The dynamics of nutrient reduction trading: A simulation-based search for effective policy design

Abstract

This paper describes a system dynamics model developed to facilitate design of nutrient reduction trading (NRT) programs. NRT is a form of cap-and-trade policy that has been strongly promoted in recent years to address diffuse source nutrient pollution. Despite its wide appeal and the enthusiasm of its proponents, very few trades have occurred to date. We propose that impediments to learning encountered in real-world pilot studies have contributed to lack of consensus in design and implementation of NRT. The model we offer is intended as a demonstration of how the design process in this instance can benefit from system dynamics modeling.

This paper describes a system dynamics model developed for the purpose of facilitating design of nutrient reduction trading (NRT) systems. NRT is a market-based “cap and trade” approach for controlling the release of polluting nutrients into aquatic and marine ecosystems. In recent years NRT has been strongly promoted by organizations such as the US Environmental Protection Agency (USEPA) and the World Resources Institute (WRI). Across the US some 37 NRT programs and pilot projects have been launched (King and Kuch 2003). Highly publicized large-scale programs have been initiated for the Chesapeake Bay and Kalamazoo watershed system in Michigan, working in close cooperation with the WRI and USEPA. Despite these efforts, and the theoretical appeal of market-based pollution abatement strategies, very few nutrient reduction trades have occurred to date (King 2005). Review of the literature (Mehan 2006, King 2005, King and Kuch 2003, Collentine 2002, Woodward et al. 2002) and discussions with an expert in the field (personal conversations, Dr. Suzie Greenhalgh, Senior Economist with the WRI) appear to point to a lack of consensus on key aspects of NRT market design and implementation. In particular King (2005) and King and Kuch (2003) point out a lack of consensus between proponents of NRT and government decision-makers that have the actual power to implement binding discharge caps and non-compliance penalties that are necessary for NRT to fulfill its potential. We propose that real-world NRT programs and trial projects present many impediments to learning due to lengthy time lags, dynamic complexity of environmental markets, and the difficulties or impossibility of implementing stringent regulations on a trial basis, and that these impediments to learning have mitigated against consensus-building among stakeholders. In this paper we describe a generic system dynamics model offering a design for NRT, with particular focus on the cap-setting decision. It is our belief that the use of similar system dynamics models can contribute to NRT program design, and potentially to consensus building and implementation, by providing experimental platforms for testing design hypotheses and accelerating the learning process.

Nutrient reduction trading

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

Since the inception of the Clean Water Act of 1972, and the investment of hundreds of billions of dollars in treatment facilities, pollution from municipal and industrial sources has been greatly reduced in the US. Despite this progress, over 40 percent of surface waters in the US remain officially impaired (Faeth 2000). The major cause of impairment today is nutrient loading, in particular of nitrogen and phosphorous, from diffuse, or "nonpoint" sources such as agricultural operations and urban runoff. Across the US it is estimated that nonpoint sources account for over 80 percent of total nitrogen and phosphorus loadings (Carpenter et al. 1998). Nonpoint nutrient sources are extremely difficult to control. Whereas point sources of nutrients such as sewage treatment plants can be accurately monitored at outlets; nonpoint nutrient discharge can only be indirectly estimated based on biogeophysical characteristics of the land and on the prevailing land use practices. These characteristics are highly variable, and land use practice is dependent on the local decisions of thousands of farmers and other land resource managers. Presently policies for nonpoint sources are largely centered on education and subsidy programs, rather than the command and control approaches imposed on point sources. Despite the difficulties, the potential gains from controlling nonpoint nutrient pollutions are huge. A case in point is the Gulf of Mexico "dead zone," a hypoxic, or oxygen depleted, zone which appears in summer months at the mouth of the Mississippi River (Greenhalgh and Sauer 2003). This dead zone occurs due to bacterial decay of dead biomass from algae blooms caused by elevated levels of nutrients from the Mississippi River primarily associated with agricultural runoff. In recent years the dead zone has expanded beyond 20,000 square kilometers. Other well known examples of nutrient induced hypoxia are found in the Chesapeake Bay and the Baltic Sea, where nonpoint sources are also the principle emitters of nutrients. There are at least 150 nutrient induced hypoxic dead zones in global waters and the number is growing (Hawn 2006).

