Presented at the Engineering Foundation and the American Society of Civil Engineers Conference on *Information & Monitoring Needs for Evaluating the Mitigating Effects of BMPs*. Snowmass, CO, August 2001.

METHODS FOR THE ASSESSMENT OF URBAN WET-WEATHER FLOW IMPACTS

Robert Pitt, M. ASCE
University of Alabama
Tuscaloosa, AL

G. Allen Burton
Wright State University
Dayton, OH

ABSTRACT

During the past decade, it has become apparent during numerous receiving water assessment studies that no one single approach (e.g., chemical-specific criteria, benthic microorganisms, or habitat surveys) can routinely be used to accurately determine or predict ecosystem health and beneficial use impairment. Each assessment approach or component has associated strengths and weaknesses. The selection of specific assessment tools and goals is highly dependent on local conditions and objectives. This paper, based on a recently published book and EPA report, outlines the major components of a receiving water assessment to evaluate urban wet-weather problems.

INTRODUCTION

A myriad of potential stressor combinations and assessment methods are possible in waters that are in human dominated watersheds, as previously described by the EPA (1989, 1996, 1999a, and 1999b), among others. In the laboratory, it would be impossible to evaluate even a small number of the possible stressor combinations, varying the magnitude, frequency, and duration of each stressor. Traditional bioassay methods simply look at one simple exposure scenario. Chemical criteria provide a benchmark from which to evaluate the significance of contaminant concentration and direct further monitoring resources. Biological assessments indicate if the aquatic community is of a pollution and/or habitat tolerant or sensitive nature by showing the effect of long-term exposures. By considering habitat influence and comparing to reference sites, evaluations of ecological integrity (health) can be made. Habitat (physical) evaluations are essential to separate point source and nonpoint source toxicity effects from physical effects. As an example, some nonpoint pollution effects from stormwater may be of a physical nature, such as habitat alteration and destruction from increased stream flow, increased suspended and bedload sediments, or elevated water temperatures. In addition, a fourth major assessment component (toxicity) is needed beyond the three components of chemical, physical, and biological integrity. Biosurvey data may not detect subtle, short-term, or recent toxic effects due to the natural variation (spatial and temporal) which occurs in aquatic communities. Toxicity testing also removes the effects of habitat problems relatively well, focusing on the availability of chemical contaminants alone.

The complexity of ecosystems dictates that these assessment tools be used in an integrated fashion. Scientists in any of the disciplines are quick to point out the multitude of ecosystem complexities associated with their science. Many of these complexities influence chemical fate and effects and, more importantly, affect natural and anthropogenic stressor fate and effects. For example, it is well documented that many natural factors may act as significant stressors to organisms in aquatic systems, including light, temperature, flow, dissolved oxygen, sediment particle size, suspended solids, habitat quality, ammonia, salinity, food quality and quantity, predators, parasites and pathogens. In addition, ecotoxicologists have long been aware of the differences existing between species and their life stages in regards to toxicant sensitivity. Unfortunately, toxicity information only exists for a minuscule fraction of the many millions of species in the world. This reality makes extrapolations between species and chemical tenuous at best. Despite these many and often interacting complexities, some excellent and proven tools exist for conducting ecologically relevant assessments of contamination.

The necessity of using each of the above assessment components and the degree to which each is utilized is a site-specific issue. At sites of extensive chemical pollution, extreme habitat destruction, or absence of desirable aquatic organisms, the impact can be clearly established with only one or two components, or simply qualitative measures. However, at most study sites, there will be “gray” areas where the ecosystem’s integrity (quality) is less clear and should be measured via multiple components, using a weight-of-evidence approach to evaluate adverse effects.

A recently completed book, and associated EPA report, has compiled substantial guidance information related to assessing urban receiving waters affected by urban wet-weather flows (Burton and Pitt 2001). In addition to outlining various suitable approaches, detailed case studies are also presented. Specific experimental designs, plus sampling, laboratory, and statistical tools, and overall assessment methods, comprise most of the content of the book. The report is divided into three main sections, as shown below:

UNIT 1: THE PROBLEM OF STORMWATER RUNOFF
Chapter 1 Introduction
Chapter 2 Receiving Water Uses, Impairments, and Sources of Stormwater Pollutants
Chapter 3 Stressor Categories and Their Effects on Humans and Ecosystems

UNIT 2: COMPONENTS OF THE ASSESSMENT
Chapter 4 Overview of Assessment Problem Formulations

Chapter 5 Sampling Effort and Collection Methods
Chapter 6 Ecosystem Component Characterization
Chapter 7 Statistical Analyses of Receiving Water Data
Chapter 8 Data Interpretation

UNIT 3: Tool Box of Assessment Methods
Appendix A Habitat Characterization
Appendix B Benthic Community Assessment
Appendix C Fish Community Assessment
Appendix D Toxicity and Bioaccumulation Testing
Appendix E Laboratory Safety, Waste Disposal, and Chemical Analyses Methods
Appendix F Sampling Requirements For Paired Tests
Appendix G Water Quality Criteria
Appendix H Watershed and Receiving Water Modeling
Appendix I Glossary
Appendix J Vendors of Supplies and Equipment Used in Receiving Water Monitoring

This paper is a summary of selected material contained in Chapter 4 of the report.

