Expanding the Use of Outcome Measurement for EPA's Office of Enforcement and Compliance Assurance

Report to OMB

Prepared by
Office of Enforcement and Compliance Assurance
U.S. Environmental Protection Agency

July 31, 2006

EXPANDING THE USE OF OUTCOME MEASUREMENT FOR EPA'S ENFORCEMENT AND COMPLIANCE ASSURANCE PROGRAM

Table of Contents

EXPANDING THE USE OF OUTCOME MEASUREMENT FOR EPA'S ENFORCEMENT AND COMPLIANCE ASSURANCE PROGRAM

EXECUTIVE SUMMARY

Focus of the Report

This report was developed to respond to the OMB instruction in the EPA FY07 budget passback to "review existing State performance metrics and pick the best to be incorporated into a lessons learned report that can serve as the basis for improved EPA metrics." In addition, OECA broadened the focus of the report to solicit ideas for additional outcome-based performance measures from a wide audience of co-regulators, stakeholders, and experts.

Actions Taken to Conduct Research

OECA took the following steps to conduct research for this report: reviewed 57 state environmental programs to review performance measures for their compliance and enforcement programs; reviewed 40 state compliance and enforcement programs described in the ECOS report entitled, "State Contributions to Enforcement and Compliance;" conducted discussions with the ECOS Compliance Committee, academic experts on performance measurement and compliance management, industry and environmental group stakeholders, and indicators experts from the International Network for Environmental Compliance and Enforcement (INECE); reviewed the OMB PART assessments of 16 federal regulatory programs, and the strategic plans and performance measures of nine federal enforcement programs.

Major Findings

1. Although state environmental agencies have developed useful, and, in some cases, sophisticated combinations of outcome measures to assess overall environmental conditions in their respective states, their measurement of outcomes from environmental compliance and enforcement programs is not very advanced.

2. No state environmental agencies are currently using sampling approaches to develop representative, statistically-valid compliance rates. The use of statistically-valid rates by state environmental agencies is occurring only in the circumstance when 80% or more of a regulated population is inspected, thereby creating a rate based on a "near census."

3. No state environmental agency is currently using, or planning to use, a "universal compliance rate" that aggregates data about all violations in the entire regulated universe under one rate.

4. Among state environmental agencies, the use of outcome measures to assess and improve performance or to report to the public is very rare.

5. Development of a universal compliance rate for the nation will not meet two criteria important in the selection and implementation of performance measures – value and feasibility – and would not meet the GAO test of "balancing ideal performance measurement systems against real-world considerations such as the cost and effort involved in gathering and analyzing data."

6. OECA's current GPRA strategic architecture, built around the four primary tools of the compliance and enforcement program (assistance, incentives, monitoring and enforcement) does not fully demonstrate how program outputs and outcomes are contributing to the reduction or elimination of environmental problems.

7. The use of approaches such as near census, near census plus modest investment, multi-year rates, and regional or state rates for specific sectors, as described in Section III, can enable OECA to expand its effort to develop and use statistically-valid rates.

8. The review of federal regulatory and enforcement programs did not identify performance measurement or compliance rate practices that were applicable and useful for managing EPA's compliance and enforcement programs.

9. Deterrence research holds promise for compliance and enforcement practitioners because it can help illuminate the factors that contribute to compliance and improve understanding about which tools or combination of tools produces the most effective results.

Recommendations

1. Recommendation: OECA does not recommend development of a "universal compliance rate." The cost and effort involved in developing and using a statistically-valid, universal compliance rate for the nation far outweighs the likelihood of any benefit such a rate might bring to program management and improved performance.

2. Recommendation: OECA will expand development of statistically-valid compliance rates for specific noncompliance patterns, when there is a clear management need that can be served by the use of a compliance rate. Development of rates will focus on national priorities or other important problem areas such as those identified as part of the transition to the problem-oriented approach. In addition, in choosing whether to develop a rate, OECA will weigh the relative importance and value of the rate and the resources and effort necessary to produce the rate.

3. Recommendation: OECA will continue making progress on characterizing public health benefits from air pollution reductions from enforcement cases by working with OAR to use their BenMAP model on OECA's air enforcement

case data. The model estimates the health benefits of pollutant reductions and reports those benefits in terms of reductions in premature mortality, chronic and acute bronchitis, myocardial infarctions, hospitalizations, respiratory symptoms, and/or work loss days. OECA will continue to explore how pollutant characterization might be used to target program activities on the most harmful pollutants, especially as part of its transition to the problem-oriented approach described in #7 below.

4. Recommendation: By September 2006, OECA will complete an analysis of its experience with the recidivism measure used from 1999-2004 and the chronic non-compliers data it has developed in the last two years. The re-evaluation will make recommendations about using one of those approaches for a pilot project to be conducted in FY 2007 to develop a repeat offenders measure for the civil enforcement program.

5. Recommendation: Guided by the results of a feasibility study in FY 2007, EPA will conduct a survey every three years of a statistically-valid sample of compliance assistance recipients to measure behavior changes resulting from compliance assistance. The survey would begin in 2008 and would be used as both a management tool to guide and improve compliance assistance programs as well as a publicly-reported measure of program performance.

6. Recommendation: EPA will begin formulating a research approach to learn more about the deterrent effect and how it might be measured. Among the possible options for this research, OECA will review the feasibility of a periodic survey of representatives of regulated facilities and industries to better understand what factors motivate compliance behavior and what combinations of tools and practices would maximize such behavior.

7. Recommendation: By November 30, 2006, OECA will develop an implementation plan for a problem-oriented approach for measuring outcomes of the national compliance and enforcement program. By developing a plan, OECA will be able to conduct outreach and consultation with regional and headquarters compliance and enforcement managers, state co-regulators, academic experts, and stakeholders. The plan would identify tasks, costs, and schedules for at least five stages: identification of problems; evaluation and characterization of problems; development of problem-specific measures; identifying data system implications; development of GPRA architecture.

SECTION I. PURPOSE AND CONTEXT

A. Purpose of the Report

In its preparation of the FY 2007 President's Budget, the Office of Management and Budget (OMB) directed EPA's Office of Enforcement and Compliance Assurance (OECA) to

> "… submit a report through the Deputy Administrator by July 31, 2006 that reviews existing State performance metrics and picks the best to be incorporated into a 'lessons learned' report that can serve as the basis for improved EPA metrics. It is envisioned that this report will look at using metrics that are statistically-valid, including compliance rates, reductions in pollution characterized as to risk, recidivism, contamination incidents, and other tools."

In carrying out the review requested by OMB, OECA chose to broaden and expand the scope to also: 1) solicit ideas for additional outcome-based performance measures from a wide audience of co-regulators and stakeholders; and 2) evaluate potential new outcome measures for implementation by OECA. In implementing its original National Performance Measures Strategy, OECA in FY98 engaged a wide variety of stakeholders regarding potential measures for its consideration. OECA viewed the current request from OMB as an opportunity to solicit fresh input about performance measurement from stakeholders in order to maximize the benefit of this review to OECA, the Agency and all other interested parties.

B. OECA's Approach to Performance Measurement

Beginning with its National Performance Measures Strategy report completed in FY1998, OECA has had an active and conscientious effort underway to identify, develop, and use performance indicators to manage the national compliance and enforcement program and enhance accountability to the public. In particular, OECA has focused on measuring not just outputs or activities (e.g., the number of enforcement actions), but the intermediate and end outcomes or results of those activities (e.g., pounds of pollution reduced through enforcement actions). In addition, OECA has expanded its measures to encompass <u>all</u> of its principal program activities, including assistance and incentives rather than only monitoring and enforcement. OECA's efforts have included periodic assessment and updating of performance measures based on actual experience in using measures for management and accountability purposes. At this point in the evolution of OECA performance measurement, the biggest measurement challenges are: 1) measuring outcomes from compliance assistance; 2) measuring the deterrent effect of enforcement; and 3) improving the use of statistically-valid compliance rates.

The performance measures that OECA has identified, developed, and used serve Goal 5 (Compliance and Environmental Stewardship) of the EPA Strategic Plan required by the Government Performance and Results Act (GPRA). Within Goal 5, the

compliance objective and its three sub-objectives (for assistance, incentives, and monitoring and enforcement) form a management architecture that helps guide the national compliance and enforcement program. Appendix 1 provides the GPRA architecture, including the performance indicators associated with each sub-objective. The objective and sub-objectives provide numerical targets for increasing specific end and intermediate outcomes resulting from the activities of the national program.

Over the past ten years, OECA has gained significant experience in performance measurement. The managers of EPA's national compliance and enforcement program, in their efforts to identify, develop, and use performance measures as an important tool for managing and improving program performance have been guided by a set of principles.

1. A combination of output and outcome measures is necessary to manage and improve the national compliance and enforcement program.

Discussion: OECA program managers need to ensure that certain activities or outputs are being carried out, and assess whether those activities are producing the results or outcomes that are most important for protecting the environment. By using a combination of output and outcome measures, OECA is able to identify patterns between the production of activities and the occurrence of outcomes, and adjust our strategies to ensure that we are achieving the right types and amounts of outcomes with the resources available.

2. Intermediate outcome measures provide valid and useful information that helps identify and correct performance issues.

Discussion: OECA measures pounds of pollution reduced, an end outcome. But intermediate outcome measures have several advantages over end or final outcome measures that are often overlooked by evaluators. First, most intermediate outcomes (e.g., changes in behavior) can be directly attributed to the activities of the compliance and enforcement program. The causal link between program activities and intermediate outcomes is very strong. Second, intermediate outcomes almost always manifest themselves more quickly than end outcomes which often focus on changes in large-scale environmental conditions. Third, intermediate outcomes are almost always less costly to measure than end outcomes which often require expensive monitoring systems.

3. Performance measures should be used for more than reporting results to the public. Their most important use is as a management tool to identify and correct performance issues and improve program effectiveness.

Discussion: Since developing an expanded set of performance measures over the last several years, OECA has utilized its measures to manage the national enforcement and compliance program. For example, performance measures are used to develop regular management reports to senior headquarters and regional managers to track progress on key outputs and outcomes. Performance measures are also used for reports given to OECA's Assistant Administrator to conduct management reviews of regional

compliance and enforcement programs. The measures are also the basis for the State Review Framework which utilizes a set of agreed-upon performance metrics to examine the adequacy of state compliance and enforcement programs and EPA's direct implementation efforts in those instances in which programs have not been delegated.

4. Expansion of existing measures or development of new measures must take resource implications into account.

Discussion: In making decisions about adding new or expanding existing performance measures, OECA has been guided by two criteria: the value of the measure and the cost of implementing and maintaining the measure. As stated by the Government Accountability Office (GAO), agencies "must balance their ideal performance measurement systems against real-world considerations such as the cost and effort involved in gathering and analyzing data (GAO, 1996)." This weighing of importance versus feasibility of measures has figured prominently in OECA's approach to performance measurement.

SECTION II. REPORT METHODOLOGY

To address OMB's request to review existing state performance measures and incorporate them into a "lessons learned" report, and to solicit ideas for additional outcome measures from a broad audience, OECA staff and managers gathered and analyzed information from a wide range of sources through a variety of means. Specifically, OECA took the following steps to develop the findings and recommendations in this report.

A. Actions Taken to Conduct Research

Performance Measurement by State and Federal Programs

- ✓ Researched 57 state environmental programs (several states had both a Department of Natural Resources and a Department of Environmental Protection) to review performance measures for their enforcement and compliance programs.

- ✓ Reviewed the information about performance measurement for 40 state compliance and enforcement programs described in the report, "State Contributions to Enforcement and Compliance," soon to be published by the Environmental Council of the States (ECOS).

- ✓ Conducted conference call with the ECOS Compliance Committee to discuss the current use of compliance rates and whether states were interested in adopting statistically-valid compliance rates.

- ✓ Reviewed the work of the Compliance Consortium, a network of state compliance and enforcement programs that has shared information and best practices about performance measurement among member states.

- ✓ Met with the EPA Office of the Inspector General (OIG) staff and managers to learn whether their ongoing review of best practices for compliance rates had identified specific states using statistically-valid compliance rates.

- ✓ Reviewed the performance measures and strategic plans of 16 federal agency regulatory-based programs that had received PART ratings of adequate, moderately effective, or effective, and another nine federal enforcement programs that have not yet been subject to PART review.

Measurement Practices in Other Countries

- ✓ Met with members of the Expert Working Group on Performance Indicators of the International Network for Environmental Compliance and Enforcement (INECE). Members represent individual countries, the Organization for

Economic Cooperation and Development (OECD), the World Bank Institute, the European Union, and the INECE Secretariat.

- ✓ Met with managers and staff of Environment Canada's compliance program to discuss their use of compliance rates.

Views of Academic Experts

- ✓ Met with two performance measurement experts, Harry Hatry of the Urban Institute, and Kathryn Newcomer of George Washington University.

- ✓ Met with two experts on management of compliance programs, Malcolm Sparrow of the Kennedy School of Government at Harvard University, and Lee Paddock of Pace University Law School.

Consultations with Stakeholders

- ✓ Met with representatives of the American Chemistry Council to discuss their proposal to develop compliance rates measuring the number of compliance obligations for categories of regulated facilities.

