THE TMDL PROGRAM RESULTS ANALYSIS PROJECT: MATCHING RESULTS MEASURES WITH PROGRAM EXPECTATIONS

Douglas J. Norton*[1], Dwight Atkinson[1], Valentina Cabrera-Stagno[1], Bruce Cleland[2], Sarah Furtak[1], Cary McElhinney[3] and Eric Monschein[1]
(US Environmental Protection Agency, [1]: Office of Water, Washington DC; [2]: Region 10, Seattle WA; [3]: Region 5, Chicago, IL)
*US Environmental Protection Agency, Office of Water (4503T)
1200 Pennsylvania Ave NW
Washington DC 20460

ABSTRACT

The evaluation of environmental results from federal programs has increased significantly in recent years, both due to government-wide review requirements and agencies' own interest in identifying potential program improvements. The USEPA's Clean Water Act TMDL program has undertaken a self-evaluation of TMDL environmental results, program milestones, and explanatory factors that can collectively improve understanding of program results, provide a feedback loop for program managers, and lead to results-based program improvements. An effective approach to results analysis initially required the clarification of TMDL program goals and expectations, followed by identification of measurable results concepts that relate directly to progress toward program goals. We reviewed several recent analyses of the TMDL program and similar restoration programs to identify lines of evidence that showed possible results patterns that could be monitored with available databases. From the available evidence we developed evaluation metrics of three types. *Response measures* relate to environmental outcomes – interim improvements and full recovery in waters that have been addressed by the TMDL program. *Programmatic measures* track key milestones in the several-step TMDL program process from TMDL development and implementation through recovery. *Explanatory measures* assess linkages between environmental response results and potential causes; these measures are crucial for revealing why certain results are occurring and focusing attention on specific causes that should be addressed in program improvements. As this five-year effort accumulates new insights into TMDL results and their driving factors, the TMDL results analysis project will provide insights and evidence for improving the TMDL program. This paper provides the conceptual overview and unifying context for several papers comprising a special session on TMDL results analysis at the WEF TMDL 2007 conference.

KEYWORDS

results analysis, assessment, program evaluation, effectiveness monitoring, indicators, TMDL, 303(d), USEPA

INTRODUCTION

The evaluation of environmental results from federal programs has increased significantly in recent years. Strategic planning at the US Environmental Protection Agency (USEPA) relies on tracking results measures, and program evaluation by the Office of Management and Budget (OMB) emphasizes environmental outcomes instead of purely programmatic milestones. The USEPA's Watershed Branch is the national program manager for Clean Water Act Section 303(d) (FWPCA 1972) and in this role it oversees state listing of impaired waters and the development of TMDLs. TMDL program results analysis, one of the Branch's four main work themes, focuses on assessing program effectiveness and providing insights for improvement. The purpose of this paper is to summarize the conceptual basis for an ongoing, multi-year Watershed Branch project in program results analysis: the programmatic setting and expectations, the state of our current knowledge about TMDL program results, a gap analysis of the information we need to know, and the analytical approach for increasing our understanding of results and ultimately improving the TMDL program. Further, we address specific activities designed to carry out the TMDL program results analysis project and our progress to date. This paper provides the conceptual overview and unifying context for several papers comprising a special session on TMDL results analysis at the WEF TMDL 2007 conference (Benham et al. 2007, Cabrera-Stagno et al. 2007, Cleland 2007, Monschein and Mann 2007, Wickham and Norton 2007).