SUMMARY OF ASSESSMENT TOOLS

Almost all states using bioassessment tools have relied on EPA guidance as the basis for their programs. Common components of these bioassessment programs (in general order of popularity) include:

- macroinvertebrate surveys (almost all programs, but with varying identification and sampling efforts)
- habitat surveys (almost all programs)
- some simple water quality analyses
- some watershed characterizations
- few fish surveys
- limited sediment quality analyses
- limited stream flow analyses
- hardly any toxicity testing
- hardly any comprehensive water quality analyses

Normally, numerous metrics are used, typically only based on macroinvertebrate survey results, which are then assembled into a composite index. Many researchers have identified correlations between these composite index values and habitat conditions. Water quality analyses in many of these assessments are seldom comprehensive, a possible over-reaction to conventional and very costly programs that have typically resulted in minimally worthwhile information. We recommend a more balanced assessment approach, using toxicity testing and carefully selected water and sediment analyses to supplement the needed biological and habitat monitoring activities. A multi-component assessment enables a more complete evaluation of causative factors and potential mitigation approaches.

BEGINNING THE ASSESSMENT

Designing and implementing an assessment study requires careful and methodical planning in order to ensure that the study objectives will be successfully accomplished. The main objectives of most environmental monitoring studies may be divided into two general categories: characterization, and/or comparisons. Characterization pertains to quantifying a few simple attributes of the parameter of interest. As an example, the concentration of copper in the sediment near an outfall may be of concern. The important question would be “What is the most likely concentration of the copper?” Other questions of interest include changes in the copper concentrations between surface deposits and buried deposits, or in upstream vs. downstream locations. These additional questions are considered in the second category, namely comparisons. Other comparison questions may relate to comparing the observed copper concentrations with criteria or standards. Finally, many researchers would also be interested in quantifying trends in the copper concentrations. This extends beyond the above comparison category, as treads usually consider more than just two locations or conditions. Examples of trend analyses would examine copper gradients along the receiving stream, or trends of copper concentrations with time. Another type of analysis related to comparisons is the identification of hot spots, where the gradient of concentrations in an area is used to identify areas having unusually high concentrations.

An adequate experimental design enables a researcher to efficiently investigate a study hypothesis. The results of the experiments will theoretically either prove or disprove the hypothesis. In reality, the experiments will tend to shed some light on the real problem and will probably result in many more questions that need addressing. In many cases, the real question may not have even been recognized initially. Therefore, even though it is very important to formally have a study hypothesis and appropriate experimental design, it may be important to save sufficient study resources in reserve to enable additional unanticipated experiments.

Experimental design covers several aspects of a monitoring program. The most important aspect of an experimental design is being able to write down the study objectives and why the data is needed. The quality of the data (accuracy of the measurements) must also be known. Allowable errors need to be identified based on how the information will change a conclusion. Specifically, how sensitive is the data that is to be collected in defining the needed answer? A logical experimental process that can be used to set up an assessment of receiving waters consists of several steps:

1) Establish clear study objectives and goals (hypothesis to be tested, calibration of equation or model to be used, etc.);
2) Initial site assessment and preliminary problem identification;
3) Review historical site data. Collect information on the physical conditions of the system to be studied (watershed characteristics, etc.) and estimate the time and space variabilities of the parameters of interest (assumed, based on prior knowledge, or other methods);
4) Formulate a conceptual framework (e.g., the EPA ecological risk framework);

5) Determine optimal assessment parameters. Determine the sampling plan (strata and relationships that need to be defined), including the number of samples needed (when and where, within budget restraints);
6) Establish data quality objectives (DQO) and procedures needed for QA/QC during sample collection, processing, analysis, data management, and data analyses;
7) Locate sampling sites;
8) Establish field procedures, including the sampling specifics (volumes, bottle types, preservatives, samplers to be used, etc.);
9) Review QA/QC issues;
10) Construct data analysis plan by determining the statistical procedures that will be used to analyze the data (including field data sheets and laboratory QA/QC plan); and finally,
11) Study implementation.

Preliminary project data obtained at the beginning of the project should be analyzed to verify assumptions used in the experimental design process. However, one needs to be cautious and not make major changes until sufficient data has been collected to verify new assumptions. After the data has been analyzed and evaluated, it is likely that follow-up monitoring could be conducted to address new concerns uncovered during the project.

All of these elements are described in detail in the book and EPA report. If any of these process components are inadequately addressed, the study outputs may not achieve the necessary study goals and objectives and/or lead to erroneous conclusions. An early paper by Green (1979) lists principles (Table 1) that are still valid for preparing environmental study designs.

Table 1. Principles for Designing Successful Environmental Studies (from Green 1979)

1. State concisely to someone what question you are asking. Your results will be as coherent and as comprehensible as your initial conception of the problem.

2. Take replicate samples within each combination of time, location, and any other controlled variable. Differences between groups can only be demonstrated by comparison to differences within groups.

3. To test whether a condition has an effect, collect samples both where the condition is present and where the condition is absent (reference site) but all else is the same. An effect can only be demonstrated by comparison with a control.

4. Carry out some preliminary sampling to provide a basis for evaluation of sampling design and statistical analysis options. Deleting this step to save time usually results in losing time.

5. Verify that the sampling device or method is sampling the population it should be sampling, and with equal and adequate efficiency over the entire range of sampling conditions to be encountered. Variation in efficiency of sampling from area to area biases among-area comparisons.

6. If the area to be sampled has a large-scale environmental pattern, break the area up into relatively homogeneous subareas and allocate samples to each in proportion to the size of the subarea. If it is an estimate of total abundance over the entire area that is desired, make the allocation proportional to the number of organisms in the subarea.

7. Verify that the sample unit size is appropriate to the size, densities, and spatial distributions of the organisms being sampled. Then estimate the number of replicate samples required to obtain the needed precision.