- ✓ Met with representatives of U.S. Public Interest Research Group, the Environmental Integrity Project, Natural Resources Defense Council, Environmental Working Group, and OMB Watch to get their perspectives about compliance rates and other performance measures.

B. Summary of Research and Discussions

The results of the research about state compliance and enforcement programs, as well as the discussions with the ECOS Compliance Committee and the OIG can be found in Section III, C. (States' Use of Compliance Rates), and in Section IV, A. (Outcome Measurement in State Compliance and Enforcement Programs).

The results of research and discussions about the compliance rate approach suggested by the American Chemistry Council and one used by Environment Canada are in Section III, D. (Other Approaches to Compliance Rates).

The academic experts consulted by OECA provided many useful insights. See Appendix 4 for a list of questions posed to these experts. Harry Hatry of the Urban Institute discussed ways to identify significant environmental outcomes to which compliance and enforcement programs might contribute, and reiterated (as stated in his books and articles) the value of intermediate outcomes as a source of information useful to managers. Kathryn Newcomer of George Washington University discussed the strengths, weaknesses, and technical issues associated with compliance rates, presented the concept of "plausible attribution" as the standard that government programs need to meet to link their activities to broader social outcomes, and commented about the need

for a combination of output and outcome measures to aid program management. Malcolm Sparrow of Harvard University discussed the distinct advantages of a "problem-oriented" approach to measurement and management of compliance and enforcement programs, the futility of a "universal" compliance rate for a compliance and enforcement program as complex as EPA's, and how rates that are not statistically-valid can mislead programs in which they are used as a management tool.

Representatives of environmental organizations also provided many comments and reactions that can be summarized as follows. First, they did not place much value on compliance rates as a measure of performance for either government or regulated facilities, claiming that other measures (including basic output measures about the number of inspections and enforcement actions) were more important for their purposes. Second, they were concerned that if OECA expanded the use of statistically-valid rates, too many inspections targeted at known or suspected violators would be directed to random inspections for the sake of constructing a compliance rate. Third, they expressed concern that the total volume of performance measurement being required of the compliance and enforcement program had reached a point of impeding the capacity to carry out the program's mission.

SECTION III. COMPLIANCE RATES AS A PERFORMANCE MEASURE

A. Rates: Types and Issues

Compliance rates, when developed properly using statistical techniques, can provide useful information about one of the significant outcomes of a compliance and enforcement program, i.e., the percentage of a regulated universe (or some portion of it) in compliance at the time the rate was developed. Compliance rates can be based on inspections or on-site visits, or on self-reported data (e.g. Discharge Monitoring Reports reported under the Clean Water Act). Compliance rates can be statistically-valid—that is, representative of the population being measured—either because observations have been gathered from at least 80% of the population (a kind of "near census" rate), or by collecting observations of a representative sample of the population. Most inspection-based compliance rates being used by environmental compliance and enforcement programs tend not to be statistically-valid because they do not cover 80% or more of the population.

There are a number of different ways that compliance rates can be structured or defined, and the complexity and level of effort that go into developing a compliance rate depend, to a large degree, on the following factors:

- The size and geographic distribution of the population of interest: the number of inspections required grows as the size of the population in question grows. As a practical matter, as the population is more broadly distributed the cost of reaching them increases.

- The number of obligations for which a compliance determination must be made: as the number of compliance determinations that must be made increases so does the time and effort to make them. If the types of compliance obligations are dissimilar, then multiple inspectors or media experts may be needed to make the determinations. In addition, numerous and dissimilar compliance obligations may require the development of multiple or stratified random samples, both of which would increase the sample size.

- The relative difficulty of making the compliance determination: the difficulty of making compliance determinations varies greatly. At its simplest, compliance can be determined through a visual inspection or review of records. At its most complex a compliance determination may require sampling and testing, or a detailed engineering review of facility operations. As the level of difficulty increases so does the time and cost involved with making the determination.

The following examples illustrate how these factors impact the ability to develop different types of compliance rates.

1. Multimedia Facility Rate

$$\text{Multimedia Facility Rate} = \frac{\text{\# of facilities with one or more violations}}{\text{\# of regulated facilities}}$$

This type of universal multimedia rate is difficult to develop given the number of statutory and regulatory requirements involved and the compliance determinations that would have to be made. Depending on the size and complexity of a facility, this could require tens to hundreds of compliance determinations. Further complicating this type of rate is the difficulty of making many of the compliance determinations, and the geographic distribution of facilities.

2. Specific Regulation Rate

$$\text{Specific Regulation Rate} = \frac{\text{\# of facilities violating the specific regulation}}{\text{\# of facilities to which the regulation applies}}$$

Limiting the rate to a specific regulation or a specific compliance requirement can simplify the development of a compliance rate by narrowing the scope and clearly defining the types of compliance determinations that must be made. Good information on the regulated universe must be available in order to develop a representative sample. Many populations may still be so large as to make developing a rate very resource-intensive.

3. Specific Regulation, Specific Population Rate

$$\text{Specific Regulation, Specific Population Rate} = \frac{\text{\# of facilities violating the specific regulation}}{\text{\# of facilities in the specific population}}$$

Compliance rates can be further specified by limiting the rate to a specific population (e.g., industrial sector, or geography) as well as a specific regulation or compliance requirement. Limiting the population in this way can also help to reduce the cost of developing a rate. All of the compliance rates that OECA has developed to date have been of this type.

Unfortunately, producing representative, statistically-valid rates has proven very difficult for compliance and enforcement programs, as will be discussed in detail later in this section. The primary causes of this difficulty are the finite and limited number of inspections that can be conducted given available resources, and the growing need to target those inspections at regulated entities that are known or suspected violators. The movement toward targeted inspections has been in response to changing views of how to improve effectiveness of compliance and enforcement programs (i.e., by focusing resources on identified problems rather than on general monitoring of the entire universe) and the pressure such programs are under to produce valuable outcomes.

B. States' Use of Compliance Rates

Based on the information contained in the ECOS report and OECA's research on state compliance and enforcement programs, the vast majority of compliance rates used by state programs were calculated based on the following ratio:

$$\frac{\text{number of facilities with violations}}{\text{number of facilities inspected}}$$

Given that these inspections are usually conducted at less than 80% of the facilities in the population and are primarily targeted rather than random – a fact confirmed by the ECOS report, OECA's research, and the discussion with the ECOS Compliance Committee -- rates using this numerator and denominator are not statistically-valid because they are not representative of the population being measured. Such rates are really a kind of "hit rate," measuring whether targeted inspections are succeeding at identifying violations.

The ECOS report points out that "the largest use for compliance rates is program targeting," and that states use rate information to develop sector initiatives, target compliance and technical assistance, develop and target inspection initiatives, target information/outreach to the regulated community, target follow-up enforcement, and to develop industry-wide compliance characterizations.

Where states were developing or using statistically-valid rates, it was usually a "near census" rate. For example, the ECOS report pointed out that many states are developing rates for drinking water and RCRA UST using self-reported data from the population, and in response to a federal requirement to do so.

Although the vast majority of states employing rates as metrics for their programs are using non-statistically-valid inspection-based rates, there are a handful of states that are applying the use of statistics to track performance within their innovative permitting projects.

There are currently 16 states implementing an Environmental Results Program (ERP). The ERP is an alternative to traditional permitting that incorporates inspections, compliance assistance, self-certification and statistically-based performance measurement in order to reduce the environmental impacts of business. ERP uses statistics to estimate performance (e.g., compliance with select requirements and adherence to best practices) for a large group of facilities where inspection resources are limited. This is similar to the approach that OECA has taken with its SVNCR project wherein we assess performance against a key set of federal requirements, except in the ERP the state has added some non-regulatory performance indicators such as best practices and pollution prevention.

Massachusetts and Rhode Island have been the leading states in the implementation of Environmental Results Programs and the other 14 states have been working to follow their example. EPA did not, nor did ECOS in their recent report on

state measurement activities, analyze whether the states are following the Massachusetts model in their use of statistical techniques.

We could find no instances of any state developing or using statistically-valid compliance rates through the use of random inspections or by combining targeted and random inspections. Based on the discussion with the ECOS Compliance Committee, the principal reason for this seemed to be the unwillingness to sacrifice targeted inspections for the random inspections necessary to produce a statistically-valid rate. When queried about their interest in moving toward statistically-valid rates as a performance measure, considerable resistance was expressed. As one member put it, "should inspections be used to *measure* compliance or to *achieve* compliance?" (emphasis added)

Appendix 3 provides information about state compliance rate information.

C. EPA's Use of Statistically-Valid Compliance Rates

Over the past seven years, OECA has piloted various methods for calculating statistically-valid noncompliance rates (SVNCR) for industry sectors and specific segments of regulated populations.

In FY 1999 OECA worked with a PhD statistician from George Mason University to develop methodologies for calculating representative (statistically-valid) noncompliance rates. Methodologies for calculating statistically-valid noncompliance rates were developed for data derived from on-site inspections, and self-reported data such as Clean Water Act Discharge Monitoring Reports (DMRs). The methodology for rates derived from on-site inspections was designed to allow EPA to combine targeted inspections with an increment of randomly-selected inspections. OECA began piloting these methodologies in FY 2000; the table below shows the sectors covered and methods used from FY 2000 through FY 2004.

Year	Sector and Noncompliance Rate	Method
FY 2000, 2001, 2002	Petroleum Refining: Ammonia, zinc and lead violations with more than 20% over NPDES limit	self-reported DMR data
FY 2000, 2001, 2002	Iron and Steel: Ammonia, zinc and lead violations with more than 20% over NPDES limit	self-reported DMR data
FY 2000, 2001, 2002	Municipalities: BOD and TSS violations with more than 40% over NPDES limit	self-reported DMR data
FY 2001	Organic Chemical Manufacturing: RCRA Small Quantity Generator Compliance	statistically-valid inspections
FY 2001	Iron and Steel and Metal Services: DMR Accuracy Audit	statistically-valid inspections
FY 2002	Ethylene Oxide Manufacturers: MACT Compliance	statistically-valid inspections

Year	Sector and Noncompliance Rate	Method
FY 2002	Combined Sewer Municipalities: CSO Nine Minimum Control Policy Compliance (baseline)	statistically-valid inspections
FY 2004	Combined Sewer Municipalities: CSO Nine Minimum Control Policy Compliance (reevaluation)	statistically-valid inspections
FY 2004	RCRA Foundries: Compliance with RCRA Regulations	statistically-valid inspections
FY 2006	Feasibility determination for compliance rates development from data in legacy systems (AFS, PCS, RCRAInfo)	Near-census based rate development based on violation data reported by states and EPA during routine and targeted required inspections

Though a number of noncompliance rates have been calculated using self-reported DMR data, the majority of work to date has gone into piloting the process for calculating noncompliance rates based on combining targeted and randomly-selected compliance inspections. A number of factors were used to select populations for statistical noncompliance rate analysis, including: the size of the population, significant environment risks presented by the population, and the ease with which noncompliance could be determined. Following the selection of a sector, the regulatory requirements against which compliance should be measured are identified and instruments to collect the inspection results are developed. In consultation with a PhD statistician, a statistical sample size is identified for each sector, and a prescribed number of random and targeted facilities are identified to be inspected, by region. Written guidance is provided to the regions and the overall effort coordinated as part of the annual work planning process.

Examples of Statistically-Valid Rates for Self-Reported Violations

`	Petroleum Refining	Iron and Steel	Iron and Steel	Municipal	Municipal
Parameter	**Ammonia**	**Zinc**	**Lead**	**BOD**	**TSS**
FY '02 Noncompliance Rate*	4.85%	15.52%	1.79%	10.97%	14.43%
FY'01 Noncompliance Rate	6.60%	22.22%	5.0%	12.98%	15.79&
FY '00 Noncompliance Rate	9.35%	13.04%	7.94%	12.2%	15.53%

*Noncompliance is defined as >20% over NPDES limit for toxic pollutants and > 40% for BOD and TSS.

Examples of Statistically-Valid Inspection-based Rates

Sector	Regulation	Noncompliance Rate*
Organic Chemical Manufacturing	RCRA Small Quantity Generator Regulations	34.3% (+/- 8.1%)
Iron and Steel	DMR Data Accuracy	6.25% discrepancy rate
Metal Services	DMR Data Accuracy	44.2% discrepancy rate
Ethylene Oxide Manufacturers	CAA MACT Standards: Sterilizer Vent Regulations and Aeration Room Regulations	46% (sterilizer vent) 33.3% (aeration room) 49.2% (overall)
Combined Sewer Municipalities (2002 Baseline)	CWA Nine Minimum Controls	61.4%

*Noncompliance rate is defined as having a minimum of one violation with any given requirement examined during the inspection. Margin of error is +/- 5% unless noted.

What can be learned from the SVNCR pilot projects conducted by OECA?

1. The methodology developed for combining targeted and random inspections was successful in producing statistically-valid rates.

2. To conduct the random inspections necessary to produce statistically-valid rates, a number of targeted inspections were sacrificed. For example, the CSO (2002) project required 215 random inspections, the ethylene oxide rate needed 67 random inspections, the CSO (2004) project required 214 random inspections, and the RCRA foundries rate required 112 random inspections. These random inspections reduced targeted inspections by a corresponding number.

