LobTag2

An R Based Approach to Tagging Data Management

Installation

Enter these lines into the R console to install the package:

```
> install.packages("devtools", dependencies = T)
> devtools::install_github("yutannihilation/ggsflabel")
> devtools::install_github("LobsterScience/LobTag2")
```

You may need to update some of your packages while installing if you have older versions of dependencies. After installation, load the package:

```
> library(LobTag2)
```

On loading the package, the following help message will print showing the location of this user guide and the template xlsx files you will use to load your data:

```
> library(LobTag2)
Welcome to LobTag2! To get started, your User Guide and data entry templates can be found in: C:/LOBTAG/extdata
Good luck with your tagging project!
```

Read this user guide, then begin loading your data!

Package Description

This package is intended as a tool for researchers studying aquatic animal movement using public reporting of tagged animals. The package runs in R and allows the user to call functions which facilitate intuitive inputting and organization of release and recapture data into an SQL based database (local or Oracle), as well as generating plausible paths of animal movement, with the end products being data tables, user generated movement maps, and other user generated summary maps. Fundamentally, the package provides a standardized method for maintaining data quality in tagging projects and allowing easy access to these data for further investigations.

R

• This package was written for R version 4.4.1. It is assumed that the user has this or a later version of R installed.

Using Locally on Hard Drive (db = "local")

For local operation, the package uses RSQLite to create and write an SQL based database file on the user's local hard drive (no external network connections required). This file is stored at C:/LOBTAG/LOBTAG.db

Local operation should be specified when running functions in the app, for example, when using the upload_releases() function locally, the user will enter

```
> upload_releases(db = "local")
```

Deleting the LOBTAG.db file will delete your entire database, so it is advised that you backup this file after building your database.

Data tables in LOBTAG.db can be accessed, queried and manipulated using SQL code much like any SQL based database (such as Oracle).

Pro tip: If you're working in R and want to interact with your database in ways outside the functionality of this package, RSQLite is a valuable R package for this.

Using with Oracle

This package has the option to be used to build and interact with Oracle databases. To use with Oracle, the user must enter their Oracle credentials as arguments when calling functions. To use Oracle with any of the functions in this package, enter the following when calling the function:

- db = "Oracle"
- oracle.user = "username"
- oracle.password = "password"
- oracle.dbname = "database/server name"
- > upload_releases(db="Oracle", oracle.user = "madeupuser", oracle.password = "madeuppassword",
 oracle.dbname = "madeupdatabasename")

Pro tip: If you're regularly working with Oracle, saving your credentials as the following variables in your R environment will prevent you having to enter them each time you call a function:

```
> oracle.lobtag.user = "your user name"
> oracle.lobtag.password = "your password"
> oracle.lobtag.server = "your server name"
```

Included data files

The following required files are installed with the package and are stored in the LOBTAG folder (files in green should be opened by the user while using the package normally, files in red are used automatically by the package and should only be opened/manipulated by experienced users):

data:

- gebco_2024.tif
- knit_rewards.Rmd
- new_land_coords.rds
- user guide.pdf

extdata:

- releases_template.xlsx
- recaptures_template.xlsx

Mapping Token

The package's mapping functionality uses Mapbox to generate map tiles. To use this functionality, you will need to create an account with Mapbox in order to get a mapping token. This is quick and easy, just go to:

https://account.mapbox.com/auth/signup/

and create an account. Once you log into your account, look under "Access tokens" and you should see your "Default public token" which is a long string of characters. Will need to enter this value later when using the generate maps() function.