8. Test the data to determine whether the error variation is homogeneous, normally distributed, and independent of the mean. If it is not, as will be the case for most field data, then *(a)* appropriately transform the data, *(b)* use a distribution-free (nonparametric procedure, *(c)* use an appropriate sequential sampling design, or *(d)* test against simulated H_0 data.

9. Having chosen the best statistical method to test the hypothesis, stick with the result. An unexpected or undesired result is *not* a valid reason for rejecting the method and searching for a "better" one.

SELECTING OPTIMAL ASSESSMENT PARAMETERS (ENDPOINTS)

Characterization of the ecosystem should allow for differentiation of its present “natural” status from its present condition caused by polluted discharges and/or other anthropogenic stressors. This requires that a number of chemical, biological, and physical parameters be monitored, including flow and habitat. There are a wide variety of potentially useful study parameters which vary in importance with the study objectives and program needs. Many of the chemical endpoints would be specifically selected based on the likely pollutant sources in the watershed.

The selection of the specific endpoints for monitoring should be based on expected/known receiving water problems. The parameters being monitored should confirm if these uses are being impaired. If they are, then more detailed investigations can be conducted to understand the discharges of the problem pollutants, or the other factors, causing the documented problems. Finally, control programs can be designed, implemented, and monitored for success. Therefore, any receiving water investigation should proceed in stages if at all possible. It is much more cost-effective to begin with a relatively simple and inexpensive monitoring program to document the problems that may exist in a receiving water than it is to conduct a large and comprehensive monitoring program with little prior knowledge. Without having information on the potential existing problems, the initial list of parameters to be monitored has to be based on best judgment. The parameters to be monitored can be grouped into general categories depending on expected beneficial use impairments, as follows:

- Flooding and drainage: debris and obstructions affecting flow conveyance are parameters of concern.

- Biological integrity: habitat destruction, high/low flows, inappropriate discharges, polluted sediment (SOD and toxicants), benthic macroinvertebrate and fish species impairment (toxicity and bioaccumulation of contaminants) and wet-weather quality (toxicants, nutrients, DO) are key parameters.
- Non-contact recreation: odors, trash, high/low flows, aesthetics, and public access are the key parameters.
- Swimming and other contact recreation: pathogens, and above listed non-contact parameters, are key parameters.
- Water supply: water quality standards (especially pathogens and toxicants) are key parameters.
- Shellfish harvesting and other consumptive fishing: pathogens, toxicants, and those listed under biological integrity, are key parameters.

Point source discharges, stormwater runoff, snowmelt, base flows in receiving waters, sediments, and biological specimens may all need to be sampled and analyzed to obtain a complete understanding of receiving water effects from pollutant discharges. The following paragraphs briefly describe a long list of analytes that could be monitored in urban receiving waters. It is expected that the list could be significantly reduced in most cases through screening analyses and better selections based on site-specific conditions.

Selection of Biological Endpoints for Monitoring

The optimal assessment parameters which should be included depend on the project objectives. These parameters can be defined as measured characteristics, responses, or endpoints. For example, if the affected stream is classified as a high quality water and cold water fishery, then possible assessment or measured responses (endpoints) could include trout survival and hatchability, population and community indices (e.g., species richness), spawning area quantity and quality, dissolved oxygen, suspended solids, and water temperature. Endpoints vary dramatically in their sensitivity to pollutants and ecological relevance. The endpoints which are more sensitive are often more variable or respond to natural “nonpollutant” factors, so that adverse effects (stressors) are more difficult to classify with certainty.

Aquatic ecosystems are quite complex, consisting of a wide variety of organisms. These organisms each have their own unique function in the ecosystem and are directly or indirectly linked with other organisms. For example, bacteria, fungi, insects, and other invertebrates that inhabit the bottom of the waterways each need the other to assist in the decomposition of organic matter (such as leaves) so that they may consume it as food. If any one of these groups of organisms is lost or reduced, then the others will also be adversely affected. For instance, if the invertebrates are lost, then their fish predators will be impacted. These groups are made up of a number of species with varying tolerance levels to stressors, and each possess unique or overlapping functional characteristics (e.g., organic matter processing, nitrogen cycling). By carefully selecting the biological monitoring parameters, a broad range of relevant and sensitive indicator organisms can be used to efficiently assess ecosystem quality.

The most commonly used biological groups in aquatic assessments are fish, benthic macroinvertebrates, zooplankton and algae. In lotic (flowing water) systems, fish and benthic macroinvertebrates are often chosen as monitoring tools. Benthic refers to sediment or bottom surfaces (organic and inorganic). Macroinvertebrates are typically classified as those organisms which are retained in sieves larger than 0.3 to 0.5 mm. They include a wide range of invertebrates, such as worms, insect larvae, snails, and bivalves. They are excellent indicators of water quality because they are relatively sedentary and do not move between different parts of a stream or lake, as fish do. In addition, a great deal is known about their life histories and pollution sensitivity. Algae, zooplankton, and fish are used more in lentic (lake) environments. Of these, fish are most often used (both in lotic and lentic habitats). Fish are transient, moving between sites, therefore it is more difficult to determine their source of exposure to stressors; however, they are excellent indicators of water quality and provide a direct link to human health and wildlife consumption advisories. Rooted macrophytes and terrestrial plant species are good wetland health indicators, but are used less frequently.

