

CWT Data Status for 2016

Regional Committee on Marking and Tagging

As of 4/22/2016

Dan Webb

The focus of this discussion is to review the reporting status of CWT database and identify any obstacles to reporting, validating and processing of the data.

Changes in Data Reporting:

California Department of Fish and Wildlife (CDFW) – Melodie Palmer will continue to provide CDFW catch/sample and recovery data possible up through mid-2017 however, replacement staff and training is already being addressed

Oregon Department of Fish and Wildlife (ODFW) – Gabe Garza has taken on responsibilities' to report ODFW data.

Locations

All locations necessary for Coded Wire Tag (CWT) validation and processing purposes are believed to be up to date.

Releases

California Department of Fish and Wildlife (CDFW), US Fish and Wildlife (FWS), Northwest Indian Fisheries Commission (NIFC), Washington Department of Fish and Wildlife (WDFW) and Yakama Nation (YAKA), have all made additional release submission within the past 2 weeks. Columbia River Inter-Tribal Fish Commission (CRFC) and Stillaguamish Tribe (STIL) have data sets last reported in August 2015. All other agencies have updated releases since December 2015.

Review of counts per year per agency for the past 10 years show most agencies are consistent with recent trends. California Department of Fish and Wildlife (CDFW) releases are mostly reported however, apx. 4 mil from Trinity Hatchery (~2.5 mil fall, ~1.5 mil spring) have been delayed.

There are currently no releases in the CWT database with the MIDYEAR status.

Recoveries

Recoveries were reviewed for data sets with outstanding errors and years missing up to and including the 2014 run year. 2015 run year data sets are not yet evaluated as missing since reporting of the run year may be dependent upon escapement data that is still unavailable.

Alaska Department of Fish and Game (ADFG), Department of Fisheries and Oceans, Canada (CDFO), California Department of Fish and Wildlife (CDFW), U.S. Fish and Wildlife Service (FWS), Idaho Department of Fish and Game (IDFG), Nez Perce Tribe (NEZP), National Marine Fisheries Service (NMFS), Oregon Department of Fish and Wildlife (ODFW), & Washington Department of Fish and Wildlife (WDFW) have reported recovery data sets for run years up to and including 2015.

Northwest Indian Fisheries Commission (NIFC), National Marine Fisheries Service – Northwest Region (NMFSNWR), & Stillaguamish Tribe (STIL) have reported recovery data sets for run years up to and including 2014.

Yakama Nation (YAKA) have not reported any recovery data sets since the RCMT meeting last year however, they do have run year 2014 reported.

Recovery years with data sets missing or containing errors since last submission:

National Marine Fisheries Service (NMFS) and U.S. Fish and Wildlife Service have datasets submitted up to and including 2015, however for run_year 2010 NMFS has 3 minor tag type discrepancy errors and for run_year 2014 FWS has 228 errors for tag discrepancies.

The following agencies have no status change since last year:

California Department of Fish and Wildlife - Klamath/Trinity (CDFWKT) – 2008 thru 2011 and 2014 data sets are missing. 2012 and 2013 are present.

Columbia River Inter-Tribal Fish Commission (CRFC) – 2002 thru present data sets are now being reported by individual tribal agencies. Their only remaining data sets are for 2000 & 2001 and are for the Yakama Nation (YAKA) sampling agency.

Quinault Department of Natural Resources (QDNR) – 2013 data set was submitted in April 2015 but has not yet passed validation. 2014 data set is missing.

Quileute Tribe (QUIL) – 2010 data errors, 2011 thru 2014 data sets are missing. They are currently working with Northwest Indian Fisheries Commission (NIFC) to resolve data entry application issues.

Yurok Tribe Fisheries Program (YTFP) – 2006 data set has errors. (NOTE: 2008 typo from last year status). 2007 thru 2011 are not present. 2012 & 2013 data for the sampling agency = has been included in the California Klamath/Trinity (CDFGKT) datasets.

Catch/Sample

Catch/Sample was reviewed with the same criteria as the recovery data for missing data sets and data sets with errors.

Alaska Department of Fish and Game (ADFG), Department of Fisheries and Oceans, Canada (CDFO), U.S. Fish and Wildlife Service (FWS), National Marine Fisheries Service (NMFS), Oregon Department of Fish and Wildlife (ODFW), & Washington Department of Fish and Wildlife (WDFW) have reported catch/sample data sets for catch years up to and including 2015.

