A Tool for Data Consolidation and Standardization

Karen Horodysky

2018 OFWIM Annual Meeting

Hood River, Oregon

5 November 2018

Aquatic Comprehensive Database Initiative

To provide **comprehensive** databases that will enhance aquatic **data storage, security, analysis, and sharing** for our staff and our constituents

Proposed Aquatic Databases

Coldwater Streams

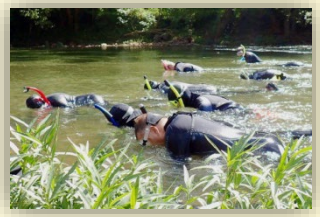
Warmwater Streams

Lakes

Creel

Nongame Fish

Nongame Invertebrates



Fish Stocking & Production

Mussel Propagation

Existing CWSS Access Database

Location Data

Collection ID
Date
Collectors
Coordinates
Reach Length
Reach Width
Water Quality
Habitat
Comments

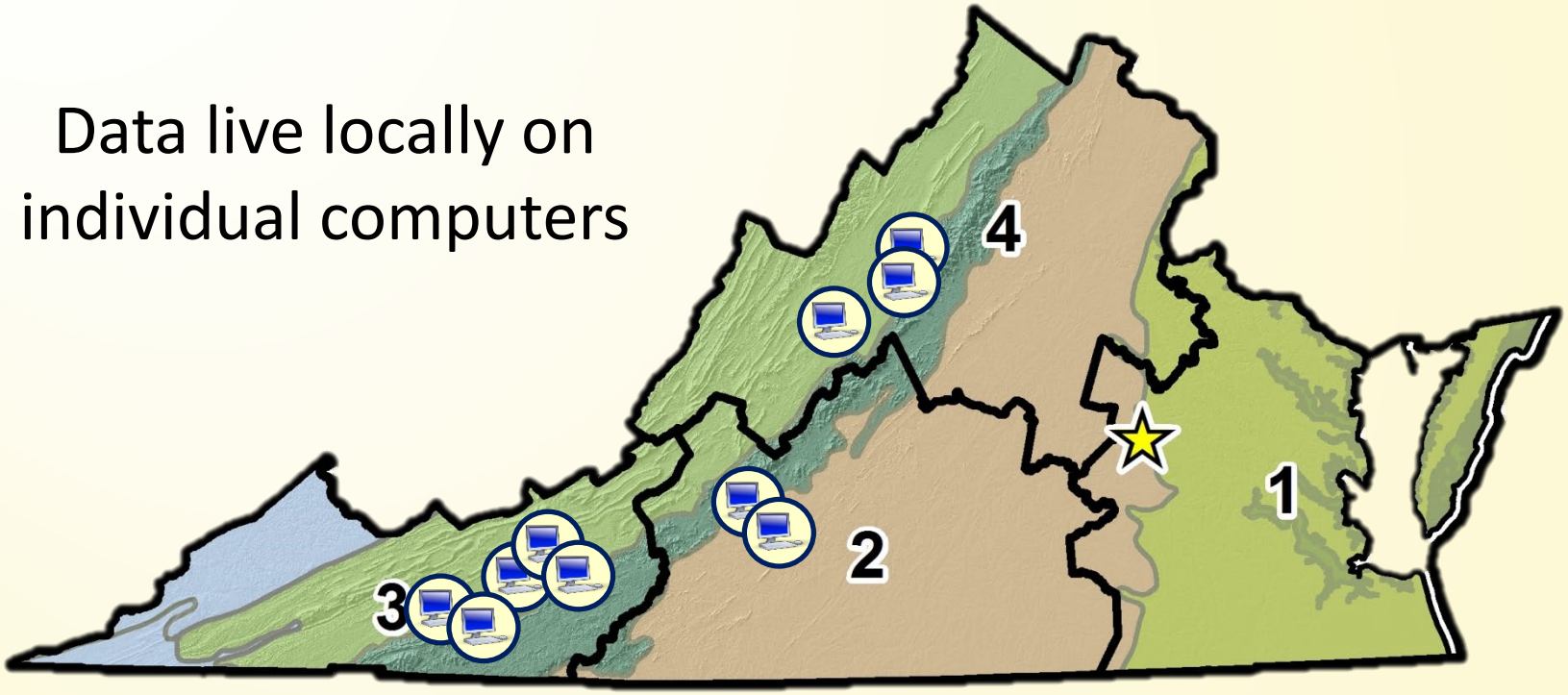
Existing CWSS Access Database

Fish Data

Collection ID
Run
Fish Code
Count
Length
Weight
Max Length
Min Length
Aggregate Weight

Existing CWSS Access Database

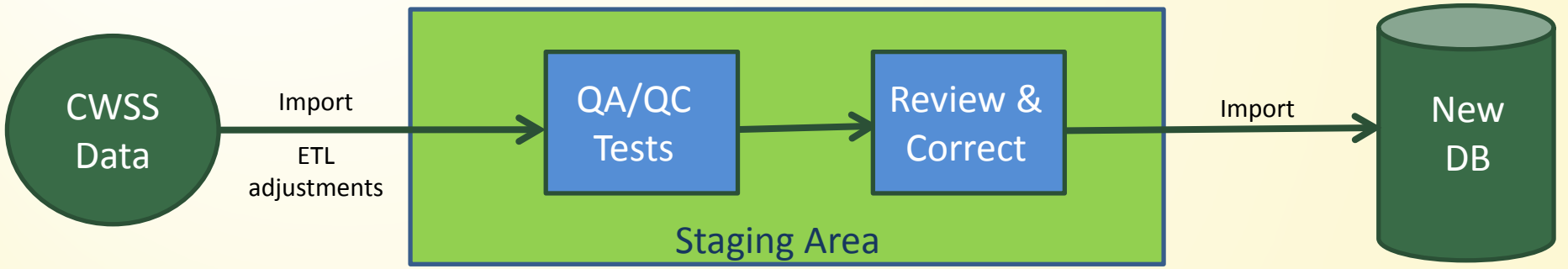
Data live locally on individual computers



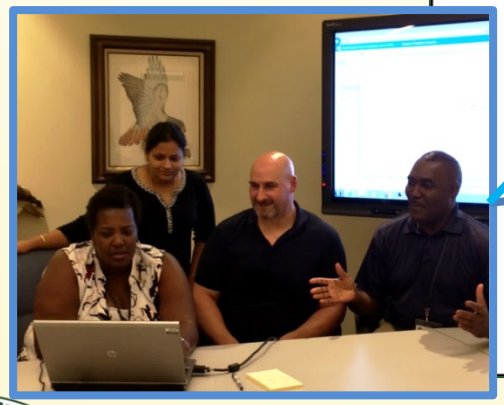
Data Consolidation & Standardization Tool

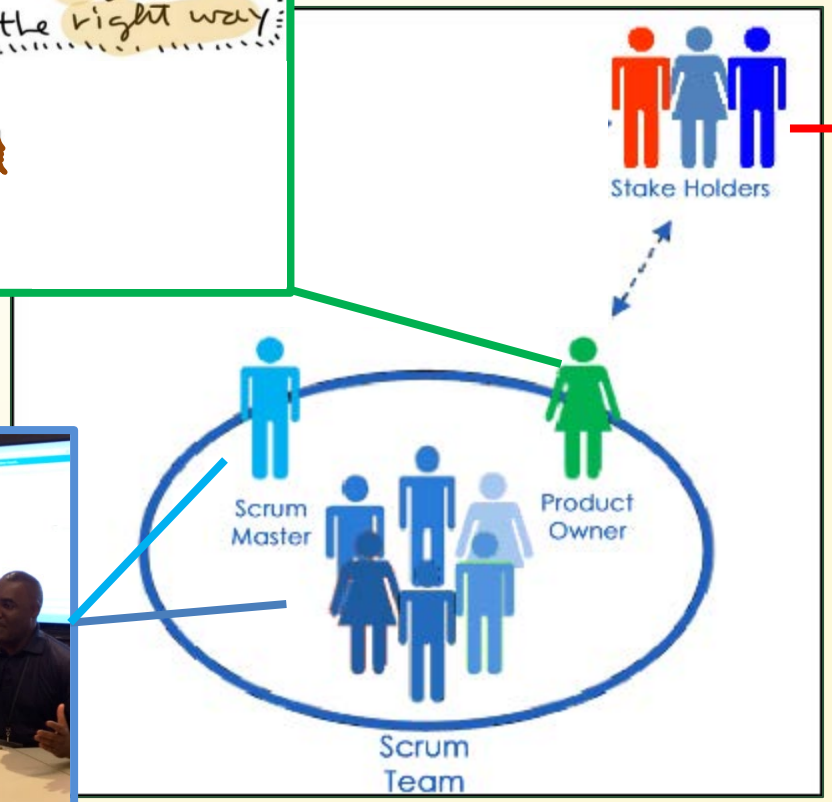
1. **Find** the problems in existing data
2. **Fix** the problems in existing data
3. **Consolidate** data to database
4. **Maintain** accountability

