Smith, K. A., Burt, T. P., Heathwaite, A. L., & Trudgill, S. T. (1993). Nitrogen cycling in agricultural soils. *Nitrate: processes, patterns and management*, 39-73.

Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., Schlesinger, W.H., and Tilman, D. G. (1997). Human alteration of the global nitrogen cycle: sources and consequences. *Ecological applications*, 7(3), 737-750.

United States Environmental Protection Agency (USEPA). (2010). Chesapeake Bay Phase 5.3 Community Watershed Model. EPA 903S10002 - CBP/TRS-303-10. U.S. Environmental Protection Agency, Chesapeake Bay Program Office, Annapolis MD.

Chapter 5: Numerical Simulation of Seasonal Changes in Base-flow Age³

5.0 INTRODUCTION

5.0.0 - Background and Motivation

The age of streamwater – and the distribution of flowpaths that constitute that age – are important integrated descriptors of how catchments store and release both rainfall and contaminants (Kirchner et al., 2001; McDonnell et al., 2010). For example, in-stream concentrations of contaminants that discharge from the subsurface, such as agricultural nitrates, are a function of both time-varying land surface inputs and time-varying hydrology. The interaction of these variations makes it challenging to assess the impact of land management action on in-stream water quality (Hirsch et al., 2010; Howden et al., 2011). A largely-unexplored dimension of these entangled drivers is the relationship between seasonal variations in base-flow discharge and the age of that discharge. That is, base-flow age is a measure of the subsurface flow paths discharging to a stream, and for time-variable land surface inputs, the contaminant concentration of those flow paths is a function of recharge time and therefore age. In many agricultural systems where land management actions should be improving stream water quality, it would be instructive to resolve the components of in-stream contaminant variability (**Figure 5.1**). These components include: long-term trends in contaminant inputs; shorter term noise from recent climate impact on crop performance (Staver and Brinsfield, 1998); temperature- and season-dependent in-stream contaminant removal processes (Mulholland et al., 2008;

³ This chapter is being prepared for submission to *Water Resources Research*.

Bohlke et al., 2009); and transient subsurface hydrology and the associated changes in the distribution of ages arriving at the stream.

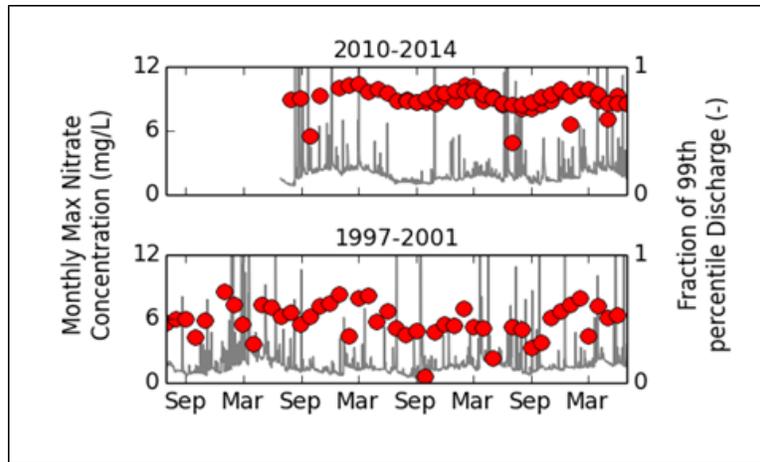


Figure 5.1 Seasonal variations in stream nitrate concentrations (•) and stream discharge for Chesterville Branch on the Maryland Eastern Shore.

Several researchers have developed theoretical frameworks for describing the age of streamwater (Botter 2012; Gomez and Wilson, 2013; Harman, 2014) or have used experimental methods to identify the age of water discharging from monitored catchments (e.g., Ogrinc et al., 2008; Rodhe et al. 1996). A preponderance of experimental studies have used the input and output signals of the stable isotope $\delta^{18}\text{O}$ to infer the underlying distribution of flowpath ages discharging to a stream (McGuire and McDonnell, 2006). In these lumped-parameter studies, the travel time distribution (TTD) is assumed to be of a known form such as an exponential or gamma distribution (Kirchner et al., 2001; Maloszewski and Zuber, 1998). With the form of the TTD assumed, the input and output signals may then be used in inverse-modeling to specify the values of the distribution parameters (e.g., the mean age parameter for an exponential distribution, or the shape and scale parameters for a gamma distribution) and thus derive the mean age (**Figure 5.2**).

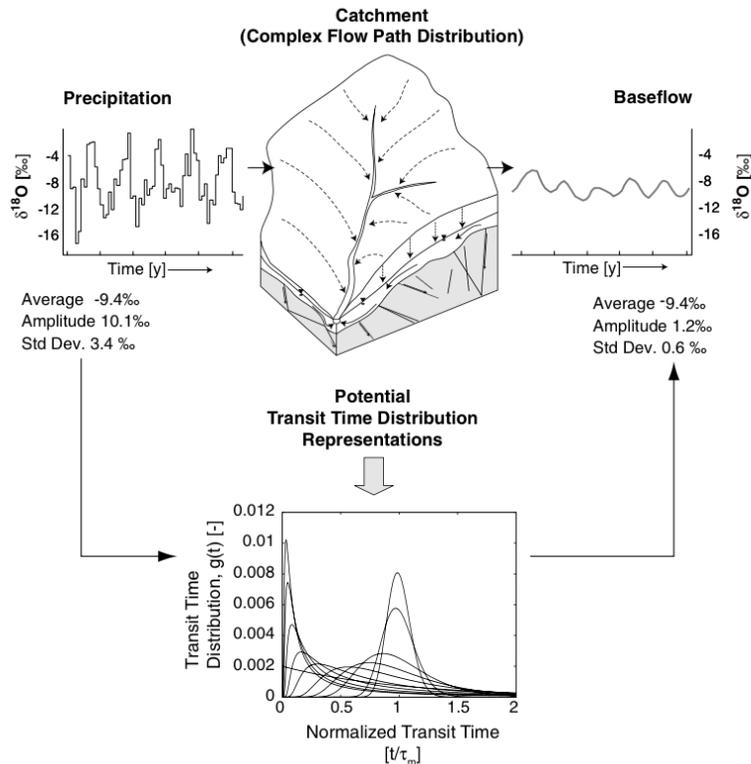


Figure 5.2 From McGuire and McDonnell (2006). Schematic illustrating the use of (i) a precipitation input signal plus (ii) the assumption that the streamflow preserves that signal to infer the TTD for a catchment.

The use of $\delta^{18}\text{O}$ as a tracer depends on seasonal variations in the $\delta^{18}\text{O}$ content of precipitation. Previous researchers have shown that the actual range of these seasonal variations effectively restricts the use of the $\delta^{18}\text{O}$ method to catchments with mean travel times of approximately 4 years or less (McGuire and McDonnell, 2006). Stewart et al. (2010) has similarly shown that because $\delta^{18}\text{O}$ cannot be used to detect the contributions of older groundwaters to stream age, the results from analysis that have relied on it are likely biased young. These limitations suggest that more methods are needed in order to better understand the contribution of base-flow to the age (and therefore contaminant response) of streamflow. This is particularly the case in low-relief, high-infiltration catchments in

which groundwater discharge dominates streamflow, and even more so when such a catchment is subject to diffuse agricultural loadings such as nitrate. For example, in some agriculturally-intensive catchments on the Delmarva Peninsula, chemical hydrograph separations suggest that groundwater discharge is responsible for as much as 85% of long-term average streamflow (Sanford, 2011).

Beyond using $\delta^{18}\text{O}$, a few studies have used tritium (^3H) individually or in combination with some other conservative tracer to date streamwater (Michel, 1992; Stewart et al., 2010). For example, Peters et al. (2013) paired (i) a high-frequency time series of in-stream dissolved silica (Si) measurements with (ii) a relationship between ^3H and Si derived from near-stream observation wells to generate a time series of stream age at the Panola Mountain (GA) catchment. It should be noted, however, that ^3H methods face their own limitations. The atmospheric concentration of ^3H has declined from the mid-20th century bomb peak towards background levels at a rate similar to the rate of ^3H radioactive decay, which means that waters of different ages may have indistinguishable ^3H concentrations (Stewart et al., 2010). Tritium is therefore more useful when coupled with either the ^3He daughter species or an altogether different tracer such as dissolved CFCs or SF_6 . These dissolved tracers have, in turn, limitations as well, since the equilibration of dissolved gaseous tracers with the atmosphere has historically required that these tracers must be measured in the saturated zone. However, more recently methods have been developed that allow atmospheric-tracer base-flow concentrations to be measured by sampling stream water (Sanford et al., in press). These methods measure diurnal changes in stream temperature and in-stream concentrations of SF_6 or CFCs to infer the stream residence time and the mean concentration of the dissolved tracer when it discharged from the

subsurface. This estimate of the discharging tracer concentration can then be used to estimate the apparent age of the discharging base-flow.

In contrast to lumped-parameter methods that use in-stream stable isotope information, the present study uses a groundwater flow and transport model calibrated against subsurface environmental tracer information to simulate the age of base-flow discharge to a stream. In other words, rather than sample some surrogate of age in the stream and infer a TTD from that, the present study develops a numerical groundwater flow and transport model and uses it to simulate the transport of age through the system.

In addition to a general methodological dependence on stable isotopes, most investigations of streamwater age have focused on the calculation or measurement of a mean steady state age as a fundamental descriptor of catchment processes. These investigations include studies that have employed lumped parameter methods (e.g., Ogrinc et al., 2008; Hrachowitz et al., 2009) as well as studies that have used forward modeling methods similar to those implemented in the present work (e.g., Abrams, 2012; Sanford and Pope, 2013). However, more recent researchers have emphasized that the TTD is itself a function of time-dependent hydrological changes, such as seasonal fluctuations in storage driven by seasonal variability in precipitation and evapotranspiration. Harman (2014), building on Botter (2012), used a probabilistic framework to derive a general expression for a time-variable TTD that was then applied to the Plynlimon Experimental Catchment chloride time series in order to develop a transient, lumped-parameter model of streamflow age. Gomez and Wilson (2013) used a finite-element implementation of the governing equation for groundwater age (Ginn, 1999; Ginn et al., 2009) to simulate the impact of a harmonic head

distribution on system velocities, flow directions, and age distributions in a synthetic (i.e., Tothian – cf. Toth, 1962) regional groundwater system. The present study is complementary to the work of Gomez and Wilson (2013) in that it applies an alternative simulation method to generate time-variable TTDs. Further, the present study applies the method to estimate the seasonal changes in base-flow age for a field study.

5.0.1 – Potential methods for simulating transient base-flow age

Another vein of research has conceptualized groundwater age as a mass that accumulates with time in a subsurface system (Ginn, 1999; Goode, 1996; Gomez and Wilson, 2013;). This conceptual model may be mathematically expressed through a modified form of the advective dispersive model (ADM) and implemented with a solute transport code (e.g., MT3DMS – Zheng, 2005). Under this modeling scheme, the simulated mean age of base-flow discharging from the groundwater system to surface water is equal to the mean age (i.e., solute concentration of age) of the model element (e.g., finite-difference grid cell) at which the discharge occurs. Varni and Carrera (1998) and Gomez and Wilson (2013) describe methods for specifying, in addition to the mean, higher moments or the full TTD at a point in time. It may be noted, however, that direct simulation of age with a modified transport code (such as MT3DMS) may only allow simulation of a mean age, rather than a full TTD, and production of the full time-variable TTD may be computationally prohibitive (Gomez and Wilson, 2013).

An alternative method for simulating time-variable base-flow age is with a tracking scheme that traces the flow paths governed by system hydraulics. While multiple authors have discussed the potential shortcomings of this kinematic conception of age (e.g., Goode, 1996;

Varni and Carrera, 1998), some systems are in fact well described by advective transport and therefore well-suited for a kinematic description of age. For example, Chapter 3 of this dissertation showed that environmental tracer datasets were well-described by particle tracking. Further, kinematic methods avoid the computational burdens and numerical dispersion associated with applying the groundwater age form of the ADM to advection-dominated systems. We use kinematic methods for this study, as described in further detail in next section.

5.1 METHODS

In order to investigate the impact of both system characteristics and system inputs on base-flow age, we simulated the time-variable age of groundwater discharge to surface water under different recharge conditions for a variety of synthetic two-dimensional (2D) aquifers. We also simulated the time-variable age of base-flow to Morgan Creek in the Upper Chester catchment using a transient form of the Upper Chester model described in Chapter 2 and calibrated in Chapter 3 of this dissertation. The following sections describe the method of simulating flow and base-flow age, the suite of synthetic recharge series and synthetic aquifers used in the study.

5.1.1 – Simulation of base-flow and base-flow age

For all scenarios, the distribution of aquifer heads was calculated using a modified form of the finite-difference simulation code MODFLOW-2000 (Harbaugh et al., 2000). The code modifications were developed by S.S. Papadopolous and Associates, and the resulting simulation code is hereafter referred to as MF2K-SSPA. Briefly stated, MF2K-SSPA uses a modified form of the equation for saturated groundwater flow and a Newton-Raphson

solution method in order better accommodate fluctuating water tables (Bedekar et al., 2012). Base-flow discharge was simulated using the MODFLOW drain package as described in Chapter 2.

Simulations were performed in monthly time steps for the synthetic cases and quarterly time steps for the field study. Intra-annual variations in groundwater recharge were represented by a discretized approximation of a sinusoidal function, with the system receiving the highest recharge inputs during the late-winter/early-spring months and no recharge inputs during summer months (described in more detail below).

For all scenarios in this study, base-flow age was simulated using a modified form of the particle tracking software MODPATH 5 (Pollock, 1994) developed by S.S. Papadopolous as a companion to MF2K-SSPA and hereafter referred to as MP-SSPA. The code modifications included in MP-SSPA include improved capacity to track particles (i.e., quantify flow paths) through grid cells that dry and re-wet as the water table elevation changes due to transient forcing. A series of Python scripts was used to (i) extract the water table elevation from the MF2K-SSPA head solution for each time step at each row/column location throughout the model grid and (ii) input those time-variable water table elevations as particle starting locations for MP-SSPA. Particles were tracked forwards from the water table to the discharge location. An additional series of Python scripts was used to aggregate discharging particles at each time step and generate a kinematic TTD from the travel times associated with those discharging particles. In order to calculate the flux-weighted mean age, each discharging particle was weighted by the recharge rate associated with the time-step at which that particle entered the system. For all scenarios, the simulation included an

initialization period required to spin the model up to an annually repeating cycle of seasonally-variable mean ages. In most cases, the total simulation time required to reach this annually repeating cycle exceeded the number of time steps accommodated by the simulation codes. For these circumstances, Python scripting was used to extract the location of particles that were in transit between recharge and discharge at the end of a simulation run and re-start those particles at the same location in the model grid for the next stage of the simulation.

5.1.2 – Simulation of base-flow SF₆

In addition to simulating discharging base-flow ages, we also simulated the time variable concentration of base-flow SF₆ by weighting each discharging particle with the recharging SF₆ concentration for the year in which the particle entered the subsurface. In order to calculate the recharging SF₆ concentration for each particle, the atmospheric SF₆ partial pressure for the recharge year was converted to dissolved SF₆ concentrations using Henry's law and assuming that SF₆ recharged at 10° C and with 2 cm³/L excess air (cf. Chapter 3). Note that the simulation of base-flow SF₆ required the assignment of an arbitrary reference time. That is, while the simulation of base-flow age subject to sinusoidal recharge forcing is independent of a specific calendar year, the simulation of SF₆ is not because its atmospheric history is not steady state. For the Synthetic cases we assumed that the sampling year was 2011, and for the Morgan Creek case study we assumed that the sampling year was 2007.

For each scenario, an SF₆ apparent age was determined by relating the simulated mean base-flow SF₆ concentration to the SF₆ atmospheric time series. The 'apparent age' refers

to the age derived from some form of measured tracer under the assumption that the entire sample recharged at a single point in time. See McCallum et al. (2014) or Kazemi (2006) for more information on apparent age and related groundwater age concepts.

5.1.3 - Synthetic simulation scenarios

We simulated the seasonally-variable base-flow age for a suite of synthetic two-dimensional (2D) aquifers. For all synthetic scenarios, the aquifer was 100 m deep by 1250 m wide and discretized into 20 layers of uniform thickness equal to 5 m. For most scenarios, the aquifer was discretized into 50 columns of uniform width equal to 25 m. However, for some scenarios we re-gridded the column layout into 250 columns of equal width equal to 5 m; these re-gridded scenarios are identified in the Results and Discussion section. The spatially-uniform porosity and specific yield for each aquifer was equal to 0.25, and the specific storage for each aquifer was equal to $1 \times 10^{-5} \text{ m}^{-1}$. The land surface elevation slope was 2% (not shown on Figure 3).

The synthetic aquifer systems were subject to variation in: (i) the aquifer hydraulic conductivity field; (ii) the annually averaged recharge; (iii) the amplitude of the time-varying recharge input; and (iv) the spatial distribution of the recharge input for a given time step. These variations are summarized in **Figure 5.3** and further described as follows:

- i. The synthetic aquifer framework was either (a) homogeneous unconfined or (b) composed of an unconfined surficial aquifer overlaying a confining unit and a confined aquifer. The impact of aquifer diffusivity on the base-flow age time series was tested by varying the homogenous hydraulic conductivity field between 0.6, 1, 2.5 and 10 m/day (**Figure 5.3**). In these simulations the aquifer diffusivity is equal

to Kb/S_y , where K is the hydraulic conductivity, b is the saturated thickness and S_y is the specific yield.

- ii. Unless explicitly noted, annually averaged recharge to the subsurface was 40 cm/year. For the four separate hydraulic conductivity fields (three homogeneous plus layered) the steady state application of this recharge rate resulted in steady state mean base-flow ages of 54 to 58 years (**Figure 5.4**). In order to examine the impact of aquifer mean age on the base-flow age time series we held the aquifer conductivity field constant (homogeneous conductivity = 1 m/day) and ran a series of scenarios with different magnitudes of annually-averaged recharge, ranging from 20 to 80 cm/year. This resulted in scenarios with steady state mean base-flow ages ranging from 36 to 112 years (**Figure 5.5**).
- iii. In order to examine the impact of the amplitude of the seasonal recharge input on the base-flow time series we applied both a high and low amplitude recharge time series to the homogenous 1 m/day conductivity field (**Table 5.1**). For all other synthetic cases we generated the monthly recharge time series use the recharge factors designated as 'High Amplitude' in **Table 5.1**.
- iv. Finally, we also considered the impact of the spatial distribution of the recharge by applying both spatially uniform and spatially variable recharge to the homogeneous 1 m/day conductivity field. For the spatially variable scenario, the recharge rate increased as a linear function of distance from the discharge location (**Figure 5.3**; cf. Goode, 1996). Note that for a given scenario the same spatial distribution was applied at each time step, with the magnitude of the recharge for that time step determined by the appropriate recharge factors listed in **Table 5.1**.

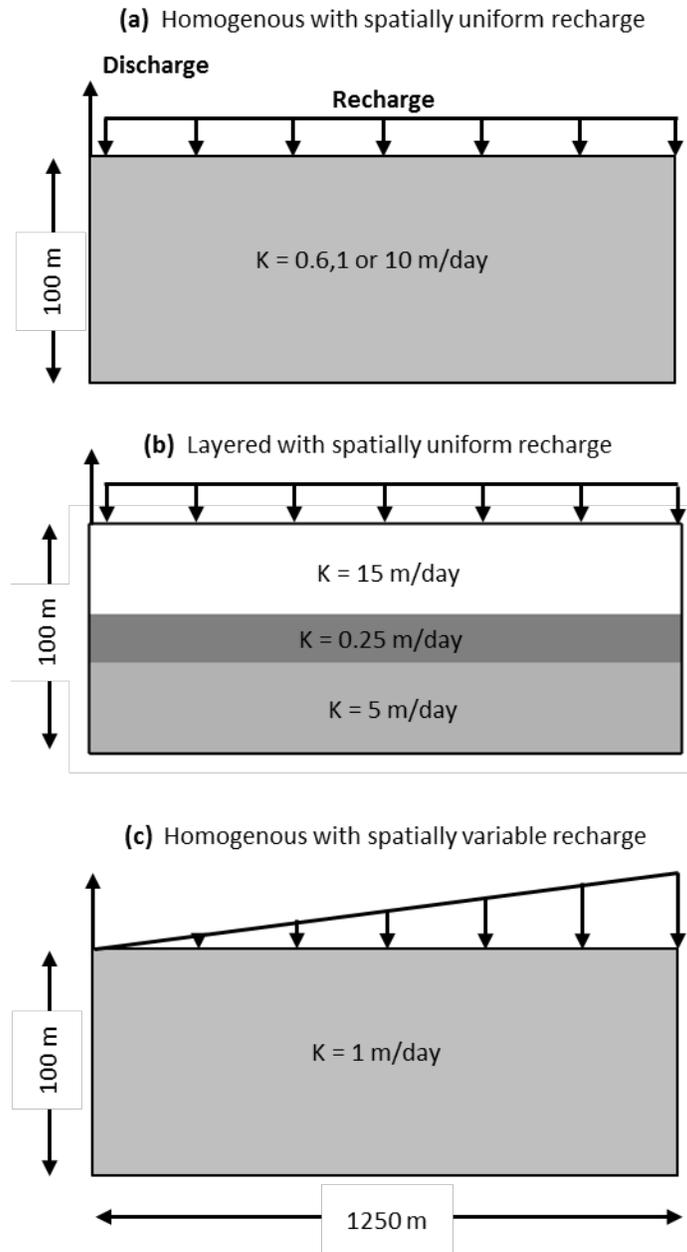


Figure 5.3 Aquifer dimensions, hydraulic conductivity fields, and recharge spatial distributions for synthetic 2D scenarios.