Nutrient reduction trading (NRT) is currently being aggressively promoted as a cost effective means of reducing aggregate nonpoint nutrient pollution. Under NRT polluting sources are able to meet compliance obligations by purchasing credits representing pollution reductions achieved by other sources. The cost effectiveness of the arrangement arises because sources with high nutrient reduction cost can purchase credits from sources with lower cost and the overall goal of reducing the aggregate nutrient load can be attained. Under the NRT programs envisioned by the US EPA (www.epa.gov/owow/watershed/trading/tradelinks.html), WRI (Faeth 2000), and other proponents, a discharge "cap" or maximum allowance is imposed on point sources of nutrient loadings. The cap is designated in mass per unit time, for example, kilograms of total nitrogen per year. Point sources discharging at a rate under their respective caps are awarded credits in accordance with their degree of over-compliance. These credits in turn represent the right for the holder to discharge a designated amount of nutrient per unit time. Point sources exceeding their caps must purchase credits to balance their excess or pay a penalty. In the longer run point sources exceeding their caps may opt to upgrade their facilities, usually a prohibitively expensive option. Nonpoint sources are not subject to a cap but can earn credits by voluntarily adopting land management practices that reduce their nutrient discharge above an officially designated baseline. These credits can then be sold to point sources. Because nutrient reductions associated with nonpoint sources are generally much less expensive than treatment plant upgrades, NRT holds

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

great potential to reduce diffuse nutrient pollution by rewarding adoption of relatively simple and inexpensive changes in land use practices.

In principle, NRT as described here is similar to greenhouse gas cap-and-trade programs that award offset credits for forest carbon sequestration or farming practices which sequester or release less green gas into the atmosphere. There are, however, some important differences. In the trade of atmospheric pollutants spatial dimensions are generally not considered. The pollutant is considered to be perfectly mixed in the receiving environment, i.e. the atmosphere. The scope of NRT is typically constrained by spatial considerations, with trades confined within a single watershed or cluster of watersheds to prevent the occurrence of pollution “hotspots.” For example, if a wastewater treatment plant chooses to purchase credits instead of implementing a plant upgrade to meet its discharge cap, and the purchased credits originate from a distant watershed, then the waters in proximity of the treatment plant would be at risk of greater than acceptable pollution. Another difference is in the viability of “banking” credits. Greenhouse gas trading programs generally allow banking of credits, i.e. credits can be held in inventory for sale at a later time. NRT programs, in contrast do not allow banking of credits because nutrients in their bioavailable forms have a much shorter residence time than greenhouse gases, due to a multitude of biophysical effects including tidal flushing, uptake by primary production, and the process of denitrification.

We believe that lack of consensus on the appropriate structure and potential efficacy of NRT is a contributing factor behind the stagnated growth of nutrient reduction trading. Part of this lack of consensus revolves around the question of market structure. The vision of researchers and proponents of NRT seems to focus on a commodity style market featuring exchange of uniform nutrient reduction credits. However, when one examines project documents many stakeholders appear to adhere to a mental model of ad hoc bilateral trade of individually approved contracts. Perhaps more important is lack of consensus between NRP proponents and governmental entities with legal and executive power. King (2005) and King and Kuch (2003) point to lack of willingness on the part of government decision-makers to mandate the binding discharge caps and stringent penalties that are necessary to drive demand for reduction credits, but that are likely to be unpopular with much of their constituencies.

A role for System Dynamics in design of NRT programs

A number of recent works have demonstrated the power of System Dynamics in facilitating design of environmental institutions and policies. Saeed (2004) developed a generic model for design of environmental mitigation banking in which he experimented with a wide range of institutional and market structures. Ford (2005) developed a system dynamics model to gain insight into design issues for renewable energy credit trading programs. Important works by Stave (2003, 2002) demonstrate the usefulness of system dynamics in participatory design of environmental programs.

System dynamics gives us the opportunity to experiment with markets and other dynamically complex systems at relative low cost. It also allows experimentation with

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

assumptions that may be impossible or excessively risky to implement in the real world. NRP is dependent on binding discharge caps and stringent penalties that cannot be effectively implemented and tested in a prototypical project. Also, failure of a trial project could have severe ramifications for the credibility of NRT, providing a disincentive to take action or conduct risky real world experiments. Lag times in the adoption and implementation of improved landuse practices and/or waste treatment upgrade may be on the order of years or even decades. All these factors suggest a role for system dynamics in NRP design considerations.

The primary purpose of the model described below is to facilitate design by encouraging discussion and debate. The work is exploratory and somewhat speculative in that fully functional NRT programs have yet to emerge in the real world. We assume a commodity-type market for standardized reductions credits. Material delays, information and perception lags are explicitly accounted for. Particular attention is paid to the cap-setting decision as a policy leverage point. The model only considers total nitrogen as a nutrient. However, the existing model could be easily adapted to consider phosphorous.