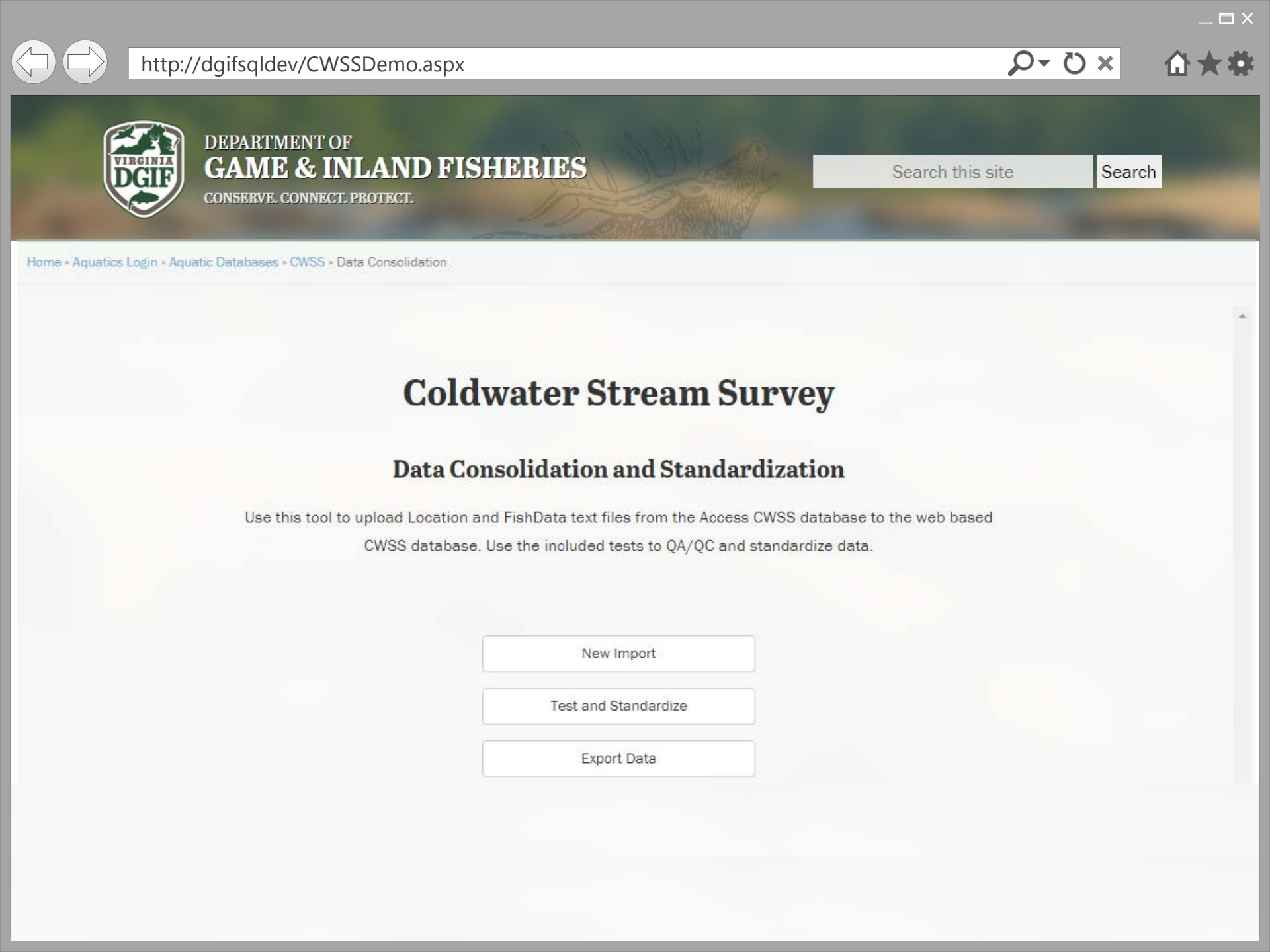
Data Consolidation & Standardization Tool



Agile Software Development
SCRUM






http://dgifsqldev/CWSSDemo.aspx
DEPARTMENT OF
GAME & INLAND FISHERIES
CONSERVE. CONNECT. PROTECT.
Search this site
Search
Home » Aquatics Login » Aquatic Databases » CWSS » Data Consolidation
Coldwater Stream Survey
Data Consolidation and Standardization
Use this tool to upload Location and FishData text files from the Access CWSS database to the web based CWSS database. Use the included tests to QA/QC and standardize data.
New Import
Test and Standardize
Export Data

Upload Data to Staging Area

Direct	File Name	File Size	Date Received	File Type	Import Flag	Selected File
\\dgifsqldev\dev$\Leroy\CWS_Updates\	fishdataDemoTest1	3453	08-16-2018	FishData		
\\dgifsqldev\dev$\Leroy\CWS_Updates\	fishdataDemoTest2	4224	08-16-2018	FishData		
\\dgifsqldev\dev$\Leroy\CWS_Updates\	fishdataHallacherJ20170420	3737116	07-25-2017	FishData		☑
\\dgifsqldev\dev$\Leroy\CWS_Updates\	fishdataHorodyskyk20170327	3088473	07-25-2017	FishData		
\\dgifsqldev\dev$\Leroy\CWS_Updates\	fishdataMaster20170301	4223713	03-01-2017	FishData		
\\dgifsqldev\dev$\Leroy\CWS_Updates\	fishdataTestData20171006	18633	10-06-2017	FishData		
\\dgifsqldev\dev$\Leroy\CWS_Updates\	fishdataTestData20171208	18633	12-08-2017	FishData		
\\dgifsqldev\dev$\Leroy\CWS_Updates\	fishdataTestData20180412	18796	04-12-2018	FishData		
\\dgifsqldev\dev$\Leroy\CWS_Updates\	fishdataTestData20180611	12802	08-16-2018	FishData		
\\dgifsqldev\dev$\Leroy\CWS_Updates\	locationsDemoTest1	20136	05-16-2018	Locations		
\\dgifsqldev\dev$\Leroy\CWS_Updates\	locationsDemoTest2	25582	05-16-2018	Locations		
\\dgifsqldev\dev$\Leroy\CWS_Updates\	locationsHallacherJ20170420	20768784	07-25-2017	Locations		Select

Data QA/QC Tests

- Codes
- Coordinates
- Values
- Fish Data Scenarios

Count	Test Type	Tested Table
10	BOVA ID Bulk Update	Fish Data
1	BOVA ID Invalid	Fish Data
3	Failed Fish Data Scenarios	Fish Data
2	Length Invalid	Fish Data
2	Collection ID Format Invalid	Location Data
1	Collection ID Format T Invalid	Location Data
1	DD Lat Invalid	Location Data
1	DD Long Invalid	Location Data
1	DMS Lat Invalid	Location Data
1	DMS Long Invalid	Location Data
1	DO Invalid	Location Data
1	Pool % Invalid	Location Data
1	Reach Code Invalid	Location Data
1	Reach Code Mismatch	Location Data
1	Water Temp Invalid	Location Data

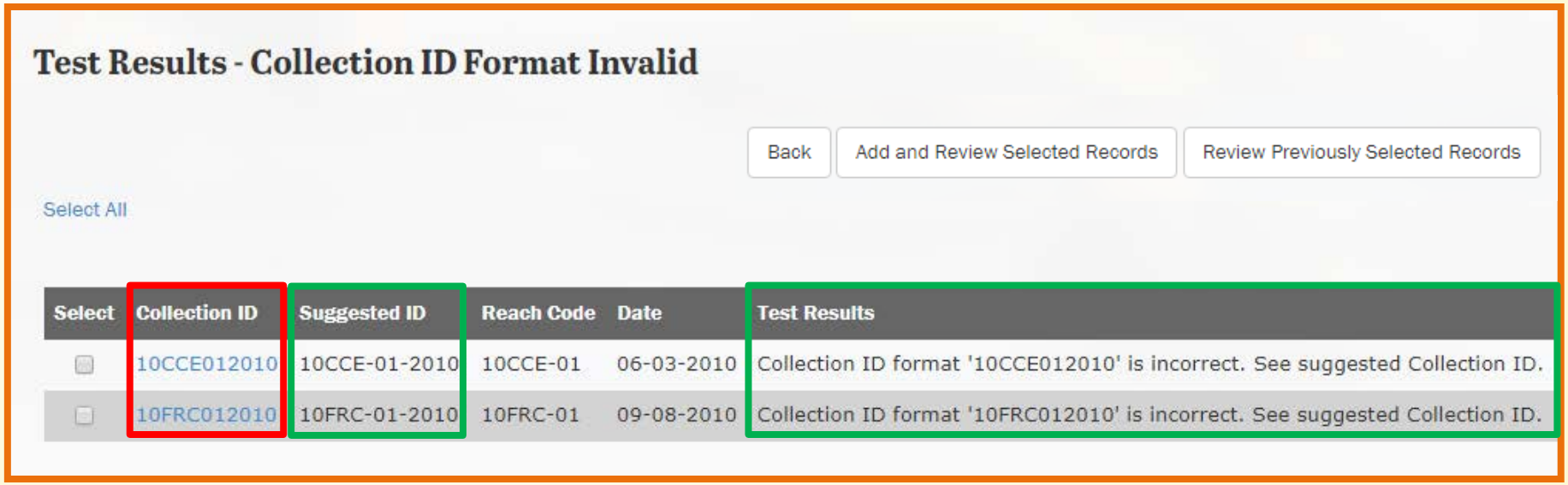
Data QA/QC Tests

- **Codes**
- Coordinates
- Values
- Fish Data Scenarios

Count	Test Type	Tested Table
10	BOVA ID Bulk Update	Fish Data
1	BOVA ID Invalid	Fish Data
3	Failed Fish Data Scenarios	Fish Data
2	Length Invalid	Fish Data
2	Collection ID Format Invalid	Location Data
1	Collection ID Format T Invalid	Location Data
1	DD Lat Invalid	Location Data
1	DD Long Invalid	Location Data
1	DMS Lat Invalid	Location Data
1	DMS Long Invalid	Location Data
1	DO Invalid	Location Data
1	Pool % Invalid	Location Data
1	Reach Code Invalid	Location Data
1	Reach Code Mismatch	Location Data
1	Water Temp Invalid	Location Data