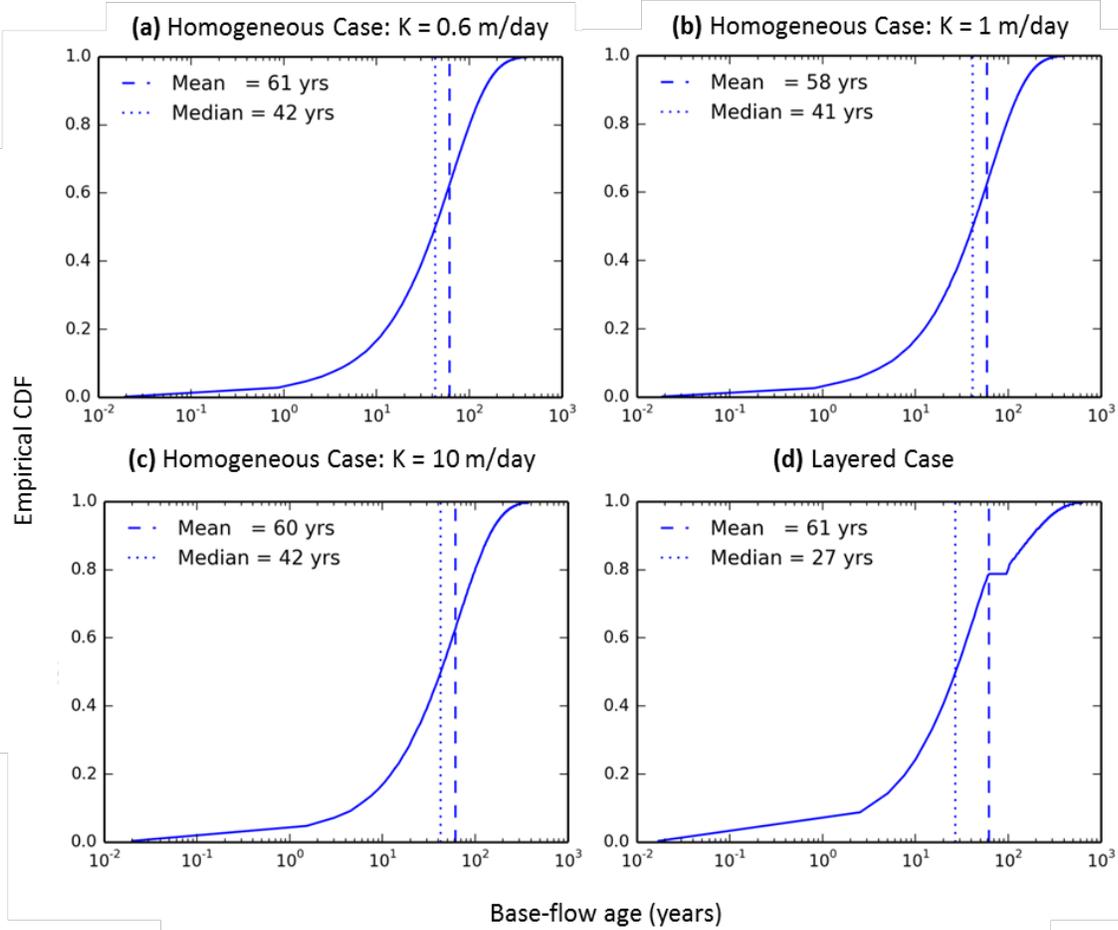


Figure 5.4 Steady state base-flow age distributions for different aquifer scenarios subject to 40 cm/year recharge.

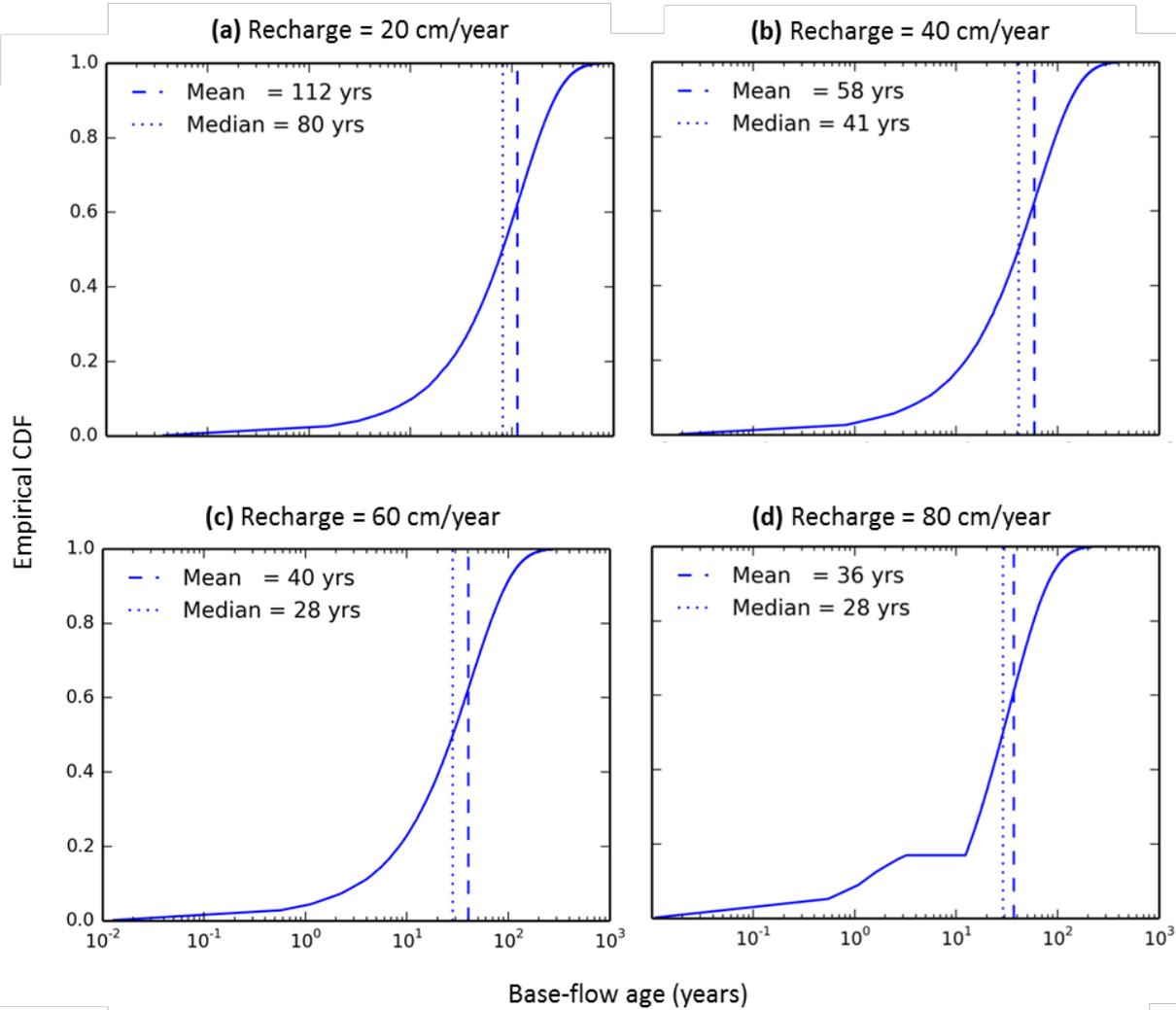


Figure 5.5 Steady state base-flow age distributions for the homogenous 1 m/day conductivity field subject to different recharge rates.

Table 5.1 Multiples of steady state recharge used to decompose the annually averaged recharge into seasonally varied recharge.

Month	<u>Monthly Factors</u>		<u>Quarterly Factors</u>
	High Amplitude	Low Amplitude	
January	2.01	1.15	3.2
February	2.17	1.17	
March	2.01	1.15	
April	1.57	1.09	0.4
May	0.97	1.00	
June	0.36	0.91	
July	0.00	0.84	0.0
August	0.00	0.82	
September	0.00	0.84	
October	0.36	0.91	0.4
November	0.97	1.00	
December	1.57	1.09	

5.2 RESULTS AND DISCUSSION

5.2.1 - Synthetic Scenarios

5.2.1a - Comparison of variable conductivity field

Figures 5.6 and **5.7** shows the seasonally varying recharge, base-flow, age, and SF₆ time series for various aquifer configurations. Each configuration in **Figure 5.6** and **Figure 5.7** received 40 cm/year recharge distributed according to the High Amplitude factors listed in

Table 5.1 (maximum recharge in February and zero recharge in July, August, and September). For each scenario we computed the recharge-weighted base-flow age for each month and then fitted a sinusoidal curve to the resulting monthly series using least squares regression. For each scenario, the vertical dashed lines indicate the temporal relationship between the minimum base-flow discharge and the maximum base-flow mean age. Note that the recharge input, base-flow discharge, and sinusoidal best-fit mean age are annually-repeating series; the vertical lines are shown only for a single year.

As expected, the base-flow discharge time series is a damped and lagged form of the sinusoidal recharge input. The amplitude of the base-flow discharge series varied with system conductivity. For the homogeneous cases (**Figures 5.6a, 5.6b, and 5.6c**) the lag between low recharge and low base-flow was constant; low base-flow occurred in October, which is a one-month lag from the September zero recharge. For the layered case (**Figure 5.7**), which had a much lower median age due to the shallow unconfined aquifer, the base-flow discharge responded more quickly to changes in recharge and the low base-flow discharge occurred in September.

Within a single year, the mean base-flow ages for the homogeneous cases vary by 7-12 years, but for the layered case that included a confining unit and lower aquifer the mean base-flow ages vary by nearly 20 years during a single year (**Table 5.2**). This larger range of ages for the layered case is a function of the changing contributions of older water from below the confining unit during the contrasting high and low base-flow regimes. At high base-flow, the base-flow age TTD for the layered scenario is dominated by the shallower

flow-paths of the surficial aquifer, while at low base-flow the TTD is more heavily influenced by discharge from below the confining unit.

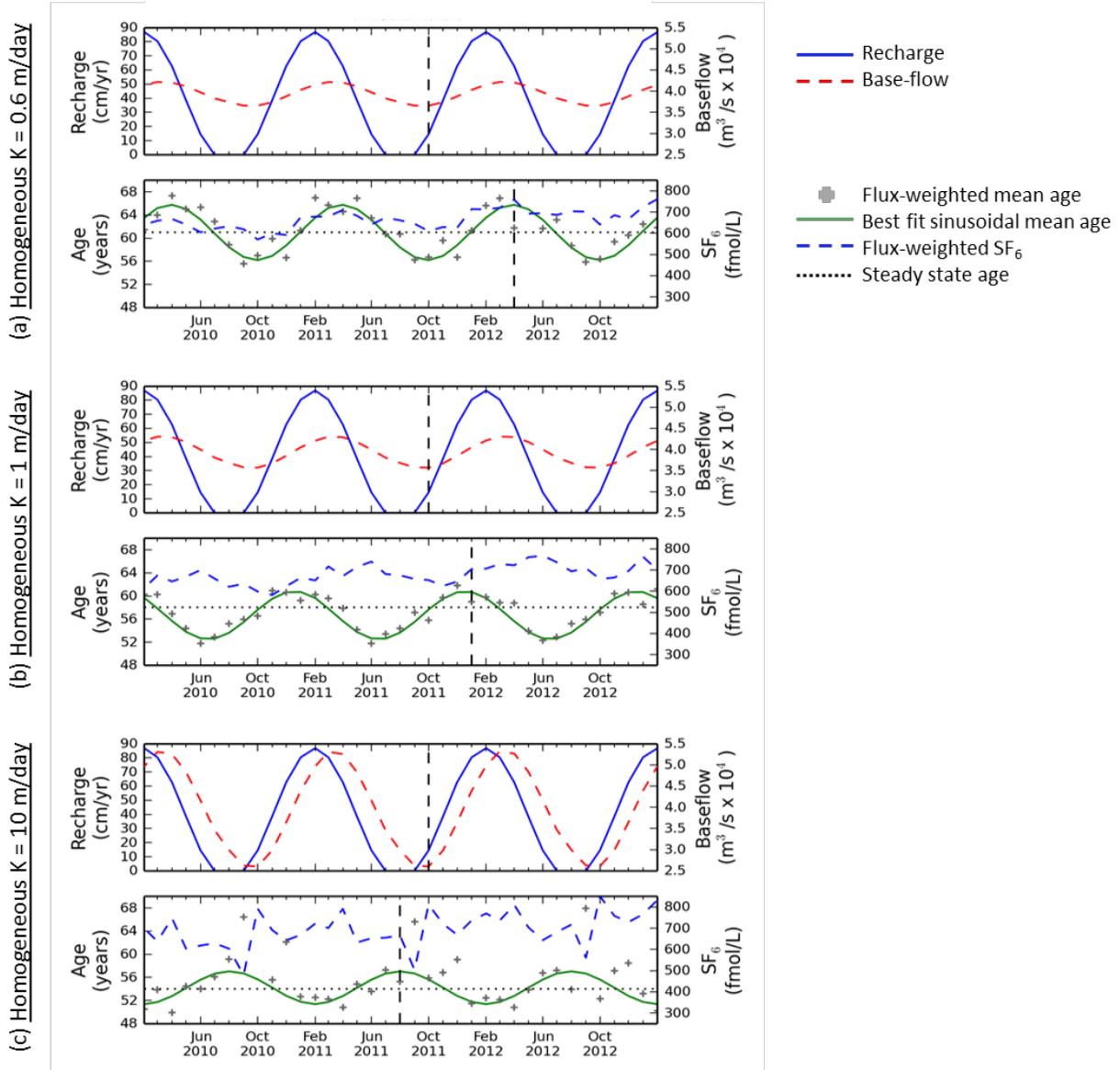


Figure 5.6 Comparison of base-flow age time series resulting from changes in the aquifer conductivity field. The vertical dashed lines mark the timing of minimum base-flow and peak sinusoidal best fit age (in top and bottom panels, respectively, for scenarios a-c).

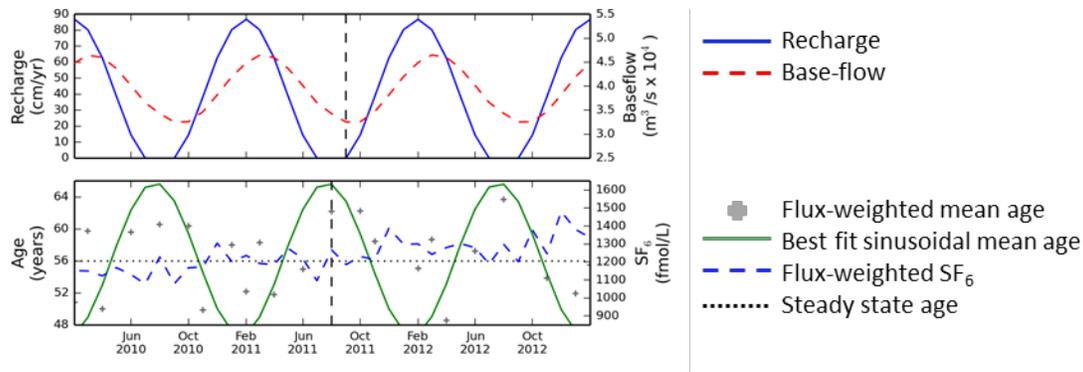


Figure 5.7 Flow and age time series for the layered conductivity scenario with annually averaged recharge = 40 cm/year.

For the homogeneous case with hydraulic conductivity = 0.6 m/day (**Figure 5.6a**) the maximum base-flow age is lagged by 5-6 months from the minimum base-flow discharge; the peak simulated mean age occurred in March and the peak of the best-fit sinusoidal occurred in April. For the case in which homogeneous hydraulic conductivity = 1 m/day (**Figure 5.6b**) the lag between minimum base-flow and maximum age is reduced to 1-2 months. With homogeneous conductivity = 10 m/day or with layered conductivity (**Figures 5.6c** and **5.7**, respectively) the maximum base-flow age is closely coordinated with the minimum recharge. These results suggest that the distribution of discharging ages is a function of both (i) velocity, which is controlled by the recharge rate and (ii) diffusivity, which controls the rate at which changes in the potential energy field are translated through the system. For both (i) the shallow, fast surficial system in the layered conductivity field and (ii) the highest homogeneous conductivity (10 m/day) the ages are more closely synced with rising and falling recharge velocities, while for the lower hydraulic conductivities the response of discharging age to incoming recharge is more delayed.

5.2.1b - Consideration of computational noise

For all cases the long-term average recharge weighted age for the transient simulation was close to the steady-state base-flow age; this supports the basic utility of this method for identifying seasonal variability in ages. However, for some scenarios the simulated base-flow age time series included significant deviation from a sinusoidal form (e.g., **Figure 5.6c**). A full characterization of this noise, including its dependence on the spatial or temporal simulation discretization or the aquifer characteristics, has not been completed. However, we offer preliminary observations here. It is possible that the noise includes artifacts introduced by approximations of the velocity field as generated by the particle-tracking routine. MODPATH calculates advective velocities by interpolating between heads at adjacent nodes (Pollock, 1994), and the resolution of the interpolated velocity field is determined by the number of particles. However, preliminary sensitivity studies indicated that refining the resolution of the velocity field by increasing the number of particles introduced during each time step does not of itself improve the fit of the simulated ages to a sinusoidal function. Similarly, we did not find that reducing the time step of the transient flow solution improved the fit. We did find evidence that reducing the grid size was necessary to reduce the noise in simulations with higher flow velocities. **Figure 5.8** compares the seasonal base-flow ages generated by simulation using column widths equal to 25 m and 5 m for the homogeneous aquifer with hydraulic conductivity equal to 10 m/day. Note that **Figure 5.8a** is identical to **Figure 5.6c**. Note that when compared to the 25 m resolution, the reduction of column width to 5 m had some effect on the simulated water table elevation and thus increased the mean steady state age. Reducing the column width to 5 m significantly improved the fit of the simulated mean ages to a sinusoidal

function. However, we also observe that the phase of the sinusoidal best-fit function is identical for each grid resolution, suggesting that the noisy base-flow age time series is also identifying the temporal relationship between recharge, base-flow, and discharging ages.

While the computational questions that we are describing here should not be confused with the numerical challenges associated with the advection-dispersion equation (ADE) for solute transport (Zheng and Bennett, 2002), it is plausible that the computational noise encountered in some of the simulation conditions in this study is a function of the same elements that constitute the Courant number, which is one of the stability criterion for numerical solutions to the ADE (I am grateful to Mohamed Morsy for suggesting this paradigm). The Courant number is a combined indicator of advective velocity, simulation time step, and grid discretization (Zheng and Bennett, 2002). To our knowledge, no formal accuracy criterion has been formulated for MODPATH and should be the subject of further investigation in support of the method described in this chapter.