Idaho Department of Fish and Game (IDFG) and Nez Perce Tribe (NEZP) submit recovery datasets but do not submit catch/sample datasets.

National Marine Fisheries Service – Northwest Region (NMFSNWR) & Stillaguamish Tribe (STIL) have reported catch/sample data sets for catch years up to and including 2014.

Yakama Nation (YAKA) – 2008 data has 1 error, 2009 & 2010 are missing. No additional catch/sample data have been reported since the RCMT meeting last year however, 2014 data is present.

Catch/Sample years with data sets missing or containing errors since last submission:

California Department of Fish and Wildlife (CDFW) has datasets submitted up to and including 2015 however, 2013 data set has 2 errors for escapement and fishery conflicts.

Northwest Indian Fisheries Commission (NIFC) has datasets submitted up to and including 2014 however, one 2014 data error where number recovered & decoded exceeds the sample number.

The following agencies have no status change since last year:

California Department of Fish and Wildlife – Klamath/Trinity (CDFWKT) – Same as Recoveries. 2008 thru 2011 data sets are missing. 2012 & 2013 are present and 2014 is missing.

Columbia River Inter-Tribal Fish Commission (CRFC) – 2002 thru present data sets are now being reported by individual tribal agencies. Only has data sets for 2000 & 2001 remaining. All the catch/sample are for the Yakama Nation (YAKA) sampling agency. Same as Recoveries.

Quinault Department of Natural Resources (QDNR) - 2013 data set has errors and 2014 is missing.

Quileute Tribe (QUIL) – 2009 data has 2 errors, 2011 thru 2014 is missing.

Yurok Tribe Fisheries Program (YTFP) – The same status as the recovery data.

Missing Tag Codes

Some agencies have taken the initiative to contact other agencies directly to resolve missing CWT tag codes. Other agencies as well as individuals using the CWT database for research purposes contact us (RMPC) directly for assistance in tracking down missing release records.

The Tag Prefix Contact List indicates the primary agency associated with each tag prefix.

In the event missing tags are not resolved in a timely manner we can now refer to a master “Known Good Tag Code” list provided to us by Northwest Marine Technology, Inc. (NMT). If after checking the KGT list, we do not find the tag code, we can then contact NMT directly to determine if the tag code has been manufactured since the most recent download of the list. The process speeds up tracking missing tag codes and points out bogus tag codes that may have been introduced by being mis-read or mis-recorded. The “Known Good Tag Code” list allows us to determine the correct tag type when discrepancies occur between release and recovery records.

At last year’s RCMT there were 32 instances considered to be “missing tag codes”. Since then 25 have been resolved through contacting the tag coordinators, 3 have been determined to be bogus codes and only 4 tag codes are still considered to be missing after referring to the KGT list and RMIS.

DEMO - “Known Tag Code” web service.

For those 4 missing tag codes we can contact the tag coordinator with the purchase date information and have them track down to missing release information.

191191– Bogus

220519– Bogus

222022– Bogus

055778 – Still Unreported – Tag Type 12 – Purchased 2014-09-15

633782 – Still Unreported – Tag Type 12 – Purchased 2006-10-05

635067 – Still Unreported – Tag Type 12 – Purchased 2009-03-02

635575 – Still Unreported – Tag Type 12 – Purchased 2010-06-01

½ Length Tag Code Issues

See references for tag type diagrams

Decimal Coded Wire Tags <http://www.nmt.us/products/cwt/readinginstructions.pdf>

Binary Coded Wire Tags http://www.nmt.us/support/appnotes/binary/binary_formats.pdf

Issue – some ½ length tags have the same etched code as full length tags for tag type 3 – “6 Word half-length tags” and type 15 – “Half-length Alphanumeric, includes Decimal (0.5mm)” tags. For the tag manufacturers NMT - “Northwest Marine Technologies” the tag code and tag type are used as a unique key, however in the RMIS - “Regional Mark Information System” and CWT - “Coded Wire Tag” database only the tag code itself was being used. Without including the tag type as part of the unique key in the CWT database we have ended up with duplicate tag codes between ½ length and full length tag. Ensuring that the tag type is recorded correctly has become critical to identifying distinct codes within RMIS and the CWT database.