In order to effectively and accurately evaluate ecosystem integrity, biosurveys should use two to three types of organisms which have different roles (functions) in the ecosystem, such as decomposers (bacteria, producers, primary to tertiary consumers). This same approach should be used in toxicity testing (Burton, *et al.* 1989, 1996; Burton 1991). This increases the power of the assessment, providing greater certainty that if there is a type of organism(s) (species, population, or community) in the ecosystem being adversely affected, either directly or indirectly, then it will be detected. This also allows for better predictions of effects, such as in food chain bioaccumulation with subsequent risk to fish eating organisms (e.g., birds, wildlife, humans). A large database exists for many useful indicator species concerning their life history, distribution, abundance in specific habitats or ecoregions, ecological function and pollutant (stressor) sensitivity.

In the monitoring of fish and benthic macroinvertebrate communities, a wide variety of approaches have been used. A particularly popular approach recommended by the U.S. EPA, Ohio EPA, state volunteer monitoring programs, and other agencies is a multi-metric approach. The multi-metric approach uses the basic data of which organisms are present at the site and analyzes the data using a number of different metrics, such as richness (number of species present), abundance (number of individuals present), and groups types of pollution sensitive and resistant species. The various metrics provide unique and sometimes overlapping information on the quality of the aquatic community. Structural metrics describe the composition of a community, that is the number and abundance of different species, with associated tolerance rankings. Functional metrics may measure photosynthesis, respiration, enzymatic activity, nutrient cycling or proportions of feeding groups, such as omnivores, herbivores, insectivores, shredders, collectors, and grazers.

The Microtox™ (from Azur) toxicity screening test has been successfully used in numerous studies to indicate the sources and variability of toxicant discharges. However, these tests have not been standardized by the U.S. EPA or state environmental agencies. More typically, whole effluent toxicity test methods are employed (see review by Burton, *et al*. 2000). These tests may miss toxicant pulses and do not reflect real-world exposure dynamics. Many of the *in-situ* toxicity tests, especially in conjunction with biological surveys (at least habitat and benthic macroinvertebrate evaluations) and sediment chemical analyses, can provide more useful information to document actual receiving water toxicity problems than relying on water analyses alone. If a water body is shown to have toxicant problems, it is best to conduct a toxicity identification evaluation (TIE) study to attempt to isolate the specific problematic compounds (or groups of compounds) before long lists of toxicants are routinely analyzed.

Selection of Chemical Endpoints for Monitoring

An initial monitoring program needs to include parameters associated with the above beneficial uses. However, as the receiving water study progresses, it is likely that many locations and some beneficial uses may not be found to be problematic. This would enable a reduction in the list of parameters to be routinely monitored. Similarly, additional problems may also become evident with time, possibly requiring an expansion of the monitoring program. The following paragraphs briefly describe the main chemical monitoring parameters that could be included for the beneficial use impact categories for a receiving water only affected by stormwater. However, it may be a good idea to periodically conduct a more detailed analysis as a screening tool to observe less obvious, but persistent problems. If industrial or municipal point discharges, or other nonpoint discharges (such as from agriculture, forestry, or mining activities) also affect the receiving water under study, additional constituents may need to be added to this list.

Obviously, chemical analyses can be very expensive. Therefore, care should be taken to select an appropriate list of parameters for monitoring. However, the appropriate number of samples need to be collected (using statistically-based experimental design equations) to ensure reliable conclusions. Chemical analyses of sediments may be more informative of many receiving water problems (especially related to toxicants) than chemical analyses of water samples. This is fortunate because sediment chemical characteristics do not change much with time, so fewer sediment samples generally need to be analyzed during a study period compared to water samples. However, the chemical characteristics of sediments tend to vary greatly with location, including depth. The concentrations of many of the constituents are much higher in sediment samples than in water samples, requiring less expensive methods for analyses. Unfortunately, sediment sample preparation (especially extractions for organic toxicant analyses and digestions for heavy metal analyses) can be much more difficult for sediments than for water.

Sediment Chemical Analyses

The basic list for chemical analyses for sediment samples, depending on beneficial use impairments, includes: toxicants and sediment oxygen demand. The toxicants should include heavy metals (likely routine analyses for copper, zinc, lead, and cadmium, in addition to periodic ICP analyses for a broad list of metals). Acid volatile sulfides (AVS) are sometimes also analyzed to better understand the availability of the sediment heavy metals. Other sediment toxicant analyses may include PAHs and pesticides. Particle size analyses should also be routinely conducted on the sediment samples collected. Sediment oxygen demand analyses, in addition to an indication of sediment organic content (preferably particulate organic carbon, or at least COD and volatile solids), and nutrient analyses, are important in areas having nutrient enrichment or oxygen depletion problems. Microorganisms (*E. coli,* enterococci, and fecal coliforms) should also be evaluated in sediments in areas having likely pathogen problems (all urban areas). Interstitial water may also need to be periodically sampled and analyzed at important locations for the above constituents.

Water Chemical Analyses

The basic list for analyses for water samples, depending on beneficial use impairments, includes: toxicants, nutrients, dissolved oxygen, and pathogens.

The list of specific toxicants is similar as for the sediments (copper, zinc, lead, and cadmium, plus PAHs and pesticides). However, because of the generally lower concentrations of the constituents in the sample extracts for these analyses, more difficult analytical methods are generally needed, but the extraction and digestion processes are usually less complex than for sediments. In addition, because of the high variability of the constituent concentrations with time, many water samples are usually required to be analyzed for acceptable error levels. Therefore, less costly screening methods should be stressed for indicating toxicants in water. Because of the their strong associations with particulates, the toxicants should also be periodically analyzed in both their total and filterable forms. This increases the laboratory costs, but is necessary to understand the fates and controllability of the toxicant discharges. Typical chemical analyses for stormwater toxicants may include:

- metals (lead, copper, cadmium, and zinc using graphite furnace atomic adsorption spectrophotometer, or other methods having comparable detection limits), periodic total and filtered sample analyses;
- organics (PAHs, phenols, and phthalate esters using GC/MSD with SIM, or HPLC), pesticides (using GC/ECD, or immunoassays), periodic total and filtered sample analyses.

