Chesapeake Bay Program
A Watershed Partnership

Update on FAST_DUET_ESAR Process Development for the NTWQM NTN Project and Request to Discuss Critical Elements to Complete Process

Michael Koterba and Mike Mallonee

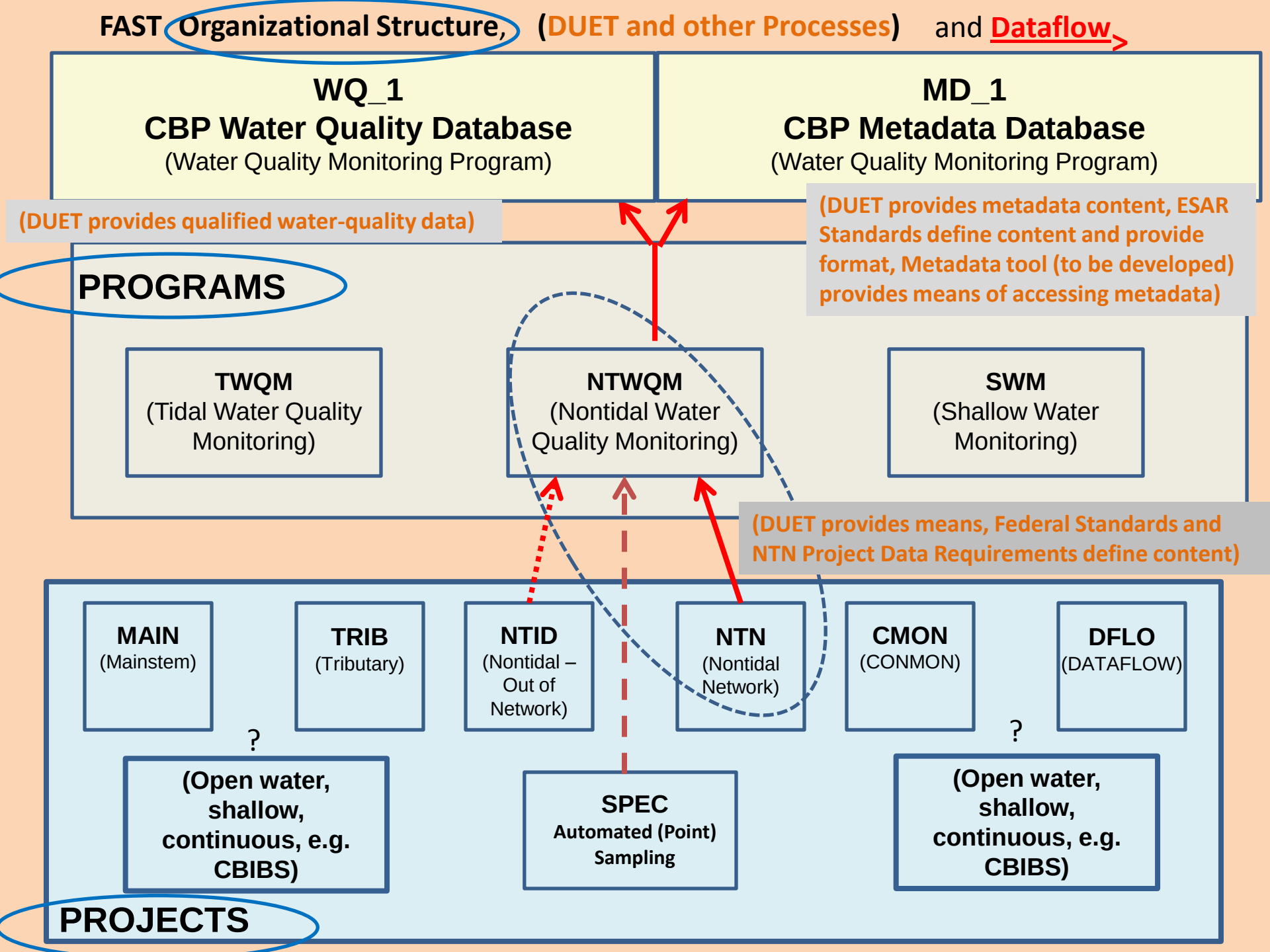
June 13, 2012

BACKGROUND INFORMATION: Please skim last seven slides if possible before meeting today … provides reasoning behind the development of FAST_DUET_ESAR Process.

