

1 July 2020

To: Forest Practices Board sub-committee on Water Typing
From: Chris Mendoza, Conservation Caucus
Subject: Methods used for QA/QC screening of CMER Eastern Washington fish distribution data from 2001, 2002, 2005.

Forest Practices Board sub-committee members,

In response to your request to the water typing technical group at your May 5, 2020 meeting to continue screening the CMER eastern Washington fish distribution data, below are details on the methods used for screening the data for QA/QC.

First, all of the archived fish distribution data files from the CMER 2001 report (Terrapin), 2002 report (ABR, interannual survey comparing 2001 Terrapin data), and 2005 report (ABR, Seasonal fish distribution) were emailed to the technical group from DNR staff (Marc Ratcliff). All of the data files were "read only" and therefore, converted to MS Excel (MS Office 360, 2020) prior to screening to better facilitate data screening and analyses.

The technical group used four main criteria for the initial screening based on input from DNR staff, and the Board sub-committee at their May 5, 2020 meeting. The criteria used to screen out sites include those containing: 1) culvert blockages in close proximity to the Last Fish / Last Habitat point (Type F/N water type break), 2) Large Wood Debris (LWD, a deformable temporary barrier) associated with the Last Fish / Last Habitat point, 3) Last Fish / Last Habitat stream surveys conducted after the July 15 window for water typing, and 4) surveys that contained stream reaches that were dry during the time of the survey for at least 100 meters in close proximity to Last Fish / Last Habitat point.

After the initial screen was finished using the four criteria above, an additional more detailed screen was applied by technical group member Chris Mendoza (Conservation Caucus) in an attempt to salvage some of the sites that had been initially screened out of the pool. While this was not a request from the Board sub-committee, Mr. Mendoza determined that some stream surveys had recorded enough detailed channel profile information to prevent them from being screened out of the pool (e.g. surveys conducted after July 15, but had a permanent natural barrier / waterfall above which no fish were detected during the survey).

The initial QA/QC screen using the 4 criteria left approximately 26 useable Type F/ N break points out of 314. The additional more detailed screen used to recover potential sites initially screened out resulted in 150 useable Type F/N break points out of 314, the majority of which were recovered from surveys conducted after July 15, but contain permanent natural barriers above which no fish were detected, or contain channel gradients 20% or greater for at least 100 meters directly upstream of the Last Fish / Last Habitat point.

The following is a detailed description of the QA/QC process sent by Mr. Mendoza to technical group members soliciting their review and feedback via email. For the purpose of reviewing and interpreting the color-coded Excel spreadsheets, read the Email below dated May 12, 2020 as the color codes apply to all survey data years (2001, 2002, 2005).

From: Christopher Mendoza <cmendoza2@comcast.net>

Date: Tuesday, May 12, 2020 at 3:29 PM

To: Jim Matthews <matj@yakamafish-nsn.gov>, "Heimburg, John R (DFW)" <John.Heimburg@dfw.wa.gov>, Brian Fransen <brianfransen@outlook.com>, "RATCLIFF, MARC (DNR)" <MARC.RATCLIFF@dnr.wa.gov>

Cc: "ENGEL, MARC (DNR)" <MARC.ENGEL@dnr.wa.gov>

Subject: QA/QC screening of CMER 2001 Type F/N break data from Terrapin

Attached is the 2001 Terrapin (Eddie C.) data screened using the four criteria the Board sub-committee on water typing asked us to finish up at their last meeting. Below is a summary of my work thus far.

As per previous versions, the rows have been colored using the following screening criteria: Culverts (code #8, Barrier type column W) = **Tan not yellow**, LWD (code #7, Barrier type column W) = **Light Green**, Survey conducted past July 15 (no code but date of survey in column AD) = **red**, surveys with loss of surface flow for at least 100 meters of dry channel (code #4, Last habitat feature column U) = **light blue**. **That was my initial screen as a first step** using Eddie's "Data code book" tab (second tab in spreadsheet). Some of these definitions are duplicated with letter codes also (F,G,I, J), again see the "Data code book" tab, but are not consistently labeled with both a number and a letter code (Column AG - Combined last habitat feature; Column AH – Combined last fish feature) .

