Update Log - Oklahoma Fishery Analysis Validation Application

1. 3/29/2019 – Dan
   1. Problem: Ashley found some older gears used different numbers for effort and she felt it would be better to allow any value of effort >0 as acceptable rather than > 24 | Effort < 0.1
   2. Solution: changed this in the code (but commented old code out so it could be reinstituted if we want to in the future after all old data have been validated and added).
2. 12/30/2020 – Dan
   1. Problem: App was not catching “.” in station field as missing station despite having sampleData$nullStation[sampleData$Station=="."] <- TRUE code (and I also tried adding sampleData$Station[sampleData$Station=="."] <- NA prior to line that creates the nullStation field (checks for NA’s) and this did not work.
   2. Solution: added sampleData$Station[sampleData$Station=="."] <- NA to code where app reads in the sampleData file. This still allows for checking for blanks instead of “.” in fields that are allowed to be blank (I think ODWC wanted to require this to make sure the blank is intentional), but prevents a file from being validated without a station.
   3. Other changes made: Reformatted UI for validation app so file upload of both SSP and Age tabsto make the template download less prominent and instead put the upload button front and center.
3. 3/2/2020 – Dan
   1. Problem: update of shiny app software broke renderDataTable function. Research into this indicates this function is being depreciated and the same function from the DT package is to be used instead
   2. Solution: added DT package to the app and specifically call DT::renderDataTable to make tables
4. 3/2/2020 – Dan
   1. Problem: ODWC wanted app to work such that people need to add something to blank cells to ensure they are not accidentally left blank
   2. Solution: Added explicit check for blank cells. Unfortunately, R will automatically convert blank cells to NA in cases of numeric or logical data types…making it impossible to tell an NA from a blank. As such, we had to institute a rule that all blank cells should be marked with a period. I recoded app so that blanks and NA’s are converted to NA when the csv file is read in, but “.” Is left. I then check for NA’s in all fields and add a row to the error report indicating if any NA’s were found (tells user to use periods instead). Then as the file is downloaded, the periods are turned into NA’s.
5. 3/2/2020 – Dan
   1. Added table to display uploaded age data. We had already done this for the sample data verification tab, but apparently never did this for the age verification tab
6. 1/31/2022 – Dan
   1. Add several validation checks. All but Wr and Station checks (which only occurs on sample data) were made on both sample data and age data. Changes were as follows:
      1. Test for abnormally large or small individuals (using table ODWC gave with min and max sizes as thresholds)…these are quite conservative and will flag a lot of data, so may need adjusting.
      2. Explicit test for column names and order (will allow additional columns to be added to SSP database, but only as approved and in right order with right names)
      3. Test if month, day, year make an actual date
      4. Tightened up range for acceptable Wr from 20-200 to 60-120 (was ODWC’s idea for finding abnormally light or heavy fish)
      5. Check if number of different Station codes and number of time codes match…prevents using same station code for more than one sample.
   2. Added code that disables download button unless all validations are “Okay” with the exception of abnormally large/small fish and Wr as these could have exceptions. This was done on both sample and age validation tabs.
   3. Added a datatable at the bottom of the age validation tab to display data with row numbers to make it easier to see what the validation problem might be.
7. 6/22/2022 – Dan
   1. Added new validation rules decided on with Teams meeting with Ashley Nealis and Cliff Sager discussing issues of fish size/Wr on new data upload
      1. No check for abnormally small fish from seines (gear 10)
      2. Max size for abnormally large fish for Gizzard Shad (spp 501) 483 mm TL (had been 345 mm)
      3. Max size for abnormally large fish for White Bass (spp 109) 450 mm TL (had been 363 mm)
      4. Abnormal Wr now will be <50 or >150 (had been <20 and >200, later changed to <60 and >120…hopefully this is now a good cuttoff that is perhaps a little generous, but will at least catch the most extreme errors).
   2. Added new column called Verified.TL.Wr to template that will have “verified” or “Verified” for any row with abnormal TL/ Wr that user has now checked and wants to keep. I have then made it so you cannot download a file that has abnormal length or Wr without a “verified” flag set. (Note there is no check on Wt directly…only tested via Wr value).