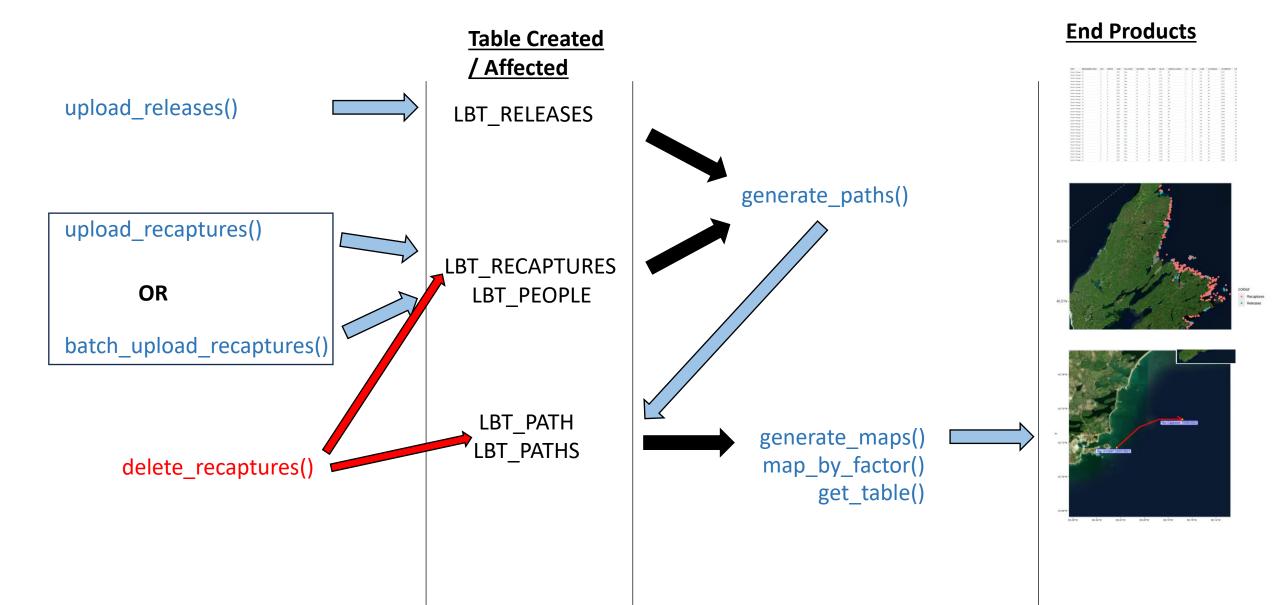
Functions (Overview)

The package is made up of the following main functions used to create and manipulate the tagging database, as well as generate maps from tag release and recapture data entered by the user:

```
upload_releases()
upload_recaptures()
batch upload recaptures()
generate paths()
generate maps()
delete recaptures()
get table()
map by factor()
```

tags2kmz()

Functions (Overview)



Formatting Releases Data:

The first phase of most tagging projects will be releasing the tagged animals. At bare minimum, this will produce data for **location** and **time of release** for each tag. This data will need to be formatted before using with the package.

Use the template file provided to format your release data into an xlsx spreadsheet for uploading with the package. These template files are found in C:\LOBTAG\extdata or can also be found by entering:

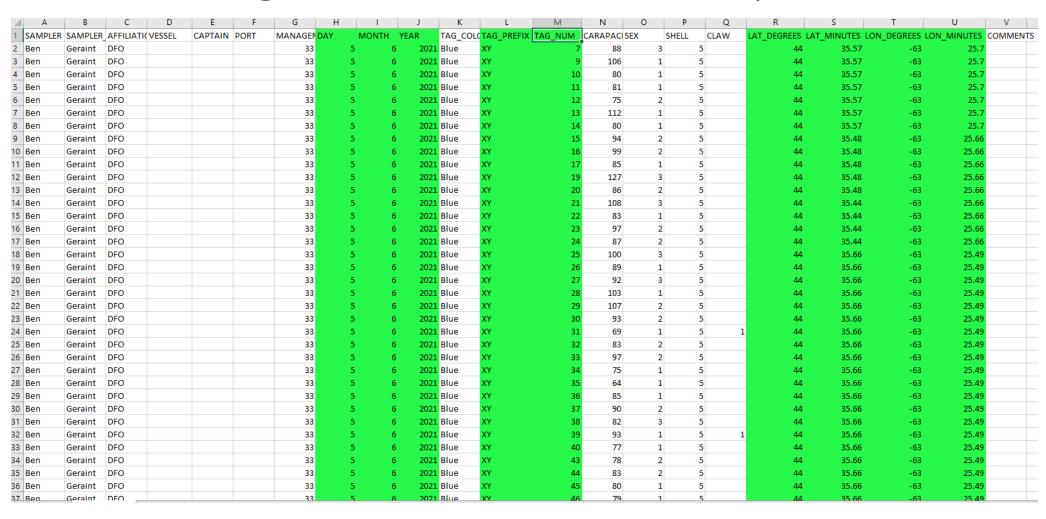
> system.file("extdata", package = "LobTag2")
Which will return a path to the folder with your template files.