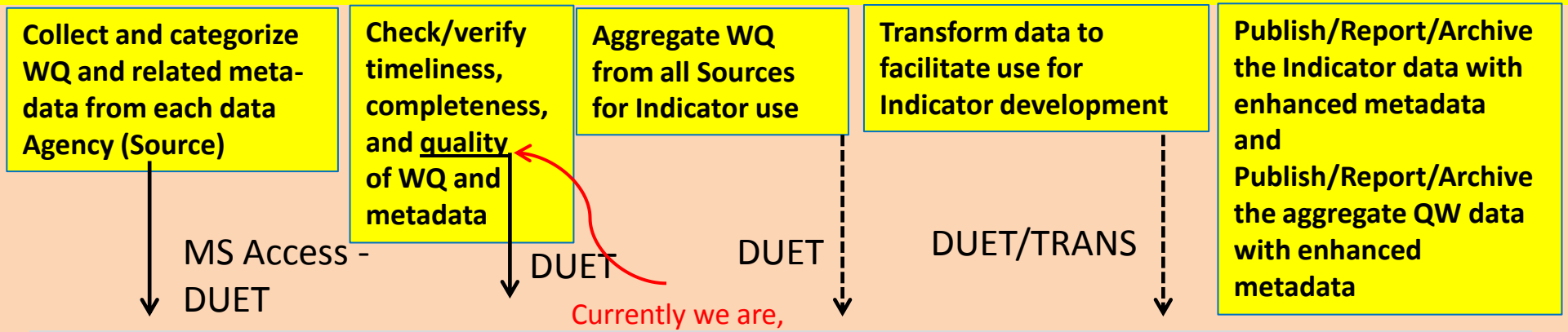
UPDATE

Status of FAST_DUET_ESAR Process for the Nontidal Water Quality Monitoring Program Nontidal Network Project

- Update on DUET Development
- Update on ESAR Standard Content Development
- Four Critical Issues to Discuss Complete Development


FAST Organizational Structure, (DUET and other Processes) and Dataflow
WQ_1
CBP Water Quality Database
(Water Quality Monitoring Program)
MD_1
CBP Metadata Database
(Water Quality Monitoring Program)
(DUET provides qualified water-quality data)
(DUET provides metadata content, ESAR Standards define content and provide format, Metadata tool (to be developed) provides means of accessing metadata)
PROGRAMS
TWQM
(Tidal Water Quality Monitoring)
NTWQM
(Nontidal Water Quality Monitoring)
SWM
(Shallow Water Monitoring)
(DUET provides means, Federal Standards and NTN Project Data Requirements define content)
MAIN
(Mainstem)
TRIB
(Tributary)
NTID
(Nontidal – Out of Network)
NTN
(Nontidal Network)
CMON
(CONMON)
DFLO
(DATAFLOW)
?
?
(Open water, shallow, continuous, e.g. CBIBS)
SPEC
Automated (Point) Sampling
(Open water, shallow, continuous, e.g. CBIBS)
PROJECTS

Generalized flow schematic of NTN Water Quality Monitoring data life cycle and development of FAST _ DUET _ESAR DATA Processing