   3. Also added code to deal with situation if there is no valid Gear.Code or if >1 Gear.Code was in file…was giving error for Gear.Length and Effort row numbers in far right check boxes that display offending rows…I changed to add a message about needing one and only one valid gear code instead of a list of row numbers.
8. 9/1/2022 – Dan
   1. Changed Verified.TL.Wr to separate Verified.TL and Verified.Wr columns…realized people might fix a TL issue by marking it verified when the Wr value still indicates a weight error that would then be skipped.
9. 6/20/2023 – Dan
   1. Corrected typo in ui.R of validation app that mis-represented acceptable Wr values as <60 or >120 and corrected to <50 or >150, which is what the app is actually testing.
10. 1/18/2024 (first 4 items from summer 2023 but were not pushed then)-Dan
    1. Force only “Validated” or “validated” for Verified.TL and Verified.Wr columns. Added text to right column indicating what to type in these fields as a reminder.
    2. Changed download of verified data to add lake year and gear so the format of file name is now: lake\_year\_gear.code\_Sample\_verified\_date.csv
    3. Added clarifying text about “needing period in any blank cell” in right hand column text related to this.
    4. Require TL\_mm (SSP data) and TLmm (Age data) to be integers (S-central region measures in inches and converts to mm...this will require them to round to whole number to pass validation). This was important as it changes the data type and creates problems when merging into the main databases if these columns are not integer data types.
    5. Added lake codes for VERLIN and SMITH (requested by Austin Griffin and approved by Ashley Nealis)
    6. Removed code (“if(input$validateSamp != 0){”) from Render sample error table##### section so we no longer need to press “validate data” to get the error table to load after the file is first uploaded. I then cleaned up the ui.R elements (removed the button and renumbered the steps to use) This simplifies the user experience as validation tests just automatically run once the file is uploaded. Also edited comparable lines on the age data validation code sections.
    7. Removed all\_equal() function that tests if all column have same name and in the same order as it has been depreciated. Now use identical() which works essentially as a drop in replacement given how I had all\_equal() set up.
11. 1/24/2024 – Dan bug fixes found by Kendal Robbins related to new “verified” columns not working
    1. Had some “or” statements where I needed “and”
    2. Did not anticipate people writing “verified” with quote marks (but I can see from the instructions why they would...I also changed instructions to indicate quotes not needed), so added code to pick up “verified” with quotes as marked verified.
    3. Found/fixed code issue in the non\_integerTL code I had...was minor but produced an error if you clicked on the button to get row numbers when there were no offending rows.
12. 2/13/2024 – Dan added code to refine ages allowed on age data validation tab
    1. 0-20 for most species now (was up to 40)
    2. 0-40 for large catfish spp (spp code 114, 115, 401)
    3. 0-100 for long lived species (spp code 302-307, 311, 321, 325 or 402-405)
13. 5/29/2024 – Dan bug fix for issue found by Madison Mitchell where having blank gear length caused app to crash. This resulted from section of code that checks gear length and gear effort to be sure they are reasonable values
    1. Dray had coded this to identify offending rows using a new column (NAlength or NAeffort) where “TRUE” indicates missing values that are required, filtered by those rows, then used a column number to delete the new variable once filtered...but now that we are adding Verified.TL and Verified.Wr, there are 2 more columns to the left of the column to be deleted, so this was not deleting the right column resulting in error when rbind was used to merge this dataframe with one that has gear length or effort values that were present but out of the legal range.
    2. Fix:
       1. I replaced with a dplyr select() function to drop the column by column name instead of number
       2. Also searched code for [, or ,] to try to find/replace all other instances where column numbers instead of names had been used....found the following and replaced with select() function:
          1. In code to download validated sample, column numbers were used to rearrange column order as last step before writing file
          2. Code checking for number of individuals (NOI) is not 0 or NA
          3. Block of code creating output$sampleLength that is used to print row numbers when checkbox is checked.
          4. Block of code creating output$sampleEffort that is used to print row numbers when checkbox is checked.
          5. Code evaluating if NOI is 0 or NA in age data (2 locations in code)
          6. Code evaluating if TL was blank on age data (NATLAge) (2 locations in code)
    3. Added some code to put up a modal box if an Age data file were uploaded for sampling validation or if a Sampling data file were uploaded for age data validation. This prevents crashing app if the wrong file type is loaded.