



Do You Know a Successful Trading Program When You See It?

NATIONAL NETWORK ON WATER QUALITY TRADING

A REPORT FROM THE SPRING 2016 DIALOGUE

The National Network on Water Quality Trading

The National Network on Water Quality Trading (Network) was established to improve consistency and integrity across WQT programs by providing a community of WQT practitioners to articulate shared principles, core trading program design elements, recommendations for implementing and operating trading programs, and lessons learned from experience.

Formed in 2013, the Network started with 18 participating organizations (below) representing the diversity found in most emerging trading programs in the country, including agriculture, wastewater and stormwater utilities, environmental groups, regulatory agencies, and practitioners delivering water quality trading programs. The Network's strength is derived in part from its diversity of experience and viewpoints. Thus, all Network participants are free to maintain their own individual positions on any issues or documents discussed or published by the Network as a whole.

The Network uses a facilitated dialogue structure to connect participants in order to share lessons learned and work collaboratively on research agendas, tools, and resources that can benefit the water quality trading community. The Network welcomes new participants to ensure that expertise and diversity are well represented at each dialogue workshop. Any entity that has read the National Network guide [Building a Water Quality Trading Program: Options and Considerations](#), and can commit to participate actively in a dialogue may consider themselves a Network participant.

The National Network is guided and governed by a steering committee that sets the agenda and objectives for upcoming dialogues, selects the Network Coordinator, and identifies experts that can guide and help convene Network Dialogues.

National Network Steering Committee

American Farmland Trust	Maryland Department of Agriculture
Association of Clean Water Administrators	Mississippi River Collaborative
Electric Power Research Institute	The Ohio Farm Bureau Federation
National Association of Clean Water Agencies	Troutman Sanders
National Milk Producers Federation	Willamette Partnership – Coordinator
The Freshwater Trust	World Resources Institute – Coordinator
Kieser & Associates, LLC	U.S. Department of Agriculture – Technical Advisor

Table of Contents

The National Network on Water Quality Trading	i
Executive Summary.....	iii
Spring 2016 Dialogue	iii
AGENDA	iv
Dialogue Summary.....	1
1. Introduction to program evaluation and performance measurement in water quality trading.....	1
2. Approaches and methods for measuring and reporting progress.....	2
3. Scorecards and metrics for water quality trading	6
4. Identifying common metrics roundtable discussion.....	8
5. Communicating and Incorporating Results into Decision Making.....	9
6. Final Thoughts from the Spring Dialogue Workshop.....	12
Continuing the Conversation	13
APPENDIX I: 2016 Spring Dialogue Participants.....	14
Appendix II: Speaker Bios.....	16
APPENDIX III: Water Quality Trading Performance Metric Brainstorm Results	17

Executive Summary

Spring 2016 Dialogue

When asked about the issues most important to the future of water quality trading, participants in the Network quickly identified challenges related to how water quality trading programs track progress and measure success. It can be difficult to document claims of trading program advantages and disadvantages not only because many programs are small or new, but there is also a lot of diversity in what *success* means. Three factors drive variation in the definition of success:

- **Trading programs are often based in the Clean Water Act (CWA) and National Pollution Discharge Elimination System (NPDES) permit program.** So is success permit compliance? Or is success meeting water quality standards? Or somewhere in between?
- **Point-nonpoint trading is often suitable where the proportion of nonpoint pollution loading is higher in the watershed than point source loading.** So can the contribution of trading to in-stream water quality be detected? Or is success defined as reducing pollutant loads?
- **Trading builds new relationships between cities, farms, and other stakeholders.** So how important is improved trust and satisfaction relative to meeting environmental goals?

The focus of the Spring 2016 Dialogue looked at the different elements and considerations of program evaluation and performance measurement in the context of water quality trading. The goal of the dialogue was to identify the resources and tools that would help trading programs determine how evaluation and performance metrics can fit their program and ultimately answer questions about success.

Over the two day dialogue workshop participants recognized the value of program evaluation and performance measurement as water quality trading programs across the country are asked to demonstrate effectiveness. However, when considering the elements of program evaluation and performance measurement, administrators expressed concerns about the resources available to create such assessments, as well as the feasibility to develop metrics that could be employed uniformly across programs. Participants felt that the discussion of program evaluation and performance measurement as applied to water quality trading is in its infancy with little direct experience to draw from. To help administrators explore the potential for program evaluation and performance measurement further it was recommended that broad tools such as a logic and maturity model could prove to be valuable starting points. Over the coming months, Willamette Partnership with assistance from participants of the Spring 2016 Dialogue will work to develop these example models for further consideration and revision.

How Do You Know a Successful Trading Program When You See It?

Measuring Progress in Water Quality Trading

May 24-26, 2016 St. Louis, MO

AGENDA

May 24 - Reception and Introductions

5:30 – 7:00pm	Welcome and introductions Bobby Cochran, Willamette Partnership
	Framing the issue: Measuring progress in water quality trading How do you know a successful trading program when you see it? Goals of the Dialogue Bobby Cochran, Willamette Partnership Jessica Fox, Electric Power Research Institute
	Participant networking session/reception

May 25 – Dialogue

8:30-9:00am	Introductory remarks from Conveners Introductions and reiteration of dialogue goals – Bobby Cochran, Willamette Partnership Defining language – Jessica Fox, Electric Power Research Institute and Neil Crescenti, Willamette Partnership.
	9:00 – 10:15am The Business case for evaluation: If you don't know where you're going, any road will get you there. <u>Presentation:</u> An introduction to evaluation and performance measurement. Mackenzie Staub, Washington University <u>Group discussion:</u> What is the value of evaluation and performance measurement? Who are the audiences most interested?