Model structure and boundary

The model is divided into 6 sectors each embodying distinct decision making processes. The sectors are represented by hexagons in Figure 1. Flows and information linkages between the sectors are indicated by labeled arrows. A glance at Figure 1 shows that the sectors interact endogenously with the exception of the Influent generation sector.

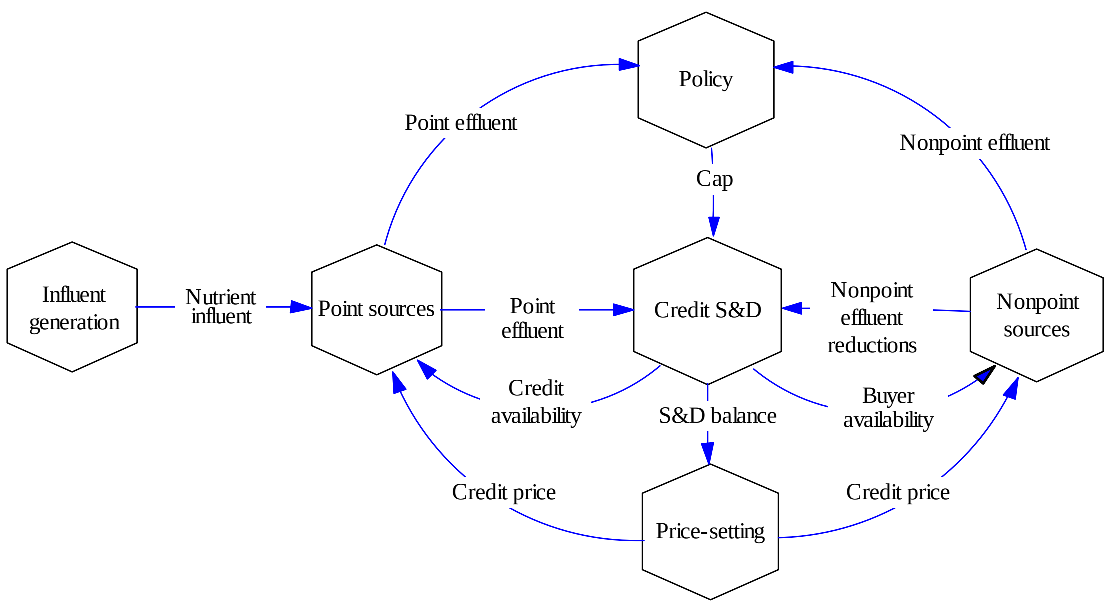

Figure 1. Model sector structure. Hexagons represent model sectors embodying distinct decision-making structures. Information cues and physical flows linking the sectors are indicated by the labeled arrows.

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

Sector assumptions

The *Influent generation sector* assumes that all nitrogen influent to point sources arises from household wastewater. Population is represented by a single stock and increases exponentially. The sector assumes that waste water generation per person is constant and that nitrogen concentration in wastewater remains constant.

The *Point sources sector* assumes that all point sources of nutrient pollution are wastewater treatment plants (WTPs) treating the household influent. The WTPs are placed into two categories, “old” and “new.” Old WTPs release wastewater effluent at a nutrient concentration of 8 mg/liter. New WTPs with superior technology release at 3 mg/liter. These concentrations are in line with figures from a large database of wastewater treatment plants maintained by the World Resources Institute.