3. Because targeted inspections had to be sacrificed, OECA conducted each pilot and applied the methodology to small populations where it was feasible to produce a statistically-valid rate without sacrificing large numbers of targeted inspections As a result, the rates tended to focus on noncompliance patterns of lesser strategic importance. The one exception was the project for combined sewer municipalities. But to conduct that project, OECA not only sacrificed 214 targeted inspections, it also directed $150,000 through a managed competition to states to conduct some of the random inspections. (In the RCRA foundries project, OECA had to spend $86,000 to pay for the hazardous waste sampling to make a compliance determination.)

4. Future compliance rate projects should be driven by the need to develop a rate to measure compliance with an important regulatory requirement in a significant population of the regulated universe, not by the opportunity to test the methodology as was the case with the pilot projects.

D. Other Approaches to Compliance Rates

In addition to reviewing state use of compliance rates, OECA also reviewed an approach offered by the American Chemistry Council (the trade association for the chemical industry), and Environment Canada's efforts to develop and use compliance rates.

1. American Chemistry Council (ACC)

ACC's approach is built on the premise that a meaningful measurement of compliance should take into account the total number of compliance obligations applicable to a particular facility, or at least some scale of magnitude of those obligations. According to ACC, to be a "true rate," a compliance rate must contain both a numerator – the number of violations observed – and a comparable denominator – the number of possible violations during the observation period. ACC and the American Petroleum Institute produced an analysis that estimates the number of federal obligations that apply to various sizes of chemical plants and oil refineries. The analysis concluded that the number of compliance obligations at these facilities ranges from tens of thousands (10^4) to millions (10^6).

The ACC approach would replace the customary type of information generated by compliance rates, e.g., a facility with five instances of noncompliance with air program requirements over a given time period, instead expressing the level of noncompliance by saying that a facility had five violations out of 100 or 10,000 or 100,000 compliance obligations over that same period. ACC also acknowledges that it would be advisable to develop some way of assigning relative significance to violations since, for example, a violation for failure to report information might be less serious than an emission that exceeds a permit limit.

Producing compliance rates using the ACC approach raises various concerns. First, gathering the information about the number of facility obligations would require a massive collection of information by EPA and industry. Even if conducted using an approach that produces a statistically-valid sample, the number of distinct types of facilities to which federal environmental requirements apply and their variation by number of obligations would require a very large sample. Second, as ACC points out, their approach may require that a greater proportion of inspection resources be diverted from traditional "for cause," i.e. targeted, inspections. Third, our research did not identify any agency using compliance obligations as the denominator. Finally, such an approach applied to enforcement functions would yield information of little utility. For example, a traffic violation rate using the compliance obligation approach would indicate that a particular driver violated the law only once last year due to running a red light, could have violated traffic laws in thousands of additional instances, but apparently did not. The seriousness of that violation could be minor, i.e., the car proceeded through the intersection without incident, or it could be extremely serious, i.e. the car struck three other cars, damaged property, and caused injuries or fatalities.

2. Environment Canada

Discussions with Environment Canada did not identify an approach that could be useful for producing either universal or more focused rates that are statistically-valid. Environment Canada is not developing or using a universal rate, instead focusing on rates for specific industry sectors or in connection with new regulations. Their approach is a ratio of the number of "compliant inspections" and the total number of inspections in the population they are measuring. This approach is similar to the approach of state compliance and enforcement programs in that such rates are based on targeted inspections, do not include 80% or more of the population, and are therefore biased. This limitation has kept Environment Canada from reporting their rates publicly. Their use has been limited to internal targeting purposes. Environment Canada has also attempted to use rates (though not statistically-valid) to determine whether compliance improved after a targeted campaign of compliance assistance.

E. Prospects for Future Use of Compliance Rates

The principal constraint on developing and using statistically-valid compliance rates that rely on random inspections or the combination of targeted and random inspections is that inspection resources are finite and sacrificing targeted inspections in significant numbers can have an adverse impact on fulfilling the law enforcement mission.

In his book entitled, The Regulatory Craft (Brookings Institution Press, 2000), Malcolm Sparrow of Harvard University describes the challenges associated with developing and using meaningful compliance rates in regulatory programs and comments specifically on the use of targeted inspections to develop rates:

> Another difficulty relates to the fact that most readily available compliance data come from focused or biased inspection programs, which either deliberately target high-risk facilities or respond to incoming reports or complaints. Such focused or biased inspection programs help deal with specific risks that are already identified, *but they cannot provide statistically-valid estimates of general compliance behavior* or reveal emergent risks. These purposes *require representative sampling (either random or comprehensive) and require diversion of inspection or audit resources away from focused or complaint-oriented programs* [emphasis added]. (p.290)

To illustrate the impact of this diversion of inspection resources away from focused or targeted programs, Table 3.1 presents various combinations of targeted and random inspections for fixed levels of inspections of a regulated population of 250, and further demonstrates the sacrifice in rate precision due to resource constraints which limit the number of inspections conducted.

Table 3.1 Illustrating the Sacrifice in Precision When Allocating a Fixed Number of Inspections to Targeted and Randomly Selected Sites for Four Values of Fixed Number of Inspections

Population Size	FIXED Number of Inspections	Targeted Inspections	Random Inspections	Half Width 95% Conf. Interval
250	**200**	0	200	3.1%
250	**200**	25	175	3.1%
250	**200**	75	125	3.3%
250	**200**	125	75	3.6%
250	**200**	175	25	4.8%
250	**200**	200	0	--
250	**150**	0	150	5.1%
250	**150**	25	125	5.3%
250	**150**	75	75	6.0%
250	**150**	100	50	6.8%
250	**150**	125	25	8.8%
250	**150**	150	0	--
250	**125**	0	125	6.2%
250	**125**	25	100	6.6%
250	**125**	50	75	7.2%
250	**125**	75	50	8.2%
250	**125**	100	25	10.7%
250	**125**	125	0	--
250	**100**	0	100	7.6%
250	**100**	25	75	8.3%
250	**100**	50	50	9.6%
250	**100**	75	25	12.7%
250	**100**	100	0	--

The table demonstrates the difficulty of allocation problem; i.e., the tradeoff between random inspections conducted in order to make a statistical inference about a population, and targeted inspections to address a specific risk or noncompliance problem. The table illustrates how this tradeoff, for a population size of 250, affects the precision of the estimate of the compliance level for different allocations of random and targeted inspections. For example, if 200 random inspections are conducted (first grouping at top of table) and a compliance level of R is observed, the 95% half-width confidence interval would be 3.1%, typically expressed as (R ±3.1%).[1] If only 25 of the 200 inspections are random, and the remaining 175 are targeted, then the half-width confidence interval (i.e., precision of the estimate) is denigrated to 4.8%, which is still reasonable. However, if there are 100 inspections, all of which are random (last grouping in the table), then the half-width confidence interval is 7.6%, a level of precision that brings into question the utility of the estimate. When 25 of the 100 inspections are random, the precision worsens

[1] A 95% confidence interval is generated by a process that is right 95% of the time – so, if we were to repeat the process of generating compliance rates and confidence intervals many times, the actual compliance level would fall within the confidence interval 95% of the time.

considerably, resulting in a half-width confidence interval of 12.7%, a level of precision that we would not want to base program management decisions.

Ideas and Options for Expanding Use of Compliance Rates

In FY 2006 OECA contracted with a statistical consultant, NuStats, to review our current methodology for developing statistically-valid compliance rates and provide recommendations for ways to develop compliance rates in a more economical fashion in order to expand their use. NuStats offered the recommendations summarized below as having potential to help expand the enforcement and compliance program's use of compliance rates.

Near-Census Rates – Statistically-valid compliance rates can be developed for a sector when the inspection coverage in a given year reaches near-census levels; i.e., 80% inspection coverage or better. OECA reviewed programmatic inspection coverage requirements and developed a list of candidate sectors or statutes where the inspection coverage may be high enough to develop a near-census compliance rate. After reviewing the actual inspection coverage (state and federal combined) four areas showed promise for developing this type of rate. They are: Clean Air Act Title V major inspections, RCRA 3007 (d), RCRA 3007(e) – TSDF facilities, and Federal Facility TSDF inspections.

OECA also reviewed other Agency programs to see if they are developing compliance rates. The drinking water and Underground Storage Tank (UST) programs are developing compliance rates based on self-reported data, but OECA has not evaluated the statistical validity of the rates.

Near Census + Modest Investment – For sectors where the inspection coverage level is close to, but not at census levels, an additional increment of random inspections can be added to develop a statistically-valid compliance rate. The number of additional random inspections would be much lower than if a rate was developed only using random inspections.

Multi-Year Rates – In order to spread the cost of developing rates from a random sample the inspections can be spread over more than one year, and OECA has piloted this type of rate in the past. The distinction between a rate based on random inspections done in a year or less, and a multi-year rate is that the multi-year rate is an average compliance rate over the years the inspections were conducted. The multi-year strategy can be used in combination with others, such as the near-census rate. In fact, the Clean Air Act Title V rate mentioned above in the near census discussion is a two-year rate.

The obvious downside to an average compliance rate over a number of years is that as the time period gets longer, there is less confidence that the average represents the current state of compliance. This decreases the utility of the rate as a tool to manage the program and make resource decisions.

Regional & State Rates – Though not in the NuStats report, another opportunity to reduce the burden of developing statistically-valid compliance rates is to limit their scope and focus on a universe that is not national. For example, rates could be developed for an individual or a small subset of regions or states. Narrowing the universe in this way reduces the number of inspections that must be conducted to develop a compliance rate. This strategy could also be combined with those mentioned above.

These four approaches provide some relief from the resource burdens (e.g., reduced number of targeted inspections, effort necessary to conduct random inspections) imposed by the development of statistically-valid compliance rates, and might be helpful in expanding the use of such rates.

Given these possible approaches, what are the options for expanding the enforcement and compliance program's use of compliance rates in the future?

Strategic Compliance Rate Development – Compliance rates should only be developed after a clear management need has been identified and a plan for how and by whom the compliance rate will be used. When a decision is made to develop a rate for strategic purposes the recommendations listed above should be employed to reduce the burden of developing the rate.

OECA's national priorities are strategically significant to the enforcement and compliance program and represent an obvious area where compliance rates could be applied. The national priority strategy implementation teams are working to develop compliance rates for those strategies where there is a specific compliance goal, or it will assist with implementation of the strategy. Currently, rates are planned for the CSO and Mineral Processing priorities, and possible rate development is being considered for the Stormwater, Air Toxics, and portions of the Tribal priorities.

Compliance Correlation with Human Health and Ecological Benefits – Conducting analysis to identify the correlation between compliance with a given environmental requirement and the potential for yielding human health and environmental benefits may assist with prioritizing areas where compliance rates should be developed. Compliance with environmental requirements varies with regard to its potential to produce human health and environmental benefits (i.e., compliance with a requirement to label a hazardous waste will not produce the same benefits as compliance with a CAA regulation that requires the installation of pollution control equipment). Once sectors or environmental regulations correlated with the greatest potential to yield human health and environmental benefits are identified they could be used to prioritize where statistically-valid compliance rates should be developed.

Compliance Volatility Characterization - Additional research to characterize the degree to which compliance is volatile for a given population would enhance the use of compliance rates. For regulated populations where compliance is highly volatile the inspections should be conducted over a short time period to ensure that the resultant rate is as representative as possible (i.e., multi-year rates would not be recommended).

Populations with volatile compliance status also argues for there being an immediate use for the rate once it is available, since the rates ability to represent actual compliance in the regulated population will decay rapidly with time.

Knowing that compliance in a regulated population is less volatile suggests different approaches. Populations with more stable compliance patterns are better candidates for multi-year rates since the average compliance will be more representative than in a volatile compliance population. In addition, if it is known that compliance is more stable, then compliance rates can be updated less frequently, which will save resources.

SECTION IV. FINDINGS

A. Outcome Measurement in State Compliance and Enforcement Programs

1. Although state environmental agencies have developed useful, and, in some cases, sophisticated combinations of outcome measures to assess overall environmental conditions in their respective states, their measurement of outcomes from environmental compliance and enforcement programs is not very advanced.

Discussion: Many states have been using outcome indicators to assess and report about environmental conditions in their states for a number of years. However, states have not made a similar effort to develop and use outcome measures to assess the performance of their enforcement and compliance programs. Instead, they are using basic output measures as their principal performance measures for their compliance and enforcement programs. This may be partly due to the fact that in most state environmental agencies, compliance and enforcement programs are built into the structure of the air, water, and waste programs rather than consolidated into one compliance and enforcement program.

2. No state environmental agencies are currently using sampling approaches to develop representative, statistically-valid compliance rates. The use of statistically-valid rates by state environmental agencies is occurring only in the circumstance when 80% or more of a regulated population is inspected, thereby creating a rate based on a "near census."

Discussion: Developing statistically-valid rates through sampling approaches generally relies on the use of on-site inspections to gather data about compliance at facilities. The number of inspections conducted by states is limited by resource constraints, and states largely target those inspections at facilities where there is reason to believe (from current or historical information) there might be noncompliance. States have been unwilling to trade targeted inspections for the randomly-selected inspections necessary to develop statistically-valid rates.

