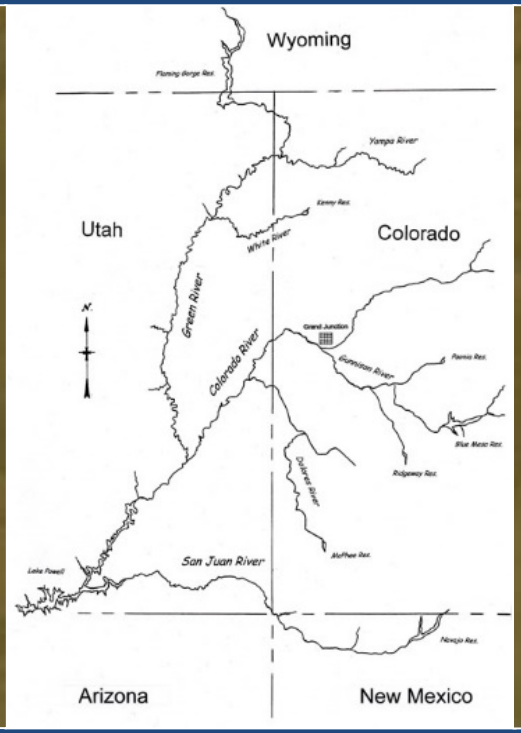
STReaMS: Species Tagging, Research & Monitoring System

Monthly Webinar
6/16/2016



Amy Greenwell & Puja Gurung
Colorado Natural Heritage Program
Warner College of Natural Resources
Colorado State University
www.cnhp.colostate.edu

Ben Johnke
Research Software Facility
Colorado State University

Colorado Natural Heritage Program

Warner College of Natural Resources
Colorado State University

Agenda

- Announcements
- Review of New Features
- Summary of Batch Import Process
- PIT Tag Lots vs PIT Tags
- Storing data for Stocking Events
- Assigning Sample Numbers

Data Developers

CNHP has hired a new Database and Web Developer, Puja Gurung. She will be joining the STReaMS team.

Upcoming tasks:

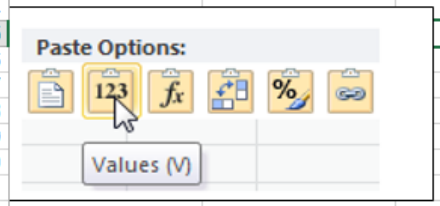
- Reconciling Data – Ben Johnke
- Analyzing Data – Puja Gurung
- Other Tasks – David Hu
  - New role for public access
  - Enhancements to user interface
  - Bug fixes, etc.

Batch Uploads – Overview of Process

- Datasheets get filled out in the template
- Data Leads use Batch Upload Tool to upload data
  - Data Leads ensure Study is in STReaMS
- Data in the template are Analyzed for errors
- Data are compared to existing data in STReaMS (Reconciling data)
- Records are imported, flagged, or diverted to a temporary table for review
- Problem records are investigated by a Database Manager