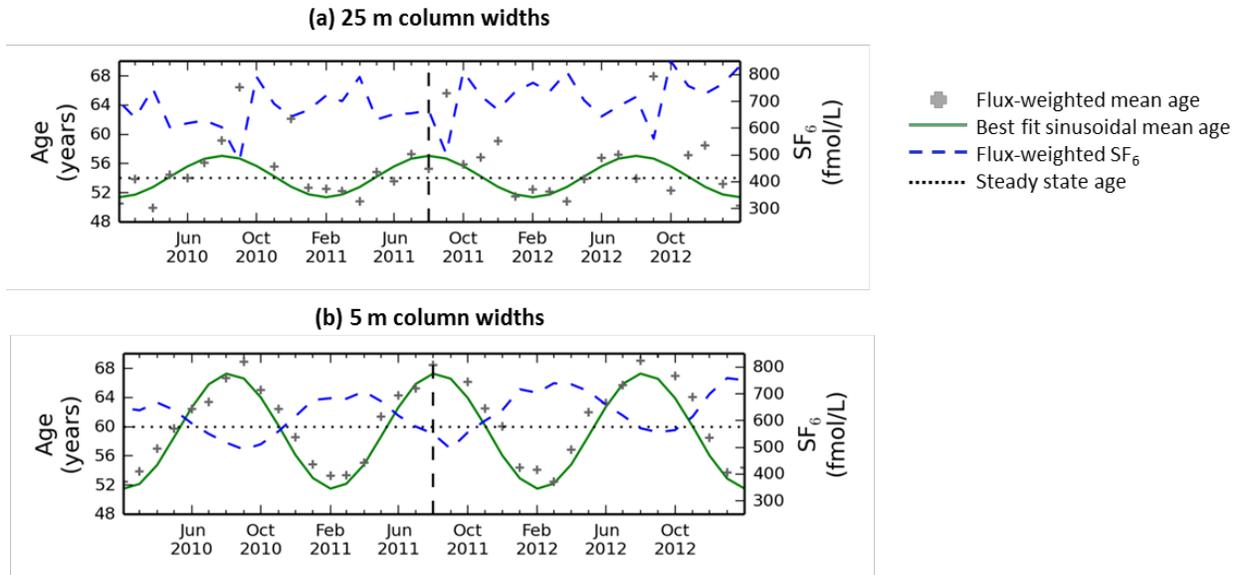


Figure 5.8 Effects of simulated grid size on the fit of simulated mean ages to a sinusoidal function for homogeneous aquifer with conductivity = 10 m/day and annual recharge = 40 cm/year. The dashed vertical line marks the timing of the maximum value of the sinusoidal for each panel.

5.2.1c - Comparison of the seasonal range of base-flow ages to ages derived from steady state scenarios

In order to further inspect the simulated range of seasonal ages we consider the variations in age that would result under two simplifying assumptions. The first assumption considers the aquifer to be completely mixed in the vertical direction such that age transport only occurs in one-dimension (1D) through a constant saturated thickness. Implementation of the groundwater transport equations can demonstrate that for 1D flow and transport perpendicular to the stream with spatially distributed recharge that varies cyclically in time, the variations in mean age are negligible. For the homogeneous aquifer with conductivity = 1 m/day and high amplitude recharge series, such a 1D analysis showed that the mean age varied by < 1% (results not shown). A contrasting simplifying assumption would consider poor mixing and strict stratification between ages in the aquifer, such that all variation in base-flow discharge occurred within a young fraction of

the discharge. Under this second assumption, the seasonal high in base-flow age would result from the absence of some young fraction at low base-flow conditions. **Figure 5.6b** shows that for the homogeneous aquifer with conductivity = 1 m/day, the base-flow discharge varies approximately $\pm 10\%$ around the mean annual value, such that at low base-flow the discharge is 20% less than at high base-flow. Using the assumption just described we may consider the impact of removing the youngest 20% of water from a steady state age distribution that is generated by a steady state recharge rate which reproduces the maximum base-flow discharge. Note that this assumes that the full distribution of base-flow ages is discharging at maximum base-flow; this assumption is not exact. This condition is assumed in **Figure 5.9a**. For the homogeneous aquifer with hydraulic conductivity = 1 m/day, we estimated the age distribution at maximum base-flow using the steady state age distribution associated with that base-flow level. This results in a mean base-flow age of 54 years; if, in order to move to low base-flow conditions, we remove the youngest 20% of water, the resulting mean age is equal to 67 years. This age is approximately 10% older than the maximum base-flow age from our transient simulations for the same system (61 years - see **Figure 5.6b**). We also note that we may vary the amount of water that we remove from the steady state age distribution at maximum base-flow in order to match the range of simulated ages from the transient simulation. We found that removing 11% of the youngest water from the steady state distribution at maximum base-flow discharge resulted in a mean age of 61 years; removing 1% of the oldest water resulted in a mean age of 52 years. In sum, the two assumptions of (i) vertically well-mixed and (ii) strongly stratified provide two end-members estimating how the base-flow age

might respond to seasonal changes in recharge and base-flow. The simulations in this study demonstrate that the change that occurs in real systems is somewhere in between.

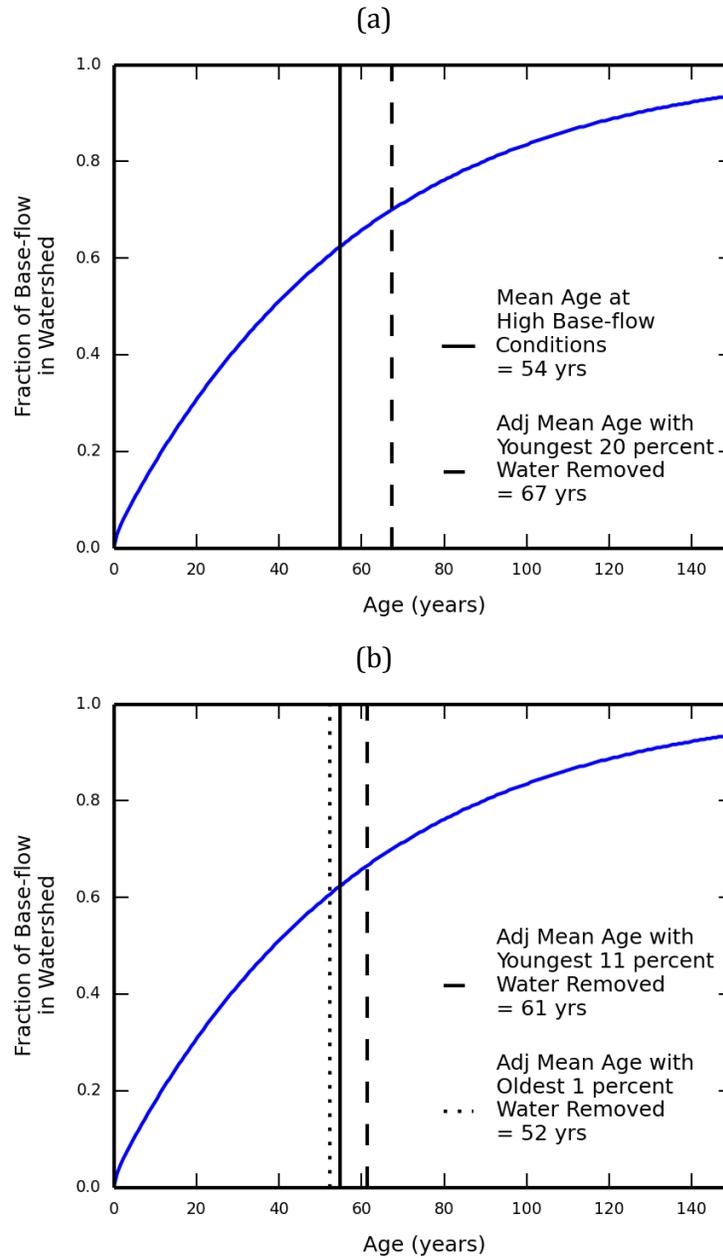


Figure 5.9 Steady state age distribution at maximum base-flow conditions for homogeneous aquifer with hydraulic conductivity = 1 m/day. Solid line shows mean base-flow age at 43 cm/year; this is the amount of steady state recharge that will reproduce the maximum base-flow from the transient simulation (**Figure 5.6b**). Dashed line shows the adjusted mean age that results from removing young ages. Dotted line shows the adjusted mean age that results from removing the old ages.

5.2.1d - Comparison of simulated mean ages and apparent ages from SF₆ concentration

For all cases shown in **Figure 5.6** and **Figure 5.7**, apparent ages derived from SF₆ measurements were younger than the mean ages (**Table 5.2**). The bias is the largest for the layered case. This is due to the contrasting ages in the surficial and lower aquifers and the absence of SF₆ from the lower aquifer. That is, the elevated SF₆ concentrations of the younger water mask the presence of older, tracer-free water. Stewart et al. (2010) describes a similar bias associated with the use of δ¹⁸O as an age tracer. For all systems, base-flow ages derived from the SF₆ concentrations more closely correlated to median than mean ages (**Table 5.2**).

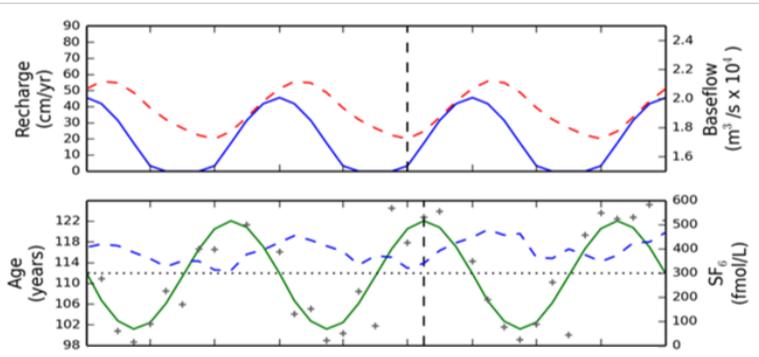
Table 5.2 Comparison of base-flow SF₆ apparent ages and mean age of discharging particle distribution. For Synthetic cases, SF₆ apparent ages calculated during 2011 (see **Figure 5.6**). For Morgan Creek, SF₆ apparent ages calculated during 2007 (see **Figure 5.11** and discussion later in the text).

Scenario	Seasonal Range of Mean Ages (years)	Seasonal Range of SF ₆ Apparent Ages (years)	SF ₆ Age Bias (Relative to Mean) (%)	Steady State Median Age (years)
Homogenous 0.6 m/day	56-66	27-30	52-55	42
Homogeneous 1 m/day	53-61	27-29	49-52	41
Homogeneous 10 m/day	51-67	27-32	44-52	42
Layered	47-66	13-18	72-73	17
Morgan Creek	36-39	13-14	63-64	12

5.2.1e - Comparison of variable annually averaged recharge

Figure 5.10 further considers the variability in the lag between low base-flow and maximum mean age for homogeneous hydraulic conductivity fields ($K = 1$ m/day) with different depths of recharge and resultant mean ages. Note that **Figure 5.10b** is identical to **Figure 5.6b**, though the vertical axes have been rescaled. Note also that for **Figures 5.10a-5.10d**, since the focus of the analyses is the relative timing of peaks, the vertical axes are not consistently scaled between subfigures. For all cases in **Figure 5.10** the lag between the minimum recharge and the minimum volumetric base-flow discharge was identical such that the minimum base-flow discharge occurred in November. The month at which the peak best-fit sinusoidal age occurred varied between August (synced with minimum recharge) and January (2 month lag following minimum base-flow). Interestingly, the amplitude of the intra-annual variation in the base-flow age time series decreased with the increases in annually averaged recharge from 20 to 40 cm then increased with increases in recharge to 60 and 80 cm. The amplitude of the SF_6 concentration time series increased dramatically with increasing recharge suggesting the uncertainty of inferences made from point measurements of SF_6 in systems with both (i) high annually averaged recharge and (ii) high inter-annual recharge variability.

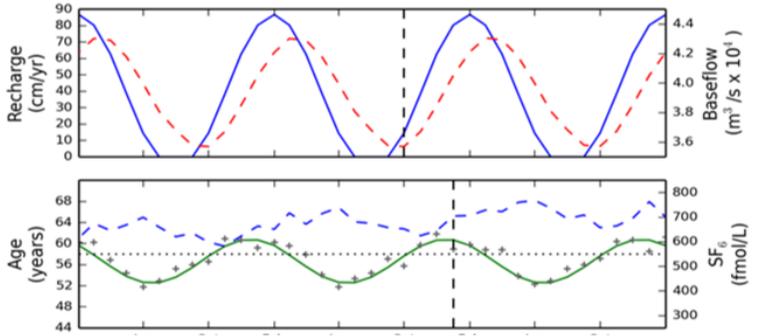
(a) Annual Recharge = 20 cm



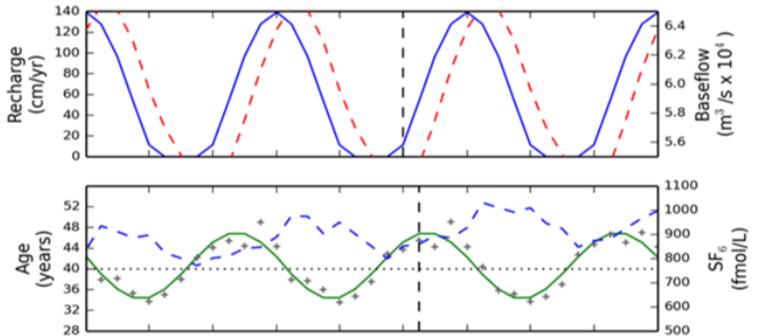
— Recharge
- - Base-flow

+ Flux-weighted mean age
— Best fit sinusoidal mean age
- - Flux-weighted SF₆
..... Steady state age

(b) Annual Recharge = 40 cm



(c) Annual Recharge = 60 cm



(d) Annual Recharge = 80 cm

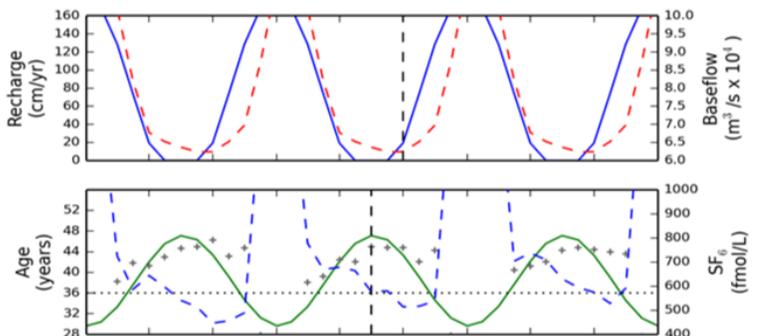


Figure 5.10 Comparison of base-flow age time series resulting from changes in annually averaged recharge and aquifer mean age.

5.2.1f - Comparison of recharge amplitude and recharge spatial distribution

Figure 5.11 compares the results of simulations that varied both the amplitude of the seasonally variable recharge (**Table 5.1**) and the spatial distribution of that recharge at each time step. Note that **Figure 5.11c** is identical to **Figure 5.6b**. The time series of discharging mean ages is very noisy under the Low Amplitude recharge scenarios (**Figures 5.11a** and **5.11b**), and under the Low Amplitude/Spatially Uniform scenario (**Figure 5.11a**) no periodic trend in discharging mean ages could be detected. For spatially variable recharge, the recharge at the discharge location is equal to 0 and the recharge at the point most distant from the discharge location (i.e., the watershed divide) is equal to twice the average recharge for that time step (**Figure 5.3**). Relative to the spatially uniform scenarios, this spatial configuration of recharge decreases the velocities and increases the travel time of the shortest flow paths but increases the velocities and decreases the travel time of the longest flow paths. For the Low Amplitude/Spatially Varied scenario (**Figure 5.11b**) this had no net effect on the long-term average mean base-flow age (compare to **Figure 5.11a**). However, when compared to the Spatially Uniform scenario (**Figure 5.11c**), the High Amplitude/Spatially Varied scenario resulted in a reduced long-term average mean base-flow and a larger amplitude of the base-flow age time series (**Figure 5.11d**).

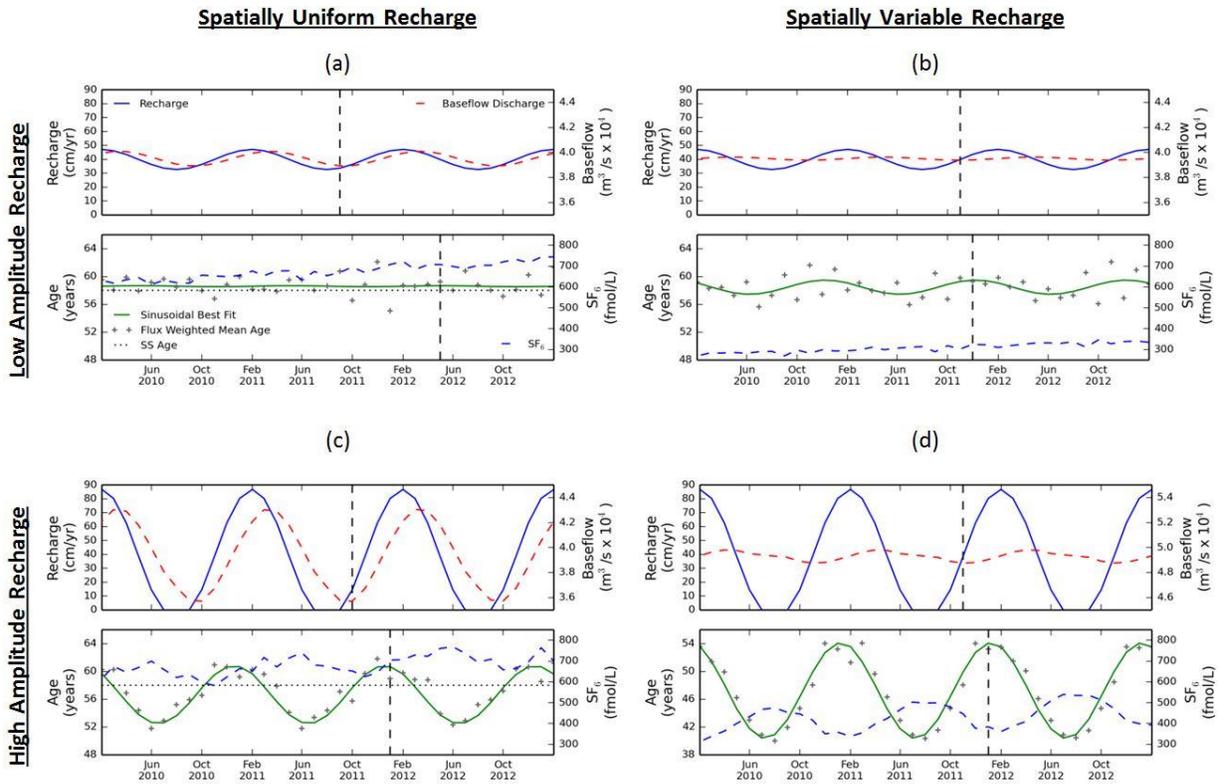


Figure 5.11 Comparison of base-flow age time series resulting from changes in the amplitude and spatial distribution of the seasonally-varying recharge. For all cases the homogeneous hydraulic conductivity = 1 m/day and the annually averaged recharge = 40 cm/year. The vertical dashed lines mark the timing of minimum base-flow and peak sinusoidal best fit age (in top and bottom panels, respectively, for scenarios a-d).

5.2.2 - Case study

Figure 5.12 shows the results of the Morgan Creek base-flow age simulation. The hydrogeology of Morgan Creek is similar to the layered synthetic case discussed above (**Figure 5.3b**) with a shallow surficial aquifer underlain by a semi-confining unit and a lower confined aquifer. This creates contrasting age regimes contributing to the base-flow discharge to the stream. According to the calibrated flow and transport model the steady state base-flow age is 37 years (see Chapter 3). See **Table 5.1** for the decomposition of the calibrated steady state recharge into quarterly time steps. **Figure 5.12** shows that, for the

calibrated flow and transport model, the range of seasonally-varying base-flow ages predicted by the particle-tracking method used in this study is 35-41 years, and the range of the best-fit sinusoidal trend is 36-39 years. As with the synthetic 2D layered system, the peak of the sinusoidal trend in base-flow mean age is synced with the minimum of the recharge signal (cf. **Figure 5.7**), though the maximum simulated monthly age typically occurs the following quarter in sync with the minimum base-flow discharge.

Also similar to the layered 2D scenario, our model suggests that the Morgan Creek system is subject to significant biases in apparent ages inferred from SF₆ concentrations. Apparent ages associated with simulated concentrations of SF₆ reflect the median, rather than the mean, steady state base-flow age (**Table 5.2**). This disconnect between SF₆ apparent ages and mean base-flow age could complicate attempts to link in-stream water quality with time-variable inputs at the land surface.