10:15 – 10:30am	BREAK
10:30am-12:30pm	<p>Approaches and methods for measuring and reporting progress</p> <p><u>Presentation:</u> Common approaches or practices to building an evaluation and measurement system – Sarah Bobmeyer, Washington University</p> <p><u>Presentation:</u> Scorecards and metrics for water quality trading: – Jessica Fox, Electric Power Research Institute</p> <p><u>Brainstorming session:</u> What are common goals of water quality trading programs?</p>
12:30-1:30pm	LUNCH
1:30 – 3:00pm	Roundtable discussion and exercise to identify common metrics – Jessica Fox, Electric Power Research Institute
3:00-3:15pm	BREAK
3:15-4:30pm	<p>Communicating and incorporating results into decision-making</p> <p><u>Presentation:</u> Lake Tahoe Environmental Improvement Program – Kim Caringer, Tahoe Regional Planning Agency.</p> <p><u>Panel Discussion:</u> EPA Evaluation of Trading and Offset programs in the Chesapeake Bay – Bob Rose, EPA, Jason Keppler-Maryland Department of Agriculture, Brian Schlauderaff- Pennsylvania Environmental Protection Department</p> <p><u>Question and answer session</u></p>
4:30pm	ADJOURN
May 26 – Dialogue Conclusion and Next Steps	
8:30-9:00am	<p>Welcome and thoughts from yesterday</p> <p>Bobby Cochran, Willamette Partnership Jessica Fox, Electric Power Research Institute</p>
9:00am –12:00pm	<p>Where do we take this next?</p> <p>Facilitated discussion with participants on next steps, outcomes and products, roles and responsibilities.</p> <p>Bobby Cochran, Willamette Partnership Jessica Fox, Electric Power Research Institute</p>
12:00pm	ADJOURN DIALOGUE

Dialogue Summary

1. Introduction to program evaluation and performance measurement in water quality trading.

Speakers

Jessica Fox, Electric Power Research Institute
Bobby Cochran, Willamette Partnership

Presentation: Welcome and Introductions

How can programs show that trading is contributing to water quality improvement? Why are trades occurring in some programs and not others? How do you know a successful program when you see it? Despite the interest in these questions, program evaluation and performance measurement are not widely implemented in the field of water quality trading. To help programs meet the demands of transparency and accounting with the understanding that they are also faced with limited resources, the goal of the Spring Dialogue is to identify the resources and tools that can help administrators determine how evaluation and performance measurement can fit their specific water quality trading program and ultimately answer questions about success.

Participant discussions and questions

With feedback from participants the objectives of the dialogue workshop were identified as:

- Identify a menu of trading program performance metrics and characteristics of metrics, how they can be used, when is it appropriate to compare across programs;
- Investigate and assess how evaluation systems can be tailored for different stakeholder groups and contexts;
- Using evaluation and performance measurement as a communication tool with the public;
- Develop a product that can be used by water quality trading programs interested in evaluating and reporting out on their performance;
- How to prioritize metrics based on target audiences and program objectives; and
- How to link program evaluation and performance measurement to goals and outcomes early in the program design process.

2. Approaches and methods for measuring and reporting progress

Speakers

Mackenzie Staub, Washington University

Presentation: The Business Case for Program Evaluation

Before jumping into *what* program evaluation looks like or *how* evaluation works, it is important to start with the *why*. The answer to the *why* question is best captured by a quote from Yogi Berra who said “If you don’t know where you’re going, you’ll end up somewhere else”. Program evaluation is the systematic and intentional process of gathering and analyzing data to learn, inform decision-making, and to take action towards program improvement.

There are several common concerns that have been expressed as reasons not to undertake evaluation including diverting program funding, overly complex processes, additional workload for staff, and the potential reactions to negative results. Program impacts such as funding and staff commitment can easily be addressed through appropriate design that aligns with an organization’s resources and priorities. As for negative results, finding out what doesn’t work so well is just as important as finding out what is working well.

Program evaluation explores whether outcomes were achieved in order to improve program processes. Where research seeks to *prove*, evaluation seeks to *improve*. In addition to process improvement, evaluation can measure the effectiveness of a program, information that is likely to be of great interest to external audiences such as funders, stakeholders, and the public. Ongoing evaluations can contribute to a program’s sustainability overtime and allow programs to adapt to changing external conditions. Finally by identifying what is working well and what isn’t, program evaluation has the potential to improve the field as a whole by contributing knowledge to others based on an organization’s experience.

Evaluations are often categorized as *formative* –those that provide information that helps a program improve and *summative* – those that demonstrate program results to various audiences. While they may have separate objectives, both summative and formative reviews can be completed as part of one program evaluation.