Batch Uploads - Template

	A	B	C
1	**Field Name**	**Format**	**Description**
2	PIT TAG 400	Text	400 khz PIT tag number
3	PIT TAG 134	Text	134 khz PIT tag number
4	REC NUM	Text	A unique, alpha-numeric code used to track the data to the original file. Use the following format: Year, Study Code or Description, and Sequential Number (ex. 2016COPowell00001, 2016COPowell00002, etc.)
5	AGENCY	Text	Agency doing the work: FWS GJ, FWS V, UDWR V, UDWR M, CDOW GJ, LFL, other?
6	PRINCIPAL	Text	Last name of principal biologist responsible for the work
7	FILENAME	Text	The original file that the data was delivered in (i.e. excel spreadsheet)
8	RIVER	Text	River where sampling occurred; Colorado=CO, Green=GR, Yampa=YA, White=WH, see code list for others.
9	RMI	Numeric, 1 decimals	River mile where fish was collected.
10	YEAR	Numeric, 0 decimals	Year fish was collected.
11	DATE TIME	Date Time	Date and time (if known) fish was collected; in the format yyyy-mm-dd hh:mm:ss (if time is not collected just enter yyyy-mm-dd)
12	SAMPLE NUMBER	Text	Sample number associated with capture sample
13	SPECIES	Text	Species collected; Colorado pikeminnow=CS, razorback sucker=RZ, humpback chub=HB, bonytail=BT, see code list for others.
14	LENGTH	Numeric, 1 decimals	Total length of fish (mm).
15	WEIGHT	Numeric, 1 decimals	Weight of fish (g).
16	SEX	Text	Sex of fish: male=M, female=F, indeterminate = I
17	RIPE	Text	Did the fish exude eggs or milt? Yes=Y, No=N.
18	TUBERCLES	Text	Did the fish have tubercles anywhere on its body? Yes=Y, No=N.
19	RAYS	Text	Number of dorsal/anal fin rays (primarily counted on Gila e.g. 10/11 or 9 anal or 10 dorsal)
20	RECAP 400 KHZ	Text	Was this a recapture of a fish with a 400 khz PIT tag? Yes=Y, No=N
21	RECAP 134 KHZ	Text	Was this a recapture of a fish with a 134 khz PIT tag? Yes=Y, No=N
22	NEW TAG	Text	Was a new tag inserted? Yes=Y, No=N
23	DISP	Text	Disposition of the fish, released alive = RA, see code list for others.
24	GEAR	Text	Sampling gear used to collect fish. See code list
25	HABITAT	Text	Habitat where fish was collected. MCBA=main channel backwater, see code list for others.
26	HAB 1	Text	Primary habitat where fish was collected
27	HAB 2	Text	Secondary habitat where fish was collected
28	UTM ZONE	Numeric, 0 decimal	Number of UTM zone (eg, 12 or 13)
29	UTM X	Float	In the format: 699408
30	UTM Y	Float	In the format: 4332467
31	NOTES	Text	Relevent notes
32			
33	*Fields in the datasheet are pre-formatted. If you copy and paste data into the template, you must use Paste Special and select Values to maintain this special formatting		
34			
35			
36			
37			
38			
39			
40			
41			
42			



Latest templates live on STReaMS
vMM-DD-YYYY

Batch Uploads – Analyzing Data File

	A	B
1	**Field Name**	**Format**
2	PIT TAG 400	Text
3	PIT TAG 134	Text
4	REC NUM	Text
5	AGENCY	Text
6	PRINCIPAL	Text
7	FILENAME	Text
8	RIVER	Text
9	RMI	Numeric, 1 decimals
10	YEAR	Numeric, 0 decimals
11	DATE TIME	Date Time
12	SAMPLE NUMBER	Text
13	SPECIES	Text
14	LENGTH	Numeric, 1 decimals
15	WEIGHT	Numeric, 1 decimals
16	SEX	Text
17	RIPE	Text
18	TUBERCLES	Text
19	RAYS	Text
20	RECAP 400 KHZ	Text
21	RECAP 134 KHZ	Text
22	NEW TAG	Text
23	DISP	Text
24	GEAR	Text
25	HABITAT	Text
26	HAB 1	Text
27	HAB 2	Text
28	UTM ZONE	Numeric, 0 decimal
29	UTM X	Float
30	UTM Y	Float
31	NOTES	Text
32		
33	*Fields in the datasheet are pre-formatted. If	
34		
35	Paste Options:	
36		
37		
38		
39		
40	Values (V)	
41		
42		

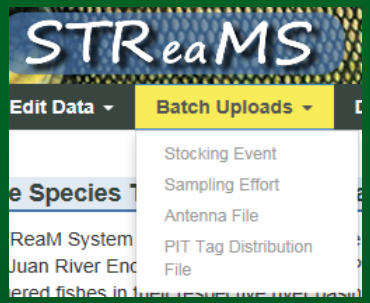
Analyzing data ensures all of the data in your file are valid. This process checks validity of the data **within** the data file.