Pesticides in urban stormwater have recently started to receive more attention (USGS 1999). The USGS's National Water Quality Assessment (NAWQA) program has extensively sampled urban and rural waters throughout the nation. Herbicides commonly detected in urban water samples include: Simazine, Prometon, 2,4-D, Diuron, and Tebuthiuron. These herbicides are extensively used in urban areas. However, other herbicides frequently found in urban waters are used in agricultural areas almost exclusively (and likely drift in to urban lands from adjacent farm lands) and include: Atrazine, Metolachlor, Deethylatrazine, Alachlor, Cyanezine, and EPTC. Insecticides commonly detected in urban waters include: Diazinon, Carbaryl, Chlorpyrifos, and Malathion.

Nutrient analyses are also important when evaluating several beneficial uses. These analyses are not as complex as the above listed toxicants and are therefore much less expensive. However, relatively large numbers of analyses are still required. Water analyses may include the following typical nutrients: total phosphorus, inorganic phosphates (and, by difference, organic phosphates), ammonia, Kjeldahl nitrogen (or the new HACH total nitrogen method), nitrate plus nitrite, and TOC. Periodic analyses for total and filtered forms of the phosphorus, total nitrogen, and TOC should also be conducted.

Dissolved oxygen is a basic water quality parameter and is important for several beneficial uses. Historical discharge limits have typically been set based on expected DO conditions in the receiving water. The typical approach is to use a portable DO meter for grab analyses of DO. Continuous *in-situ* monitors are much more useful, especially the new units that have much more stable DO monitoring capabilities and can also frequently record temperature, specific conductance, turbidity, pH, and ORP. These long-term analyses are especially useful when evaluating diurnal variations or storm-induced discharges.

Pathogens should be frequently monitored in most receiving waters. Both urban and rural streams are apparently much more contaminated by problematic pathogenic conditions than have been previously assumed. Historically monitored organisms (such as fecal coliforms), in addition to *E. coli* and enterococci which are now more commonly monitored, can be very high and persistent in urban streams. Specific pathogens (such as *Pseudomonas aeruginosa* and *Shigella*) can also be more easily monitored now than in the past. Most monitoring efforts would probably focus on fecal coliforms, *E. coli,* and enterococci.

Additional conventional parameters affecting fates and effects of pollutants in receiving waters should also be routinely monitored, including hardness, alkalinity, pH, specific conductivity, COD, turbidity, suspended solids (SS), volatile suspended solids (VSS), and dissolved solids (TDS).

Selection of Additional Endpoints Needed for Monitoring

Several other stream parameters also need to be evaluated when investigating beneficial uses. These may include: debris and flow obstructions, high/low flow variations, inappropriate discharges, aesthetics (odors and trash), and public access.

DATA QUALITY OBJECTIVES AND QUALITY ASSURANCE ISSUES

For each study parameter, the precision and accuracy needed to meet the project objectives should be defined. After this is accomplished, the procedures for monitoring and controlling data quality must be specific and incorporated within all aspects of the assessment, including sample collection, processing, analysis, data management and statistical procedures.

When designing a plan one should look at the study objectives and ask:

- how will the data be used to arrive at conclusions?
- what will the resulting actions be? and
- what are the allowable errors?

This process establishes the Data Quality Objectives (DQOs) which determine the level of uncertainty that the manager is willing to accept in the results. DQOs, in theory, require the study designers (decision makers and technical staff) to decide what are allowable probabilities for Type I and II errors (false positive and false negative errors) and issues such as what difference in replicate means is significant. The DQO process is a pragmatic approach to environmental studies, where limited resources prevent the collection of data nonessential to the decision making process. Uncertainty in ecological impact assessments is natural due to variability and unknowns, sampling measurement errors and data interpretation errors. Determining the degree of uncertainty in any of these areas can be difficult or impractical. Yet an understanding of these uncertainties and their relative magnitudes is critical to the QA objectives of producing meaningful, reliable and representative data. The more traditional practices of QA/QC should be expanded to encompass these objectives and thus help achieve valid conclusions on the test ecosystem's health (Burton 1992).

The first stage in developing DQOs requires the decision-makers to determine what information is needed, reasons for the need, how it will be used, and specify time and resource limits. During the second stage, the problem is clarified and constraints on data collection are identified. The third stage develops alternative approaches to data selection, selecting the optimal approach, and establishing the DQOs (EPA 1986).

EXAMPLE OUTLINE OF A COMPREHENSIVE RUNOFF EFFECT STUDY

The following is an outline of the specific steps that need to be generally followed when designing and conducting a receiving water investigation. Some specific examples of monitored parameters are listed, but these would need to be modified based on local conditions.

Step 1. What's the Question?

For example: Does site runoff degrade the quality of the receiving stream ecosystem? Pitt (1995 and 2001) presents a summary of documented receiving water problems associated with urban stormwater, for example. Knowing the problems that have been identified and studied elsewhere will enable the investigators to identify the likely problems that may be occurring in their own local receiving waters, and to identify the likely causative factors.