3. No state environmental agency is currently using, or planning to use, a "universal compliance rate" that aggregates data about all violations in the entire regulated universe under one rate.

Discussion: Our review of the ECOS report, the work of the Compliance Consortium, and our discussions with the ECOS Compliance Committee and the OIG identified no current use of or future support for a universal rate. There are significant practical difficulties with developing a consolidated compliance rate from among multiple statutes, with numerous distinct programs under each statute, myriad requirements under each program, applied to segments of a large and diverse regulated universe. These practical difficulties far outweigh the value that would be produced by using an overall compliance rate as a management tool for an enforcement and compliance program.

4. In those instances in which state environmental agencies are using compliance rates not based on a "near census", these rates are based on only the inspected portion of the population, are biased since the inspections are targeted at problem facilities, and are used primarily as a targeting tool rather than a valid measure representing the actual state of compliance in regulated populations.

Discussion: States are making use of compliance rates that are limited by their bias toward targeted inspections where noncompliance is likely to be found. Given these limitations, the use of rates seems to be primarily for internal management purposes to help identify where additional compliance assistance or enforcement might be needed (the "program targeting" use described in the ECOS report), and to determine whether inspections strategies are succeeding in finding noncompliance.

5. Among state environmental agencies, the use of outcome measures to assess and improve performance or to report to the public is very rare.

Discussion: There are very few examples of state environmental compliance and enforcement programs developing and using outcome measures to manage their programs, improve program performance, or report to the public. Our research indicated that only four states, Ohio, Minnesota, New Hampshire, and Texas have developed end outcome measures. Each of these states' attempts to capture the environmental benefit of their enforcement actions in a manner similar to EPA's measurement of pounds of pollution reduced. Our research did not identify any state currently using, or planning to use, recidivism measures.

B. Outcome Measurement in EPA's Compliance and Enforcement Program

1. Development of a universal compliance rate for the nation would not meet two criteria important in the selection and implementation of performance measures – value and feasibility – and would not meet the GAO test of "balancing ideal performance measurement systems against real-world considerations such as the cost and effort involved in gathering and analyzing data."

Discussion: EPA's national compliance and enforcement program is responsible, along with the states, for maximizing compliance with 12 environmental statutes, 28 distinct programs under those statutes, and dozens of regulatory requirements under those 28 programs which apply in various combinations to a universe of 40 million regulated entities. As a compliance rate takes on more aggregation, it is subject to more influences, and less able to provide specificity or even useful information. Year-to-year comparisons could mask major swings in portions of the universe. Minor but widely applicable requirements could drive the direction of the rate. Moreover, the cost of developing and implementing the sampling methodology necessary for a universal rate would be prohibitive both in terms of dollars and the number of targeted inspections sacrificed to conduct randomly-selected inspections. While a universal rate might provide a crude device in making high-level evaluative judgments based on one program outcome, it

would not be specific enough to be useful in making management decisions about program operations.

2. The use of approaches such as near census, near census plus modest investment, multi-year rates, and regional or state rates for specific industry sectors can be used to expand development and use of meaningful compliance rates.

Discussion: These approaches to developing compliance rates, described in Section III, can enable EPA to apply compliance rates to a larger number of noncompliance patterns, regulated populations, and environmental problems. Efficiencies offered by these approaches should help relieve some of the resource burden associated with producing compliance rates.

3. OECA's current GPRA architecture, built around the four primary tools of the compliance and enforcement program (assistance, incentives, monitoring and enforcement) does not fully demonstrate how program outputs and outcomes are contributing to the reduction or elimination of environmental problems.

Discussion: Measuring the intermediate and end outcomes of assistance, incentives, monitoring, and enforcement enables OECA to roll up national totals for all of its activities in the 28 distinct compliance and enforcement programs for which it is responsible. But these measures do not reveal whether important environmental problems are being reduced or eliminated, a shortcoming shared also by compliance rates. Moving to a problem-oriented approach which focuses on environmental risks and noncompliance patterns would require development of a new objective, sub-objectives, and performance measures tailored to each environmental problem.

4. Deterrence research holds promise for compliance and enforcement practitioners because it can help illuminate the factors that contribute to compliance and improve understanding about which tools or combination of tools produces the most effective results.

Discussion: OECA has often described its compliance and enforcement programs as providing both specific and general deterrence. Inspections and other forms of compliance monitoring and enforcement are undertaken not only to identify specific violators and return them to compliance, but also to deter all other similarly situated regulated entities from future noncompliance. OECA would benefit from an expanded effort to conduct research to better understand the relationship between its activities and the behavioral responses of the regulated universe.

5. OMB's and EPA's preferences regarding performance measures for the compliance and enforcement program are not in alignment because the two agencies use measures for very different purposes.

Discussion: In his article entitled, "Why Measure Performance? Different Purposes Require Different Measures," Robert Behn of Harvard University points out that there are eight different purposes for measuring programs. For OMB, the purpose of measuring performance is evaluative, i.e., to make high-level judgments about program performance and decisions about resource allocations within the EPA budget. For OECA program managers, the purpose of measuring performance is to monitor and control operations, manage resources wisely, motivate personnel, learn what is working or not working, and improve program effectiveness. These two very different sets of purposes tend to lead to divergent views of the value of individual measures and the appropriate way to judge overall program performance.

C. Outcome Measurement in Other Federal Programs

The review of federal regulatory programs did not identify performance measurement or compliance rate practices that were applicable and useful for managing EPA's compliance and enforcement programs.

Discussion: Only 3 of the 25 programs reviewed had measures that could be categorized as end outcomes. Two of these were non-enforcement programs at EPA and the third was a program involving a small universe of DOI managed watershed lands. Review of nine federal enforcement programs found two compliance rates being used by the Department of Homeland Security and the Corps of Engineers, both of which were based on inspections that were not designed to produce a statistically-valid rate. Examples of purportedly statistically-valid rates identified by the EPA OIG were actually one-time evaluations that employed statistical techniques but were not producing regular compliance rates.

SECTION V. RECOMMENDATIONS FOR IMPROVING OECA OUTCOME MEASURES

A. Using Compliance Rates

Based on the research for this report and OECA's experience developing and using statistically-valid rates over the last seven years, it is clear that such rates can produce valuable information. When statistically-valid rates require sampling and random inspections, they can demand considerable effort and resource investment, and require sacrificing inspections targeted at known or suspected violators.

In spite of these concerns, OECA believes it can be beneficial to expand the development and use of statistically-valid rates for the appropriate purposes. Rates that focus on specific programs or noncompliance patterns are worth the investment if they are selected carefully and developed efficiently.

Recommendation: OECA will expand development of statistically-valid compliance rates for specific noncompliance patterns, driven by a clear management need that can be served by the use of a compliance rate. Development of rates will focus on national priorities or other important problem areas such as those identified as part of the transition to the problem-oriented approach. (See the recommendation in Section F below.) In addition, in choosing whether to develop a rate, OECA will weigh the relative importance and value of the rate and the resources and effort necessary to produce the rate.

Regarding the development of a universal rate, OECA believes that a universal rate would not be a useful tool for managing the national program. It is very telling that no state or federal enforcement program is either using or developing a universal compliance rate.

Recommendation: OECA does not recommend development of a universal compliance rate. The cost and effort involved in developing and using a statistically-valid, universal compliance rate for the nation far outweighs the likelihood of any benefit such a rate might bring to program management and improved performance.

B. Characterizing Pollutant Reductions by Hazard and Exposure

As part of its 2004 Program Assessment Rating Tool (PART) review the Civil Enforcement Program submitted a Measures Implementation Plan (MIP) to improve the current pollutant reduction measure by adding a characterization of hazard and exposure. In addition to adding information about human health risks, the MIP also seeks to qualify pollution reduction measures with ecological indicators that reflect damage to the environment and provide a more accurate picture of the total effect of each pollutant.

The pollutants reduced measure includes hundreds of different pollutants in air, water, and land media. The measure is reported in pounds of pollutants reduced. OECA chose to begin work on the MIP by exploring hazard and exposure calculation approaches for air pollutant reductions. There were two reasons for this. First, of all the pollutants reduced in FY2004, air pollutants accounted for 76% of the total for all media (land, water, and air). And second, the methodologies and modeling for air pollutants are the most advanced at the national level. OECA began working with the Office of Air and Radiation (OAR) to model the environmental and health benefits associated with our air enforcement cases. Using FY 2005 case data, OAR provided an estimate of billions of dollars in health benefits resulting from OECA's ten largest air cases. See Appendix 4 for a summary of the results of this exercise. This estimate was calculated by extrapolating from existing Clean Air Interstate Rule (CAIR) BenMAP (OAR's benefits model) results.

OMB has stated that the OECA's Civil Enforcement Program's primary outcome measure "needs to be characterized as to risk, so that enforcement can focus on the most harmful pollutants first." In effort to identify the "most harmful" pollutants, OECA summarized the human health and ecological impacts of each of the top ten air pollutants reduced in 2005. Together, the top ten pollutants account for 99% of the air pollutants reduced. An example of a characterization for one air pollutant, nitrogen oxide, is provided in Appendix 5. Each pollutant has been characterized as to the health effects (respiratory, cardiac, neurological, reproductive/developmental, or carcinogenic) and ecological effects (impact on vegetation, contribution to acid rain, water quality, greenhouse gases, smog, ozone depletion, toxic to animals, visibility, bioaccumulative, and impact on the built environment).

To use this type of characterization to focus the compliance and enforcement program on the most harmful pollutants, progress will need to be made in answering difficult questions. How do you define most harmful? Is a pollutant more harmful if it:

- Has greater acute toxicity? e.g., dioxin would be prioritized over nitrogen oxides.
- Bioaccumulates, or persists in the environment without degrading for a longer period of time?
- Contributes to the most prevalent health problems? Any pollutant contributing to asthma should be reduced first?
- Is released from a facility in an area with high population density? This would increase the public's exposure.

In addition, how should a pollutant that degrades the environment, but has less direct effect on human health be prioritized? How does ecological harm compare to harm to human health?

By accessing the OAR model to assess the human health impacts (and resulting economic burden) avoided due to enforcement actions, OECA is making use of a peer-reviewed model, the best the Agency has to offer. However, it is important to understand

that the model's results can be influenced by many forces from year to year (e.g., fluctuations in number and type of enforcement actions, changes in factors used in the model). These forces may preclude the use of the model's results for setting annual performance targets for the compliance and enforcement program.

Recommendation: OECA will continue making progress on characterizing air pollution reductions from enforcement cases by working with OAR to expand the use of their BenMAP model on OECA's air enforcement case data. The model estimates the health benefits of pollutant reductions and reports those benefits in terms of reductions in premature mortality, chronic and acute bronchitis, myocardial infarctions, hospitalizations, respiratory symptoms, and/or work loss days. OECA will continue to explore how pollutant characterization might be used to target program activities on the most harmful pollutants, especially as part of the transition to the problem-oriented approach. (See Section F below.)

OECA will provide reduction quantities for pollutants and facility locations for its air cases. BenMAP then combines air pollution monitoring data, air quality modeling data, census data, and population projections to calculate a population's potential exposure to ambient air pollution. Given these exposure estimates, BenMAP calculates the associated change in health effect incidence using health impact functions derived from the epidemiological literature.

Can OECA use an approach similar to the air modeling with OAR for water and land/soil pollutant reductions? Assessing hazard and exposure in soil and water is more complex than in air. There are models that are used locally for water and soil contamination assessments. With air, estimates can be made assuming certain distributions in the environment and acknowledging that all people are exposed to ambient air. For pollutants in water or soil, modeling would need to take into account the properties of the pollutants themselves and their fate and transport in the environment, includes well as the geology and hydrology of each site. The potential for national estimates from prospective soil or water models will be slower to develop.

C. Use of Recidivism Rates as a Performance Measure

Recidivism is defined as "a tendency to relapse into a previous condition or mode of behavior; *especially*: relapse into criminal behavior." OECA first developed a recidivism measure for the Civil Enforcement Program in 1998. OECA developed three recidivism measures that focused on two-year recidivism rates for Significant Noncompliance (SNC) for the Clean Water Act (CWA) and Resource Conservation and Recovery Act (RCRA), and High Priority Violators (HPV) in the Clean Air Act. The three measures tracked recidivism at the facility level, and a facility was counted as a recidivist if it returned to SNC or HPV status within two years of returning to compliance from that status. OECA tracked these measures from 1999 through 2004, at which time the measures were discontinued due to a number of concerns with data quality, and utility of the measure with regard to program management.

Over the last two years, OECA has been utilizing another measure which might be modified to serve the purpose of a recidivism measure, i.e., to track and reduce the number of repeat offenders. OECA has gathered data about "chronic non-compliers," regulated facilities that have been in some form of noncompliance for five or more of the previous ten quarters. The purpose for gathering this information has been to ensure that facilities are not languishing in noncompliance without a federal or state response (e.g., an enforcement action) to move them back into compliance. Data about chronic non-compliers might be developed into a measure to track facilities by amount and type of noncompliance.