Older type – 1 “Half tags (H type)” and type -2 “Half tags (B type)” tag codes in the CWT database begin with the letters “H” or “B”, and are not considered to conflict with any other tag type codes.

The current approach to resolve the issues of identifying ½ length tags and to ensuring their uniqueness from full length tags is to:

- 1) Expand the 6 digit ½ length codes into 10 digit codes when reported to the RMIS and CWT database for tag types 3 and 15
- 2) Enforce validation rules that require the recovery tag type to match the release tag type for ½ length tag types 3 and 15.
- 3) Refer to an NMT supplied master list of “Known Good Tag Codes” - KGT to resolve tag type discrepancies.

NMT has agreed to restrict any further production of ½ length tag codes that could be confused with full length tag codes when used in the RMIS and CWT database.

NOTE: If actual 10 digit tag codes are produced in the future, the current expanded 10 digit ½ length codes will also need to be blocked out to avoid further duplication of codes.

Outstanding issues with ½ length tag types in Releases

6 digit ½ length tag codes

Zero 6-digit (less than 10 characters) tag type “3” or “15” releases

2 – FWS 6 digit tag codes with RL tag type “12” and 2 RC tag type “15” recoveries that should probably all be tag type “0” (tag codes 054024 & 054026

NOTE:0504000204 & 0504000206 Not Found In KGT List).

Tag Type Mis-Match with Releases and “Known Good Tag Code” list

40 tag codes with tag type “3” or “15” that Do Not Match KGT list.

NOTE: all are 10-digit codes with mis-matched types

- 7 - WDFW tags prior to 2002 with RL tag type “0” instead of type “3”.
 - 4 of the tag codes have duplicate 6-digit tag type “0” codes and 1,070 Tag Status 1 Recoveries for the 10 digit codes and 398 Status 1 Recoveries for the 6-digit codes.
 - 3 of the tag codes had no duplicate 6-digit codes found in the current KGT list and 48 Tag Status 1 Recoveries
- 33 - CDFO tags since 2000 with RL tag type “0”, “12” or “3” instead of “15”.
 - 19 of the tag codes have duplicate 6-digit tag type “0” codes and 1 Tag Status 1 Recovery for the 10 digit codes and 1750 Tag Status 1 Recoveries for the 6-digit codes.
 - 14 of the tag codes had no duplicate 6 digit codes found in the current KGT list and had Zero Recoveries.

Outstanding issues with ½ length tag types in Recoveries

The recoveries are far more complicated than releases, are dependent upon the accuracy of the release reporting and have been found to have tag types inconsistent with Release as well as the KGT list. At this time we feel that it is most important to first resolve any tag type and duplicate half vs full length tag code discrepancies in releases before further analyzing the impacts to recoveries.

Initial review of recoveries show apx 10,000 records from apx 1987 to present OR apx 4,000 records within the past 10 years with tag type discrepancies between the release and recovery records for the targeted ½ length tag types. Once release data are correct a more accurate evaluation of recoveries can occur.

Efforts to resolve existing tag code conflicts within the database

The primary focus is to perform additional validation for any tag codes in which a release record OR recovery record (with tag status “1”) has been reported with either a tag type “3” or “15”. The additional validation ensures that the ½ length tags are reported as 10 character codes AND that the recovery tag type must match the release tag type for the code and length of the code. Failure to meet the additional validation rules will result as an “error” - |ERR| in the validation log files and will block the records from being added or updated in the CWT database. For tag types other than “3” or “15” where the tag type does not match between releases and recoveries a “warning” - |MSG| will occur in the validation log files to inform the data provider of the mis-match, however the records will not be blocked from insertion or update in the CWT database. Integrity of other tag types between releases and recoveries has not historically been enforced, are inconsistent between use and reporting and will now impede to ability to continue loading data into the CWT database.

NOTE the new requirement that any resubmissions where either the release or recovery have a tag type “3” or “15” may require manual review and update of a record if the release tag type changes and tag status “1” recoveries already exist.

Scope of the project?

For Releases, do we resolve all release records with tag type “3” or “15” or limit the resolution to the past 10 years?

