Guidance in Understanding and Developing the Data Quality Objectives

CERP QAOT G-002

Prepared by the
CERP Quality Oversight Team
Effective 8/15/2005

The Data Quality Objective (DQO) Process

Developing DQOs is an integral and important part of a systematic planning process that is designed to ensure that the final results can be used for the purpose for which the data were generated. This systematic planning process for purposes of these discussions on environmental data quality is the quality system that each organization must develop, implement and evaluate on a continuing basis.

> Quality Assurance is doing:
> The right thing,
> The right way
> The first time.

Data quality is not the sole responsibility of the laboratory or field operations - each individual that is a part of an environmental monitoring project has a responsibility for the ensuring that the final outcome meets a set of originally stated goals (DQOs). Therefore, it important that all management and technical areas be a part of the original goal-setting process so that the outcome is technically sound, economically feasible, and where applicable, meets regulatory requirements.

WHAT VALUE IS THE DQO PROCESS?

> "If you fail to plan, you plan to fail" - Anonymous

The DQO process, when properly used is designed to answer the following questions:

- What data are needed?
- Why do you need data?
- How will the data be used (decision)?
- What is the tolerance for uncertainty?
- What are the chances of making a right or wrong decision?
- And in doing so, it ensures that
- The project demonstrates fiscal responsibility since resources (time, funding, personnel, etc.) are used to collect right type, quality, and quantity of data
- There is stakeholder buy-in because affected parties are included in planning process
- Subsequent planning documents (quality assurance project plan, work plan, sampling and analysis plan) are written to reflect the specified DQOs and
- Open lines of communication are encouraged among all project participants.

WHO IS RESPONSIBLE FOR COORDINATING THE DEVELOPMENT OF DQOS?

Once a need for environmental information is established, the person who is responsible for overseeing the information gathering process bears the greatest responsibility for coordinating the QA activities. This person may be a principal investigator (PI), a project leader, a project manager or other individual who will be ultimately responsible for the project.

WHO ELSE IS INVOLVED WITH DQO DEVELOPMENT?

The DQO team must include individuals that represent key activities or interests. At a minimum, the team must have

- The project manager (described above)
- A representative from the group(s) performing field activities
- A representative from the group(s) performing the analytical/testing activities
- Representatives from data user organizations who may include regulatory agencies or decision makers
- A quality assurance specialist

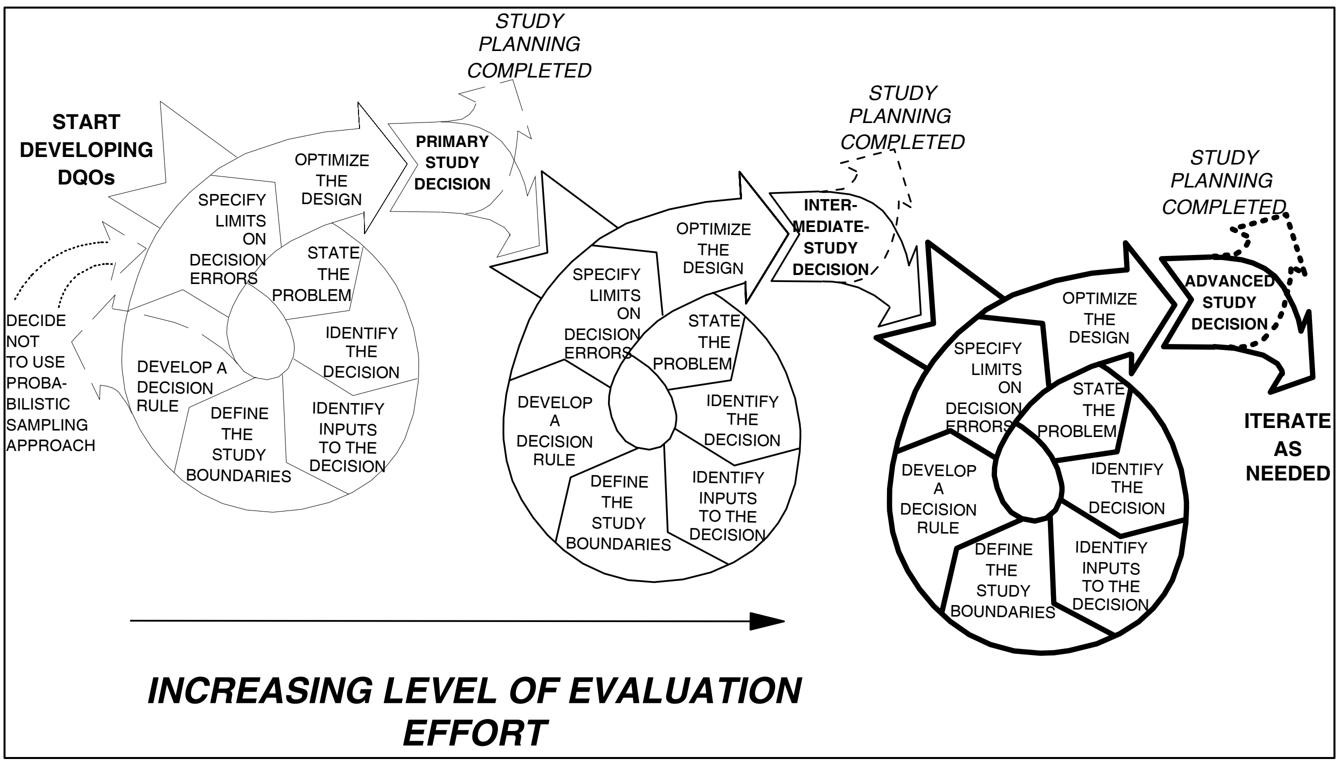
Other team members may include risk assessors, statisticians, data analysts or other individuals who will have important roles in evaluating or using the data.

WHAT ARE THE RESPONSIBILITIES OF THE DQO TEAM?

- The DQO team must first go through a systematic planning process (see EPA's 7-Step DQO Process below) to identify the project data requirements and the quality of the data.
- The Team must oversee the development of subsequent documents to ensure consistency with the DQOs.
- The Team must periodically review the project's progress to determine if the DQOs are being met, and to refine or change project requirements.
- The Team should also function as the QA oversight team to ensure that all quality assurance aspects of the project are being implemented.