Recommendation: By September 2006, OECA will complete an analysis of its experience with the recidivism measure used from 1999-2004 and the chronic non-compliers data it has developed in the last two years. The re-evaluation will make recommendations about using one of those approaches for a pilot project to be conducted in FY 2007 to develop a repeat offenders measure for the civil enforcement program.

OECA's Criminal Enforcement Program is developing a recidivism measure in response to the Program's FY 2004 PART review. The definition of recidivism for this measure is:

> A recidivist is any "defendant" (individual, facility or business entity) who, having been found guilty (i.e., by either guilty plea or trial) of a criminal violation in an EPA/CID investigation, is *the object of a subsequent enforcement action* by EPA's civil or criminal program. Defendants are tracked for a period of five (5) years from the time of commencement of the instant offense (i.e., tracking the sentencing guidelines)

The Criminal Enforcement Program is on track to develop a baseline and set a target for this measure by October 2006.

D. Compliance Assistance Outcomes

EPA provides two types of compliance assistance: 1) direct assistance is provided by EPA personnel through workshops, training programs, and distribution of guidance documents; and 2) assistance is also provided through web-based online assistance centers, each of which is organized to meet the assistance needs of specific industry sectors.

For the last several years, EPA has counted the number of regulated entities reached through these two delivery mechanisms. For example, in FY 2005, two million regulated entities received compliance assistance from EPA. More recently, EPA has measured three types of intermediate outcomes resulting from compliance assistance delivered through its on-line assistance centers: increased understanding of regulatory requirements; implementation of improved environmental management practices; and

reduction of pollution. In FY 2005, 84% of compliance assistance center users reported increased understanding of regulatory requirements, 78% reported improved management practices, and 46% reported reducing pollution as a result of compliance assistance received from EPA. Ninety-one percent of recipients of direct compliance assistance reported increased understanding of regulatory requirements, 51% reported improved environmental management practices, and 13% reported that they reduced pollution as a result of direct compliance assistance received from EPA.

These intermediate outcome measures have provided valuable insight into the results achieved by compliance assistance, and have helped guide EPA's efforts to design and deliver effective assistance programs. Our review of measurement practices in state and federal compliance assistance programs indicates that no other compliance and enforcement program has advanced as far as EPA in measuring outcomes of compliance assistance. Nevertheless, OMB has raised issues about the value and validity of these measures, requested that EPA address these issues, and has further urged that these measures be removed entirely from EPA's set of publicly-reported performance measures.

EPA believes these measures should continue to be used by EPA and reported to the public on an annual basis. EPA shares OMB's interest in improving the validity of these measures. For example, EPA hopes to be able to ensure that a representative sample of assistance recipients is used to measure with more certainty the results achieved by compliance assistance.

Recommendation: Guided by the results of a feasibility study in FY 2007, EPA will conduct a survey every three years of a statistically-valid sample of compliance assistance recipients to measure behavior changes resulting from compliance assistance. The survey would begin in 2008 and would be used as both a management tool to guide and improve compliance assistance programs as well as a publicly-reported measure of program performance.

E. Measuring the Deterrent Effect of Enforcement

Law enforcement agencies have long believed that their actions have an impact beyond the person or entity that is the subject of an enforcement action. Beyond specific deterrence that identifies individual violators, returns them to compliance, and motivates (deters) continuing behavior, there is the strongly-held view among law enforcers that individual enforcement actions also bring about general deterrence of violations among other entities subject to the same or similar laws or requirements. There is a significant body of literature about deterrence theory which examines the forces that seem to influence compliant and non-compliant behavior, and there are now a number of studies documenting and analyzing the deterrent, motivational, and performance-related effects of compliance monitoring, enforcement, technical assistance, incentives and other government interventions, market forces, community pressure, and other factors such as corporate policy, organization, and compensation.

The conventional wisdom is that the effectiveness of a legal threat depends on three related factors: 1) the likelihood that a lawbreaker will be caught; 2) the nature and severity of the punishment; 3) and the speed of apprehension and punishment. It is believed that increases in each dimension correlate with increased deterrence and compliance. Arguably, a fourth factor would be the lawbreaker's perception of the first three.

Compliance and enforcement managers in OECA would benefit greatly from better understanding the relationship among these factors and their applicability to varying types of requirements and regulated entities. Knowing more about the interplay of these factors might lead to both subtle and significant changes in how assistance, incentives, monitoring, and enforcement are targeted to address specific environmental risks and noncompliance patterns.

Recommendation: EPA will begin formulating a research approach to learn more about the deterrent effect and how it might be measured in the context of OECA's work. Among the possible options for this research, OECA will review the feasibility of a periodic survey of representatives of regulated facilities and industries to better understand what factors motivate compliance behavior and what combinations of tools and practices would maximize such behavior.

F. Moving from Tool-Oriented to Problem-Oriented Objectives and Measures

As described in Section I, the current GPRA architecture (shown in Appendix 1) builds sub-objectives around and therefore measures specific intermediate outcomes produced by the major activities of the national program (i.e., assistance, incentives, monitoring and enforcement). One major advantage of this tool-based approach is that it provides the ability to aggregate outputs and outcomes across all 28 distinct compliance and enforcement programs to produce national totals. The tool-based approach reveals how a particular tool was used to produce an intermediate outcome, a set of intermediate outcomes, or an end outcome.

A disadvantage of the tool-based approach is that it does not reveal whether particular environmental problems (i.e., an environmental risk or a noncompliance pattern) were addressed by the activities of the national program. The reliance on the tool-based approach has resulted in recent criticism from the GAO. They stated:

> "EPA needs to … intensify its efforts to move from a performance management system focused on costly and often unproductive activities towards a system focused on achieving measurable improvements to the environment…" (GAO letter to Lyons Gray, EPA's Chief Financial Officer, April 26, 2006)

Moving from a tool-based approach to a problem-oriented approach would have certain benefits. First, it would focus a larger portion of the resources and activities of the program on important national risks and noncompliance patterns. Second, a problem-

oriented approach would produce an account of program performance which could be presented to the public and others in more compelling language of problem identification and remediation.

Specific challenges would have to be overcome to move toward a problem-oriented framework. Performance measures associated with a specific problem would need to be tailored to that problem. This would require development of new measures to assess progress in addressing each problem. Because these measures would vary from one problem to the next, they could not be aggregated into national trends.

Recommendation: By November 30, 2006, OECA will develop an implementation plan for transitioning to a problem-oriented approach for measuring outcomes of the national compliance and enforcement program. By developing a plan, OECA will be able to conduct outreach and consultation with regional and headquarters compliance and enforcement managers, state co-regulators, academic experts, and stakeholders. The plan would identify tasks, costs, and schedules for at least the following five stages:

1. *Identification of problems.* Problems (i.e., environmental risks and noncompliance patterns) would be identified from four sources: OECA's own national priorities; the objectives and sub-objectives of the air, water and waste goals of the EPA Strategic Plan; key environmental conditions assessed in EPA's Report on the Environment; and multi-regional priorities identified through the "eco-region" effort.

2. *Evaluation and characterization of problems.* A list of candidate problems would be developed. Risk characterization of pollutants would be a criterion in choosing problems, as would the "plausible attribution" link between OECA activities and the reduction or elimination of problems.

3. *Development of problem-specific measures.* Performance measures for the problems selected would be developed and implemented. For some problems there would likely be specific goals against which to measure, for others there might be measures to monitor the OECA contribution to addressing the problem.

4. *Implications for data systems.* For each measure, an analysis of the need for new data collection and reporting would be conducted. Changes and enhancements to data systems would need to be identified and their costs estimated.

5. *Development of GPRA architecture.* The problems selected would then be translated into an objective and a set of sub-objectives that would eventually supplement the current tool-oriented GPRA architecture.

References

Behn, Robert D., "Why Measure Performance? Different Purposes Require Different Measures," *Public Administration Review*, September/October 2003, Vol. 63, No. 5, p. 586-606.

Environmental Council of the States. State Environmental Agency Contributions to Enforcement and Compliance 2000-2003. Final Review Draft, April 2006.

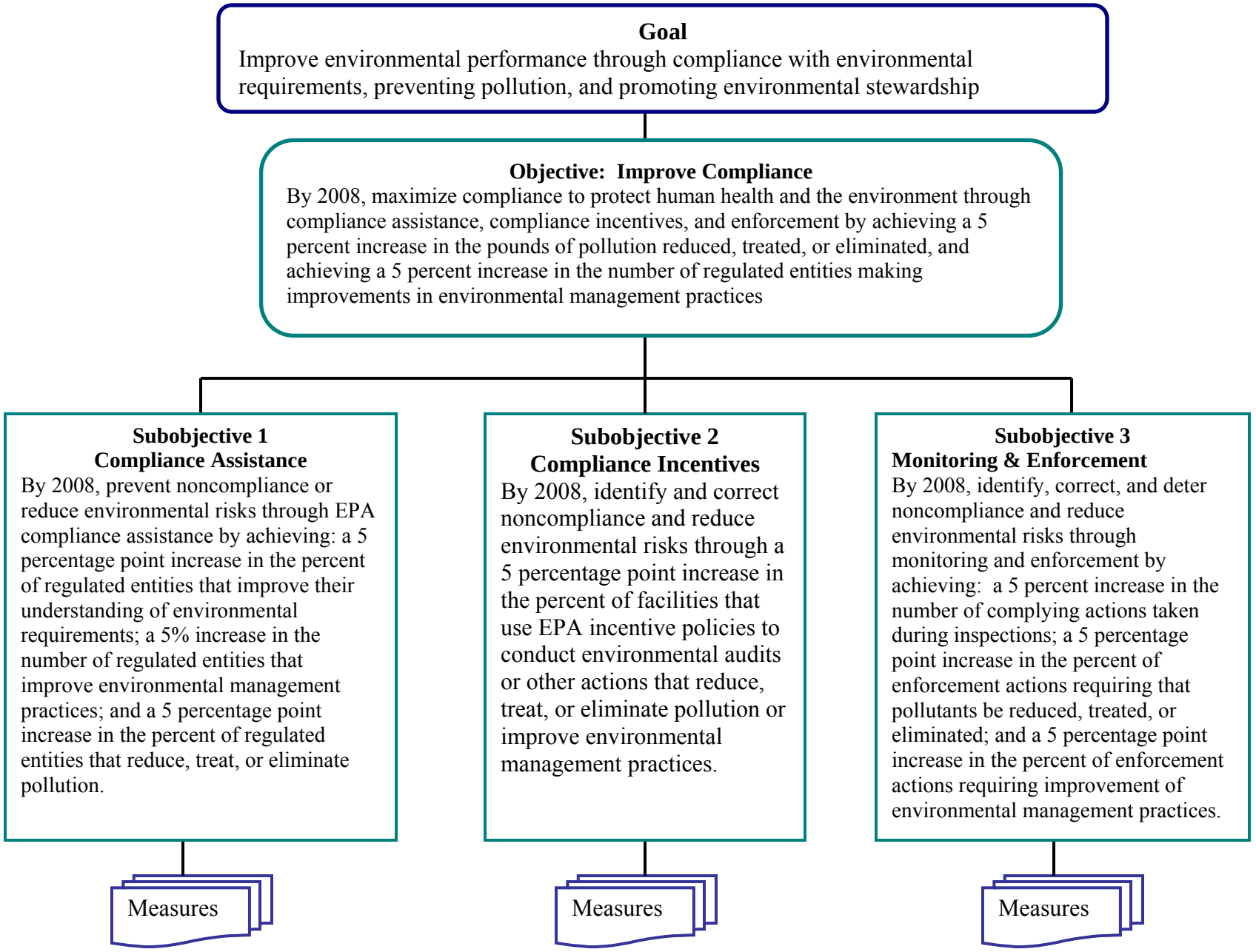
Office of Management and Budget, Examples of Performance Measures. http://www.whitehouse.gov/omb/part/performance_measure_examples.html