The meaning of **_TMDL Program Results Analysis_** as the focus of this project is complex. We include in the scope of the **_TMDL Program_** the Clean Water Act (CWA) Section 303(d) impaired waters listing process, TMDL development and approval, USEPA oversight of approved TMDL alternatives (e.g., Category 4b projects), implementation of management practices through TMDLs, and post-project evaluation of TMDL outcomes. **_Results_** primarily means environmental outcomes (e.g., the recovery of formerly impaired waters) but also includes many causally related factors, including programmatic milestones, that are associated with the influence of the TMDL program on the environment; without linking the probable causes to the outcomes observed, programs would be poorly equipped to make improvements. **_Results Analysis_** involves assessing measures of environmental outcomes, their probable causes, and program progress in order to generate the insights needed to improve our programs.
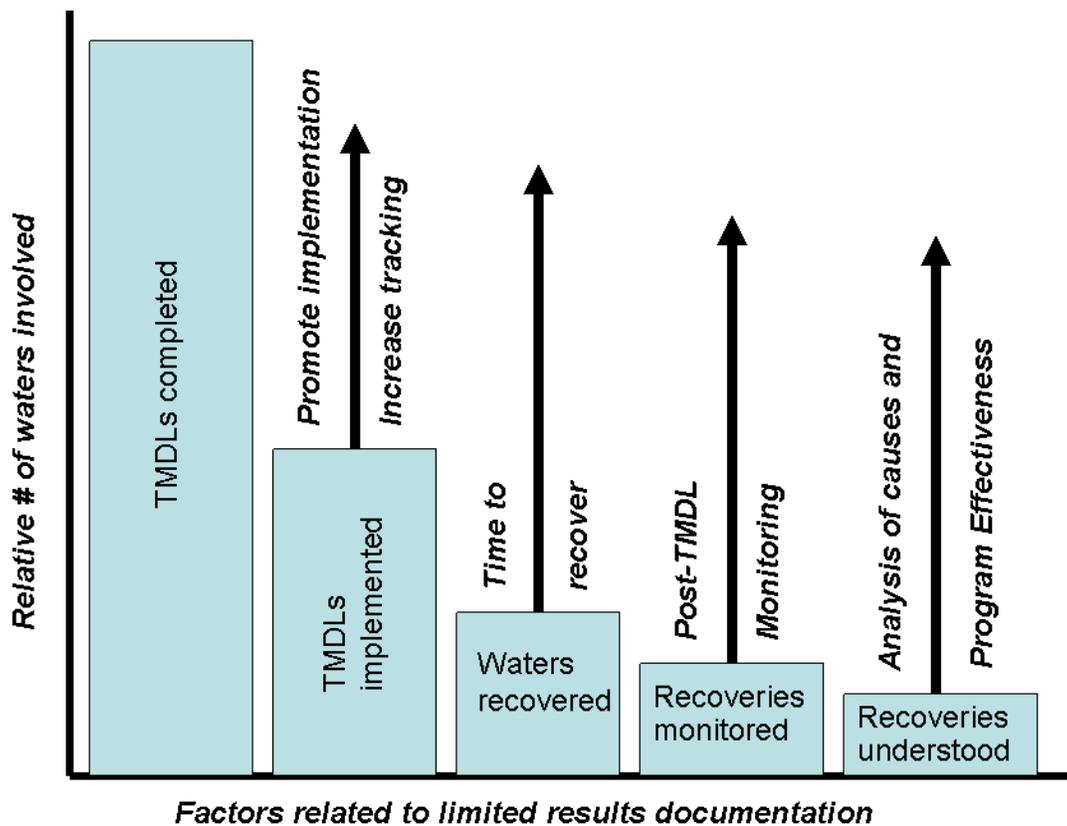
Background: the Need for Results Analysis and Main Goals of this Effort
The rate of development of TMDLs has increased markedly in recent years and over 24,000 TMDLs have now been completed (USEPA 2007). Numerous impairment characteristics have been well documented in USEPA and state databases over the years. Yet, over the TMDL program's history, it has been difficult to analyze results in terms of the changes in impaired waters attributed to the TMDL program. Several factors (see Figure 1) explain why. Mainly, the development of large numbers of TMDLs is recent; over half of current TMDLs are less than three years old. Large numbers of TMDLs may have been developed recently, but far fewer have yet been fully implemented. Fewer still have had enough time for implemented practices to achieve a recovery that can often take several years or even decades. Additional factors include limited post-TMDL monitoring information necessary to document results, limited analysis of the causal factors related to recovery, and the past tendency of USEPA to self-evaluate its programs using programmatic milestones (e.g., number of impaired waters listed, number of

TMDLs completed) instead of desired environmental outcomes such as waters reattaining Water Quality Standards (WQS). Whereas these programmatic milestones were consistent with the extent of USEPA regulatory authority – for example, USEPA oversees and approves TMDL development by states but cannot require their implementation – these types of measures fell short of providing insights about the environmental results attributable to the program.

With insights on actual results thus limited, past changes meant to improve the TMDL program have been most influenced by a mix of political issues, court cases, and best professional judgment. By addressing the gaps in results information, a greater investment in results analysis is intended to contribute to making TMDL program improvements based on evidence of environmental results and their probable causes rather than on limited data.

Figure 1 - Factors related to limited understanding of final TMDL outcomes despite the abundance of TMDLs completed (conceptual - bar heights are not to scale). Labeled arrows indicate what can be done to move increasing numbers of waters (i.e. raise the bar height) toward the right end.

As the USEPA's national program manager for TMDLs, the Watershed Branch is assessing TMDL program results and incrementally applying what we learn to improve our program. The goals of the project are:

- to improve the state of our understanding about how the TMDL program and closely related activities (e.g. restoration projects) lead to varying patterns of results;
- explain these patterns and where possible their causes; and
- help improve the TMDL program by applying these findings at state, regional and national program levels through technical, programmatic and policy actions.

Many ongoing USEPA activities are relevant to results analysis and can contribute to meeting these goals as well as benefit in turn from this project's findings. Among these is USEPA Office of Water (OW) strategic planning (USEPA 2006). This activity is the primary source of organized, national reporting and tracking of TMDL program output and outcome measures submitted periodically by state water programs. OW strategic planning shares with results analysis a focus on environmental outcomes and program milestones, robust data analysis and tracking of specific results measures, and an underlying purpose of strengthening programs by providing new results understanding. Both activities also provide new facts and figures for communicating EPA programs' successes to the public and decision makers, and offer a venue for integrating cross-program efforts to track results and propose results-based improvements. But these efforts also differ in important ways, one being the results analysis project's focus on the TMDL program alone versus strategic planning's focus on the entire surface water protection program. Moreover, strategic planning's relatively few measures require data of consistent, national coverage, whereas the TMDL results analysis project will gain its insights from a broader assemblage of results metrics at statewide or regional-scale along with the national datasets and measures. The most significant difference is that results analysis will address the probable causes of the results we find along with environmental responses and program milestones. At this point, OW strategic planning measures are currently limited to program milestones and environmental responses without looking at causes. Assessing causal linkages is crucial for translating results information into program improvements.

Although the numbers of improving and restored waters will continue to grow as more implemented TMDLs mature, USEPA currently faces a shortage of site-specific environmental results documentation. Limited data may constrain this project's ultimate purpose until detailed documentation of partially to fully recovered waters numbers in the hundreds or thousands. Achieving sufficient documentation of TMDL program outcomes is likely to take more time for the reasons already discussed. Accordingly, we plan to build up the information on successes in restoring impaired waters using a number of techniques in addition to improved documentation of TMDL outcomes.

