Doing Water Quality Credit Trading Right

Alexandra Dapolito Dunn and Elise Bacon

The U.S. Environmental Protection Agency's (EPA's) Water Quality Trading Policy (Policy), 68 Fed. Reg. 1608 (2003), encourages states to adopt voluntary water quality credit trading programs and highlights how trading programs can facilitate the implementation of Total Maximum Daily Loads (TMDLs), reduce Clean Water Act compliance costs, promote voluntary pollutant reductions, and further other important watershed-based initiatives. In the two years that have passed since EPA released its Policy, many public clean water agencies have explored and then embarked on trading programs in an effort to resolve complex water quality challenges consistent with the Policy. Taking place in many different parts of the country, these projects have sought not only to build on lessons learned from the vast number of early trading projects, but also to explore new territory as encouraged by the Policy. Without question, today there truly is renewed energy and interest around water quality credit trading on the part of utility managers, watershed stakeholders, and the general public.

One might think that with this regulatory and policy climate, the time is ideal to consider developing a water quality credit trading program, and that the pathway to implementation might be quite smooth. In reality, however, many of the same core challenges that confronted almost every early trading program still remain. These problems largely come from several inherent inequities that are present in any watershed-based trading environment, and which cannot be completely eliminated because they derive from existing federal regulatory provisions. These inequities include differing regulatory authority over the parties in a trading program, the parties' varying need to make a trade in order to meet enforceable permit limits, and the related willingness of and pressures on different parties to come to the table and participate in a water quality credit trading program.

For example, as regulated point sources, municipalities must assure compliance with the effluent limitations in their National Pollutant Discharge Elimination System (NPDES) permits. As public entities responsible for providing an essential public health service, municipalities as a practical matter almost necessarily must come to the trading table for both regulatory and public outreach reasons. For NPDES permittees *not* interested in trading, regulatory authorities always retain the ability to simply exercise their enforcement powers to address a pollutant of concern. In contrast, trading is more discretionary for other participants in a trading program, such as unpermitted nonpoint sources.

Recent experience reveals that when these inequities are openly acknowledged at the start of a trading program and are not used by one party to disadvantage another, the water quality credit trading program has the potential to be quite successful. However, if any party to a trading program seeks to capitalize on the inherent inequities—either by asking too much from another party or by trying to force a particular result—the trading program is likely to crumble under its own weight.

To make trading work in the real world, the four key principles addressed in this article should be considered in the up-front design and subsequent implementation of a program. First, it should be understood that the participants in a trading program may freely use the credits or offsets generated under the trading program as a part of their regular compliance efforts. The opposite view suggests that the use of legitimately created credits by a program participant is evidence of ultimate noncompliance. Under this approach, the trader is viewed and ultimately treated as a violator rather than as an innovator, and the credit is characterized as mitigation for the violation. Programs adopting this "violator view" encounter difficulties by suggesting that sources should reduce effluent concentrations so greatly that the ultimate use of a credit is unnecessary, rendering trading moot. Moreover, treating credit trading as mitigation creates significant regulatory and public relations disincentives for participation in what is intended to be a voluntary effort.

Second, cost savings are a benefit. A source should not be viewed skeptically merely because it achieves cost savings through trading. In fact, EPA's Policy rejects the notion that sources ultimately saving money through trading cannot truly be benefiting the environment. Trading programs built upon the "skeptical view" ultimately will founder if stakeholders are so overly con-

Ms. Dunn is general counsel of the National Association of Clean Water Agencies in Washington, D.C., and can be reached at adunn@nacwa.org. Ms. Bacon is a principal technologist in CH2M HILL's Water Business Group in Herndon, Virginia, and can be reached at Lisa.Bacon@CH2M.com.

cerned about the savings sought by permitted sources through trading that the economic incentives for participation are kept below what they should be to draw participation.

