ACTION TEAM MEMBERS

Requested Approach to Prepare for Action Team Meeting

Please read through background material (pp. 1-4)

Please read through Table 1 to become familiar with the issues (pp. 5-16)

Use the Attached Xcell File Titled **WY2012 NTN WQ Stream, Duplicate, Field Blank Samples.10Oct2012** as a reference where indicated.

**Data Collectors:** Revisit each issue in Table 1, and as a Data Collector, please summarize with respect to each of the issues, your ability, and relative to only your data collection, whether or not you could provide the listed Data Required for the WY2012 DUET Data Upload, and, if not the reason(s) why. If the required data are not provided in the WY2012 data upload, then it is assumed that the NTWG will resolves any issues related to that requirement ASAP in order to collect and provide the required data for the WY2013 data upload. A summary table is provided for your summary results (Xcell File: **Summary Table of Action Team Data Collector Response to Issues**. *If possible, please provide completed summary table to the CBP G4 (KF, MM, MK, and MLEY) by Monday Morning Oct 15th, so that we can summarize results for everyone for our Monday Oct 15th afternoon meeting. If not, please have completed Table to us by the meeting.*

# Resolving NTN Data Collection, Review, Qualifying of WQ Data and Metadata for FAST\_DUET\_ESAR Process

# Establishing the Monitoring Data Requirements for the Development of FAST Process, DUET, and U.S. EPA ESAR Standards for the Nontidal Water Quality Monitoring Program Nontidal Network

Developing a formal, automated, standardized, and timely (FAST) process, Data Upload and Evaluation Tool (DUET), and Environmental Sampling, Analysis, and Results (ESAR) Standards to obtain, review, qualify, aggregate, and publish annual (WY) water quality and related ancillary (meta) data for Nontidal Network monitoring requires defining data requirements, mainly for DUET programming. DUET can’t process data from an uploaded data submission if it can’t identify who collected it, who provided it, what type of data are being submitted, whether there are any known issues with selected data, etc., which in turn drive how DUET will review and process the data.

1. Timing, Software, and Technical Data requirements FAST\_DUET\_ESAR process were derived by following MRAT Response to STAC and the FLC Response to the Presidential EO:[[1]](#footnote-1)
	1. Maintain the continuity while increasing the stewardship of the current CB Nontidal water-quality network (NTN) and its data, and improve the management and accessibility of the data to the science and management communities (MRAT, 2009); and
	2. Construct a Chesapeake Bay Data Enterprise from the existing Chesapeake Information Management System (CIMS), to improve the archival of critical environmental information (mostly water quality data) used to assess progress toward environmental goals and support computer models, and to help ensure the availability and sharing of data, …, between CB partners to support the expanded computer models and assessment of progress toward goals (Federal Leadership Committee for the Chesapeake Bay, 2010).

Taking these recommendations, we then asked:

1. Key end users (USGS and CBP Nontidal and Tidal Modelers) what WQ data and metadata they needed and when they needed that data to help maintain continuity;
2. The USGS G4, notably Mike Langland, what procedures, processes, and criteria he (they) used to create the most recent NTN WQ datasets to help maintain continuity;
3. The CBP Quality Assurance Coordinator what QC data are required to assess quality of WQ and metadata to help improve CB partner comprehension and use of the NTN Data;
4. The CBP Data Manager what software and data are required given DUQAT (Data Upload and Quality Assurance Tool) is leveraged to create DUET, which is an integral part of improving the management (upload, review, qualifying, and aggregation) of the NTN data;
5. The CBP Enterprise Architecture staff what metadata were required given external reviews of EA and federal audits of the CBP, which also need to be met to improve assessment of progress towards goals;
6. What the ESAR standards required in terms of the metadata that would have to come from Data Collectors or from DUET processing of uploaded WQ and metadata, in order to meet required federal standards for data;

1. What the most recent NTN guidance (General, QAPPS, and 2008 NTN Field Procedures) indicated were the WQ and inherently metadata data likely being collected, a critical check on continuity and consistency; and
2. What NTN Data Collectors actually are collecting for WQ data and metadata and with some additional management could provide for WY2012 relative to the above data requirements, which all reflect increasing stewardship of the current NTN and its data, and the broader MRAT and FLC responses.

Given all the above, the CBP G4 (KF, MM, ML, MK) scaled down the timing, software, and technical data requirements needed to develop FAST, DUET, and the ESAR standards. However, relative to the technical data requirements, mainly for DUET development, we identified some discrepancies between the NTN guidance (mainly the most recent (Nov 2008) NTN Field Procedures and what different Data Collectors currently collect and could provide (Table 1).

An Action Team is to review these discrepancies and categorized them from being possibly relatively easy to more challenging to resolve. Resolution of these discrepancies could require modifying the NTN guidance, one or more Data Collectors modifying what they are doing, or both. All have some bearing on the development of FAST\_DUET\_ESAR process, most notably DUET, for the WY2012 data upload.

The NTWG is to then review the listed discrepancies (Table 1) to improve the NTN Monitoring in several key ways. From a high level perspective, the value of the NTN Monitoring data is ever increasing. The NTWG has an expanded set of NTWQM Program objectives that require these data (MRAT, 2009):

1. Measure and assess the status and trends[[2]](#footnote-2) of nutrient and sediment concentrations and loads in:
	1. Major tributaries and sub watersheds
	2. Selected tributary strategy basins;
2. Provide data suitable for the assessment of factors affecting nutrient and sediment status and trends from major pollutant source sectors;
3. Measure and assess the effects of targeted management and land-use change;
4. Improve calibration and verification of partners’ watershed models; and
5. Support spatial and topical prioritization of restoration and preservation.