Table 1. Types of Evaluation

Evaluation Type	Focus	Purpose	Types
Formative	<ul style="list-style-type: none">• Program activities• Outputs• Short-term outcomes	<ul style="list-style-type: none">• Monitoring progress• Make mid-course corrections• Bring suggestions for improvement	<ul style="list-style-type: none">• Needs assessment• Process evaluation• Implementation Evaluation
Summative	<ul style="list-style-type: none">• Intermediate outcomes• Impact	<ul style="list-style-type: none">• Determine the value of the program.• Describe the quality & effectiveness of the program	<ul style="list-style-type: none">• Outcome evaluation• Impact evaluation• Cost-effectiveness analysis

Participant discussions and questions

- Evaluation of water quality trading programs may be difficult, particularly at the state level because they are likely a small portion of a much larger program.
- As so many water quality trading programs are still relatively new, there may be greater value in focusing on outputs and short-term outcomes (changes in knowledge, attitude, and skills) that let you assume that you are on the right track towards longer term outcomes.
- In addition to water quality improvements, there is also an evaluation and communication goal related to political and economic benefits associated with water quality trading as a flexible approach to achieving regulatory objectives.

Speakers

Sara Bobmeyer, Washington University

Presentation: The Evaluation Planning Process

The first step to initiating a program evaluation is to identify the program's stakeholders, those who are served by the program, involved in operations, funders, or those impacted by the outcomes of the program. Identifying a program's stakeholders helps to know which evaluation questions to ask, and how to communicate results. Engaging stakeholders early in the evaluation process increases awareness in the process, can reduce suspicion or fear, and increase the likelihood that the findings will be used and accepted.



Figure 1-1. The Planning Process.
Credit: Sarah Bobmeyer, Washington University

In addition to identifying stakeholders, an initial step in the planning process is to describe the program. This description includes the goals and objectives, as well as the need or problem statement of the program. Creating a logic model provides a picture of these critical elements of a program and how they work together.

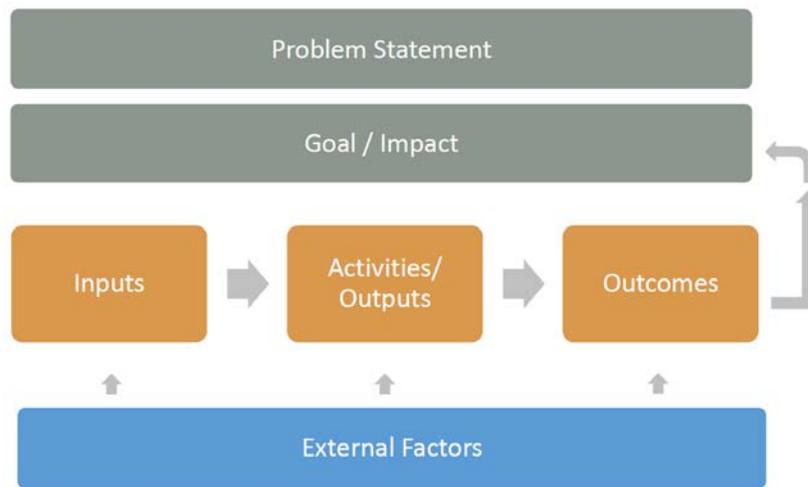


Figure 1-2. Logic Model.
Credit: Sarah Bobmeyer, Washington University

Elements of a logic model include:

- **Problem statement:** What the program seeks to solve.
- **Goals:** The intended aim or impact over the life of the program.
- **Inputs:** Resources such as staff, money, materials, equipment, facilities and volunteer time.
- **Activities:** Process, tools, events, technology, actions.
- **Outputs:** Direct products of program activities.
- **Outcomes:**
 - Short-term: First benefits or changes most closely influenced by program outputs.
 - Intermediate: Links initial to long-term outcomes, often changes in behavior.
 - Long-term: What a program desires to ultimately achieve.
- **External Factors:** Those factors outside the program’s process and control that are likely to influence the various elements of the logic model.

Evaluation questions should be focused by referring to the logic model and thinking about those elements that are likely to be of most interest to the stakeholders identified at the outset of the process. In selecting questions, program administrators should consider those whose findings are going to be most useful, relevant, feasible, and credible. Evaluation questions typically fall into one of four categories including process, value, outcome, and next steps.

Table 2. Types of Evaluation Questions

Evaluation Question Type	Example
Process	How well was the project designed and implemented?
Value/Worth	Did the value of the outcomes outweigh the value of the resources used?
Outcome	How valuable were the outcomes to participants?
Next Steps	How sustainable is the program over time?

To answer evaluation questions indicators are needed that represent the changes that show the progress a program is making toward achieving a specific outcome. Characteristics of good indicators include meaningful, direct, useful, reliable, and practical to collect. Data collection is the process and methods for acquiring the relevant information to calculate or populate the indicator. Specific methods will depend on the indicator and question that is being answered, common data collection techniques include interviews, focus groups, case studies, monitoring and observation, and literature review. Each of these data collection techniques has its own methodology, which is beyond the scope of this presentation. However, there are several considerations in selecting indicators and data collection methods such as the availability of data sources, timing, resources and credibility. As evaluation programs are ongoing, it is also important to consider how data is managed over time and to review the collection strategy periodically.

Participant discussions and questions

- An important consideration in designing an evaluation is the stage of the program's maturity which may include early stages of design, program development, early stages of program implementation, and established implementation. A program's maturity should be used to calibrate expectations for program performance (e.g., a program in the development stage is not likely to show measurable changes for in-stream water quality).
- There are likely to be different stakeholders at different stages of a program maturity. This is true both for stakeholders participating in a trading program, and for the audiences wanting/needing to see evaluation information.
- Getting stakeholders involved first and incorporating their input on the important evaluation questions will build buy-in or support for the evaluation.
- Examples of potential WQT evaluation question categories¹:
 - Stakeholder attitudes and why?
 - What are the location specific conditions conducive to water quality trading?
 - What would have happened without this program – counter factual?
- Indicators: Characteristics or changes that show progress toward a specific outcome. Should be meaningful, useful, reliable, and practical to collect.
- Data collection methods: The methods and procedures should match the questions and indicators: interviews, focus groups, case studies, monitoring systems, questionnaires.
 - Think about what you are going to do with it before you start – what question will it answer, how will it be managed.
 - Review data collection strategies periodically.
 - Disseminating and utilizing results: Who's your audience, why disseminate, what needs to be disseminated, how will you do it?