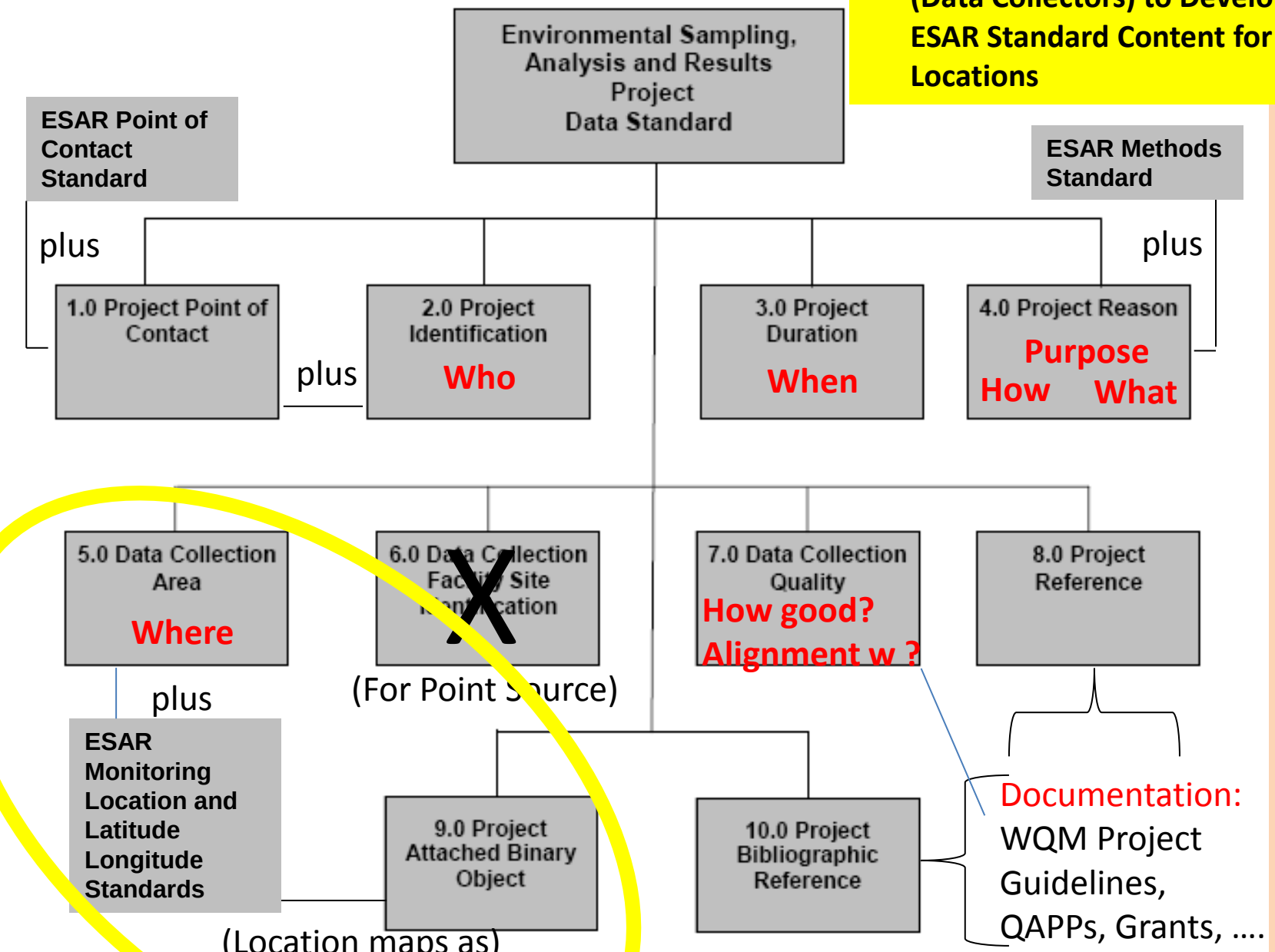
DUET Development Update:

- ❑ Have completed DUET data requirements and programming to
  - Collect and categorize data from Agencies (Data Providers) and Sources (Data Collectors) given expected data
  - Check/verify timeliness and completeness of data submission given expected data
- ❑ Have completed data requirements for QAC review to develop DUET programming.
  - DUET programming for QAC review will commence this week
- ❑ Have identified three DUET data requirements that require discussion with CBP Staff NTWQM Workgroup staff, and possibly selected TWQM and SWM staff.
  - Routine submission of replicate water-quality and field-blank data, including metadata, to globally assess and document the quality of submitted water quality data
  - Routine submission of metadata as necessary to help identify sampling locations other than normal, or sampling conducted under flow conditions different from those planned.
  - Routine reviews WQ parameters, and submission of metadata reflecting review results, that is needed by DUET to guide and WQ parameters calculated by DUET.

