

These are WRD Summary Comments made to all Networks after Review of the Water Quality Monitoring Components of the first 12 Phase 3 Monitoring Plans, and recommendations following the Austin, Texas February 2005 “Meeting of the Networks”. (Section 1, Monitoring Plan Recommended Examples; Section 2, Plan Suggestions per Checklist; Section 3, WRD Reviewer Comments on First 12 Plans; Section 4, Part B “lite”; Section 5, Estimated Flow Guidance Revision; Section 6, Protocol Content “Short List”; Section 7, Data Management Illustration and Discussion)

WRD Summary Comments to all Networks after Review of the Water Quality Monitoring Components of the first 12 Phase 3 Monitoring Plans and Austin, TX. National I & M Meeting.

Prepared by WRD Staff

March 8, 2005

Table of Contents:

Introduction.....	1
Section 1: Relatively Good Monitoring Plan examples from the Standpoint of Water Quality Monitoring	3
Section 2: Suggestions Organized According to Original VS Monitoring Plan Checklist.....	5
Section 3: Additional Comments from WRD Reviewers on the First 12 Phase III Plans.....	14
Section 4: Part B lite (Just the Basics) Review Checklist for Protocols and SOPs	15
Section 5: Revision of Estimated Flow Guidance	27
Section 6: Protocol/Protocol Narrative Content “Short List”	29
Section 7: Data Management Flow Illustration and Options Discussion	31

Introduction

WRD staff reviewed the first 12 Phase 3 monitoring plans at the Florida review-a-thon. The following is a compilation of lessons learned and summary thoughts (water quality aspects) that we believe will be helpful not only to those 12 networks, but also to other networks following along at various stages.

In general, most of the network monitoring plans met the content criteria specified in the Fancy-provided VS guidance review and checklist. However, some plans seemed more effective, concise, or complete than others. The plans varied in approach and to some degree in comprehensiveness. In general, the thickness of the main document and smallness of the print used may be inversely proportional to the numbers of people that will choose to read the document. As a rule of thumb, it may be best to move any non-critical information to an appendix or protocol or SOP and just summarize this information in the text whenever feasible.

One conclusion is that many of the critical details and much of the content is not in the broader plan text itself but in the protocols and SOPs. The first five chapters of the plans

all lead up to developing protocols. For those networks that have developed draft protocols, the real proof of how good some of the first phases of the planning process have been can be found in the protocols and attached SOPs. If protocols are inadequate, then much of the value of the hard work that led up to them will be negated.

So we not only reviewed the plans but also concentrated on finding relatively good examples of draft (developing, none are totally complete yet) water quality protocols and attached SOPs. The best protocol narratives were those that seemed “complete” and flowed well together. They tended to repeat and elaborate on issues summarized more broadly in Plan sections on the meaning of past data, questions to be answered, the type of issues tackled (regulatory or not), the target population, how the measurements relate to values to be protected, and how the random (or not) study design will assure that the samples are representative of the target population. The best examples also had all basic QA/QC aspects summarized in a QC SOP. If monitoring changes (emphasis or approach) do occur in the future, the protocol narrative should be key in understanding the basis for the initial monitoring decisions and how these changes may affect future data.

In the following sections, we provide some “good examples,” some additional summary comments, Part B lite (a checklist of just the basics, for those who think part B is too long), a revision of qualitative flow classifications to make them compatible with STORET, a short list of items to consider when preparing a protocol narrative and the steps we envision in water quality data management and various options at the network level.

We hope you find the following helpful.

Section 1: Relatively Good Monitoring Plan examples from the Standpoint of Water Quality Monitoring

The following are suggested as relatively good examples at this stage of development. They are not perfect or finished. All of them envision future improvements, but these seemed relatively good for this point in time and on the right path. If your documents are not on the list, it does not mean they were bad, but some of the documents just stood out in our initial, admittedly hurried (and not totally complete) 3.5 day first review. This list will be fluid and change as we become aware of additional examples as the networks develop them:

Water Quality Portions of the Central Monitoring Plan:

- GRYN: It had short concise technical writing, good illustrations, and thoughtful presentation of representativeness and random vs. non random study designs

Water Quality Protocol Development Summaries:

- NCPN Appendix T

Nearly Completed (Draft) Water Quality Protocols:

- SFAN Freshwater Quality Protocol, for a general model. Good discussions of all final checklist issues, comprehensive in content, includes all “Part b lite”(see below) issues
- GRYN Regulatory Water Quality Protocol (Early Drafts are available for review but are not yet in the Plan officially). This is a good example of incorporating State SOPs and looking specifically at GPRA issues. The other GRYN water quality protocols are further behind in development but show study design promise and may try to standardize many SOPs with the regulatory protocol.
- APHN is a good example of doing things the USGS way, protocols contain some very good insights/considerations in implementation, but needs some “part b lite” beef ups to become a really good example.

Best discussions of study design tradeoffs: Random vs. Non Random Approaches and Why the Study Design was Chosen:

- GRYN
- SFAN
- NCBN Estuarine Eutrophication Protocol

Good Park Specific WATER QUALITY vital signs discussions:

- NCPN, for a regulatory/CWA approach w/ comprehensive historical data analysis & waterbody summaries (see their appendices R & S)
- GRYN short concise technical writing with good illustrations

- SFAN/CAKN/GRYN - Protocols and SOPs (CAKN are early drafts but good start)
- NCCN & CAKN – Generally good monitoring plan content overall, rationale for approach etc.

Some additional lessons learned comments:

Most networks have chosen to “integrate” water quality monitoring with the other vital signs in the Phase 3 monitoring plans. Integration of all vital signs in such an overarching monitoring plan clearly adds some value. However, this integration has sometimes resulted in water quality monitoring details being hard to find or deferred to the protocol. This overarching, multi-component plan approach has made it somewhat difficult to understand how all the water quality monitoring pieces will fit together into a cohesive, implement-able plan (our responsibility and initial review goal).

Our focus has therefore now shifted towards the draft protocols (particularly the protocol narratives and attached SOPs). These should summarize how the water quality pieces fit together (where, what, when, why, who, costs, equipment used etc.). The protocol narratives and SOPs should provide the required metadata for archiving the data in STORET with specifics provided about data flow and management from the field to the network and WRD. This should include: type of labs used, sources of data for populating NP STORET/EDD, processing of data & QA/QC that will occur at the network level, and if a locally managed WQ database is planned for development at the network/park level (see Section 7 for generalized WQ Data Flow Diagram showing various options). It is also important that the water quality protocols repeat and further detail decisions on questions to be answered, target populations, how the study design will optimally answer questions and the various QA/QC requirements listed in the check list (see part b lite below).

There are several topic areas of special importance in water quality monitoring that need to be covered in sufficient detail in the protocols and attached SOPs. After reviewing the suggestions herein, many networks may need (or choose) to make changes in the way they approach the water quality component of their phase reports. The first 12 networks may choose to make only minor revisions to the monitoring plan and instead focus on providing more details in their aquatic protocols and SOPs. The following bolded checklist questions (from the vital signs generic checklist used to evaluate monitoring plans), either should be or have been addressed in the central monitoring plan but also will typically need additional development and documentation in the water quality/aquatic resource protocols and SOPs.

Section 2: Suggestions Organized According to Original VS Monitoring Plan Checklist

Overall Organization and Presentation of Monitoring Plan

If you produced a separate report for the water quality monitoring component, did you closely follow the five-part guidance developed by WRD? If the water quality section was integrated, did you provide QA/QC details and other needed water quality pieces considered important to WRD somewhere in the plan, protocols, or appendices, attachments, or “pointed to” references?

The easiest way to address the QA/QC details is to attach a separate QA/QC SOP to each protocol that adequately documents (per “Part b lite” below and SFAN example) how each of the following checklist items will be controlled and documented for individual parameters or suites of parameters: A. Representativeness, Target Population, and Completeness, B. Data Comparability (Internal/NPS and External/other regional data), C. Measurement Sensitivity (usually MDL and PQL) Detection Limits, D. Measurement Precision as Reproducibility and or Repeatability, E. Measurement Systematic Error/Bias/Percent Recovery (still wrongly called accuracy by some), and F. Blank Control Bias (for chemical lab work only)? Some of the discussions on representativeness and target populations should also be touched on in the plan chapter 4 (Sampling Design) and in the protocol narrative, but site-specific and parameter specific issues related to representativeness and target populations should be covered in those sections of a QC SOP in each water quality protocol.