For selected objectives, the use of NTN data as actual “indicators” of improvements in or deterioration in CB nontidal WQ are increasing. As an example, a recent presentations made in the last NTWG meeting illustrated how analyses of the low, moderate, and high flow-dependent trends in yields either align or differ, and how such analyses combined with CBP or Sparrow modeling could factor in to the 2017 assessment. In addition, the NTN has markedly expanded in station numbers, which means the NTN WY WQ datasets are increasing in value to other CB partners at regional, state, and local levels. Also with the above improvements came a considerable increase in federal funding for the NTN.

From a purely pragmatic perspective, resolving these discrepancies will lead to consistency between actual and described monitoring practices. It will provide all Data Collectors the opportunity to be on the same page, and provide some Data Collectors the chance to benefit from how other Data Collectors have addressed selected discrepancies. It will enable DUET data requirements to be clearly defined and DUET programming to proceed for the WY2012 data uploads. In addition, the increase in the value and investment in the NTN Monitoring data comes with an increased demand for access to the NTN data and accountability. The FAST\_DUET\_ESAR process directly addresses both these issues, and does considerably more in that it

1. Routinely obtains NTWQM Program-NTN Project Water Year (WY) monitoring data in a timely manner;
2. Routinely constructs fully reviewed, qualified, and standardized NTN Project water quality and ancillary (meta) data databases in a timely manner;
3. Routinely provides these databases to all core and extended CBP partners for internal and/or external use as appropriate in a timely manner; and
4. Routinely addresses the following fundamental questions often asked of any long-term monitoring program for each WY in a timely manner at more than an individual Data Collector level: (a) what is the purpose of the monitoring being conducted? (b) what is monitored? (c) where is monitoring conducted? (d) when is monitoring conducted? (e) how is monitoring conducted? (f) who is conducting monitoring? (g) what is the quality of the monitoring data, and (h) do the monitoring data meet monitoring data requirements?

Given all the above potential benefits, resolving these discrepancies for all Data Collectors can markedly increase the value of the NTN Data among core CBP (Goal Improvement Team) monitoring programs and many other CB partners.

However, currently the attached discrepancies highlight differences in the way WY2012 data were collected, managed, reviewed, and qualified among current Data Collectors, as well as in relation to the existing NTN Monitoring guidance. By reviewing these discrepancies, the NTWG, and in particular all NTN Data Collectors, can recognize where there needs to be modifications in the guidance, in their data activities, or both to better align the two; in addition, they can establish what each Data Collector can and cannot provide DUET in relation to the WY2012 data upload, which by default if unresolved, means they need to still be addressed with respect to WY2013 Data Collection (now just beginning) and the WY2013 data upload.

To help in this effort, Table 1 (below) provide a description of each discrepancy, why it is an issue, how it might be addressed, and how it relates to what DUET would require for the WY2012 data upload. Finally, in relation to the latter, a number of DUET requirements reflect the management of data, which no recent NTN guidance has sufficiently addressed. Thus, the resources required for each NTN Data Collector to not only collect but manage their data for DUET uploads are largely unknown and need to be quantified. The data management work required for the WY2012 data upload can provide each Data Collector the means to quantify the resources required for future data uploads, which should be considered as an integral part of the cost of providing the CBP the requested NTN WQ data and metadata.

**Table 1. Discrepancies in the NTN Monitoring Design between current and described practices, related to data collection, management, review, qualifying, and submission of NTN data which limit the completion of the development of the FAST process, DUET, and ESAR Standards.**

------------------------------------------------------------------------------------------------------------------------------- Discrepancies reflect differences and gaps between current (WY2012) NTN Data Collector activities and 2008 NTN Field Procedures (Document attached) or other NTN Project-level guidance, or simply a lack of sufficient NTN Project level guidance on practices related to the collection, management, review, qualifying, and reporting of NTN WQ and meta data. -------------------------------------------------------------------------------------------------------------------------------

***SAMPLING ISSUES (1, 2, and 3)***

Several discrepancies occur between the described sampling design and sampling conducted in WY2012.

Each is described below.

1. Monthly fixed time interval sampling is to include both routine and routine storm impacted samples, with the stream sample equipment used, and data collected, being dependent on the actual flow condition antecedent to and during sampling. When the stream sample equipment used does not match flow conditions, a representative sample of the quality of stream flow for that flow condition is not likely to be obtained; and selected WQ data are likely to be biased. For example use of sampling equipment designed mainly for a Routine Event under Storm Impacted conditions, implies using a sampling device that is lighter, and has a smaller diameter orifice, than that of the equipment designed for a Storm impacted event and flow conditions. Both of the above will lead to low biased samples for suspended sediment content, which can influence the concentrations of dissolved, particulate, and total N.P, S, and C, which all have suspended forms in stream water. In addition, when data are collected for a Routine event under Storm Impacted Conditions, selected WQP data are missing, which according to the NTN Guidance are to be collected for such an event. **Key issues are**
2. The degree to which the above is done routinely at selected WQ stations appears to be in conflict with 2008 NTN Guidance; occasionally the above possibly cannot be avoided given streamflow conditions at the time of sampling cannot be predicted with absolute certainty, and
3. Regardless of the frequency with which the sampling equipment and/or data collected may not meet that described for actual streamflow conditions, it is important to document such discrepancies when they occur, as part of the DUET data upload.
4. Monthly fixed interval sampling in combination with seasonal storm sampling provides the means ***to provide a relatively unbiased representation of the quality of stream water throughout the WY.*** This design works well in wet years but may not work as well in dry years to achieve the stated objective. During drought periods, storm flows are less likely. Routine Storm Impacted flows also are apt to be less likely, particularly during the monthly randomly selected fixed time intervals. However, that is not to say the latter won’t occur. As the WY progresses under these conditions, one or more quarters of the WY could pass with few Storms, and even fewer Routine Storm Impacted flows being sampled. Under such conditions, selected, but not all, Data Collectors have conducted additional interim monthly (not within fixed time interval) Routine or Routine Storm Impacted sampling, and increased their efforts to obtain additional storm samples to obtain a sufficient number of total samples per WY.