¹ See Appendix III for a list of potential water quality trading evaluation metrics identified by dialogue participants.

3. Scorecards and metrics for water quality trading

Speakers

Jessica Fox, Electric Power Research Institute
 Becca Madsen, Electric Power Research Institute

Presentation: Metrics, Benchmarking, and Scorecards

EPRI has developed a scorecard system by which individual electric power utilities can measure and communicate their sustainability programs, as well as benchmark this performance against other utilities. This scorecard or metrics of performance are part of a *road map* for identifying and advancing sustainability within the industry. The steps along the roadmap include 1) identifying the issues; 2) determine the maturity and goals; 3) measure performance; 4) benchmark; and 5) communicate results.

For the electric power industry issues related to sustainability are tied to three pillars of environmental, social, and economic characteristics. These characteristics are further categorized by maturity levels that range from awareness of the issue (level 1) to achieving sustainability (level 5).

The characteristics used to determine position and goals are then used to identify the appropriate metrics of sustainability. In a general search of relevant sustainability indicators, EPRI found 448 that could have been potentially selected. By focusing on the “right” metrics that capture the characteristics at each stage of maturity, the list was reduced to 78 metrics across the twelve issues originally identified.

Model Framework		GHGs	Affordability	Water Availability	Reliability
Maturity Levels	5	Achieving sustainability	<ul style="list-style-type: none"> Strategy is continuously improved and re-imagined Developing innovative solutions to new and future challenges Measured progress of metrics suggests sustainability is achieved Achieving benefits of shared value both internally and externally 		
	4	Optimizing for sustainability	<ul style="list-style-type: none"> Strategy is implemented and optimized, as necessary Actions and solutions now integrated across the business units Measuring widespread results Realizing early benefits of shared value 		
	3	Managing toward sustainability	<ul style="list-style-type: none"> Implementation of strategy is well underway Actions and solutions are deployed in relevant business units Beginning to achieve measurable results Initiating specific shared value activities 		
	2	Initiating sustainability	<ul style="list-style-type: none"> Developing vision, strategy and setting objectives Researching and testing possible actions or solutions Establishing metrics for measuring progress Planning for shared value 		
	1	Awareness of the issue	<ul style="list-style-type: none"> Acknowledgment that there is an issue deserving a strategy Compiling possible actions & solutions Collecting data to understand current performance Considering opportunities for shared value 		

Figure 1-2. EPRI Sample Maturity Model for the Electric Power Utility Industry

While it is important for an organization to evaluate its own performance, it is equally important to understand how performance measures against others within the industry. Benchmarking can be

challenging as it requires organizations to have similar context and measurement procedures. For water quality trading this could prove difficult because of the variety in watershed context. In addition, it can be hard to find a clear normalizing denominator that makes progress or program status comparable in different contexts.

In communicating results of sustainability programs, many companies use scorecards, which present metric results in graphic and visual formats. Great examples of corporate scorecards include Coca-Cola, Campbell's, and Target. A good example of an environmental scorecard is the *State of the Bay Report*, produced by the Chesapeake Bay Foundation².

In thinking about metrics for water quality, it is first important to understand what they are and what they are not. Metrics are a measure of performance in the material issue area (i.e., air and water), a measure of past performance (lagging metric), a baseline for setting goals, and a catalyst for internal dialogue. Metrics are not predictive of future performance (leading metric), a measure of whether or not a company is sustainable, and indicative of links between material issues (correlation, not causality).

No metric tells a complete story of the issue, but rather each metric is likely to serve a different purpose and speak to a different audience. Similarly, communicating or reporting results is a continuum from narrative to quantitative and equates to whether the purpose of reporting is for internal, external or external auditing.

EPRI is currently developing an audit/scorecard for water quality trading as part of the Ohio River Basin program. The drivers of this process are to evaluate the broad goals for: sustainability of water quality trading, adaptive management, grant reporting, and transparency for stakeholders. The audit scorecard will utilize data from their online credit registry, accounting data, reports on project team labor hours, and survey data. EPRI is exploring the use of the National Network's *Options and Considerations Guide* as a basis for a scorecard framework³. Advantages of this approach is that the Network's guide is established and has consensus amongst most program administrators (as to program characteristics). However, many of the elements are focused on the "how", are not quantitative, not benchmark-able, and generally do not address the social and economic aspects of water quality trading.

With respects to the ability to benchmark across programs, it will be challenging to develop a normalized denominator, ensure the same data quality across programs and compare across watersheds with significantly different characteristics.