ESAR STANDARD CONTENT DEVELPMENT

Status: Most elements undergoing minor revisions continuously

- ❑ **Need Assistance of NTWQM Sources (Data Collectors) to Develop Uniform ESAR Standard Content for Monitoring Locations**



DUET data requirement issue #1. Submission of monthly field-blank and field-duplicate data, including related metadata, to assess quality of monitoring (water quality) data

NTN Project monitoring guidelines (Chapter V, Final Draft, Non-Tidal Water Quality Monitoring Procedures, 2008, p. V-12 p.):

6. Quality Control Samples

6.1. Field blanks: A field blank is an aliquot of DI water that mimics the sampling procedures. Prepare at least one field-filtered blank per month for both dissolved and particulate parameters. Pour the DI water into the sampler bottle, churn splitter, graduated cylinder and through the filters, exactly the same as the samples. If the concentration of an analyte in the field blank exceeds the laboratory MDL, equipment contamination should be suspected and investigated to identify the source of contamination.

6.2. Field duplicate: Two representative portions are taken from one homogeneous churn sample and processed identically. Field duplicate data provide an indicator of sample preparation and analytical reproducibility (precision). The minimum frequency for collecting a field duplicate sample is one per month or once for every 20 samples.

- To enable DUET to help assess and document data quality in relation to (a) contamination bias (field through laboratory) and (b) precision of water-quality measurements throughout the WQ data obtained each WY.

- Request 1-hour discussion that would focus on (a) the proposed DUET QAC review and use of WY2012 data and (b) possible potential improvements in the collection and use of blank and duplicate water-quality samples for future (WY2013 and beyond). There would be background information sent in advance to facilitate discussion.

- Discussion would hopefully include all NTN Project Sources (13 Data Collector Agencies) and selected CBP Staff (Monitoring, NTWQM Program, and QA Coordinator, WQ_1 Data Manager, and EA Technical Assistant)

DUET data requirement issue #2. Submission of routine metadata to help qualify how well actual monitoring meets or differs from selected monitoring requirements and data-quality objectives.

NTN Project monitoring guidelines (Chapter V, Final Draft, Non-Tidal Water Quality Monitoring Procedures, 2008, p. V-12 p.):

4.2. *Site Specifications*: As of 2007, 65 Primary NTN stations have been fully implemented, and 21 sites partially implemented (known as supplemental sites).

4.2.1. Primary sites are characterized as having:

4.2.1.1. A close proximity to a continuous stream flow gage so that the water-quality and discharge information are comparable,

4.2.1.2. Twenty water chemistry samples collected per year over a range of flow conditions (12 routine + 8 storm flow),

4.2.1.3. Total nitrogen, total phosphorous, ammonium, nitrate, phosphate and total suspended solids analyses.

4.2.1.4. Storm samples must also include analyses of suspended sediment concentrations, and each quarter, a sand/fine particle size analysis, and

4.2.1.5. Equal-width increment (EWI), isokinetic, depth-integrated sampling techniques to obtain representative samples.

4.2.2. Supplemental stations do not have storm sampling, but follow primary station criteria such as:

4.2.2.1. Sites are associated with a stream-flow gage to allow computation of loadings trends,

4.2.2.2. Samples must be collected at least monthly,

4.2.2.3. Parameters shall include total nitrogen, total phosphorous and total suspended solids, and

4.2.2.4. Use of isokinetic, depth-integrated sampling techniques to obtain representative samples.

Submission of routine metadata to help qualify how when actual monitoring meets or differs from selected monitoring requirements and data-quality objectives…continued

- Routinely obtain metadata from Data Sources to enable DUET to help categorize (a) when a sampling location other than that normally associated with a given streamgage is required because of low flow, high flow, or other issues, and (b) when sampling techniques and data collection reflect one flow condition, but actual streamflow conditions ultimately indicate another streamflow condition. (Use of (a) would be in tandem with ESAR Monitoring Location data; see below).