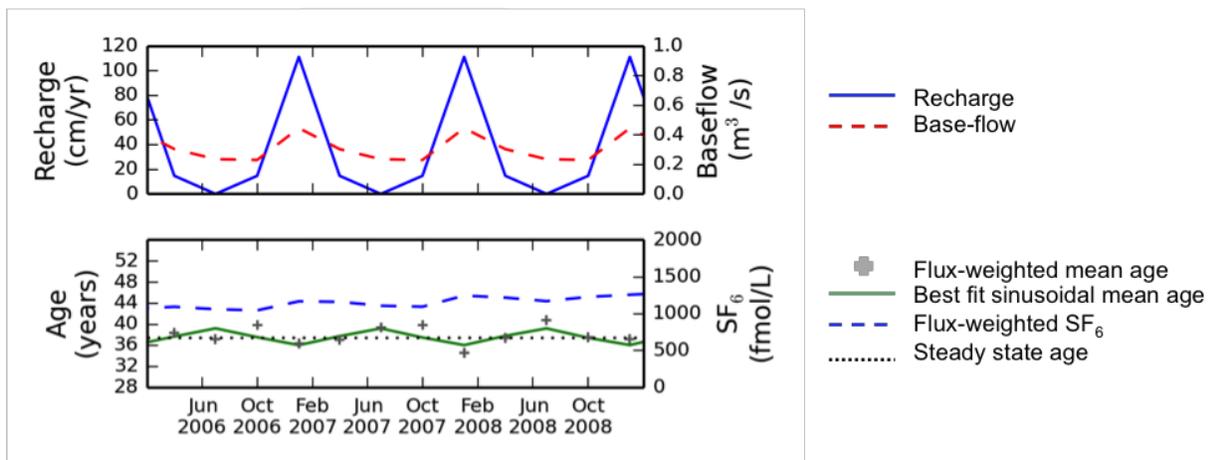


Figure 5.12 Simulated seasonal base-flow age and SF₆ concentration for Morgan Creek.

5.3 CONCLUSIONS AND FUTURE WORK

We simulated the transient delivery of base-flow age from subsurface to receiving stream as a function of seasonal changes in hydrology and aquifer storage for a variety of synthetic 2D aquifer configurations and for a fully 3D case study. We found that the timing of maximum base-flow age varied with both the hydraulic conductivity field and the annually averaged recharge, which determines the system mean age. The two assumptions of (i) an aquifer in which ages are vertically well-mixed and (ii) an aquifer in which ages are strongly stratified provide two end-members for estimation of how the base-flow age might respond to seasonal changes in recharge and base-flow. The simulations in this study found that the change that occurs in real systems is somewhere in between. For the cases that we investigated, apparent ages inferred from SF₆ measurements were biased young, with biases especially pronounced with layered hydrogeology in which discharge consists of shallow surficial flow mixed with a contrasting regime of much older water. Future work with this modeling technique should include further characterization of the sources and remedies for computational noise. This should include further investigation of the role played in computational noise by both (i) the spatial and temporal resolution of the flow simulation as well as (ii) the spatial and temporal distribution of particle starting locations.

5.4 REFERENCES

- Abrams, D. (2012). Correcting transit time distributions in coarse MODFLOW-MODPATH models. *Groundwater*, 51(3), 474-478.
- Bedekar, V., Niswonger, R. G., Kipp, K., Panday, S., & Tonkin, M. (2012). Approaches to the simulation of unconfined flow and perched groundwater flow in MODFLOW. *Groundwater*, 50(2), 187-198.

- Böhlke, J. K., Antweiler, R. C., Harvey, J. W., Laursen, A. E., Smith, L. K., Smith, R. L., & Voytek, M. A. (2009). Multi-scale measurements and modeling of denitrification in streams with varying flow and nitrate concentration in the Upper Mississippi River Basin, USA. *Biogeochemistry*, 93(1-2), 117-141.
- Botter, G. (2012). Catchment mixing processes and travel time distributions. *Water Resources Research*, 48(5).
- Ginn, T. R. (1999). On the distribution of multicomponent mixtures over generalized exposure time in subsurface flow and reactive transport: Foundations, and formulations for groundwater age, chemical heterogeneity, and biodegradation. *Water Resources Research*, 35(5), 1395-1407.
- Ginn, T. R., Haeri, H., Massoudieh, A., & Foglia, L. (2009). Notes on groundwater age in forward and inverse modeling. *Transport in Porous Media*, 79(1), 117-134.
- Gomez, J.D., & Wilson, J.L. (2013). Age distributions and dynamically changing hydrologic systems: Exploring topography-driven flow. *Water Resources Research*, 49(3), 1503-1522.
- Goode, D. J. (1996). Direct simulation of groundwater age. *Water Resources Research*, 32(2), 289-296.
- Harbaugh, A. W., Banta, E. R., Hill, M. C., & McDonald, M. G. (2000). *MODFLOW-2000, the US geological survey modular ground-water model: User guide to modularization concepts and the ground-water flow process*. US Geological Survey Reston, VA.
- Harman, C. J. (2014). Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed. *Water Resources Research*, 51(1), 1-30.
- Hirsch, R. M., Moyer, D. L., & Archfield, S. A. (2010). Weighted regressions on time, discharge, and season (WRTDS), with an application to chesapeake bay river inputs1. *JAWRA Journal of the American Water Resources Association*, 46(5), 857-880.
- Howden, N. J., Burt, T. P., Worrall, F., Mathias, S., & Whelan, M. J. (2011). Nitrate pollution in intensively farmed regions: What are the prospects for sustaining high-quality groundwater? *Water Resources Research*, 47(6).
- Hrachowitz, M., Soulsby, C., Tetzlaff, D., Dawson, J. J. C., Dunn, S. M., & Malcolm, I. A. (2009). Using long-term data sets to understand transit times in contrasting headwater catchments. *Journal of Hydrology*, 367(3), 237-248.
- Kazemi, G. A., Lehr, J. H., & Perrochet, P. (2006). *Groundwater age*. John Wiley & Sons.
- Kirchner, J. W., Feng, X., & Neal, C. (2001). Catchment-scale advection and dispersion as a mechanism for fractal scaling in stream tracer concentrations. *Journal of Hydrology*, 254(1), 82-101.
- Małoszewski, P., & Zuber, A. (1982). Determining the turnover time of groundwater systems with the aid of environmental tracers: 1. Models and their applicability. *Journal of Hydrology*, 57(3), 207-231.
- McCallum, J. L., Cook, P. G., Simmons, C. T., & Werner, A. D. (2014a). Bias of apparent tracer ages in heterogeneous environments. *Groundwater*, 52(2), 239-250.

- Maloszewski, P., & Zuber, A. (1998). A general lumped parameter model for the interpretation of tracer data and transit time calculation in hydrologic systems (journal of hydrology 179 (1996) 1-21) comments. *Journal of Hydrology*, 204(1-4), 297-300.
- McDonnell, J.J., McGuire, K., Aggarwal, P., Beven, K.J., Biondi, D., Destouni, G., Dunn, S., James, A., Kirchner, J., Kraft, P., Lyon, S., Maloszewski, P., Newman, B., Pfister, L., Rinaldo, A., Rodhe, A., Sayama, T., Seibert, J., Solomon, K., Soulsby, C., Stewart, M., Tetzlaff, D., Tobin, C., Troch, P., Weiler, M., Western, A., Worman, A., & Wrede, S. (2010). How old is streamwater? Open questions in catchment transit time conceptualization, modelling and analysis. *Hydrological Processes*, 24(12), 1745-1754.
- McGuire, K. J., & McDonnell, J. J. (2006). A review and evaluation of catchment transit time modeling. *Journal of Hydrology*, 330(3-4), 543-563.
- Michel, R. L. (1992). Residence times in river basins as determined by analysis of long-term tritium records. *Journal of Hydrology*, 130(1), 367-378.
- Mulholland, P. J., Helton, A. M., Poole, G. C., Hall, R. O., Hamilton, S. K., Peterson, B. J., Dahm, C. N. (2008). Stream denitrification across biomes and its response to anthropogenic nitrate loading. *Nature*, 452(7184), 202-205.
- Ogrinc, N., Kanduč, T., Stichler, W., & Vreča, P. (2008). Spatial and seasonal variations in $\delta^{18}\text{O}$ and δD values in the river Sava in Slovenia. *Journal of Hydrology*, 359(3-4), 303-312.
- Peters, N. E., Burns, D. A., & Aulenbach, B. T. (2013). Evaluation of high-frequency mean streamwater transit-time estimates using groundwater age and dissolved silica concentrations in a small forested watershed. *Aquatic Geochemistry*, 20(2-3), 183-202.
- Pollock, D.W., (1994). User's Guide for MODPATH/MODPATH-PLOT Version 3: A particle tracking post-processing package for MODFLOW, the U.S. Geological Survey finite-difference ground-water flow model: U.S. Geological Survey Open-File Report 94-464, 234 p.
- Rodhe, A., Nyberg, L., & Bishop, K. (1996). Transit times for water in a small till catchment from a step shift in the oxygen 18 content of the water input. *Water Resources Research*, 32(12), 3497-3511.
- Sanford, W. E., Nelms, D. L., Pope, J. P., & Selnick, D. L. (2011). *Quantifying components of the hydrologic cycle in Virginia using chemical hydrograph separation and multiple regression analysis*. US Department of the Interior, US Geological Survey.
- Sanford, W. E., & Pope, J. P. (2013). Quantifying groundwater's role in delaying improvements to Chesapeake Bay water quality. *Environmental Science and Technology*, 47(23), 13330-13338.
- Sanford, W.E., Casile, G., and Haase, K.B. (2015). Dating base flow in streams using dissolved gases and diurnal temperature changes. *Water Resources Research* (in press).
- Staver, K. W., & Brinsfield, R. B. (1998). Using cereal grain winter cover crops to reduce groundwater nitrate contamination in the mid-atlantic coastal plain. *Journal of Soil and Water Conservation*, 53(3), 230-240.

Stewart, M.K., Morgenstern, U., & McDonnell, J.J. (2010). Truncation of stream residence time: how the use of stable isotopes has skewed our concept of streamwater age and origin. *Hydrological Processes*, 24(12), 1646-1659.

Toth, J. (1962). A theory of groundwater motion in small drainage basins in central Alberta, Canada. *Journal of Geophysical Research*, 67(11), 4375-4387.

Varni, M., & Carrera, J. (1998). Simulation of groundwater age distributions. *Water Resources Research*, 34(12), 3271-3281.

Zheng, C., & Bennett, G. D. (2002). *Applied contaminant transport modeling* (2nd ed.). New York: Wiley-Interscience.

Chapter 6: The Impact of Kinematic Assumptions on Parameter Estimation for Advective-Dispersive Groundwater Systems⁴

6.0 INTRODUCTION

As discussed in Chapter 3 of this dissertation, information about groundwater age can be an important means of establishing rates of recharge and subsurface velocities and – through inference from these - the parameters critical to the prediction of contaminant transport and management impact (Sanford, 2011). While some researchers have developed methods for deriving dispersivities from the distribution of hydraulic conductivities (e.g., Sudicky, 1986; Rehfeldt et al., 1992), in most applications the distribution of conductivities is not known with the detail required for such a method and the dispersivity, like the effective porosity, must be estimated through model calibration (Zheng and Bennett, 2002).

The integration of groundwater age information into modeling efforts requires decisions about two fundamental questions, as well as attention to the model's sensitivity to those decisions. These decisions are summarized in **Table 6.1**. First, how is the age information represented in the parameter estimation process? For example, is the calibration target for an inverse modeling procedure a sampled tracer concentration, or is it instead a measure of age inferred from the tracer? The advantage of the former is that fewer interpretations are interposed between the model and its constraining data, since the conversion of a sampled tracer concentration to an age is itself a modeling exercise with built-in assumptions. For instance, lumped parameter models assume some distribution to describe a mixture of ages

⁴ This chapter is being prepared for submission to *Groundwater*.

at a sampling point (Maloszewski and Zuber, 1982; Cook and Bohlke, 2000; Massoudieh and Ginn, 2011). Alternatively, the simplest assumption treats the water sample as an unmixed packet of a single age; under this assumption the age inferred from the tracer concentration is often termed the apparent age (McCallum et al., 2014b). If the calibration data is represented as interpreted age, a second question follows: how should the model be compared to the data? That is, given the representation of groundwater age data as an age (rather than as a tracer concentration), how should the model simulate age? As described below, age transport (like the transport of other solutes) may be simulated as an advection-only process or as a more complicated process involving dispersion. Most basically stated: what are some possible consequences of treating the system as advection-only when in fact dispersion is present? These questions are of computational significance since, for a large three-dimensional model, the time required to simulate an advection-only, particle tracking measure of age is much faster than the full numerical solution of an ADM, making advection-only simulations of age much more efficient given the likely need to run the age simulation many times during the calibration process.

Table 6.1 Options for inclusion of groundwater age information in model calibration

<u>Calibration Target</u>	<u>Simulated Output</u>	<u>Transport Simulation Type</u>
Tracer Concentration	Tracer Concentration	Kinematic or ADM
Interpreted Age	Age	Kinematic or ADM

6.0.1 - Potential problems with kinematic measures of age

An advection-only measure of groundwater age assumes that a tracer measurement represents the mean age at a point and that the tracer mass arrived at that point along a single flow path and without mixing. This conceptualization of the elapsed time for groundwater flow may be referred to as the kinematic age (Varni and Carrera, 1998). Kinematic-age assumptions allow the calibration of porosity against the age information using an advection-only particle tracking simulation tool such as MODPATH (Pollock, 2012). Szabo et al. (1996) and Hunt et al. (2006) used kinematic assumptions and apparent ages derived from CFCs and ^3H to calibrate groundwater flow models for a New Jersey coastal aquifer and a Wisconsin lake system, respectively. Sanford and Pope (2012) similarly used kinematic assumptions and apparent ages derived from CFCs and ^3H to calibrate the porosity for a groundwater model of the Delmarva Peninsula. Varni and Carrera (1998) demonstrated the shortcomings of the kinematic conception of age for heterogeneous aquifers, namely that kinematic ages are unstable with respect to location and that a groundwater sample at an observation point should in fact be represented as a distribution of ages. Several researchers have developed methods for simulating the age of groundwater subject to other transport factors. Common to these methods is the

recognition that a sample from any point in a non-homogenous aquifer is a composite of multiple converging flow paths subject to exchanges due to macrodispersion and diffusive exchange with waters of varying age (McCallum, 2014b). The age of any groundwater sample should therefore be more rigorously viewed as the mean of a mixture – i.e., a distribution – of the travel times accumulated along those flow paths and subject to dispersive exchange. After various authors we refer to the distribution of ages at a point as the travel time distribution (TTD) and we define ‘kinematic TTD’ as the distribution of advection-only travel times collected at a location of interest such as a well screen or receiving stream. In contrast to the kinematic TTD, multiple methods exist for generating a TTD that accounts for both advective and dispersive effects at an observation location in an aquifer. Varni and Carrera (1998) developed a recursive method for identifying moments of the TTD (but not the distribution itself). Neupauer and Wilson (2004) derived the backward adjoint method, which for a given flow and transport model simulates the TTD as a distribution of probable travel times and transport pathways to an observation point. Jury and Roth (1990) showed that for a steady state system the TTD at a point is equivalent to the breakthrough curve due to a unit pulse of a conservative substance at the inflow boundary. For finite difference approximation of the groundwater flow and transport equations, the unit pulse boundary condition may be implemented using a solute transport code such as MT3DMS (Zheng, 1999), with an initial unit mass dissolved in the mobile zone of each grid cell containing the water table. In addition to these methods of generating a travel time distribution at a point, Goode (1996) showed that mean groundwater age may be calculated using a modified form of the ADM; this is commonly referred to as the direct age method. For purposes of numerical simulation the direct age method produces a single

age rather than a TTD at each model node, such that multiple model nodes must be combined to create a TTD. Gusyev et al. (2014) used this method to generate the TTD for rivers using a groundwater flow and transport model that simulates the direct age at all cells contributing baseflow discharge to the receiving stream.

6.0.2 – Examining the calibration impact of assuming kinematic age

While the potential limitations associated with the simplifying assumptions of the kinematic conception of age are well known, few if any studies have considered the impact of kinematic assumptions on automated parameter estimation processes or compared a model calibrated with only kinematic representations of age to those calibrated with fully advective-dispersive representations of age. The present study uses the automated calibration of several synthetic two-dimensional (2D) aquifers to observe the consequences of kinematic assumptions for the resulting parameter estimates. To isolate the impacts of these kinematic assumptions on calibration, we assume for the synthetic cases that accurate age determinations are available. That is, for the synthetic cases we do not explore the variety of tracer-specific biases inherent in the derivation of an age measure from a sampled tracer concentration, including those biases that result from the shape of the atmospheric time series and the exacerbation of those biases by aquifer heterogeneity. These complexities have been explored in two recent papers by McCallum et al. (2014a and 2014b).

Further, the synthetic case studies considered in this chapter assume that the ADM and its Fickian approximation of the exchange among flowpaths can accurately represent subsurface solute transport in many systems and, by extension, the hydrodynamic

dispersive mixing of groundwaters of different ages. Several authors have shown the limits of the ADM assumption that (i) variability in velocity is generated by (ii) random variability in hydraulic conductivity (Zheng and Bennett, 2002). Perhaps the most notable example of these limits is the highly heterogeneous MADE test site, where the inability of the ADM to explain the most significant spatial and temporal features of intensively-monitored experimental plumes led to a more satisfactory model of mass exchange between mobile and immobile zones (Feehley et al., 2000; Harvey and Gorelick, 2000; Zheng, et al., 2011). However, in contrast to the MADE site, the results of tracer tests at more homogenous sites like Borden (Freyberg, 1986) and Cape Cod (Garabedian et al., 1991) are well-described by an ADM. These latter sites serve as the paradigm for the synthetic cases that are considered here.

As a supporting question for our synthetic studies, the present chapter also considers whether the use of a registration volume – i.e., enlarging the number of advective pathlines contributing to the TTD at an observation location – can serve as a surrogate for dispersive effects. In their analytical work on TTDs, Varni and Carrera (1998) described the TTD for a point in an advective-dispersive system as “the distribution of kinematic ages one would obtain by perturbing the measurement location”, which suggests that kinematic tools such as particle tracking routines may be able to approximate the true TTD at a point through the construction of an ensemble of kinematic ages collected within a registration volume around that point (**Figure 6.1**). Troldborg et al. (2008) used a registration volume and kinematic ages to construct a TTD that they convolved with a particle-based contaminant input function in order to predict the contaminant concentration at an observation point. Weissman et al. (2002) used backward particle tracking from a registration volume with

the addition of random walk perturbations to simulate dispersion of environmental tracers in an aquifer system. Both of these studies assigned a fixed registration volume; neither calibrated transport parameters against groundwater age and consequently did not consider the impact of the registration volume on the parameter estimation process. In addition to using fixed registration volumes to construct kinematic TTDs, the present study also investigates the use of adjustable registration volumes as an option to improve calibration. These methods are described in Section 6.1.

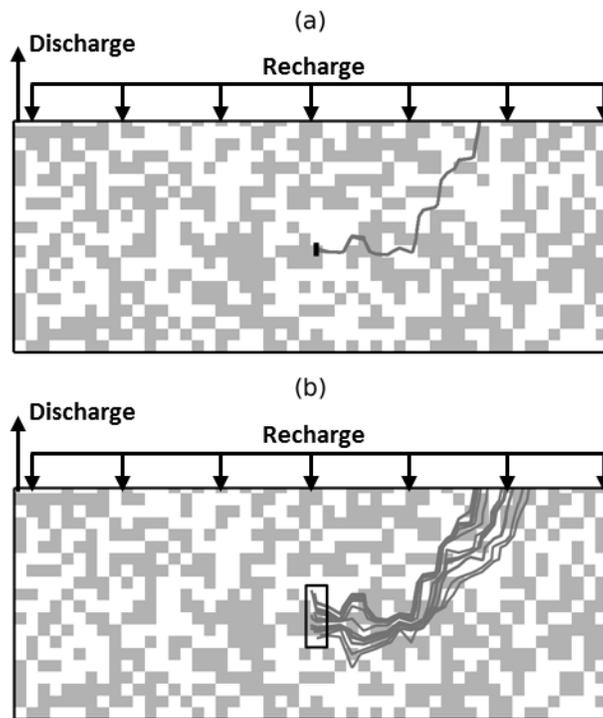


Figure 6.1 Comparison of simulated advective flowpaths to (a) a wellscreen and (b) a registration volume surrounding the well screen. Flowpaths do not include dispersive effects and are solely subject to system hydraulics. The spatially-correlated binary conductivity field is described in the text.