Chapter 1 – Introduction and Background

Is there either a list of monitoring objectives, or a list of monitoring questions, that have an obvious connection to the monitoring goals and provide additional focus and understanding of the purpose of the network’s monitoring program?

Some questions or objectives are typically given in chapter 1, but in the protocol narratives and SOPs, the generalized questions typically need to be further detailed in both time and space. This is necessary to judge whether or not the questions make sense in comparison with summary discussions for representativeness and named target population(s) about which inferences will be made. Some of detailed questions are site specific and/or parameter specific. In these cases, the protocol should include the questions to be answered by long-term monitoring at these sites.

For water quality monitoring, is there a table or some clear, thorough presentation of all water bodies within the network that are listed on State 303d list, or are Outstanding Natural Resource Waters or have other special protective status?

Again, some generalizations are typically placed in chapter 1 of the plan, but more detail will typically be needed in the protocol narratives. What type(s) of monitoring is being considered? This year we have been told that that networks will need to pay more attention to Clean Water Act & GPRA goals (303d listed,

Outstanding Natural Resource Waters and/or other pristine or high value non-degradation waters (Category I or regulatory monitoring, see also Part A guidance). Is the monitoring being done for reasons related to the 303d or protection of non-degraded waters, or is the monitoring being done for a more general ecological vital sign reason (Category II)?

Part A update (Identification of priority impaired and pristine waters for the water quality vital signs monitoring component): Networks beginning Phase 1 data mining work are provided with the following updated information that replaces table 1 in Part A of the water quality monitoring guidance. Updated and more complete information for the identification of State-listed 303d waters is now contained for all parks in our Designated Use and Impairments database which can be accessed at <http://www1.nrintra.nps.gov/wrd/dui/>

. The site contains information in a geographic information system framework that includes: (1) CWA State-designated uses; (2) CWA 303(d) quality impaired waters and causes; (3) special designations recognizing waters of exceptional quality as defined in State water quality standards; and (4) hydrographic statistics based on the United States Geological Survey (USGS) National Hydrography Dataset (NHD). Hydrographic and 303(d) impairment statistics based on a combination of 1:100,000 and 1:24,000 scale NHD have been completed on Servicewide basis and are available for viewing here. Of these elements, the impaired waters inventory is the most complete and current. Information on State-designated uses and waters of exceptional quality are only available for a limited number of parks at this time and can be found in the completed park reports below. Networks are advised to still review the most current version of State water quality standards and regulations for designated Outstanding National Resource Waters or other special designations.

For water quality monitoring, has information content of available past aquatic data (for each water body being considered for monitoring) been adequately summarized in terms of hints of trends or other important issues of concern?

Some generalities need to be in Chapter 1, but most of the first 12 network plans were deficient in this area. Repeating hints (of trends or impacts) and making them more detailed in relationship to individual water bodies and questions would be appropriate in the protocol narratives.

Does the monitoring plan describe well the process that was used to determine monitoring objectives or questions, develop potential vital signs, and then prioritize and select vital signs to be monitored (additional detail on the process and criteria for ranking vital signs should be put in Chapter 3)?

Some generalities need to be in Chapter 1, but the initial protocols that were most successful repeated some of the basics and made them more detailed in the protocol narratives. If the monitoring is responding to 303d or GPRA, the protocol should state that on a site specific basis. If responding to a more broad,

park or network identified ecosystem concern or ecological tie-in, which should also be made clear.

Chapter 2 – No General Comments

Chapter 3 – Vital Signs

Is there a single list of vital signs that is consistent with the vital signs framework scheme and clearly shows the resulting “short list” of vital signs, including vital signs monitored by other programs and agencies? The list may include vital signs that the network hopes to implement in the foreseeable future, but may not currently be able to fund?

The short list should be at the end of chapter 3, and more detail about which vital sign characteristics/parameters will be measured at each site should optimally be summarized in the protocol narrative. For example, what aquatic biological monitoring (if any) is to be done at each water quality monitoring site and will that monitoring be sufficient to respond to state biocriteria? Will any continuous monitoring be performed to establish variability on a diurnal, seasonal, or annual basis?

Is there some obvious connection between the conceptual models and the high-priority vital signs that were selected for implementation? Are the high-priority vital signs all adequately justified through either the narrative or conceptual models, such that the average reader will be convinced of the value of them being monitored?

The basics should be in chapter 3, but for the protocols to flow together it would be desirable to include more detail on the relationship between the resources to be protected and/or desired future condition on a site specific basis in the protocol narrative or tables therein. Do protocol narratives repeat/make clear the resource to be protected, justification for the measurement selected and tie it to the measurements to be made on the target population?

Chapter 4 – Sampling Design

Has the network provided an overall sampling design that promotes integration of the various monitoring components over the long term and allows inferences to be made beyond the areas actually sampled? Is there an adequate description of any decisions to stratify or not stratify the network/park for various monitoring components?

Good example general discussions of study design tradeoffs: Random vs. Non Random Approaches and Why the Study Design was Chosen include the GRYN, SFAN examples in Chapter 4. Good examples of this being repeated and elaborated on in a protocol included the SFAN freshwater quality protocol narrative and the NCBN Estuarine Eutrophication Protocol.

One example of the detail included in each protocol narrative or SOP would be: What will be the **sampling frequency** by characteristic/parameter and site?

For water quality monitoring, does the broad plan contain a network map that shows the location of waterbodies to be sampled and an accompanying table that briefly summarizes

the characteristic(s)/parameter(s) to be sampled at each site, sample frequencies, who will collect the samples, and the protocol(s) to be used? Additional protocol-specific details (such as a map that shows the detailed location sampling sites within each waterbody and the associated latitude and longitude coordinates) should be placed in each protocol, but a brief overview of the overall sampling design (within the network as a whole) should be included in this chapter.

The type of additional detail that needs to be in the protocol narrative includes the exact location of the sites and water bodies to be monitored for individual characteristics/parameters or characteristic/parameter suites. Whereas Chapter 4 might include general maps and tables, the protocol narrative should include more detailed versions, including latitude and longitude coordinates (with the datum and method used to obtain the coordinates) and other GIS information along with a detailed map of each station location and some description of access, and why the exact sites were selected. Does the protocol adequately justify why the sampling scheme (random, judgmental, etc.) is optimal to help answer the detailed questions? Does the protocol narrative or a statistical SOP document the level of change that can be detected in light of the natural range of a variable versus the potential anthropogenic induced changes outside that range? Think through or determine (if possible) how will that be ascertained?

For water quality monitoring, data representativeness typically must be documented as a quality assurance basic. Does the plan (or the protocol for water quality monitoring that is referenced by the plan) adequately explain how the sampling scheme chosen will insure that the values obtained will be representative of the target population being studied? Is the sampling design appropriate to help answer previously identified questions?

The basics need to be in this chapter (4), with more details in the representativeness portion of the protocol narrative and/or the representativeness portion of the QA/QC SOP. The most successful draft protocols have the details in one of these two places. What characteristics/parameters are to be measured at each site and what is known about their natural variability in both time and space? Given what is known about that variability, how will representativeness be assured? Does the protocol narrative or an SOP document that the natural diel, seasonal, and annual variation in parameter measurements have been adequately considered in planning the monitoring? Have baseline conditions (parameter measurement ranges) been established at the monitoring sites (daily, seasonal, annual)?

Chapter 5 – Sampling Protocols

For each protocol, has the target population or “sampling frame”, and the sampling units, been identified?

The generalized basics need to be in chapter 5, with more site specific details as they relate to representativeness in the representativeness portion of the protocol narrative and/or the representativeness portion of the QA/QC SOP. The most successful draft protocols have the details in one of these two places.

For water quality monitoring, is there a quality control SOP associated with each protocol that adequately documents QC objectives for measurement sensitivity (detection limits), measurement precision, measurement systematic error (bias as percent recovery), data completeness (including adequacy of planned sample sizes and statistical power), and (if applicable for lab measurements only) blank control? Are instrument calibration details included either in the QC SOP or in a separate calibration SOP?