Third, a trading program needs to be truly voluntary. Under semivoluntary or nonvoluntary approaches, regulators implicitly ask—or explicitly attempt—to force point sources to take on patently inequitable additional pollutant-reduction obligations in "exchange" for the privilege of trading. Many programs struggle to achieve the proper balance between a voluntary trading program that leads to pollutant reductions, and a less than voluntary program that is based on the regulator's ultimate ability to force a reduction by reducing a point source's permit limitation for a particular pollutant. This balance is difficult to achieve due to the significant differences in the Clean Water Act's approach to controlling point and nonpoint sources of pollution. Point sources in a trading program always are responsible at the end of the day for permit compliance, and the specter of enforcement looms in the background. On the other hand, because nonpoint sources do not require a permit, they can implement their reductions in a more flexible and iterative manner.

Fourth, and finally, the level of pollutant reduction needed to generate credits must be set fairly and with certainty. Ambiguous baselines for defining creditable reductions can undermine a trading program in two ways. Moving credit "targets" can lead to a perception that there will be no measurable pollutant reductions with trading, delaying TMDL compliance or making the ultimate attainment of water quality goals mathematically impossible. Shifting baselines also can leave credits vulnerable to diminution or invalidation, thereby presenting an unacceptable level of risk to prospective credit buyers and sellers.

If any party to a trading program seeks to capitalize on the inherent inequities . . . the trading program is likely to crumble under its own weight.

The next section of this article explores how these four principles have been successfully followed in trading programs to date. This article then analyzes a current trading initiative that lies in limbo because these principles are being ignored. Finally, we summarize key observations and speculate generally on the future viability of water quality trading.

Trading Done Right

A quick review of the trading programs that are being actively implemented across the United States today reveals that most are good, if not excellent, examples of how principles of equity and fairness can yield stakeholder satisfaction and sufficient support to take a trading program from design to successful implementation. Many of the benchmark programs in this regard were among the first formal functioning trading programs. They developed in the mid- to late-1990s, benefited from at least a five-year gestation period, and were launched by the early 2000s. Several newer programs also are characterized by equity and fairness between stakeholders.

The benchmark group includes established programs for these watershed areas: Lower Boise River, Idaho (point–point or point–nonpoint phosphorus credit trading); Cherry Creek, Colorado (predominantly point–nonpoint phosphorus credit trading); Long Island Sound, Connecticut (point–point nitrogen credit trading); San Joaquin Valley Grasslands, California (selenium credit trading among irrigation districts); and Tar Pamlico River, North Carolina (point–point or point–nonpoint nitrogen credit trading). Among the more recent programs, three stand out. One is the watershed permit and nitrogen trading program for certain Neuse River, North Carolina dischargers. Another is the watershed permit that supports point–point dissolved oxygen credit trading and point–nonpoint temperature credit trading by a single utility, Clean Water Services, with four wastewater treatment plants on the Tualatin River in Oregon. The third is the Virginia Nutrient Credit Exchange Program that will cover all municipal and industrial point sources in five major river basins that are tributaries to the Chesapeake Bay under a general permit and enable point–point trading among them, as well as point–nonpoint trading in specified limited circumstances. Certainly, other programs could be included in either group, but collectively these demonstrate how to do water quality trading right.

So how exactly—by circumstance or choice—do these programs illustrate the four principles introduced at the outset of this article? First and foremost, they clearly treat compliance achieved through trading as "equal" to compliance achieved without trading. Within these regulatory, policy, and philosophical frameworks, trading is legally, practically, and morally equivalent to not trading—not worse and not better, so long as effluent limits or loading caps are met, subject of course to specific rules governing eligibility, trading transactions and tracking and reporting of such transactions.