Codes

Collection ID = 10CCE-01-2010
Reach Code Year



Test Results - Collection ID Format Invalid

Back | Add and Review Selected Records | Review Previously Selected Records

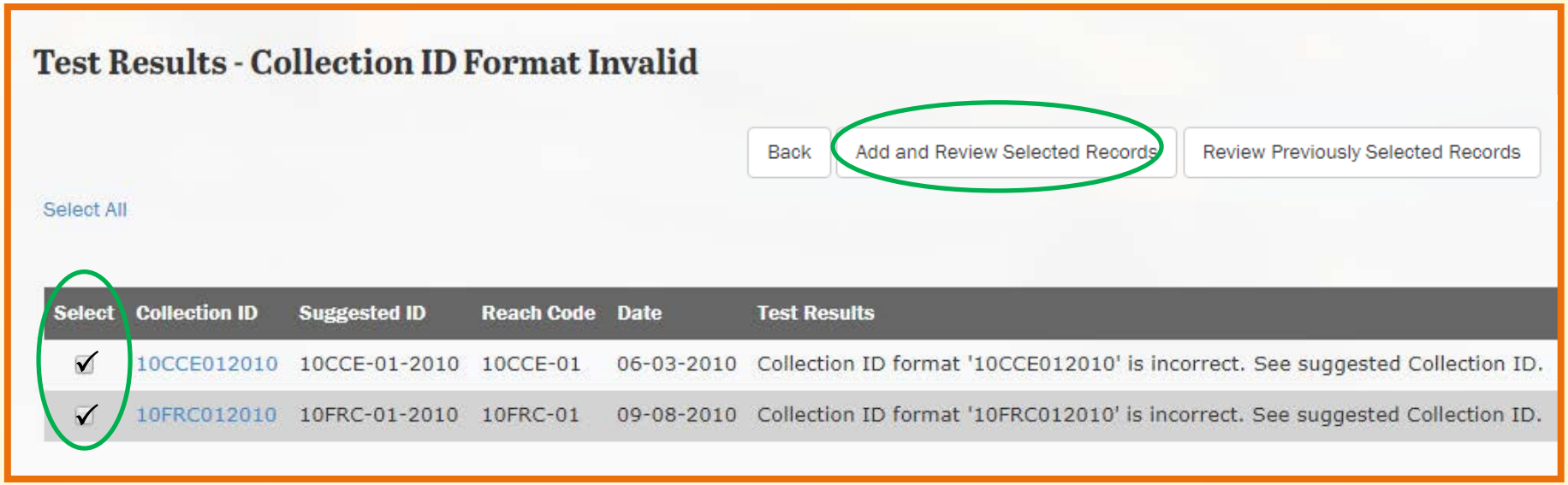
Select All

Select	Collection ID	Suggested ID	Reach Code	Date	Test Results
☐	10CCE012010	10CCE-01-2010	10CCE-01	06-03-2010	Collection ID format '10CCE012010' is incorrect. See suggested Collection ID.
☐	10FRC012010	10FRC-01-2010	10FRC-01	09-08-2010	Collection ID format '10FRC012010' is incorrect. See suggested Collection ID.

Codes

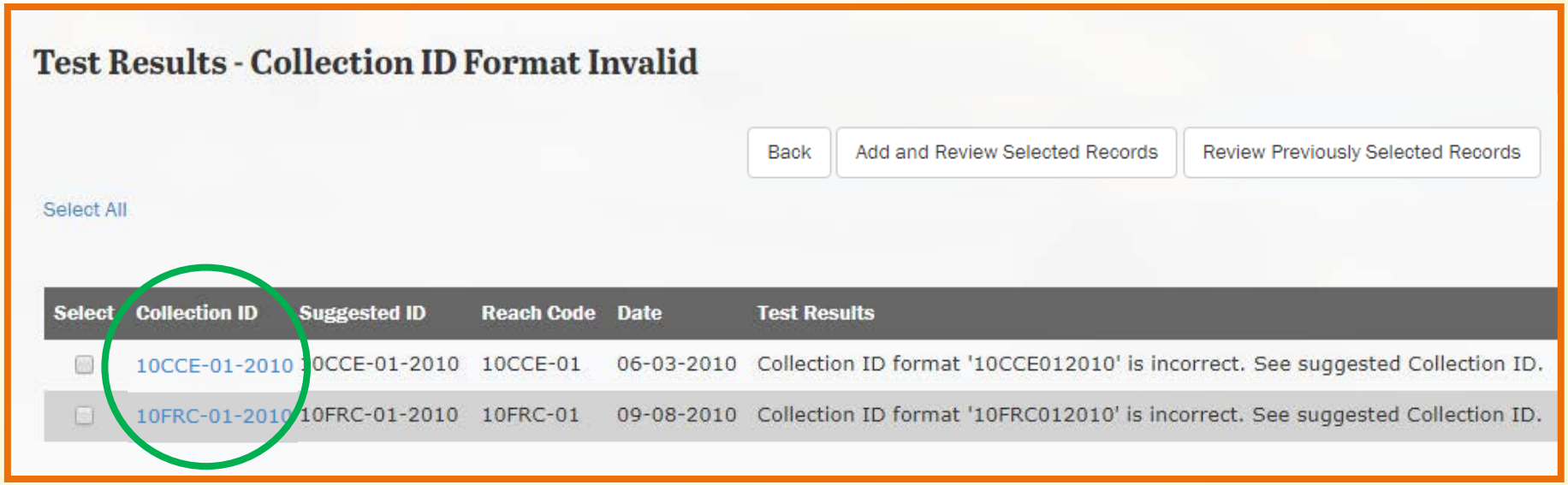
Collection ID = 10CCE-01-2010

Reach Code Year



Codes

Collection ID = 10CCE-01-2010
Reach Code Year



Test Results - Collection ID Format Invalid

Back | Add and Review Selected Records | Review Previously Selected Records

Select All

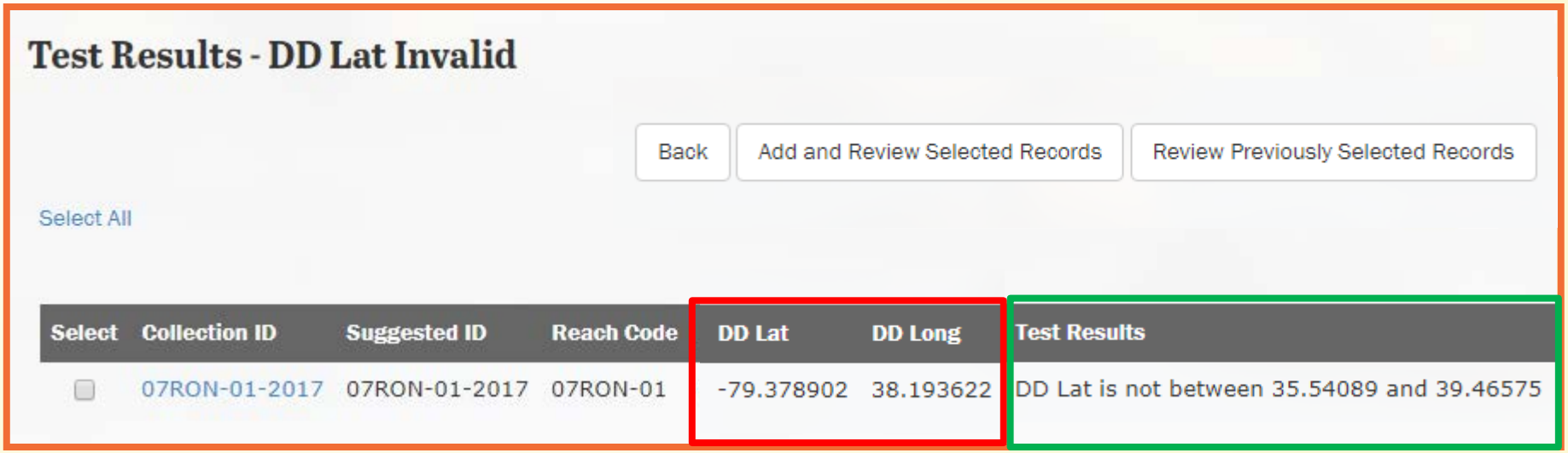
Select	Collection ID	Suggested ID	Reach Code	Date	Test Results
☐	10CCE-01-2010	10CCE-01-2010	10CCE-01	06-03-2010	Collection ID format '10CCE012010' is incorrect. See suggested Collection ID.
☐	10FRC-01-2010	10FRC-01-2010	10FRC-01	09-08-2010	Collection ID format '10FRC012010' is incorrect. See suggested Collection ID.