See Part B lite (below) for the basics that need to be in each QC SOP.

Do water quality sections of the plan or protocols sections include an explanation of how data comparability (a quality assurance basic) was considered in choosing which protocols and chemical labs to utilize? Do protocol SOPs contain enough field and lab method details to allow others to determine if data produced is comparable enough to other regional data sets to be considered credible by regulatory agencies interested in the data?

See Part B lite (below) for the basics that need to be in each QC SOP. Document which State Water Quality Regulatory Agencies have peer reviewed the Protocols and SOPs for data comparability and regulatory credibility aspects. Make sure the lab and field method SOPs are detailed enough not only to determine comparability aspects, but also are detailed enough that others could reproduce the methods exactly through changes in field crews and time.

Do the aquatic protocol SOPs adequately describe the details of all Sampling Protocols (Field and Laboratory), as well as equipment needs and operation, sampling techniques, sample preservation and handling and logistics?

Among the details appropriate for the protocol narrative: Who will be doing the sampling (network, park, cooperator, contractor, other agency etc.) at each site?

In the SOPs, among the details that are not mentioned in the checklist but whose documentation would be optimal are the following:

What specialized field equipment will be used (multi-parameter sondes, use of automated samplers, telemetry systems, other methods/instruments, field test kits of a screening level nature etc.).

Is any continuous monitoring on a site rotational basis a component of your program? Which type of flow/level monitoring is being considered at each site (how quantitatively measured or how qualitatively estimated – e.g. using “flow severity index” of EPA STORET for streams). Will continuous monitoring of core parameters (and possibly other) water quality parameters be a component of your program to establish baseline variation that synoptic sampling does not?

What key or critical instrumentation, automated monitoring systems, telemetry etc. will be considered or used to collect field measurements? Should that be considered for the more remote sites from a cost/benefit sense? Are there any anticipated future applications of these developing technologies to your monitoring program as it develops further?

Has the procedural complexity of the protocols and SOPs been given sufficient consideration? Will there be adequate staff time allocated to read, comprehend, and implement (particularly under changing or rotating personnel conditions) the required field sampling procedures, and populating field forms with acquired data and metadata and the subsequent transfer of the information to NPSTORET (templates) or through the STORET Electronic Data Deliverable (EDD) file format specifications to WRD?

Does each of the protocols include a narrative and one or more SOPs for data management and analysis that describe how data will be managed from the field to its final permanent repository (STORET)? Do the protocols contain examples of field forms that will be used, and describe how data will be entered, managed and archived in NPSTORET or other relational database (e.g. network/park customized database) and be transmitted annually to WRD (WRD STORET)?

The data management SOP in each water quality protocol should address how results obtained in the field and from the lab will be entered into NPSTORET and transmitted annually to the Water Resources Division for upload to STORET. For networks that elect to store their water quality data in their own local data system in addition to NPSTORET, the data management protocol should address how the network will generate the required STORET Electronic Data Deliverable (EDD) file format specifications.

Documentation and planning in the SOP needs to include matching the network's characteristics/parameters with the official standardized EPA list of 337,378 (as of 1/5/2005) characteristics (found in tblDef_TSRCHAR in NPSTORET or at <http://nrdata.nps.gov/Programs/Water/storetcharacteristics/storetcharacteristics.zip>). Performing this match to the EPA's standardized characteristic list when protocols/SOPs are being written will provide sufficient lead time for the NPS to request EPA to add a missing characteristic. Should you not find a characteristic/parameter that your network will be monitoring in the EPA's list, contact Dean Tucker (970-22-3516 Dean_Tucker@NPS.GOV) so he can formally request that EPA add it. Some common standardized (spelling must be exact) characteristics that many networks will be monitoring include:

WRD Core Parameters:

- pH
- Dissolved oxygen (DO)
- Specific conductance
- Temperature, water
- Flow (quantitative)

- Flow, severity (choice list); New estimated/qualitative flow guidance (see Section 5 below)

Some Other Examples

- Turbidity
- Fecal Coliform
- All ITIS taxa

In addition to itemizing the standardized characteristics/parameters that will be monitored, networks should explicitly define each characteristic (as appropriate using STORET terminology) in terms of:

- Medium (water, air, biological, sediment, soil, etc.)
- Sample Fraction (total, dissolved, suspended etc.)
- Unit of Measure
- Value Type (Actual, Calculated, Estimated)
- Field or Lab Measured
- Statistic Type (mean, max, min, mode, MPN, etc.)
- Duration (# of hours or days)
- Weight Basis (wet, dry, ash-free dry)
- Temperature Basis (5° to 95°C in 5°C degree increments)
- Particle Size Basis
- Detection Limit (MDL or similar)
- Lower Quantification Limit (PQL or similar)
- Upper Quantification Limit
- Field/Lab Analytical Procedure and/or Equipment
- Sample Collection Procedure
- Sample Handling Procedure
- Lab Sample Preparation Procedure
- Name/Contact Information for the Lab
- Is Lab EPA Certified for the Characteristic?

The field/lab analytical procedure/methods SOP and the data management SOP need to be consistent and detailed. The analytical procedure and/or equipment used to produce a result should include a reference such as “ASTM D1688(C) Copper in Water by GFAA”, “EPA 245.1 Mercury in Water by CVAA”, “USGS B0051 Fecal Coliform Bacteria-Presumptive Test-MPN Method”, or “Hach 8156 pH in Water” where the procedure/method is a recognized standard. For “non-standard” analytical procedures/method, a thorough description is required. For characteristics that required the extraction of a sample, a sample collection procedure should be specified for each characteristic. Additionally, a sample handling (preservation, transport, and storage procedure), a lab sample preparation procedure, the name/contact information for the lab, and whether the lab is EPA certified for the analyzing the characteristic should be provided.

Chapter 6 – Data Management and Archiving

Is the full data management plan for the network attached as an appendix or supporting document? Specifically for water quality monitoring data, does the plan specify how that data will be reported to the Water Resources Division annually via either NPSTORET or the STORET Electronic Data Deliverable (EDD) file format specifications for upload into the NPS’ Servicewide water quality archive, WRD STORET?

Among the details needed in Chapter 6 for water quality:

There should be an explanation of the data management process. How will the water quality specific issues in the data management SOP in each water quality protocol mesh with the overall data flow between park/network, WRD, STORET (use of field forms, electronic capture, NPSTORET templates, STORET Electronic Data Deliverable (EDD) file format specifications, data entry review and error checking etc.)?

Data Managers should ensure that all the metadata provided about characteristics/parameters in the sampling protocols in chapter 5 above is complete. If not, Data Managers should work with the protocol writers to specify the missing information to facilitate data documentation and upload to STORET. For documentation purposes, it might be useful to create a spreadsheet with the characteristics/parameters as row headers and the characteristic attributes (a. through s. above in chapter 5) as column headers. Have the water quality protocol writers fill in each attribute for each characteristic as appropriate so you have your metadata in one location. NPSTORET allows you to enter this information on its Metadata Template.

For additional information on vital signs water quality data management and archiving, consult: <http://www.nature.nps.gov/water/infoanddata/index.htm>.

Chapter 7 – Data Analysis and Reporting

The basics of the following should be included in chapter 7, with more details in protocol SOPs:

How will the data be analyzed, who will do the analysis and how often, and how will it be reported, summarized?

Has the planned data analysis – time series, trend detection, significance of change, summary statistics, etc. been adequately summarized and does it make sense in light of the questions and study design? Has adequate staff time been allocated to perform such analyses?

Chapters 8 & 9: No general comments.

Chapter 10 – Budget

Are costs for all the water quality monitoring activities adequately broken out in this chapter or best left to the protocols once the details are known? Are costs realistic when all aquatic monitoring activities are considered? Are costs for equipment replacement and maintenance on an annual basis along with consumables (e.g. calibration standards) been considered in the water quality monitoring cost estimates?

Section 3: Additional Comments from WRD Reviewers on the First 12 Phase III Plans

Protocol developments of the same general type should include as much collusion among networks as possible to avoid re-inventing the wheel and duplicating costs for development of largely the same/similar protocols. Every network will likely have their own specific protocol (unique aspects) needs, but there should be a lot of overlap. In water quality and macroinvertebrates, state protocols sharing a common ecoregion should be fairly similar as would EPA and NAWQA broader program protocols. NCPN and SCPN may be a good example of such collusion.