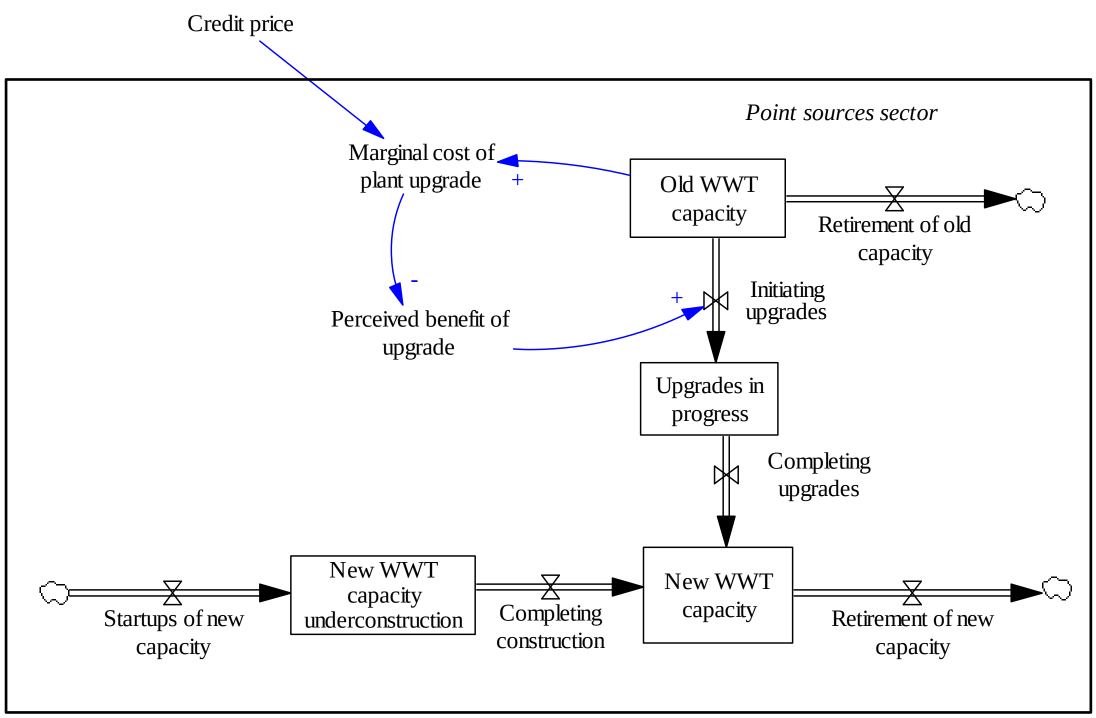

Figure 2. Simplified diagram of the Point sources sector. The variable shown outside the large rectangle is developed in another sector.

The two categories of WTPs are represented by capacity supply chains as shown in Figure 2. When old WTPs are retired they are replaced by new plants at the higher technology level. Old WTPs have the option to upgrade to the new technology level if the perceived benefit is adequate. Review of unit upgrade cost of WWT plants in the Chesapeake region very wide variation in cost of upgrades. We assume that the least expensive upgrades would be selected first and that marginal upgrade costs increase as plants are upgraded. The startup rate of new WTPs takes into account the retirement of old and new WTPs and uses a simple forcasting heuristic to anticipate necessary capacity expansion to accommodate the growing population.

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

The *Nonpoint sources sector* (Figure 3) makes use of a supply chain to track the adoption and implementation of nutrient reduction investments. We assume that these investments occur on agricultural lands. Possibilities for reducing nonpoint nutrient discharge include the planting of grass or forest filter strips, riparian vegetation, the adoption of cover cropping, adoption of conservation tillage, or combinations of these options.

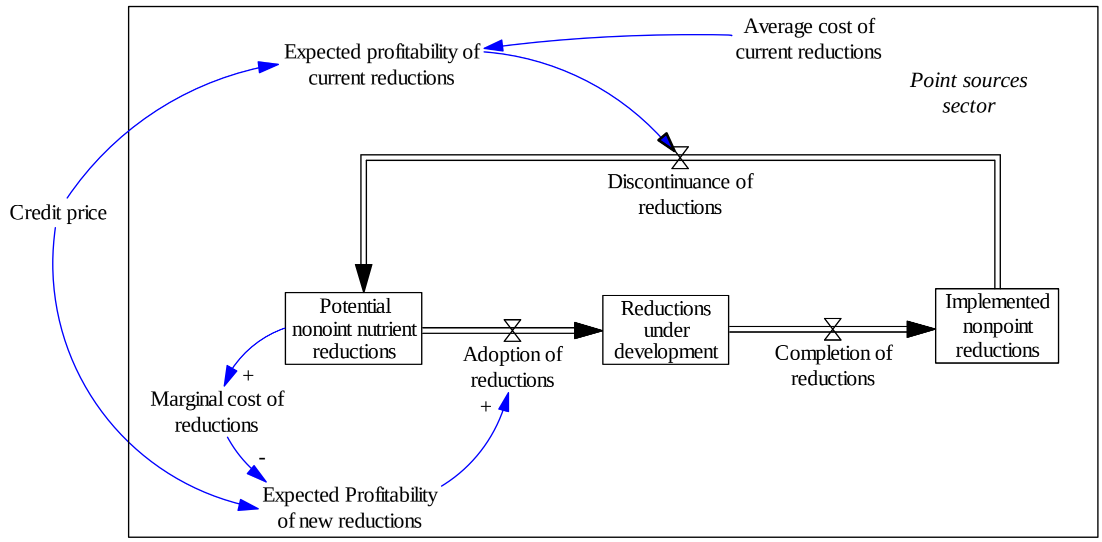

Figure 3. Simplified diagram of Nonpoint source sector.