Data QA/QC Tests

- Codes
- **Coordinates**
- Values
- Fish Data Scenarios

Count	Test Type	Tested Table
10	BOVA ID Bulk Update	Fish Data
1	BOVA ID Invalid	Fish Data
3	Failed Fish Data Scenarios	Fish Data
2	Length Invalid	Fish Data
2	Collection ID Format Invalid	Location Data
1	Collection ID Format T Invalid	Location Data
1	DD Lat Invalid	Location Data
1	DD Long Invalid	Location Data
1	DMS Lat Invalid	Location Data
1	DMS Long Invalid	Location Data
1	DO Invalid	Location Data
1	Pool % Invalid	Location Data
1	Reach Code Invalid	Location Data
1	Reach Code Mismatch	Location Data
1	Water Temp Invalid	Location Data

Coordinates



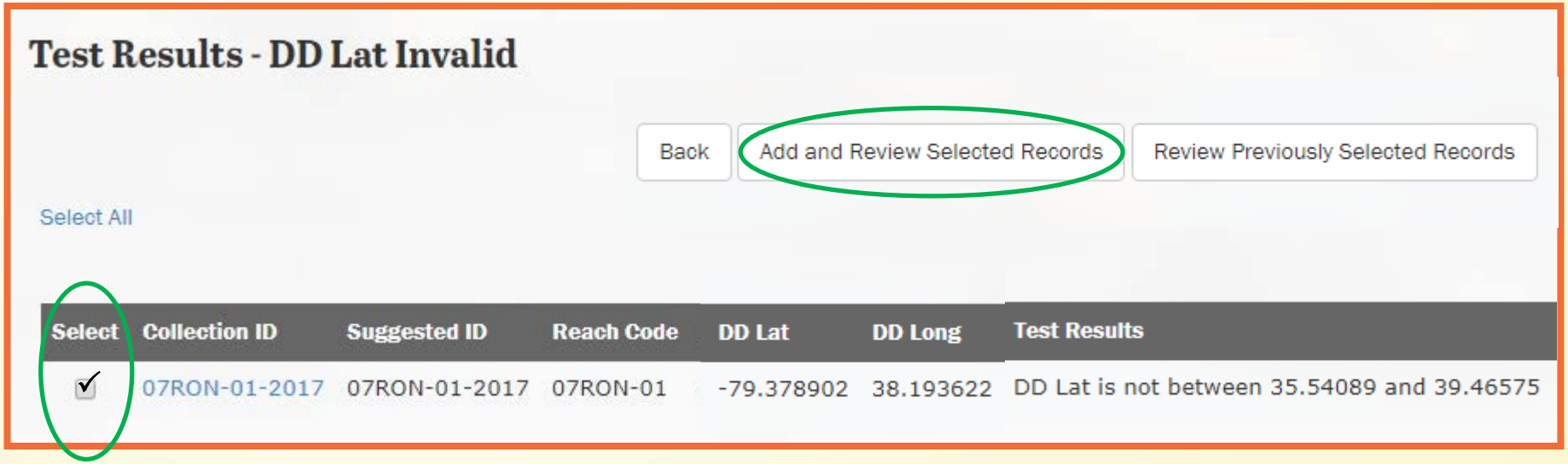
Test Results - DD Lat Invalid

Back | Add and Review Selected Records | Review Previously Selected Records

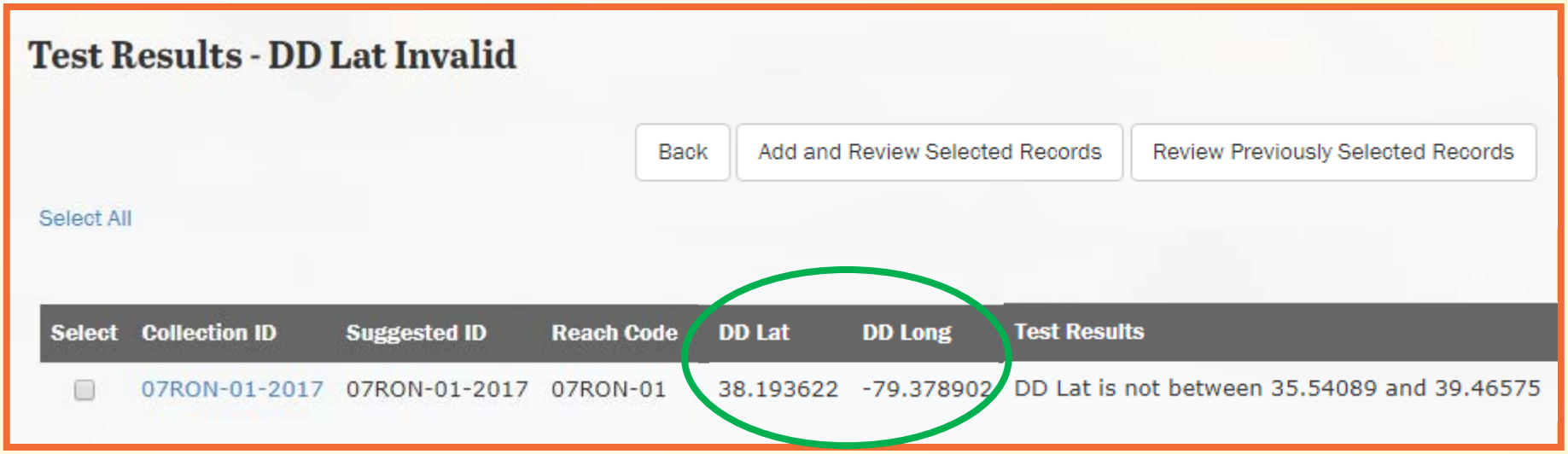
Select All

Select	Collection ID	Suggested ID	Reach Code	DD Lat	DD Long	Test Results
☐	07RON-01-2017	07RON-01-2017	07RON-01	-79.378902	38.193622	DD Lat is not between 35.54089 and 39.46575

Coordinates



Coordinates



Data QA/QC Tests

- Codes
- Coordinates
- **Values**
- Fish Data Scenarios

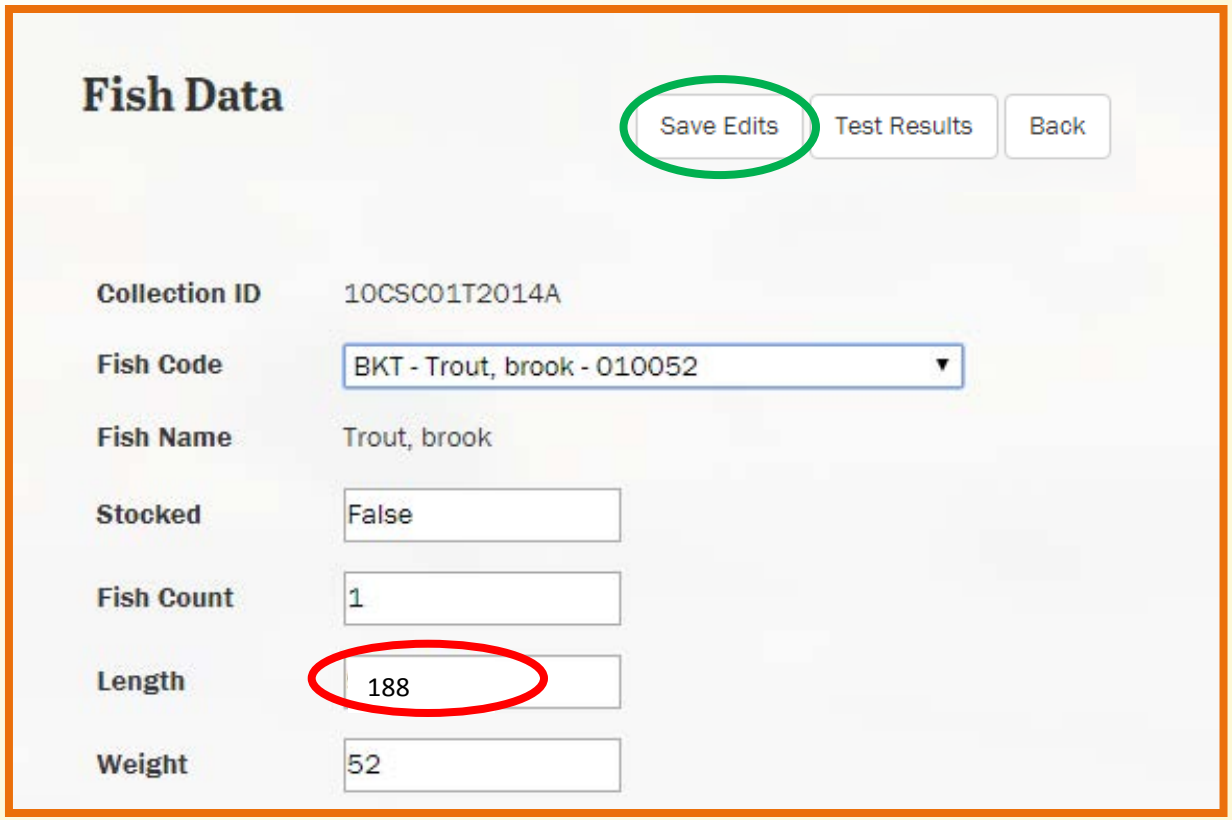
Count	Test Type	Tested Table
10	BOVA ID Bulk Update	Fish Data
1	BOVA ID Invalid	Fish Data
3	Failed Fish Data Scenarios	Fish Data
2	Length Invalid	Fish Data
2	Collection ID Format Invalid	Location Data
1	Collection ID Format T Invalid	Location Data
1	DD Lat Invalid	Location Data
1	DD Long Invalid	Location Data
1	DMS Lat Invalid	Location Data
1	DMS Long Invalid	Location Data
1	DO Invalid	Location Data
1	Pool % Invalid	Location Data
1	Reach Code Invalid	Location Data
1	Reach Code Mismatch	Location Data
1	Water Temp Invalid	Location Data