Spending time and providing details on the specifics of vital signs monitoring implementation, covered in the protocol narratives will be a key component of the success of the program but brevity in these narratives should be less an emphasis than content and comprehensiveness. If all the basics or key elements are in the protocols and SOPs, it is easier for WRD to determine if all the pieces of water quality monitoring fit together and will be implement-able (i.e. does the proposed implementation meet the reality test from an overall staffing, cost, equipment use, site access, cooperator involvement etc. analysis).

We feel the ultimate success of water quality/aquatic monitoring in each network will depend a great deal on planning, preparation and thought given to development of protocols, and SOPs. It is also important that concise direction and training be given to monitoring staff particularly, through each protocol narrative (restating of goals, objectives and monitoring questions to be answered, and justification for the methods selected etc.) in one cohesive document.

Changing/rotating monitoring staff will likely exhibit a wide range of training and experience levels over the course of a long-term program so clear direction laid out in protocols should be emphasized. Thus, our focus beyond this review will be the details of the protocol documents as they become finalized in the coming months to determine if sufficient clear direction is given.

Section 4: Part B lite (Just the Basics) Review Checklist for Protocols and SOPs

Roy Irwin, NPS, WRD

March 3, 2005 Draft

Please send peer review comments to roy_irwin@nps.gov

Introduction:

Note: The following is condensed from the much longer version of Part B at [http://science.nature.nps.gov/im/monitor/protocols/water qualityPartB.doc](http://science.nature.nps.gov/im/monitor/protocols/water%20qualityPartB.doc).

As suggested in generic VS guidance (K.L. Oakley, L.P. Thomas, and S.G. Fancy, 2003. Wildlife Society Bulletin 31(4), reprint at <http://science.nature.nps.gov/im/monitor/protocols/ProtocolGuidelines.doc>, all sampling protocols will include three basic sections:

- A. Protocol Narrative
- B. Protocol Standard Operating Procedures (SOPs), and
- C. Protocol Supplementary Materials.

So one item on any protocol checklist is whether or not the protocol follows the organization above, is complete, and has a table of contents that helps one determine where things are. Either the protocol narrative or a separate SOP should include a discussion of who will do the monitoring and who will train them and how often (recurrent training and is Quality Assurance/QA basic). Is there a SOP that clearly defines protocol variables and how to measure them?

The following text summarizes the basics of what has to be in water quality and other aquatic protocol SOPs (usually including a QA/QC SOP or QAPP SOP) to meet checklist (“Checklist for Review of Vital Signs Monitoring Plans,” on the Internet at <http://science.nature.nps.gov/im/monitor/docs/MonitoringPlanChecklist.doc>, hereafter referred to as “the checklist”) review requirements. It can also be used for the basics that should be included in Phase 1, 2, and 3 monitoring plan chapters. In most cases, the planning process is iterative, with very general statements in the plan chapters becoming more detailed in the protocols and SOPs.

Among the basics that need to be covered in the narrative and SOPs are (adapted from USGS wildlife monitoring website at <http://testweb-pwrc.er.usgs.gov/monmanual/>):

- WHAT are you going to measure?
- WHERE are you going to put your sampling points?
- HOW are you going to measure it?
- WHEN (HOW FREQUENTLY) are you going to measure it?

Some have found Part b to be too long, so Part b lite is an effort to provide networks with an extremely condensed version they can use for a last minute check to see if all the checklist basics have been covered in protocol narratives and attached SOPs.

The full version of part B can still be used for much more detailed discussions of the entire planning process and how all the pieces (review of past data, questions, target populations, representativeness, precision, detection limits, etc.) all fit together.

Summary of Information from Past Data

For water quality monitoring, has information content of available past aquatic data (for each waterbody being considered for monitoring) been adequately summarized in terms of hints of trends or other important issues of concern? The emphasis should not be on who is monitoring where, but what does it all mean.

Objectives and Questions:

It is easiest to plan monitoring if general objectives are rephrased into more detailed questions in the protocol narrative. The monitoring is then planned in such a way that questions can be answered with the data collected. In Chapter 1 of the monitoring plan itself questions can be somewhat general, but the questions should be made more detailed in both time and space when they are addressed in more detail each protocol narrative. These detailed questions should make sense in comparison with summary discussions for representativeness and named target population(s) about which inferences will be made.

This part of the protocol narrative is where the following checklist item should be addressed “Does the protocol narrative identify specific measurable objectives such as thresholds or trigger points for management actions?” If this issue is not addressed here, there should be a “point to” hyperlink showing where it is addressed.

This part of the protocol narrative should also summarize which questions and/or sites were selected to ensure monitoring of a 303d impaired water body or a very pristine water body that the park wants to keep that way. WRD has suggested that at roughly 2/3 of the sites should be in one of those two categories (see Part A guidance). What monitoring will be done to help answer GPRA reporting goals? A table listing 303d waters, with as much spatial detail (from where to where) as possible, should be included along with a note that the most recent WRD Designated Use and Impairments database (at <http://www1.nrintra.nps.gov/wrd/dui/>) has been consulted and that any differences with Network versions of the 303d lists have been logically reconciled.

Include Detailed SOPs for All Field and Lab Methods

Field Collection or Measurement SOPs should detail exactly what will be done in the field. Exactly how will field measures be done?

Lab SOPs should detail exactly how everything is done in the lab. If a standard State of USGS or EPA method is used, it should be written out or attached in its entirety. If the agency (EPA, USGS, etc.) changes the method, will the NPS also change in the same way? The SOPs should be detailed enough to allow third parties to reproduce the

methods and to allow determinations of data comparability (see further discussions below).

Examples of details to be included in protocols rather than in the central monitoring plan:

More details on sampling locations and method specifics.

For example, in chapter 4 a table might say that Chlorophyll *a* was the parameter to be monitored. Chapter 5 would give the method details, Chlorophyll *a* is to be monitored using field water collection procedures of the USGS field manual and lab method x (USGS Schedule 1637 method, or EPA method 445.0, or APHA method 10200H-4, or whatever is selected).

If flow or water level is to be recorded, will it be qualitative or quantitative?

What field instrumentation will be needed?

What pre and post season activities are required?

How will samples be collected and preserved, what containers will be used, and what minimum holding times will be used? Unless otherwise justified, use suggestions in 40 CFR Part 136 to 136.3 and appendices.

Attach a QA/QC SOP to Each Aquatic Protocol

A key quote in the checklist that “For water quality monitoring, there should be a quality control SOP associated with each protocol that adequately documents quality control (QC) objectives for measurement sensitivity (detection limits), measurement precision, measurement systematic error (bias as percent recovery), data completeness (including adequacy of planned sample sizes and statistical power), and (if applicable for lab measurements only) blank control.”

Since the checklist item (just above) is so brief, more information on what the WRD will be looking for in a QC SOP is provided herein.

For aquatic projects, and any other project where quality assurance and quality control (QA/QC) is important (where is quality not important?), QA/QC details should ordinarily be included in a separate QA/QC SOP attached to each protocol. It could also be called a QC SOP or a Quality Assurance Project Plan (QAPP) SOP. Whatever it is called, it should include sections explaining how each the following will be controlled and documented for individual parameters or suites of parameters:

- 1) Representativeness, Target Population, And Completeness
- 2) Data Comparability (Internal/NPS And External/Other Regional Data)
- 3) Measurement Sensitivity (Usually MDL And PQL) Detection Limits
- 4) Measurement Precision As Reproducibility And/Or Repeatability
- 5) Measurement Systematic Error/Bias/Percent Recovery (Still Wrongly Called Accuracy By Some)

6) Blank Control Bias (Usually For Chemical Lab Work Only)

If additional detail on any of the topics listed above is located anywhere other than in the QA/QC SOP, a summary of what will be done to control each of the issues listed above should be included in the QA/QC SOP and it should be made clear exactly where the other detail is in the monitoring plan or protocol. For example, if representativeness and target populations are fully explained in the protocol narrative or in the chapter on sampling design in the plan, then the representativeness section of the QA/QC SOP should clearly “point to” the section where the subject is fully covered. However, often the protocol narrative text will be more general, with the details related to individual sites and individual measurements/parameters (characteristic in STORET-speak) being more fully explained in the representativeness section of the QA/QC SOP.