Participant discussion and questions

- There is/should be a bigger process going on to evaluate the water quality program, of which trading is a small part. So can we use that process to evaluate water quality trading rather than develop a standalone process? Or do so with the addition of a couple more metrics?
- The question of what is getting evaluated, and at what scale was an important theme throughout the dialogue. Participants recognized that trading is one tool connected to a broader set of strategies for meeting water quality goals (e.g., wastewater technology improvement, voluntary nonpoint reduction strategies through the Farm Bill, etc.). In instances where point

² State of the Bay Report. Available at: <http://www.cbf.org/about-the-bay/state-of-the-bay-report-2014>

³ Options and Consideration Guide. Available at: <http://nnwqt.org/products/>

sources represent a small percentage of the overall loading (e.g., ~30%), broader evaluation of water quality strategies may be more appropriate. This raises the question as to whether trading evaluations fit better within broader TMDL implementation evaluations or state nutrient strategies. This may be the case, but stakeholders have expectations for seeing results of trading programs themselves, so the question remains how best to calibrate expectations given the role trading might play in a watershed?

4. Identifying common metrics roundtable discussion

Facilitator

Jessica Fox, Electric Power Research Institute

Dialogue participants engaged in a brainstorming session to identify potential performance metrics for water quality trading programs. Following a similar process to that described in the presentation provided by EPRI, groups identified characteristics of water quality trading programs and quantitative metrics. At the end of the brainstorming session, participants were asked to rank the metrics and characteristics that they felt were the “right” metrics for a water quality trading scorecard. The following are the top ten metrics from each category with the full list and votes shown in Appendix I.

Table 3. Water Quality Trading Program Characteristics

Characteristics	Votes
Presence of Market demand drivers	11
Goals of the crediting project are articulated	10
Credit quantification methods, type of water quality model, attenuation factors.	10
Stakeholder and participant confidence in the market	7
Method for expressing credit project baseline	7
Clear articulation of verification standards	6
Program maturity	5
Quality of data,	3
Buyer/seller satisfaction in the trading process	3
Opportunity for public involvement and input in the development and execution of trades	3

Table 4. Metrics of Water Quality Trading Program Performance

Metrics	Votes
Cost savings/avoided costs	20
Acres/miles of conservation/restoration by type of credit project/BMP	14
Credits sold/utilization by type of credit (e.g., TN, TP)	13
Ancillary benefits (other pollutant reductions, environmental benefit, job creation)	6
Dollars invested by type (public-state/federal, private)	6
Percentage of pollutant loading addressed by water quality trading	5
Acres meeting load allocations	3
Innovation (new BMPs, rate of adoption)	3
BMP performance	3

5. Communicating and incorporating results into decision making

Speakers

Kimberly Caringer, Tahoe Regional Planning Agency

Presentation: Tahoe Regional Planning Agency: Measuring Progress

The Tahoe Regional Planning Agency (TRPA) was formed in 1969 through a Bi-State Compact (California and Nevada) to protect and restore the environment of the Lake Tahoe Basin through land-use planning, enforcement of regulations, and restoration. As part of the compact, the TRPA is responsible for meeting environmental threshold carrying capacities for nine categories – air quality, water quality, fisheries, soil conservation, wildlife, vegetation, scenic resources, recreation, and noise. There are three types of threshold standards that have been adopted including:

- **Numeric standards** – a numerical target for an indicator
- **Management standards** – adopted best practice, regulation norm or requirement
- **Policy statements** – provide direction to the Agency in developing the Regional Plan.

Of the nine threshold categories there are over 150 standards intended to be measured and tracked as indicators of the Lake Tahoe environment's health. It is intended that the attainment of these threshold standards is achieved through the implementation of the Regional Plan, permitting and regulatory compliance, and restoration projects and programs.

The Environmental Improvement Program (EIP) is the restoration element of the TRPA's mission and is currently structured into six program areas:

- Watersheds, habitat, and water quality
- Air quality and transportation
- Forest management
- Recreation and scenic resources
- Applied science
- Program support

Although TRPA is the lead program administrator, the EIP program is a partnership of over 40 agencies and has received a significant amount of state and federal funding to implement restoration in the Lake Tahoe Basin. As with other performance programs, the EIP program has been challenged to measure progress and program outcomes and report these to the public, stakeholders, and funders.

The reporting process of both environmental thresholds and the EIP program have evolved over time. In its latest iteration, the *Environmental Threshold Report* now provides a status and determination trend and is much more visual in its presentation of information in order to reach a broader audience⁴. Specifically for the EIP program, the Agency has developed an on-line project tracking database, which has helped to standardize reporting information and provide the public with an easy to access resource. The EIP project tracker provides information on individual projects, as well as reporting on 33 performance measures across the nine threshold categories. The EIP-tracker has provided valuable information on program performance and has increased the accountability of the program.

⁴ Environmental Threshold Report. Available at: <http://www.trpa.org/regional-plan/threshold-evaluation/>

There are some lessons that the TRPA has learned in the evolution of its performance measurement program and considerations to share with others considering developing a program. First, TRPA has the mandate to measure progress and therefore has dedicated resources and staff to meeting that mandate. Seeking external funding for measurement and evaluation is difficult as most sources focus on project implementation and do not include funds for ongoing monitoring. In establishing measures, there is a continual need to review metrics, the methodologies of data gathering practices and monitoring protocols that needs to be considered. Finally, while increased access to measurement data improves transparency and makes reporting more efficient, it can also create new questions or additional requests for information that may not have been anticipated. Therefore documentation of the data collection and measurement process becomes more and more important.