6.0.2 – Examining the calibration impact of aggregating time-variable tracer information into steady state measures of age

As a related question, we additionally consider the impact of aggregating time-varying data from multiple tracer species into a single set of steady-state calibration targets that are expressed as age rather than tracer concentration. Thus we examine the informational effect of converting multiple tracer concentrations at an observation site into a single measure of inferred age at that site and subsequently simulating the transport of age rather than the transport of a tracer (cf. **Table 6.1**). As already discussed, this conversion requires some assumptions about the underlying age distribution of the groundwater sample (Cook and Böhlke, 2000; Maloszewski and Zuber, 1982; Massoudieh and Ginn, 2011; McCallum et al., 2014a). Most studies that have used inferred ages as calibration targets have simultaneously assumed a piston-flow mechanism of subsurface tracer transport is responsible for the underlying tracer concentration and aggregated multiple tracer observations into a single age measure (e.g. Reilly et al., 1994; Szabo et al., 1996; Hunt et al., 2006; Sanford and Pope, 2012). Turnadge and Smerdon (2014) have recently argued that, due to the uncertainties associated with the tracer-to-age conversion of tracer observations, the tracer itself rather than interpreted ages should be modeled. To our knowledge, no studies have considered the implications of this conversion for the calibration of a field model; the present study consequently examines this aggregation by comparing the calibrated flow and transport model that results from using a consolidated dataset of apparent ages to models that are calibrated against the original datasets of tracer concentrations as described in Chapter 3 of this dissertation.

6.1 METHODS

6.1.1 – Synthetic Aquifer Configurations and ‘True’ System Values

Three synthetic 2D confined aquifers with different spatial distributions of heterogeneity were used to examine the impact of how age is modeled (i.e., simulated kinematic age or simulated advective-dispersive age) on calibration results. **Figure 6.2** shows the boundary conditions and the three synthetic heterogeneity cases that were considered by variously representing the subsurface as: ‘Binary’, which used a spatially-correlated random binary field; ‘Lenses’, which used an aquifer matrix with randomly located high permeability lenses (with each lens randomly assigned one of two conductivities, for a total of three conductivities for the system); or ‘Stratified’, which used a stratified sequence of three conductivities (aquifer – confining unit – aquifer). It is important to note that for each case the spatial distribution of hydrogeologic facies were assumed to be known, so that during the inverse modeling process only the parameters assigned to each unit – rather than any spatial characterization of the units - were calibrated. This includes the random binary case; for that scenario the spatial description was stochastically generated, but once generated was considered to be known.

Each of the three heterogeneity configurations had a ‘true’ constant porosity of 0.30. For each configuration, three different constant longitudinal dispersivities (0, 10, or 25 m) were tested. The heterogeneity configuration (Binary, Lenses, or Stratified) and the system dispersivity (0, 10, or 25 m) are hereafter designated by the form *Configuration_Dispersivity* (e.g., ‘Binary_0’ refers to the case of the randomly generated binary heterogeneity configuration with no longitudinal dispersivity).

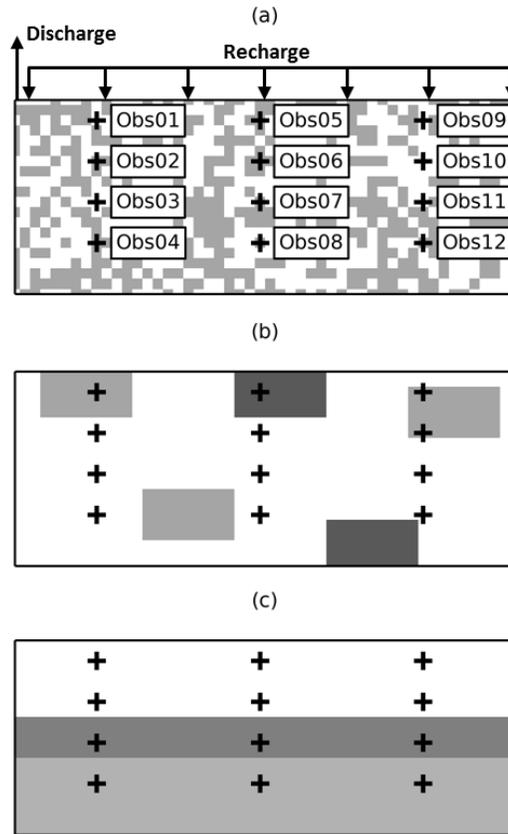


Figure 6.2 Aquifer conductivity configurations and locations of calibration targets for synthetic cases. (a) Binary case; (b) Lenses case; (c) Stratified case. The boundary conditions and observation names shown for the Binary case apply to all three cases. Model grid dimensions are 50 columns x 20 layers. Model cells are 25 m wide by 5 m thick.

For each of the nine possible combinations of heterogeneity and dispersivity, calibration targets included system discharge as well as heads and groundwater ages at 12 observation locations (**Figure 6.2**). For each scenario, the values of these calibration targets were generated by steady state simulation with the addition of Gaussian noise to represent sampling uncertainty. We will refer to these synthetically-generated target values as the ‘true’ representation of the system. Weights for each observation type and location were assigned as an inverse function of the standard deviation of the sampled set. Heads and

baseflow discharge were simulated with MODFLOW-2005 (Harbaugh, 2005). For the zero dispersivity case, true groundwater ages were simulated by backward-tracking particles from a 1-meter screen length using MODPATH (Pollock, 2012). For the cases with dispersivity, true groundwater ages were simulated with MT3DMS (Zheng and Wang, 1999) using the total variation diminishing (TVD) solver and a zeroth-order production term to implement the direct age method (Zheng, 2009; Goode, 1996). In order to generate the steady state groundwater age field with this method the transport simulation was run with a steady state flow field until the spatial distribution of age reached steady state. The TVD solver was used to generate the target ages in order to minimize the effects of numerical dispersion, though it is important to note that in the ADM calibration tests, the finite difference solver was used based on the presumption of a dispersive system.

6.1.2 - Calibration Method, Measures of Simulated Groundwater Age, and Performance Metrics

Automated calibration was performed using UCODE (Poeter et al., 2005) as described in **Section 3.2.3** of this dissertation. The parameter values for recharge, hydraulic conductivities, porosity, and (when the ADM was employed) dispersivity were estimated for each of the nine aquifer configurations by calibrating the model against the generated targets. For all aquifer cases, steady recharge was first derived from the noisy discharge value and then this recharge was assigned during the estimation of the other parameters.

For each model, flow and transport parameters were calibrated using both simultaneous and sequential methods. The simultaneous methods used all available head and age information in order to calibrate flow and transport parameters at the same time. The sequential method first calibrated hydraulic conductivities against head observations then

assigned those conductivity values in order to calibrate porosity against age measurements. The simultaneous methods are distinguished from one another by the method of simulating the travel time distribution at the observation location (**Table 6.2**). Three of these methods used kinematic measures of age and the fourth used an ADM measure of age. 'WellOnly' refers to the use of particle tracking to simulate groundwater age by backward tracking advective pathlines from the 1-m well screen to the recharge location and then computing the mean age of the resulting kinematic TTD. 'RegVol' refers to the use of backward particle tracking from a registration volume centered on the well to similarly generate a kinematic TTD and an associated mean age; for the 2D synthetic examples the registration volume reduces to a rectangle (**Figure 6.1**). For the RegVol method, the dimensions of the registration area were calibrated parameters. As a comparison, and in order to simplify the calibration by reducing the number of estimated parameters, a RegVol_Fixed method was also tested; this method generated kinematic TTDs using a fixed 10-m by 10-m area centered on the well screen. 'UnitPulse' refers to the use of a unit pulse boundary condition with an ADM to generate a travel time distribution and an associated mean age. Mean ages were computed as the 1st moment of the TTD. As mentioned above, the UnitPulse method used the finite different (FD) solver to simulate the transport of age with dispersion. For the sequential calibration method (labeled 'Sequential'), age was simulated using the WellOnly method (i.e., backward particle tracking from the well screen to the recharge point in order to construct the kinematic TTD and derive the mean age).

Table 6.2 Calibration Methods and Calibrated Parameters for Synthetic Cases

<u>Name</u>	<u>TTD simulation method</u>	<u>Calibrated Parameters</u> ^a
WellOnly	Kinematic	Simultaneous: Conductivities, Porosity
RegVol	Kinematic	Simultaneous: Conductivities, Porosity, Size of Registration Area
RegVol_Fixed	Kinematic	Simultaneous: Conductivities, Porosity (Registration Area Assigned)
UnitPulse	ADM	Simultaneous: Conductivities, Porosity, Dispersivity
Sequential	Kinematic	Sequential: Conductivities, then Porosity

^aFor all cases the recharge parameter was estimated independently prior to the estimation of remaining parameters

6.1.3 - Calibration of the Upper Chester model using ages inferred from environmental tracers

In order to examine the impact of aggregating multiple tracer measurements into single measures of steady state age we calibrated the Upper Chester flow and transport model (see Chapter 3 of this dissertation) against interpreted steady state ages. For calibration methods that used aggregated interpreted ages rather than tracer concentrations as transport observations we simulated the age at each well by calculating the mean of the age distribution $g(\tau)$ (Eq. 3.2). To consolidate these multiple environmental tracer concentrations into units of steady state age we aggregated the tracer information available at each well into a single inferred age for that well. In order to consider the impact of the aggregation scheme on calibration results, we used two methods of estimating the steady state age. We refer to these estimates as the Curated and Uncurated steady state age targets. Curated targets used the apparent ages and associated uncertainty (i.e., sample standard deviation) estimated by Busenberg and Plummer (2000). Curated targets also privileged $^3\text{H}/^3\text{He}$ and SF_6 information when consolidating multiple apparent ages at a single observation site. $^3\text{H}/^3\text{He}$ was privileged because it approximates an ideal tracer; if a

$^3\text{H}/^3\text{He}$ measurement existed for a particular location, the $^3\text{H}/^3\text{He}$ -derived age was used as the steady state age. SF_6 was privileged for the curated ages because the atmospheric concentration of SF_6 has continued to increase in recent decades as atmospheric CFCs have plateaued and declined; the SF_6 signal in young waters is more distinct, and the SF_6 -derived ages were therefore used for groundwaters apparently recharged after 1988. Furthermore, SF_6 is assumed to be less subject to mass loss (see discussion in **Chapter 3** of this dissertation). An unweighted average of ages derived from CFC and SF_6 data were used for remaining samples. The standard deviations of the age that were estimated by Busenberg and Plummer (2000) were combined into a single measurement variance and used to weight the observations during calibration. In contrast to the Curated calibration targets, Uncurated targets for each site were constructed as an unweighted average of all available SF_6 , CFC, and $^3\text{H}/^3\text{He}$ measurements. Each dissolved tracer concentration was converted to an apparent age by matching the sampled concentration to a corresponding year on the annually averaged atmospheric input time series (**Figure 3.1**). This is simply the inversion of the method described in the Section 3.2 for calculating the tracer concentrations at each well as a function of the travel times; as such, the calculation of age from tracer concentration also assumed advective-only transport with recharge conditions of 10 C and 2 cm^3/L excess air. Standard deviations for each Uncurated steady state age were calculated by multiplying the derived mean age by the coefficient of variation of the tracer sample.

In general, the mean Curated ages tend to be slightly younger than the Uncurated ages (**Table 6.3**). This is mostly due to privileging the tritium-based ages that were derived by Busenberg and Plummer (2000). More significantly, the Uncurated targets include much

larger standard deviations for older ages. That is, for water samples with very low CFC concentrations, the ages interpreted by Busenberg and Plummer included a relatively high confidence that the trace CFC measurement is in fact a signal from the beginning of the CFC time series, rather than noise from contamination or non-simulated, non-advective mixing (see Table 3 in Busenberg and Plummer, 2000). Because the inverse of the standard deviations were used to weight observations during the calibration procedure, increasing the uncurated standard deviations decreased the relative importance of older, deeper observations.

Table 6.3 Calibration targets for Upper Chester calibrations using steady state interpreted age.

Site Name	Curated Targets		Uncurated Targets	
	Mean	Std	Mean	Std
KEBe50	8.6 ^b	2.6	6.1	3.8
KEBe52	7.1 ^a	2.0	8.4	2.8
KEBe53	2.6 ^b	2.6	5.3	3.1
KEBe59	5.9 ^b	1.0	5.4	0.6
KEBe61	16.5 ^a	2.5	18.7	4.3
KEBe62	3.1 ^a	2.9	5.4	2.8
KEBe63	21.5 ^c	2.6	21.3	3.3
KEBe64	10.2 ^c	3.5	10.2	4.1
KEBe158	8.5 ^c	3.6	8.5	4.1
KEBe159	34.6 ^c	9.4	33.6	28.3
KEBe160	19.8 ^a	4.1	25.7	6.3
KEBe161	14.4 ^a	2.4	11.6	2.7
KEBe162	31.6 ^c	6.9	29.8	8.7
KEBe163	7.8 ^a	2.2	9.4	2.6
KEBe164	21.9 ^c	3.0	20.9	4.4
KEBe165	18.4 ^a	4.2	22.7	4.9
KEBe166	9.2 ^c	2.8	7.3	3.0
KEBe167	6.8 ^c	2.8	2.6	2.9
KEBe189	43.6 ^c	7.7	43.2	29.7
KEBe192	4.1 ^c	3.5	7.1	1.5
KEBe195	20.9 ^c	1.2	21.1	2.1
KEBe206	17.6 ^c	1.3	18.4	1.4
KEBe207	20.6 ^c	1.4	20.9	2.4
KEBe208	31.1 ^c	0.7	31.1	1.7
KEBe210	40.4 ^c	14.6	45.3	43.6
KEBe211	45.4 ^c	1.5	46.0	16.2
KEBe212	8.9 ^c	2.1	9.6	1.8
KEBe216	9.8 ^c	3.2	10.8	1.9

^a ³H + ³He age used^b SF₆ age used^c Average tracer age used (no ³H + ³He data)

6.2 RESULTS AND DISCUSSION

6.2.1 – Results and Discussion for Synthetic Cases

In this section the performances of the various calibration methods using the Synthetic cases are evaluated using the following metrics for comparison: (1) the weighted sum of squared errors (WSSE); (2) the accuracy of the porosity estimate; and (3) the accuracy of the simulated mean baseflow age (i.e., age of the system discharge). In surface water systems where contaminant delivery is dominated by baseflow, the travel time distribution of baseflow ages provides an important integrated hydrological metric for the susceptibility of receiving waters to contamination. The capacity of the calibrated model to simulate the ‘true’ baseflow age may therefore be understood as an indicator of the model’s capacity to make accurate transport predictions.

Table 6.4 and **Table 6.5** show the calibrated porosity values and simulated discharge ages, respectively, for each scenario. Because the RegVol_Fixed method performed more poorly than the RegVol method it is therefore not included in the results or discussion. **Figure 6.3** shows the WSSE for each scenario as a function of calibration iteration. As a reminder of the problem of non-uniqueness in inverse modeling, it is useful to note that the WSSE is not necessarily an indicator of relative model accuracy. For the Stratified_10 scenario the UnitPulse method resulted in the lowest WSSE (**Figure 6.3h**) but the RegVol and WellOnly methods resulted in more accurate porosity estimates (**Table 6.4**) and simulated baseflow age (**Table 6.5**). For this case, the dispersivity parameter appears to have enabled better fit to the noise in the age observations while worsening the estimation of the actual parameter set.

Table 6.4 Calibrated system porosities for different calibration methods and Synthetic aquifer configurations. The true system porosity is 0.30. See Table 2 and Figure 2 for descriptions of the calibration methods and aquifer configurations, respectively.

Configuration	Calibration Method			
	WellOnly	RegVol	UnitPulse ^a	Sequential
Binary_0	0.32	0.25	0.32 [0]	0.27
Binary_10	0.44	0.33	0.34 [4]	0.39
Binary_25	0.41	0.37	0.37 [8]	0.39
Lenses_0	0.34	0.31	0.52 [0]	0.31
Lenses_10	0.44	0.43	0.43 [0]	0.33
Lenses_25	0.51	0.46	0.41 [15]	0.33
Stratified_0	0.32	0.30	0.36 [0]	0.26
Stratified_10	0.28	0.27	0.23 [15]	0.28
Stratified_25	0.41	0.41	0.26 [37]	0.46

^a Values in brackets indicate the estimated dispersivity (m)

For all simultaneous scenarios, the WellOnly method was more accurate than the UnitPulse method at estimating porosities for zero dispersivity cases and the UnitPulse method was more accurate than the WellOnly method at estimates for the highest dispersivity cases. The relative performances of those two methods were mixed for the intermediate dispersivity cases. This suggests that as true system dispersivity increases, the capacity of kinematic ages to translate the available system information into accurate parameter estimates decreases. It is important to note, however, that for the Lenses scenarios

involving dispersion, the sequential use of a kinematic method resulted in a better porosity estimate and simulated baseflow age than the simultaneous use of an ADM method (**Table 6.4** and **Table 6.5**); in fact, for the Lenses scenarios, the Sequential method resulted in a more accurate porosity and baseflow age estimate than any of the simultaneous methods. This is in contrast with previous literature that has shown that the simultaneous inclusion of flow and transport information improves inverse modeling (Anderman et al., 1996; Ginn et al., 2009). The superior performance of the Sequential method in the Lenses scenarios here suggests that the benefits of simultaneous inclusion of flow and transport information may apply more to some aquifer configurations or calibration routines than others.

6.2.1.a - Impact of applying kinematic-only methods to dispersive systems

The effect of dispersion on groundwater age at a point in space is a function of the subsurface conductivity field, the direction of flow, the amount of dispersion, and the relative location of boundary conditions. The impact of system dispersion on a kinematic calibration method therefore depends on the effect of dispersion at the most influential observation points. The influence of a single observation on a single parameter estimate can be quantified with regression statistics such as Cook's D or DFBETAS (Hill and Tiedeman, 2006).

With the exception of the Stratified_10 scenario, the kinematic WellOnly method resulted in porosity estimates and baseflow age predictions that were 30-70% too high when applied to dispersive systems (**Table 6.4** and **Table 6.5**). The Binary_10 scenario illustrates and explains the reasons for this general trend. According to the DFBETAS statistics for the Binary_10 cases (not shown), the porosity estimate using the WellOnly method was most

strongly influenced by age observations at Obs04 and Obs08 (see **Figure 6.2** for locations). For each of these locations, the true groundwater age at the observation increased with dispersion due to the mixing of older water (**Table 6.6**). In an effort to better fit the age at those observation locations, but with only kinematic flowpaths available to account for the TTD, the WellOnly calibration method overestimated the porosity in order to reduce groundwater velocities (and thus simulate older groundwater ages). In contrast, the UnitPulse method was able to account for the older ages through dispersive mixing and thus more accurately estimate the porosity (**Table 6.4**).

Table 6.5 True and simulated mean base-flow age (years) for different calibration methods and Synthetic aquifer configurations. See Table 6.2 and Figure 6.2 for descriptions of the calibration methods and aquifer configurations, respectively.