Each field has a different set of rules. The system will check these rules to determine whether or not an entry is valid

You will receive an error report highlighting data records with invalid entries (ex. tag code with the wrong number of digits, species code that does not exist in the look up table, etc.)

Errors must be fixed before committing data and starting the reconciliation process

This step is active, meaning the user must be logged in and interacting with the website



"Analyze" code changes based on upload type

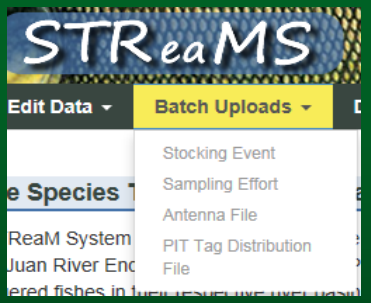
Batch Uploads – Reconciling Data

Once your data are analyzed and error-free, you will commit your file and the system will start reconciling your new data with the existing data in STReaMS

Reconciling data involves two main tasks:
Identifying Insert Logic
Conflict Detection

The system will run a quick scan to make sure the first 20 records are not duplicates (use REC NUM?), which may indicate a file has already been uploaded into STReaMS

This step is passive, meaning once the user commits the data file and approves the prompt to proceed (after the quick duplicate scan) the user can log out of the website if he/she wants to, and will be sent an email when the upload is complete.



"Reconciliation" code is the same, regardless of the upload type

Batch Uploads – Insert Logic

During the reconciliation process, records will fall into one of four categories. The system will search existing data and identify which category a record falls into. This will determine which tables are updated in STReaMS.

Tag does not exist, but Fish does.
Insert a new Tag, assign it to existing Individual, create new Encounter

Tag exists, and is associated with a Fish.
Create new Encounter

Tag and Fish do not exist in STReaMS.
Insert a new Tag, create a new Individual, assign it to the new Tag, create new Encounter

Tag exists, but is not associated with a Fish. (Tag Lots)
Create a new Individual, assign it to the existing Tag, create new Encounter

Batch Uploads – Insert Logic

During the reconciliation process, records will fall into one of four categories. The system will search existing data and identify which category a record falls into. This will determine which tables are updated in STReaMS.

A. Tag and Fish do not exist in STReaMS.
Insert a new Tag, create a new Individual, assign it to the new Tag, create new Encounter

B. Tag exists, but is not associated with a Fish. (Tag Lots)
Create a new Individual, assign it to the existing Tag, create new Encounter

C. Tag exists, and is associated with a Fish.
Create new Encounter

D. Tag does not exist, but Fish does.
Insert a new Tag, assign it to existing Individual, create new Encounter

Batch Uploads – Conflict Detection

During the reconciliation process, new records will be compared to existing data to identify conflicts. There are four possible types of conflicts.

Tag Conflicts – *Example: One tag associated with multiple individuals*
Species Conflicts – *Example: One Individual associated with multiple species*
Disposition Conflicts – *Example: An Individual coming back from the dead*
Sex Conflicts – *Example: One individual associated with multiple sexes*

Depending on the severity of the conflict, records with conflicts get imported and flagged OR not imported and diverted to a temporary table

Batch Uploads – Conflict Summary

Not imported, placed in temporary holding area

Species Conflict (not chub): The Individual is currently assigned to a valid species and the new Encounter has a different, but valid species. This record will go into a "temporary holding area" to be reviewed by a Database Manager. There could be other problems causing the mismatch, like a bad tag code, and the Data Manager will have the skills to track down the issue

Tag Conflict 1: 400khz tag and 134 khz tag are already in the system assigned to two different Individuals. This record will go into a "temporary holding area" to be reviewed by a Database Manager. There could be a bad tag code, or other problems causing the mismatch, and the Data Manager will have the skills to track down the issue

Tag Conflict 2: New Tag = Y and the tag code listed in PIT TAG 134 is already in the system associated with an Individual. This record will go into a "temporary holding area" to be reviewed by a Database Manager. There could be a bad tag code, or other problems causing the mismatch, and the Data Manager will have the skills to track down the issue