Speakers

Neil Crescenti, Willamette Partnership - Facilitator

Bob Rose, Policy Analyst- EPA Office of Water

Jason Keppler, Watershed Implementation Program Manager - Maryland Office of Resource Conservation

Brian Schlauderaff, Environmental Group Manager – Pennsylvania Department of Environmental Protection.

Panel Discussion

Neil Crescenti:

In 2011 EPA conducted an evaluation of all nutrient trading programs within the Chesapeake Bay watershed. Although the evaluation is unique in that all of the state program evaluated share a common TMDL, it provides the only example of where a program evaluation has occurred across programs (although programs were not compared directly). Each state program was evaluated based on elements identified as common across trading programs, which included program type-trading vs offset, authority, baseline, minimum controls, eligibility, credit calculation and verification, safeguards, certification and enforceability, accountability and tracking, nutrient impairment segments, credit banking and growth. Each program element was rated as

- Necessary measures not in place;
- Partial (e.g., Legislation drafted or steps have been taken to implement but not fully in place, some details still to be determined but framework is largely established);
- Jurisdiction has measures in place and in effect;
- Jurisdiction is evaluating the issue but has taken no formal measures to implement anything specifically; or
- Not applicable.

The evaluation process was a collaborative dialogue with each program's administrative agency and the final evaluations were circulated for public comment. Each state was provided a summary report which outlined the evaluations results and recommended changes.

Bob Rose:

The 2011 program review was part of EPA's oversight responsibilities under the Clean Water Act and ongoing review of the Chesapeake Bay TMDL which tries to determine whether the load reductions can be reasonably met. The evaluation was formative and involved iterative dialogue with each of the state

program administrators on a range of issues, such as baseline. Based on the assessment, EPA has asked each state to implement changes, which are categorized into tier 1 recommendations- those that require statutory or regulatory changes or tier 2-programmatic changes.

Jason Keppler:

The Maryland trading program trades nitrogen, phosphorus, and sediment. The program was developed in three phases and has seen relatively few trades. In 2008 the program started with point-to-point trades, but no longer relevant because most treatment plants have been upgraded through funding provided by Maryland's "flush tax", which collects fees for all properties regardless of whether they are on sewer or septic. Phase 2, which is nonpoint to point trading has not seen any trades. A third phase has begun where those neighboring states in the Chesapeake Bay that comply with Maryland's trading protocols are eligible to sell credits in Maryland.

With regards to the evaluation, since EPA conducted it in an ongoing dialogue with the jurisdictions, it was easy for the State to know what the issues were and provide input on the process. Since the evaluation was completed, the Maryland Department of the Environment has taken steps to complete the recommended actions of the summary report.

Brian Schlauderaff:

Pennsylvania's nutrient trading program was established in 2006 prior to the current Chesapeake Bay TMDL and allows trading of nitrogen and phosphorus. The nutrient trading program has seen a number of transactions, mostly point-to-point. The state administers the program and provides transaction information to the public through its website⁵. In 2010 DEP published nutrient trading program regulations, but in 2014 EPA objected to several permits under these regulations because of concerns that DEP had not made quantitative demonstration that these requirements achieve the load allocations for agricultural sources under the Chesapeake Bay TMDL. Pennsylvania has already taken additional steps to update the Code to meet the legal requirements of the TMDL and allow for trading under NPDES permits.

⁵ <http://www.dep.pa.gov/Business/Water/PointNonPointMgmt/NutrientTrading/Pages/default.aspx>

6. Final thoughts from the Spring 2016 Dialogue

The Spring 2016 Dialogue brought together a variety of water quality trading stakeholders, program administrators, practitioners, and those from the broader program evaluation field. Over the course of the two days the group had great conversations around what program evaluation and performance measurement are, what they can mean for water quality trading programs, and how they could potentially be adopted by trading programs. Based on the information presented and discussions that followed, there were several points that the group identified as important considerations for advancing program evaluation and performance measurement for water quality trading.

The business case for program evaluation and performance measurement. Most water quality trading programs are small and administered as part of larger water quality management programs. In addition, most programs are still new in terms of transactions experienced and years of operation. Yet there is also the continued call for the field of water quality trading to demonstrate success towards the goals and benefits that are often used to justify it. Program evaluation and performance measurement have proven themselves to be valuable tools in other arenas both within natural resource management and outside in fields such as healthcare and education. Program evaluation and performance measurement have not been developed in great detail for water quality trading programs to date. However, through the presentations and discussions held, National Network participants expressed that program evaluation and performance measurement, if designed appropriately and efficiently, can provide significant value to individual programs and the field as a whole.

Program maturity is very meaningful when it comes to evaluation. The information presented by Jessica Fox of EPRI on maturity models resonated with many dialogue participants. Many water quality trading programs are either in the development or early stages of implementation (in terms of the number of transactions). In considering developing an evaluation and measurement program, the phase or stage of maturity of the program is seen as a critical component not only to selecting the right evaluation questions and performance metrics, but to also setting appropriate expectations. Building example maturity models specific to water quality trading was identified during the dialogue as an area for further development that would help advance program evaluation for water quality trading.

Comparing across programs will be challenging. Benchmarking can provide additional information as to a program's performance. However, there are significant challenges that require further consideration in developing benchmarking for water quality trading programs. These include variation across watersheds, quantification metrics, and the program driver (TMDL, permit, voluntary). As water quality trading programs are still being established, evolving, and growing, the time does not appear ripe to explore benchmarking further. Rather program evaluations should be formative – focusing on improving program performance and improvement.