Configuration	True Baseflow Age	Baseflow Age of Calibrated Parameter Set			
		WellOnly	RegVol	UnitPulse	Sequential
Binary_0	77	79	59	61	64
Binary_10	82	109	76	83	89
Binary_25	82	99	88	90	93
Lenses_0	81	90	81	76	82
Lenses_10	82	116	116	115	89
Lenses_25	81	136	123	108	88
Stratified_0	42	45	42	18	36
Stratified_10	40	39	37	30	39
Stratified_25	40	57	56	33	64

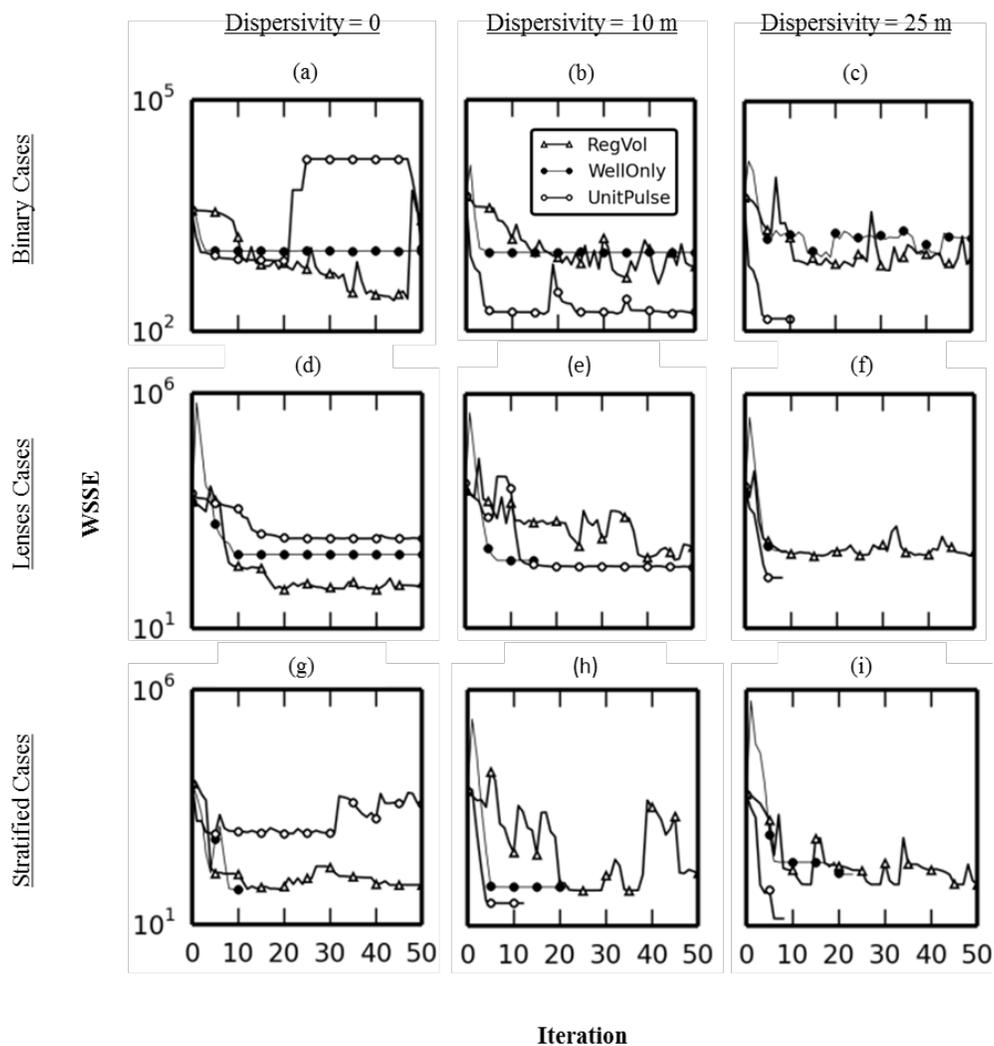


Figure 6.3. WSSE progression for different heterogeneity scenarios and calibration methods. Incomplete time series indicate convergence of the calibration process prior to 50th iteration. Porosity estimates reported in Table 3 are from iteration with minimum WSSE (not necessarily the 50th iteration). Legend in panel (b) applies to all panels.

The Stratified_25 case provides a slightly more complicated example and shows the impact of ignoring dispersion on both the porosity estimate and the conductivity estimate. For the Stratified_25 WellOnly method, the porosity estimate was most strongly influenced by site Obs01 in the upper aquifer (see **Figure 6.2**). Because the true dispersive age at Obs01 is older than the true kinematic age (**Table 6.6**), the WellOnly method over-estimated the porosity in order to reduce kinematic velocities and increase travel time (**Table 6.4**). By comparing the WellOnly calibration results for all three Stratified scenarios (**Figure 6.4**), we may note that the upper aquifer conductivity is consistently predicted regardless of the magnitude of dispersivity. Since age varies substantially with dispersivity (see **Table 6.6**), the calibrated upper aquifer conductivity is insensitive to age information and therefore not sensitive to dispersive effects on age. However, this was not true for the estimates of the confining layer and lower aquifer conductivities. For the Stratified_25 scenario, those parameters are sensitive to the age information at sites Obs03 and Obs07. These sites are located in the confining layer, where the effect of dispersion is to reduce the groundwater age (**Table 6.6**). To accommodate these younger ages, the regression raised the conductivity estimates for both the confining layer and the lower aquifer in order to compensate for the reduced velocities associated with the exaggerated porosity.

Table 6.6 True age (years) as a function of dispersivity at select observation sites. See **Figure 6.2** for location of sites.

	<u>Dispersivity (m)</u>		
	0	10	25
Binary Obs03	41	60	72
Binary Obs04	71	88	103
Binary Obs08	102	208	152
Stratified Obs01	8	6	13
Stratified Obs03	267	211	161
Stratified Obs04	201	235	258
Stratified Obs07	100	85	88

6.2.1.b - Impact of applying dispersive methods

While the most basic aim of this study was to assess the calibration impacts of kinematic assumptions in advective-dispersive systems, we may also consider the impact of applying an ADM method to the same scenarios. As noted, for most scenarios involving actual system dispersion, the use of the UnitPulse ADM method improved upon the use of kinematic methods. However, the accuracy of the optimized parameter set that was identified using the UnitPulse method varied widely across scenarios. Both porosity and baseflow age predictions were most accurate for the Binary scenarios and least accurate for the Lenses scenarios. For the Lenses scenarios the porosity was overestimated by 36-75% (**Table 6.4**). For the Lenses_10 and Lenses_25 scenarios this resulted in similar overestimates of baseflow age (**Table 6.5**); for the Lenses_0 scenario the baseflow age prediction is more accurate because the calibration compensated for the exaggerated porosity by over-estimating the aquifer conductivity by 80% (not shown). Two particular

trends are evident from the porosity estimate and the accompanying dispersivity estimate for the UnitPulse methods (**Table 6.4**): (i) for zero dispersivity scenarios, the calibration correctly estimated the system dispersivity parameter but over-estimated system porosity; and (ii) for the scenarios with dispersivity, the calibration either over-estimated porosity while under-estimating dispersivity or under-estimated porosity while over-estimating dispersivity. **Figure 6.3** shows that the UnitPulse had the lowest WSSE for all scenarios with dispersivity; in most cases the UnitPulse WSSE was substantially lower. This suggests that the calibration routine used the combination of porosity and dispersivity to improve the objective function but was largely unable to accurately represent the system.

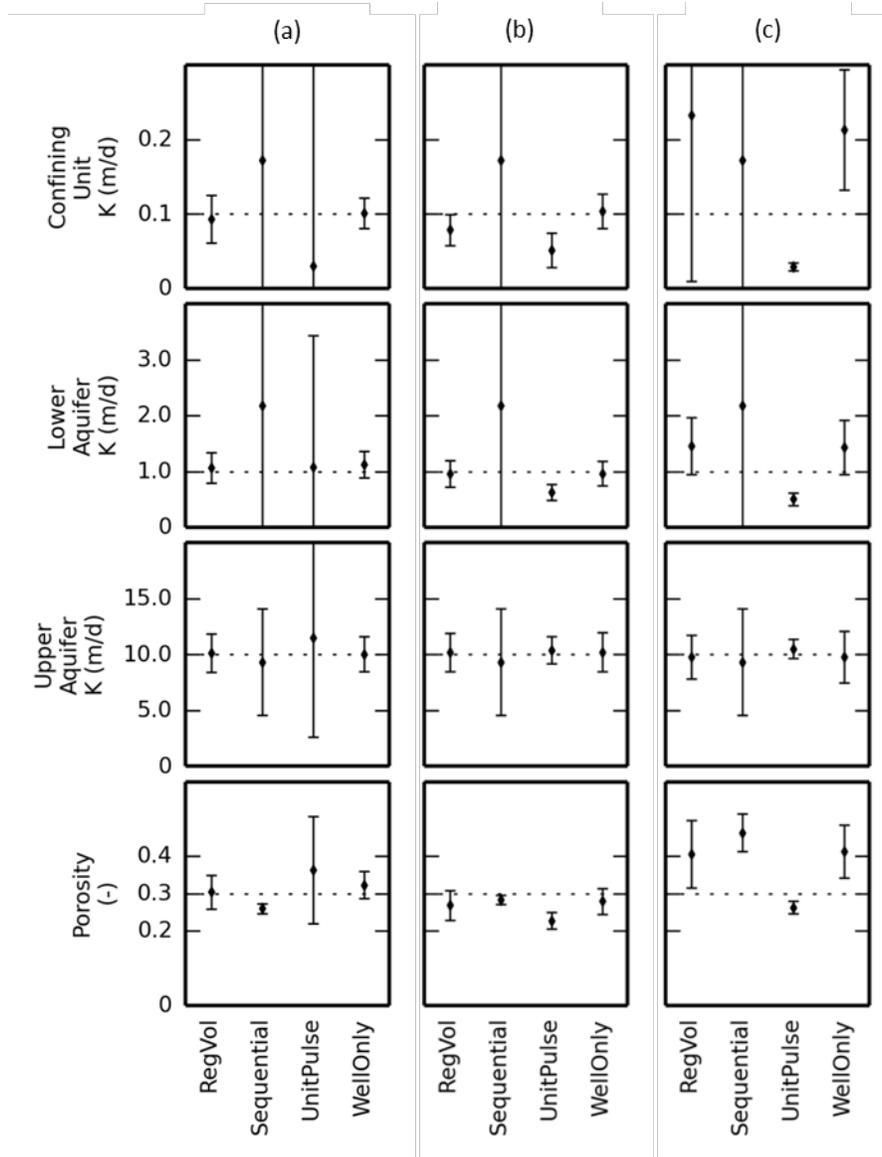


Figure 6.4 Parameter estimates for the scenarios (a) Stratified_0, (b) Stratified_10, and (c) Stratified_25. Error bars show 95% confidence intervals generated by UCODE. Dashed line indicates true parameter value.

While a full examination of the numerical challenges to transport modeling, and the feedbacks between those challenges and the inverse modeling algorithm, is beyond the scope of this paper, we may consider the possible effects of numerical dispersion on the results of calibrations using the UnitPulse method. Avoiding or accounting for the presence

of numerical dispersion is a well-known challenge to advective-dispersive transport simulation (Zheng and Bennett, 2002; Noorishad et al., 1993; Fletcher, 1991). When using the FD solver while calibrating against age information, the effects of dispersivity on age at a point may be a function of both numerical dispersion as well as the calibrated value of the dispersivity input parameter. It is difficult to quantify the cumulative effect of numerical dispersion on a system (Zheng and Bennett, 2002). However, in order to estimate this effect for our steady state 2D system, we may calculate the one-dimensional (1D) dispersivity that is a result of numerical dispersion in both the horizontal and vertical directions at each model node. For the solution method applied in our study, which used an implicit approximation of the ADM with upstream weighting, 1D numerical dispersivity α_{num} is calculated as follows:

$$\alpha_{num} = \frac{v\Delta x}{2} \left(1 + \frac{v\Delta t}{\Delta x} \right) \quad (6.1)$$

where v is the velocity in the direction of flow; Δx is the grid dimension in direction of flow; and Δt is the time step (Zheng and Bennett, 2002). As noted above, for our model Δx was equal to 25 m for the horizontal direction and 5 m for the vertical direction. For all scenarios, Δt was equal to 10 days. For the horizontal and vertical velocities associated with the ‘true’ parameter sets, the average numerical dispersivity at each node for each scenario was less than 0.5 m. For the Stratified scenarios, α_{num} at a point was as high as 4 m; for the Binary and Lense scenarios, α_{num} at a point was as high as 10 m. Noting that (i) these 1D approximations do not account for the cumulative effect of dispersivity on the system and (ii) the velocities and therefore the numerical dispersivities change for each

iteration of each calibration, we may nonetheless assume that numerical dispersion contributed to inaccurate parameter estimates for the zero dispersivity cases. This is because, for each zero dispersivity case, numerical dispersion continued to have dispersive effects on age in spite of the fact that UCODE had set the dispersivity input parameter to 0.

The influence of numerical dispersion may explain, for example, the UnitPulse porosity estimate for the Stratified_0 scenario. Analysis of the regression residuals and the DFBETAS statistics (not shown) for the optimized UnitPulse parameter set indicates that the age data for the confining layer was the greatest influence on the estimated porosity. In the presence of numerical dispersion, the simulated age of the confining layer was younger than it would be without dispersion because of the mixing of younger, upper aquifer water; this in turn confounded the identification of porosity, which was estimated too high as a compensatory measure meant to reduce advective velocities and thus increase the age.

In contrast, for the Stratified_10 scenario, the porosity estimate was most influenced by a combination of lower and upper aquifer observations (Obs04 and Obs10). More research is needed to assess any numerical reasons that drive the under-estimation of porosity. It may be noted, however, that for the Upper Chester model, with hydrogeology closely resembling the Stratified cases here, the use of ADM methods to simulate age resulted in unreasonably low calibrated porosities (results not shown). The low porosity estimates derived from ADM inverse modeling at the field site appear to be artifacts of the simulation method rather than a true representation of the system. In contrast, the use of the WellOnly method in the field case resulted in porosity estimates that are consistent with the analysis of sediment cores and porosity estimates based on further analysis of an extensive set of

environmental tracer data (including transport simulations of multiple environmental tracers).

We may conclude from the UnitPulse cases that even in situations where the ADM method may accurately estimate the dispersivity (as in the Binary_0, Lenses_0, and Stratified_0 scenarios), the estimates of the other parameters are not necessarily accurate. However, keeping in mind that we do not necessarily know beforehand that we are applying an ADM method to a low dispersivity system, for the zero dispersivity scenarios the ADM dispersivity estimate served as a useful signal that kinematic methods should instead be used. In contrast, though, note that the UnitPulse method wrongly estimated zero dispersivity for the Lenses_10 scenario. Interestingly, for the Lenses_10 case both the kinematic methods and the UnitPulse estimated nearly identical parameter sets (not shown). This may suggest that, if an ADM method estimates a very low dispersivity system, but a kinematic method estimates a very similar parameter set, then both methods are wrong, and a sequential method is more appropriate (see **Table 6.4** and **Table 6.5** and discussion above). More research is required to investigate this further.

6.2.1.c – Potential benefits of the RegVol method

For the Lenses_0 scenario, the RegVol method resulted in a more accurate model than the WellOnly method as measured by the porosity estimate (**Table 6.4**) and the baseflow age prediction (**Table 6.5**). This is a curious result because (i) the WellOnly method was itself used to generate the ‘true’ values and (ii) the dimensions of the registration rectangle calibrated by the RegVol were effectively zero. This latter aspect means that at the optimal parameter set, the RegVol method was only calculating the simulated age at the well screen

– i.e., acting as the WellOnly method – , yet as a method it was more effective than the WellOnly method. **Figure 6.3d** shows that the Lenses_0 WellOnly method quickly reached a local minimum on the WSSE response surface and made little change for the remainder of the 50 iterations. The RegVol method, however, continued to explore the parameter space. These contrasting behaviors are perhaps due to the kinematic volatility that Varni and Carrera (1998) identified for heterogeneous conductivity fields. As noted above, for a given parameter set, small variations in location may have large implications for a kinematic age. Similarly, for a given location, small variations in parameter values may have large implications for kinematic age. This sensitivity of the age to the parameter set is in part a function of grid dimensions and the characteristics of the simulated advective velocity field (e.g., the number of particles that are used and the interpolation of heads and velocities that is required to place the particles on individual advective flow paths). For a gradient search optimization scheme such as UCODE, which uses a sensitivity analysis at each iteration to determine the direction of parameter change, the sensitivity of simulated kinematic ages to the changes in parameter values may limit the precision with which the calibration can approach the true values. It may be possible to address this using UCODE controls such as the amount each parameter is perturbed during the sensitivity analysis. The RegVol method, in contrast, by extending the number of advective pathlines contributing to each observation location, appears to dampen this kinematic volatility and allow the optimization algorithm to make a more refined search of the parameter space.

6.2.2 – Results and Discussion for Case Study

Figure 6.5 and Figure 6.6 compare the results of (i) flow and transport parameter estimation using aggregated interpreted ages as calibration data with (ii) parameter estimation using tracer concentrations (see Chapter 3 of this dissertation).

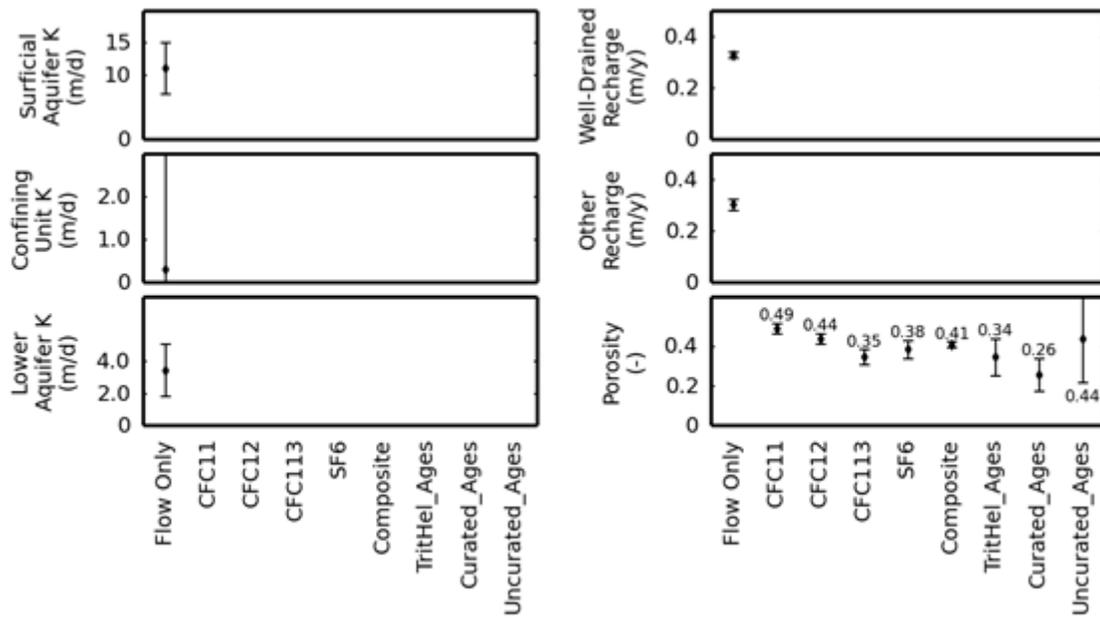


Figure 6.5 Sequentially-calibrated parameter estimates for Upper Chester. Error bars show 95% confidence intervals calculated by UCODE. The porosity values that are displayed are the calibrated values (rather than the upper confidence interval).

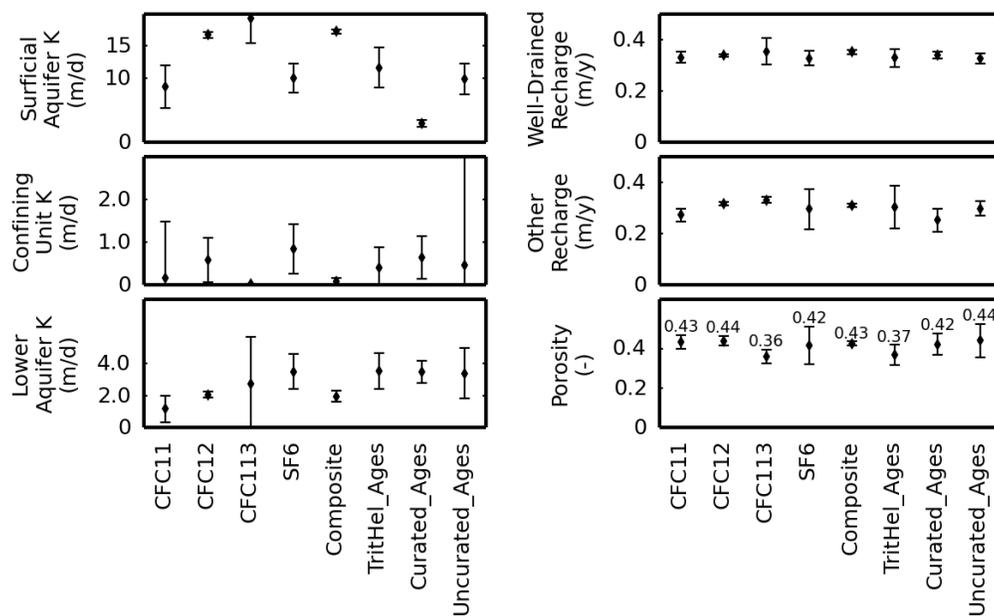


Figure 6.6 Simultaneously-calibrated parameter estimates for Upper Chester. Error bars show 95% confidence intervals calculated by UCODE. The porosity values that are displayed are the calibrated values (rather than the upper confidence interval).