Disposition Conflict: The Individual is currently presumed dead and the new Encounter has a code indicating the fish is alive (HA, RA, RT or TL code). This record will go into a "temporary holding area" to be reviewed by a Database Manager. There could be other problems causing the mismatch, like a bad tag code, and the Data Manager will have the skills to track down the issue

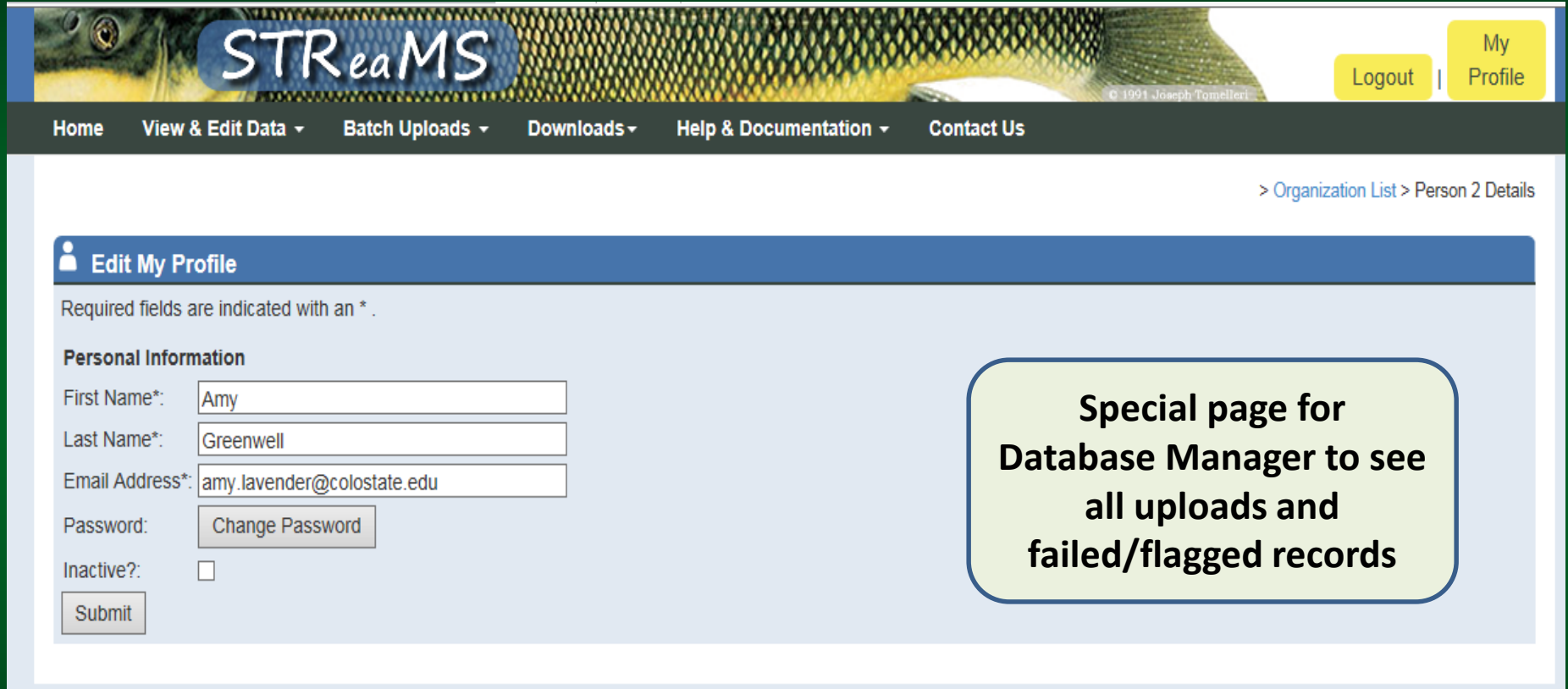
Imported, but flagged

Sex Conflict: The Individual is currently assigned to a valid Sex and the new Encounter has a different, but valid Sex (Ex. Species is Female, new encounter says it is Male)

No changes will be made to the Sex of the Individual. This record will get flagged for review by a Database Manager.