It would be shortsighted to assume that insights on restoring impaired waters that could improve the TMDL program can come only from studying existing TMDL outcomes. Three primary sources for such insights include:
- empirical evidence from improving or recovered waters;
- expert judgment and experience of TMDL practitioners; and
- published literature addressing TMDLs or TMDL-like projects.

Thus, we also plan to examine additional literature and TMDL-like datasets, such as the documented results of watershed-based CWA Section 319 non-point source restoration and Category 4b projects, or studies of Best Management Practice (BMP) effectiveness that are typical of TMDL practices but may not have been connected to an actual TMDL. Including supplementary TMDL-like data can increase the numbers of waters under study, which should particularly strengthen our analysis of probable causes. Another supplementary source of results insight could come from practitioners' accumulated knowledge in TMDL and similar programs. Our broad approach is appropriate because there are many types of results that are expected of,
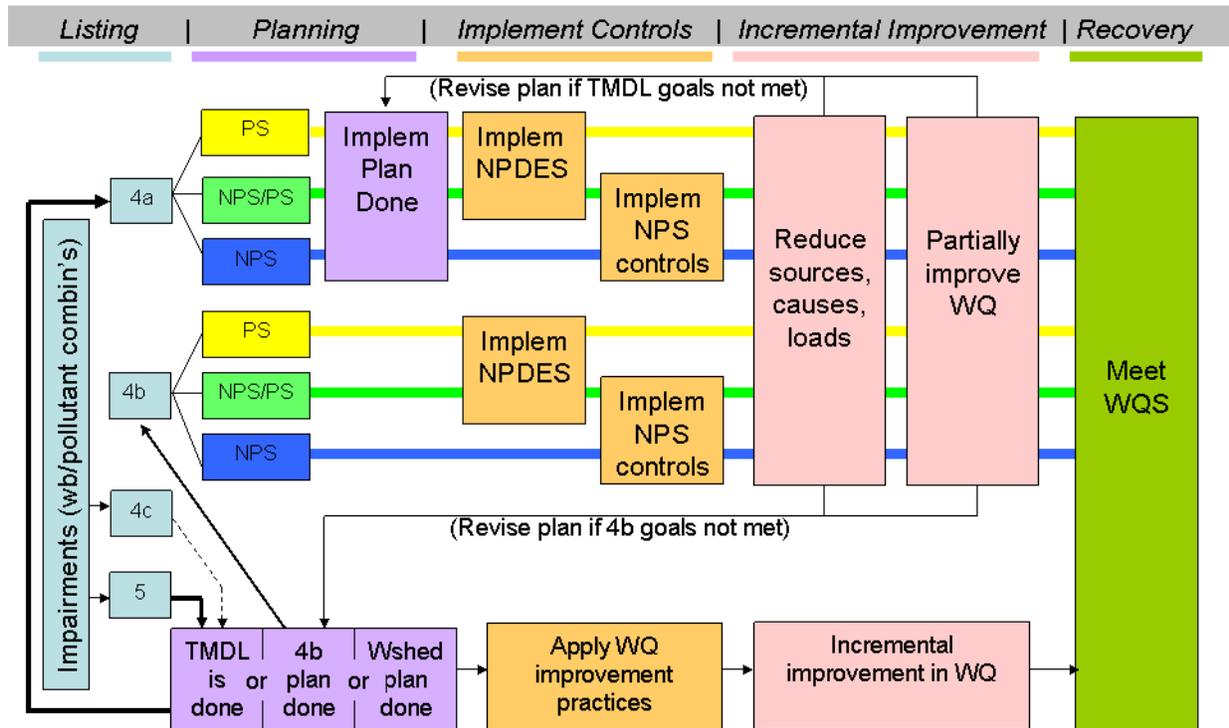
and brought about by, the TMDL program. Further, the public and our decision makers expect us to be able to explain the causes of these results.

METHODOLOGY

The "TMDL Program Pipeline" – a Conceptual Framework for Results Analysis
Program analysis, like mechanical diagnosis and repair, profits from a good diagram of component parts and their interrelationships. A flow diagram of the TMDL program process is useful for visualizing specific program steps and stages of impaired waters recovery that are at once critical milestones of the TMDL process and also amenable to measurement and evaluation. Figure 2 summarizes the stages that an impaired water (i.e., a single waterbody segment-pollutant combination) passes through in traversing the TMDL process from start to finish. The upper section (gray boxes) of the figure represent five general stages of the process, and the more complex colored boxes beneath each of the five stages show more details and interrelationships within the process.

Figure 2 – The TMDL Program Pipeline. General stages (gray boxes along top) are passed through in varying ways by different categories of impaired waters (colored boxes along left to right pathways below) en route to full recovery (4a: TMDLs are done; 4b: other controls expected to meet WQS; 4c: impairment not caused by a pollutant (TMDL optional); 5: impaired and TMDL is needed, will move to 4a after TMDL is done.)

In its simplest form, the TMDL pipeline consists of listing impaired waters (303(d) process), planning (development of a TMDL and an implementation plan, or a 4b alternative plan), application of controls (NPDES permits, or non-point source BMPs, or both), incremental

improvement in condition (progress short of meeting WQS), and recovery (meeting WQS). Whereas this simple sequence is easy to follow, it does not capture the complexities of the different pathways through the TMDL process that are followed by point source-impaired, NPS-impaired, and mixed waters, and in the TMDL alternative, category 4b process – hence the added detail in the colored boxes and arrows. This pipeline diagram serves as a central organizing device for selecting and assessing measures that relate to specific results as well as to key parts of the TMDL program (i.e., boxes in the pipeline).