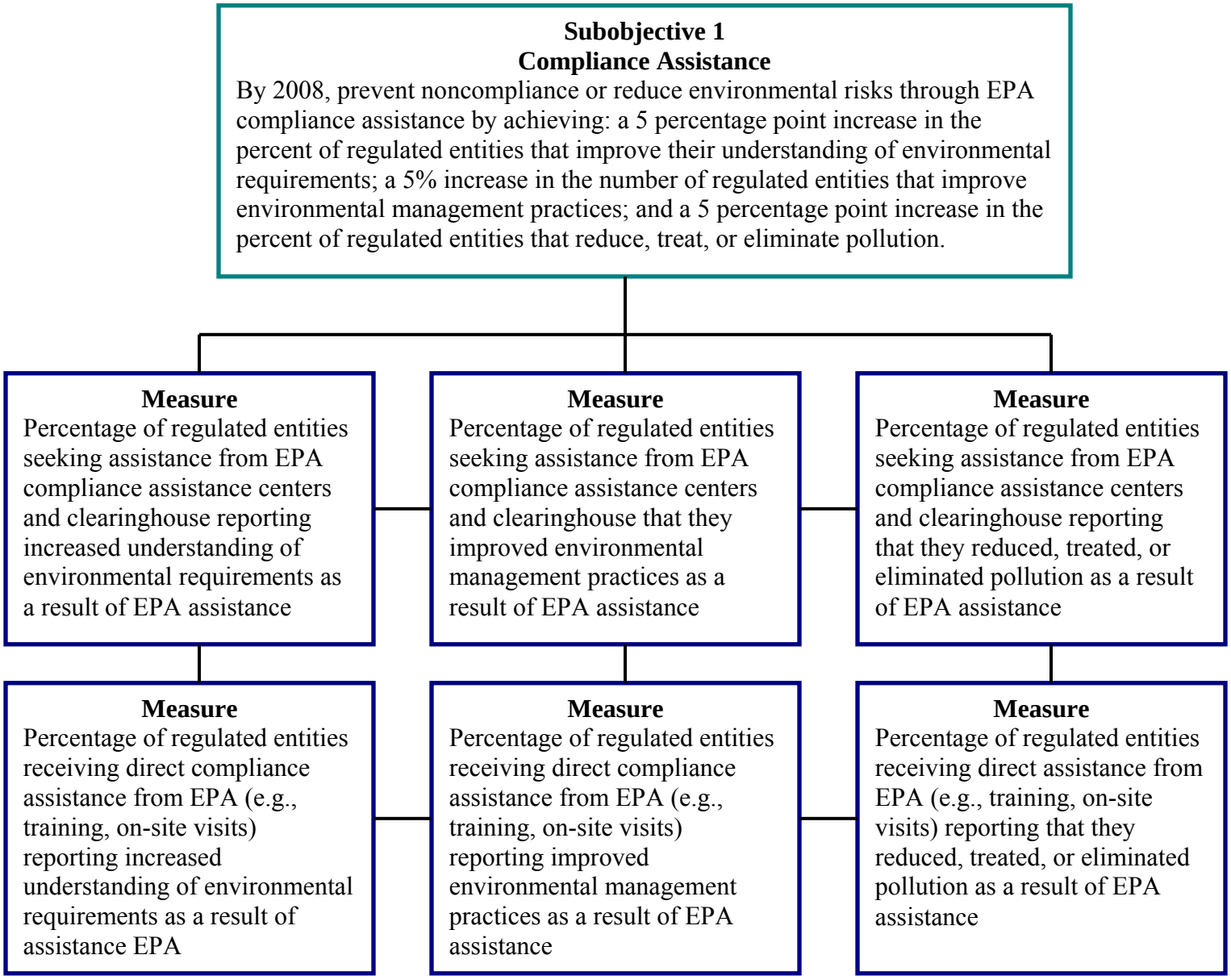
Silberman, Jonathan, "Does Environmental Deterrence Work? Evidence and Experience Say Yes, But We Need to Understand How and Why, "in *Environmental Law Reporter*, Vol. 30, 10523, 2000.

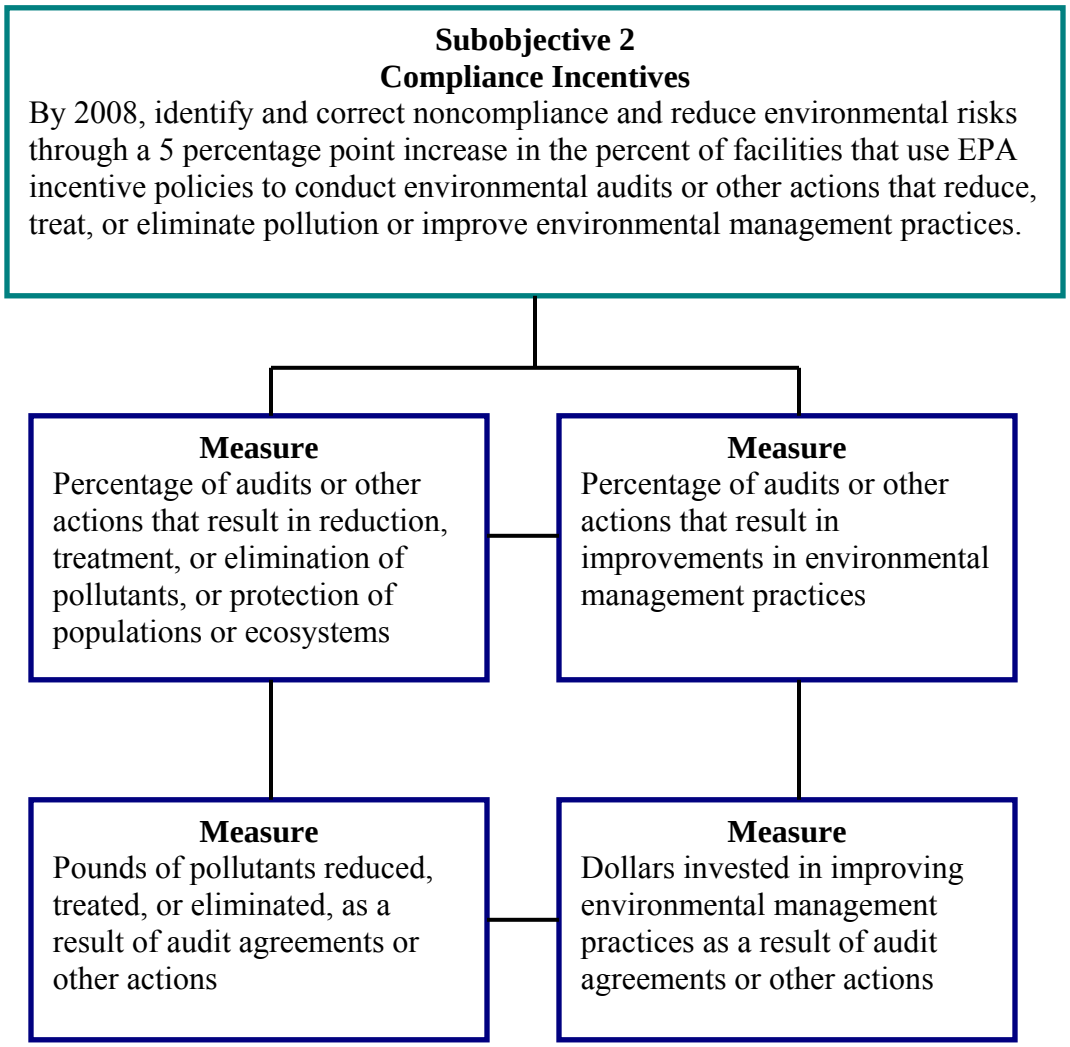
Sparrow, Malcolm. *The Regulatory Craft*. The Brookings Institution Press, 2000.

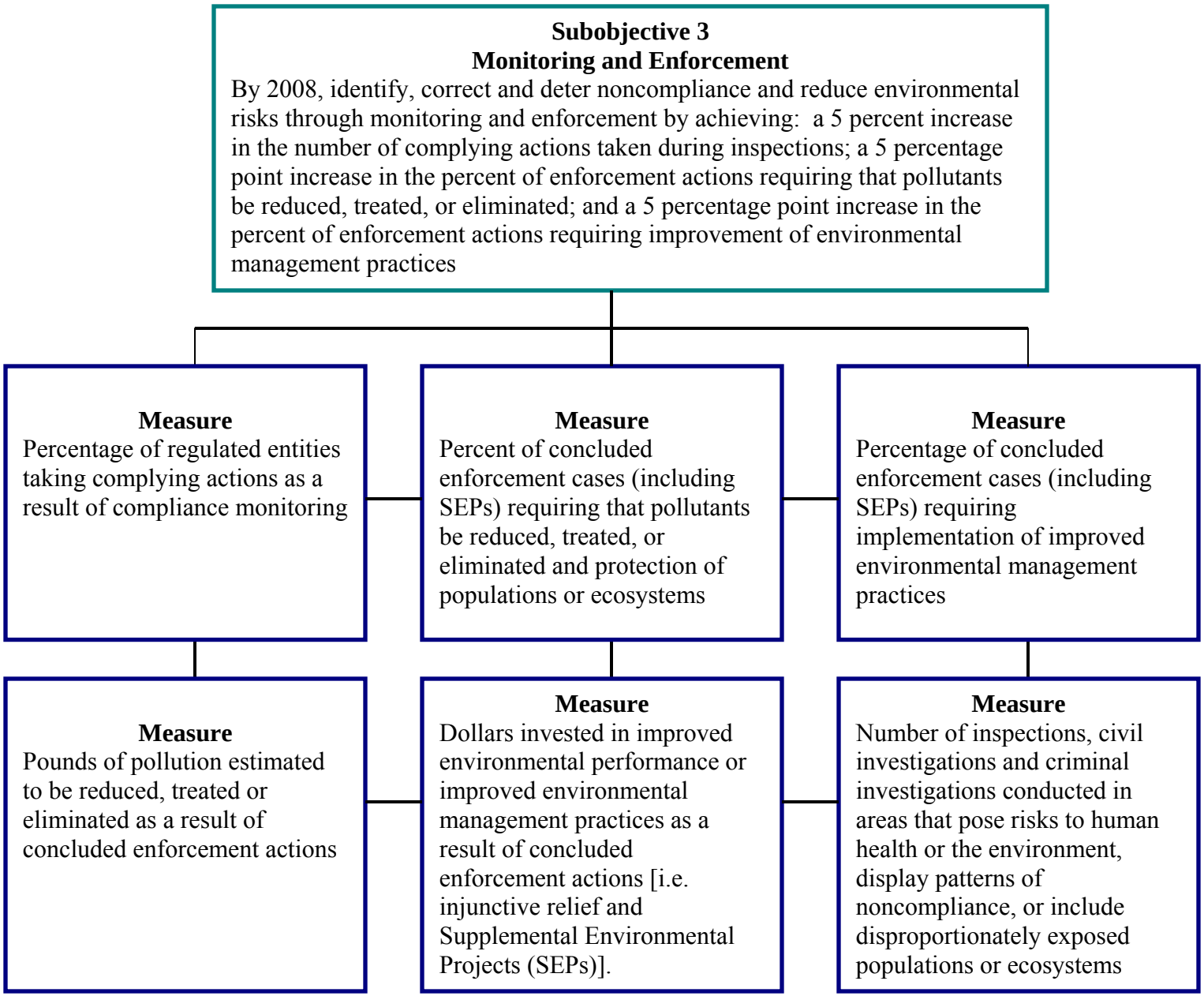
U.S. Government Accounting Office. Executive Guide—Effectively Implementing the Government Performance and Results Act. GAO-GGD-96-118. June, 1996.

Appendix 1. OECA Strategic Goal Architecture




Subobjective 1
Compliance Assistance
By 2008, prevent noncompliance or reduce environmental risks through EPA compliance assistance by achieving: a 5 percentage point increase in the percent of regulated entities that improve their understanding of environmental requirements; a 5% increase in the number of regulated entities that improve environmental management practices; and a 5 percentage point increase in the percent of regulated entities that reduce, treat, or eliminate pollution.
Measure
Percentage of regulated entities seeking assistance from EPA compliance assistance centers and clearinghouse reporting increased understanding of environmental requirements as a result of EPA assistance
Measure
Percentage of regulated entities seeking assistance from EPA compliance assistance centers and clearinghouse that they improved environmental management practices as a result of EPA assistance
Measure
Percentage of regulated entities seeking assistance from EPA compliance assistance centers and clearinghouse reporting that they reduced, treated, or eliminated pollution as a result of EPA assistance
Measure
Percentage of regulated entities receiving direct compliance assistance from EPA (e.g., training, on-site visits) reporting increased understanding of environmental requirements as a result of assistance EPA
Measure
Percentage of regulated entities receiving direct compliance assistance from EPA (e.g., training, on-site visits) reporting improved environmental management practices as a result of EPA assistance
Measure
Percentage of regulated entities receiving direct assistance from EPA (e.g., training, on-site visits) reporting that they reduced, treated, or eliminated pollution as a result of EPA assistance


Subobjective 2
Compliance Incentives
By 2008, identify and correct noncompliance and reduce environmental risks through a 5 percentage point increase in the percent of facilities that use EPA incentive policies to conduct environmental audits or other actions that reduce, treat, or eliminate pollution or improve environmental management practices.
Measure
Percentage of audits or other actions that result in reduction, treatment, or elimination of pollutants, or protection of populations or ecosystems
Measure
Percentage of audits or other actions that result in improvements in environmental management practices
Measure
Pounds of pollutants reduced, treated, or eliminated, as a result of audit agreements or other actions
Measure
Dollars invested in improving environmental management practices as a result of audit agreements or other actions


Subobjective 3
Monitoring and Enforcement
By 2008, identify, correct and deter noncompliance and reduce environmental risks through monitoring and enforcement by achieving: a 5 percent increase in the number of complying actions taken during inspections; a 5 percentage point increase in the percent of enforcement actions requiring that pollutants be reduced, treated, or eliminated; and a 5 percentage point increase in the percent of enforcement actions requiring improvement of environmental management practices
Measure
Percentage of regulated entities taking complying actions as a result of compliance monitoring
Measure
Percent of concluded enforcement cases (including SEPs) requiring that pollutants be reduced, treated, or eliminated and protection of populations or ecosystems
Measure
Percentage of concluded enforcement cases (including SEPs) requiring implementation of improved environmental management practices
Measure
Pounds of pollution estimated to be reduced, treated or eliminated as a result of concluded enforcement actions
Measure
Dollars invested in improved environmental performance or improved environmental management practices as a result of concluded enforcement actions [i.e. injunctive relief and Supplemental Environmental Projects (SEPs)].
Measure
Number of inspections, civil investigations and criminal investigations conducted in areas that pose risks to human health or the environment, display patterns of noncompliance, or include disproportionately exposed populations or ecosystems

Appendix 2. Questions for Discussions with Academic Experts

Academic experts in the field of performance measurement for public programs were consulted for their views on performance measurement for enforcement and compliance activities, particularly the value and feasibility of compliance rates and higher-order outcomes. Individuals were selected who are well-known and highly respected among leaders in their field of expertise.