Step 2. Decide on Problem Formulation:

Candidate experimental designs can be organized in one of the following basic patterns:

1. Parallel watersheds (developed and undeveloped)
2. Upstream and downstream of a city
3. Long-term trend
4. Preferably most elements of all of the above approaches combined in a staged approach

Another important issue is determining the appropriate study duration. In most cases, at least one year should be planned in order to examine seasonal variations, but a longer duration may be needed if unusual or dynamic conditions are present. However, trend analyses can require many years. In addition, variations in the parameters being investigated will require specific numbers of observations in order to obtain the necessary levels of errors in the program. If the numbers of observations need relate to events (such as runoff events), then the study will need to last for the duration necessary to observe and monitor the required number of events.

Step 3. Project Design

3.1. Qualitative watershed characterization

3.1.1. Establish degree of residential, commercial, and industrial areas to predict potential stressors. Typically, elevated solids, flows and temperatures are stressors common to all urban land uses. The following lists typical problem pollutants that may be associated with each of these land uses:

- Residential: nutrients, pesticides, fecal pathogens, PAHs and metals
- Commercial: petroleum compounds, metals
- Industrial: petroleum compounds, other organics, metals
- Construction: suspended solids

Topographical maps are also used to determine watershed areas and drainage patterns.

3.2. Stream characterization
3.2.1. Identify potential upstream stressor sources and potential stressors and photograph and describe sites.
3.2.2. Survey upstream and downstream (from outfall to 1 km minimum) quality. Record observations on physical characteristics including: channel morphology (pools, riffles, runs, modification), flow levels, habitat (for fish and benthos), riparian zone, sediment type, organic matter, oil sheens, and odors. Record observations on biological communities, such as waterfowl, fish eating birds or mammals, fish, benthic invertebrate, algal blooms, benthic algae, and filamentous bacteria.
3.2.3. Identify appropriate reference site upstream and/or in a similar sized watershed with same ecoregion.
3.2.4. Collect any historical data on water quality and flows.

3.3. Select Monitoring Parameters
3.3.1. Habitat Evaluation. Should be conducted at project initiation and termination. Includes Quantitative Habitat Evaluation Index (QHEI), bed instability survey (bed lining materials and channel cross-sectional area changes), aesthetic/litter survey, inappropriate discharges (field screening), etc.
3.3.2. Stressors and their indicators:
3.3.2.1. Physical: flow, temperature, turbidity. Determine at intervals throughout base to high flow conditions.
3.3.2.2. Chemical: conductivity, dissolved oxygen, hardness, alkalinity, pH, nutrients (nitrates, ammonia, ortho-phosphates), metals (cadmium, copper, lead, and zinc) and immunoassays (pesticides and polycyclic aromatic hydrocarbons) and/or toxicity screening (Microtox). The necessity of doing nutrients, metals, and organics will be dependent on the watershed characteristics. Determine at intervals throughout base to high flow conditions.
3.3.2.3. Biological: benthic community structure (e.g., RBP), fish community structure and tissue residues (confirmatory studies only). Benthic structure should be determined at the end of the project. Sediment bioaccumulation potential can be determined using the benthic invertebrate, *Lumbriculus variegatus*.

3.3.2.4. Toxicity: short-term chronic toxicity assays of stream water, outfalls, and sediment. Sediment should be sampled during base flow conditions and tested prior and after a high flow event. Water samples should be collected during base flow and during pre-crest levels. Expose test chambers with and without sunlight simulating light (containing ultraviolet light wavelengths) to detect PAH toxicity. *In situ* toxicity assays should be deployed in the stream for confirmatory studies during base and high flow periods.

3.4. Data Quality Objectives. Determine the kinds of data needed and the levels of accuracy and precision necessary to meet the project objectives. These decisions must consider that there typically is a large amount of spatial and temporal variation associated with runoff study parameters. Guidance is available that relates sampling efforts associated with actual variability and accuracy and precision goals. This requires additional resources for adequate quantification compared to simple preliminary surveys.

3.5. Triggers and Tiered Testing. Establish the trigger levels or criteria which will be used to determine when there is a significant effect, when the objective has been answered, and/or when additional testing is required. Appropriate trigger levels may include:

- An arbitrary 20% difference in the test site sample, as compared to the reference site, might constitute a significant effect (a difference this small will be difficult and therefore expensive to detect because of the natural variability for many parameters),
- An exceedence of the 95% statistical confidence intervals as compared to the reference sample,
- High toxicity in the test site sample, measured as Toxic Units (TUs) (e.g., 1/LC50),
- Exceedence of biotic integrity, sediment or water quality criteria, guidelines, or standards at the test site, and/or
- Exceedence of a hazard quotient of 1 (e.g., site concentration/environmental effect or background concentration).

A tiered or a phased testing approach is most cost-effective, if time permits. A qualitative or semi-quantitative study may include a greater number of indicator or screening parameters, such as: turbidity, temperature, DO, specific conductivity, and pH using a continuous recording water quality sonde, plus artificial substrate macroinvertebrate colonization tests, and “quick” sediment toxicity tests. If possible, Microtox™ screening toxicity tests, immunoassay tests for pesticides and PAHs, and sediment metal analyses should also be added to this initial effort. These simple tests can be conducted with more widespread sampling to better focus later tiers on quantifying appropriate stressors in critical sampling areas and times. Final project tiers can identify specific stressors, their contribution to the problem, their sources, or simply confirm the ecological significance of the observed effects.

3.6. Sampling Station Selection. Select the study sites, such as upstream reference sites, outfall(s), and downstream impacted sites. In the selection of the upstream/reference and downstream sites, consider flow dynamics, stressor sources, and reference habitat similarities.

3.7. Quality Assurance Project Plans (QAPP). It is essential that the quality of the project be ensured with adequate quality assurance and quality control measures. This will include routine laboratory and field documentation of operator and instrumentation performance, chain-of-custody procedures, adequate sample replication, QA/QC samples (blanks and spikes, etc.), performance criteria, and ensuring data validity. Appropriate experimental design (study design and sampling efforts) are also critical components of a QAPP.