HOW MUCH TIME IS NEEDED?

The amount of time spent in the process is dependent on the complexity of the problem, and how successful the DQO Team was in initially characterizing the problem. The DQO process is iterative. It may take several attempts to fine-tune the objectives, study design and data gathering activities:



WHAT PROCESS IS USED TO DEVELOP DQOS?

The EPA 7-step DQO process is recommended:

1. State the Problem
2. Identify the Decision
3. Identify the Inputs to the Decision
4. Define the Boundaries of the Study
5. Develop a Decision Rule
6. Specify Tolerable Limits on Decision Errors
7. Optimize the Design

WHAT IF I WANT TO USE ANOTHER APPROACH?

The DQO process is one version of what scientists call the "Scientific Method" or statisticians term the "Test of Hypotheses." Use any process that systematically approaches a monitoring project to determine:

- The purpose of the data gathering exercise
- The anticipated outcomes and the decision that will be made on the data
- The consequences of making an incorrect decision
- The spatial, analytical and time restrictions of the project
- The quality of the data needed to make a correct decision

The EPA process works well to organize a complex problem so that each component of the problem contributes to the final problem resolution. If there is a regulatory component (permit, consent order, etc.), then parts of the DQO process may have been predetermined.

THE DQO PROCESS

STEP 1 - STATE THE PROBLEM

1. Develop a succinct statement of the problem. In doing so, determine the external and internal economic, political and ecological interests may impact resolving the stated problem.
2. Consider:
    a. Availability of resources
    b. Time allocation for all aspects (including data review and evaluation)
    c. Impacts – social, political, economic, environmental
    d. External interests: public, regulatory, etc.

STEP 2 - IDENTIFY THE DECISION

1. Based on the problem, identify the principal study question(s). The question generally characterizes how the problem will be solved.
2. Discuss how the results of the study will be used to answer the question.
3. Determine the anticipated outcome(s) (results) and what course of action(s) will be taken based on the results. To the extent possible, anticipate all outcomes and how these would affect the decision.
4. Define the decision statement that must be resolved to address the problem.
5. Combine the principal study question and the alternative actions into a specific decision statement.

STEP 3 – IDENTIFY INPUTS

1. Focus on the information needed for the decision
    a. Parameters/characteristics to be monitored
2. Identify the variable/characteristic(s) to be measured.
    a. How will the information be reduced for the decision? (e.g., trend analysis, mean, highest contamination, etc.)
3. Identify the information needed to establish/meet the action level
    a. Permit limits, environmental action levels, etc may predetermine action levels.
    b. In some cases the study itself may provide information to set an action level. In these cases, rather than identifying a specific numerical value, determine the mechanisms to be used to establish the level (e.g., risk assessment, technology-based standards, negotiations between stakeholders, etc.)

STEP 4 - DEFINE PROJECT BOUNDARIES

1. Define the spatial boundary
    a. Define geographical area which decision will apply
    b. Define type of media
    c. Divide each medium into homogenous strata
2. Define temporal boundary
    a. Time frame which decision applies
    b. Determine when to sample
3. Define scale of decision making – how large/small is the unit that will be impacted by the decision?
4. Identify practical constraints on data collection

STEP 5 - DEVELOP A DECISION RULE

1. Develop “if/then” statement that incorporates:
2. Identify the statistic used to describe the population (e.g. mean, maximum, percentile)
3. Determine the scale of decision making (e.g. residential lot, pond, watershed)
4. Decide on the value that will trigger action
5. Discuss alternative actions

STEP 6 – SPECIFY LIMITS ON DECISION ERROR

1. Determine possible range of parameter of interest (e.g. total lead from 5 to 200 ug/L)
2. Identify decision errors and choose null hypothesis testing
3. Specify range of parameter values where consequences of decision error are minor (gray region)

Assign probability limits above/below the gray region that reflect tolerable probability for decision error

<u>Step 6.1 – Establish the Anticipated range of values</u>

Establish the expected or plausible range of values based on historical data, available information or professional judgment.

<u>Step 6.2a - Identify Potential Error Sources</u>

1. Identify the sources of potential errors in the sample data set that may lead to a false or faulty conclusion. Consider:
2. Sampling Design Error
    a. Temporal and spatial population variability
    b. Sample Collection Design
    c. Number of Samples
3. Measurement Error:
    a. Measurement Process (instrumental error)
    b. Sample Handling (collection, preparation)
    c. Data Reduction
    d. Storage
4. Specify and evaluate the potential consequence of each decision error
    a. Health risks
    b. Ecological risks
    c. Political risks

> **Considerations for Error**
>
> In general more error is introduced during the sampling process than the analytical process
>
> The greater the variability in the population being sampled, the greater the number of samples that must be collected to characterize that population
>
> A statistical sampling design must be used in order to develop quantitative decision error limits

d. Social risks
e. Resource risk

Step 6.2b - Determine the Baseline Condition

Determine baseline condition (null hypothesis) and alternative hypothesis and assign "false positive" and "false negative" to the decision error

Step 6.2c – Outline the consequences of a Decision Error

Step 6.3 - Establish which error has more severe consequence near the action level and the "gray area".

The "gray region" is the area where the consequences of making an incorrect decision are relatively minor (too close to call). Typically this area uses the action level as the lower limit, with the upper limit being the value at which the consequence of making a decision error is significant.

Step 6.4 Assign probability limits above/below the gray region that reflect tolerable probability for decision error

Construct a "what if" table or

Use a statistical approach to graph the decision process (for discussions on decision performance see the EPA document on DQOs). A statistical analysis will also document the justification for the decision making process, as well as the tolerable error rate for making either a false rejection or false acceptance decision.

DQO STEP 7 – OPTIMIZE THE DESIGN

Collect right amount and type of samples to give an accurate picture of the question you are trying to answer.