Values

Test Results - Length Invalid

Collection ID	Run	Fish Code	BOVA ID	Fish Name	Row	Length	Weight	Test Results
10CSC01T2014A	1	BKT	999999	Trout, brook	63670	987.6543	52	Length is not a whole number and/or not between 1 and 800
99RRR-01-2017	1	BKT	010052	Trout, brook	63639	899	52	Length is not a whole number and/or not between 1 and 800

Values

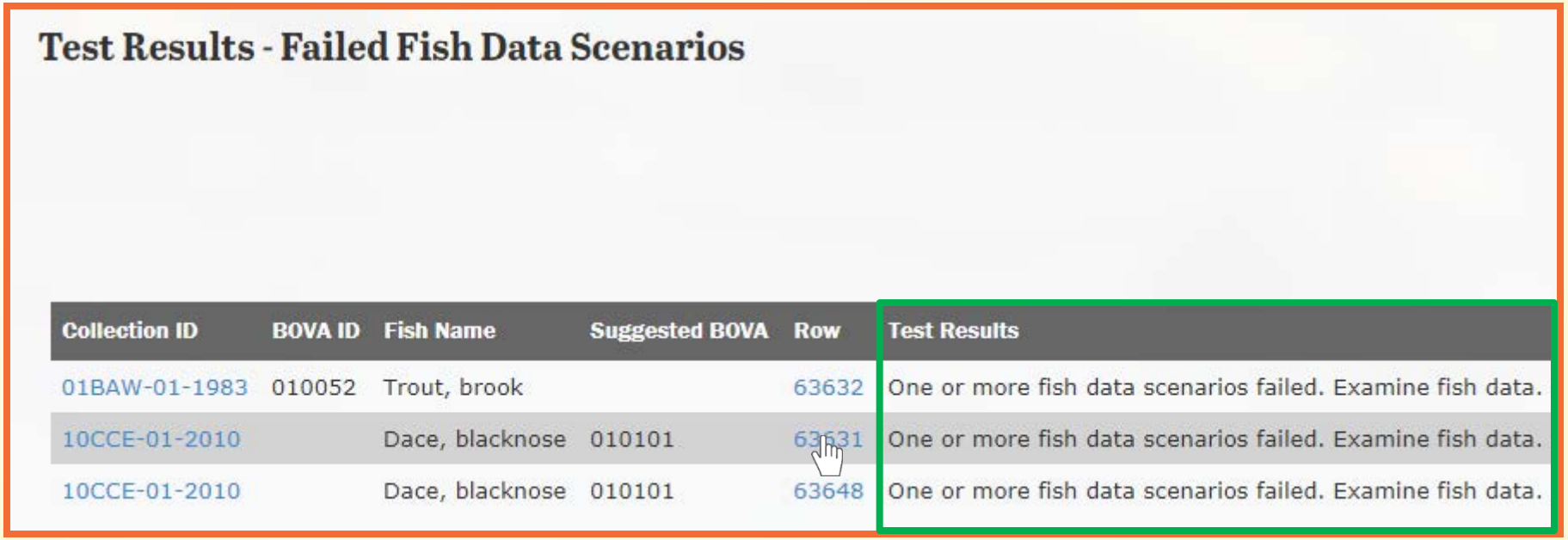

Fish Data
Save Edits
Test Results
Back
Collection ID
10CSC01T2014A
Fish Code
BKT - Trout, brook - 010052
Fish Name
Trout, brook
Stocked
False
Fish Count
1
Length
188
Weight
52

Data QA/QC Tests

- Codes
- Coordinates
- Values
- **Fish Data Scenarios**

Count	Test Type	Tested Table
10	BOVA ID Bulk Update	Fish Data
1	BOVA ID Invalid	Fish Data
3	Failed Fish Data Scenarios	Fish Data
2	Length Invalid	Fish Data
2	Collection ID Format Invalid	Location Data
1	Collection ID Format T Invalid	Location Data
1	DD Lat Invalid	Location Data
1	DD Long Invalid	Location Data
1	DMS Lat Invalid	Location Data
1	DMS Long Invalid	Location Data
1	DO Invalid	Location Data
1	Pool % Invalid	Location Data
1	Reach Code Invalid	Location Data
1	Reach Code Mismatch	Location Data
1	Water Temp Invalid	Location Data

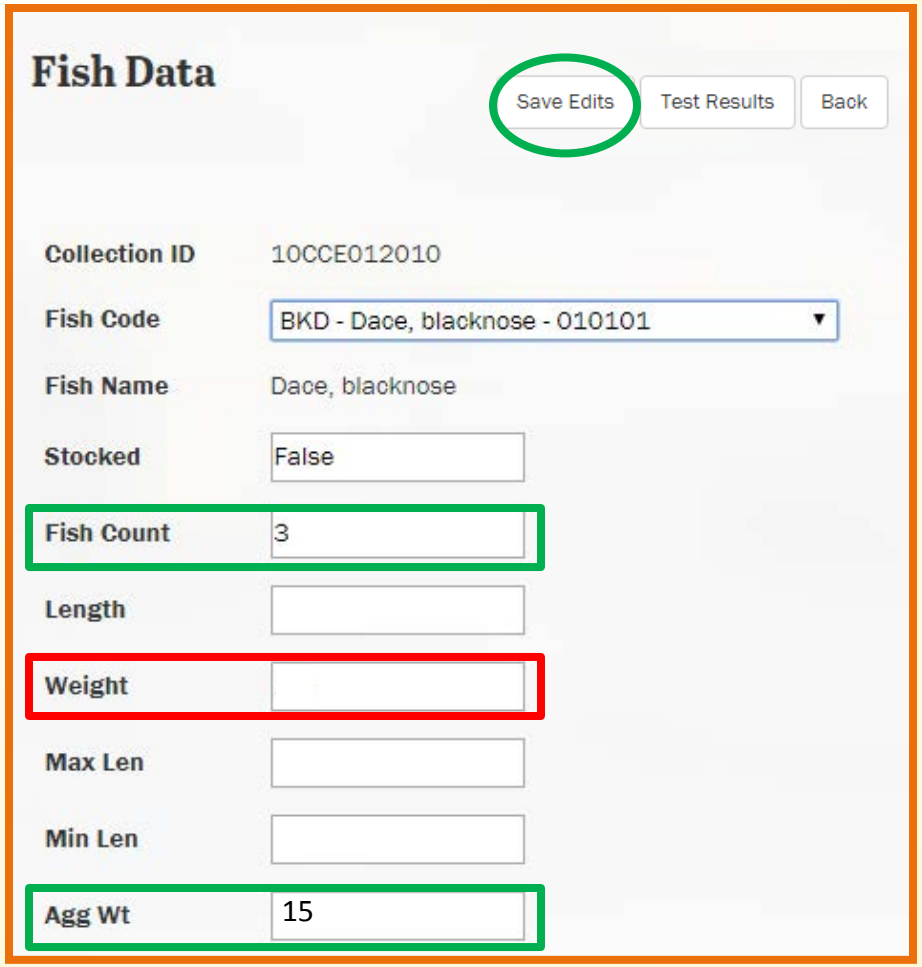
Fish Data Scenarios



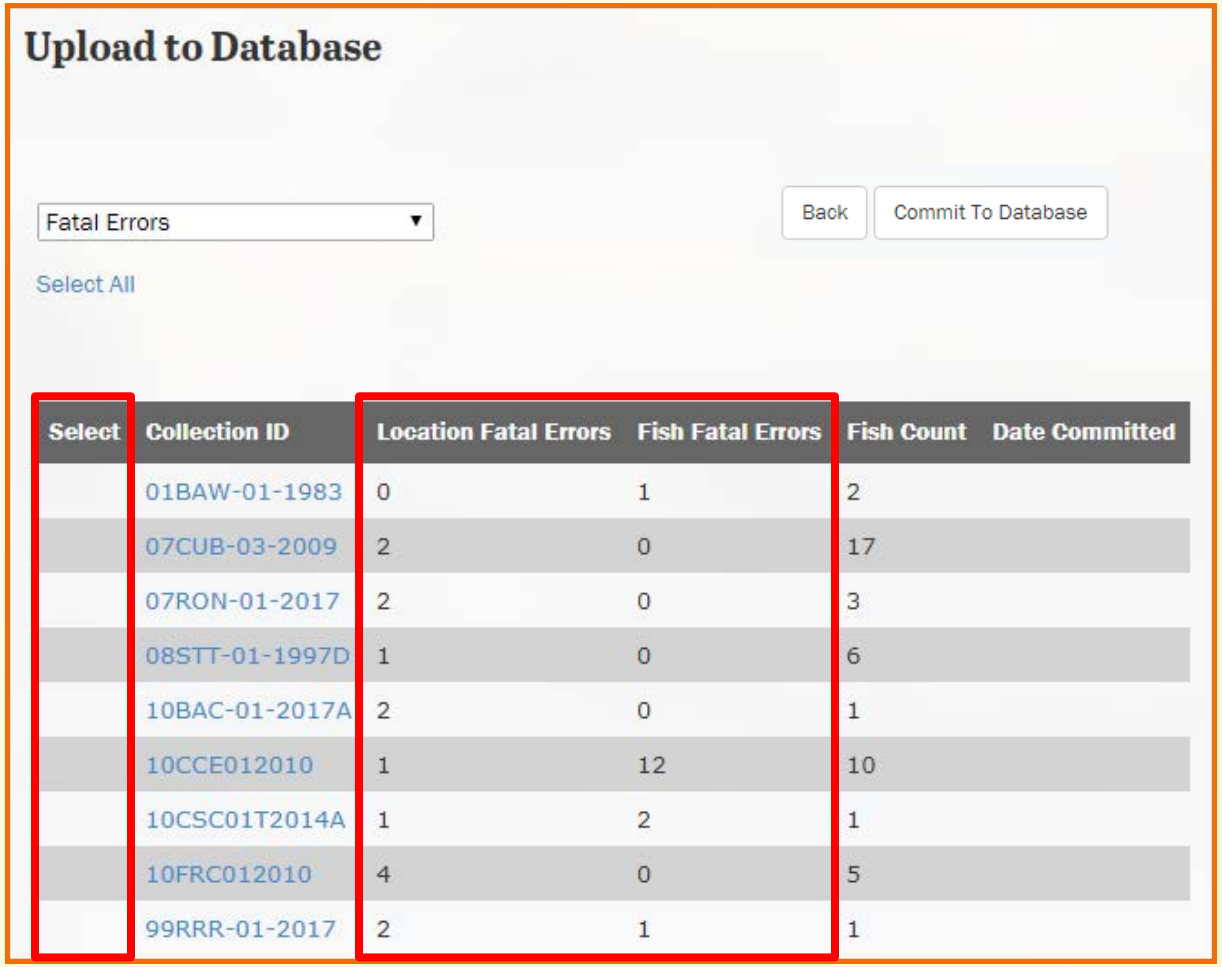
Test Results - Failed Fish Data Scenarios