To obtain data comparability, it is OK and even desirable to use well established QA/QC procedures of another federal agency (USGS, NOAA, EPA, NAWQA, or EMAP) or a State agency, but the source and the measurement quality objectives and SOP details for sensitivity/detection limits, precision, systematic error/bias, and blank control (for chemical lab work only) need to be listed in the QA/QC SOP because the other groups may change their SOPs as time goes along, and we need to have solid documentation of what we started with and/or what will be used in our long term monitoring.

QC details such as measurement quality objectives will often be different for individual parameters/vital signs/characteristics to be measured. However, there may be cases where the same QC measurement quality objective might be given for several parameters in a suite of vital signs included in one protocol. For example, if a network decided to use EPA marine EMAP QC SOPs to obtain maximum data comparability with EPA and State marine EMAP data, they could specify a precision repeatability measurement quality objective of a relative percent difference (sample size 2) or a relative standard deviation (sample size 3 or more) of 10% for several parameters to be measured in the field, including include pH, temperature, DO, specific conductance, salinity depth, light transmittance (PAR), turbidity, and Secchi depth. However, in many other cases measurement quality objectives would be different for different parameters and could simply be listed in a QC SOP table included as part of each protocol. Thus a protocol for water column parameters measured in the field would typically have different measurement quality objectives than a protocol for nutrient parameters measured in the lab. However, in both cases, a table in a separate QC SOP in each protocol could list the measurement quality objectives for each applicable parameter.

If the QA/QC of another agency is not adopted, or if the QC details come from multiple sources or are brand new, before completing the QA/QC SOP, a final check should be made to make sure that the measurement process will be controlled in some documented and defensible manner. The networks need to document what will be done for each of the issues listed below (doing nothing is not an option), but the networks need not “go overboard.”

At minimum, reviewers will be looking for the following basics.

Representativeness, Target Population, and Completeness:

Chapter 4 requirements for study design are well covered in the generic VS study plan guidance and the checklist (op cit.). Of particular interest in aquatic projects are the water monitoring-specific portions of the checklist. The checklist calls for a map of water bodies & sites to be monitored and an accompanying table of frequency of sampling events, and the number and different types of samples to be taken during each sampling event, all as part of chapter four.

The plan should briefly describe what is known about average values and variability in the various strata and how the sampling scheme will insure that the value obtained will be representative of the target population being studied (checklist, op cit.).

Typically the basics of the study design should be reiterated in the protocol narrative, usually with more detail. So whereas maps in chapter 4 of the central monitoring plan might show where water bodies are in the network, maps in the protocols should show location details of each sampling site, reach, (or area or strata where random selection will occur) in greater detail. What types/kinds of sampling will be done (what is involved spatially and vertically by waterbody)? Will vertical profiling (e.g. non-flowing water bodies/lakes) or stream cross sections, or continuous monitoring be used? Where will randomness be included in the process? Will grab sampling or compositing be done? Will sites be rotated through on some yearly or multi-year basis? Some details could go into tables in the protocol narrative.

In the protocol narrative, explain what the target population is in time and space detail. Explain how sampling will be done in a manner that insures the data collected will be representative of the target population and useful in answer identified questions. Justify the sampling design (random, spatially-balanced random with unequal (but not zero) chance of selection, other EMAP variants, biased design to sample historical bridge sites, etc.) and how it insures representativeness. Where in the process is randomness or stratified random sampling involved? Even if a historical bridge site is sampled, there should be some randomness in exactly where a stream is sampled near the bridge. How many samples will be taken and what percentage can fail and not result in sample sizes too small to answer the identified questions? These issues are just as important for biological monitoring as for water chemistry monitoring.

Good examples of probabilistic sample designs that allow inference to a broader target population include the GRTS design of the HTLN network. Another good solution and variant on spatially balanced probability design is the Systematic, Unequal Probability Sampling Design utilized by the Northern Colorado Plateau Network (NCPN). The SFAN freshwater quality protocol is a good example of reiterating the basics with more detail in the protocol narrative.

It would be appropriate to mention in the representativeness section what statistics are planned and how those fit in with representativeness, the target population, and the questions to answered, with a “point to” to lead one to more detail).

If some of the detailed explanations are in the plan itself and the decision is made that they need not be repeated in detail in the protocol, briefly recap them in the QC SOP section of representativeness and target populations, and provide a hyperlink or other “point to” in the QC SOP so the reader will know where to find the additional (more detailed) discussions.

Data completeness goals are typically given as percentages in tables and are developed estimating both needed sample sizes and then by working backwards from

estimates of the number of projected samples that are not likely to produce usable data, recognizing that it is very rare that 100% of all planned samples are successfully obtained and also pass all data acceptance criteria. So extra samples are planned and completeness goals are usually 80-90%.

Data Comparability (Internal/NPS and External/other regional data)

For Internal Data Comparability: What will be done to maximize temporal and methodological consistency in NPS data? Control typically involves limiting changes in internal NPS methods or timing of sampling to help insure our own newer data is comparable to our older data. When methods are changed, bias from the change should be documented in the data analysis SOP (see last section for more detail).

For External Data Comparability: What will be done to achieve comparability with other regional data sets (USGS, States, NOAA, CERCLA, etc)? Again, what will be done to insure our NPS data are comparable enough to the data from other state and federal agencies that need to be convinced our data is credible and comparable, given our purposes for monitoring? Is our NPS data comparable enough to other important outside data sets that the two sets of data could be combined for purposes of determining trends or making management or regulatory decisions? Has the chemical lab being used passed federal round robin blind sample checks or been approved by NELAP or the applicable State regulatory water quality agency?

A final check should be made to ensure that both the lab and the field method SOPs attached to the protocol are detailed enough to allow for reproducibility of exactly the same methods by third parties. Are they also detailed enough to allow judgments about the comparability of the data with the data of other agencies? Perfectly comparable data can be merged and analyzed together without introducing problems.

These issues are just as important for biological monitoring as for water chemistry monitoring.

Measurement Sensitivity (usually MDL and PQL Detection Limits)

Minimum Requirement: List (a table is fine) pre-project semi-quantitative and quantitative detection limits and how often they will be estimated once monitoring begins.

Unless otherwise justified, for lab measured chemical parameters, and any other parameters where very low levels can be encountered, list the target standard EPA method detection limit (MDL), semi-quantitative detection limit (For details See part B and Appendix B to 40 CFR Part 136).

Most labs and even some EPA staff and published EPA methods do not always use all the steps suggested by EPA to calculate an MDL, but most at least eventually use the central equation of Method Detection Limit (MDL) = t times S, where, t = the one-sided Student's t-value for seven replicate (precision repeatability) samples. In this equation $t = 3.143$, so $MDL = 3.143$ times the sample standard deviation for 7 replicate measurements of a blank. According to EPA, the MDL is the lowest value we really believe with 99% confidence is different than zero (a one sided comparison).

Unless otherwise justified, for lab measured chemical parameters, list the standard practical quantitation limit (PQL, same as LQL in STORET) as 3.18 times the MDL as the (low level) quantitative detection limit. If the network is going to use USGS alternatives such as reporting levels (RLs), explain how and how often they will be calculated and reported.

If a network wants to standardize with a regional governmental agency defining a PQL as 5 (rather than 3.18) times the MDL, the network could justify doing so on that basis. However, some EPA methods specify the use of 3.18. Top experts in the field now consider 3.18 to be sufficiently high to protect against false negatives, to be the value most commonly used, and to have other advantages over 5 (for details see Part B, or D. Helsel. 2005. Nondetects and Data Analysis: Statistics for Censored Environmental Data. Wiley. 288 pp., <http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471671738.html>, or Environmental Protection Agency. 2003. Technical Support Document for the Assessment of Detection and Quantitation Concepts, EPA publication no. EPA-821-R-03-005, available at <http://www.epa.gov/waterscience/methods/det/dqch1-3.pdf>).

No matter what detection limits are used, how they are calculated should be explained in the text and (once monitoring begins) in STORET metadata.