Enter your releases data into the columns provided in the releases_template.xlsx table. The following columns are mandatory (must contain data for the upload to work):

DAY, MONTH, YEAR, TAG_PREFIX, TAG_NUM, LAT_DEGREES, LAT_MINUTES, LON_DEGREES, LON_MINUTES

The remaining columns are optional and originate from the lobster tagging program for which the package was originally written.

Formatting Release Data: Example



Green = mandatory columns

Formatting Release Data: Mandatory Columns TAG_PREFIX

Most tagging programs use a Prefix-Number system to identify unique tags. This package assumes that tag numbers have a prefix to separate them from other tagging programs. This is usually a simple pair of characters (for example, "XY" such as in tag ID "XY1234") but can be any character combination entered by the user. Even if your tagging program does not use a prefix, one **must** be entered when uploading data with this package. This ensures that the database can hold multiple tagging programs without duplicate tag numbers causing issues.

Formatting Release Data: Mandatory Columns TAG_NUM

The tag number following the tag prefix is entered as a simple number in its own column called "TAG_NUM".

Formatting Release Data: Mandatory Columns Latitude and Longitude

All release coordinates are uploaded in the format Degrees Decimal Minutes (DDM), split into 4 columns: LAT_DEGREES, LAT_MINUTES, LON_DEGREES, LON_MINUTES. Columns ending in "DEGREES" receive the degree value, and those ending in "MINUTES" receive the MINUTES value as a decimal. For cases when Longitude is west (for example, 58°30.50 W), the proper "-" **must** be included in front of the degree value to denote this, otherwise the package will not know that the

longitudes are westerly, for example:

R	S	Т	U	V
LAT_DEGREES	LAT_MINUTES	LON_DEGREES	LON_MINUTES	COMM
44	35.57	-63	25.7	
44	35.57	-63	25.7	
44	35.57	-63	25.7	
44	35.57	-63	25.7	
44	35.57	-63	25.7	
44	35.57	-63	25.7	

Release Uploading: Calling the function upload_releases()

Begin the process of uploading release data by calling function upload_releases() which will then prompt you to upload an excel file with the data. By now you should have formatted your release data with the columns found in the provided template file (releases_template.xlsx).

Example:

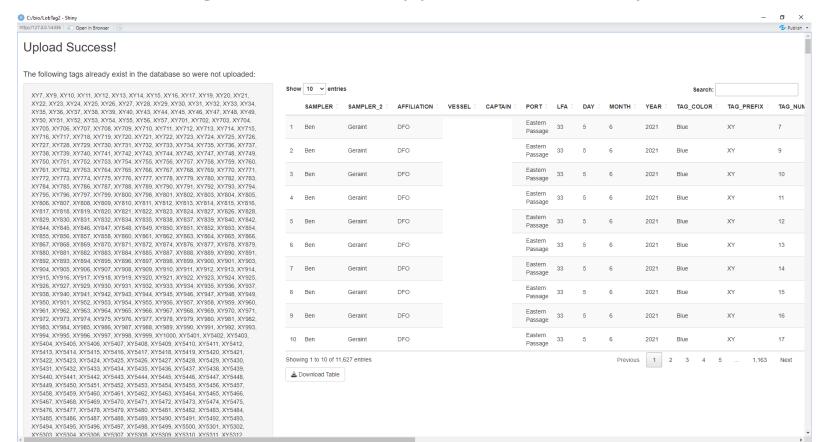
```
> upload_releases(db = "local")
```

Release Uploading: Error Checking and Upload Completion

When the user has run upload releases() and selected a data file, the function will perform some general error checking. If there are crucial errors found in the mandatory columns, a dialogue box will appear alerting the user to the location of errors and instructing them to fix these before attempting the upload again. If there are no errors, the user will receive a dialogue alerting them that the upload has completed without errors. This means that the database table LBT RELEASES has been created and now contains the release data. If there are issues with the data that are concerning but don't prevent uploading, a different dialogue box will appear, warning of these issues and asking the user if they want to proceed with the upload.