1. Determine feasibility of data collection design and alternatives
2. Select optimal sample size while considering decision error tolerance and available resources
3. Review DQOs for Consistency
4. Document sample design and assumptions in a QAPP

Step 7.1 and 7.2 - Sampling Design

Glossary

Quality System – a structured and documented management system describing the policies, objectives, principles, organizational authority, responsibility, accountability and implementation plan of an organization for ensuring quality in its work processes, products (items), and services. The quality system provides the framework for planning, implementing and assessing work performed by the organization and for carrying out required QA and QC. (ANSI/ASQC E4-1994)

Quality Assurance (QA) – an integrated system of management activities involving planning, implementation, assessment, reporting, and quality improvement to ensure that a process, item, or service is of the type and quality needed and expected by the customer.

Quality Control (QC) – the overall system of technical activities that measures the attributes and performance of a process, item, or service against defined standards to verify that they meet the stated requirements established by the customer; operational techniques and activities that are used to fulfill requirements for quality.

Data Quality Objectives (DQO) – Qualitative and quantitative statements derived from the DQO Process that clarify study technical and quality objectives, define the appropriate type of data, and specify tolerable levels of potential decision error that will be used as the basis for establishing the quality and quantity of data needed to support decisions.

Data Quality Objectives Process – a systematic strategic planning tool based on the scientific method that identifies and defines the type, quality, and quantity of data needed to satisfy a specified use.

Example 1 - **BACTERIAL EXPOSURE AT SUNSET BEACH**

Citizens, county officials and environmental regulators are concerned that individuals using a recreational area, Sunset Beach, may be exposed to unacceptable levels of pathogens (disease causing organisms) while swimming. Sunset Beach is located on a tidal bay where the Green River discharges. A recently established concentrated animal feeding operation (CAFO) for swine is located approximately 2 miles upriver from the bay. There is concern that a large rainfall event, leaking waste treatment lagoon, or other adverse event at this farm could cause a discharge of animal wastes into the Green River resulting in exposure of beach users to pathogens.

A Citizens Advisory Group (CAG) represents the citizens and has indicated that beach users want to make their own decisions about swimming at Sunset Beach on any given day. CAG wants the County department of Health to provide daily guidance on conditions at the beach and realizes this will have to be based on sampling and analyzing the beach water.

CAG recognizes that measuring one or two organisms that are indicators of pathogen density is a practical and defensible method to indicate the levels of potentially pathogenic bacteria in the water. They realize information is also needed on risk-based standards or regulatory levels of indicator organisms that indicate low, moderate, and high risk of infection from swimming at the beach.

CAG has also been informed that the variability in pathogen density measurements among beach water samples can be large. Their technical adviser has informed them that this variability is caused in part by differences in the indicator organism density in beach water samples collected at different distances from shore, at different depths, at different times of the day, and on different days of the week. Other sources of variability may include water temperature, the Green River flow rate, tidal flow, weather conditions including wind speed and direction, the number of people in the water, and the number of domestic and wild animal present upstream of the bay.

Of concern to the CAG is the possibility that sample results may not accurately determine the risk of infection if the samples taken are not truly representative of the conditions of the beach. Specifically CAG want the county to post a sign each day that would indicate that

- the beach is open because the risk of infection is low,
- the beach is open but bathers should swim at their own risk because the risk of infection is increased,
- the beach is closed due to a high risk of infection as predicted by the bacterial indicator organisms.

CAG is also sensitive to the fact that the number of water samples that should be collected will depend in part on the public's tolerance for how often the sampling program yields data that result in incorrect decisions. Severe consequences to human health and potential loss of revenue from reduced beach use indicate that the probabilities of both types of decision errors must be very low.

Facts to consider

1. The State has established the following risk based health criteria for fecal coliform densities in marine/estuarine water samples:

- Safe for swimming $\leq$ 200 CFU (colony forming units)/100 mL,
- Unsafe for swimming $\geq$ 400 CFU/100 mL.

2. Three samples collected on July 4 of the previous summer showed fecal coliform values of 40, 160, and 840 CFU / 100mL.

DQO Step 1 - State the Problem

Problem: There is concern that contamination from pathogens may result in illness due to exposure.

Resources: County will fund sampling and analysis

Time Frame: Study should be conducted before the summer recreational season

Social or political Impacts: High public interest due to health risks. News media is involved as well as some state legislators.

DQO Step 2 - Identify the Decision

Study Question: Do indicator organisms exceed health-based standards in the waters adjacent to the beach area:
Actions that would result: No action, warning signs posted, beach closures
Decision Statement to resolve the problem: If indicators exceed health standards, then the beach must be closed.

DQO Step 3 – Identify Inputs

Information needed for decision: Which bacterium will be used as an indicator parameter. What is the action level?
Measurement characteristics: Geometric mean or maximum measurement.
Action Level: State water quality standard

DQO Step 4 - Define Project Boundaries

Spatial boundary: Beach area to a distance of X feet from a specified point.
Media: Water
Homogeneous Strata: Distance from shore; depth; salinity; etc.
Temporal boundary: Study before and during the recreational season
Time Frame: Decisions will be made during the recreational season.
When to Sample: Rain events, tidal flows before and during the recreational season
Practical Constraints: Sampling resources, magnitude of sampling effort, lab resources, holding time delays

DQO Step 5 - Develop a Decision Rule

Statistic used to describe the population: geometric mean or maximum value
Scale of decision making: swimming area
Action-triggering value: 400 CFU/100 mL
Decision Rule: If the geometric mean, over the summer recreational season, then signs will be posted around the swimming area to advise users of the risk.
Alternative: If the maximum value on a single day exceeds 400 CFU/100 mL, the swimming area will be closed until the maximum count falls below 400 CFU/100 mL.

DQO Step 6 – Specify Limits on Decision Error

<u>DQO Step 6.1 – Establish the Anticipated range of values</u>

Possible range of parameter: 3 samples collected on July 4 of the previous summers for FC values of 40, 160, and 840 CFU/100 mL.

<u>DQO Step 6.2a - Identify Potential Error Sources</u>

Facts:

- Swimming in waters with fecal coliform counts of 400 CFU/100 ml is unsafe
- Swimming in waters with fecal coliform counts 200 CFU/100 is safe
- There is a lag time between sample collection and final results.