Continuing the Conversation

At the conclusion of the dialogue, participants were asked how the National Network could help to advance program evaluation and performance measurement for water quality trading. A challenge to this question is that many programs are still in the development stage or are only a single strategy in a much larger program or suite of strategies to achieve water quality goals. Participants recognized the value of evaluating programs and measuring performance. It was noted that the convening report provided important foundational information that integrating information presented at the workshop could be a valuable resource for those program administrators interested in exploring the topic further.

A theme that resonated throughout the two days was the importance of recognizing a program's stage of development or maturity. Participants felt that although there isn't a lot of program experience to draw from, that the development of a water quality specific maturity model would benefit the discussion of program evaluation and performance measurement. An identified product that the Network could develop out of the dialogue was a framework based on a program evaluation logic model similar to what was presented that would also define the stages of maturity likely to be important water quality trading programs.

Next steps following the Spring 2016 Dialogue will be to integrate the information presented at the workshop along with the input collected during the discussions into the convening report. In addition to the foundational information on program evaluation and performance measurement, a sample logic-maturity model will be developed along with a sample suite of metrics for performance evaluation specific to water quality trading. These examples will be intended to serve as a template framework from which water quality trading administrators can build their own more detailed evaluation and performance measurement program.

Several participants from the dialogue have volunteered to assist in the development of the logic-maturity model and to provide feedback on the evaluation framework. The report has been completed and vetted, it will be made available on the National Network's website at www.NNWQT.org.

APPENDIX I: 2016 Spring Dialogue Participants

Participant Name	Organization	E-mail
Julia Anastasio	Association of Clean Water Administrators	janastasio@acwa-us.org
Todd Blank	Missouri Department of Natural Resources	DONNA.BELFIORE@DNR.MO.GOV
Sarah Bobmeyer	Evaluation Center-Brown School of Social Work-Washington University	sbobmeyer@wustl.edu
Joe Boland	Missouri Environmental Improvement and Energy Resources Authority	joe.boland@eiera.mo.gov
Brian Brandt	American Farmland Trust	bbrandt@farmland.org
Courtney Brown	Oregon Department of Environmental Quality	brown.courtney@deq.state.or.us
Kimberly Caringer	Tahoe Regional Planning Agency	kcaringer@trpa.org
Kari Cohen	USDA Natural Resources Conservation Service	kari.cohen@wdc.usda.gov
Bobby Cochran	Willamette Partnership	cochran@willamettepartnership.org
Jeff Corbin	Restoration Systems	jcorbin@restorationsystems.com
Neil Crescenti	Willamette Partnership	crescenti@willamettepartnership.org
Clay Detlefsen	National Milk Producers Federation	cdetlefsen@nmpf.org
Alex Echols	Ecosystem Services Exchange	alex@ecoexch.com
Jessica Fox	Electric Power Research Institute	jfox@epri.com
Coleman Garrison	National Association of Conservation Districts	coleman-garrison@nacdn.org
Allan Gates	Mitchell Williams Selig Gates & Woodyard	agates@mvlaw.com
Jason Keppler	Maryland Department of Agriculture	jason.keppler@maryland.gov
David Kuszmar	California North Coast Regional Water Quality Control Board	david.kuszmar@waterboards.ca.gov
Lance Holloway	Idaho Department of Environmental Quality	lance.holloway@deq.idaho.gov
Elizabeth Hubertz	Interdisciplinary Environmental Clinic, Washington University School of Law	ejhubertz@wustl.edu
Kolby Jones	Ecosystem Services Exchange	esther@ecoexch.com
James Klang	Kieser & Associates, LLC	jklang@kieser-associates.com

Harbans Lal	USDA Natural Resources Conservation Service	Harbans.Lal@por.usda.gov
Amelia Letnes	US Environmental Protection Agency	aletnes@gmail.com
Alicia Claire Lloyd	Missouri Coalition for the Environment	alloyd@moenviron.org
Becca Madsen	Electric Power Research Institute	bmadsen@epri.com
Mark Patrick McGuire	Association of Clean Water Administrators	mmpmcguire@acwa-us.org
MARK OSBORN	Missouri Department of Natural Resources	mark.osborn@dnr.mo.gov
Susan Payne	Maryland Department of Agriculture	susan.payne@maryland.gov
William Payne	City of Chattanooga	ebishop@chattanooga.gov
Bob Rose	US Environmental Protection Agency	rose.bob@epa.gov
Carrie Sannemen	Willamette Partnership	sanneman@willamettepartnership.org
Katherine Sever	Colorado State University	kat.sever@colostate.edu
Brian Schlauderaff	Pennsylvania Bureau of Clean Water	bschlauder@pa.gov
William Schleizer	Delta Institute	wschleizer@delta-institute.org
Adam Schnieders	Iowa Department of Natural Resources	adam.schnieders@dnr.iowa.gov
Mark Shumar	Idaho Department of Environmental Quality	mark.shumar@deq.idaho.gov
Mackenzie Staub	Evaluation Center-Brown School of Social Work-Washington University	mstaub@wustl.edu
Kristina Surfus	National Association of Clean Water Administrators	ksurfus@nacwa.org
George Houtven Van	RTI International	gvh@rti.org
Sara Walker	World Resources Institute	swalker@wri.org
Sunny Wellesley	US Environmental Protection Agency	wellesley.sunny@epa.gov
Nicole Zacharda	Great Lakes Commission	nzacharda@glc.org

Appendix II: Speaker Bios

Jessica Fox is a Program Manager at the Electric Power Research Institute (EPRI) where she leads efforts on water quality trading, ecosystem services, sustainability and related work. Ms. Fox created and manages the Ohio River Basin Water Quality Trading Project, which is the world's first interstate trading program for nutrients. Ms. Fox has hosted and facilitated numerous workshops and has authored and coauthored over 40 reports and publications. A scientific background in conservation genetics, sustainable development, sustainable agriculture, human genome mapping and sequencing, and molecular biology provides Ms. Fox a strong foundation for understanding complex environmental challenges and innovative methods for addressing them.