Background materials were provided in advance of each meeting to provide context for discussion, including a summary of this study and its purpose, OECA's current goals and measures, an example accomplishments report using current measures, a short paper describing OECA's efforts to develop and use statistically-valid compliance rates, a hierarchy of performance measures developed by GAO and used by OMB, and a set of questions for discussion.

Questions for discussion included:

1. OMB has been urging EPA's compliance program to develop more outcome measures in categories 3-6 (i.e., end outcome measures) of the GAO's performance measurement hierarchy.
   What are your views about whether such measures are appropriate for a compliance program?

2. OMB has been urging EPA's compliance program to develop a statistically-valid compliance rate that characterizes compliance of the entire regulated universe subject to federal environmental requirements.
   What are your views about the feasibility and utility of an overall compliance rate?

3. EPA has been developing statistically-valid compliance rates for selected industry sectors to characterize their compliance with specific regulatory requirements.
   Do you have any suggestions about how EPA can improve and expand its use of compliance rates as a performance measure?

4. EPA has been developing and using intermediate outcome measures (e.g., behavior changes at regulated facilities) to assess program performance.
   What are your views about the advantages and disadvantages of using intermediate outcomes to measure performance?

5. EPA has been using a combination of output and outcome measures to manage its compliance program and report annually to the public about program accomplishments.
   What are your views about using both output and outcome measures (as opposed to relying solely on outcomes) for these purposes?

6. Given EPA's current set of performance measures for its compliance program and OMB's desire for the use of more end outcome measures, can you suggest any outcome measures which would be both valuable and feasible?

Appendix 3. State Compliance Rate Calculation Method, Compliance Rates, and Reporting Methods

State	Program	Compliance Rate Calculation Method	Is it Statistically-Valid/Census based or other?	States' Use of the Measure	OECA Comments	Source
AK	All	Percent of violations versus inspections conducted in a given fiscal year	Not statistically-valid based on inspected universe	2004 is the first year data was collected to calculate the compliance rate. Could be used to target enforcement efforts.		http://www.ecos.org/section/2006_state_enforcement__report/, http://www.dec.state.ak.us/das/pdfs/enfreport.pdf, http://www.dec.state.ak.us/das/pdfs/enfreport.pdf
CO	Hazardous Waste	Number of facilities with violations/the total number inspected in a given year.	Not statistically-valid based on inspected universe	Use the data to target compliance assistance delivery, enforcement follow up, sector initiatives, and outcome performance indicators.	Not statistically-valid either in sample selection or results. Sample is self-selected and subject to non-response bias.	http://www.ecos.org/section/2006_state_enforcement__report/, Documents sent from CPHE, http://www.cdphe.state.co.us/hm/scorepilotprojectfinal.pdf,
CT	Air	Percent of noncompliant facilities out of number of facilities	Not statistically-valid based on inspected universe			http://www.ecos.org/section/2006_state_enforcement__report/
DE	Air	Percentage of facilities in compliance at time of review	Not statistically-valid based on inspected universe	Compliance rates are a performance measure for the Department's budget review process. Data is included in an annual report to the state legislature, Governor, and public is published.		http://www.ecos.org/section/2006_state_enforcement__report/
FL	No longer produced due to budget cuts				They are exploring a new methodology for developing statistically-valid rates	http://www.ecos.org/section/2006_state_enforcement__report/
IL	Drinking Water	Percentage of the population served by community water supplies receiving drinking water with no short-term (acute) or long-term chronic) adverse health effects	Census-based using on self-reported data	Required by EPA.	This data has been used in our Annual Compliance Reports.	http://www.ecos.org/section/2006_state_enforcement__report/

State	Program	Compliance Rate Calculation Method	Is it Statistically-Valid/Census based or other?	States' Use of the Measure	OECA Comments	Source
IN	Drinking Water	Total number of system with any type of violation during a calendar year versus the total number of active public water system	Census-based using on self-reported data	Based on the data, IDEM is able to shift priorities in terms of identifying what systems are more vulnerable and what are the major types of violations incurred. Most of the violations during each calendar year are related to monitoring and reporting violations.		http://www.ecos.org/section/2006_state_enforcement__report/
LA	Drinking Water	Reported health-based violations of drinking water standards divided by population of community systems in compliance with SDWA.	Census-based using on self-reported data			http://www.ecos.org/section/2006_state_enforcement__report/
MD	All	100% - # with significant violations/ # of INSPECTED sites	Not statistically-valid based on inspected universe.	We believe that this data is only relevant in a per program historical analysis and ONLY in conjunction with other measures. "Compliance rate" is a measure of what we see at the places we get to. Unless you can calculate some measure of what we call "coverage" (how many of the regulated universe do we see) it may not be relevant.	% of inspected sites/facilities in significant compliance. Annual. Required by legislature for 15 programs. Not statistically-significant. Often rates are produced where coverage is 1 or 2%. No definition of significant compliance provided.	http://www.ecos.org/section/2006_state_enforcement__report/

State	Program	Compliance Rate Calculation Method	Is it Statistically-Valid/Census based or other?	States' Use of the Measure	OECA Comments	Source
MA	All From 2004 Enf. Report. Compliance rates used in drinking water, water management act which are both fully reporting populations. SNC rates are used for NPDES.	Number of facilities with documented noncompliance divided by the total number of known regulated facilities	General rate are not statistically-valid and are inspection based. Rates for ERP program are statistically-valid.	We have used compliance rates to evaluate the efficacy of C/E programs, to assess the areas within a regulated sector where the most problematic noncompliance is occurring, and to develop strategies to address significant noncompliance, which can include compliance goal setting, facility targeting and technical assistance	Statistically-valid sample. ERP -- based on key indicators of compliance. To verify results of program to replace permits with self-certification. Program verification and to identify worst non-compliance problems. Non-ERP compliance rate data is based on self-reporting populations	http://www.ecos.org/section/2006_state_enforcement__report/
MI	Hazardous Waste	100% - Number of facilities in noncompliance /the total number of evaluations performed for the period	Not statistically-valid based on inspected universe	The compliance rates are used to view general trends in the compliance/noncompliance. Inspection staff is consulted to determine inspection initiatives.	% of inspected sites/facilities in significant compliance	http://www.ecos.org/section/2006_state_enforcement__report/
MN	Air	((# of facilities evaluated - # of facilities with violations identified in these evaluations)/ # of facilities evaluated)x 100	Not statistically-valid based on inspected universe	These data are used by management to track program performance		http://www.ecos.org/section/2006_state_enforcement__report/, Documents sent from MN PCA
MO	Surface and Ground Water	Quarterly Noncompliance Rate: Number of Major and 92-500 facilities found in violation divided by the total number of Major and 92-500 facilities in state. Annual Noncompliance Rate: Number of Minor Facilities found in violation divided by the total number of Minor Facilities.	Not statistically-valid based on inspected universe	Used to focus enforcement effort on compliance in our work plan.		http://www.ecos.org/section/2006_state_enforcement__report/

State	Program	Compliance Rate Calculation Method	Is it Statistically-Valid/Census based or other?	States' Use of the Measure	OECA Comments	Source
MT	Underground Storage Tanks*	Significant operational compliance per EPA standards for leak prevention, leak detection, and both.	Not statistically-valid based on inspected universe	Used as a federal reporting measures	Required by EPA	http://www.ecos.org/section/2006_state_enforcement__report/
NV	All	Rate of inspected facilities found to be in substantial compliance over a specified time	Not statistically-valid based on inspected universe	Provides a rough indication of the degree of compliance by regulated facilities and the effectiveness of outreach and assistance efforts.		http://www.ecos.org/section/2006_state_enforcement__report/
NJ	Air	New Jersey employs the Environmental Compliance Consortium standard definition of compliance rate: 100*(1 - (# of facilities with violations / # of facilities with compliance evaluations)	Not statistically-valid based on inspected universe	The Air program uses these data to target inspections in sectors of the regulated universe where compliance has been low. Other programs exploring its use include Water, Hazardous Waste, Solid Waste, and UST. Compliance rates were further used to analyze geographic areas to target a major enforcement sweep in Dec. 2003. The eventual location selected was Patterson, NJ.		http://www.ecos.org/section/2006_state_enforcement__report/
NY	Drinking Water**	Performance measures based on percent of systems providing water data without specific types of violations	Census-based using self-reported data	Used to establish performance goals, track related improvement, and to identify non-reporting facilities.	EPA requirement	http://www.ecos.org/section/2006_state_enforcement__report/
NC	All	(# violation-free initial routine inspections/total # initial routine inspections) x 100	Not statistically-valid based on inspected universe	Programs have used this number to target sectors and to develop better compliance assistance.		http://www.ecos.org/section/2006_state_enforcement__report/, Documents sent from NC DENR, http://www.enr.state.nc.us/assets/applets/compliance2004.pdf
ND	Drinking Water** and RCRA	Two programs calculate compliance rates. One uses the number of systems with no violations divided by total number of systems and the other program uses on-site inspection data.	Census based for both, based on self-reported data for drinking water and on-site inspections for RCRA UST*	Drinking water program publishes this information in an annual compliance report. The UST program provides this information to US EPA.	Both are conducted due to federal requirements.*	http://www.ecos.org/section/2006_state_enforcement__report/

State	Program	Compliance Rate Calculation Method	Is it Statistically-Valid/Census based or other?	States' Use of the Measure	OECA Comments	Source
OH	Hazardous Waste	Percentage of permitted TSD facilities and LQG facilities that are in substantial compliance	Not statistically-valid based on inspected universe	We use these compliance rates to determine whether we need to increase our inspection efforts, develop guidance, or just to inform the specific regulated community of the results.		http://www.ecos.org/section/2006_state_enforcement__report/, Documents sent from OH EPA
OK	Drinking Water**	% = 100 - (# of systems with violations / # of systems)	Census-based using self-reported data	For EPA required Annual Compliance Report	EPA requirement	http://www.ecos.org/section/2006_state_enforcement__report/
OR	Underground Storage Tanks*	Specific EPA criteria as outlined in their significant operational compliance (SOC) guidelines	Not statistically-valid based on inspected universe	DEQ reports the data to the US EPA, comparing the success of our program to national averages and gauging the improvement of Oregon's program.	EPA requirement	http://www.ecos.org/section/2006_state_enforcement__report/
PA	All	Inspections w/o violations / Total inspections	Not statistically-valid based on inspected universe	It is only used for internal program management.		http://www.ecos.org/section/2006_state_enforcement__report/
RI	Drinking Water**	Indicator Value = SUM [(PWS Population Served) x (Days in Compliance With MCLs and Treatment Technique Requirements)] divided by the SUM [(PWS Population Served) x (Total Days in Operation)]	Census-based using self-reported data	We simply use it as a running indicator of program efficiency. The public can track the indicator in our annual reports	EPA requirement	http://www.ecos.org/section/2006_state_enforcement__report/
SC	Underground Storage Tanks*	EPA Significant Operational Compliance Formula (SOC)	Census-based using self-reported data	Data is reported to EPA Region 4 on a quarterly basis. The UST Program uses data to target outreach efforts.	EPA requirement	http://www.ecos.org/section/2006_state_enforcement__report/
SD	Hazardous Waste	Number of enforcement actions divided by total number of inspections conducted	Not statistically-valid based on inspected universe	Mostly the data used is for internal reporting to management and reporting to US EPA		http://www.ecos.org/section/2006_state_enforcement__report/
TN	Surface and Ground Water	Percent of major NPDES permittees in significant compliance	Statistically-valid SNC rate based on self-reported NPDES data.	Used as a performance measure	Same as EPA SNC rate	http://www.ecos.org/section/2006_state_enforcement__report/