- The measurement error for the test is =/- 5% (380 – 420 CFU/1000) ***(DQO)***

The sampling design error associated with the sampling is dependent on the number of samples to be collected and the location of those samples. Acceptable error will be established at 10% (360 – 440) ***(DQO)***

The total acceptable error rate is 15 % (340 – 460) ***(DQO)***

DQO Step 6.2b - Determine the Baseline Condition

Baseline statement: The Maximum FC levels are < 400 CFU/100 mL
Alternative Statement: The maximum FC levels are > 400 CFU/100 mL

Actual State	Baseline is True	Alternative is true
Decisions based on data		
Decide Baseline is true	Correct decision	Error (false acceptance)
Decide Alternative is True	Error (false rejection)	Correct Decision

DQO Step 6.2c – Outline the consequences of a Decision Error

Consequences of Decision Error:

False Acceptance:

Decision: Based on the data, the FC levels are less than the action level, when in reality, they are higher.

Consequence: People become sick because of exposure to the bacteria.

False Rejection:

Decision: Based on the data, the FC levels are higher than the action level, when, in reality, they are lower.

Consequence: Beach is close unnecessarily, which causes economic impacts to the merchants that depend of beach traffic.

DQO Step 6.3 - Establish which error has more severe consequence near the action level and the "gray area".

Since swimming in unsafe waters is detrimental, health risks are identified as having the greatest consequences if the decision to close or not close the beach is made.

The impact on local businesses would cause economic problems, if the beaches were closed inadvertently for long periods of time.

The established error rate is 340-460 CFU/100 mL. Therefore any count above 340 has the potential of being a violation.

Colony counts that approach the maximum level are problematic. As a conservative approach, the monitoring frequency will be increased, and warning signs posted about the potential health hazard.

DQO Step 6.4 Assign probability limits above/below the gray region that reflect tolerable probability for decision error

Measured Concentration	Decision	True Concentration Range	Error Type	Aversion	Tolerable Probability*
>400	Close Beach	0-340	False Rejection	Severe	Minimal
>400	Close Beach	340-360	False Rejection	Moderate	Possible
> 400	Close Beach	380 - 399	False Rejection	Minor	Gray Region
<400	Leave	400-420	False Acceptance	Minor	Gray Region

<400	Leave	420 – 440	False Acceptance	Moderate?	Possible
<400	Leave	440+	False Acceptance	Severe	Minimal

*May be expressed in terms of percentages

This table was constructed based on the error rates established in 6.2a above:

If the reported concentration >400 when in reality the true concentration is less than 340, the beach would be unnecessarily closed. The consequences would be economic (lost revenue) and political (loss of confidence). However, the chances of this occurring are minimal.

If the reported concentration is >400 when the true concentration is lies somewhere between 340 – 360 and the beach is closed, there are moderate consequences, primarily economic, although prolonged values can severely affect economics. While possible, this is not probable since the error associated with the test procedure is 5%.

If the reported concentration is >400, and the true concentration range is 380 – 399, there are minor consequences since the actual value is so close to the action level. However, because of the error associated with the sampling design, the situation may is likely to occur at some frequency.

If the reported concentration is < 400, the true value is actually between 400 and 420, and the beach remains open, there is a small chance that someone will become ill, and this situation is likely to occur.

If the reported concentration is < 400 while the true concentration is between 420 and 440, there is a greater chance of illness due to exposure, and a likelihood this might occur.

If the reported concentration is <400, the actual concentration is greater that 440, and the beach remains open, the consequences to human health are greatly increased. However, this situation is unlikely to occur.

DQO STEP 7 – OPTIMIZE THE DESIGN

Step 7.1 and 7.2 - Sampling Design

Establish a sampling strategy based on the spatial and temporal boundaries identified in 4 above.

We established that monitoring would occur in the vicinity of the beach area during the time period just be fore and during the recreational season. The monitoring before will establish if warning signs need to be posted. The monitoring during the season will determine daily beach use.

Ideally sampling locations should be established at the perimeter of the swimming area and at points that would be most influenced by upstream activities.

The baseline sampling frequency must address tidal fluctuations, and rain events, to determine the impact that these occurrences have on the water quality in the swimming area. Therefore, the time frame for collecting background samples should consider the frequency with which each type of event occurs. This will establish the time period for collecting samples before the recreational season. Rainfall events are typical from early spring through the recreational season. Therefore sample needs to begin before the anticipated first rain event.

It may be beneficial to see if there is a correlation between the suspected sources and levels of bacteria above 400 CFU/100 mL. If a correlation can be established, monitoring downstream of sources but away from the beach area may provide a warning signal before a problem occurs rather than after and during a violation.

During the season, daily updates are necessary. Therefore, sampling needs to occur at key locations daily.

DQO Step 7.3 – Review DQOs against sampling strategy

Thus far, we have established these DQOs:

- Action level of 400 CFU/100 mL
- Use of a method to test for fecal coliforms
- Error rate (bias of laboratory test) at +/- 5%
- Acceptable sampling error rate of +/-10%
- Overall acceptable error rate of +/- 15%.

We developed a table that outlines consequences if a false acceptance or false rejection occurs. We also know that there is a lag time for laboratory analysis.

Based on these facts, we need to design a sampling plan that will meet these objectives. We may also want to consider reducing the false acceptance rate. This may be done by adjusting certain of the inputs:

- Increasing the number of sites to be sampled will decrease the sampling error rate or increasing the frequency of sampling will also decrease the sampling error rate. However, any increase in samples will mean a proportional increase in laboratory analysis. If there is limited budget, this may not be feasible.
- We could also consider using a screening test (presence/absence) so that decisions to close could be made more quickly. However, by using a screening test, the analysis error may increase which would mean a potential increase in the false acceptance or false rejection rate.
- We could also consider monitoring critical locations with both laboratory tests, so that there is a faster decision response, verified with the longer test.
- If there is a good correlation between when elevated bacteria counts occur at points downstream of potential sources, and when (or if) the elevated counts are reflected in counts greater than 400 at the beach stations within a specified time period, some of the beach stations could be eliminated, using data from upstream locations to make evaluate daily reports.

Finally, one can adjust or minimize false acceptance or rejection errors by changing the decision point concentration, adding more test sites or manipulating other sampling or analysis variables.

DQO Step 7.4 – Documenting the Process

The entire data quality objectives process including the assumptions, acceptance/tolerance limits and sampling plan must be recorded in a document. This could be a quality assurance project plan, sampling and analysis plan, research proposal, or some other document.