The second screening I did was more detailed and resulted in changing colored rows back to white background for inclusion in the dataset based on more specific channel condition information supplied by the surveyor(s). **Example:** If a survey was conducted after July 15 (**red** row), but a permanent natural barrier was present (code # 5 or 6, column W) above which no fish were observed that survey is now included. Also, if the average channel gradient was 20% or greater for at least 100 meters directly upstream of the type F/N break (column AE) that row series was included. There's other useful information in the column categories describing channel conditions near the end of the survey that likely allows for inclusion (e.g. BFW < 2ft. wide, no defined channel, poorly defined channel, etc.) indicating the LF/LH location may have been near FP rule defined default physical criteria for Eastern WA. By contrast, if a survey included at least 100 meters of dry channel near the Type F/N break but didn't have any of the above channel attributes than it remained highlighted **light blue**.

The "dry channel for 100 meters" and "outside survey window July 15" were by far the most time consuming screening criteria portion of this exercise as channels can often go dry / intermittent near their headwaters that, regardless of flow, would not be considered fish habitat based on substantial reductions in channel size and abrupt changes in channel gradient

equal to or greater than the current rule. And, as stated above, surveys after July 15 may contain similar channel attributes that would not prevent them from being included.

The other 2 screening criteria are more straight forward mainly because; culverts are easily identified, but we could also look at whether or not they were in close proximity to the end of the survey (some are over 1000 ft downstream and fish were found above them). LWD got screened out if listed as a last habitat LH or last fish LF feature as a barrier type with no further screening as per prior discussion. Note: there are a few coded permanent natural barriers (# 5, 6) listed in column W that state they are actually LWD or “wood” steps in the “notes” (column AC). The “notes” are very helpful and contain details about the survey’s progress so please make use of them.

Lastly, I highlighted all the rows the same color that have the same Stream ID (column B) for one of the 4 criteria. **Some rows meet more than one criteria** (e.g. surveyed past July 16, with culvert) **but only contain one color**. Every survey has a story to tell about how the channel conditions changed as the survey progressed upstream (see Stream Physical Characteristics, Columns G through O) as you can follow changes in channel width, flow and gradient. Column E gives “distance to reference point” allowing one to see the distance between LF and LH when those points are not coincident. I recommend looking at all these columns when piecing the survey together before providing feedback.

From: Christopher Mendoza <cmendoza2@comcast.net>

Date: Friday, June 12, 2020 at 9:54 AM

Subject: FW: QA/QC screening of CMER 2001 Type F/N break data from Terrapin

Attached is the final screen for the Terrapin 2001 raw data sheets based on input from all of you. The last screen are of rows that contain “0” in the BFW column for stream characteristics, and code “15” as “no defined channel” and have been **Colored purple**. The other color codes are listed below that you’re already familiar with. Keep in mind the rationale for doing this is that many of these rows also have “0” listed as a distance from reference point and “0” for BFW, and code “15” = “no defined channel.” So they don’t meet the DNR definition of a typed water. If they did, they surveyors did not list BFW. Also keep in mind that many of them were not used in the final Terrapin data set and don’t have measurements for the other channel attributes.

John, I left the channel gradient at 20% as a threshold and did not have time to find the ones that you’re concerned about not meeting the rule at 21% since you didn’t highlight them. Again the channels were also E-fished using protocol surveys so default physicals are not required, and not knowing if the basin area is < 175 acres < means that the threshold could be as low as 16% channel gradient for surveys done past July 15, so I think 20% is still a fairly good precautionary measure. If you could go through the spreadsheet and list the row number, Watershed, and Stream ID with the 20% channel gradients you are concerned about that would be helpful. We could go over them at the meeting next week.

I will send the other two spreadsheet later today 1) Terrapin 2001 LF/ LH compared with ABR 2001 LF/LH with listed Type F/N break furthest upstream, and 2) 2005 ABR seasonal variability sample points spreadsheet compared to 1) with the highest point for all three. Both of these will also be screened using the 4 criteria listed below.

From: Christopher Mendoza <cmendoza2@comcast.net>

Date: Friday, June 12, 2020 at 2:49 PM

Subject: QA/QC screening of ABR 2005 seasonal variability data w/ highest fish.

Attached is the ABR seasonal variability spreadsheet that 1) uses the upper most fish point for those 100 samples and 2) has been color coded with screening criteria. Note the highest fish related columns (I, J, K) contain the survey year, distance, and notes. These columns were added by Jim M. and are labels with 2020 do distinguish them as newly added information.