A Generic Results Analysis Approach
The complexity of Figure 2 indicates that results analysis could pursue many different measures of multiple parts of the program -- more than funding, or information reporting, will likely support. Our analytical approach for results assessment includes four basic steps (see Table 1):

1. Identify expected program results
- Determine what results are expected of the program and examine the state of program knowledge about them. Program expectations provide the basis for initially identifying and evaluating program results. Sources of expectations broadly include the Clean Water Act itself along with perspectives of past TMDL program reviews, public opinion, states' expectations, and views of TMDL program staff and managers. Expectations often go beyond the relatively narrow 'letter of the law' and include such elements as expectations for public outreach, data management and program tracking. Expectations from the TMDL program implicit in external reviews and public comments, for example, go far beyond completing TMDLs and include achieving water quality improvements, detailed database management, public outreach, and technical knowledge of what causes a variety of outcomes. The fact that the TMDL program operates in a thoroughly integrated fashion with many other elements of CWA surface waters protection (e.g. permits, non-point source control, monitoring) also tends to blur perceptions of what the TMDL program (versus the broader CWA surface waters protection program) is meant to accomplish, and results in higher expectations from the TMDL program than its mandate includes.

Table 1 - the basic steps of results analysis
1. Identify expected program results
2. Identify appropriate program measures
3. Carry out measures analysis
4. Assess implications for program

- Conduct a gap analysis of results information needs. After identifying all program results of concern, it is evident that we know far less about achieving some types of results than others. Setting the scope of a specific results assessment, then, involves selecting the type of results that a given assessment will target and setting a goal of increased understanding about those results. This planning step must verify what information is now accessible, and any key information gaps that need to be addressed to fully inform program improvements. One significant constraint is the substantial time and effort it would take to track and report on very useful elements of results information (for example, tracking the highly variable and scattered implementation of all BMPs installed as part of TMDLs). Realistically, state capabilities for tracking key results information are limited and many data gaps will likely continue to exist.

2. Identify appropriate program measures

- <u>Identify measures of the expected results.</u> Results measures (indicators, metrics) are carefully selected based on their direct or indirect association with a given result of interest and their ability to be measured on available, suitable data. TMDL measures may address fully successful program outcomes (formerly listed waters now attaining WQS), incremental progress (waters showing improvements after TMDL implementation), maintenance (waters not further degrading), explanatory variables (building a case that specific factors that may have caused an observed result), or programmatic milestones (e.g., number/type of implementation plans activated), among other concepts. Narrowly focused analyses such as OMB program review may concentrate on just one expected result and one measure; current USEPA strategic planning focuses on multiple types of results. By allowing for regional or state-level analyses and assessment of sub-national subsets of information rather than strictly national data, the results analysis project will increase the potential number and variety of measures we can learn from.

- <u>Find data sources from which these measures can be analyzed.</u> A large sample of waters must be studied to determine consistent relationships of measures to program results, thus either large-area data (statewide, regional or national scale geographic coverage) or large-number site data (hundreds to thousands of waters tracked in USEPA data systems) are generally required. With these kinds of numbers currently unavailable concerning full recoveries after TMDLs, this project plans to seek more empirical outcome data while building insights from supplementary sources. Given that results insights can come from empirical outcome data, from expert opinions, or from the literature, all three approaches can be tapped. The results analysis effort is using existing USEPA databases, from which new forms of data extraction and compilation are expected. Further, our project has conducted analyses of practitioner experience and the restoration literature as supplementary sources that build lines of evidence while more empirical outcome data are accumulated.

3. Carry out measures analysis

- <u>Compile evidence of results through data analysis.</u> Here, the specific results measures chosen above are analyzed using suitable data sources. The analysis should produce its findings as quantitative summaries of observable results seen in the waters under study – for example, 'In state X, 72% of waters with TMDLs implemented more than three years ago and recent monitoring show some form of water quality improvement.'

- <u>Analyze underlying, explanatory factors where possible.</u> Observed results are of far greater value for program improvement when they can be explained. This phase of results analysis reexamines a given dataset that has yielded a results finding to explain probable causes for that result. For example, an assessment may find that 'the majority of waters with implemented TMDLs and improved waterbody conditions also have active watershed organizations and funding for implementation.' Establishing absolute causal linkages is very useful, but often difficult. In the absence of sufficient TMDL case study data, the scientific literature and practitioner experience are important supporting sources of identifying explanatory factors.

- <u>Document observable patterns.</u> Patterns detected by results analysis may reveal a consistent environmental outcome or related explanatory factors. Results analysis reports should summarize observed patterns in ways that can help decision-makers apply new

information to program improvements and support program elements that are already working well.

4. Assess implications for the program

- Assess policy and program implications of these patterns. Analysis becomes assessment when data summaries are translated into policy-relevant findings. TMDL results analysis reports would be further studied to identify program alternatives that respond to the new understanding of results the program has been producing. In this stage, programmatic measures may need to be reexamined while evaluating the feasibility and impact of possible changes.

- Develop recommended actions for program improvement. Of the program action alternatives considered, the most feasible and desirable actions should be built into future work plans. In the TMDL program these may eventually take the form of revisions to listing and TMDL guidance, updated research priorities, new funding initiatives, improved tools and data systems, or regulatory updates.

RESULTS

The remainder of this paper reiterates the generic results analysis process (Table 1) in reporting on our progress to date for each step.