Collection ID	BOVA ID	Fish Name	Suggested BOVA	Row	Test Results
01BAW-01-1983	010052	Trout, brook		63632	One or more fish data scenarios failed. Examine fish data.
10CCE-01-2010		Dace, blacknose	010101	63631	One or more fish data scenarios failed. Examine fish data.
10CCE-01-2010		Dace, blacknose	010101	63648	One or more fish data scenarios failed. Examine fish data.

Data QA/QC Tests



Import to New Database



Upload to Database

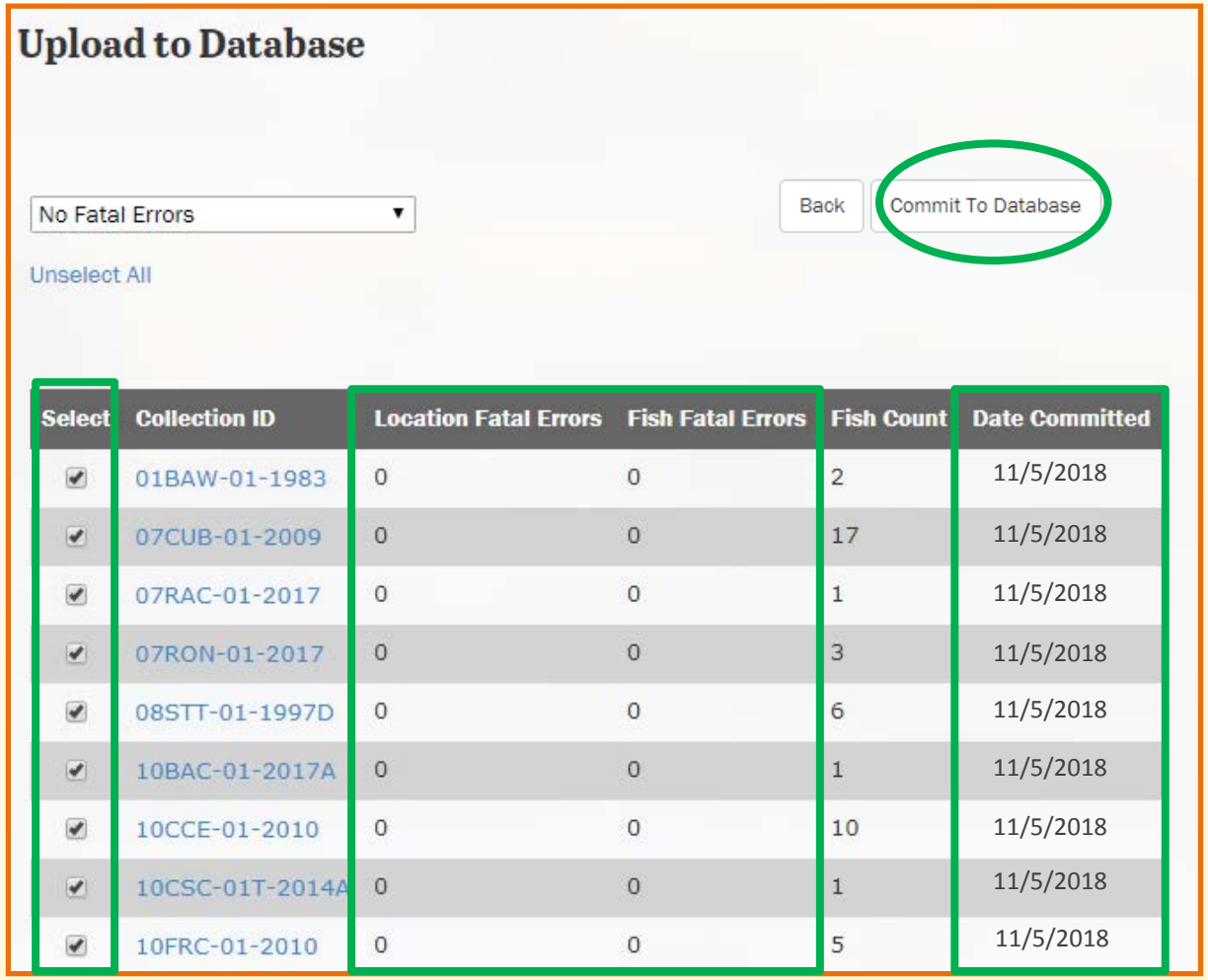
Fatal Errors

Back | Commit To Database

Select All

Select	Collection ID	Location Fatal Errors	Fish Fatal Errors	Fish Count	Date Committed
	01BAW-01-1983	0	1	2	
	07CUB-03-2009	2	0	17	
	07RON-01-2017	2	0	3	
	08STT-01-1997D	1	0	6	
	10BAC-01-2017A	2	0	1	
	10CCE012010	1	12	10	
	10CSC01T2014A	1	2	1	
	10FRC012010	4	0	5	
	99RRR-01-2017	2	1	1	

Import to New Database

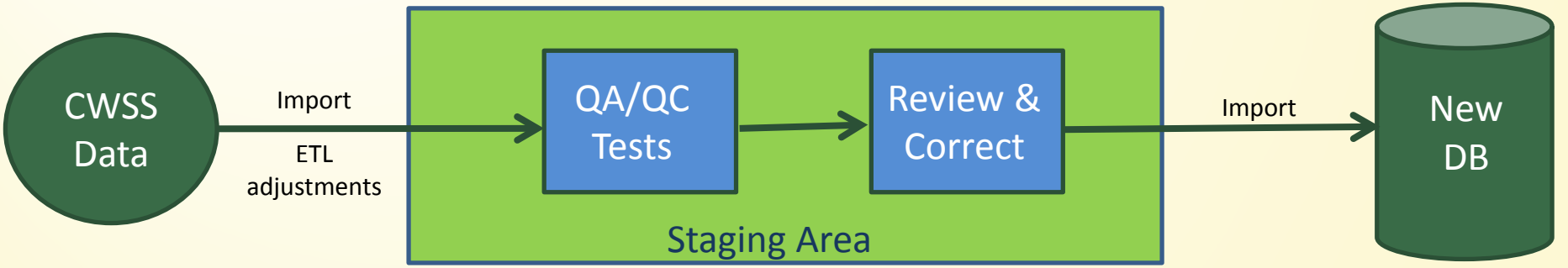


Select	Collection ID	Location Fatal Errors	Fish Fatal Errors	Fish Count	Date Committed
☑	01BAW-01-1983	0	0	2	11/5/2018
☑	07CUB-01-2009	0	0	17	11/5/2018
☑	07RAC-01-2017	0	0	1	11/5/2018
☑	07RON-01-2017	0	0	3	11/5/2018
☑	08STT-01-1997D	0	0	6	11/5/2018
☑	10BAC-01-2017A	0	0	1	11/5/2018
☑	10CCE-01-2010	0	0	10	11/5/2018
☑	10CSC-01T-2014A	0	0	1	11/5/2018
☑	10FRC-01-2010	0	0	5	11/5/2018