Release Uploading: Existing Tags

The function also checks if any of the tags being uploaded are the same as tags already existing in the database, defined as having both identical **prefix and number**. Existing tags will not be uploaded again. If the upload file contains existing tags, the upload will still proceed, but the user will receive a dialogue box alerting them of the tags that were skipped because they were found in LBT_RELEASES.



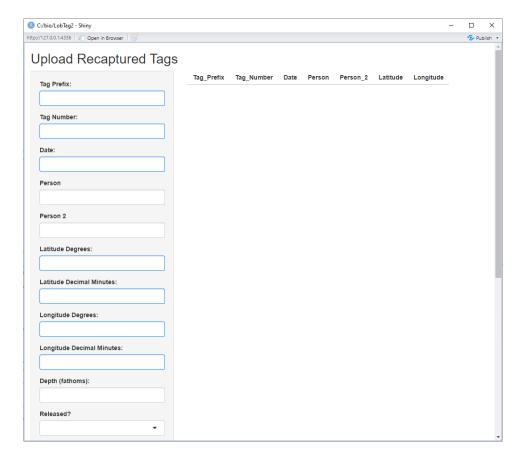
Individual Recapture Uploading:

upload_recaptures()

Generally, once tags are released, recapture reports will then come in individually over time. For public programs, these data may be highly variable in format and quality. Running the function upload_recaptures() provides the user with a data entry window which standardizes the entry of these data into the database. Mandatory fields are highlighted in blue.

Example:

> upload_recaptures(db = "local")



Individual Recapture Uploading: Tag Checking

The user interface for uploading recaptures communicates with the database in real-time, allowing it to check if the tag's release data exists in LBT_RELEASES. If the Prefix and Number combination are not found, the user will be alerted:

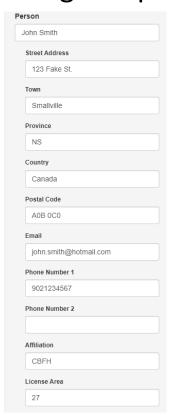
Upload Recaptured Tags

BB		
Tag Numbe	er:	
76		

This is just a warning, and the user may still choose to upload this tag. However, if release data are not eventually entered for this tag, then path generation will not be possible.

Individual Recapture Uploading: Person Info

Some tagging programs may wish to collect contact information for the person reporting tag recaptures. For this purpose, the database includes the LBT_PEOPLE table, which is created/updated when a name is entered in the "PERSON" field during recapture uploading. When this happens, additional fields will appear



prompting the user for more contact information. Since the user interface communicates with the database, these will auto-fill with any existing information for the entered name so this does not need to be re-entered each time the person reports a recapture.

Individual Recapture Uploading: Latitude and Longitude

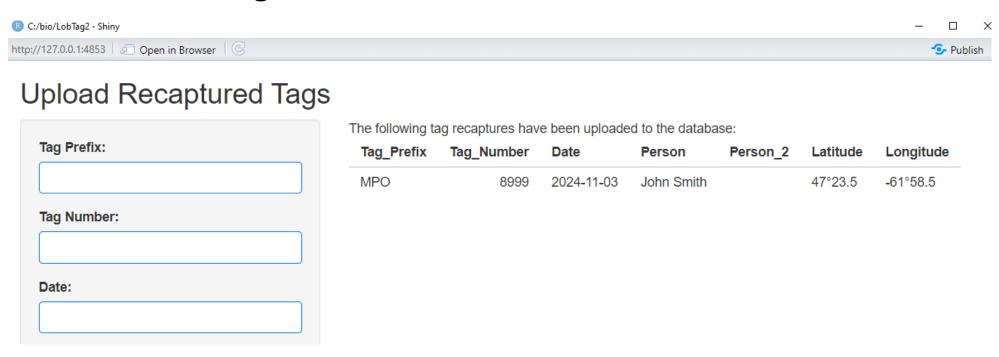
Latitude Degrees:	
46	
Latitude Decimal Minutes:	
26.98	
Longitude Degrees:	
60	
Careful! Degrees Longitude must be negative for western hemisphere! Longitude Decimal Minutes:	