With the exception of the sequential Curated Ages method, parameter estimates obtained using the aggregated steady state apparent ages were very similar to estimates obtained using the full tracer datasets (**Figures 6.5 and 6.6**). It is notable that for the simultaneous cases, the porosity estimates from both steady state age methods are very close to the porosity estimate from the Composite tracer method. The steady state age methods result in larger confidence intervals than the Composite method because they have fewer observations. The general similarity of porosity estimates suggests that the use of apparent ages as calibration targets does not appreciably impact the resulting parameter estimates. However, this interchangeability of tracer and age calibration targets is likely very sensitive to the validity of the transport model as well as the travel time distribution that is assumed in the interpretation of ages from tracer data. That is, for the Upper Chester, the advection-

only transport model seems to describe the system well, such that advection-only assumptions could be invoked both when simulating tracer concentrations and when inferring age from individual tracer concentrations. However, in systems where the tracer movement requires a more sophisticated transport model (e.g., ADM or dual domain), it is unlikely that the lumped parameter model used to generate age targets and the transport model used to describe tracer movement would be as well-coordinated as they are for our study site.

It is apparent from **Figure 6.5** that the aggregation method (i.e., Curated or Uncurated) impacted the sequential steady state methods; the optimized porosity was typical of the other methods when calibrated against uncurated age targets but very low (0.26) when calibrated against curated age targets. The reduced porosity estimate derived from the curated ages is attributable in part to the greater weight assigned to observations at lower depths (**Table 6.3**); in order to satisfy the younger curated ages reported at sites below the confining layer, but with the confining layer conductivity adjusted, the sequential method reduced the effective porosity estimate. It should also here be noted here that the conversion of a very low tracer concentration to an age changes the potential impact of that site on the objective function during the parameter estimation. Consider for example, a site at which a very low CFC-12 concentration suggests an apparent recharge date of 1950, which is at the threshold of groundwater age dateable with the CFC-12 method (**Figure 3.1**). If during a calibration iteration the porosity is overestimated, such that the simulated recharge date at that site is 1930, then the simulated CFC-12 concentration will be zero and the concentration residual for that site will simply be equal to the sampled concentration,

which was small. If the simulated recharge date for that site is 1910, the residual does not change because the simulated CFC-12 concentration is still equal to zero. In other words, for simulation of tracer concentrations, the calibration is insensitive to over-estimation of age once the simulated velocity field results in a groundwater age at a site that is older than the beginning of the tracer time series. However, in cases where age, rather than tracer concentration, is simulated, the age residual will continue to grow as the simulated velocity field gets slower and slower. This phenomenon may of course be impacted by the relative weights placed on age and head data, since the velocity field is a function of conductivity and recharge as well as porosity.

In addition to the influence of older, deeper wells on the porosity estimate from the sequential Curated Age method, the reduced influence of CFC information in the curated dataset (**Table 6.3**) may also be a factor. Though, in contrast to the sequential method, the simultaneous inclusion of curated age information resulted in: (i) a much higher porosity estimate, which increases the age at lower depths; (ii) a higher confining unit conductivity estimate, which reduces the age at lower depths; and (iii) a lower upland recharge estimate, which increases the age at lower depths (**Figure 6.6**). The aggregation method had less impact on the simultaneous steady state method (**Figure 6.6**). Relative to the simultaneous Uncurated Age method, the simultaneous Curated Age method resulted in a lower porosity estimate, which reduces simulated age, and a lower Surficial Aquifer conductivity estimate, which increases simulated age. It is notable, however, that the simultaneous inclusion of uncurated age information resulted in essentially the same flow model as when sequential calibration was used; the sequential flow parameter values were used as the initial estimates of the flow parameter values for the simultaneous calibration.

6.3 CONCLUSIONS

The automated calibrations of several synthetic 2D aquifer configurations were used to investigate the impact of model simulation of groundwater age (i.e., whether age was simulated as a kinematic or ADM process) on the estimated parameter sets and simulated system baseflow age. As true system dispersivity increased, the capacity of kinematic simulations of age to translate the available system information into accurate parameter estimates decreased. With the exception of the Stratified_10 scenario, this resulted in estimates of system porosity and mean baseflow age that were too high. For higher-dispersivity systems, the UnitPulse ADM method improved on the performance of kinematic methods. However, in general the ADM method did not accurately estimate the system porosity or baseflow age. The ADM calibration methods correctly estimated the dispersivity parameter for zero-dispersivity systems. However, other parameters were not correctly estimated, possibly as a result of the numerical dispersion associated with the finite difference solution to the ADM. For some scenarios, the use of a registration volume of kinematic ages appeared to assist the calibration algorithm in finding a more accurate solution. We also found for the Upper Chester flow and transport model that parameter estimates obtained using the aggregated steady state apparent ages as calibration data were very similar to estimates obtained using the full tracer datasets.

6.4 REFERENCES

- Anderman, E. R., Hill, M. C., & Poeter, E. P. (1996). Two-dimensional advective transport in ground-water flow parameter estimation. *Groundwater*, 34(6), 1001-1009.
- Busenberg, E., & Plummer, L. N. (2000). Dating young groundwater with sulfur hexafluoride: Natural and anthropogenic sources of sulfur hexafluoride. *Water Resources Research*, 36(10), 3011-3030.
- Cook, P. G., & Böhlke, J. K. (2000). Determining timescales for groundwater flow and solute transport. In *Environmental tracers in subsurface hydrology* (pp. 1-30). Springer.
- Feehley, C. E., Zheng, C., & Molz, F. J. (2000). A dual-domain mass transfer approach for modeling solute transport in heterogeneous aquifers: Application to the Macrodispersion Experiment (MADE) site. *Water Resources Research*, 36(9), 2501-2515.
- Fletcher, C. A. J. (1991). *Computational techniques for fluid dynamics*. Springer.
- Freyberg, D. L. (1986). A natural gradient experiment on solute transport in a sand aquifer: 2. Spatial moments and the advection and dispersion of nonreactive tracers. *Water Resources Research*, 22(13), 2031-2046.
- Garabedian, S. P., LeBlanc, D. R., Gelhar, L. W., & Celia, M. A. (1991). Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 2. Analysis of spatial moments for a nonreactive tracer. *Water Resources Research*, 27(5), 911-924.
- Ginn, T. R., Haeri, H., Massoudieh, A., & Foglia, L. (2009). Notes on groundwater age in forward and inverse modeling. *Transport in Porous Media*, 79(1), 117-134.
- Goode, D. J. (1996). Direct simulation of groundwater age. *Water Resources Research*, 32(2), 289-296.
- Gusyev, M. A., Abrams, D., Toews, M. W., Morgenstern, U., & Stewart, M. K. (2014). A comparison of particle-tracking and solute transport methods for simulation of tritium concentrations and groundwater transit times in river water. *Hydrology and Earth System Sciences*, 18(8), 3109-3119.
- Harbaugh, A. W. (2005). *MODFLOW-2005, the U.S. Geological Survey modular ground-water model—The ground-water flow process*, U.S. Geological Survey, Technical Methods, Book 6, Chap. A16.
- Harvey, C., & Gorelick, S. M. (2000). Rate-limited mass transfer or macrodispersion: Which dominates plume evolution at the macrodispersion experiment (MADE) site? *Water Resources Research*, 36(3), 637-650.
- Hill, M. C., & Tiedeman, C. R. (2006). *Effective groundwater model calibration: with analysis of data, sensitivities, predictions, and uncertainty*. John Wiley & Sons.

- Hunt, R. J., Feinstein, D. T., Pint, C. D., & Anderson, M. P. (2006). The importance of diverse data types to calibrate a watershed model of the Trout Lake Basin, Northern Wisconsin, USA. *Journal of Hydrology*, 321(1), 286-296.
- Jury, W. A., & Roth, K. (1990). *Transfer functions and solute movement through soil: theory and applications*. Birkhäuser Verlag AG.
- Kazemi, G. A., Lehr, J. H., & Perrochet, P. (2006). *Groundwater age*. John Wiley & Sons.
- Małoszewski, P., & Zuber, A. (1982). Determining the turnover time of groundwater systems with the aid of environmental tracers: 1. Models and their applicability. *Journal of Hydrology*, 57(3), 207-231.
- Massoudieh, A., & Ginn, T. R. (2011). The theoretical relation between unstable solutes and groundwater age. *Water Resources Research*, 47(10).
- McCallum, J. L., Cook, P. G., Simmons, C. T., & Werner, A. D. (2014a). Bias of apparent tracer ages in heterogeneous environments. *Groundwater*, 52(2), 239-250.
- McCallum, J. L., Cook, P. G., & Simmons, C. T. (2014b). Limitations of the use of environmental tracers to infer groundwater age. *Groundwater*, 53(S1), 56-70.
- Noorishad, J., Tsang, C. F., Perrochet, P., & Musy, A. (1992). A perspective on the numerical solution of convection-dominated transport problems: A price to pay for the easy way out. *Water Resources Research*, 28(2), 551-561.
- Neupauer, R. M., & Wilson, J. L. (2004). Numerical implementation of a backward probabilistic model of ground water contamination. *Groundwater*, 42(2), 175-189.
- Pollock, D. (2012). *User guide for MODPATH version 6: a particle tracking model for MODFLOW*. US Geological Survey Techniques and Methods 6-A41.
- Poeter, E. P., Hill, M. C., Banta, E. R., Mehl, S. W., & Christensen, S. (2005). *UCODE- 2005 and six other computer codes for universal sensitivity analysis, inverse modeling, and uncertainty evaluation*. USGS Technical Methods, 6-A11.
- Rehfeldt, K. R., Boggs, J. M., & Gelhar, L. W. (1992). Field study of dispersion in a heterogeneous aquifer: 3. Geostatistical analysis of hydraulic conductivity. *Water Resources Research*, 28(12), 3309-3324.
- Reilly, T. E., Plummer, L. N., Phillips, P. J., & Busenberg, E. (1994). The use of simulation and multiple environmental tracers to quantify groundwater flow in a shallow aquifer. *Water Resources Research*, 30(2), 421-433.

- Sanford, W.E. (2011). Calibration of models using groundwater age. *Hydrogeology Journal*, 19(1), 13-16.
- Sanford, W.E., Pope, J.P., Selnick, D.L., & Stumvoll, R.F. (2012). *Simulation of groundwater flow in the shallow aquifer system of the Delmarva Peninsula, Maryland and Delaware*. US Department of the Interior, US Geological Survey.
- Sudicky, E. A. (1986). A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process. *Water Resources Research*, 22(13), 2069-2082.
- Szabo, Z., Rice, D. E., Plummer, L. N., Busenberg, E., Drenkard, S., & Schlosser, P. (1996). Age dating of shallow groundwater with chlorofluorocarbons, tritium/helium: 3, and flow path analysis, Southern New Jersey Coastal Plain. *Water Resources Research*, 32(4), 1023-1038.
- Troldborg, L., Jensen, K. H., Engesgaard, P., Refsgaard, J. C., & Hinsby, K. (2008). Using environmental tracers in modeling flow in a complex shallow aquifer system. *Journal of Hydrologic Engineering*, 13(11), 1037-1048.
- Varni, M., & Carrera, J. (1998). Simulation of groundwater age distributions. *Water Resources Research*, 34(12), 3271-3281.
- Weissmann, G. S., Zhang, Y., LaBolle, E. M., & Fogg, G. E. (2002). Dispersion of groundwater age in an alluvial aquifer system. *Water Resources Research*, 38(10), 16-1.
- Zheng, C., & Wang, P. P. (1999). *MT3DMS: a modular three-dimensional multispecies transport model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems; documentation and user's guide*. Department of Geological Sciences, University of Alabama, Tuscaloosa, Alabama.
- Zheng, C., & Bennett, G. D. (2002). *Applied contaminant transport modeling* (2nd ed.). New York: Wiley-Interscience.
- Zheng, C. (2009). *MT3DMS v5.3 Supplemental User's Guide*. Department of Geological Sciences, University of Alabama, Tuscaloosa, Alabama.
- Zheng, C., Bianchi, M., & Gorelick, S. M. (2011). Lessons learned from 25 years of research at the MADE site. *Groundwater*, 49(5), 649-662.

Chapter 7: Conclusions and Future Work

In this dissertation we examined a series of questions related to understanding and managing nitrogen in agricultural systems, including fundamental methodological questions associated with the calibration of groundwater simulation tools as well as the application of those tools in the description of nitrate transport and removal processes. In Chapter 3 we used CFC, SF₆, and ³H/³He datasets individually and in combination to calibrate the groundwater flow and transport model for the Upper Chester targeted watershed. Our study demonstrated that while tracer data can provide necessary supplemental information for the calibration of flow and transport models, the use of data from a single tracer or from a small tracer set may be insufficient to fully interpret the information content of the tracers. While the use of multiple weighting schemes with datasets of individual tracers may be important for suggesting the range of possible models, the combined use of multiple tracers is less sensitive to the weighting scheme and results in more confident parameter estimates. Further, we showed that the use of tracer information in parameter estimation can help reduce uncertainty in the characterization of important hydrogeological features such as the semi-confining unit in the Upper Chester.

In Chapter 4 we used the travel time distributions generated by the calibrated flow model in order to relate land surface loadings to stream responses and resolve the key components of the catchment nitrogen budget. We examined adjacent agricultural subcatchments with similar land use histories but disparate nitrate export signatures in order to quantify the removal fractions of various removal mechanisms. We showed that in

spite of spatial and temporal uncertainty in loading, multiple calibration scenarios agreed that in-stream nitrate removal efficiencies vary significantly between the two streams. These site specific findings of disparate removal mechanisms highlighted more general, and potentially important, questions about the complexity of predicting and managing agricultural nitrate export at a regional scale, including the challenges of integrating agricultural catchments with highly heterogeneous nitrogen processing and export characteristics into nutrient trading frameworks.

In Chapter 5 we developed a method for simulating the transient delivery of base-flow age from subsurface to receiving stream as a function of seasonal changes in hydrology and aquifer storage, and we applied the method to a variety of synthetic two-dimensional (2D) aquifers as well as to the Upper Chester site. We found that the timing of maximum base-flow age relative to the timing of minimum base-flow discharge varied with both the hydraulic conductivity field and the annually averaged recharge, which determines the system mean age. The two assumptions of (i) an aquifer in which ages are vertically well-mixed and (ii) an aquifer in which ages are strongly stratified provide two end-members for estimation of how the base-flow age might respond to seasonal changes in recharge and base-flow, and the simulations in this study found that the change that occurs in real systems is somewhere in between. For the cases that we investigated, apparent ages inferred from SF₆ measurements while assuming piston-flow transport assumptions for the SF₆ were biased young, with biases especially pronounced with layered hydrogeology in which discharge consists of shallow surficial flow mixed with a contrasting regime of much older water. For one of the subcatchments in our Maryland study site we found that seasonal changes in recharge may only result in changes in base-flow age of 3 to 4 years,

but that SF₆ apparent ages based on piston-flow assumptions of tracer transport may underestimate the mean base-flow age by 60% and more closely resemble the median system age. The conclusions from Chapter 5 and the further development of the modeling methodology developed in that chapter will be important for interpreting water quality changes in streams where base-flow delivery of solute is an important component of stream solute concentrations and where the system is subject to seasonal changes in the hydrologic regime.

Finally, in Chapter 6 we considered the impact of assumptions about age transport on the automated calibration of groundwater flow and transport models. Specifically, we used the automated calibration of several synthetic aquifers to investigate the impact on the resulting calibrated model of the assumption that no dispersion is present under a variety of heterogeneity and dispersivity scenarios. We showed that as true system dispersivity increases, the capacity of kinematic simulations of age to translate the available system information into accurate parameter estimates decreases, and that in general the automated calibration routines compensated for the un-modeled dispersivity with elevated porosity estimates. This work provides a cautionary tale that complements the work of Chapter 3 and the insights from that chapter into (i) the uncertainty that may accompany the integration of groundwater age information into model calibration as well as (ii) the benefits of the information when properly employed.

In sum, the studies in this dissertation collectively illuminate the general task of groundwater modeling and the particular task of using groundwater modeling to describe nitrogen transport in agricultural systems. Important future site-specific work includes the

integration of improved land use information and more spatially distributed watershed monitoring into the description of nitrate transport in the Upper Chester, with the ultimate aim of discerning the impact on water quality of management improvements that have begun in the last 5-10 years. Given the lag times between land surface action and stream response, several years of water quality monitoring will be required to fully leverage the higher resolution land use data now being collected. However, more immediate further investigations should consider whether the emerging land use dataset provides any information upon which reconstructions of historical loading may be conditioned, thus reducing the loading uncertainty described in Chapter 4.

Appendix A: Scripting tools developed for this work

This dissertation work included the development of several tools that were written in the Python language for purposes of:

- i. generating the model framework from geospatial information (e.g., land surface elevations and hydrogeological unit locations);
- ii. developing input files for the MODFLOW, MT3DMS, and MODPATH numerical simulation engines;
- iii. integrating data from sources such as the US Geological Survey National Water Information System into simulation calibration routines;
- iv. post-processing (e.g., summarizing and visualizing) model output; and
- v. coordinating output from and updating input information for simulation engines during model calibration.

Documentation for these Python tools is currently being updated in preparation for making the tools publicly available through GitHub or a similar distribution method. **Figure A1** illustrates the use of some of these scripting tools in the nitrate modeling workflow used in Chapter 4 of this dissertation.

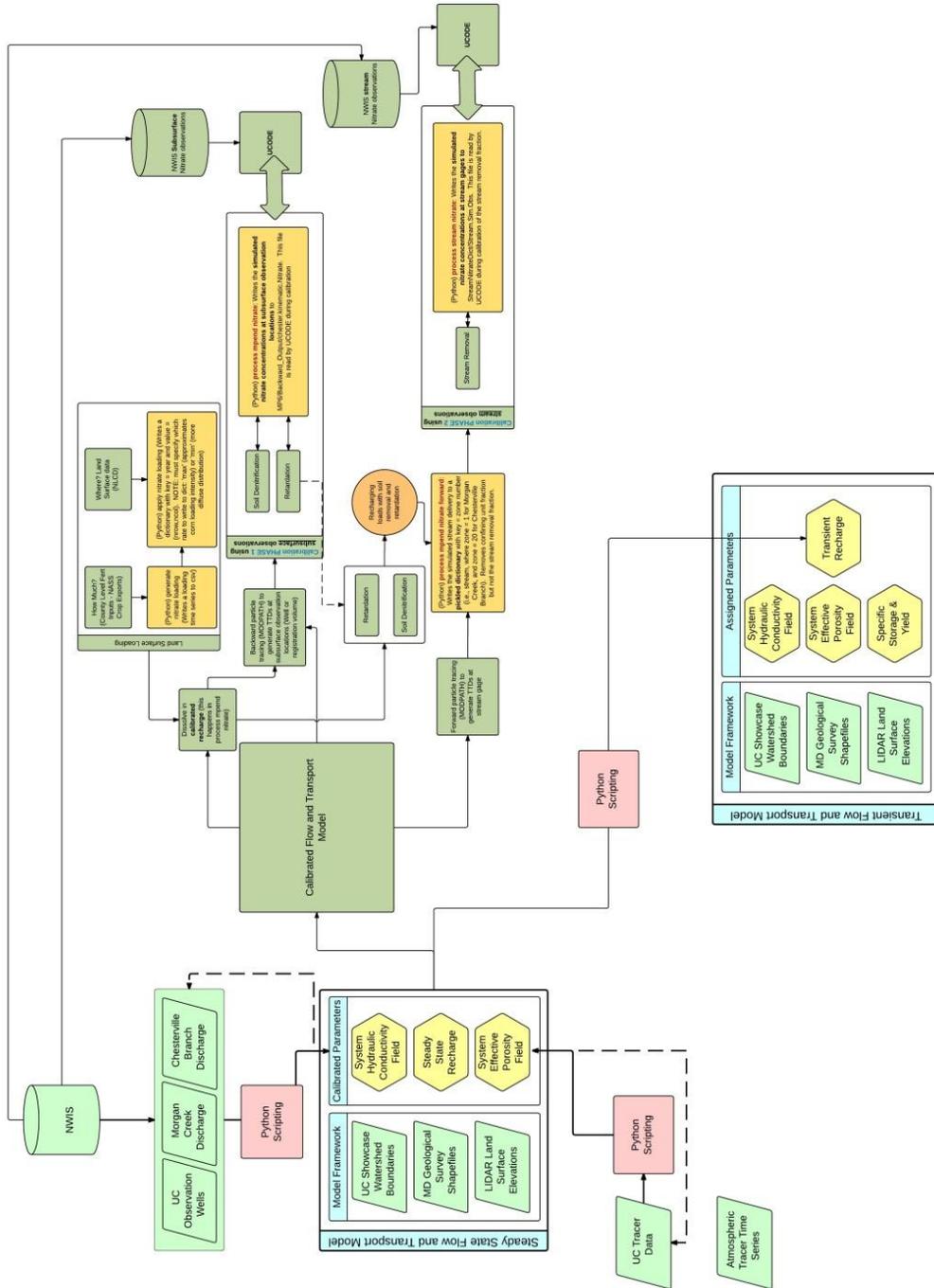


Figure A1. Nitrate transport modeling workflow.