Step 4. Project Implementation (Routine Initial Semi-Quantitative Survey)

4.1. Base Flow Conditions
4.1.1. Habitat Survey (e.g., Qualitative Habitat Evaluation Index)
4.1.2. Benthic RBP
4.1.3. Test water and sediment from all test sites for short-term chronic toxicity with two species.
4.1.4. Establish spatial and diurnal variation (YSI 6000 for several weeks, plus grab samples or time composites).
4.1.5. Set up automatic stream samplers/monitors, stream depth gauges, and rain gauges.
4.1.6. Establish local contacts to oversee field equipment and provide rain event notification.
4.1.7. Conduct field screening survey at outfalls to identify sources of dry weather flows.

4.2. High Flow Conditions
4.2.1. Confirm that the samplers and monitors are operational. Collect grab samples if necessary (for microbiological and VOC analyses, for example).
4.2.2. Deploy *in situ* toxicity test assays.
4.2.3. Measure flow and note staff gauge depth, using manual or automatic samplers and flow recorders. Repeat flow measurements at intervals of 0.5 to 1.0 ft stream depth intervals as the stream rises, noting time and depth. Focus on rising limb to crest period.
4.2.4. Measure D.O., temperature, turbidity, conductivity, and stage at each station following each flow measurement. Establish spatial variance. May use continuous recording water quality sondes.
4.2.5. Collect flow-weighted composited (or combine many discrete) samples for other analyses.

4.3. Sample Analyses
4.3.1. Filter, preserve and chill samples, as required.
4.3.2. Deliver samples to analytical laboratories with chain of custody forms.
4.3.3. Initiate toxicity testing and other chemical and microbiological analyses within required time period since sample collection.
4.3.4. Document QA/QC.

4.4. Follow-Up (Post-Event) Monitoring

4.4.1. Check *in situ* assay chambers at 24 and 48 and at 7 and 14 days if deployed.
4.4.2. Conduct benthic RBP.
4.4.3. Conduct QHEI, noting bed load movement.
4.4.4. Collect fish for tissue residue analyses.

Step 5. Data Evaluation

5.1. Plot flow vs. physical and chemical analysis results.

5.2. Statistically compare responses/loadings during base, rising limb, and post-crest conditions. This will provide a characterization of flow dynamics and its affect on stressor profiles.

5.3. Statistically compare stations (instantaneous, mean periods) for significant differences and correlations.

5.4. Calculate and compare physical, chemical and toxicity (using Toxicity Units) loadings. This will show the relative load contribution of stressors from reference (upstream) vs. impacted (downstream) reach.

5.5. Identify magnitude and duration of trigger exceedences.

5.6. Identify sources of uncertainty.

5.7. Identify potential sources of pollutants and stressors.

5.8. Determine literature value thresholds for key stressors on key indigenous species.

Step 6. Confirmatory Assessment (Optional Tier 2 Testing)

6.1. Repeat Steps 2 and 3 using Tier 1 information to select fewer test parameters with increased sampling frequency and/or select more descriptive methods. Increased sampling will better quantify the magnitude and duration of stressor dynamics. Expanded sampling will better document the quality of the receiving water. More definitive testing could include:

- Short-term chronic toxicity testing with additional species (lab and *in situ*),
- Increased testing of toxicants,
- Characterizing fish, plankton, periphyton, or mussel populations,
- Measuring assimilative capacity via long term BOD and SOD testing, and/or
- Measuring productivity with light/dark bottle BOD *in situ* tests.

6.2. Conduct Toxicity Identification Evaluation (TIE) study of water, outfalls, and/or sediment to determine contribution of each stressor to total toxicity. This information can better determine which stressors are important to control and can also identify sources of toxicity.

6.3. Conduct bioaccumulation testing of site sediments. Some pollutants, such as highly chlorinated organic compounds (e.g., chlordane, DDT, PCBs, dioxins) are readily bioaccumulated, yet may not be detected using the above study design. The EPA has a benthic invertebrate 28-day assay to measure sediment bioaccumulation potential. Also SPMDs (semi-permeable membrane devices) may be used.

6.4. Indigenous Biological Community Characterization and Tissue Analysis. More in-depth quantification of benthic and/or fish community structure on a seasonal basis will better identify significant ecological effects. Tissue sampling of fish for contaminants will provide information on bioaccumulative pollutants and potential food web or human health effects from consumption.

Step 7. Project Conclusions

7.1. List probable stressors.

7.2. Document trigger exceedences.

7.3. Discuss relative contribution of stressors(s) to ecosystem degradation. Support documentation may include:

- Literature threshold values,
- Criteria exceedences,
- Toxicity observed (from TIE, photo-activation, or *in situ* assays), and/or
- Bioaccumulation factors and potential for food web contamination.

7.4. Provide recommendations for stressor reduction and ecosystem enhancement.

7.5. Include suggestions on habitat improvement, flow reduction, turbidity removal and reduced siltation.

CONCLUSIONS

The specifics for any receiving monitoring program would be determined by the study objectives and the site conditions. As an example, Table 2 summarizes some general parameters that should be included in an urban water use evaluation study, depending on the specific beneficial uses of interest. Of course, the final parameters selected for study would vary for specific site conditions and historical information. As expected, an investigation of drainage uses (the primary use for an urban waterway) would be relatively straight-forward compared to studies of other use impairments. However, investigations of drainage problems can be expensive and time-consuming. When the other uses are added to the list of potential objectives, the necessary data collection effort can become very comprehensive and expensive. Therefore, a staged approach is usually recommended, with a fairly simple initial effort used to obtain basic information. This information can then be used to develop specific experimental designs for later study stages.