Supporters and participants in the benchmark programs were able to arrive at an approach that treats

trading as equal to not trading in part by agreeing on the spatial and temporal relationship between pollutant loadings and reductions at one source location relative to another. Trading ratios are the most common way watershed stakeholders have accounted for differences, known and unknown, between the impact of a pollutant contributed by Source A at Point A and Time A, and the benefit of a pollutant reduction by Source N at Point N and Time N. A wide variety in analytical approach and rigor is represented among the trading ratio methodologies of the benchmark programs. Several programs made extensive use of water quality models and demonstration project monitoring. Others elected to use or had to make do with best professional judgment using what local data was available. Still others simply adopted ratios that had been accepted elsewhere and were acceptable to the relevant authorities and decision-makers. Seasonally variable ratios and seasonally bounded trading periods are two other ways used successfully to establish "equivalency" between a pound of pollutant traded and a pound not traded.

A critical confirmation of the fundamental view that the trading option, once granted and if properly exercised, is an equal and not lesser outcome than pollutant reductions achieved by other means, is the complete lack of mitigation requirements or penalties levied on NPDES dischargers electing to use credits as permitted. In fact, most of the permits have carefully constructed limits, where the permittee is determined to meet its limits if: (1) its effluent concentrations or pollutant loadings are below allowed limits *or* (2) it holds, or otherwise obtains within sufficient time, enough valid credits to offset exceedences above the applicable limit. Of course, they must also meet other applicable requirements for compliance in the broadest sense. These requirements appear either with the dual effluent limits or elsewhere in the permit as a blanket condition governing all aspects of permit compliance. Together, these permit provisions maintain regulatory authority and ensure that compliance is judged not just by having enough credits but also by following all the rules that may be associated with securing and applying them. The net effect is that so long as (1) *or* (2) above, and the applicable requirements, are met there is never a moment—not even a second—where a permittee using credits as offsets is not in compliance with the Clean Water Act. Therefore, such permittees have created no harm requiring a penalty, mitigation, or remedy.

Second, within this framework of equivalency between trading or not trading, a primary driver for establishing a trading option in the first place almost always is to lower the cost of achieving an individual or collective water quality-related target, such as water quality-based effluent limits, wasteload and load allocations of TMDLs or other types of pollutant load caps or reduction goals. Stakeholders know or believe that more cost-effective ways to control pollutant loads exist than would be required or captured by a traditional regulatory approach. In a classic market sense, trading options harness differences in relative control costs among sources. Well-designed programs create economic incentives for sources that can do more than required, when they can sell those credits to sources with higher relative control costs. Credit buyers have incentive to participate in trading when they can save money over what they would have spent to achieve compliance without trading.

Well-designed programs create economic incentives for sources that can do more than required, when they can sell those credits to sources with higher relative control costs.

In a practical sense, many of the benchmark program credit markets are "managed," in that prices (and by implication profits) are set or controlled. Most are closed markets, in that only certain types of or named point sources are allowed to buy or sell credits, and eligible nonpoint source credits are preapproved by category, location or transaction mechanism. Despite the lack of a "pure" free market among the benchmark programs, they all are unapologetically encouraging, seeking—and finding—equal or better environmental results for less money through trading.

Importantly, no benchmark program exists where regulators or other stakeholders purposefully attempted to constrain the level of cost-savings that a point source buying or otherwise securing credits could realize. Certainly, some trading rules have the effect of reducing cost-savings compared to what point sources would see under different rules. But these outcomes generally have been reached as a matter of course through a bottom-up approach to program design as it pertains to cost-effectiveness and potential cost-savings. For example, trading ratios have a direct impact on the cost of a credit to a buyer that can be significant: double the trading ratio, double the cost of a credit. Thus, at a ratio of 1.5:1, a point source would need one and a half credits, that might cost $10 each, to offset one pound at its plant, for a total cost of $15, whereas at a ratio of 3:1, the point source would need three credits, for a total cost of $30. As mentioned above, most of the benchmark programs have followed a science-based approach to setting trading ratios, to the extent avail-

able tools and data allowed. It has not been suggested in any of these programs that the ratios were set arbitrarily high to limit cost-savings.