We assume that the cheapest options are taken first and that the marginal cost of adoption increases as options are depleted. The cost curve for marginal reduction cost is developed with data from “Nutrientnet” an online facility developed and maintained by the WRI to aid farmers considering nutrient reduction options. Nonpoint source managers adopt reduction investments on the basis of expected profitability. To model this decision process we have adopted the archetypical capacity investment structure described by Sterman (2000, Chapter 20). We assume that managers may disinvest if current operations are perceived as unprofitable. A co-flow structure is used to model average cost of current operations (not shown in Figure 3).

In the *Credit supply and demand (S&D) sector* restoration credits are awarded to point sources exceeding their nutrient discharge caps, and to implemented nonpoint nutrient reductions. Demand for credits is defined by point nutrient discharge in excess of the discharge cap. The credits are considered to be valid for only one year, thus there is no explicit inventory of credits in the model, reflecting the disallowance of credit banking in most NRT programs. For point sources one kilogram per year of nutrient discharge equals one reduction credit. For nonpoint sources three kilograms of nutrient reduction per year are required to earn one credit, a three to one “trading ratio.” This is a typical value for the trading ratios applied to nonpoint sources in NRT programs (King and Kuch 2003). The trading ratio compensates for uncertainties inherent in nonpoint nutrient reductions and provides for net environmental gains. The ratio of credit demand to supply

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

effects price (modeled with a power law function), mimicking the aggregate effect of bids and offers.

The *Price-setting sector* employs the anchoring and adjustment pricing heuristic described by Sterman in which traders'expected price adjusts gradually to current price and, in turn, provides the anchor for the current price (Sterman 2000, Chapter 20). If a point source discharges nutrients in excess of its cap and cannot purchase reduction credits (due to lack of availability) then that point source must pay a penalty for each unit of discharge over the cap. In the model this penalty is set to 1.5 times the expected maximum nonpoint credit acquisition cost (developed in the Point sources sector). The penalty then serves as a maximum credit price because no under-compliant point source would be expected to pay a credit higher than the penalty. The current credit price is then formulated as:

MIN(Expected Price*Effect of S&D Balance, Maximum Price).

Where Maximum Price is equal to the penalty amount.

The *Policy sector* contains the cap-setting decision. The cap is defined in an aggregate sense and applies only to the point sources as discussed above. The cap is distributed between the old and new WWT plants on the basis of their respective fractions of total treatment capacity. In the real world this each individual WWT plant would be assigned its own cap, perhaps on the basis of its fraction of total capacity. It is important to keep in mind that the cap is *not* the overall maximum desired rate of nutrient discharge, but is a policy lever to generate credit demand and thereby encourage adoption of nonpoint nutrient reductions. The discussions of the cap-setting decision are scant in the literature. This is surprising considering the importance of the discharge cap in a market for nutrient reduction credits. The Policy sector is setup up in a way that allows experimentation with a variety of cap-setting heuristics. These include a fixed cap, a fixed fractional cap reduction, and an adaptive cap that responds to the sate of the system. These options will further discussed in the next section. The policy sector accounts for delays in perception of nutrient discharge levels within the trading area. The overall goal for aggregate discharge reduction is set to forty percent which matches suggested discharge goals for the Chesapeake Bay and Mississippi River watersheds (Hawn 2006, Nishizawa 2003).

Model boundary

Nutrient reduction trading is highly complex dynamically and in detail. Two important boundary assumptions of the present work are as follows:

1) The work does not consider environmental impacts or feedbacks directly from the environment. Rather, expert recommendations for the required degree of nutrient reduction are accepted. We then attempt a NRT design to achieve those recommended reductions.
2) As mentioned above the model is based on a commodity market assumption. Many of the discussions of NRT assume an ad hoc market based on bilaterally negotiated contracts. In some cases bilateral trade may be more appropriate. We

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

believe, however, that commodity style trading of standardized reduction credits will be necessary for NRT to succeed on a large scale, for example in the Mississippi drainage system.

Model behavior

The model is simulated over a time horizon of 75 years. Well beyond the planning horizon of NRT programs but sufficiently long to display dynamics of interest. Returning to our discussion of the cap-setting decision, Figure 4 shows percent nutrient reduction and credit price patterns under fixed cap and fixed cap reduction policies.