The user interface largely standardizes the input of coordinate values, making it impossible to enter a "non-coordinate", however, the user must be aware and check that:

- Coordinates are being entered in **DDM** format
- Westerly values for "Longitude Degrees" are negative. The scroller for this value goes down for negative values. This is the only way for the functions to know that coordinates are in the western hemisphere. Since the package is likely to be used in the western hemisphere, a warning message will appear if the user enters a positive value here.

Individual Recapture Uploading: Submitting Recaptures

When the user clicks "Submit" in the recapture uploading window, the information for the submitted tag appears listed in the upper right, so the user can see important information for all the tags that they have submitted during the session.



Batch Recapture Uploading: Formatting

If the tagging program already has many recapture data stored in tables, these can be batch uploaded to the database much like the releases data. Like the releases data, these will first need to be formatted using the provided template (recaptures_template.xlsx). For recapture formatting, the mandatory columns are:

TAG_PREFIX, TAG_NUMER, DAY, MONTH, YEAR,

And either: LAT_DEGREE, LAT_MINUTE, LON_DEGREE, LON_MINUTE

Or: LAT_DD, LON_DD (these last two columns allow the user the option to enter coordinates in decimal degree (DD) format if this is the format your coordinates are in. Uploading will work as long one format of coordinate is fully entered).

Batch Recapture Uploading: Formatting Example

	С	D	E	F	G	Н	1	J	K	L	М	N	0	Р	Q	R		S	Т		U	V	W	X
1 D.	AY I	MONTH	YEAR	PERSON	CIVIC	TOWN	PROV	COUNTR	Y POST	EMAIL	PHO1	PHO2	AFFILIATION 1	LICENSE_	PERSON_	LAT_DEGRE	E LAT_M	IINUTE L	ON_DEGR	EE LON	_MINUTE	LAT_DD	LON_DD	FATHOMS REL
2	21	4	2023	3) l	i je	sh F	C NS	Canada	B0J 1N0	L	ic S	3563		LFA32	NA	4	4	33.55		63	15	44.5577	-63.2037	0
3	2	6	2023	<mark>3</mark> l n	Į P	la V	a NS	Canada	B0H 1T0	C	er	il.com	GCIFA	LFA31A	NA							45.2162	-61.1658	0
4	19	6	2023	3 ()	ι										NA							46.32	-60.2455	6.1
5	3	7	2023	3)	ι										NA							46	-60.2308	8.1
6	5	7	2023	3 ()	ι										NA							46.3065	-60.2612	2.6
7	30	6	2023	<mark>3</mark> I N	/										NA							46.079	-59.8428	0
8	6	7	2023	<mark>3</mark> (N	/										NA							46.1047	-59.8182	0
9	15	6	2023	3 I r											NA							46.1325	-59.8113	9.6
10	15	6	2023	<mark>3</mark> 31	r										NA							46.0833	-59.85	0
11	27	6	2023	<mark>3</mark> 31	r										NA							46.0833	-59.85	0
12	20	6	2023	3 3	C										NA							46.0082	-59.8888	0
13	17	6	2023	3 : a	C SI	r									NA							46.252	-60.1647	0
14	17	6	2023	3 : a	c Si	r									NA							46.252	60.1647	0
15	15	6	2023	3 B	}										NA							46.2035	59.7417	8
16	16	6	2023	3 B	}										NA							46.1863	-59.7652	10
17	20	6	2023	3 II	r										NA							46.2067	-59.9167	0
18	20	6	2023	3 II	r										NA							46.2067	-59.9167	0
19	25	6	2023	<mark>. l</mark> a	Ē										NA							46.1917	-59.8517	3
20	27	6	2023	<u>:</u>	\ sh	1									NA							45.7443	-60.21	0
21	25	6	2023	<mark>3</mark> ()	8										NA							46.015	-59.8433	2
22	27	6	2023	<mark>3</mark> ()	8										NA							46.0083	-59.8767	3
23	3	7	2023	<mark>3</mark> . a	r tt	t									NA							46.2572	-60.069	2
24	3	7	2023	3 , e	E										NA							46.3438	-60.2988	7.1
25	19	6	2023	Auaiii əili	Bipe										NA							46.0082	-59.6995	0