Example 2 - Freshwater Inflow and Utilization of the Estuarine Tributaries of Estero Bay

Introduction:

This project supports development of a minimum flow and level (MFL) for Estero Bay that is mandated by Florida law (Section 373.042(1), F.S.). Estero Bay is a small, shallow bar-built estuary located on the southwest coast of Florida. Surficial freshwater inflow comes from five major creeks that are distributed along the eastern shore of the bay. From north to south these are Hendry Creek, Mullock Creek, the Estero River, Spring Creek and the Imperial River (Figure 1).

Florida law requires water management districts to develop a priority list and schedule for the establishment of minimum flows and levels (MFL) for surface waters and aquifers within their jurisdiction. This list, included in the District Water Management Plan for the South Florida Water Management District (SFWMD 2000) requires that minimum flows and levels for Estero Bay be established. The minimum flow is defined as the "…limit at which further withdrawals would be significantly harmful to the water resources or ecology of the area." The concept of a minimum flow of freshwater is especially relevant to estuarine systems. By definition, estuaries are transitional systems where freshwater from the land mixes with salt water from the ocean. Without freshwater an estuary ceases to exist.

Much of the volume of the Estero Bay is exchanged with the ocean during each tidal cycle. Salinity in the main Bay can be near that of the ocean, with low salinity water confined to the tributaries. Significant penetration of freshwater into the main Bay may occur at high discharges. Thus, each tributary may be described as a sub-estuary of the main system, as each may possess its own salinity gradient. Establishing a minimum flow is problematic. A minimum flow for the Bay as a whole might be established, with the total flow somehow apportioned between the five major creeks and rivers. This approach might protect the main portion of Estero Bay, but fail to protect the creek resources. Alternatively, a minimum flow might be established for each creek and river. This approach might protect creek and river resources, but fail to protect Estero Bay proper.

The South Florida Water Management District is using a resource based methodology to establish a MFL for Estero Bay. The approach is similar to the Valued Ecosystem Component Approach (VEC), which is the general name given to a method developed by the U.S. Environmental Protection Agency (USEPA, 1987) to guide monitoring programs in the National Estuary Program. The approach has been modified to focus on providing critical estuarine habitat. In some cases that habitat or VEC might be physical, such as an open water oligohaline zone. In other cases the habitat or VEC is biological and typified by one or more prominent species (e.g. an oyster bar or bed of submerged aquatic vegetation). A key aspect of the method is to relate the physiological and ecological condition of a VEC to freshwater inflow or its surrogate, salinity. These relationships form the basis for establishing a MFL. The data required to quantify the relationship between VEC and inflow or salinity in Estero Bay are lacking. This project directly addresses this data gap.

As applied here the VEC approach assumes that (a) environmental conditions suitable for VEC will also be suitable for other species and (b) that enhancement of VEC will lead to enhancement of other species. The use of multiple VECs reduces the uncertainty in the MFL estimate. In other words, as the number of VEC increases, the level of confidence that a MFL will protect the estuarine ecosystem increases.

Many species depend on estuaries during some part of their life cycle. One of the more salient ecological or resource functions attributed to estuaries is to serve as nursery areas for larval and juvenile life stages of many commercially important fish and shellfish. Within estuaries three nursery habitats have been recognized in the literature: wetlands, including salt marshes, mangroves, and mudflats, the low salinity region at the head of the estuary, and grass beds (Day et al. 1989). Freshwater inflow directly influences the temporal occurrence and spatial extent of the oligohaline or low salinity zone (LSZ) of an estuary. The District will use LSZ habitat as a VEC to establish a MFL for Estero Bay.

This RFP solicits proposals to examine (1) the role of open water, including the LSZ, in the Estuarine tributaries of Estero Bay as a habitat and nursery for larval and early juvenile fish and shellfish (crabs and shrimp) and (2) develop quantitative relationships between freshwater input or its proxy, salinity, and utilization of open water habitat by these and other important components of the pelagic food web. These quantitative relationships will be used to establish minimum flows necessary to maintain a nursery function for Estero Bay.

Background:

The interaction between freshwater input and the nursery function of an estuary is complex and has been observed on varying temporal and spatial scales. At estuary wide spatial scales and inter-annual time scales, empirical correlations between freshwater discharge and subsequent yields of fish and shellfish some years hence (Sutcliffe 1972; 1973; Wilbur 1992) have been found. In Chesapeake Bay, North and Houde (2001) reported a positive correlation between abundance of young of the year striped bass and white perch and discharge from the Susquehanna River during the spring. While such large scale relationships emphasize the contribution of freshwater input to successful recruitment they shed no light on mechanism.

Successful larval development and survival in the oligohaline or LSZ may determine the magnitude of recruitment into the adult population (North and Houde 2001). The oligohaline or low salinity zone (LSZ) of an estuary is created by freshwater input and has long been recognized as an important nursery area for fish and shellfish (Pearse and Gunter 1957; Gunter 1961; Day et al 1989). This region receives eggs, larvae and young from freshwater spawners, anadromous fish, estuarine spawners and even larvae spawned in the more saline lower estuary and ocean (Day et al. 1989).

The oligohaline or LSZ salinity range is typically defined as 0.5 to 5.0 ppt. However, some investigators extend the range to 10 or 11 ppt (Holmes et al 2000, North and Houde, 2001). The LSZ is often associated with the estuarine turbidity maximum (Jassby et al, 1995, North and Houde 2001).

Several factors appear to make the oligohaline zone a viable nursery. This zone may provide optimal salinity and temperature conditions for growth and development of larvae and juveniles (Pearce and Gunter 1957; Gunter 1961; North and Houde 2001). Low salinity itself and/or the high turbidity often characteristic of this region may provide a refuge from predation (Turner and Chadwick 1972; Chesney 1989). Lastly, the dissolved nutrients and detritus associated with freshwater input make the oligohaline zone highly productive. Freshwater input supports the detrital and phytoplankton based food chains that include larval and juvenile fish and shell fish (Holmes et al. 2000; North and Houde 2001; Turner and Chadwick 1972).