Batch Uploads – Tracking Your Uploads

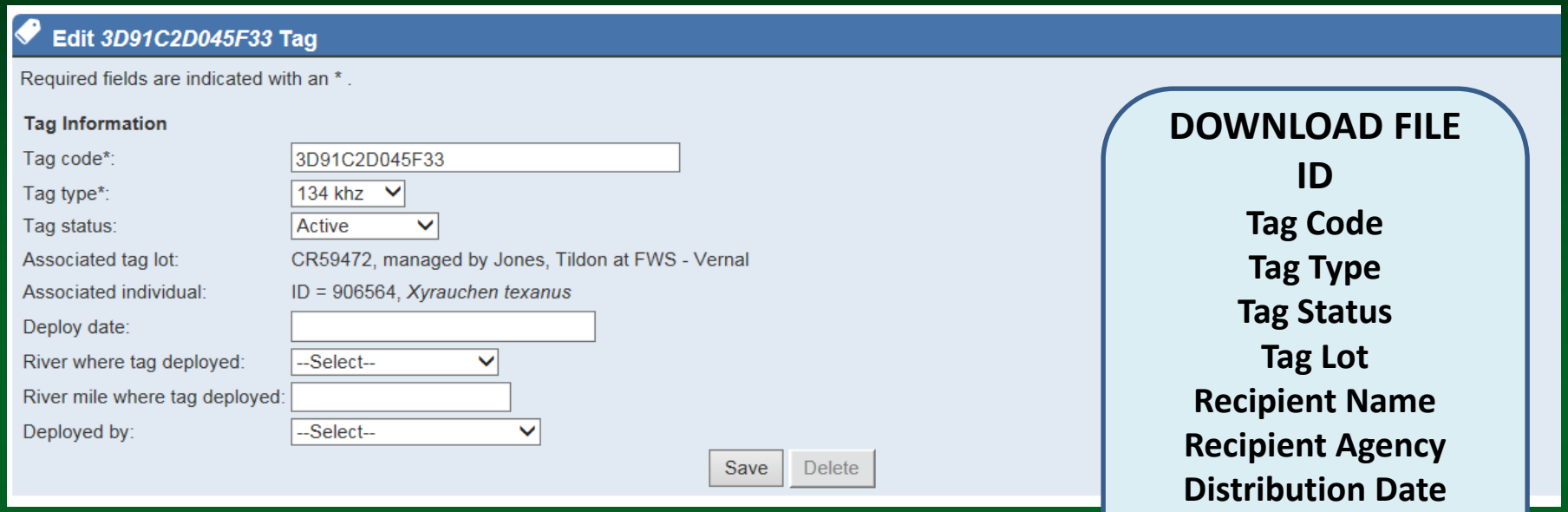
- Keep track of your Uploads in My Profile



My Uploads

- Your raw data file, date uploaded, Hyperlink to Study
  - -View Details
    - -Stats: # records imported, flagged or diverted to temp table
    - -List of Encounters imported from this file