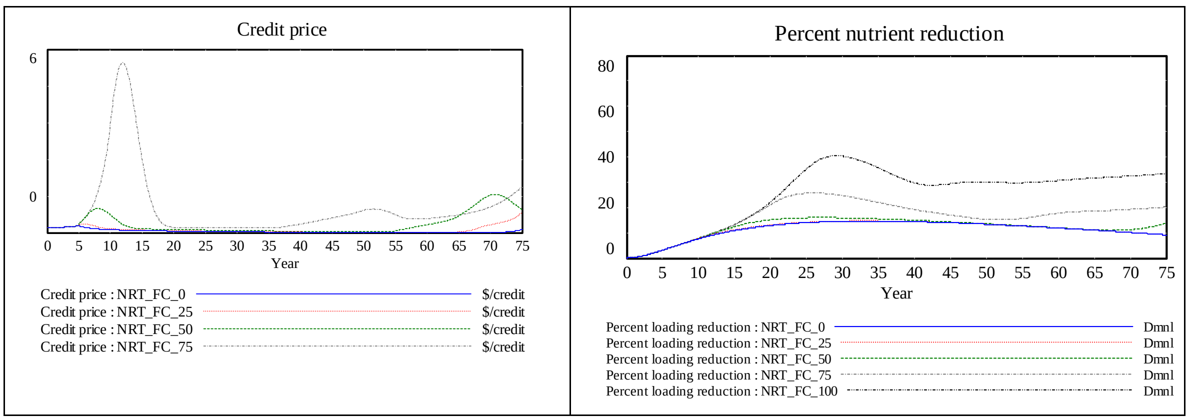

Figure 4. Credit price and Percent nutrient reduction simulated under fixed cap assumptions. The caps are set according to the initial supply and demand balance they create. When cap is set at 0, this indicates that initial S&D is equal. If cap is set at 25, then initial demand is 25 percent greater than supply. NRT commences at year 5.

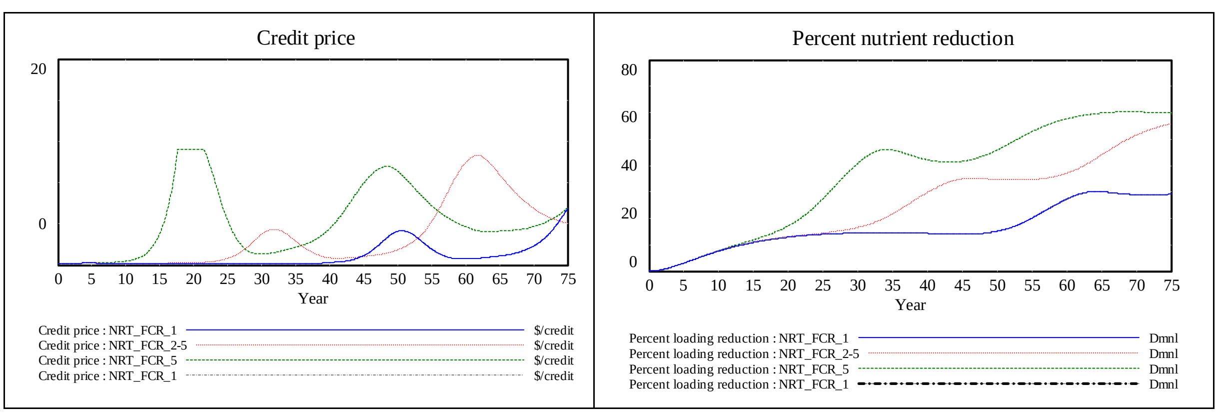

Figure 5. Credit price and Percent nutrient reduction under a range of fixed fractional reduction rates commencing at year 5.

The graphs show that fixed caps or fixed cap reduction strategies could achieve or even surpass the goal of 40 percent nutrient discharge reduction. However, patterns of credit price differ substantially. Fiddaman, in his system dynamics analyses of climate change

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

policy, points out that there is no reason to expect any particular constant tax or permit level to be optimal (2002). He terms such a fixed policy as a “ballistic strategy…that must fit all uncertain futures (2002 p. 261,)” and suggests that a “feedback control rule” or adaptive approach would be more useful (2002).

Figure 5 shows simulations of an adaptive cap that responds to perceived total nutrient reduction. Under adaptive cap-setting the cap adjusts to an indicated level. We assume that the adjustment time is 2 years. The indicated cap is defined on the basis of perceived nutrient discharge reduction achieved. As long as perceived nutrient discharge is below the desired level the cap is gradually reduced. When the desired level is achieved cap reduction ceases. If the desired level is exceeded then the cap is raised, reducing the compliance cost of point polluters.