Becca Madsen is a Project Manager at the Electric Power Research Institute (EPRI) where she executes technical projects and research related to sustainability, endangered species, water quality trading, other environmental markets, and ecosystem services as they pertain to the electric power industry and its customers. Ms. Madsen was a Peace Corps volunteer, a Doris Duke Conservation Fellow, and a Presidential Management Fellow with the US Forest Service.

Sarah Bobmeyer is the Associate Director of the Brown School Evaluation Center. Sarah provides leadership, strategy, and operational oversight for the Center. She manages the Center's daily operations, supervises and mentors Center team members, leads evaluation projects and proposal development, helps identify and develop internal and external partnerships, and assists the Director in implementing and monitoring the Center's strategic plan. Sarah has extensive experience with evaluation planning and implementation; development of evaluation protocols; development and management of data collection tools and systems; data analysis; report, paper, and presentation development that promote utilization of evaluation findings; and design and implementation of evaluation capacity-building activities, such as training institutes, workshops, and one-on-one technical assistance.

Mackenzie Staub is a Project Coordinator at the Brown School Evaluation Center. Mackenzie helps to manage several of the Center's projects and contributes to daily operations, working closely with both Center leadership and student research assistants. Mackenzie has experience in both local and international program evaluation and her primary interests include community development, global health, and health promotion. Most recently, Mackenzie worked with the Orphans and Vulnerable Children Unit at the United States Agency for International Development (USAID) in South Africa to help develop program strengthening and evaluation recommendations for a local child protection agency. Mackenzie recently graduated from the dual masters program in social work and public health at Washington University.

Kimberly Caringer is the Environmental Improvement Division Manager at the Tahoe Regional Planning Agency where she is responsible for the oversight, coordination and reporting of the Lake Tahoe Environmental Improvement Program. Prior to moving to Tahoe she was an environmental mediator with the U.S. Institute for Environmental Conflict Resolution in Tucson, Arizona. Kimberly has over 10 years of experience in environmental policy, collaboration, and implementing environmental restoration projects through partnerships.

APPENDIX III: Water Quality Trading Performance Metric Brainstorm Results

The following program characteristics (presence/absence) and program performance metrics were identified during a brainstorming session. Dialogue participants followed up by allocating 5 “votes” to the metrics they felt were most important. The table below reflects the metrics, ordered by the number of votes they received.

Trading Program Characteristics	Votes
Market Demand Driver	11
Quantification methods (attenuation, type of model)	10
Goal of project	10
Baseline	7
Confidence in market	7
Verification standards (clear, who signs	6
Program maturity level	5
Opportunity for public involvement	3
Quality of data	3
Buyer/seller satisfaction	3
Transparency of projects (location, buyer, seller)	2
Urban/rural interactions	2
Benefits to community	2
Ancillary benefits	2
Trust	2
Environmental justice	2
Rewards for early adopters/incentive for laggards	2
Reduced conflicts	2
Linkability to other programs	2
Public Acceptance	2
Winners and losers	1
Stakeholder satisfaction	1
Change in attitudes of nonpoint	1

Replicability	0
Ratios	0
Comparing to other programs	0
Thorough program description	0
Full permit that includes trading	0
Presence/absence, along a spectrum	0
Are adaptive management procedures in place	0
Number of BMP quality standards (including practice standards, measurement requirements, ect.)	0
Regional water quality monitoring program	0
Who does what – administration	0
Partnerships/networks – quantity/quality	0
Program sustainability	0
Changing perception of value of good stuff and property	0
Attenuation	0
Volume of data	0

Program Performance Metrics	Vote
Cost savings/avoided costs vs alternative opportunity cost	20
Acres/miles of conservation	14
Credits sold/utilization by type (of credit)	13
Dollars invested by investor (public –feds/state, private)	6
Ancillary benefits- other pollutants, environmental benefits, job creation	6
Percentage of pollutant loading addressed by water quality trading	5
BMP performance by type	3
Pounds of reduction (To target TMDL, by County, HUC, Source, term/permanent)	3
Acres meeting load allocations	3
Innovation (new BMPs, increased adoption rate)	3
Transaction costs	2

Price of credits	2
Uncertainty of quantification/trading ratio	2
Number of sellers/contracts	1
Average (when there is volume)	1
Time savings to meet permit limit	1
Biological improvements in-stream benefits	1
Compliance rate	1
Number of projects	0
Number of buyers	0
Number of buyer/seller inquiries/ profile, diversity	0
Population growth accommodated	0
Number of 303(d) waters delisted	0
Benefits beyond expected reductions	0
Annual credit demand/supply within watershed.	0
Project deficiency rate/percentage success/failure	0
Credit Demand -registered, generated	0
Monitoring activity	0
Contracts that continue beyond project life	0
Trust in program	0
Partnerships	0
Outside dollars leveraged	0