PIT Tag Downloads

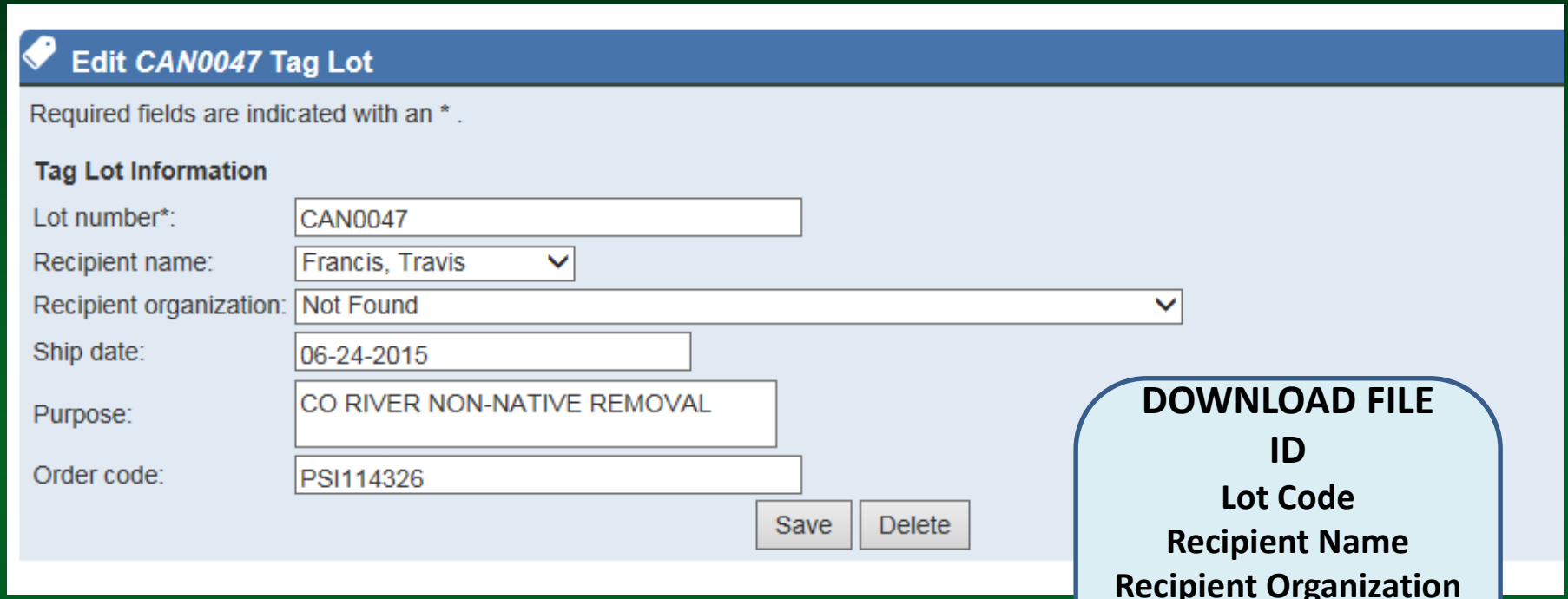


DOWNLOAD FILE

ID
Tag Code
Tag Type
Tag Status
Tag Lot
Recipient Name
Recipient Agency
Distribution Date
Purpose
Order Code
Tag Codes
Deploy Date
Deploy River
Deploy RMI
Individual ID?

Tag download files will give the user a complete set of data for the Tag in question.

PIT Tag Lot Downloads



Tag Lot download files will have a record for each tag in the lot (each tag lot will have 100 records, one for each unique tag in the lot). This gives the user a complete set of data for the Tag Lot(s) in question.

DOWNLOAD FILE

ID
Lot Code
Recipient Name
Recipient Organization
Ship Date
Purpose
Order Code
Tag Codes
Tag Status? (whether it has been deployed or not)