The text should explain how data below any of the listed detection limits will be handled, not only for reporting into data bases, but also for data analyses. The default recommendation (and one already adopted by some networks) is to handle them according to recommendations in the recent Helsel Book (D. Helsel. 2005. Nondetects and Data Analysis: Statistics for Censored Environmental Data. Wiley. 288 pp., <http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471671738.html>).

The Helsel book (op cit.) considers the modernized STORET default recommendation for writing to a database to be fully acceptable, so we are adopting this as a default NPS recommendation. Modernized STORET suggests that we not report into a database any value higher than the MDL but lower than the PQL. Instead, the detection condition field is set to "Present, below Quantification Limit". With that detection condition, STORET automatically enters "*Present <QL" in the result field. A major advantage of this approach is that no "estimates" are treated as quantitative when in fact they are not quantitative. In later statistical analyses these data between the two limits are best interpreted using either an interval-censored method (parametric), or a rank-based method (nonparametric) where all in-between values are represented as the same tied rank. The older recommendation of censoring to half the MDL is clearly no longer recommended. Helsel also gives recommendations for how not to report into data bases (for example, never report values below the Method Detection Limit as a negative ("-" sign followed by the actual MDL value, because someone invariably decides it really is a negative number (Helsel, op cit.)).

Values above the PQL are classified in EPA's modernized STORET database with the detection condition of "Detected and Quantified" This is ideal, and according to EPA STORET Staff, this is optimally the only choice which permits reporting a single number.

If an alternative USGS LT (long term) MDL is used for the semi-quantitative detection limit, document how it is calculated and how often

The MDL and PQL values are often not very relevant for field measurements situations where one is always up in the quantitative measurement range (above the PQL). It is not ok to use the lower part of the measurement range (in manufacturers

specifications) as an estimate of the MDL and for reporting to STORET, since the value is often zero. It is also not optimal to use the resolution specifications of the manufacturer, since there is no consistency in how resolution is estimated.

When one really needs a low level MDL for a field electronic instrument, one should dispense with the rough guestimate and calculate a standard MDL using the central equation on a blank. If no MDL or PQLs are entered into STORET, all measured values can be entered, but no STORET detection limit classification codes can be entered.

Since measurement sensitivity is a QC basic that should not be ignored, when no blank or low level solutions are available, or when one is always in the quantitative measurement range, how should we estimate measurement sensitivity? The answer is, estimate precision, but do so in a more rigorous way (based on 7 measurements) than you do when controlling precision more frequently for each sample batch (based on 2 samples) as explained in the next chapter (on precision).

Unless otherwise justified, for field measurements situations where one is always up in the quantitative measurement range (above the PQL), whether MDLs or PQLs are listed or not, it is appropriate to estimate an additional (more relevant) form of measurement sensitivity in the STORET metadata “analytical procedure description” text box. This alternative estimate of measurement sensitivity is a calculated estimate of NIST/ISO/APHA “measurement precision expanded uncertainty for a 99% level of confidence” (N. Taylor and C. E. Kuyatt. 1994. Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results NIST Publication TN 1297 (<http://physics.nist.gov/Document/tn1297.pdf>).

For estimating measurement precision expanded uncertainty, one samples one NORMAL (not a blank) environmental sample seven times, then the sample standard deviation is multiplied times 3.708 (the 99% confidence middle t value for sample size 7). This should be done at least once every sampling season, and more often until a reasonably consistent range is developed. The result should be called measurement precision expanded uncertainty to distinguish it from standard low-level MDLs and PQL estimates of sensitivity. The result of this calculation is functionally analogous to a PQL but uses the two sided t value because we are not only interested in one side (distinguishing a value from zero) but in the two-sided issue of how large of a difference between two individual values we can justify as actually being a true difference from one another and not due to measurement process noise. The result is used to logically estimate how many significant figures one should carry in a final result after all calculations are done.

Many biological inventory and monitoring projects have not historically estimated measurement sensitivity. However, there is often no reason why measurement of one sample could not be repeated seven times, or perhaps counted or measured by 7 different staff members. It may take some ingenuity in difficult cases, but one can usually find a common-sense way to adapt the thinking behind the paragraph above to come up with a functional analog for various types of biological monitoring scenarios. One can then use the 7 measures of one sample to (at least every so often) estimate sensitivity as measurement precision expanded uncertainty as explained in the paragraph above. The key is to try to make sense while still addressing the issue. Estimating measurement sensitivity is a QC basic.

Although not recommended in the Helsel book (op cit), for the special case of NPS analyses of “precautionary principle” comparisons with standards or criteria, one might choose to censor all data below the PQL to the exact value of the PQL, but that is only a very special (worst case, trying to totally avoid false negatives) case of data analysis, and one would not substitute the value of the PQL in a long term storage data base field for measured concentrations.

Measurement Precision as Reproducibility and/or Repeatability

Precision is not only estimated in a rigorous way every now and then to estimate sensitivity (as discussed above), it is also estimated much more often in a much less rigorous way (sample size 2 rather than 7, using environmental samples rather than blanks) to control measurement precision often enough to document that the measurement process is “in control” (not out of MQO performance specifications for precision).

Check list of items to be included in a separate QC SOP for each protocol is provided as follows. For each measurement done in the field or lab, are the following adequately covered?

- A measurement quality objective (MQO).
- Will the MQO be used as a data acceptance performance standard?
- What is the data comparability source of the MQO (State, USGS, EPA-EMAP, RCRA, CERCLA, CWA, etc.)?
- Is precision being controlled in the context of repeatability, reproducibility, reproducibility plus, or some combination (specify)?
- How will precision be calculated and reported?
- Will the raw number results be reported in addition to summary statistics like relative percent differences or sample variances? (Both should be reported).
- How often will precision be estimated and reported?

When controlling precision in batches or groups, QC samples are usually collected every 20 samples or so and then measured twice (duplicates) to control measurement precision on a regular basis.

Standard (short term, batch-specific or every 20 sample frequency) MQOs can often be summarized in a table, sometimes along with systematic error, method detection limits, and blank control MQOs (See SFAN freshwater quality protocol example).

Precision in context of repeatability is the scenario where nothing in the measurement process changes. For chemical lab measurements, repeatability MQOs are typically used, but if multiple labs, multiple instruments, or multiple staff become factors, it becomes precision in the context of reproducibility.

Precision in the context of reproducibility is typical of long term monitoring, since there will typically be changes in staff and instruments, or sometimes different staff and different instruments are even used in the same network during one season. So often precision is controlled as reproducibility.

Precision “Reproducibility Plus” is our NPS terminology for field duplicates when two samples that are not exactly the same are taken in close proximity in time

and/or space. Since the samples may not be identical, the “plus” part of the phrase is a tip off that an additional potential source of variability is present. In this case, two potential sources of variability are being lumped, lack of perfect measurement precision plus potential true sample heterogeneity. When taking this approach, it should typically be done in addition to (rather than instead of) more conventional estimates of measurement precision. The text or tables should explain how often both precision and precision plus will be estimated to determine how much of the variability is due to random error in the measurement process vs. how much is due to true sample heterogeneity. If only one type of precision is to be estimated and controlled, unless otherwise justified, it should not be precision reproducibility plus.

Regardless of the type of precision controlled, usually duplicate samples every 10-20 samples or every sampling batch, or every field sampling day (specify which) are used as precision QC samples.

The text should document the extent to which precision MQOs will be used as data rejection criteria (unless otherwise justified, they should be data rejection criteria). So if the MQO for a particular parameter is that a relative percent difference cannot exceed a plus or minus 30%, if the samples exceed that value, all values associated with that batch or that QC sample should be discarded, and recalibration or other adjustments should be done until the MQO can be met.

Many types of biological inventory and monitoring projects have not historically estimated measurement precision. However, there is growing recognition of the need to do so, and one can usually find a common-sense way to control and document measurement precision in biological projects. Often one can simply measure something twice to get a duplicate answer (see Part B for more detail).

Measurement Systematic Error/Bias/Percent Recovery (Still wrongly called accuracy by some)

For each measurement done in the field or lab, are the following adequately covered?