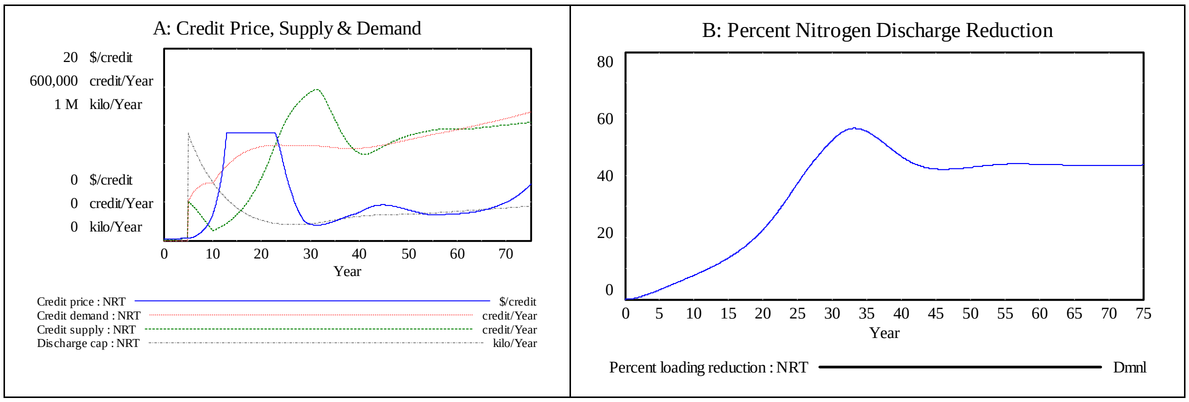

Figure 5. Simulation under adaptive cap. Graph A shows price, supply, demand, and the level of the aggregate discharge cap. Graph B shows percentage reduction in total nutrient discharge.

Graph 5A shows price, supply and demand, and the cap level. When the cap is implemented supply and demand are approximately equal. As the cap is lowered demand outstrips supply and the credit price rapidly rises to the penalty price. Nonpoint reductions are made in response to the attractive price but time delays in adoption and implementation prevent supply from meeting demand until about time 20. At this point an overshoot occurs as many reductions are moving from development to implementation. The credit price sharply drops and a portion of the nonpoint nutrient reductions are discontinued and the system enters an oscillating phase. As the nutrient discharge goal is met and exceeded the cap is gradually raised. The price pattern is similar to the price patterns for tradable green energy certificates modeled by Ford (2006); and the dampening oscillations are what we would expect to observe in a commodity system.

The goal is achieved and the system performance may appear reasonable, but what happens when we alter base assumptions that might be expected to vary between geographical regions?

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

Figure 6 shows the results when the adaptive cap is simulated with varying rates of population increase.

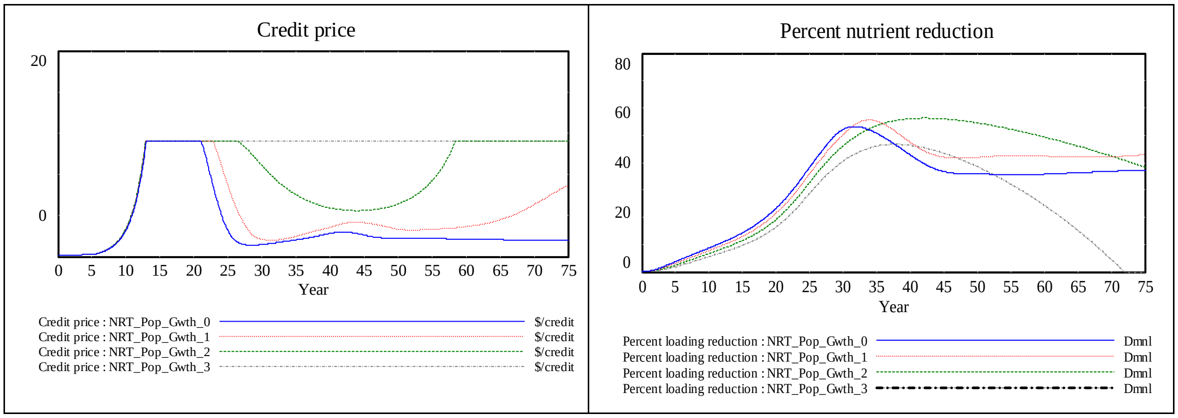

Figure 6. Patterns of credit price and percent nutrient reduction under a range of population growth rates.

When the population growth fraction is zero the results are not unlike the base case (in which the growth fraction is one percent). When set to 2 percent credit demand is greater and the initial price surge remains at the penalty level for a longer period. The price drop is much less than with the base case and price rises to and remains at its maximum as demand continues to grow with increasing population. At 2 percent growth nutrient discharge reduction begins a steady decline after reaching a peak around year 45. The decline occurs because the absolute rate of nutrient discharge from point sources is increasing despite the occurrence of plant upgrades and the gradual transition to higher treatment technology. Also nonpoint reduction potential is depleted. The increase in wastewater from the growing population is overwhelming improved treatment technology and the potential of nonpoint reductions. When population grows at 3 percent the credit price remains at its maximum. Total nutrient reduction peaks at a lower level and drops off much more quickly than in the 2 percent growth scenario.