The process necessary for optimal growth, development and survival of larvae and juveniles within the oligohaline zone has been explained by the match/mismatch (Cushing 1975) or the critical period hypothesis (Hjort 1926; Fortier and Legget 1982). Thus, the presence of an adequate food supply, tolerable environmental conditions, and/or a refuge from predation when required ensures successful development and survival (North and Houde 2001).

Freshwater input has been hypothesized to influence success of early life stages in the LSZ and thus recruitment into the adult population. In the stratified upper Chesapeake Bay, high flow is hypothesized to create a well defined low salinity entrapment zone with an organic rich turbidity maximum that supports an abundance of zooplankton prey (North and Houde 2001). Fish larvae are retained in an optimal salinity environment that provides a rich food supply, and a refuge from predation through high turbidity (North and Houde 2001). During low flow, the entrapment zone is weaker and more diffuse. The turbidity maximum is less well defined and relatively depleted in organic matter. Production of zooplankton prey is limited and larvae may experience suboptimal salinities owing to reduced retention capacity (North and Houde 2001).

By contrast, in the well mixed estuary of the Parker River, phytoplankton bloom in the oligohaline zone during low flow conditions (river discharge <0.5 m^3/sec) and support a productive pelagic food chain. A long hydraulic residence time allows phytoplankton and zooplankton to accumulate in the upper estuarine LSZ. During high flow conditions, the bloom is flushed down the estuary (Holmes et al. 2000).

Project Objectives:

One of the prominent resource functions of an estuary that depends on freshwater inflow is to provide a low salinity zone where the early life stages of many organisms can grow and mature. Among these organisms are many commercially important fish and shellfish. The objectives of the project are to determine (1) the role of open water habitat in the estuarine tributaries of Estero Bay as a habitat and nursery for larval and early juvenile fish and shellfish (crabs and shrimp) as well as commercially important blue crabs and (2) develop quantitative relationships between freshwater input or its proxy, salinity, and utilization of open water habitat by these and other important components of the pelagic food web (jellyfish and arrow worms) and 3) describe seasonal fluctuations in the abundance and distribution of these organisms. Data that quantifies the response of biota to changes in freshwater inflow is lacking for Estero Bay and its estuarine tributaries. The information generated by this project will help fill this gap and will be used to establish minimum flows necessary to maintain a nursery function for Estero Bay and its tributaries

Scope of Work:

This is the first phase of a 2.5 year study, subject to District Governing Board appropriation, to evaluate the nursery function of Estero Bay and the role of freshwater inflow in maintaining this function. Two full years of sampling are expected and should be adequate to meet project objectives. The project involves identification and enumeration of planktonic organisms. This can be time consuming. Six months are allotted for start-up and to finish processing samples and report writing at the end of the project.

The work proposed by respondents must fulfill the project objectives. It is up to the respondent to propose an appropriate experimental design. However, at a minimum, the work should conform to the following constraints.

1) **Geographic areas to be sampled.** Approximate station locations are given in Figure 1. The distribution intends to describe both the Bay and its tributaries, representing a compromise between sampling only one or the other. The final number and location of stations will be determined by the contractor and district Staff.

2) **Parameters to be sampled at each station.** Biological parameters will be sampled using techniques that are generally accepted by the scientific community and will include:

a) Larval fish and fish eggs

b) Shrimp and Crab larvae

c) Early juvenile fish

d) Ctenophores and Arrow worms

e) Blue crabs, including adult and juvenile stages

f) Physiochemical parameters - Vertical profiles of salinity, temperature, dissolved oxygen using standard electronic sensors calibrated to manufacturer's specifications. Contractor must meet and follow the minimum requirements in the FDEP field Sampling SOPs for field testing (FT), field documentation (FD), and field quality control (FQ). Depth intervals shall be 0.1 m, 0.25 m, 0.5 m and continue at 0.5 m intervals to the bottom. A log of pre-calibration and post-calibration checks along with records of yearly maintenance of the instruments cables and sensors should be kept. Copies of these logs must be submitted with each data submission.

g) Florescence/chlorophyll - Electronic sensor measurement of florescence to estimate Chl.-*a* profiles that correspond to the above physiochemical measurements. It is recommended that the sensor be part of the instrument package (e.g., YSI) included with the physiochemical sensors. Grab samples to measure chl.-*a* will be needed to correlated with florescence measurements. Florescence sensors shall be calibrated to manufacturer's specifications and calibration records should be kept. Copies of these logs must be submitted with each data submission.

3) **Sampling frequency:** The temporal frequency of sampling should be minimal but still capture seasonal variation in all parameters measured and a range of freshwater inflow and salinity conditions. Sampling frequency can differ between seasons and species to adequately represent each species (group's) unique temporal variation.

4) **Sampling and reporting units**:

Taxonomy and enumeration - Fish eggs, larvae, and juveniles, shrimp and crab larvae, ctenophores and arrow worms shall be enumerated and identified to the lowest possible taxon. Density shall be reported in number per cubic meter of water when tow nets are used and a flow meter can me employed. The Contractor shall suggest other quantitative measuring units for other types of proposed sampling gear and methods (e.g., seines or traps for juvenile fish).

Blue crabs - Blue crabs may be sampled using traps. Organisms will be measured (carapace width) and sexed. Quantitative measuring units could be in terms of catch per unit effort or estimate of population size if a catch and release approach or other method is employed.

Other Parameters - Data for other parameters (including physiochemical and chlorophyll) will be reported in standard units.

5) **Methods:** Proposals should be specific regarding sampling methods, equipment, number of stations to be sampled and frequency of sampling. Methods should be well documented and commonly used, or established (reference for less commonly used methods should be included in proposal). Respondents also should provide a description of the expected level of taxonomic detail (e.g. eggs will be identified to at least the level of family). Major species of ecological and commercial importance should be identified to species, including: sciaenids; anchovy; clupeids; and blue crabs (genus level at a minimum).

Note: The thought process set out below is based on the provided information, and is not a group endeavor. The outcomes and conclusions may not be relevant or accurate based on assumptions. This example is primarily to provide an insight into the types of thought processes and logic patterns that can be used in the DQO process.

DQO Step 1 - State the Problem

Problem: Controlling the fresh water flow into Estero Bay may cause harmful impacts on the water resources or the area ecology.

Resources: South Florida Water Management District will fund a study to determine the minimum flow and level for Estero Bay.