- A systematic error/bias measurement quality objective (MQO), such as %recovery must be within 80-120%.
- Will the MQO be used as a data acceptance performance standard?
- What is the data comparability source of the MQO (State, USGS, EPA-EMAP, RCRA, CERCLA, CWA, etc.)?
- How will systematic error/bias be calculated and reported?
- How often will systematic error/bias be estimated and reported?

Unless otherwise justified, MQOs should be rejection criteria. So if the MQO for a particular parameter is that a % recovery cannot be worse than 70-130%, then if the recovery 60% or 140%, all values associated with that batch should be discarded, and recalibration or other adjustments should be done until the MQO can be met.

If one value (say water color) is being measured to estimate another value (say chlorophyll a), how will bias and accuracy (including a precision component) be controlled and estimated? Will average observed to expected ratios be used, and how?

Will root mean square error techniques be used, and exactly how and how will the result compare to detection limits that use multiples of the standard deviation?

Many types of biological inventory and monitoring projects have not historically estimated measurement systematic error/bias. However, there is growing recognition of the need to do so, and one can usually find a common-sense way to control and document measurement bias in biological projects. One strategy sometimes used is to consider a senior expert's answer right or expected (100%) and a rookie trainee's answer wrong (see Part B for recommended strategies).

Blank Control Bias (usually applicable to chemical lab work only)

Unless otherwise justified, for lab chemical measurements of toxic chemicals, metals, pesticides, or nutrients, MQOs for blank control shall be listed in the QA/QC SOP.

For each chemical measurement done in the lab are the following adequately covered?

1. A blank control measurement quality objective (MQO), if applicable.
2. What types of blanks will be controlled (trip blanks, lab blanks, etc.)
3. Will the MQO be used as a data acceptance performance standard?
4. Will data reported be adjusted by adding concentrations found in blanks? If not, how will blank control be accomplished (reduce contamination and re-run the samples?).
5. What is the data comparability source of the MQO (State, USGS, EPA-EMAP, RCRA, CERCLA, CWA, etc.)?
6. How will blank control be calculated and reported?
7. How often will blank control be estimated and reported?

Biological inventory and monitoring projects have not historically done blank control. However, if the scenario of wrongly assigning a number value when the true value is zero seems likely, it might be possible to develop a common-sense way to control bias from blanks.

Include Calibration Details

Instrument calibration details should be included either in the QC SOP or in a separate calibration SOP (checklist, op cit.). If these details are somewhere else (perhaps in the field or lab method SOPs), there should be a "point to" in the QC SOP so that the reader will be able to find them.

Include a Data Analysis SOP

Are there recommendations for routine data summaries and statistical analysis to detect change? How often will reporting and trend analyses be done? Does this SOP or the protocol narrative describe the frequency of testing and review of protocol effectiveness?

The data analysis SOP should include a discussion of the data analyses (including statistics) planned, who will do them, how often, and ensure that adequate staff time and project funding is set aside for this very important task. Most of the proposed statistics should be worked out with a statistician before protocols are completed. Some networks have also appropriately said that statistics will usually be handled according to the recommendations of the following two text books:

Helsel, D.R. and R.M. Hirsch 1992. *Statistical Methods in Water Resources*. Studies in Environmental Science 49, Elsevier Publishing, NY, this one is on the net for free at <http://water.usgs.gov/pubs/twri/twri4a3/pdf/twri4a3.pdf>.

D. Helsel. 2005. *Nondetects and Data Analysis: Statistics for Censored Environmental Data*. Wiley (op cit., see detection limit section).

Archive Cumulative Bias from Method Changes in the Data Analysis SOP

Method, equipment, and personnel changes are inevitable in long term monitoring. The requirement of overlapping old and new measurement methods is in Oakley et al (op cit.), but is often overlooked. How long will the old and new methods be overlapped to determine changes in measurement bias or precision? It is suggested that single (identical) samples be measured by both the old and new methods (or by old and new personnel) at least 7 times (whenever possible), with the results of the average positive or negative bias (as well as any changes in precision as reproducibility expanded uncertainty) archived in a place that future data users can easily find it. It is suggested that the cumulative results of the bias over the years be detailed in the Data Analysis SOP in each protocol, with “point to” hyperlinks from other places people might look, such as the protocol revision log, each field and lab SOP for methods, the data management SOP, the data management section or data acquisition parts of the central monitoring plan, and the precision and bias discussions in the QC SOP. Even small changes in measurement bias can accumulate and become significant over time.

The goal would be for someone 100 years later to be able to discover the effect of the various changes in measurement bias. If a future user could look in the data analysis SOP and discover that 90 years ago there was a method change that resulted +2% change (on identical samples) from the previous method, then 80 years ago there was another method change that resulted in another change of +4%, then later a plus 3%, then later a -1%, they would be better equipped to separate true trends from the cumulative effects of numerous measurement method changes. This issue is important enough that some redundancy provided by the multiple “point to” links from other places seems prudent.

Include STORET Details in a Data Management SOP

Documentation and planning in the SOP needs to include matching the network’s characteristics/parameters with the official standardized EPA list of 337,378 (as of 1/5/2005) characteristics (found in tblDef_TSRCHAR in NPSTORET (for questions, contact Dean_Tucker@NPS.GOV) or at <http://nrdata.nps.gov/Programs/Water/storetcharacteristics/storetcharacteristics.zip>).

End of Part B lite. More detail on each of the topics in Part B lite is found in the long version of Part B.

Section 5: Revision of Estimated Flow Guidance

NOTE: New Guidance

(This is a revision to WRD Freshwater Workgroup White Paper RE: Qualitative or Estimated Flow Description)

In the WRD Freshwater Work Group Subcommittee White Paper "Recommendations for Core Water Quality Monitoring Parameters and Other Key Elements of the NPS Vital Signs Program, Water Quality Monitoring Component" (located at <http://science.nature.nps.gov/im/monitor/vsmTG.htm#TechGuide>), WRD developed a reporting scale and discussion for estimating flow of flowing water bodies. This estimated flow was based on a scale using relative % of bank full. At the time we were not aware that EPA also had proposed a qualitative flow severity index or scale ("flow severity") and this index is already built into the new STORET reporting scheme. To be consistent Servicewide and because many states have adopted the EPA flow severity index for qualitative flow reporting purposes (estimating flow) and loading to STORET for sites where quantitative flow measurements are not available, it makes more sense for the NPS Vital Signs program to adopt this same scale (flow severity) as well, rather than use another of our own creation.

Please make note and pass this qualitative flow scale along to your staff/cooperators involved in developing the monitoring plan or working on the water quality component protocols. You may remember, along with the core parameters, at a minimum, an estimate of flow/level is required at all water quality monitoring stations (in lieu of a quantitative flow measurement which is normally preferred). The EPA "flow severity" scale is listed below and should replace previous estimated flow guidance (per the Freshwater Workgroup Subcommittee White Paper). Please plan to use the text codes of Modern/New STORET and not the associated numerical codes used by Old STORET.

Table Starts on Next Page

(1) Flow, Severity (Choice list for estimation of flow in flowing water bodies). Note an estimation (at a minimum) of flow (e.g. rivers and streams) for flowing water bodies is a required characteristic/parameter where a quantitative flow measurement (preferred) is not available or obtained.

Flow, Severity (CHOICE LIST)

The choices are:

New STORET Code	WRD Description	Old STORET Code
Dry	No visible water in stream (typical of dry period for an ephemeral/intermittent stream).	1
No Flow	Discrete pools of water with no apparent connecting flow (at surface).	No old code
Low	Base flow for a stream or flow within roughly 10% to 20% of base flow condition.	2
Normal	When stream flow is considered normal (greatest time that stream is characterized by this in terms of flow quantity, level, or general range of flow during a falling or rising hydroperiod, but above base flow).	3
Above Normal	Bank full flow or approaching bank full (generally within upper 20% of bank full flow condition).	5
Flood	Flow extends outside normal bank full condition or spreads across floodplain.	4

Note: We have received feedback from NPS Hydrologists that documenting whether the sample was taken under rising or falling hydrograph conditions can also be very important (e.g. in case of normal or above normal flow conditions and possibly others), so this should be documented when known. The STORET characteristic for hydrograph status is called "Hydrograph Limb (Choice List).