I've also attached the spreadsheet containing the surveys for all of ABR 2005 seasonal samples. Note that they are repeated 3 times over the year so their grouped by season (rows 1-100 spring, 101-200 summer, 201-300 fall). As per above, some of Jim's notes in column "K" of that file note that some sites were not located in the 50 lats, 50 term Excel tabs. This forced me to have to screen this spreadsheet that contains all the data for all 3 years. It took some extra time, but I also color coded the rows. The main color in the spring survey is LWD . I also highlighted the same Stream ID in the summer and fall survey rows. I wanted to see if the LWD was consistently recorded in the same location. Most was, but not all, so some LWD was transported downstream. One site had LWD in the spring, none in the summer survey, then LWD again in the fall.

Lastly I included the ABR data sheet so you can cross reference the codes. Important: remember that **Terrapin and ABR used different codes for the same instream channel feature** so if you have both of the spreadsheets open at the same time, don't get confused. The Terrapin 2001 spreadsheet I sent out this morning has the codes contained in a tab / sheet within the same spreadsheet. The ABR does not, and I'm attaching it separately – open the far right tab labeled "codes".

Please review, along with the Terrapin 2001 Excel spreadsheet I sent out this morning, in preparation for our meeting next Wednesday. I have one more last spreadsheet I'm working on finishing up today, but you may get it Monday AM – The Terrapin 2001 data compared to the ABR 2002 data (interannual survey) with a new highest last fish column (Recall that this one will have ~ 300 sites, where the ABR 2005 is a sub-sample of that with only 100 sites). They will also be color screened.

Christopher Mendoza <cmendoza2@comcast.net>

Monday, June 15, 2020 at 9:08 AM

Attached is the last spreadsheet comparing the 2001 Terrapin data with the ABR 2002 data (interannual survey). Column "J" has the survey year with the highest Last Fish (LF), and Column

“I” has the distance between the two. The 314 rows have been screened using the same color codes as the Terrapin 2001 raw data screen I emailed last week (~2000 rows), and the Terrapin 2005 seasonal fish data (sub-sample of 50 laterals, 50 terminals) I emailed last Friday. See the initial email at the bottom of this chain for color code criteria definitions.

Additional Notes for interpreting that attached spreadsheet:

1. Some of the sites in ABR 2002 surveys were conducted after the July 15 sampling window, but the corresponding site surveyed by Terrapin in 2001 was conducted before the July 15 window. In those cases I chose the site inside the survey window (before July 15) if it met all other screening criteria.
2. Some of the ABR 2002 surveys did not record natural barriers (Waterfalls or steep cascades) for the same survey sites that Terrapin 2001 surveys recorded, so again I defaulted to the Terrapin 2001 site.
3. In cases where LWD was recorded by ABR 2002, but the “notes” for the corresponding site in Terrapin 2001 recorded the LWD in combination with a natural barrier (e.g. cascade), I defaulted to the Terrapin site. If both survey years consistently recorded LWD at the same site, that site was screened out.
4. Recall that a survey may contain more than one criterion, e.g., LWD and surveyed outside the window.

This process resulted in screening out approximately half of 314 surveys.

Please review in preparation for Wednesday's meeting where I will be glad to answer any questions. If you see a specific survey(s) you have questions about please write down the 1) Survey year/ file name, 2) row number, 3) Watershed name, 4) Stream ID.

From: Christopher Mendoza <cmendoza2@comcast.net>

To: Jim Matthews <matj@yakamafish-nsn.gov>; RATCLIFF, MARC (DNR)

<MARC.RATCLIFF@dnr.wa.gov>; Heimburg, John R (DFW) <John.Heimburg@dfw.wa.gov>;

'Brian Fransen' <brianfransen@outlook.com>

Date: Friday, June 19, 2020 at 9:21 AM

Jim,

Thanks once again for taking to time to review my QA/QC work. I appreciate your feedback. I incorporated all of your color coded recommendations into the 2001 Terrapin raw data / spreadsheet, save LeClerc Stream ID 290 (row 1254) as your note was taken from row 1255, not 1254, so you may have gotten those crossed up (happens to the best of us!). I did not make any changes to the ABR 2002 interannual survey data / spreadsheet, and provided an explanation as to why in column F (attached). And thanks for providing the Excel row # AND the Stream ID # so I could find the surveys more efficiently.

Marc, attached is an updated Terrapin raw data spreadsheet based on Jim's comments with today's date. Please use this as the latest version.