Time Frame: 2.5 years

Social or political Impacts:

Estuaries are historical nursery habitats for fish and shellfish. If conditions do not support these types of habitats there are potential impacts on the seafood industry (commercial and recreational) if the population declines.

Section 373.042(1), F.S. requires development of minimum flow and level for the bay.

External Interests: Commercial seafood community, recreational boating and fishing community, merchants or services that the commercial and boating community (marinas, etc.), upstream communities that may be affected if the flow of fresh water into the bay is significantly restricted (flooding, dams, structures, etc.)

DQO STEP 2 - IDENTIFY THE DECISION

Major Study Question: What is the ideal minimum flow and level for Estero Bay
Study Question for this Phase: Is there a defined relationship between physiochemical parameters and the diversity and population of selected species (eggs, larvae, juvenile and adult stages).
Actions that would result: The Water Management District would use the relationship to determine fresh water inflow so that conditions are ideal to sustain a nursery habitat.
Decision Statement to resolve the problem: A strong correlation exists between selected physiochemical properties (primarily salinity) and the population density and diversity of selected species.

DQO STEP 3 – IDENTIFY INPUTS

Information needed for decision:

1. Physiochemical parameters: Dissolved oxygen, salinity, temperature and chlorophyll a. *(**food for thought**: should turbidity be added to this list?)*
2. In the water column quantify: types of fish eggs, larvae and juveniles; shrimp and crab larvae, juveniles and adults; ctenophores and arrow worms. *(**comment**: should be prepared to justify why these species were selected)*
3. Weather conditions
4. Tide Stage
5. Flow/volume of fresh water from each freshwater input

Proposed Measurements:

1. Enumerate all selected species and life stages. Identify to the lowest possible taxon. Estimate population density in terms of number per cubic meter. *(**comment**: should identify alternative ways of measuring density)*
2. Document the size and sex of the adult blue crab population and estimate population density based on number per square meter. *(**comment**: should identify alternative ways of measuring density)*
3. Measure and/or document meteorological and stage conditions during each sampling event.
4. Measure physiochemical parameters.

Measurement characteristics: Determine statistical correlation between physiochemical parameters (salinity, temperature, dissolved oxygen and chlorophyll) and population abundance of the eggs, larvae, juveniles of selected species. *[note: this needs a lot more thought – proper statistics to be used, proper input (average physiochemical measurements a specified time; average of same measurements in the entire basin, discrete measurements or average in certain salinity zones, population averages over the basin, how to quantify population density and diversity, how to take into account life cycles, etc.)]*
Action Level: Quantified freshwater flow may be determined by the results of this study.

DQO STEP 4 - DEFINE PROJECT BOUNDARIES

Spatial boundary: Estero Bay and the five major creeks along the eastern shore of the bay
Media: Water
Homogeneous Strata: Oligohaline zone with vertical profiles of physiochemical parameters that are monitored at the surface at 0.1, 0.25 and 0.5 m, and continue to the bottom at 0.5 m increments.
Temporal boundary: Study must be conducted to address tidal and seasonal changes in water flow, salinity conditions, weather, and species-specific temporal life cycle variation.
Time Frame: Study must be performed for at least one year.
When to Sample: All tidal stages under both flood and low flow conditions.
Scale of Decision: *Note the areal extent of the oligohaline zone (assumed to be most conducive to nursery habitat) needs to be decided. It is the whole bay, or some part extending out from the freshwater sources?*
Practical Constraints:

1. Sampling resources and magnitude of sampling effort.
2. Unusual weather conditions during planned study period (prolonged drought, prolonged rain, hurricane, etc.)

DQO STEP 5 - DEVELOP A DECISION RULE

Statistic used to describe the population: ***Food for thought****: see measurement characteristics under DQO Step 3 above. How the data are grouped and compared will have an impact on the outcome.*
Scale of decision making: *This is again dependent on the ideal areal extent of the oligohaline zone.*
Action-triggering value: Physiochemical parameters with at least a 75% correlation to the greatest species diversity, and population density will be selected for phase II of this project. *Needs further refinement – consider different species during different seasons, or base on the specific temporal life cycles; is 75% correlation good enough; are there too many variables, etc.*
Decision Rule: If there is a defined relationship (at least 75% correlation) between the selected population and the average daily values of any of the physiochemical measurements, then these measurements will be used as markers to determine the amount of fresh water inflow necessary to sustain the nursery population.
Alternative: There is no quantifiable relationship between the studies biota community and the physiochemical parameters.

DQO STEP 6 – SPECIFY LIMITS ON DECISION ERROR

<u>DQO Step 6.1 – Establish the Anticipated range of values</u>

Possible range of parameter:

1. salinity range of 0.01 to 30 parts per thousand depending on tide cycle and proximity to freshwater inflows.
2. current population statistics over time
3. literature searches relating population densities to “ideal nursery habitat” conditions

<u>DQO Step 6.2a - Identify Potential Error Sources</u>

1. Selected species are either too sensitive to changes or have a high tolerance to environmental change. *(This goes back to DQO step three and evaluating species tolerance. Also researchers may elect to select sensitive species, since these would have the best potential correlation).*
2. Sampling locations are not representative of the species population or bay conditions.
3. Sampling technique may bias population density, diversity and distribution.
4. Tide and stream stage, weather and other fluctuations are not adequately characterized.
5. Population taxonomy identification is incorrect.
6. Population metrics are incorrectly estimated.

DQO Step 6.2b - Determine the Baseline Condition

Baseline statement: There is a direct correlation of greater that 75% between the species population, density and diversity and the selected physiochemical parameter. ***Note:*** *The 75% is something that needs to be determined through the DQO process – it was used here for training purposes.*
Alternative Statement: There is no correlation between the species population, density and diversity and the selected physiochemical parameter.