- CBP Watershed Modeling Workgroup have also requested information on the location of all monitoring locations relative to their assigned USGS streamgage.

- Request 1-hour discussion that would focus on (a) use of 5 new Event_Remark codes to characterize either (a) or (b) if and when either occurs, and (b) the intended use of the metadata. There would be background information sent in advance of the discussion to facilitate discussion.

- Discussion would hopefully include all NTN Project Sources (13 Data Collector Agencies) and selected CBP Staff (Monitoring, NTWQM Program, and QA Coordinator, WQ_1 Data Manager, and EA Technical Assistant) .

DUET data requirement issue 2. Calculation of selected WQ parameters by Duet Requires Antecedent Review of WQ Parameters Used in Calculations with Metadata that reflect review findings. Best if this review and findings are provided by Source (Data Collector) Laboratory.

DUET potentially can utilize over 24 different equations to calculate NTN WQ parameters related to N, P, S, and C. Calculated WQ parameters include those for TN and TP, which are required for selected Sources to meet NTN data requirements (see table previous slide). Other calculated parameters in combination with measured concentrations for these same WQ parameters, would provide data for that WQ parameter throughout one or more Agency domains (states). NTN Loads and Trends and CBP Watershed Modeling could need and make use of these calculated WQ parameters.

Concomitant with the calculation, is the need to assess whether or not the calculation is warranted. Selected calculations involve subtractions of measured WQ parameter concentrations. All measurements are to a degree imprecise. The calculated WQ parameter concentrations can be negative in value. To address this issue, the Tidal and Shallow Water Monitoring Programs have utilized their analytical laboratories and developed techniques under DUQAT to assess whether such negative values are within the expected precision of measured WQ parameters, or possibly reflect an underlying and undetected bias in the measured WQ parameters. DUET could utilize a similar or modified approach for NTN calculated WQ parameters. Regardless of the approach taken, additional metadata would need to be provided by each NTN Source (Data Collector) Laboratory for the WQ parameters calculated for their Stations. Laboratories that carry out the above, can rerun samples before holding times are exceeded.

- Request 1-hour discussion that would focus on (a) the proposed DUET calculated WQ parameters, (b) the metadata requirements needed from Sources to adequately perform and qualify calculated WQ parameters, and (c) proposed recommendations for future (WY2013) that if implemented could make the calculation of WQ parameters in the future (WY2013 and beyond), less problematic, and possibly consistent across the three WQMPs. There would be background information sent in advance of the discussion to facilitate discussion. (Same attendees as for previous two DUET issues.)

ESAR Monitoring Location Standard Data Requirement Issue: Develop Uniform Metadata Content For All NTN Monitoring Locations

Metadata content in this ESAR Final Project and supporting standard addresses the fundamental question of where NTN monitoring occurs. Required support standards qualify location information.

To address the data requirements of the above standards for all current (approximately 120) NTN WQ monitoring stations, and do so in a manner that provides simple, uniform, and consistent metadata for all 120 Stations, requires each Data Source to identify (plot) the location(s) at which they conduct their WQ monitoring on a digital coverage provided by the CBP. These locations include the normal WQ monitoring location, and if applicable, the locations for monitoring low flows, high flows, or an alternate location used whenever other conditions arise that render the normal monitoring location site inaccessible. This is a one-time operation, but the methodology developed also could be used for any new monitoring locations.

- Request 1-hour discussion that would focus on (a) the required standard content, (b) the proposed procedure to provide digital maps for to plot monitoring locations, and (c) the use of this information in relation to DUET Event-Remarks codes described above. There would be background information sent in advance of the discussion to facilitate discussion. (Same attendees as for previous three DUET issues.)

BACKGROUND

Chesapeake Bay Program
A Watershed Partnership

WHY ARE WE MODIFYING OUR NTN Project DATA LIFE CYCLE PROCESS TO A FAST_DUET_ESAR PROCESS ?