The simulations suggest that population growth is an important consideration in the design of NRT programs. Watershed systems with high rates of growth may not be able to sustain nutrient reductions under a NRT program and point sources may be subjected to high compliance cost with little environmental benefit. In the case of high growth watersheds it may be important to expand the allowable geographical range of trading to include more rural watersheds that flow into the same water body of interest, even at the expense of water quality in the immediate vicinity.

Limitations

The model has many limitations which in some instances point to areas for further research. Some of these are:

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

- Treatment technology for new WWT plants is fixed in the present model. Technological innovation in treatment technology driven by NRT could be investigated in future modeling work.
- The model exclusively takes a commodity view of the NRT market. There is much interest in bilateral contract trade. Other forms of markets should be modeled. Information on categories of markets for NRT can be found in work by Woodward et al (2002).
- The model does not take into account transaction cost or the cost of maintaining institutions to perform monitoring, accreditation, and enforcement of regulations. These are all necessary components of NRT and may warrant explicit modeling, especially under assumptions of bilateral trade where transaction costs are likely to be substantial.
- The model does not explicitly consider spatial aspects. Even a simple spatial arrangement of lower, middle, and upper watershed may help gain credibility with potential clients.
- The present model has been developed with limited stakeholder participation. To make an eventual impact this model, or similar models developed in future, must be challenged and improved through stakeholder involvement.

Discussion

NRT programs now appear to be in a stagnated state of development (King 2005). NRT has been promoted in part on the basis of atmospheric pollutant trading programs which may be of limited applicability to NRT. Learning through real-world pilot studies appears to be difficult, as is the case with many dynamic systems (Sterman 2000, Chapter 1). We suggest that system dynamics modeling could play an important in facilitating consensus and design of these programs. The model described above is a work in progress and should not be considered as offering any definitive answers for program design. It is best thought of as a demonstration of the potential of system dynamics to contribute to NRT and similar projects.

References

Collentine D. 2002. Search for the northwest passage: the assignation of NSP (non-point source pollution) rights in nutrient trading programs. *Water Science and Technology*. **45** (9): 227-234.

Faeth P. 2000. Fertile ground: nutrient trading's potential to cost-effectively improve water quality. World resources Institute.

Fiddaman T. 2002. Exploring options with a behavioral climate-economy model. *System Dynamics Review*. **18** (2): 243-267.

Ford A, 2006. Simulating price patterns for tradable green certificates to promote electricity generation from wind. *Energy Policy*. **35** (1): 91-111.

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

Greenhalgh S, Sauer A. 2003. Awakening the dead zone: an investment in agriculture, water quality, and climate change. World Resources Institute.

Greenhalgh S, Faeth P. 2001. Trading on water. *Forum for Applied Research and Policy*. Spring 2001: 71-77.

Hawn, A. Nutrient trading and dead zones – can they wake each other up? The Ecosystem Marketplace. www.ecosystemmarketplace.com

King D. 2005. Crunch time for water quality trading. *Choices*. (A publication of the American Agricultural Economics Association). 1st Quarter, 2005: 71-75.

King D, Kuch P. 2003. Will nutrient credit trading ever work? An assessment of supply and demand problems and institutional obstacles. *Environmental Law Reporter*. 33 ELR 10352

Mehan T. 2006. Water quality trading: A guide for the perplexed. The Environmental Forum. May/June 2006: 4-5.

Nishizawa E. 2003. Effluent trading for water quality management: Concept and application to the Chesapeake Bay watershed. Marine Pollution Bulletin. **47** (2003): 169-174.

Saeed K. 2004. Designing an environmental mitigation banking institution for linking the size of economic activity to environmental capacity. Journal of Economic Issues. **38** (4): 909-937.

Stave K. 2003. A system dynamics model to facilitate public understanding of water management options in Las Vegas, Nevada. Journal of Environmental Management. **67**(2003): 303-313.

Stave K. 2002. Using system dynamics to improve public participation in environmental decisions. System Dynamics Review. **18** (2): 139-167.

Sterman J. 2000. Business Dynamics: Systems Thinking and Modeling for a Complex World. Irwin: Boston.

Woodward R, Kaiser A, Wicks A. 2002. The structure and practice of water quality trading markets. Journal of the American Water Resources Association. **38** (4): 967-979.

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --

-- Converted from Word to PDF for free by Fast PDF -- www.fastpdf.com --