Green = Mandatory

Blue = Coordinate format option 1

Purple = Coordinate format option 2

Batch Recapture Uploading: Calling the function batch_upload_recaptures()

Begin the process of batch uploading recapture data by calling function batch_upload_releases() which will then prompt you to upload an xlsx file with the data. By now you should have formatted your recapture data with the columns found in the provided template file (recaptures_template.xlsx).

Example:

> batch_upload_recaptures(db = "local")

Error checking and warnings: when the file is chosen, error checking will proceed much like with the releases upload. If crucial errors are found, the user will be told to fix these and start over (run batch_upload_recaptures() again). If non-crucial issues are found, the user will be warned and asked if they want to ignore these and proceed with the upload.

Path Generation: generate_paths()

Once the user has completed uploads of releases and at least one recapture, they can begin generating movement paths between these. Running the function <code>generate_paths()</code> will calculate plausible paths for all recaptures found in LBT_RECAPTURES. This function can be run as many times as needed whenever new recaptures are added; recaptures with existing paths will simply be skipped. Pathing uses a "least cost" method while avoiding land to calculate plausible paths of movement between tag release and each sequential (chronologically) recapture event.

Example:

- > generate_paths(db = "local")
- You can also generate (or regenerate) paths for specific tags only:
- > generate_paths(db = "local", tags = c("XY123","XY124"), regen.paths = T)
 (without the regen.paths = T argument, tags that already have paths won't be re-pathed).

Note: If a new recapture is added from an earlier date than the most recent recapture for that tag, running generate_paths() will simply delete and regenerate an entirely new path for that tag. This means you can always upload new recaptures from any point in time and still regenerate paths to get your most up-to-date paths.

Path Generation: LBT_PATH & LBT_PATHS

Pathing information generated by generate_paths() is stored in two tables, LBT_PATH and LBT_PATHS. In LBT_PATH, the DIST column contains lengths in km of the total plausible paths (based on land avoidance) between each real observation of the tag. In this table, the column CID (Capture ID) contains the chronological numbering of each recapture event. LBT_PATHS contains the individual plausible movement path segments for each tag. The POS (Position) column contains the sequence for each coordinate making up the plausible path from the last known real location, up to the next known recapture location. Rows representing the recapture event will contain date and name values in the REC_DATE (Recapture Date) and REC_PERSON (Recapture Person) columns.

Note: These tables are intended for reference by the package when generating maps, but they may also be useful for further analyses. Users wanting to avoid SQL querying of the database have the option of saving all the pathing data in a single useful spreadsheet when generating paths by using the save.table = T argument:

```
> generate_paths(db = "local", save.table = T)
```

This line can be run even after all paths have already been generated; as long as you don't use the regen.paths = T argument, then no paths will be regenerated, but an xlsx file of all existing paths will still be saved (a pop-up window will ask the user where they want to save this file).

Checking/Importing your tables: get_table()

If you're not comfortable accessing your new database tables with SQL language, you can use get_table(), choosing the table you want to view in R Studio.

Options:

```
> get_table("releases")
> get_table("recaptures")
> get_table("path")
> get_table("paths")
> get_table("people")
```

These options will bring the table into your R environment to work with. For example, if you then wanted to save the releases table as a csv file, use write.csv(releases, "yourfilepath/releases.csv")

Map Generation: generate_maps()

Once paths have been created, these can be presented graphically using the generate_maps() function. This function has a number of additional arguments that change the output. Some of these are:

- map.token = Mapping requires a public mapping token for Mapbox.
- **people** = This specifies the recapture person(s) for whom to generate maps. This will cause maps to be generated showing releases and each recapture location (and the plausible connecting path) for each tag ID reported by the chosen person(s). This can be a single name, or a vector of names (Default = **NULL**).
- all.people = (Default = FALSE), if changed to all.people = T, maps for every person who has reported recaptures will be generated.