1. The state of our TMDL program results knowledge and gap analysis

Results expectations from the TMDL program. As a key component of the CWA goal to 'restore and maintain' the Nation's waters, the common expectations for TMDL program results are obvious in part. The recovery of formerly impaired waters is clearly the dominant results goal, but several other types of expectations exist. We compiled these expected results (Table 2) upon considering the Clean Water Act itself along with perspectives of past TMDL program reviews, public opinion, dialogues with states, and views of TMDL program staff and managers. These varied expectations include environmental outcomes and programmatic features as well as causes. These are the starting point for identifying results measures that can link to one or more of the listed expectations and to specific elements of the TMDL program. It is worth noting that some widely held expectations of the TMDL program exceed the statutory authority of USEPA's program as determined by the Clean Water Act; examples include TMDL implementation, over which USEPA has no regulatory authority, and the maintenance of unimpaired waters, which is not directly addressed by the TMDL program.

Table 2 - Common expectations from the TMDL program.

Environmental Outcomes
<u>**more recovered waters**</u>: *more formerly impaired waters fully meet their Water Quality Standards (WQS).*
<u>**more incremental progress on recoveries**</u>: *more waters partially meet more of their WQS or fewer stressors impair these waters.*
<u>**minimal new impairment rates**</u>: *few waters exhibit new impairments, signifying maintenance of water quality.*
<u>**optimized timing to maintain resources**</u>: *priorities and schedules for TMDLs and restoration favor sequencing actions to conserve and protect the most important aquatic goods and services.*

Programmatic Features
<u>**adequate program milestones and tracking**</u>: *Timely 303(d) lists, TMDLs completed, TMDL implementation and other critical results information on program progress is tracked and accessible.*
<u>**documented condition of impairments & recovery cases**</u>: *all impaired waters, as well as all full and partial recoveries, are well known and documented in accessible USEPA data systems.*
<u>**efficient program spending**</u>: *restoration resources are not squandered on waters with no potential for recovery.*
<u>**restoration partners**</u>: *appropriate collaboration of USEPA with state, federal and watershed-based organizations and other partners at restoration sites.*
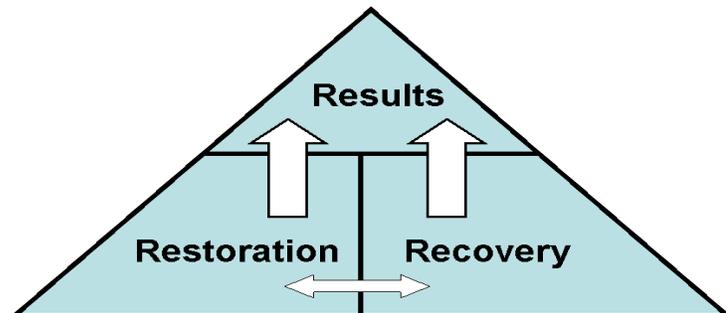
Causal Linkages
<u>**scientific explanations of recoveries**</u>: *factors responsible for water recoveries or maintenance have been revealed by technical analysis methods.*
<u>**plausible links of outcomes to CWA/TMDLs**</u>: *evidence of CWA and TMDL programs wholly or partly causing the outcome is documented.*

<u>*A Gap Analysis of TMDL Program Results Information*</u>. We chose to summarize the state of our knowledge and related informational gaps about TMDL program results in terms of three key components (Figure 3). The "three Rs" -- restoration, recovery, and results -- are tightly linked and fundamental to the success of water programs such as the TMDL program. These three elements form a logical, integrative structure for assessing and revealing gaps in information needed for program improvements. In brief, to fully understand and improve results, we will need to understand restoration and recovery as well.

Figure 3 – The three R's. *Restoration* (actions taken to improve condition) and *Recovery* (the improvements occurring in the waters) are the main drivers that bring about *Results* in the TMDL program.

Restoration has expanded in common-use meaning in recent years to the point that it encompasses all actions taken to improve the condition of US waters. Our knowledge of restoration centers on restoration practices, their appropriate use, and their effectiveness.

Recovery is the water body's improvement in condition to meet a previously unmet water quality standard, essentially the goal of most water programs. Recovery is what we all want to see. Our knowledge of recovery is based on studies of waters that have regained functionality and the ecological, stressor, and socio-economic factors that explain these recoveries.

Results is the effects of our program activities that are intended to restore and maintain the integrity of US waters. Our still-limited knowledge of results is based on empirical observations of TMDL success stories, literature, and practitioner experience.

Although closely related, these terms differ in useful ways. Results is about the effects of programs and policies; restoration is about the actions taken; recovery is about the changes in water bodies themselves.

The TMDL program strives to maximize the implementation of **restoration** actions in seeking impaired waters **recovery** as the desired **result**. In order to use the three Rs strategically, we must not only have more information but must organize it to support action. Merely knowing results without understanding their underlying causes does not support action; learning more about restoration and recovery is also essential. Integrating the analysis of results with examining the nature of recovery and evaluating restoration effectiveness will reveal what we need to know and what we need to do with that knowledge. Toward this goal, we conducted a results information gap analysis of the TMDL program in terms of the three Rs (Table 3).