For Recoveries , how much of the recovery data needs to be resolved? All Recoveries, the past 10 years or only upon re-reporting of datasets?

Using editors to modify CWT data

On occasion we receive data files that do not validate, process well or load correctly with the CWT database. Often we find the use of various word processors, data formatting tools and spreadsheets will introduce anomalies that do not conform to basic “plain text csv files” or the data standards specifications. We have noticed an increasing trend of tools used to modify data that introduce anomalies in the data submitted for validation. When encountered we have modified our processes to ensure the anomalies do not impede data flow or impact the CWT database. The anomalies have been occurring with enough frequency that we wish to address some of the issues.

There are many processors, document formatting tools, editors and data entry tools that can and do modify, reformat and introduce unprintable special characters within data fields. RMPC would like to advise avoiding the use of such tools to modify data sets intended for submission to the CWT database.

Tools to avoid include any that format text, convert fields between text, numeric and other datatypes or use formulas to modify data.

Errors introduced by word processors include special characters for formatting and printing such as bold, italic, indents, paragraph marks or any other control characters used to change the look or output of the content. When encountered in data files intended to be transmitted, validated, processed or loaded as “plain text csv files” any special “unprintable” characters can cause anything from failure to fully transfer the complete datafile or to only partially load the records into the CWT database.

Errors introduced by spreadsheets and tools that convert fields between text and numeric (or other data types) include modifying values by removing leading zeros, removing necessary repeating spaces, introducing special export field delimiting characters (such as tabs or “|” characters for various types of delimited files) or even sometimes possibly introducing summary formulas or values within the fields. These types of modification will generally not result in a failure to fully transfer the complete datafile. Instead they introduce incorrect corrupted data values that can occasionally pass through validation and potentially end up within the CWT database. We have frequently seen examples where tag codes, location codes, mark codes (adipose and other clips) can be modified by the spreadsheets usually removing any leading zeros. Cases where values are validated and compared to specific acceptable codes usually fail validation.

If it is necessary to update content after it is extracted from your local database it is highly recommended to use basic text editors that allow changes without any formatting or conversions of data types. Tools such as Notepad could be used. If spreadsheets are used, all columns in the spreadsheet should be defined as “text” fields prior to uploading records to the spreadsheet or modifying the data values within them.

It is best to correct content within the original database application before it is exported to ensure the edits are persistent and do not require the edits be repeated every time the data is prepared for submission. We frequently see the same data error re-introduced with subsequent submissions of data sets. These types of repeated errors demonstrate that corrections are not occurring in the original source.

Incorrect Data Record

"T","4.1","20151207","XXX","XXX","07","501050103","15",,,,,"1","1","2011","20120619","20120619","3F42001 451116 R","3F42001 290131 H02","3F42001 451116 S","S","H","O","FR","6.81","92","N","0","27646",,,,,"0","47",,,,,"B","0.002","1","588",,"LW SCS 29 White River"