Pro tip: You can avoid having to enter your map.token each time by saving it in your R environment as mapbox.token:

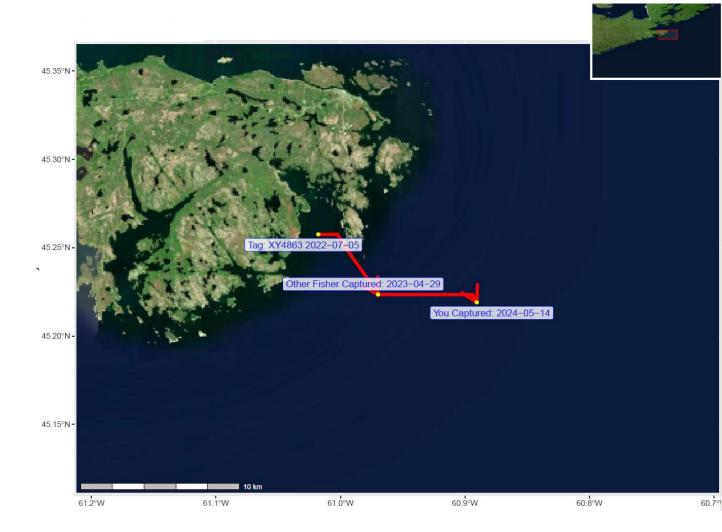
- > mapbox.token = "Pk.dgfAhSgffsgeHGhwhgZd" -- Just enter this once! Then:
- > generate_maps(db = "local", people = c("John Smith", "Jane Doe"))

Map Generation: generate_maps() Additional Options:

- tags = This specifies the tag IDs (prefix and number) to generate maps for. This will cause maps to be generated showing releases and each recapture location (and the plausible connecting path) for each tag ID chosen. This can be a single tag ID, or a vector of IDs (Default = NULL).
- all.tags = (Default = FALSE), if changed to all.tags = T, maps for every tag with recaptures will be generated.
- > generate_maps(db = "local", tags = c("XY123","XY124")))
- > generate_maps(db = "local", all.tags = T)
- map.res = and max.pixels = Can be used to set the resolution of the maps being generated (can be useful when lower resolution images are desired for faster processing) Defaults: map.res = 0.9, max.pixels = 800000.

Maps

- Show the locations of tag releases and all subsequent recaptures for which paths have been generated.
- Red arrows show the direction of movement based on date of recapture.
 Paths are generated using a "least-cost" method while avoiding land (movement patterns are the same as a Queen piece moving across a chessboard, connecting points on the grid with the smallest number of moves).
- Smaller map inset in top right corner shows larger geographic area. The user is given the option to define this area manually when running generate_maps() otherwise, it is sized automatically (turn off this option with inset.option = F or turn off inset maps entirely with inset.map = F)



Deleting Recaptures

Since paths may be generated from multiple recaptures for a single tag, it is important not to directly delete data from LBT_RECAPTURES. If you want to remove a recapture event from the database, you can use the delete_recaptures() function. Running the function will bring up a dialogue box asking the user to search for the tag prefix, number and date of the recapture. If the correct recapture event is found, the user can then click the **Delete** button to remove the recapture. This will delete all information from LBT_RECAPTURES, LBT_PATH and LBT_PATHS for this recapture, and if there are other recaptures for this tag, all paths will be deleted and regenerated.

> delete_recaptures(db = "local")

Tag Prefix XY Tag Number 7833 Date Caught 2023-06-28 Search Delete

Recapture Deletion

XY	
Tag Number	
4729	
Date Caught	
2023-05-26	

There are other recaptures for this tag! Regenerating paths for this tag, please wait ... DO NOT CLOSE THIS WINDOW!

map_by_factor() : A General mapping function for visualizing the tagging data

- This function provides arguments to filter and group for variables within either the LBT_RELEASES or LBT_RECAPTURES.
- Can be useful for general reports and visualizations of tagging efforts / yields.