To develop a formal, standardized, automated and timely (FAST) process, which permits an efficient, timely, and complete Data Upload and Evaluation (DUET) of NTN Project monitoring water quality data and metadata, where NTN Project and process, water quality data, and metadata can be described in accordance with federal standards (ESAR).

To provide the information (metadata), as well as monitoring data, necessary to address the CBP NTN Project Data-Requirements and Data-Quality Objectives, as well as the business, enterprise architecture, monitoring alignment, and other recommendations provided in the Executive Order 13508, the Imler (2009, 2010) Enterprise Architecture Reviews, the USEPA TMDL process (2010), and the BAH Review of Non-Tidal and Tidal WQM Program Data Life Cycle processes (2011).

BACKGROUND

Main Data-Quality Objectives for CBP WQ-1Database

CHESAPEAKE BAY PROGRAM: STANDARD OPERATING PROCEDURES FOR MANAGING WATER QUALITY MONITORING DATA CHESAPEAKE BAY PROGRAM, (January 25, 2011, p. 5-6.)

The main data quality objectives for the CBP Water Quality Databases are to:

1) Compile water quality information collected by multiple agencies in a consistent format to allow comparable, basin-wide assessments;

2) Present complete records of field and lab data generated under EPA grants;

3) Document the original sources of data and all attributes associated with monitoring so that: a) each data point and associated quality-control data are available and are traceable and b) any significant change, including method and problem codes, to a program can be detected;

4) Assure that user data downloads and retrievals are complete and accurate.

Eight Fundamental Technical Questions Any Monitoring Program Should Be Enabled to Readily and Routinely Answer

- **For what purpose is monitoring being conducted ?**
- **What is being monitored ?**
- **Where is monitoring being conducted ?**
- **When is monitoring being conducted ?**
- **How is monitoring being conducted ?**
- **Who is conducting this monitoring?**
- **How good is the quality of the monitoring data ?**
- **How well does the monitoring data meet the monitoring data requirements (as specified in guidelines, grants, QAPPs, … the essential documentation) ?**

Readily implies readily discoverable and available answers—machine programmable language

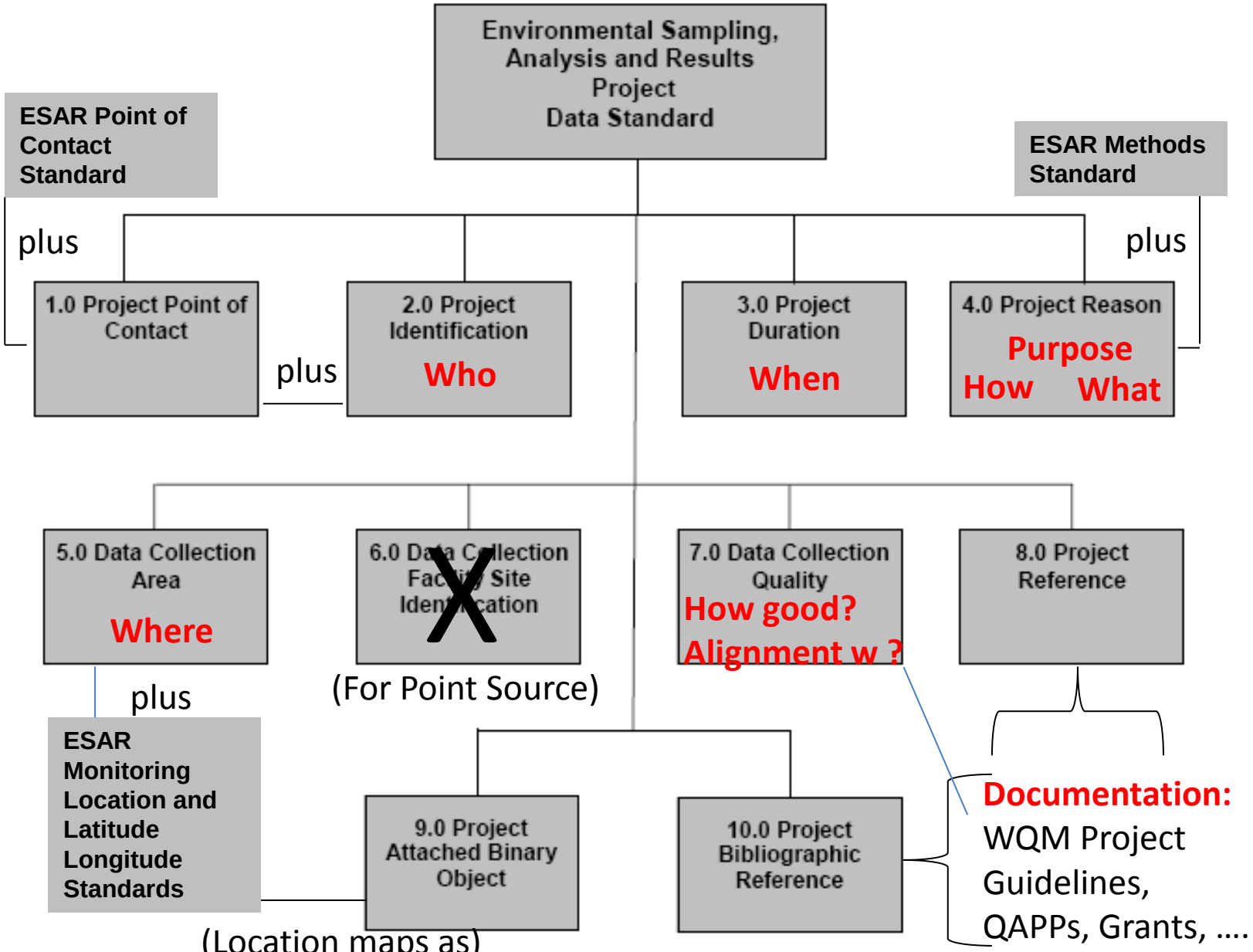
Routinely implies for the current and past data life cycles of the monitoring program