Table starts Next Page

HYDROGRAPH LIMB (CHOICE LIST)

The choices are:

New STORET Code	STORET Description
BASE	Base. Old STORET Code 1.
RISING	Rising. Old STORET Code 2.
PEAK	Peak. Old STORET Code 3.
FALLING	Falling. Old STORET Code 4.

Section 6: Protocol/Protocol Narrative Content “Short List”

This list is provided as an additional check of what should be in the protocol narrative and other than the QA/QC basics in part b lite. The focus is on very practical considerations to be included in the protocol narrative rather than the central monitoring plan. We suggest liberal use of summary tables and maps to supplement text. There is some overlap of the following with the suggestions in the original monitoring plan checklist and in part B lite above, but WRD has not previously provided water quality-specific guidance focusing on the protocol narrative by itself, so networks may find the following narrative short list helpful as an additional last minute check for the narrative only. Herein we focus on what will be measured, where it will be measured, and how it will be measured, and other very practical issues.

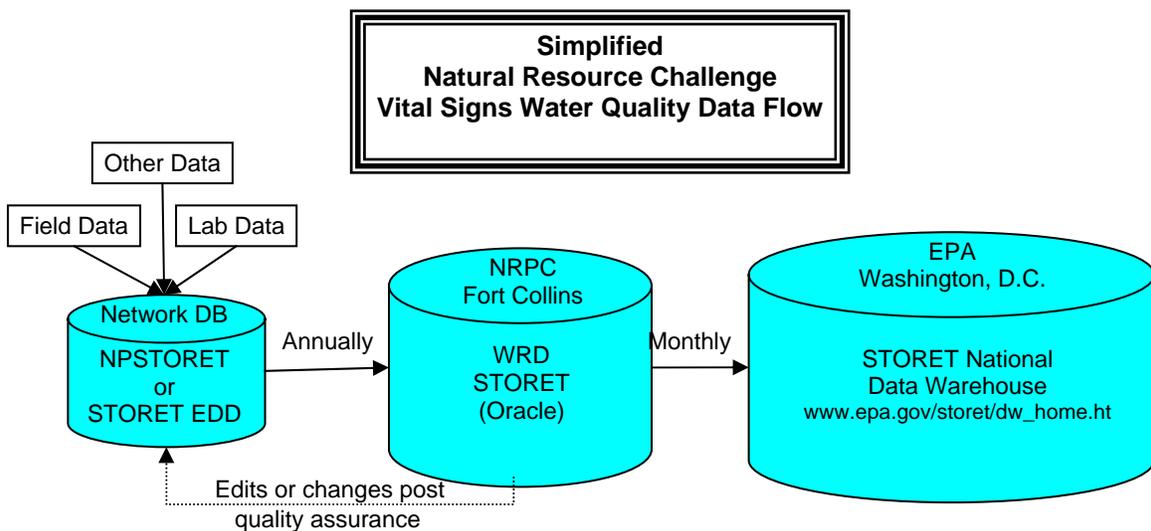
1. **What type(s) of monitoring** are being considered?
 - Clean Water Act & GPRA goal/regulatory focus -303d listed, Outstanding Natural Resource Waters and other pristine or “non-degradation” water bodies. (i.e. Category I)
 - Other aquatic monitoring Vital Signs (linkage or ecological tie-in) of a more general or “integrated” monitoring nature (i.e. Category II of Freshwater Working Group White Paper)
 - Identify and reference any existing or standard protocol(s) that will serve as the general framework for the monitoring program (NAWQA, State, EPA, other) and primary sources (if any) for SOPs.
2. **What/Where are the sites** and water body types to be monitored (provide latitude and longitude coordinates (including datum and method) along with a detailed map of each showing access) and why were they selected/basis for prioritization/stratification (what questions will be answered by long-term monitoring at these sites and what are your measurable objectives)?

3. **What STORET characteristics (water quality parameters)** are to be measured at each site and what is known about their natural variability (range)? This includes matching the network's characteristics/parameters with the official standardized EPA list of 337,378 characteristics (found in tblDef_TSRCHAR in NPSTORET at <http://nrdata.nps.gov/Programs/Water/storetcharacteristics/storetcharacteristics.zip>). All characteristics/parameters must be explicitly defined (metadata provided) as specified above in Chapter 5 – Sampling Protocols or as outlined at <http://www.nature.nps.gov/water/infoanddata/index.htm>. Identify which will be field measurements and which will be lab measurements.
4. **What will be the sampling frequency** by characteristic/parameter and site?
5. **What is the target population and how will the sample site frequencies be chosen to assure representativeness of the target population?** The reason for choosing the study design (random, judgmental, Generalized Random Tessellation Stratified Survey Design, etc.) and how it assures representativeness should be summarized in the narrative, with additional detail in the representativeness portion of the QC SOP for each water quality protocol (see part B and part B lite for more detail on this and some other items that should be in the protocol narrative). State what your measurable objectives are.
6. **Who will be doing the sampling** (network, park, cooperator, contractor, other agency etc.) and what coordination that will be necessary? Where will sampling operations be based/located within the network (one or more locations), who will be trained and by whom? What relationships have been developed with other agencies to minimize costs (lab analysis, staffing, and cost sharing opportunities)?
7. **What specialized field equipment** will be used (multi-parameter sondes, use of automated samplers, telemetry systems, other methods/instruments of a screening level nature) and is any continuous monitoring (<http://water.usgs.gov/pubs/wri/wri004252/>) on a site rotational basis a component of your program? Is a specially equipped and/or dedicated field vehicle planned for WQ sampling?
8. **Which type of flow/level monitoring** is being considered at each site (how quantitatively measured or how qualitatively estimated – e.g. using “flow severity index” of EPA STORET for streams)?
9. **What biological (aquatic) monitoring** (if any) is being considered at each site?
10. **What is the data management process** at the network/park level and how will data flow between park/network, WRD, STORET (use of field forms, electronic capture, NPSTORET, STORET Electronic Data Deliverable (EDD) file format specifications, data entry review and error checking, data archiving etc.). Parts of this should be covered in the data management SOP and/or in Chapter 6 – Data Management and Archiving and Chapter 5 – Sampling Protocols, but the basics of

how this will all work together should be in the narrative. For current information on archiving vital signs water quality data, consult <http://www.nature.nps.gov/water/infoanddata/index.htm>.

11. **How will the data be analyzed**, who will do the analysis and how often, what trend analysis methods will be used, and how will data be reported, summarized, displayed? The basics should be in the narrative, with additional detail and a data analysis/statistics SOP.

Section 7: Data Management Flow Illustration and Options Discussion



1. Note terminology change: “NPS STORET” is now referred to as “WRD STORET” to avoid confusion with NPSTORET (Access-based template database system used by networks for data entry and management prior to submission to WRD).

Discussion: Envisioned Data Management Flow (steps) and Various Network “Options”:

1. Collection of Field Data (e.g. field measured parameters, metadata, field samples for off-site lab analysis, digital photographs of site etc.)
 - Core parameter measurements “results” & metadata to Network via field forms and/or digital data sets (e.g. continuous monitoring data)
 - Lab samples to various off-site labs (private, USGS, State, other cooperating agency) for analysis
 - Select/limited samples for local processing by NPS/park-based lab (e.g. coliforms)
2. At Network Level: Merge all data from various sources (lab reports, field forms, digital files etc.) and perform QA/QC checks, data certification, etc.

- Populate NPSTORET (templates) as Network's Water Quality Database
or
 - Populate Network's alternative local Water Quality Database*
- And
- Populate and maintain locally a raw data file
 - Perform data review and analysis
3. Electronically transmit processed data (at least yearly) to WRD in one of two ways:
(1) NPSTORET or (2) STORET Electronic Data Deliverable (EDD) files.
- WRD performs limited further QA checks
 - Loads all NPS data to WRD STORET
 - WRD Electronically transmits water quality and related data monthly to EPA STORET National Data Warehouse (which is available to parks, networks, and the public for direct download)
4. Networks or parks access EPA STORET data (via downloads) to populate local databases on temporary or permanent basis to assist with management decision making. Networks and parks using NPSTORET to submit data to WRD can use its reporting, statistical, graphical, and export functionality as they are entering data to analyze data whenever desired.