Third, point–nonpoint trading programs in particular must solve problems associated with a fundamentally unlevel playing field with respect to the disparate authority that the Clean Water Act and the NPDES program impose on point sources versus nonpoint sources when it comes to requirements for controlling pollutant loadings. By contrast, point–point trading programs pose fewer problems with respect to a level playing field. Concerns about equity and fairness in programs involving only point sources usually have to do with baseline responsibilities before additional treatment generates credits in cases where significantly different levels of treatment may be in use among the dischargers.

Clearly, the federal government and states have powerful tools at their disposal to establish and change (usually tighten) wastewater treatment requirements with technology- or water quality-based effluent limits. In cases where TMDLs are in place, NPDES permits provide a strong regulatory mechanism to force compliance with wasteload allocations (WLAs). In contrast, limiting pollutant loadings from nonpoint sources is largely accomplished, if at all, through federal and state cost-share programs that help to pay for best management practice implementation. In fact, most state regulations applicable to nonpoint sources are for the development of nutrient management plans, or take the form of local ordinances of varying strength and scope that govern land use activities. Rarely are additional mechanisms created to implement the load allocations (LAs) assigned to nonpoint sources.

The disparity in regulatory leverage over point and nonpoint sources can tempt regulators and other stakeholders in a trading process to disproportionately lower point source wasteload allocations (WLAs) relative to their individual or collective contributions to the particular water quality problem. The objective is that pollutant reductions will be "successfully achieved" more quickly than through a more equitable allocation of responsibilities among point and nonpoint sources. It seems that this temptation was not acted upon in the benchmark situations, as many of the TMDLs were established prior to or independently of the trading option. While in these situations some point sources have scientific or procedural disagreements with their WLAs and/or with the overall assignment of load allocations (LAs) among point sources, and between the collective WLAs and LAs, there were no claims that a WLA or group of WLAs was set lower because of or to facilitate or force trading.

Within the point–nonpoint source trading arena, another problematic approach is setting trading ratios to achieve a specific outcome. This is done by first determining a target nonpoint source load reduction then identifying the trading ratio for a specific point source or group of dischargers that offers the best chance of hitting that target, based on analysis of relative control costs and potential demand for credits. This problem is similar to setting trading ratios artificially high in an attempt to limit cost-savings to credit buyers. Again, among the benchmark programs, trading ratios appear to have avoided these pitfalls by using a bottom-up approach working from baseline individual or collective environmental performance targets—ultimately yielding successful programs.

Stakeholders developing point–nonpoint source trading options must define carefully and clearly what nonpoint source actions are creditable to the satisfaction of participants and supporters.

The fourth principle essential to a successful program is equity and fairness in determining what actions can generate sellable or otherwise usable credits, and whether any credits are either ineligible or limited in amount or durability in some respect. By definition, every trading program needs to establish trading baselines. When loadings are below the baseline, the difference (surplus reductions) may be fully or partially creditable. When loadings are above the baseline, credits will be needed to offset the exceedence or overage. If TMDLs are in place, the default baselines are the WLAs and LAs. If TMDLs are not in place, existing individual technology-based or water quality-based effluent limits define the baselines for point sources, while existing programmatic or legal requirements (state and/or local) define the baselines for nonpoint sources.

As a practical matter, baselines for point sources are unambiguous, firm, and not optional by virtue of NPDES effluent limitations if no TMDL has been conducted or, if a TMDL is in place, via WLA assignment and implementation through permits. Baselines for nonpoint sources can be just the opposite, depending on the approach to land use management taken in a specific watershed. This means that stakeholders developing point–nonpoint source trading options must define carefully and clearly what nonpoint source actions are creditable to the sufficient satisfaction of participants and supporters. At a minimum, this establishes a level playing field within the nonpoint source sector—minimizing rewards for laggards. But more importantly, it elimi-

nates the risk to point sources using credits to comply with their loading limits that a program participant or third party will come along after the fact and claim the credits are invalid because the actions that generated them were supposed to be counted toward some other legal or voluntary obligation. The benchmark programs have successfully defined nonpoint source baselines with sufficient equity and fairness to gain support for program implementation and participation.