Use of USEPA ESAR (Environmental Sampling, Analysis, and Results) Standards Help Address these Fundamental Questions and Other Requirements

- They are federally recognized standards, with sub-elements that meet other federal standards—Data.gov and FDGC, and provide selected metadata for latter at CBP level.
- They specifically are designed to facilitate data and information exchange among partners, allowing one to adopt and complete only those standards needed to complete the necessary data and information exchange—a major activity for CBP Monitoring Programs.
- They contain minimal required sub elements, but allow for additional, sub elements to be added.
- With a minimal number of ESAR standards, and with a few sub-element additions to those standards, we can fully complete the selected standards, and answer all eight fundamental technical questions on a routine basis, for any CBP water-quality monitoring program. (Metadata content needed at monitoring source, agency (state), basin, watershed (CBP), or other scales).
- The ESAR standard content is designed to produce machine-programmable metadata (XML format), which makes the latter readily discoverable and readily available.
- The metadata content needed to routinely address all eight questions routinely must come in part from the monitoring sources (Data Collectors) through DUET, and, in part, from metadata added or generated by DUET.

BACKGROUND

CBP NWQMP Project: Map of One Data Life Cycle to ESAR Standards



BACKGROUND

Guiding Principles in Development of FAST DUET ESAR Process

- Data life cycle processes which create the water quality data and metadata for the NTN Project Bay Watershed Indicators will be as FAST as possible--formal, automated, standardized, and timely. (Provides economies of scale at many levels in human and physical resources for CBP staff and CBP Partners.)
- Formalization of these processes leverage existing automated processes (e.g., DUET process is being developed from the DUQAT process).
- Standardization at is achieved with the DUET process and ESAR (federally compliant) standards.
- Information (metadata) routinely obtained will be readily discoverable and available (e.g., USEPA ESAR Standard Content is in XML format).
- Implementation and operation of FAST on CBP Partners and their resources are to be kept minimal. (However, an increased need for resources for CBP Staff and or CBP Partners to implement or operate under FAST_DUET_ESAR processes is possible. There will be an increase in metadata requirements and management.)