Arguments:

"db =" (default = NULL)

"filter.from =" REQUIRED. Enter either "releases" or "recaptures" to choose which table your chosen filtering, mapping and factoring variables will come from (example. if you wanted to map by all the different release years you would have to choose filter.from = "releases", but if you wanted to map by recapture year you would choose "recaptures" (default = NULL).

"filter.by =" This is the variable (from your chosen table) in which you're filtering values, so if you chose factor.from = "releases" and filter.by = "YEAR" you will also need to enter a specific year value with "filter.for =" to map only releases for that year (default = NULL).

"filter.for = " Specific values of your filtering variable that you want to map. So if you chose factor.from = "releases" and filter.by = "YEAR" and you only want to show releases from 2022, then filter.for = "2022". (Default = NULL).

"map.by =" This is the variable (from your chosen table) which maps are being produced by, so if you chose factor.from = "releases" and map.by = "YEAR" you will produce a separate map for each value of YEAR found in your releases table (default = NULL).

map_by_factor(): Continued

"factor.by =" This controls the factor colouring within a map. So if you've filtered by YEAR but you want each point to be coloured by the affiliation responsible for the release, factor.by = "AFFILIATION". (Default = NULL).

"all.releases =" (Default = F). The defaults assume the user only wants to see tags that have been recaptured, even if they've selected factor.from = "releases" (often tagging programs release thousands of tags with a small percentage recaptured). If you want to see all the releases associated with your chosen factor, all.releases = T.

"show.releases = " This is different than the previous argument and should only be used if factor.from = "recaptures". Making FALSE will not allow you to see the release locations for the recaptured tags that you've mapped from recaptures (Default = T).

"show.recaptures = " Inversely to the above, making FALSE will not allow you to see the recapture locations of your tags when you're mapping from releases. (Default = T).

"tag.prefix = " If database contains multiple tagging programs, this is a quick additional filtering argument to only show tags with the chosen prefix (Default = NULL).

"add.paths = " Show paths of proposed lobster movement (as in generate maps function). (Default = F).

map_by_factor() : Continued

Additional aesthetic options:

"inset.map =" Include the inset map in top right corner showing larger geographic area (Default = T).

"inset.option = " Give user the option of whether do draw or auto-size boundaries of the inset map (Default = T). If inset.option = F, inset map will auto-size without asking the user.

"max.pixels = " Use to set number of allowed pixels in image. Can be used to reduce file size at expense of resolution (Default = 800000).

"map.res = " Use to set general resolution of map (0-1), dependent on max.pixels. (Default = 0.9).

"zoom.out = " A percent value (1-100), increases mapping area past edges defined by mapped points. (Default =1).

"point.size = " Adjusts the size of all points on the map. (Default = 1.5).

"file.type = " Choose how to save the map file. Options are "pdf" and "png" (Default = "pdf").

tags2kmz(): Creates a csv file of chosen tag table that can be used to create KMZ file in Google Earth (allows quick lookup of tags in GE)

"tag.tab =" Choose table to produced output file for (either "releases" or "recaptures")
"tag.prefix =" (Default = "XY"). Choose tag prefix to identify tagging program.

INSTRUCTIONS FOR USE:

- Use the function to create the output csv in folder of your choosing. Then import the file into Google Earth using file -> import. Click through wizard
- Choose "Import All"
- Choose "Yes" for style template (will adjust in next step). create one or use existing.
- If creating new- choose field "label" as Field Name
- Google Earth may be slow if there are many tags
- all.tag.releases.for.KMZ.csv or all.tag.recaptures.for.KMZ.csv will appear in your "Temporary Places", right click on it and choose "properties"
- Click "Style, Color" tab and click "Share Style"
- Choose label size (recommend Increase label to 1.2 scale, icon to 1.5 scale)
- Hit the icon button on the top right and choose icon you want
- Click OK