State	Program	Compliance Rate Calculation Method	Is it Statistically-Valid/Census based or other?	States' Use of the Measure	OECA Comments	Source
TX	All	Formal enforcement actions NOT required per number of facilities investigated.	Not statistically-valid based on inspected universe			http://www.ecos.org/section/2006_state_enforcement__report/, Documents sent from TCEQ, http://www.tceq.state.tx.us/assets/public/compliance/enforcement/enf_reports/AER/FY05/enfrptfy05.pdf
UT	Drinking Water** UST*	Violations/facilities reviewed	Census-based using self-reported data for drinking water. Inspection based for UST		Required by EPA	http://www.ecos.org/section/2006_state_enforcement__report/,
VT	Drinking Water** UST*	Violations/facilities reviewed	Census-based using self-reported data for drinking water and inspection-based for UST.	The Water Supply Division will be updating how we calculate this information by using the US EPA drinking water reporting tool (SDWIS), which is currently under development in Vermont. The data have been reported in the US EPA Performance Partnership Agreement.	EPA requirement***	http://www.ecos.org/section/2006_state_enforcement__report/
VA	Underground Storage Tanks*	Specific EPA criteria as outlined in their significant operational compliance (SOC) guidelines	Not statistically-valid based on inspected universe	DEQ reports the data to the US EPA, comparing the success of our program to national averages and gauging the improvement of Oregon's program.	EPA requirement	http://www.ecos.org/section/2006_state_enforcement__report/
WA	NPDES Drinking Water** UST*	Water Quality: Percentage of the number of effluent limits actually in compliance, based on total "opportunities" for compliance. Opportunities are the number of effluent limits times the number of days reported within a given time frame.	Census based on self reported data for water and NPDES. Not statistically-valid for UST, based on inspected population.	Data is used for the Water Quality Program Annual Compliance Reports; Data is used in the Underground Storage Tank program for monitoring compliance trends. Remediation: Almost all sites are voluntarily cleaned up either under an Agreed Order or under a Consent Decree; very few sites need to receive a unilateral enforcement order.	NPDES is Closest to American Chemical Council approach of violations/compliance obligations. Rest are federal requirements	http://www.ecos.org/section/2006_state_enforcement__report/

State	Program	Compliance Rate Calculation Method	Is it Statistically-Valid/Census based or other?	States' Use of the Measure	OECA Comments	Source
WY	Hazardous Waste	# of facilities in compliance/number of facilities inspected.	Not statistically-valid based on inspected universe	Currently evaluating which compliance elements to include (significant violations, minor violations, include all facilities vs. only those inspected during period, etc.)		http://www.ecos.org/section/2006_state_enforcement__report/

Total: 38 programs identified through ECOS survey. Two additional programs identified by EPA (Connecticut and Illinois)

***UST Compliance Rate**: Currently, states report UST statistics to regions; regions report them semi-annually to EPA OUST. Using those statistics, OUST publishes a report entitled "United States Environmental Protection Agency Office of Underground Storage Tanks, Semi-Annual Activity Report." The statistics in this report include: "Percentage of UST facilities in significant operational compliance with the UST release detection requirements," and "Percentage of UST facilities in signficant operational compliance with the UST release prevention (spill, overfill and correction protection) requirements. Each performance measures is a percentage based on the initial inspections at facilities during the respective reporting period. If the state were to inspect more than 80% of the universe, then one could argue that the rate was statistically representative. However, as was reported by Vermont in the ECOS report, "sample size is relatively small – 10% of the universe. Compliance rates may be biased toward higher non-compliance since inspections are targeted at facilities suspected to have compliance problems."

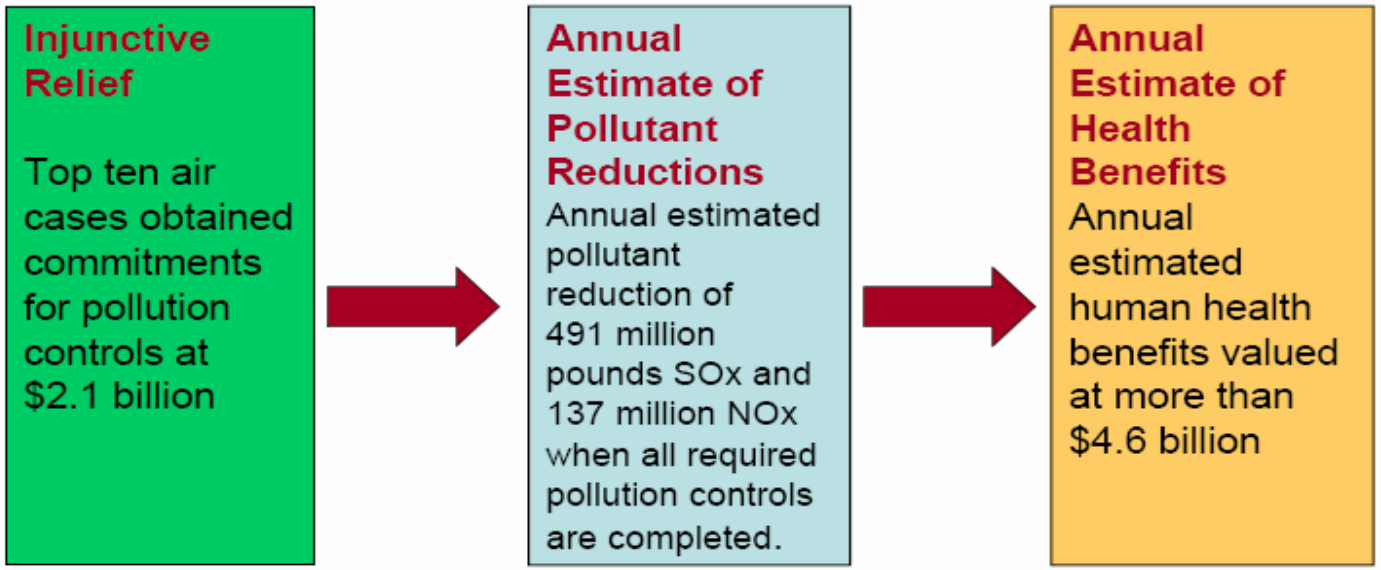
****Drinking Water Compliance Rate**: While states reported in the ECOS report that they were developing compliance rate data in the drinking water program based on the federal requirement. EPA in its own summary of this data in "Providing Safe Drinking Water in America: 2003 National Public Water Systems Compliance Report" expresses our concern with developing a rate using this data. On page 9 of the report we state, "Primacy states are required to submit data to SDWIS/FED each quarter. The data submitted includes inventory data and violation data. As part of EPA's ongoing data reliability efforts, EPA periodically examines the results of data. EPA's(most recent) review showed that most of the discrepancies between apparent and reported violations are because of unrecorded and unreported violations. This accounts for 62% of all discrepancies related to maximum contaminant level/treatment technique violations and 86% of all discrepancies related to monitoring and reporting violations."

*** Vermont reported in the ECOS report on the UST rate, "sample size is relatively small – 10% of the universe. Compliance rates may be biased toward higher non-compliance since inspections are targeted at facilities suspected to have compliance problems."

Appendix 4

FY 2005 Compliance & Enforcement Annual Results

Air Enforcement Cases Yield Human Health Benefits Valued Over $4.6 Billion Annually



This is the first time EPA is estimating health benefits from its enforcement cases. The benefits are estimated using an EPA peer-reviewed model. The estimate is for ten air cases only and is based only on NOx and SOx pollutant reductions. A portion of these benefits is expected to be realized immediately and to increase annually as the companies continue to install the pollution controls required in the judicial settlements. The benefits are expected to reach their maximum annual amount at $4.6 billion in 2012, after which they will continue to accrue. Thus, the $4.6 Billion is a very conservative estimate. The health benefits of NOx and SOx reductions quantified in this estimate include reductions in premature mortality, chronic and acute bronchitis, myocardial infarctions, hospitalizations, respiratory symptoms, and/or work loss days.

Appendix 5. Example of Air Pollutant Characterization

NITROGEN OXIDE

Rank	Pounds	# Facilities	# Cases	People Exposed (within 3 mi. radius)
2 of 10	192,203,118	31	19	1,101,370

Health Effects of Pollutant	Respiratory	Cardiac	Neurological	Reproductive/ Developmental	Carcinogenicity
NITROGEN DIOXIDE (NO2)	●	○	--	?	?

Ecological Effects of Pollutant	Vega.	Acid Rain	Water Qual.	GG	Smog	Ozone Deplete	Tox/ Animals	Visib	Bioaccum.	Built Env.
NO2	●	●	●	●	●	--	●	●	--	--

● = direct effect; ○ = indirect effect; -- = No or no data to support ; ? = unknown, not yet determined, studies are inconclusive

Definition and Uses: Nitrogen oxides (NOx) are a mixture of gases that are composed of nitrogen and oxygen. Two of the most toxicologically significant nitrogen oxides are nitric oxide and nitrogen dioxide. Nitrogen oxides are released to the air from the exhaust of motor vehicles, the burning of coal, oil, or natural gas, and during processes such as arc welding, electroplating, engraving, and dynamite blasting. They are also produced commercially by reacting nitric acid with metals or cellulose. Nitrogen oxides are used in the production of nitric acid, lacquers, dyes, and other chemicals. Nitrogen oxides are also used in rocket fuels, nitration of organic chemicals, and the manufacture of explosives.

Fate and Transport: Nitrogen oxides are broken down rapidly in the atmosphere by reacting with other substances commonly found in the air. The reaction of nitrogen dioxide with chemicals produced by sunlight leads to the formation of nitric acid, which is a major constituent of acid rain. Nitrogen dioxide also reacts with sunlight, which leads to the formation of ozone and smog conditions in the air we breathe. Small amounts of nitrogen oxides may evaporate from water, but most of it will react with water and form nitric acid. NOx reacts with ammonia, moisture, and other compounds to form nitric acid and related particles. When released to soil, small amounts of nitrogen oxides may evaporate into air. However, most of it will be converted to nitric acid or other compounds.

NOx causes a wide variety of health and environmental impacts because of various compounds and derivatives in the family of nitrogen oxides, including nitrogen dioxide, nitric acid, nitrous oxide, nitrates, and nitric oxide.

Exposure Pathways: The general population is primarily exposed to nitrogen oxides by breathing in air. People who live near combustion sources such as coal burning power plants or areas with heavy motor vehicle use may be exposed to higher levels of nitrogen oxides. Households that burn wood or use kerosene heaters and gas stoves tend to have higher levels of nitrogen oxides in them. Nitric oxide and nitrogen dioxide are found in tobacco smoke, so

people who smoke or breathe in second-hand smoke may be exposed to nitrogen oxides. Workers employed in facilities that produce nitric acid or certain explosives like dynamite and trinitrotoluene (TNT), as well as workers involved in the welding of metals may breathe in nitrogen oxides during their workday.

Human Health Effects

Low Level Exposure – Low levels of nitrogen oxides in the air can irritate your eyes, nose, throat, and lungs, possibly causing you to cough and experience shortness of breath, tiredness, and nausea. Exposure to low levels can also result in fluid build-up in the lungs 1 or 2 days after exposure.

High Level Exposure – Breathing high levels of nitrogen oxides can cause rapid burning, spasms, and swelling of tissues in the throat and upper respiratory tract, reduced oxygenation of body tissues, a build-up of fluid in your lungs, and death. If you were to come into skin or eye contact with high concentrations of nitrogen oxide gases or nitrogen dioxide liquid, you would likely experience serious burns. Children, people with lung diseases such as asthma, and people who work or exercise outside are susceptible to adverse effects of smog (NOx contribute to smog formation) such as damage to lung tissue and reduction in lung function. Millions of Americans live in areas that do not meet the health standards for ozone.

Reproductive/Developmental –Exposure of pregnant animals to nitrogen oxides has resulted in toxic effects in developing fetuses and nitrogen oxides have also caused changes in the genetic material of animal cells. We do not know if exposure to nitrogen oxides might cause developmental effects in humans.

Carcinogenicity – The Department of Health and Human Services (DHHS), the International Agency for Research on Cancer (IARC), and the EPA have not classified nitrogen oxides for potential carcinogenicity.

Susceptible Populations – Children would probably be affected by exposure to nitrogen oxides in the same ways as adults but we do not know whether children differ from adults in their susceptibility to nitrogen oxides.

Ecological Effects

Damage to Vegetation – In addition to human health impacts, nitrogen oxides play multiple roles in ecological degradation. For instance, smog, or ground-level ozone, also damages vegetation and reduced crop yields.

Acid Rain – Acid rain, formed when NOx and sulfur dioxide react with other substances in the air, causes deterioration of cars, buildings and historical monuments and causes lakes and streams to become acidic and unsuitable for many fish.

Water Quality – Increased nitrogen loading in water bodies, particularly coastal estuaries, upsets the chemical balance of nutrients used by aquatic plants and animals. Additional nitrogen accelerates "eutrophication," which leads to oxygen depletion and reduces fish and shellfish populations.

Greenhouse Gas – One member of the NOx, nitrous oxide, is a greenhouse gas. It accumulates in the atmosphere with other greenhouse gasses causing a gradual rise in the earth's temperature (global warming). Global warming will lead to increased risks to human health, a rise in the sea level, and other adverse changes to plant and animal habitat.

Smog Formation – Yes.

Toxic Effects/Biological Mutations in Animals – In addition, in the air, NOx reacts readily with common organic chemicals and even ozone, to form a wide variety of toxic products, some of which may cause biological mutations. Examples of these chemicals include the nitrate radical, nitroarenes, and nitrosamines.

Visibility Impairment – NOx are also to blame for the impairment of visibility in urban areas and on a regional scale in our national parks due to nitrate particles and nitrogen dioxide blocking the transmission of light.

Bioaccumulative – Nitrogen oxides do not build up in the food chain.

U.S. EPA Website, Six Common Air Pollutants, http://epa.gov/air/urbanair/nox/hlth.html

U.S. Center for Disease Control, Agency for Toxic Substances and Disease Registry, ToxFAQs for Nitrogen Oxides, http://www.atsdr.cdc.gov/tfacts175.html