**Key issues are**

1. As for sampling during dry years, interim sampling is not discussed in 2008 NTN Guidance. If interim sampling is to be employed it should be described for all Data Collectors, and
2. We need to identify the WQ data and metadata obtained from such interim sampling events, as part of the DUET data uploads.

It would be useful for the NTWG to resolve how best to modify the NTN Guidance to address emerging and continuing drought conditions more uniformly among all Data Collectors. For example, if such conditions are already underway at the start of the WY, perhaps interim sampling for Storm Impacted/Storm events could begin immediately to ensure at least two distinct storm events are obtained that quarter, and interim sampling continue for each subsequent quarter as long as drought conditions persist. This approach could be preferable to waiting one or two quarters before introducing interim sampling. Or if the drought begins during the WY such interim sampling could be implemented before the end of any quarter where no Storm or Storm Impacted flows have been obtained in the first half of that quarter.

*DUET can address both issues (1) and (2) for all WY2012 and future WY data uploads if* ***each*** *Data Collector can provide their WY2012 WQ data as follows:*

* + - * 1. *Identify the Event \_Type for each sampling event.* Codes have been created (See attached Event\_Type Codes) that enable a Data Collector to characterize each sampling event at a station in the WY as to the actual streamflow conditions antecedent to and during sampling. For monthly fixed time intervals as either R or RSI; for up to 2 seasonal storms as S; and for any interim monthly Routine or interim monthly Routine Storm Impacted events as IR and IRSI, respectively; and for or any interim additional seasonal Storm sample as (IAS). (In the case of SRBC, their additional sampling at six selected PA stations can be coded as Baseflow (B) and additional storm (AS) samples.)

What is required: Each Data Collector reviews the USGS stream gage flow conditions associated with each sampling event, uses the 2008 NTN Guidance decision criteria to identify whether or not those flow conditions where storm impacted (**A storm-impacted event has a rising discharge (cfs) of at least twice that of the pre-storm, average daily discharge**), and then depending on whether or not the sampling event was part of the monthly fixed time interval sampling (or the seasonal storm sampling, 1st two Storm events) or reflects additional interim sampling, choose the appropriate Event\_Type Code.

* + - 1. *Provide an appropriate Event\_Remark code for selected sampling events if needed.* If the actual sample collection for any identified Event Types was conducted in a manner that is not consistent with that defined Event\_Type (in (a) above) , an appropriate Event\_Remark code would be used, as follows:
				1. ET1, Planned and conducted sampling for non storm impacted event (R, IR or B Event\_Type); actual flow conditions during sampling reflected a storm impacted event (RSI, IRSI, S, or AS Event\_Type); OR
				2. ET2, Planned and conducted sampling for storm impacted event (RSI, IRSI, S, or AS Event\_Type); actual flow conditions indicated flow conditions during sampling reflected a nonstorm impacted event (R, IR or B Event\_Type).
1. 2008 NTN guidance indicates sampling at each monitoring location is to be conducted in close proximity to the associated U.S. Geological Survey stream gage. “In close proximity” to the gage is not explicitly defined. How close or how far away such sampling is conducted can be influenced by the physical accessibility or actual streamflow conditions at the normal sampling location. Selected Data Collectors have an alternate sampling location, which is used when the normal sampling location is inaccessible; other Data Collectors have established low and high flow sampling locations which are used when sampling cannot be conducted at the normal location because of flow conditions. To address the above, and determine where sampling occurs, all Data Collectors have been asked to identify the location(s) at which normal, low, and high flow sampling occurs, and how that determination is made; and if applicable, to identify their alternate sampling location if the normal location is simply not accessible. Changes in sampling locations, depending on the distances this implies between where a sample was collected and where stream discharge occurs is of interest to selected end users of the NTN data.

 **Key Issue:**

 Once the mapping of stream locations is completed, knowing when a location other than the normal

 location was used is all that is required for DUET WY2012 and future WY data uploads.

*DUET can address issue (3) if* ***each*** *Data Collectors can provide the sampling location as part of the Event\_Remark code for their WY2012 and all future WY WQ data:*

See Event\_Remarks Table and, as necessary, use the appropriate code:

1. LF, Low flow sampling location used given conditions at normal sampling site;
2. HF, High flow sampling location used given conditions at normal sampling site; OR
3. AL, Alternate sampling location used given inaccessibility to normal sampling site.

 **Event\_Remark Codes are provided that address both Issues (1) and (3), if needed:** LFET1, HFET1, ALET1, LFET2, HFET2, and ALET2.