Actual State	Baseline is True	Alternative is true
Decisions based on data	>75% Correlation	No Correlation
Decide Baseline is true	Correct decision	Error (false acceptance)
Decide Alternative is True	Error (false rejection)	Correct Decision

DQO Step 6.2c – Outline the consequences of a Decision Error

Consequences of Decision Error:
False Acceptance:
Decision: Conclusion that a relationship exists when there is no relationship.
Consequence: Fresh water inflow will be incorrectly controlled. The public will bear the cost of the installing structures intended to control flow, when, based on this study, none are needed.
False Rejection:
Decision: Conclusion that there is no relationship when one truly exists.
Consequence: Parameters will be excluded from use in determining input flows. If a parameter is a truly critical component to the health of the estuary, and flows are established based other considerations, the impact to the ecology will be proportional to any relationship that exists between the parameter and the selected measurement system.

DQO Step 6.3 - Establish which error has more severe consequence near the action level and the “gray area”.

The decision to not use the data will produce the greatest effect on the ecological balance in the bay and subsequent impacts on the commercial and pleasure boating and fishing communities and the commercial services that support boating.

DQO Step 6.4 Assign probability limits above/below the gray region that reflect tolerable probability for decision error

Observed % Correlation	Decision	True Correlation	Error Type	Aversion*	Tolerable Probability
0 to 70%	Don’t Use	75-100%	False Rejection	Severe	Dependent on relationship to final measurement system
70 – 74%	Don’t Use	75-100%	False Rejection	Minor	Dependent on relationship to final measurement system
80-100%	Use	0-75%	False Acceptance	Moderate	
74-80%	Use	0-75%	False Acceptance	Minor	

*Predicated on protection of natural resources.

DQO Step 7 – Optimize the Design

Step 7.1 and 7.2 - Sampling Design

Establish a sampling strategy based on the spatial and temporal boundaries identified in 4 above.
Sample collection strata must occur as described in 4 above. *Frequency must be determined. Decisions to be made: set interval (every week, month, etc) or condition driven (low flow, low tide, weather, etc). Note also that sample collection techniques need to be defined – only netting is mentioned. Each sampling technique*

needs to be evaluated so that the results represent the population of biota that is being enumerated (another food for thought: How confident do study members want to be about representativeness as it relates to the population?).
Evaluate the test organisms – do they represent the population or are they skewed to either more tolerant or intolerant? If skewed, was this deliberate?
Should some greater correlation other 75% be used? What are the consequences?

<u>DQO Step 7.4 – Documenting the Process</u>

The entire data quality objectives process including the assumptions, acceptance/tolerance limits and sampling plan must be recorded in a document. This could be a quality assurance project plan, sampling and analysis plan, research proposal, or some other document.

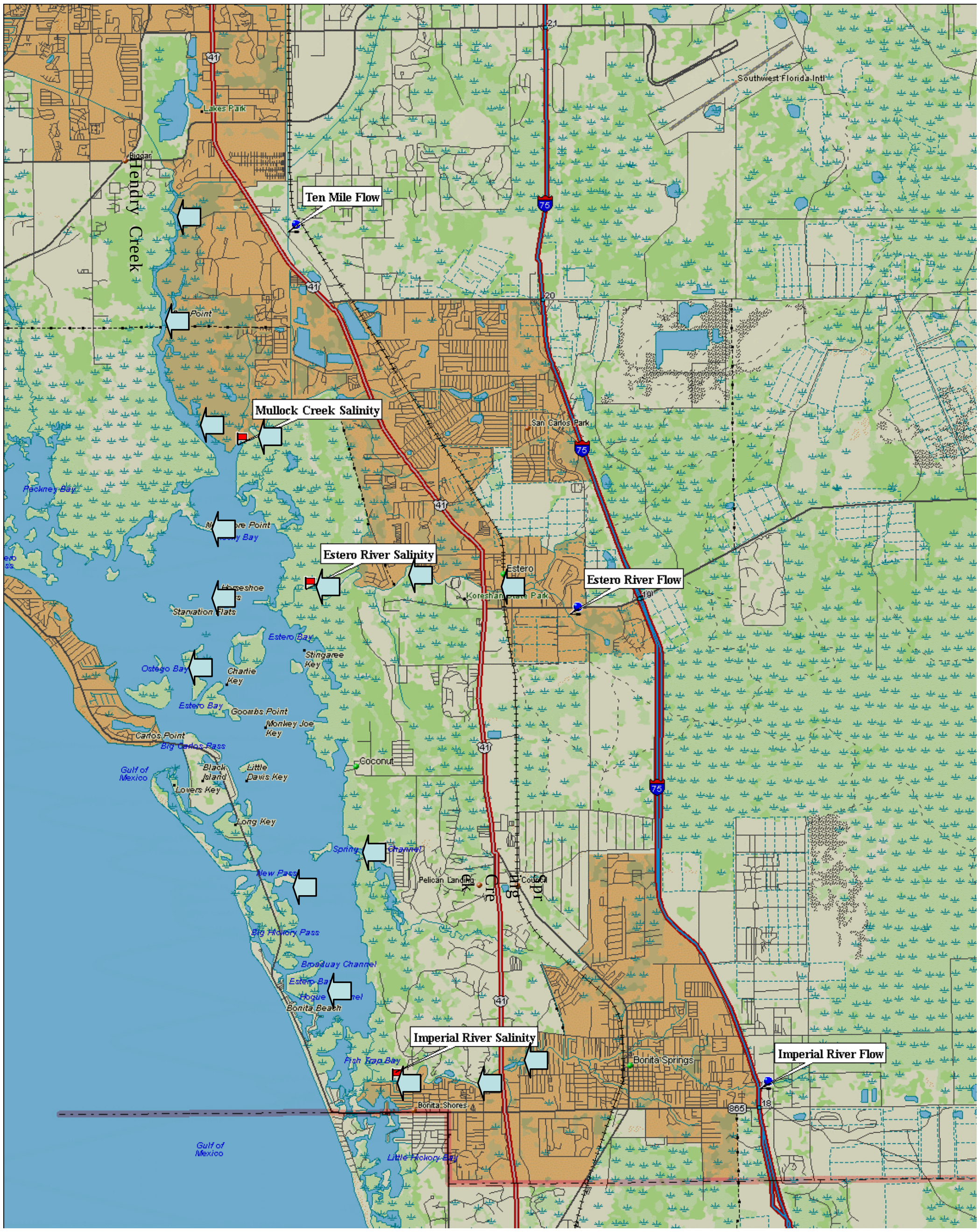


Figure 1. Proposed approximate sampling locations ( ).