Studies and Stocking Events

	Field	Format	Description
2	PIT TAG 134	Text	individual PIT tag number (134 kHz)
3	REC NUM	Text	A unique, alpha-numeric code used to track the data to the original file. Use the following format: Year, Study Code or Description, and Sequential Number (ex. 2016COPowell00001, 2016COPowell00002, etc.)
4	SPECIES	Text	razorback sucker (RZ), Colorado pikeminnow (CS), bonytail (BT)
5	STOCK YEAR	Numeric, 0 decimal	year fish was stocked (e.g. 2005)
6	STOCK DATE	Date	date fish was stocked in the format yyyy-mm-dd (e.g. 2016-05-13)
7	STOCK RIVER	Text	Colorado (CO), Gunnison (GU), Green (GR), Yampa (YA), Green River Wetland (GW), see code list for others.
8	STOCK RMI	Numeric, 1 decimal	river mile where fish was stocked, e.g. 132.5
9	LENGTH	Numeric, 0 decimal	Total length of fish (mm).
10	WEIGHT	Numeric, 0 decimal	Weight of fish (g).
11	SEX	Text	Sex of fish: male=M, female
12	YEARCLASS	Numeric, 0 decimal	year fish was spawned (if kn
13	RIVER TEMP	Numeric, 1 decimal	temperature of river, report i
14	PH RIVER	Numeric, 1 decimal	Ph of river
15	TANK TEMP	Numeric, 1 decimal	temperature of tank, report i
16	PH TANK	Numeric, 1 decimal	Ph of tank
17	TEMPERED TIME	Numeric, 0 decimal	report in minutes
18	HARVEST TYPE	Text	A = Active, P = Passive
19	RELEASE TYPE	Text	H = Hard, S = Soft
20	LOT	Text	specific lot of fish (if known)
21	SOURCE	Text	pond name or other specific
22	AGENCY	Text	Agency doing the stocking:
23	FILENAME	Text	The original file that the data
24	COMMENTS	Text	general comment about spe
25			
26			
27			
28	*Fields in the data sheet are pre-formatted. If you copy and paste data into		

Paste Options:

Values (V)

stocking metadata | stocking

READY

Edit 3 Species Surveys

2013 Study

Personnel

Samples

Sampling Events

Encounters

Add a user: --Select-- Role: Member Add Person

Show 10 entries Search:

ID	Name	Email Address	Role	Options
38	Fiorelli, Mike	mfiorelli@utah.gov	Unknown	Remove Contact

Showing 1 to 1 of 1 entries Previous 1 Next

Samples

New

Show 10 entries Search:

ID	Sample Number	Start Date/Time	End Date/Time	Gear Type	River	Options
35	OGHR TEST				White River	View Details Delete
24	SASFASFASSASAF				White River	View Details Delete
36	sdfsdfdsfdsfsdfdsfdsf			150'x4', 1/2-2□1/2" Mesh	Green River	View Details Delete

Showing 1 to 3 of 3 entries Previous 1 Next

Sampling Events

New

Show 10 entries Search:

	Description	Gear Type	Start Date	End Date	Options
s 2013					View Details Delete

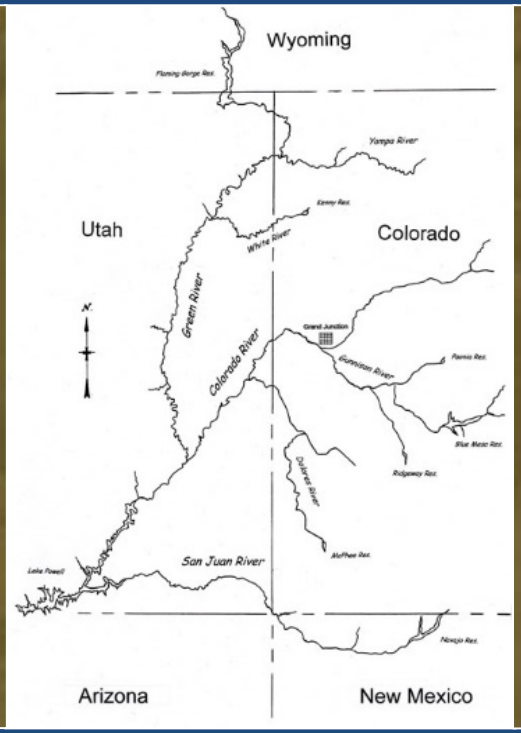
Previous 1 Next

Stocking Events don't "fit" in Samples. Should we change Sampling Events to Stocking Events?

Sample Numbers

Do we want to standardize Sample Numbers?

Questions? Comments?



Amy Greenwell & Puja Gurung
Colorado Natural Heritage Program
Warner College of Natural Resources
Colorado State University
www.cnhp.colostate.edu

Ben Johnke
Research Software Facility
Colorado State University

Colorado Natural Heritage Program

Warner College of Natural Resources
Colorado State University