**Table 1. Continued, …**

**QUALITY OF NTN WQ DATA Issues (4, 5, and 6)—Consistency (or Logic) Checks, Precision, and Bias**

DUET can perform basic quality control checks for clerical (spelling) errors in codes used for a Data Provider, a Data Collector, their Station IDs, Sample Dates, Sample Times, Event\_Types, Sample Replicate Types, and for any related Event\_Type and WQ Parameter codes (Event\_Remark, Qualifier, Problem, etc.) by ensuring in each case they are defined by a DUET code. DUET also can check to ensure that there is consistency between submitted Event\_Types and Event\_Remark Codes, and submitted WQ Parameter concentrations and associated DUET Problem or Qualifier codes. In addition, DUET can check for gross clerical errors in WQP concentrations (outliers checked), as was historically done by the USGS. Beyond the above, the quality of the NTN WQ data can be determined by (a) Consistency (or Logic) Checks between related WQP concentrations, (b) quantification of the reproducibility in measurements (or precision in measurements) of a given WQP on the basis of duplicate samples analysis, and (c) the potential for bias in WQP concentrations because of contamination introduced during or after sample collection.

**Consistency Checks and Issues**

1. WQ data needs to be normalized among all Data Collectors for N, P, C, and S to measure and assess trends, yields, and loads among NTN stations. This normalization requires the calculation of selected N, P, C, and S WQ parameters. Historically, the USGS conducted Consistency checks on WQPs, including those WQPs used to calculate related WQPs, for example, such as Total Nitrogen or Total Phosphorus. DUQAT, the current (2012) data upload and quality assurance tool is being leveraged to create DUET, also will employ such checks antecedent to conducting selected WQP calculations. Maintaining this historical process is important. However, neither the 2008 NTN Guidance nor other NTN guidance addresses Consistency Checks.

**Key issues related to Consistency Checks are:**

1. Consistency checks are best performed by individual Data Collectors in conjunction with their respective analytical laboratories rather than at the NTN project level (by DUET).
2. Selected NTN Data Collectors have performed such checks in conjunction with their laboratories, whereas others have not. Those currently performing checks possibly use different approaches.
3. Data Providers indicated they would rather have DUET perform the post WQ Parameter calculations, which is historically consistent with the USGS G4 performing these calculations.
4. Therefore, if Data Collectors perform Consistency Checks, check results need to be provided in conjunction with the WQ Parameters checked in order for DUET to determine whether or not to calculate selected WQ Parameters.

 **To address all the above Consistency Issues, the CBP G4**

* + - 1. Identified the array of NTN WY2012 N, P, C, and S WQ Parameters covered by all Data Collectors, identified 28 distinct WQ parameter calculations that could be used to fill in that array to the extent possible for each Data Collector given the WQ parameters each indicated they would provide for WY2012, and had Mike Langland review both the above, which he considered adequate to meet USGS needs.
			2. Identified Consistency checks for each Data Collector on the basis of their provided data, and in so doing, considered the WQPs that could be calculated given the expected (X) WQP data from each Data Collector (See Xcell File, 1st Worksheet: WY Expected WQ and Meta Data, Column BQ, Contingency Checks for Data Collectors).
			3. Identified 3 Problem Codes to characterize the results of any Consistency Check that can be associated with the checked WQ Parameters, as follows (see Xcell File, 2nd Worksheet Consistency CK Outcomes) :

 QQ, 'PART EXCEEDS WHOLE VALUE YET DIFFERENCE IS WITHIN ANALYTICAL PRECISION, which indicates the logical expression for that consistency check is true given the precision in WQP concentration measurements; and which is assigned to each measured WQP concentration involved in that Consistency Check;

 NQ, 'PART EXCEEDS WHOLE VALUE DIFFERENCE IS NOT WITHIN ANALYTICAL PRECISION; which indicates the logical expression is false and that one or more of the measured WQP is in error. **Except for duplicate samples (FS1 and FS2 OR S1 and S2), any measured WQP concentration assigned this NQ code also would be reported under the DUET data upload as a null value.**

IQ, 'CANNOT DETERMINE IF PART EXCEEDS WHOLE VALUE AND WHETHER OR NOT DIFFERENCE IS WITHIN ANALYTICAL PRECISION; which implies the logical expression cannot be determined to be either true or false, which can arise when because one or more of the WQPs involved in the logical expression do not have measureable values (for example, a WQP has a null or < value).

***If Data Collectors can provide the results of their Consistency Checks using the above Problem Codes for WY2012 and any WY thereafter, DUET can calculate selected WQPs for each Data Collector on the basis of the expected (X) WQ Ps and associated Consistency check Problem codes as follows:***

For WQPs with QQ codes, DUET will calculate all WQPs indicated for a given Data Collector that require those WQPs with QQ codes; in each case document the equation used, and qualify any calculated value as necessary. For any calculated WQPs whose calculated value is a negative value, DUET will assign the WQP concentration the Problem Code NV (NEGATIVE CALCULATED VALUE IS VALID GIVEN PRECISION OF MEASURED WATER QUALITY PARAMETERS; ACTUAL CONCENTRATION LIKELY IS LOW, POSSIBLY LESS THAN PARAMETER QUANTIFICATION LEVELS OF SIMILAR BUT MEASURED VALUES FOR THIS WATER QUALITY PARAMETER);

For WQPs with NQ or IQ codes; DUET will not calculate the associated WQPs, but assign the calculated WQP concentration a null value with Problem Code CC (CANNOT CALCULATE GIVEN AVAILABLE DATA).

**NOTE:** However the NTWG decides to address Consistency Checks, the only outstanding questions with respect to the proposed DUET programming are (1) Will the Data Collectors perform the checks and provide the results using the Problem Codes described above, OR will DUET be asked to perform the check and provide the results as described. There are clear downsides to having DUET perform Consistency Checks. Selected Data Collectors already have performed such checks, and although they may need to modify their procedures and/or coding to obtain consistency in how checks are performed across the NTN Project, they have established a working relation with their labs in performing these checks. Also, if DUET performs such checks, there will be no opportunity to rerun samples, and all WQ Parameter concentrations assigned NQ codes will be assigned null concentration values. This will lead to inconsistencies between Data Collectors data and DUET data. Finally, for DUET to perform these checks, the CBP NTN staff would need to work with every Data Collector and all their laboratories to set up the checks. This is far more complicated than each Data Collector working with their individual labs.