For the most part, the benchmark programs implemented to date demonstrate a variety of approaches to assembling the key trading design elements into programs that generally are considered equitable and fair by all involved. This success must be attributed in part to the technical acumen, problem-solving abilities, and commitment to multiyear program design phases by key stakeholders. An equally or probably more important hallmark of these programs is the collaboration that exists, at a minimum between the point sources and the regulators, and in many cases among a wider group of stakeholders including nongovernmental organizations (NGOs). It is true that in some cases the trading program was proposed by point sources as an alternative to a newly proposed water quality target or TMDL implementation approach, but these rebuttals became the basis for cooperatively finding a more cost-effective approach and minimizing the inherently adversarial relationship between the regulated and the regulators.

Some have suggested that trading ratios be based on the relative cost of trading versus not trading, instead of the relative environmental benefit of pollutant load reduction at one source versus another.

Trading Going Wrong? A Case Study

Somewhere between one and two dozen trading initiatives are at various stages of development around the country in addition to the designed programs that have passed into the implementation phase. For many of these, it is too early to tell where the participants will end up relative to the four principles of a good trading program. One particular situation involving point–nonpoint source trading, however, reveals how prospects for success can be compromised when the four principles are disregarded. In this case, the NPDES permittee has proposed a water quality credit trading program well aligned with the principles of equity and fairness. In contrast, some stakeholders, disregarding these key principles, are proffering a very different framework for trading than exemplified in the benchmark programs. The situation described is not doomed yet, so will remain nameless in order not to prejudge its chances for a turnaround. In this article we refer to it as the "Limbo Case," as it is in a state of limbo between continuing or dissolving.

First, in the Limbo Case, the regulatory authorities have asserted that a permittee needing to use a credit ultimately would be mitigating environmental harm via the credit. This approach is making the trading program design discussions more like a natural resources damage assessment (NRDA) discussion, in which the actions of parties who have injured the environment are subject to fines and penalties. NRDAs typically are calculated using economic methods that estimate the monetary value of the harm (e.g., in terms of clean up costs or stakeholder willingness to pay for certain benefits) and consider the savings or profits the polluter reaped by engaging in the violation. Depending on the circumstances, the required payments may include a penalty component over and above the amount necessary to return the harmed area to its original condition, or provide an equivalent environmental benefit somewhere else if it is impossible to completely mitigate the impacted site. This is an incorrect framework to apply to a trading program on a legal, regulatory or policy basis because according to our first principle, no environmental harm has occurred.

Second, in the Limbo Case, some have suggested that trading ratios be based on the relative cost of trading versus not trading, instead of the relative environmental benefit of pollutant load reduction at one source versus another. This approach arbitrarily (or not so arbitrarily) limits cost savings and is inconsistent with both EPA's Policy and the fundamental economic forces that make trading work. The trading ratio is proposed to be derived from the permittee's economic and financial data, setting the trading ratio such that a credit would cost the permittee a desired fraction of the in-plant unit control cost. Those advancing this approach have not yet stated their target limit on cost-savings the permittee would achieve through trading, except to say that the proposed credit package will be evaluated on the basis of its cost compared to in-plant treatment.

In the Limbo Case, the regulatory authorities appear to have a substantial concern with the permittee achieving an equal or better environmental result for less money. Notably, the assumption also seems to be that an in-plant option, the cost of which would be used to limit cost-savings with trading, is both technically and financially feasible. Neither has been proven yet. The permittee has pointed out that even if an in-plant option would provide the additional control level neces-

sary to achieve the loading cap without trading, the cost is so disproportionately high compared to the incremental reduction gained and to the relative contribution of the permittee's load to the total pollutant load in the waterbody system as to render this option infeasible from an affordability standpoint. As such, it should not be viewed as the default nontrading option, much less used as the basis for setting trading ratios.