Validation Errors

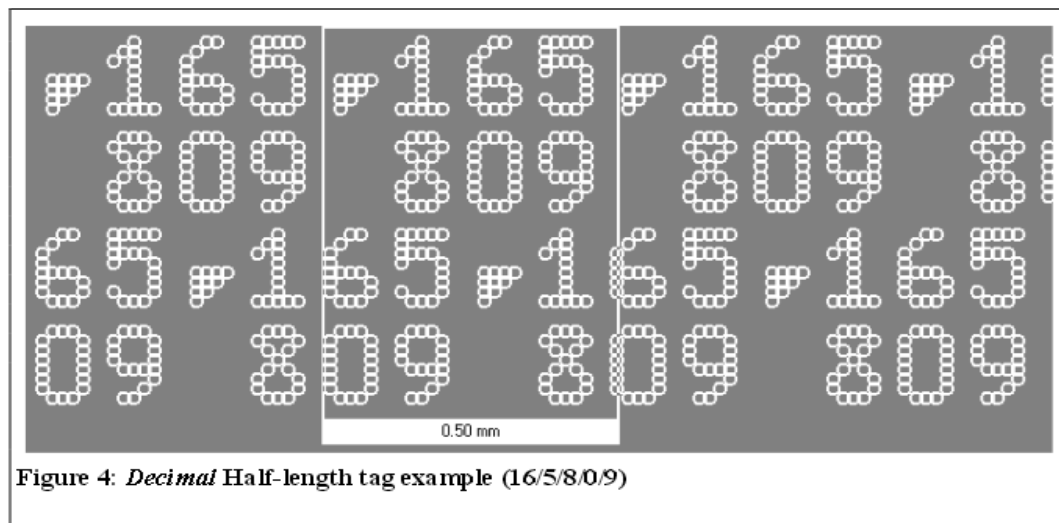
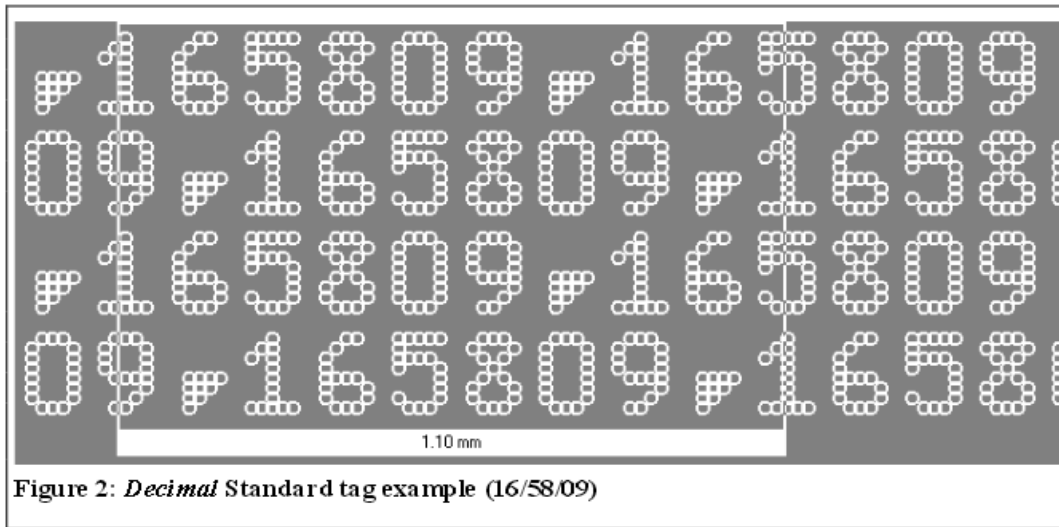
```
501050103; |ERR| tag_code_or_release_id not valid or inconsistent with
record_code; --->[501050103][T]
501050103; |ERR| tag_type inconsistent with tag_code_or_release_id for number
of recorded digits - should be 10 digits; --->[15][501050103][9]
501050103; |ERR| release_location_code not recognized; --->[3F42001 451116 R]
501050103; |ERR| hatchery_location_code not recognized; --->[3F42001 290131
H02]
501050103; |ERR| stock_location_code not recognized; --->[3F42001 451116 S]
501050103; |ERR| cwt_1st_mark not recognized; --->[0]
501050103; |ERR| non_cwt_1st_mark not recognized; --->[0]
```

Correct Data Record

"T","4.1","20151207","XXX","XXX","07","0501050103","15",,,,,"1","1","2011","20120619","20120619","3F42001 451116 R","3F42001 290131 H02","3F42001 451116 S","S","H","O","FR","6.81","92","N","0000","27646",,,,,"0000","47",,,,,"B","0.002","1","588",,"LW SCS 29 White River"

Decimal Standard vs Decimal Half Length

<http://www.nmt.us/products/cwt/readinginstructions.pdf>



Binary Standard vs Binary 6 word half length

http://www.nmt.us/support/appnotes/binary/binary_formats.pdf

BINARY CODED WIRE TAG (CWT) FORMATS

Here is a list of formats of binary tags which are currently being produced.

1. Standard – This is the most common tag. It is 1.1mm long, .25mm in diameter, is batch coded and has four words (Master, Data 1, Agency, Data 2). Data capacity is 4,096 codes for each of 64 agency codes.
2. Half-length – This tag is .5mm long, .25mm in diameter, is batch coded and has six words (Master, Data 1, Data 2, Agency, Data 3, Data 4). Half-length tags are used when fish size (less than approximately 2 grams) cannot accommodate a larger tag. Data capacity is 32,768 codes for each of 16 agency codes.