*Regardless of who performs Consistency Checks and provides check results for the WY2012 data, there will be no opportunity to rerun any samples for WQPs assigned NQ Problem Codes. However, Data Collectors will learn how to perform the checks and code the results. In addition, they can work with their individual laboratories to set up Consistency Checks for WY2013 data and have the opportunity to arrange reruns on WQPs whose initial Consistency Check results lead to NQ codes.*

*Note: For WY2012 upload, DUET already has been developed to obtain Consistency check results for identified WQ Parameters from each Data Collector, and it already has been developed to calculate all the related WQPs on the basis of those check results and the measured WQP data that each Data Collector indicated they would provide.*

**Reproducibility of Measurement (Precision) Issues**

1. 2008 NTN Guidance indicates quality control sampling is to include the collection of Duplicate Samples. Duplicate Samples are to consist of field split samples from a single churn, obtained either monthly or every 20 WQ samples.
	1. The NTN 2008 Field Guidance does not describe the purpose of Duplicate Sample collection (nor does any other NTN Project level document). The purpose of these duplicate samples presumably is twofold—to determine the reproducibility of a WQP measurement, for the expected (X) WQPs and to provide water quality data for that sampling event. This dual purpose should be in the NTN Guidance.
	2. The NTN 2008 Guidance indicates duplicate field samples are to be split samples from a single churn. (These are FS1 and FS2 Sample Replicate Types in DUET nomenclature). Selected Data Collectors collect concurrent duplicate field samples—two stream samples each deposited separate churns. (These are S1 and S2 Sample\_Replicate\_Types in DUET Nomenclature). The latter needs to be included in the NTN guidance with an explanation as to when and why it is appropriate to collect concurrent duplicate samples. Namely, when large sample volumes result from EWI sampling that exceed the capacity of a single churn. This routinely can occur when sampling at a large river (for example, RIM Station) as a result of simply river size (width and depth), and or can occur at high discharge on even a small river (for example, high storm flows).
	3. The NTN 2008 Field Guidance makes no mention of whether duplicate sampling is to be representative of the range in flow conditions (Event\_Types) or just one flow condition. Given (i) the purpose of duplicate samples (see (a) above), and (ii) that NTN WQ data reflect dissolved, particulate, and total concentrations of N, P, C, and S, and that particulate and total concentrations tend to increase with increasing streamflow, duplicate samples representative of the range in flow conditions throughout the WY is the only way to provide reasonably accurate precision measurements for particulate and total concentrations. Data Collectors likely differ in the degree to which their duplicate sampling is representative of the range in flow conditions. The degree to which duplicate sampling is to reflect the range in flow conditions needs to be resolved among all Data Collectors, and reflected in the NTN Guidance.
	4. The specified frequency of duplicate sampling in the NTN 2008 Guidance for Data Collectors with few WQ sampling stations provides too few duplicate sample data to be of use in determining the reproducibility of any WQP measurement. The NTN guidance needs to be modified to provide a reasonable amount of duplicate sample data per WY regardless of the number of stations for which data are provided.
	5. NTN 2008 Field Guidance does not specify what WQ Parameters are to be included in duplicate samples, but most (not all) Data Collectors indicated they provide the same data for all S1 and S2, or FS1and FS2 samples. If this is correct, that implies duplicate sample data for TSS, SSC\_Total, SSC\_%Fine, as well as duplicate samples for all “Other” WQ parameters are provided, including TALK, ALKF, CLW, CLF, BOD5Day, CHL-*a*, etc. This extent of coverage needs to be verified. Guidance should be modified to reflect outcome of this discussion.
	6. Although duplicate sample data must serve two purposes, the form of the data needed for each purpose differs. This difference in data required is not addressed by the 2008 NTN Guidance or any other NTN Guidance. It should be, in relation to requirements described below:
		1. To estimate the precision in WQP measurements without undue and artificial bias the NTN data for each WQP provided for FS1 and FS2 (or S1 and S2) samples should be provided in a raw form (to at least two insignificant figures). If this is not done, and data are reported to only one insignificant figure, rounding errors artificially and notably increase the imprecision in measurements for WQP concentrations near the Laboratory Method Reporting Level (RL, LRL, or PQL), which is typically used to censor the data.
		2. WQP data that historically have been provided by a Data Collector under routine data collection that includes WQP concentration above the Laboratory Method Detection Level, but below the Laboratory Reporting Level, have been used by the USGS for interpretive purpose. If this is to continue, duplicate sample data of this nature also should be provided to qualify the precision of such WQ data, and as noted above (i) data should be provided in a raw form.
		3. Duplicate sample data also must be used as the WQP data for the sampling Event\_Type for which duplicate samples were collected. Thus, Consistency Checks must be performed on the same selected WQPs for each duplicate sample, and as described earlier (see QUALITY OF NTN WQ DATA—Consistency (or Logic) Checks, above). However, in performing those checks and qualifying results, any measured WQP assigned a Problem Code NQ cannot be assigned a null value, but must keep its original concentration value when providing these data to DUET; otherwise there are no duplicate WQP concentration data to calculate the reproducibility of measuring that WQP.