Third, other reasons that might lead to a decision to base trading ratios on cost factors, and not on water quality benefit equivalencies, would be a goal of achieving a target level of credit revenue or a goal of achieving a target level of creditable pollutant load reductions. Both reasons evidence a top-down approach that generally does not respect an equitable allocation of responsibilities between point and nonpoint sources. A revenue stream from a point source's credit purchases can be an attractive source of funding in the face of a large number of unpermitted nonpoint sources in the watershed. As stated earlier, trading ratios can be set inappropriately in a manner that favors nonpoint source load reduction objectives. In the Limbo Case, the regulators are asking what the trading ratio needs to be to get the target level of funds or mass reductions. That is the wrong question according to the principles of a good trading program. The correct question is what the trading ratio needs to be to establish environmental equivalency between a pound reduced/loaded at Source A versus Sources B, C, and others. Taken to extremes, this wrong approach will lead to a situation where one or more permittees are made responsible for reductions equal to or even greater than 100 percent of the total load.

The Limbo Case also fails on the fourth principle of clarity and certainty of creditability. In this situation, a draft TMDL has been developed for some of the watersheds where prospective nonpoint source credit projects are being considered. The draft TMDLs assign LAs to selected nonpoint sources that will require significant reductions over current loads to achieve water quality standards. EPA's Policy states that when TMDLs are in place, LAs establish the nonpoint sources' baseline for trading, which by definition only makes reductions beyond that baseline—in the Limbo Case the remaining and relatively insignificant percentage—creditable. For each of the nonpoint source projects under consideration, limiting creditable reductions to a small percentage of the total will not provide a sufficient number of credits for the NPDES permittee's estimated future needs.

Several solutions have been proposed to this policy and mathematical conundrum. The LAs could be revised to leave more room for creditable reductions. Specific projects or portions of LAs could be made creditable in the TMDL implementation plan, but it is not clear how the TMDL and the implementation plan would then match up. Some have suggested crafting language for an "early actors" provision that would allow permittees to generate credits by reducing loads already assigned to other parties under the LAs, but only if the point sources acted within a specified period of time. This approach would not completely absolve those nonpoint sources of their LA responsibilities, but would create strong incentives to implement reductions as creditable, which most agree will not soon (or will not ever) happen otherwise. Finally, several time-consuming and legally cumbersome avenues appear to be open to the state to formally declare some sites with LAs "orphans," which presumably would free up those allocations as creditable. As of this writing, it remains unclear if this issue will be resolved. Open questions regarding the creditability of reductions will discourage most NPDES permittees from entering the trading program in any meaningful way.

Trading must work in tandem with other efforts to bring a watershed into compliance.

Future Directions

Where the four principles are followed—credit use without penalty, cost-savings without scorn, maintaining a truly voluntary trading program, and establishing certainty around credit generation—trading can be a successful complement to achieving water quality goals in a watershed. These principles were followed in the benchmark cases and all stakeholders—regulated, regulator, NGO, and others—largely benefited from enhanced water quality progress. However, when these principles are not followed, the result becomes less and less enticing to the potential participants. Trading must work in tandem with other efforts to bring a watershed into compliance. It cannot be singled out as a saving grace, a quick ticket, or an expedient approach. A successful trading program depends on thoughtful and equitable approaches. A trading program should not squeeze certain sources to put them in the position of taking responsibility for substantially more than their fair share of a pollutant load. If this occurs, the net result will be a trading program ultimately set on a course for economic and environmental failure. Given that it is most stakeholders' hope that trading becomes a viable part of the toolkit in watersheds nationwide, careful and open consideration of the principles outlined in this article will ensure that future programs ultimately are added to the list of successful trading programs.