Appendix B: Subsurface environmental tracer observations for the Upper Chester study site

CFC and SF6 data from Busenberg and Plummer (2000). Tritium data from Ekwurzel et al. (1994). (See Section 3.5 for full references.)

Well Identifier	Observation Date	CFC-11 (pg/kg)	CFC-11 std (pg/kg)	CFC-11 CFC-12 (pg/kg)	CFC-12 std (pg/kg)	CFC-113 (pg/kg)	CFC-113 std (pg/kg)	SF ₆ (fmol/L)	SF ₆ std (fmol/L)	³ H (Tritium units)	³ H + ³ He (Tritium units)
KEBe50	11/19/1990	--	--	279	15	--	--	--	--	--	--
KEBe50	11/14/1991	--	--	336	2	80.5	2.6	--	--	--	--
KEBe50	3/31/1992	--	--	343	8	62.3	8.5	--	--	--	--
KEBe50	7/24/1998	--	--	313	0	83.3	5.9	1.32	--	--	--
KEBe50	12/10/1998	--	--	--	--	--	--	0.98	0.01	--	--
KEBe52	11/5/1990	751	56	211	16	--	--	--	--	15.3	--
KEBe52	11/5/1991	607	11	250	4	36.2	5	--	--	16.6	24.7
KEBe52	3/31/1992	671	6	265	5	33.4	2.2	--	--	--	--
KEBe52	7/21/1993	630	11	276	17	--	--	--	--	--	--
KEBe52	7/29/1997	647	7	307	0	69.5	2.9	1.06	0.01	--	--
KEBe53	4/2/1991	778	15	266	6	--	--	--	--	16	--
KEBe53	7/21/1993	658	2	299	9	80	0.9	--	--	--	--
KEBe53	7/29/1997	655	6	312	6	84.3	2.8	1.64	0.25	--	--
KEBe59	11/19/1990	--	--	265	15	--	--	--	--	14.8	--
KEBe59	12/10/1998	--	--	--	--	--	--	1.43	0.01	--	--
KEBe61	11/5/1990	211	1	91	0	9.3	1.7	--	--	50	--
KEBe61	4/2/1991	344	100	136	49	--	--	--	--	60.3	147.2
KEBe61	11/6/1991	212	12	93	2	6.1	3	--	--	63.4	153.8
KEBe61	7/21/1993	251	3	121	0	11.9	0	--	--	--	--
KEBe61	7/29/1997	487	1	192	2	23.3	0.3	0.47	0.15	--	--
KEBe62	11/6/1990	--	--	310	34	--	--	--	--	13.7	--
KEBe62	11/6/1991	704	4	302	0	66.5	11	--	--	13.7	16.2
KEBe62	3/31/1992	747	34	297	7	51.2	3.5	--	--	--	--
KEBe62	7/21/1993	677	4	304	1	77.5	2.6	--	--	--	--
KEBe62	4/12/1995	666	9	366	10	97.4	2.3	--	--	--	--
KEBe62	7/28/1997	640	1	303	2	81.2	0	1.43	0.01	--	--
KEBe63	11/6/1990	350	7	102	6	--	--	--	--	38.5	--
KEBe63	7/29/1997	372	2	136	1	13.7	0.3	0.21	0.03	--	--
KEBe63	12/11/1998	--	--	--	--	--	--	0.2	0	--	--
KEBe64	11/6/1990	793	5	231	17	--	--	--	--	16.5	--
KEBe64	7/29/1997	604	11	259	4	53.1	0.1	0.69	0.09	--	--

Well Identifier	Observation Date	CFC-11 (pg/kg)	CFC-11 std (pg/kg)	CFC-12 (pg/kg)	CFC-12 std (pg/kg)	CFC-113 (pg/kg)	CFC-113 std (pg/kg)	SF ₆ (fmol/L)	SF ₆ std (fmol/L)	³ H (Tritium units)	³ H + ³ He (Tritium units)
KEBe64	12/11/1998	--	--	--	--	--	--	1.02	0.03	--	--
KEBe158	11/21/1990	788	13	239	20	--	--	--	--	--	--
KEBe158	7/21/1998	624	19	299	26	80.6	7.7	--	--	--	--
KEBe159	11/7/1990	44	4	1	1	--	--	--	--	0	--
KEBe159	11/7/1991	13	3	4.1	5	8.1	2.9	--	--	--	--
KEBe159	7/20/1993	0	0	0	0	0	0	--	--	--	--
KEBe159	7/28/1997	0	0	0	0	0	0	0.03	0.01	--	--
KEBe159	12/11/1998	--	--	--	--	--	--	0.05	0	--	--
KEBe160	11/19/1990	140	1	49	2	--	--	--	--	20	--
KEBe160	11/7/1991	144	3	77	7	8	1.5	--	--	40.8	123.1
KEBe160	3/31/1992	196	16	76	3	6.3	5.4	--	--	--	--
KEBe160	7/20/1993	121	3	75	0	4.1	0.3	--	--	--	--
KEBe160	7/28/1997	160	1	112	0	5.5	0.4	0.09	0.02	--	--
KEBe161	11/7/1990	601	7	186	7	--	--	--	--	19.2	--
KEBe161	4/2/1991	517	5	206	3	--	--	--	--	16.8	--
KEBe161	11/7/1991	532	18	223	5	29.4	4.9	--	--	16.8	37.4
KEBe161	11/7/1991	601	0	216	0	--	--	--	--	--	--
KEBe161	7/21/1993	552	9	251	7	41.8	4.4	--	--	--	--
KEBe161	7/28/1997	560	8	244	8	41.9	0.7	0.62	0.05	--	--
KEBe162	11/6/1990	118	4	9	2	--	--	--	--	0.9	--
KEBe162	11/6/1991	52	8	16	2	--	--	--	--	2.2	21.3
KEBe162	7/21/1993	20	1	15	0	0.6	0	--	--	--	--
KEBe162	4/12/1995	22	1	23	6	0	0	--	--	--	--
KEBe162	4/12/1995	167	84	95	42	19.9	9.6	--	--	--	--
KEBe162	7/28/1997	49	2	37	2	0	0	0.14	0.01	--	--
KEBe162	12/11/1998	--	--	--	--	--	--	0.15	0	--	--
KEBe163	11/6/1990	778	111	231	3	--	--	--	--	16.5	--
KEBe163	11/6/1991	567	4	239	3	35.5	7.5	--	--	16.7	25.9
KEBe163	7/21/1993	562	4	244	1	41.6	0.2	--	--	--	--
KEBe163	4/12/1995	600	3	291	9	64.4	0.9	--	--	--	--
KEBe163	7/28/1997	634	1	284	2	66.4	0	0.85	0.07	--	--
KEBe164	11/19/1990	280	30	64	15	--	--	--	--	31.5	--
KEBe164	12/10/1998	--	--	--	--	--	--	0.22	0.01	--	--

Well Identifier	Observation Date	CFC-11 (pg/kg)	CFC-11 std (pg/kg)	CFC-12 (pg/kg)	CFC-12 std (pg/kg)	CFC-113 (pg/kg)	CFC-113 std (pg/kg)	SF ₆ (fmol/L)	SF ₆ std (fmol/L)	³ H (Tritium units)	³ H + ³ He (Tritium units)
KEBe165	11/7/1990	344	8	98	5	--	--	--	--	--	--
KEBe165	4/2/1991	332	13	126	2	--	--	--	--	--	117.7
KEBe165	7/29/1997	269	3	114	2	7.9	1	0.16	0.04	--	--
KEBe166	11/7/1990	726	2	287	1	--	--	--	--	--	--
KEBe166	7/29/1997	665	0	324	3	84.8	0	1.06	0.04	--	--
KEBe167	7/29/1997	769	2	368	6.4	91.9	1.4	1.73	0.01	--	--
KEBe169	4/2/1991	--	--	--	--	--	--	--	--	12.2	48.4
KEBe170	4/2/1991	--	--	--	--	--	--	--	--	20	31.8
KEBe189	7/21/1998	1	1	8	1.4	0	0	0.05	0	--	--
KEBe192	7/21/1998	--	--	342	4.7	--	--	1.13	0.02	--	--
KEBe194	7/20/1998	597	1	279	2.2	65	2.3	0.92	--	--	--
KEBe195	7/20/1998	404	2	155	2.9	24.4	3.8	0.32	--	--	--
KEBe199	7/22/1998	401	15	160	12.1	16.1	0.4	0.21	--	--	--
KEBe199	12/11/1998	--	--	--	--	--	--	0.2	0.01	--	--
KEBe200	7/22/1998	9	2	19	15.7	1.2	1.8	0.07	--	--	--
KEBe200	12/11/1998	--	--	--	--	--	--	0.03	0.01	--	--
KEBe206	7/22/1998	501	4	215	4	29.7	2.5	0.37	--	--	--
KEBe206	12/10/1998	--	--	--	--	--	--	0.37	--	--	--
KEBe207	7/23/1998	408	4	175	1.4	24.1	4.5	0.39	0.09	--	--
KEBe207	12/10/1998	401	5	174	1.9	26	0.2	0.25	0.02	--	--
KEBe208	7/23/1998	101	0	58	2.1	4.5	0.3	--	--	--	--
KEBe208	12/10/1998	120	1	67	0.9	4.5	0.5	0.1	0	--	--
KEBe210	7/22/1998	3	4	9	8	0	0	0.03	--	--	--
KEBe210	12/10/1998	--	--	--	--	--	--	0.02	0	--	--
KEBe211	12/11/1998	0	0	0	0	0	0	0.02	0	--	--
KEBe212	7/23/1998	633	4	319	2.6	86.3	5.7	1.12	--	--	--
KEBe216	7/24/1998	653	5	322	5.6	85.5	4.8	0.72	--	--	--

Appendix C: Subsurface nitrate observations for the Upper Chester study site

All observations downloaded from the U.S. Geological Survey National Water Information System (NWIS).

Well Identifier	Nitrate Concentration (mg N/L)	Observation Date
KEBd150	0.06	3/15/2004
KEBd151	0.06	3/15/2004
KEBd151	0.02	5/25/2004
KEBd152	0.06	3/15/2004
KEBd153	0.06	3/15/2004
KEBd153	0.06	5/24/2004
KEBd153	0.29	9/15/2004
KEBd154	4.50	3/15/2004
KEBd155	16.10	3/11/2004
KEBd156	27.00	3/11/2004
KEBd156	25.90	5/26/2004
KEBd156	21.20	9/27/2004
KEBd157	1.65	3/10/2004
KEBd158	17.70	3/10/2004
KEBd158	17.40	5/25/2004
KEBd158	17.50	9/27/2004
KEBd159	4.50	3/10/2004
KEBd160	10.20	3/10/2004
KEBd160	9.71	5/25/2004
KEBd160	9.82	7/20/2004
KEBd160	9.83	9/23/2004
KEBd161	32.10	3/9/2004
KEBd162	31.70	3/9/2004
KEBd162	32.70	5/26/2004
KEBd162	32.10	7/22/2004
KEBd162	36.00	9/27/2004
KEBd163	0.06	3/9/2004
KEBd163	0.06	5/13/2004
KEBd163	0.04	7/22/2004
KEBd163	0.06	9/23/2004
KEBd164	0.06	3/9/2004
KEBd164	0.04	5/18/2004
KEBd164	0.06	9/23/2004
KEBd165	8.71	3/11/2004
KEBd165	9.18	5/19/2004
KEBd165	9.44	7/21/2004

Well Identifier	Nitrate Concentration (mg N/L)	Observation Date
KEBd165	9.34	9/14/2004
KEBd166	11.80	3/11/2004
KEBd166	9.55	5/18/2004
KEBd166	11.60	7/21/2004
KEBd166	12.60	9/14/2004
KEBd168	14.50	3/9/2004
KEBd168	12.70	5/18/2004
KEBd168	13.20	7/21/2004
KEBd168	13.30	9/14/2004
KEBd169	12.30	3/16/2004
KEBd170	11.30	3/16/2004
KEBd170	10.80	5/26/2004
KEBd170	11.00	7/20/2004
KEBd170	10.10	10/4/2004
KEBd171	3.62	3/10/2004
KEBd171	3.89	5/19/2004
KEBd172	15.70	3/10/2004
KEBd172	13.50	5/19/2004
KEBd172	15.00	7/15/2004
KEBd172	16.00	9/29/2004
KEBd173	14.20	3/10/2004
KEBd173	12.00	5/19/2004
KEBd173	13.10	7/15/2004
KEBd173	13.60	9/29/2004
KEBd174	6.87	3/11/2004
KEBd174	6.56	5/25/2004
KEBd174	6.37	7/14/2004
KEBd174	6.24	9/22/2004
KEBd175	12.50	3/11/2004
KEBd175	12.70	5/25/2004
KEBd175	11.60	9/22/2004
KEBd176	17.50	3/9/2004
KEBd176	18.50	5/25/2004
KEBd176	17.30	7/14/2004
KEBd176	16.90	9/22/2004
KEBd177	8.05	3/17/2004
KEBd177	10.60	5/17/2004
KEBd178	10.80	3/9/2004
KEBd178	9.46	5/18/2004
KEBd178	8.52	7/15/2004

Well Identifier	Nitrate Concentration (mg N/L)	Observation Date
KEBd178	10.50	9/23/2004
KEBd179	14.70	3/9/2004
KEBd179	13.20	5/18/2004
KEBd179	14.10	7/14/2004
KEBd179	14.10	9/23/2004
KEBd180	7.52	3/15/2004
KEBd180	7.40	5/24/2004
KEBd180	7.04	7/20/2004
KEBd180	7.23	9/15/2004
KEBe158	14.00	12/5/1991
KEBe159	1.30	11/7/1990
KEBe159	0.83	3/5/1991
KEBe159	0.81	6/4/1991
KEBe159	1.30	9/9/1992
KEBe159	1.31	9/1/1999
KEBe160	3.69	11/19/1990
KEBe160	3.80	3/5/1991
KEBe160	3.70	4/23/1991
KEBe160	3.90	5/28/1991
KEBe160	3.79	6/25/1991
KEBe160	3.60	7/22/1991
KEBe160	3.40	8/20/1991
KEBe160	3.87	9/16/1991
KEBe160	3.30	10/15/1991
KEBe160	3.60	11/13/1991
KEBe160	3.80	12/9/1991
KEBe160	3.70	1/6/1992
KEBe160	3.90	2/3/1992
KEBe160	3.80	3/2/1992
KEBe160	4.00	3/31/1992
KEBe160	6.37	9/1/1999
KEBe161	9.70	11/7/1990
KEBe161	9.80	3/5/1991
KEBe161	9.80	6/4/1991
KEBe161	9.50	10/15/1991
KEBe161	12.60	9/1/1999
KEBe162	3.30	11/6/1990
KEBe162	3.20	3/4/1991
KEBe162	3.50	6/5/1991
KEBe162	3.50	9/9/1992

Well Identifier	Nitrate Concentration (mg N/L)	Observation Date
KEBe162	4.96	9/8/1999
KEBe163	14.00	11/6/1990
KEBe163	13.00	3/4/1991
KEBe163	15.00	6/5/1991
KEBe163	15.10	9/8/1999
KEBe164	11.00	11/19/1990
KEBe164	12.10	9/2/1999
KEBe165	18.00	11/7/1990
KEBe165	22.00	10/11/1991
KEBe166	14.00	11/7/1990
KEBe167	3.60	11/7/1990
KEBe167	2.78	10/11/1991
KEBe169	10.00	3/5/1991
KEBe169	12.00	5/29/1991
KEBe170	5.00	3/5/1991
KEBe170	4.80	5/29/1991
KEBe170	10.10	9/9/1999
KEBe189	0.06	7/21/1998
KEBe189	0.06	3/12/2004
KEBe189	0.06	5/20/2004
KEBe192	11.80	7/21/1998
KEBe195	8.86	7/20/1998
KEBe195	3.69	3/15/2004
KEBe195	4.97	5/20/2004
KEBe206	11.90	7/22/1998
KEBe207	9.04	7/23/1998
KEBe207	8.80	5/24/2004
KEBe207	8.73	7/13/2004
KEBe207	14.30	9/30/2004
KEBe208	5.18	7/23/1998
KEBe210	0.07	7/22/1998
KEBe212	43.10	7/23/1998
KEBe218	12.90	3/11/2004
KEBe218	10.90	5/26/2004
KEBe218	10.30	7/12/2004
KEBe218	8.35	9/21/2004
KEBe219	5.47	3/11/2004
KEBe219	4.62	5/26/2004
KEBe219	4.70	9/21/2004
KEBe50	41.00	11/28/1988

Well Identifier	Nitrate Concentration (mg N/L)	Observation Date
KEBe50	36.00	9/20/1989
KEBe50	36.00	6/5/1990
KEBe50	34.00	11/19/1990
KEBe50	28.00	4/23/1991
KEBe50	30.00	5/28/1991
KEBe50	29.00	6/25/1991
KEBe50	29.00	7/22/1991
KEBe50	29.00	8/20/1991
KEBe50	30.80	9/16/1991
KEBe50	31.00	10/15/1991
KEBe50	29.00	11/14/1991
KEBe50	31.00	12/9/1991
KEBe50	30.00	1/6/1992
KEBe50	29.00	2/3/1992
KEBe50	29.00	3/2/1992
KEBe50	28.00	3/31/1992
KEBe52	11.00	11/30/1988
KEBe52	12.00	4/5/1989
KEBe52	12.00	6/14/1989
KEBe52	13.00	6/5/1990
KEBe52	14.00	11/5/1990
KEBe52	13.00	3/4/1991
KEBe52	14.00	4/23/1991
KEBe52	15.00	5/28/1991
KEBe52	15.00	6/25/1991
KEBe52	15.00	7/22/1991
KEBe52	15.00	8/20/1991
KEBe52	16.20	9/16/1991
KEBe52	16.00	10/15/1991
KEBe52	16.00	11/14/1991
KEBe52	16.00	12/9/1991
KEBe52	16.00	1/6/1992
KEBe52	16.00	2/3/1992
KEBe52	17.00	3/2/1992
KEBe52	17.00	3/31/1992
KEBe52	19.10	9/7/1999
KEBe53	3.08	6/5/1990
KEBe59	12.00	6/4/1990
KEBe59	13.00	11/19/1990
KEBe59	10.30	9/2/1999

Well Identifier	Nitrate Concentration (mg N/L)	Observation Date
KEBe59	6.64	7/12/2001
KEBe60	3.80	9/20/1989
KEBe60	5.60	6/7/1990
KEBe60	14.00	3/6/1992
KEBe61	9.00	7/6/1989
KEBe61	7.40	6/5/1990
KEBe61	7.60	11/5/1990
KEBe61	7.10	3/4/1991
KEBe61	8.10	6/5/1991
KEBe61	10.00	9/7/1999
KEBe62	15.00	6/6/1990
KEBe62	15.00	11/6/1990
KEBe62	14.00	3/4/1991
KEBe62	15.00	4/23/1991
KEBe62	16.00	5/28/1991
KEBe62	18.00	9/16/1991
KEBe62	17.00	10/15/1991
KEBe62	18.00	11/14/1991
KEBe62	17.00	12/9/1991
KEBe62	17.00	1/6/1992
KEBe62	16.00	2/3/1992
KEBe62	17.00	3/2/1992
KEBe62	17.00	3/31/1992
KEBe62	12.10	9/8/1999
KEBe63	4.90	9/7/1989
KEBe63	5.00	6/6/1990
KEBe63	5.00	11/6/1990
KEBe63	9.59	8/31/1999
KEBe64	11.00	9/7/1989
KEBe64	12.00	6/6/1990
KEBe64	13.00	11/6/1990
KEBe64	13.00	4/23/1991
KEBe64	13.00	5/28/1991
KEBe64	12.20	8/31/1999
KEBe65	1.30	6/13/1989
QABg62	5.37	2/26/1998
QABg62	12.30	5/11/2011