**To address all the above Precision Issues, and in relation to DUET programming for WY2012 data upload, the CBP G4 assumed the following (which Data Collectors needs to discuss to determine whether or not they can provide the WY2012 WQ or metadata required where indicated):**

* + 1. Assumed the two-fold purpose for the collection of Duplicate Samples; and in relation to these two purposes (see Xcell File 1st Worksheet: Expected WQ and Meta Data, Column BS: WY2012 Expected DUPLICATE Samples, WQPS, …)
		2. ***Assumed Data Collectors could code and provide whatever duplicate field split samples from a single churn (Sample Replicate Type coded as FS1 and FS2 samples), or concurrent duplicate samples from two churns (Sample Replicate Type coded as S1 and S2 samples) that they collected.[[3]](#footnote-3)***
		3. ***Assumed Data Collectors could submit whatever field split or concurrent duplicate sample data they actually obtained with the appropriate coding associated with any WQ sample, including designating the Data Provider, Data Collector, Sample Date, Sample Time, Event\_Type,( if needed, with Event\_Remarks,), Sample Replicate Type, etc. Initially assumed Data Collectors would provide duplicate samples reflecting R, RSI, or S (or B and SA) Event\_Types; we can adjust programming if needs be to include any interim sample Event\_Types if properly coded.***
		4. ***Assumed Data Collectors could provide duplicate samples that would include all their Expected (X) WQPs or some subset of those WQPs, such as all expected N, P, and C WQPs from each data Collector, but only selected S (TSS) and selected Other WQPs (TALK, ALK, CLW, CLF, and CHL-a); we can change DUET programming to reflect the NTN Data Collectors decision in relation to scope of WQPs included but need that decision ASAP;***
		5. ***Assumed Data Collectors could provide DUET the duplicate sample WQP data in raw form, including that WQP data between the MDL and RL, with Consistency check results for those measured WQPs, which require checks (Table 2), but with no censuring (null values) assigned to those WQPs because of check results that lead to a Problem Code NQ.***

If Data Collectors can perform the above, DUET can calculate the precision in measurements for each WQP, and archive these precision measurements and the raw WQP data in a provisional manner[[4]](#footnote-4). DUET would then round all raw WQP data to the generally reported form for these data—on the basis of the MDL and RL formats provided in the Data Collectors Laboratory Table (see Xcell File 3rd Worksheet (Template Lab\_Tab, MDDNR) and the routinely reported S1 WQP data, null the values for those WQP concentrations with NQ codes, and archive all revised WQP data as the “WQP” data for each duplicate sample for that sampling event in the CBP NTN WQ\_1 database.

**For WY2013 and beyond, the NTWG should discuss all Precision Issues (6 a through 6 f).** In relation to these issues, and in particular issue 6 c), the CBP G4 proposed a randomized procedure for NTWG and Data Collectors to obtain representative duplicate (FS1 and FS2, or S1 and S2) samples for the range in WQ and flow conditions sampled throughout the WY (see Xcell File 1st Worksheet: Expected WQ and Meta Data, Column BS: WY2012 Expected DUPLICATE Samples, WQPS, …) for WY2013 data collection.

**Contamination BIAS, Issue 6**

1. 2008 NTN Field Procedures indicate quality control sampling is to include the collection of Field Blanks (FBs). FBs are to be obtained monthly, in the field, and, the source(s) of FB contamination investigated.

 **Key BIAS issues are**

1. The 2008 NTN Guidance does not explicitly define the purpose of FBs, which presumably are used to assess whether any WQP data potentially are positively biased due to the inadvertent contamination of the samples used to obtain WQP data during or any time after these samples are collected. The NTN guidance should clearly state the purpose of FBs.
2. The NTN Guidance for Data Collectors provides no guidance on the degree to which FB sampling should be representative of WQ sampling. Also the indicated monthly frequency of FB sampling implies Data Collectors with few WQ sampling stations provide an excessive amount of FB data relative to WQ data for a given WY.
3. What constitutes a true FB was not clearly defined by 200 NTN Guidance. WY2012 collection of blanks by selected Data Collectors are true FBs. Namely, at the actual sampling site, certified blank water was poured into the stream sampling equipment that then will be used to collect the stream sample at that site and given stream flow conditions at sight. This water was then poured from the sampling equipment into the churn just as the stream sample would be. The above process was repeated the same number of times as the sampling equipment will be filled to provide the stream sample at that site. (Generally this number corresponds to the number of EWIs determined for the streamflow conditions.) Samples of blank water were then taken from the churn and processed in a manner identical to that which will be used for the corresponding stream sample to be taken from the site. Other Data Collectors have taken blanks at the sampling site, but begin by pouring certified blank water into the churn, and bypassing the sampling equipment; whereas other Data Collectors have taken blanks in the Field Office rather than at the sampling site; or have taken blanks on selected equipment, and so forth. The NTN Workgroup needs to modify the NTN guidance. True FBs are the preferred and first blanks routinely collected by many other long-term monitoring programs. Source water blanks are inherently the second blank generally taken to ensure the blank water used in FBs is not contaminated.
4. The 2008 NTN Guidance indicates that contaminated FBs need to be investigated as to the source of the contamination, presumably to determine if the source of the contamination reflects a similar source of contamination of associated WQP samples. The NTN Guidance needs to be modified to describe just how the above would be accomplished and communicated. Selected Data Collectors only consider FBs contaminated at levels above the WQP Reporting Level (RL) as requiring investigation into the sources of contamination. Contamination investigations from other monitoring programs have shown that what appear to be spurious occurrences of FB contamination above the RL, which often are not fully investigated, actually can reflect persistent contamination of FBs and WQ samples mainly at levels just below the RL, which occasionally rise to levels above the RL. Such contamination implies low level WQP data obtained during the same period as the FBs are in fact positively biased, as are any WQP concentration values reported above the MDL but below the RL.
5. If source(s) of FB contamination also likely result in contamination of associated WQP samples, the biased WQP data need to be identified. Identification of which WQP data are biased needs to consider the level of contamination. The 2008 NTN Guidance does not address this issue.