Table 3 – Results information gap analysis for the TMDL program.	
Restoration Gap Analysis	
Current knowledge	*Knowledge gaps*
- we know that there are hundreds of BMPs, restoration treatments, TMDL implementation plans, and other actions that are available as restoration techniques. - we know that active, onsite restoration techniques are capable of bringing about recovery of a water body, often in an accelerated time frame. - we know that restoration techniques are not universally applicable or uniformly successful due to different settings, water body types, and regional differences. - we know that restoration is generally an emerging discipline with uncertainties, risks, and variable success rates. - we know that our knowledge of restoration technique effectiveness and appropriate technique selection is limited, and these limitations are impediments to successfully implementing other new restoration efforts. - we know that tens of thousands of waters (most outside of the TMDL process) have undergone restoration projects, about which some documentation but limited outcome information is available. - we know little about the relative proportion of restoration-treated waters that were 303d listed, versus impaired/unlisted or showing signs of emerging problems only. - we know that the costs of restoring all impaired waters	- knowledge about the effectiveness of as many practices as possible. - the appropriate selection factors for choosing the right technique for any given site. - the relative amount of restoration effort/funding invested in 303d listed waters. - the outcomes of more restoration projects along with a basis for relating these outcomes to similar sites.

greatly exceed the funding dedicated yearly to restoration from all major sources. - we know that restoration effectiveness is a major causal factor related to results.	

Recovery Gap Analysis

Current knowledge	*Knowledge gaps*
- we know that recovery is substantially less well understood than degradation. - we know that recovery is substantially less common than degradation, at present. - we know that different waters vary substantially in their likelihood of recovery, but we rarely consider recovery potential in CWA-driven actions and decisions. - we know that some data and tools for assessing recovery potential at the statewide scale can be developed. - we know that recovery is partially dependent upon a water body=s site-specific ecological capacity to reattain functionality. - we know that recovery is partially dependent upon a specific water body=s stressor exposure in the past, present and future. - we know that recovery is partially dependent upon a specific water body=s social, economic, and political setting. - we know that recovery is partially dependent not only upon the three intrinsic factors above, but also on the management actions applied to the site (see also Restoration). - we know that recovery is a delayed reaction that almost always takes years or decades, especially for larger water bodies and more complex systems, but we can poorly predict a given water body=s needed recovery time. - we know that recovery potential is a major causal factor related to results.	- a strong scientific basis for understanding recovery in general - tools for characterizing recovery potential on a site-specific basis. - tools for characterizing recovery potential on a statewide or national basis. - demonstrations of priority-setting with recovery potential among the considerations - testing recovery assumptions in predicting recoveries of numerous water bodies

Results Gap Analysis

Current knowledge	*Knowledge gaps*
- we know very well the target populations of our efforts (impaired or threatened US waters) nationally. - we know fairly well the condition of these waters nationally, although not from a statistical design. - we see isolated instances of full recovery, but these are still in very small numbers. - we see interim progress made by individual waters that are improving. - we suspect maintenance of condition occurs in many waters despite increasing pressures. - we know fairly well the expectations of results held by the public, OMB and others. - we generally know what to measure to evaluate site specific results of a management action, but rarely encounter post-project data. - we understand many factors that may influence results, such as time delays in response to restoration and to negative impacts, or intrinsic vulnerability to degradation. - we have very limited information linking results specifically to CWA=s influence.	- real outcome information on as many as possible waters, in the near term. - better evidence of CWA positively influencing the condition of US waters. - better evidence of incremental progress on waters that have not fully recovered. - better evidence of WQ maintenance in waters that have withstood increasing pressures. - greater investment in post-management action monitoring, strategically placed. - better statistical designs for attributing observed results information to national patterns and trends. - better understanding of the contributing as well as confounding factors affecting results information (see also Recovery).

The program expectations and gap analysis described above are the primary products of step 1, including a list of expected TMDL program results, the state of program knowledge relative to these results, and the primary information gaps to be addressed in our results assessments. We also developed conceptual diagrams of the TMDL program elements useful in identifying program elements of highest interest, selecting relevant measures, designing our assessments, and relating our analyses to specific opportunities in improving individual program components. Altogether, these provide a sound conceptual basis for proceeding with our results analysis efforts.

2. Types of results measures and their analysis

As planned, we used our findings about program expectations to identify broad groups of useful results measure types, and used our gap analysis to home in on more specific areas of interest. The numerous measures potentially useful in TMDL results analysis can be loosely grouped as environmental response (outcome) measures, programmatic progress measures, and explanatory (causal) measures.

- ***Environmental response measures*** are needed to track the results and evaluate the effects of CWA programs like the TMDL program. Outcomes of our CWA surface waters protection program in general are generalized as *restoration* outcomes (making impaired waters unimpaired), *incremental progress* outcomes (making impaired waters less impaired), and *maintenance* outcomes (minimizing or preventing further impairment). These measures are the most valuable empirical evidence of program outcomes, but data for assessing these measures are in limited supply.

- ***Programmatic progress measures*** are metrics that reflect the relative amount of work completed in successive stages of the TMDL process (such as number of TMDLs developed, or percent of TMDLs with implementation plans in place.) Despite not indicating environmental outcomes, these measures are important because some directly track USEPA statutory responsibilities, and others are recognized as essential steps in restoration efforts. Further, programmatic and response measures are both considered when weighing the options for program changes and improvements. Most programmatic measures can be related clearly to the 'pipeline diagram' of the TMDL process. These measures are generally well-tracked and accessible in USEPA and state databases.