Full Process

Jason

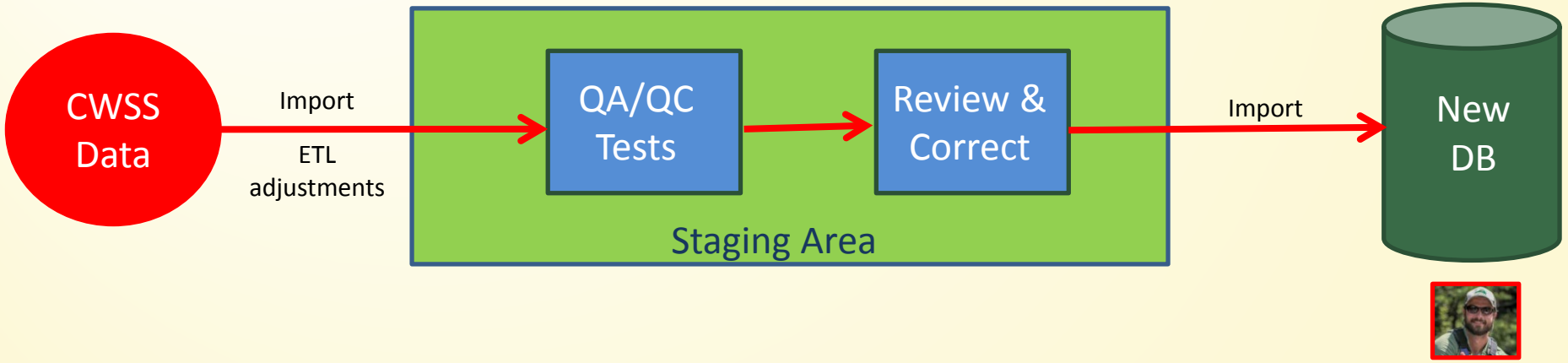
George



Full Process

Jason

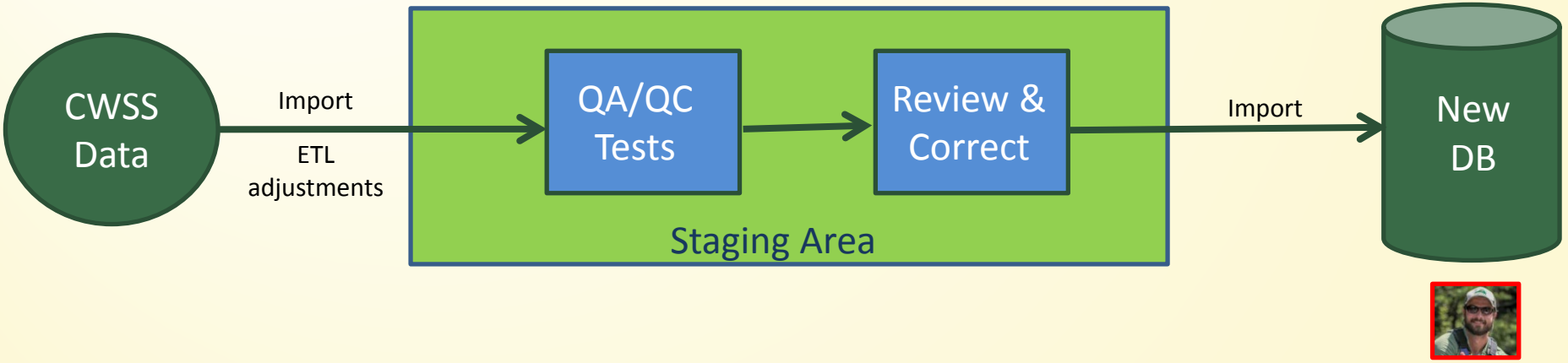
George



Full Process

Jason

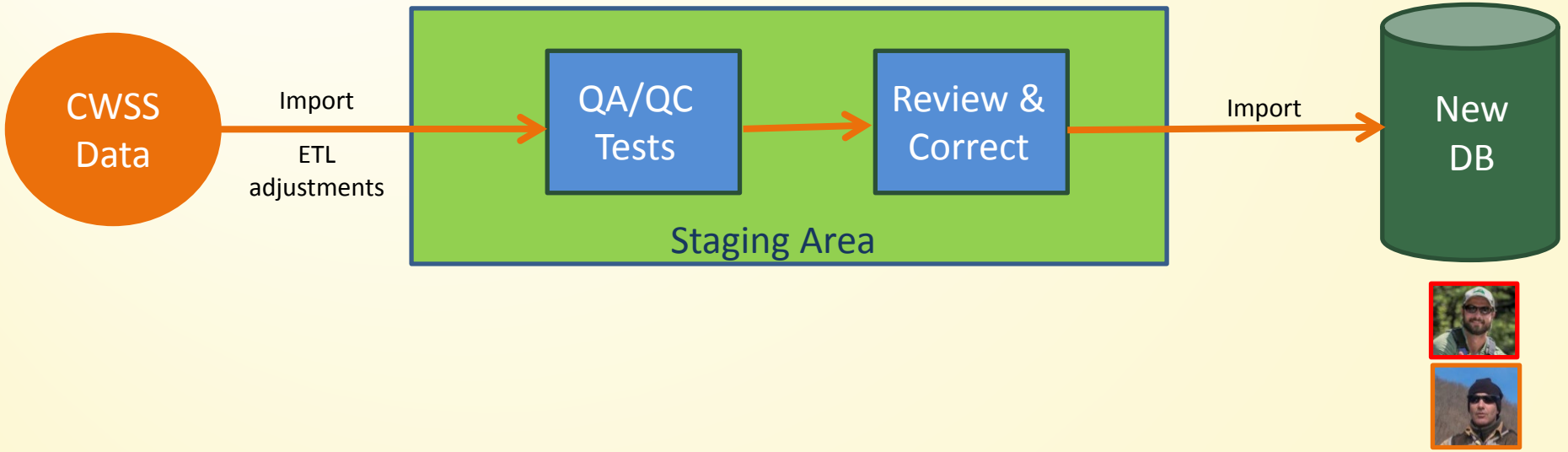
George



Full Process

Jason

George



“Ignore” data already in database

All Data

Total Count: 44

Count	Test Type	Tested Table
22	BOVA ID Bulk Update	Fish Data
1	BOVA ID Invalid	Fish Data
3	Failed Fish Data Scenarios	Fish Data
2	Length Invalid	Fish Data
3	Collection ID Format Invalid	Location Data
1	Collection ID Format T Invalid	Location Data
1	DD Lat Invalid	Location Data
1	DD Long Invalid	Location Data
1	DMS Lat Invalid	Location Data
1	DMS Long Invalid	Location Data
1	DO Invalid	Location Data
1	Pool % Invalid	Location Data
2	Quant/Qual Invalid	Location Data
2	Reach Code Invalid	Location Data
1	Reach Code Mismatch	Location Data
1	Water Temp Invalid	Location Data

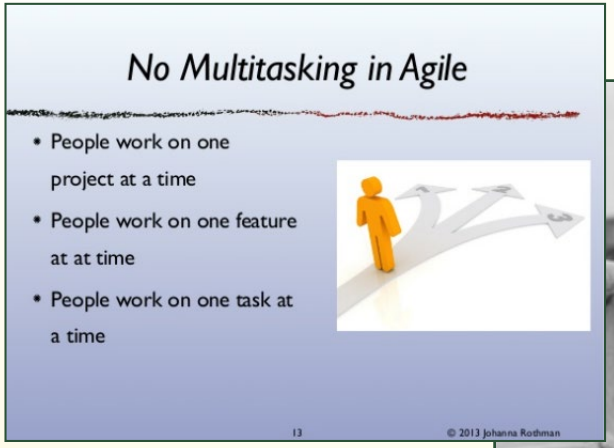
Not In Database

Total Count: 13

Count	Test Type	Tested Table
12	BOVA ID Bulk Update	Fish Data
1	Collection ID Format Invalid	Location Data

Retrospective

Multitasking: "An Illusion of Progress?"



Scrum Way

- Multitasking makes you stupid
- Do it right the first time
- Time is Limited –Treat it that way
- Throw away business cards
- Everyone knows everything
- One meeting a day
- Half Done is not done
- Working hard makes only more work
- Do not be unreasonable
- Strive for flow

FAKE SCRUM - MULTITASKING: GIVING THE TEAM A BUILT-IN EXCUSE TO FAIL

TEAM

Retrospective

Collaboration and Teamwork

Agile Mindset for People

- Collaboration
- Trust & Transparency
- Flexibility in roles and responsibilities
- Responding to change