For WY2012 Data Upload, if Data Collectors can:

* + - * 1. Provide whatever true or nearly true FB samples they collected, properly codes as to Data provider, Data Collector, Event\_Type, Event\_Remark (if needed) Sample Replicate Type (FS1 and FS2 or S1 and S2), Sample-Date, Sample\_Time, ….,;
				2. For each FB sample, provide their expected (X) WQP data, or some subset of the expected (X) WQP data—notifying CBP G4 ASAP if subset provided, what WQPs will be provided);
				3. For each FB WQP, provide the data, including measureable values above the MDL but below the RL, (non-null values and no < qualifier); and
				4. Assuming Data Collectors have investigated sources of contamination in any FB; assign one of the following Problem Codes to any measureable WQP associated with such a FB:

UB [UNDETERMINED SOURCE OF BIAS UNDER INVESTIGATION, MEASUREABLE CONCENTRATION IN FIELD BLANK); BB [SPURIOS OR PERSISTENT CONTAMINATION (BIAS), SOURCE(S) APPEAR TO REFLECT BLANK BIAS ONLY, AND RELATED TO MANNER, EQUIPMENT, OR SUPPLIES USED TO OBTAIN OR ANALYZE BLANKS); OR CB (SPURIOUS OR PERSISTENT CONTAMINATION (BIAS), SOURCE(S) APPEAR TO REFLECT MANNER, EQUIPMENT, OR SUPPLIES USED TO OBTAIN OR ANALYZE BLANKS AND ASSOCIATED WATER QUALITY SAMPLES);

DUET will summarize the FB data for each and all Data Collectors. On the basis of these summaries, the NTN Project QA Coordinator will collaborate with each and all Data Collectors to determine whether any WQP data appear to have been biased on the basis of FBs coded as CB.

Where such BIAS appears to have occurred, selected WQP values will be assigned the Problem code

BM (BIASED MEASUREMENT, CONCETRATIONCOULD REFLECT CONTAMINATION BIAS FOUND IN ASSOCIATED FIELD BLANK CONCENTRATIONS WHICH WERE WITHIN SAME ORDER OF MAGNITUDE AND EQUAL AT LEAST 10 PERCENT OF THE CONCENTRATION OF THIS WATER QUALITY PARAMETER).

Note: Assuming FBs are characterized as CB, the proposed decision criteria is that WQ samples collected in conjunction with such FBs also are likely contaminated. If the level of contamination in the FBs is equal to at least 10 percent of the measured WQP concentration, that WQP value is considered to be biased.

**As for WY2013,**

**The NTN Workgroup needs to address Bias issues raised above (6 a) through (6 e). In relation to selected issues, and to aid in the discussion and resolution of these issues, the CBP G4 proposes considering the following:**

1. The 2008 NTN Guidance does not explicitly define the purpose of FBs. Presumably are used to assess whether any WQP data potentially are positively biased due to the inadvertent contamination of the samples used to obtain WQP data during or any time after these samples are collected. The NTN guidance should clearly state the purpose of FBs.
2. Define what constitutes a FB. For example, a blank collected at the sampling site just prior to sampling where certified blank water was poured into the stream sampling equipment that then will be used to collect the stream sample at that site and given stream flow conditions at sight. This water is then poured from the sampling equipment into the churn just as the stream sample would be. The above process is repeated the same number of times as the sampling equipment will be filled to provide the stream sample at that site. (Generally this number corresponds to the number of EWIs determined for the streamflow conditions.) Samples of blank water are then taken from the churn and processed in a manner identical to that to be used for the corresponding stream sample to be taken from the site. True FBs are the preferred and routinely collected by many other long-term monitoring programs. Source water blanks also are inherently the second blank generally taken to ensure the blank water used in FBs is not contaminated. Other blanks could be produced as need to investigate the source of any FB contamination.
3. The specified monthly frequency of FB sampling in 2008 NTN Guidance for Data Collectors with few WQ sampling stations appears excessive if there is no evidence that FBs or when contaminated the source of that contamination reflects the blank water. Suggest NTN Guidance be modified to reduce FB sampling for such Data Collectors, with FB sampling increasing if FB contamination occurs and the source is not the blank water.
4. The 2008 NTN Guidance does not address whether FB sampling should be representative of all environmental conditions (all WQ sampling stations and all Sampling Event Types) or selected environmental conditions and Sampling Event\_Types. However, NTN WQ data are collected under a wide range of environmental and streamflow conditions. The norm for most long-term monitoring programs is that FBs should reflect those conditions under which WQ data are collected. Suggest NTN Guidance be clarified to indicate FB sampling design represent all WQ sampling Stations and all Sampling Event\_Types. To address c) and d) above, the CBP G4 has proposed a sampling design for FB collection tailored to each Data Collector for WY2013 (see Xcell File 1st worksheet: WY WQ and Meta Data, Column BV: WY2013 Field Blanks).
5. The 2008 NTN Guidance indicates that contaminated FBs need to be investigated as to the source of the contamination, presumably to determine if the source of the contamination reflects a similar source of contamination of associated WQP samples. The NTN Guidance needs to be modified to describe just how the above would be accomplished and communicated. Selected Data Collectors only consider FBs contaminated at levels above the WQP Reporting Level (RL) as requiring investigation into the sources of contamination. Contamination investigations from other monitoring programs have shown that what appear to be spurious occurrences of FB contamination above the RL, which often are not fully investigated, actually can reflect persistent contamination of FBs and WQ samples mainly at levels just below the RL, which occasionally rise to levels above the RL. Such contamination implies low level WQP data obtained during the same period as the FBs are in fact positively biased, as are any WQP concentration values reported above the MDL but below the RL.
6. Noting that if the source(s) of FB contamination also likely results in contamination of associated WQP samples, the biased WQP data need to be identified. Identification of which WQP data are biased needs to consider the level of contamination. The 2008 NTN Guidance does not address this issue. As a starting point, the CBP G4 has proposed that when it is apparent the WQP data likely are affected, the affected data are those where the concentration of the contaminated FBs equals at least 10 percent of the concentration value of the WQP.