- ***Explanatory measures*** generally match water quality response information with co-occurring, plausibly related stressor indicators. This type of metric intends to determine whether negative water quality effects are increasing concurrently with an increase in stressors, or if there are signs that the stressor effects are being held in check. Improvement, no degradation, or degradation slower than expected in the face of increasing stressors would be positive results. Usually such metrics must be narrowly defined (e.g., relating one land use type to effects from one pollutant type) to limit confounding variables, but they represent the primary analytical approach to increasing understanding of where and why positive results may be occurring. Understanding of the factors related to observed results is essential to making the connection to CWA as a contributing cause. Explanatory measures generally do not provide simple management

targets for strategies, which usually rely on response measures. Their greatest value is in revealing underlying drivers of successful results and causal links to CWA programs that can be used to improve the way these programs are implemented and managed.

Multiple activities and products have already contributed to the development and evaluation of TMDL results analysis measures. We participated in an OW strategic planning measures workgroup in 2005-06, which resulted in defining measures concepts well-aligned with the USEPA Strategic Plan and documenting these in a set of measures white papers. Further, we directed the development of the *TMDL Program Results Analysis Measures and Methods Report* (Research Triangle Institute 2006) as the main measures selection and planning document for TMDL results analysis. This report identified a broad array of over 100 candidate TMDL results measures, examined data availability and feasibility of analysis, and summarized recommendations for analysis of 12 specific measures. The Measures and Methods Report also built into its extensive measures review the findings of other USEPA-led TMDL program reviews including the Non-point and Stormwater Source TMDL Program Evaluation (Industrial Economics, Inc. 2006; Cabrera-Stagno et al. 2007), the Virginia Tech TMDL Center's Success Stories Report (Benham et al. 2006, 2007), and an internal USEPA 100 TMDLs Document Review. In addition, the recovery potential project has identified and demonstrated statewide screening analysis of over 80 explanatory measures linked to the likelihood of recovery of 303(d)-listed waters (Norton et al. 2007, Wickham and Norton 2007). Several of these reviews are also important sources of results measures analysis to date, as discussed below.

Despite our progress, we view identification of suitable measures as a continuing and dynamic process that has not ended. New knowledge from our assessments may continually reveal newly important results concepts and measures. USEPA's dialogue with states on TMDL program results is a crucially important source of adaptive adjustments in this project, thus we are also supporting a dialogue about TMDL results measures led by the Association of State and Interstate Water Pollution Control Administrators (ASIWPCA).

3. Measures Analysis
The project has compiled some measures analysis on an empirical dataset of fully recovered waters (Benham et al. 2007), on practitioner survey information (Cabrera-Stagno et al. 2007, Norton et al. 2007), and through literature review and synthesis (Wickham and Norton 2007). Readers are referred to companion papers in these Proceedings that describe each of these projects in greater detail. Currently, we are continuing additional measures analysis on our highest priority measures from the Measures and Methods Report.

We have also identified the databases from which more recommended measures can be analyzed as well as the pro's and con's of different data sources. Further, we have funded a grant to the National River Restoration Science Synthesis 37,000-site restoration database, are engaged in revising and improving the results-relevant data elements in USEPA's databases, and are participating in USEPA strategic planning dialogue on measures and database issues. Plans are also in place to support a USEPA Regional, multi-state analysis of TMDL implementation tracking information needs, capacity, and possible data system structure, as well as funding a second multi-state assessment of the driving factors underlying a larger sample of successful TMDL recoveries.

4. Assessing results implications for the TMDL program
At this point in our project's history, much of our data gathering still lies ahead and efforts are focused more on assessing measures than relating our findings to recommended program improvements. Nevertheless, the project has included a demonstration of explanatory analysis using TMDL-like case study data as a proof of concept on how improved understanding of results can be used to improve programs (Norton et al. 2007, Wickham and Norton 2007). The recovery potential project has used practitioner interviews and scientific literature review to explain water body and watershed traits associated with greater likelihood of recovery. GIS-based landscape analysis of over 80 recovery-linked causal measures has been demonstrated as a recovery potential screening method, using the 700+ impaired waters from the State of Illinois' 303(d) list (Norton et al. 2007). This screening tool is intended to improve the TMDL program by supporting state decisions on prioritizing their 303(d) waters for TMDL development. Statewide recovery potential data and tools may also be useful in selecting priorities for TMDL implementation, Section 319 investments and other restoration efforts. As such, the recovery potential project is an early example of how results analysis measures (particularly causal indicators) can be applied to specific program technical or procedural improvements.

DISCUSSION

As stated, the main goals of this effort are to improve insights about TMDL program effectiveness and use this knowledge to drive program improvements. The USEPA Strategic Plan has required state and USEPA regional reporting on several results measures that are partially attributable to the TMDL program, but these measures generally reflect the collective influence of the multiple programs that constitute surface waters protection under CWA (USEPA 2006). The TMDL results analysis project's measures will augment the existing strategic plan measures, and will differ by assessing features specific to the TMDL program and placing a greater emphasis on assessment of explanatory measures to the extent possible. As our measures assessments are still ongoing, this fourth and final step will occur as significant sets of findings are finalized and interpreted.

Before being able to make significant advances in assessing more empirical data on TMDL outcomes, the project will face several challenges in the form of data availability, quality and consistency, interpretation, strength of causal linkages, ability to extrapolate findings to other areas, and limited state reporting of empirical outcome data. Where these problems can be overcome, we will develop solutions; where intractable, we will document the issue and point out the potential value of resolving it, so that this may be considered in future work planning. Uncertainties will undoubtedly be present in the analyses, thus these will be openly documented along with transparent explanations of how analytical conclusions were reached. Further, the expense of investment in tracking and reporting results measures, after a point, is capable of draining funds from actual restoration, requiring careful consideration of the tradeoffs involved.