Acknowledgements

CWSS Scrum Team

CWSS Biologists

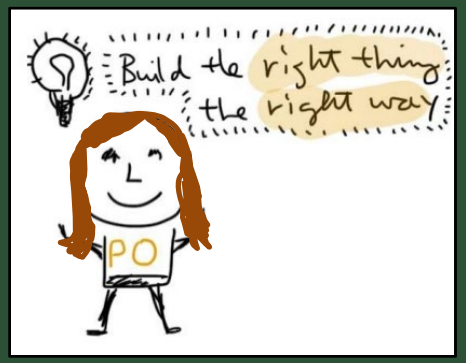
Amy Ewing

Mike Bednarski

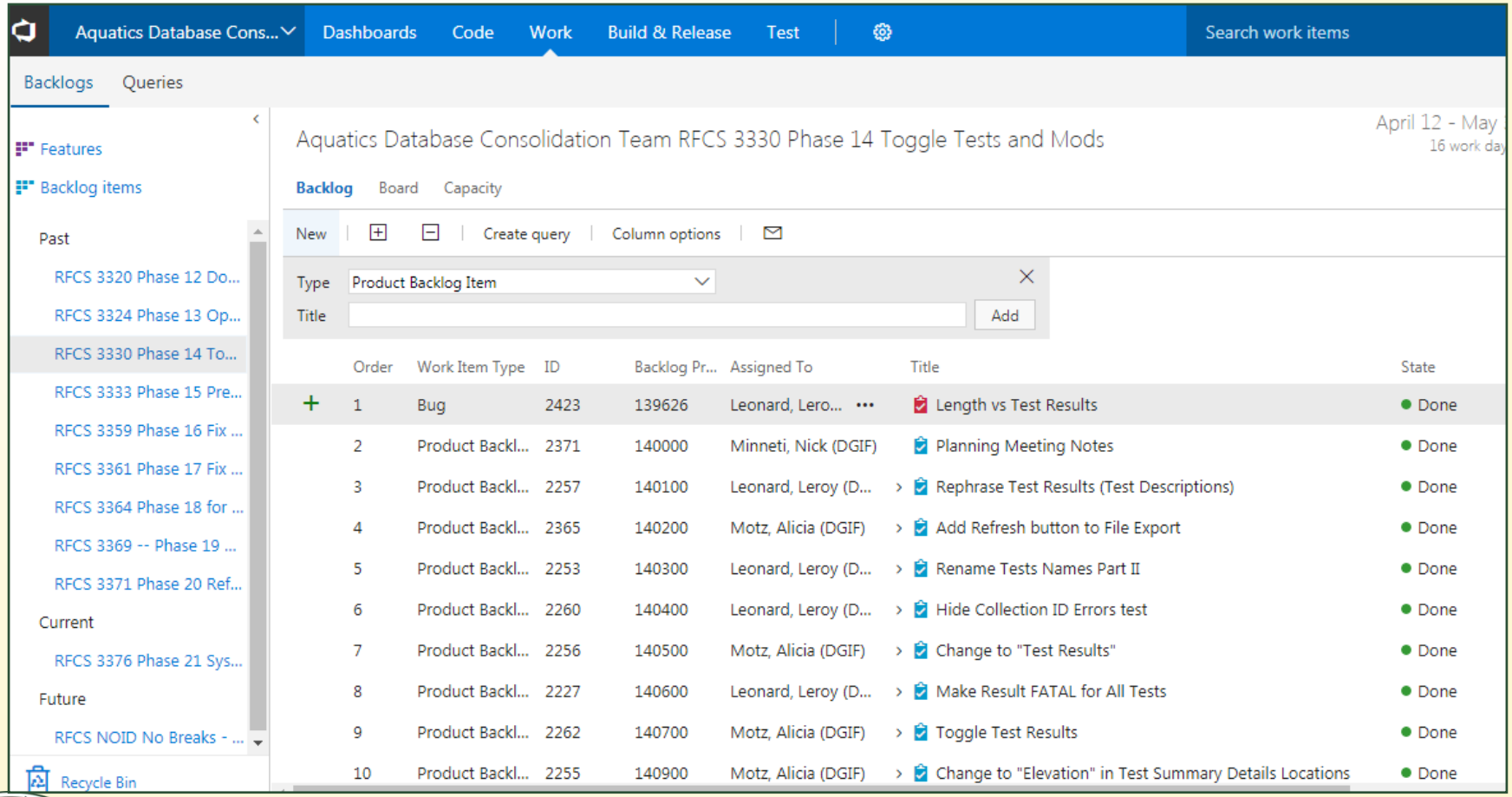
DGIF Aquatics Division

CDI Oversite Committee

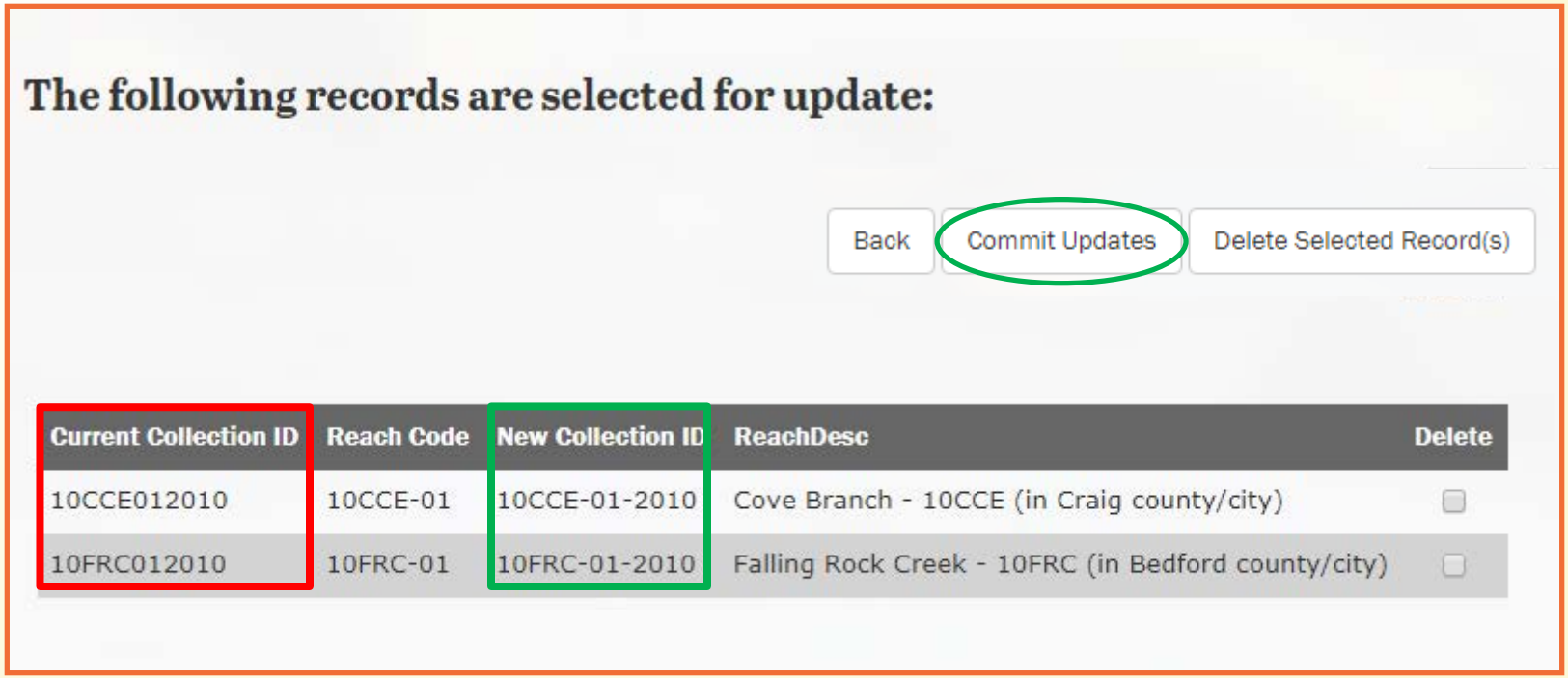
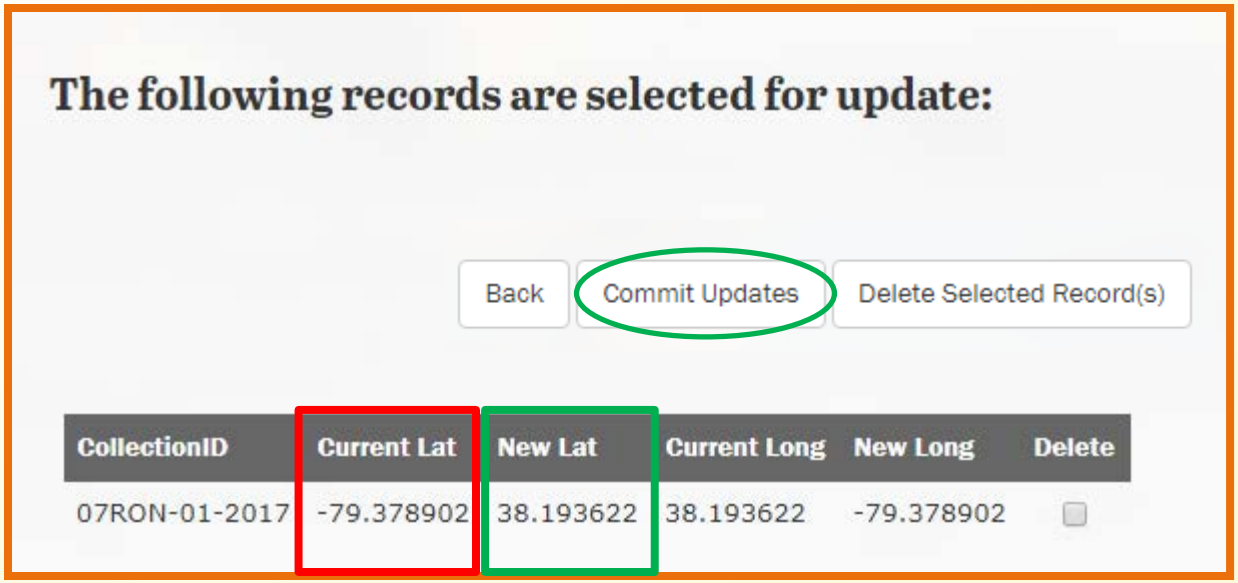
OFWIM Collaborators



Agile Software Development using SCRUM



Codes


The following records are selected for update:
Back
Commit Updates
Delete Selected Record(s)
Current Collection ID
Reach Code
New Collection ID
ReachDesc
Delete
10CCE012010
10CCE-01
10CCE-01-2010
Cove Branch - 10CCE (in Craig county/city)
10FRC012010
10FRC-01
10FRC-01-2010
Falling Rock Creek - 10FRC (in Bedford county/city)

Coordinates