**HISTORICAL ISSUES: NTN Data Before WY2012**

1. Implementation of the FAST\_DUET\_ESAR process begins with the establishment of that process for WY2012 NTN Project data, and future WY data.

**Key Issue**: The USGS has requested information on whether above could be used on historical (pre-WY2012) data. This issue can be discussed during the Action Team Meeting and subsequent NTWG meeting. In theory, the answer is yes, it is possible to go back and capture previous WY data. Pragmatically, it would require meeting the same type of WQ and metadata requirements currently being requested for WY2012 and beyond. For each WY we need the information described and requested in the Xcell File – WY Expected WQ and Meta Data, Lab\_Tab, and POC information (4th Worksheet). Whether or not each NTN Data Collector is able to provide the required historical information ultimately determines the extent to which we can go back in time.

**OUT OF NETWORK DATA ISSUES: Data currently are available from other monitoring programs that possibly could be used in meeting the NTWQM Program interpretive objectives (see INTRODUCTION)--most notably from automated sampling and to a lesser extent from continuous monitoring**

1. Automated and continuous data collection and sampling has and continues to be conducted at at least 4 non-NTN stations within the CB Nontidal watershed. Data are very similar in type to the data routinely being collected by the NTN**.**

**Key Issue:** Guidance is lacking in the form of protocols for monitoring and data management. However, as the need for increased numbers of monitoring Stations available within the CB Nontidal watershed increases, it will become increasingly difficult to obtain the human resources to be everywhere that monitoring is needed every time it is needed. Obtaining NTN data from approximately 130 NTN monitoring sites already an acute challenge when it comes to data collection for storm impacted and storm flows.

1. Automated programmable sampling has and can provide monthly randomly selected time interval Routine and Routine Storm Impacted samples, Storm Samples, and interim samples of all types. Currently (as of October, 2012), finalized guidance is lacking in the form of protocols for monitoring and data management within the USGS or MDDNR who are conducting such sampling. However, as the need for increased numbers of monitoring stations available within the CB Nontidal watershed increases, it will become increasingly difficult to obtain the human resources to be everywhere that monitoring is needed every time it is needed. It already is an acute challenge to collect sufficient data from many stations for storm and storm impacted events.

There are legitimate concerns regarding automated sampling. One major concern is how representative are the WQ data generated from automated samples in relation to WQ data generated from composite EWI samples. The latter are generated from a churn containing a composite sample composed of individual EWI vertical samples, where the latter are collected with sampling equipment designed for the stream discharge at the time of sample collection. The former are obtained from a composite sample obtained from a selected series of sequential or intermittent smaller sample volumes obtained over a portion of the storm hydrograph, with each sample being obtained from a relatively small orifice located at or near the streambed at a fixed elevation from the streambed. Determining what the relation is between WQ data obtained from each procedure is essential to answer. However that requires concurrent sampling using both procedures for a variety of flow conditions. This leads to a third concern—maintaining the frequency of concurrent sampling possibly needed to ensure that any initial relation established between a WQP using concurrent sample data possibly over a finite or full range of flow conditions does not change through time. For the above reasons, another way to possibly envision the use of automated sampling is to consider it as an intermittent aid for stations which can’t always be monitored using established NTN sampling procedures. The NTWG could consider all of the above as a quick overview to help guide its discussion.

* + - * 1. Continuous sampling is being conducted for selected WQPs, including Nitrate-nitrite nitrogen, at several stream locations in the CBP Watershed. Although the WQPs monitored are limited in terms of their coverage of WQ conditions, procedures to verify the quality of the data have been established. As in the case of automated sampling, the issues are again related to how well each monitored parameter, often measured at a single point or on water drawn from a single point represents the streamflow conditions for that parameter. For some WQPs, nitrate-nitrate nitrogen this may be less of an issue. For other WQPs such as turbidity and how that measurement serves as a surrogate measurement for suspended sediment concentration in the stream, this is an issue.
				2. Other out of network monitoring: CBP ConMon (selected nontidal sites), River Keeper Monitoring, …
1. Note both the MRAT and FLC responses were prepared wholly or in part by Nontidal Workgroup (NTWG) members with input from the entire NTWG. However, the bulk of the resources obtained by the NTN Project and NTWG after 2009 were used to expand the network. Managing the network and NTN data for continuity and improvements as described in the MRAT and FLC responses has largely been unaddressed for the last four years, and clearly an area for considerable and valuable improvements in the NTN Project. [↑](#footnote-ref-1)
2. On basis of most recent NTWG meeting, also could add yields. [↑](#footnote-ref-2)
3. Sequential duplicate samples are not considered as they are likely to reflect short-term variations in stream WQ as well as variations in the reproducibility in measuring a WQP due data-collection methods. [↑](#footnote-ref-3)
4. Provisional data would be available to any Data Collector via request to NTN Project Data Manager. [↑](#footnote-ref-4)