CONCLUSIONS

The TMDL program's results analysis project is addressing several as-yet unanswered questions about the effects of the program on impaired waters of the US. Many of the most compelling questions about environmental changes resulting from TMDL implementation cannot be definitively answered at the current time, for the reasons discussed in this paper's Introduction and Figure 1. Nevertheless, new analyses are proceeding to assess the available data in innovative ways while new results data are reported and compiled. Looking ahead, the following points about likely results information needs and program improvements are offered in a speculative sense based on our observations thus far:

- Although numbers of recovered waters are increasing, a shortage of case study data on recoveries still limits the opportunity to assess and report on why TMDLs do or do not succeed. Possibly the highest priority for results analysis is USEPA investment in documenting incremental improvements and full recovery of impaired waters.

- Based on anecdotal information from numerous sources, incremental improvements appear to be occurring on many waters – likely far more in number than the known full recoveries. Although some assumed improvements may be temporary (e.g., seasonal variation), others are likely legitimate, and all merit verification to help USEPA and states focus on understanding progress in many more waters. Sound methods for identifying incremental improvements and tracking specific waters undergoing improvements should be used.

- Increased post-TMDL monitoring will be needed to help increase the numbers of improving and recovered waters that are documented; close coordination among monitoring and TMDL programs will be essential.

- State reporting on performance measures is seen as a burden and USEPA recognizes the sensitivity of its reporting requests, but limiting the reporting burden also reduces results insights where there is no other alternative source of data. USEPA and states together should weigh the tradeoffs among the limited number of potential reporting measures and USEPA should consider alternative ways to obtain data that states are not tracking and reporting themselves.

- Whereas response measures and programmatic measures may indicate performance and progress, TMDL program improvement would rely heavily on analysis of the explanatory measures that tell why a given result is happening. Results information with unknown causes cannot help programs improve. If USEPA is to go beyond evaluating program performance to applying its new insights, it should assess causes of the results as well as measuring the program results themselves. This suggestion is pertinent to measures development for the USEPA Strategic Plan as well as in TMDL results analysis.

- Many practitioners believe that the differing regulatory pathways through the 'program pipeline' of point source-impaired versus non-point or mixed point and non-point source waters will reveal substantial differences in results, both at the stage of implementing controls and in eventual water quality improvements. USEPA should test and verify this assumption and, if major differences are evident, generally differentiate its TMDL program results data based on the relative roles of point or non-point sources.

- TMDL implementation is a key programmatic milestone. The difference in numbers of TMDLs developed to TMDLs implemented is believed to be a significant drop, but is virtually undocumented. Documentation of implementation, however, would be exceedingly challenging and costly on a national basis. This is because of the substantial number, variety, diverse participants, and lack of tracking of specific implementation

actions involved, especially nonpoint source control actions. To the extent possible, USEPA should measure and validate this assumption using feasible methods other than full national inventory and tracking, and also search for states, regions, or impairment scenarios where the ratio is better, in order to apply their approaches elsewhere.

ACKNOWLEDGMENTS

REFERENCES

Benham, B., R. Aeckoski, G. Yagow and S. Ekka (2006) *TMDL Implementation – Characteristics of Successful Projects: Final Report.* Prepared for the U.S. Environmental Protection Agency (Project X7-83156301), The Center for TMDL and Watershed Studies, Virginia Tech University.

Benham, B., R. Zeckoski and G. Yagow (2007) TMDL implementation: lessons learned. *Proceedings: Water Environment Federation TMDL 2007 Conference*, Bellevue, Washington.

Cabrera-Stagno, V., J. Heffelfinger and C. Kakoyannis (2007) An evaluation of the development of non-point source TMDLS. *Proceedings: Water Environment Federation TMDL 2007 Conference*, Bellevue, Washington.

Cleland, B. (2007) TMDL development from the "bottom up" – Part IV: connecting to stormwater management programs. *Proceedings: Water Environment Federation TMDL 2007 Conference*, Bellevue, Washington.

Federal Water Pollution Control Act (1972) Public Law 92-500, 86 Stat. 816. Amended in 1977 and 1987, referred to as the Clean Water Act, codified at 33 U.S.C. 1251-1387 (1988).

Industrial Economics, Inc. (2006*) Developing effective nonpoint source TMDLs: an evaluation of the TMDL development process.* Prepared for the U.S. Environmental Protection Agency Office of Water and Office of Policy, Economics and Innovation. Final Report.

Monschein, E. and L. Mann (2007) Category 4b – a regulatory alternative to TMDLs. *Proceedings: Water Environment Federation TMDL 2007 Conference,* Bellevue, Washington.

Norton, D.J., J. Wickham, T. Wade and J. Thomas (2007) Recovery potential: setting priorities for impaired waters restoration. (in review).

Research Triangle Institute (2006) *TMDL Program Results Analysis Measures and Methods Report.* Contract no. 68-C-02-110 report to USEPA Office of Water, Washington DC. 166 pp.

USEPA (2006) 2006-2011 *EPA strategic plan: charting our course*.   U.S. Environmental Protection Agency, Washington, DC. URL: http://www.epa.gov/ocfo/plan/plan.htm. Accessed 2/27/07; last updated 3/20/07.

USEPA. (2007) *Total Maximum Daily Loads: National Section 303(d) List Fact Sheet*. URL: http://oaspub.epa.gov/waters/national_rept.control. Accessed 2/14/07; last updated 4/3/07.

Wickham, J.D. and D.J. Norton (2007) Applying results findings: the recovery potential project. *Proceedings: Water Environment Federation TMDL 2007 Conference*, Bellevue, Washington.