The book and EPA report (to be available on the EPA's web site:) includes many examples of receiving water investigations and specific examples for conducting a multi-faceted study. Also included in the book are chapters describing experimental design procedures for determining the extend of an investigation and chapters to assist in the evaluation of the data.

Table 2. Parameters of Concern when Evaluating Different Receiving Water Uses

	Drainage	**Biological life and integrity**	**Non-contact recreation**	**Swimming and other contact recreation**	**Water supply**	**Shellfish harvesting and other consumptive fishing uses**
debris and obstructions (channel conveyance capacity)	X					
habitat destruction (channel stability, sediment scour and deposition)		X				X
high/low flows (rates and durations)		X	X	X		X
aesthetics, odors and trash			X	X		
safety (bank condition, garbage)			X	X		
public access			X	X		
inappropriate discharges		X	X	X	X	X
benthic macroinvertebrate species present		X	X			X
fish species present		X				X
polluted sediment (SOD and toxicants [1])		X				X
toxicity and bioaccumulation of toxicants [1]		X				X
health-related water quality standards (especially microorganisms [2] and toxicants [1])				X	X	X
wet-weather quality (toxicants[1], nutrients[3], DO, temperature, alkalinity, and hardness)		X				X

Primary constituents are indicated in bold/underlined and should be analyzed for most, if not all, samples. Others can be analyzed less often as screening tests. In all cases, the common constituents should also be analyzed for all samples.

[1] Toxicants (organic toxicants such as: pesticides, herbicides, and PAHs; metallic toxicants such as: **zinc**, **copper**, **lead**, cadmium, arsenic, and mercury) and toxicity tests (such as: **Microtox screening test**, plus other *in-situ* and laboratory toxicity tests)

[2] Microorganisms (indicator bacteria and selected pathogens such as: **fecal coliforms**, ***E. coli***, **enterococci**, and *Pseudomonas aeruginosa*)

[3] Nutrients (**ammonia**, TKN, **nitrates**, TP, **phosphates**)

Common constituents, added to all water quality investigations (**pH**, **conductivity**, **turbidity**, **suspended solids**, **COD**)

REFERENCES

Burton, G. A., B. L. Stemmer, K. L. Winks, P. E. Ross, and L. C. Burnett. "A Multitrophic Level Evaluation of Sediment Toxicity in Waukegan and Indiana Harbors," *Environ. Toxicol. Chem.* 8:1057-1066. 1989.

Burton G.A., Jr. Assessing freshwater sediment toxicity. *Environ. Toxicol. Chem*. 10: 1585-1627. 1991.

Burton, G. A., Jr. "Quality Assurance Issues in Assessing Receiving Waters," in *Proc. of the Conf. on Effects of Urban Runoff on Receiving Systems* J. Saxena, Ed. New York, NY: Engineering Foundation Publ. 1992.

Burton G.A., Jr., Hickey C., DeWitt T., Morrison D., Roper D., and Nipper M. *In situ* toxicity testing: Teasing out the environmental stressors. SETAC NEWS 16(5):20-22. 1996.

Burton, G.A., Jr., R. Pitt, and S. Clark. “The role of whole effluent toxicity test methods in assessing stormwater and sediment contamination.” *CRC Critical Reviews in Environmental Science & Technology.* 30: 413-447. 2000.

Burton, G.A. Jr., and R. Pitt. *Stormwater Effects Handbook: A Tool Box for Watershed Managers, Scientists, and Engineers*. CRC Press, Inc., Boca Raton, FL. 985 pgs. August 2001.

EPA (U.S. Environmental Protection Agency). "The Development of Data Quality Objectives," Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC. 1986.

EPA. *Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish*. Office of Water, U. S. Environmental Protection Agency, Washington, DC, EPA 444/4-89/001. 1989.

EPA. *Environmental Indicators of Water Quality in the United States*. Office of Water, U.S. Environmental Protection Agency. EPA 841-F-96-002. Washington, D.C., June 1996.

EPA. *Methods for Measuring the Toxicity and Bioaccumulation of Sediment-Associated Contaminants with Freshwater Invertebrates*. Second edition. EPA series number pending, Duluth, MN. 1999a.

EPA. *Rapid Bioassessment Protocols for Use in Wadeable Streams and Rivers: Periphyton, Benthic Macroinvertebrates and Fish*. Office of Water, U. S. Environmental Protection Agency, Washington, DC, EPA 841/B-99/002. 1999b.

Green, R. H. *Sampling Design and Statistical Methods for Environmental Biologists*. New York, NY: John Wiley and Sons, 1979.

Pitt, R. “Biological Effects of Urban Runoff Discharges.” In: *Stormwater Runoff and Receiving Systems: Impact, Monitoring, and Assessment*. (Edited by E.E. Herricks). Engineering Foundation and ASCE. CRC/Lewis. Boca Raton, pp. 127-162. 1995.

Pitt, R. “Receiving Water Impacts Associated with Urban Runoff.” In: *Handbook of Ecotoxicology*, 2nd edition (Edited by D.J. Hoffman, B.A. Rattner, G.A. Burton, Jr. and J.Cairns, Jr.). Lewis Publishers/CRC Press, Boca Raton. Publication in 2001.

USGS. *The Quality of Our Nation’s Waters-Nutrients and Pesticides*. U.S. Geological Survey Circular 1225. Denver, CO. 1999.