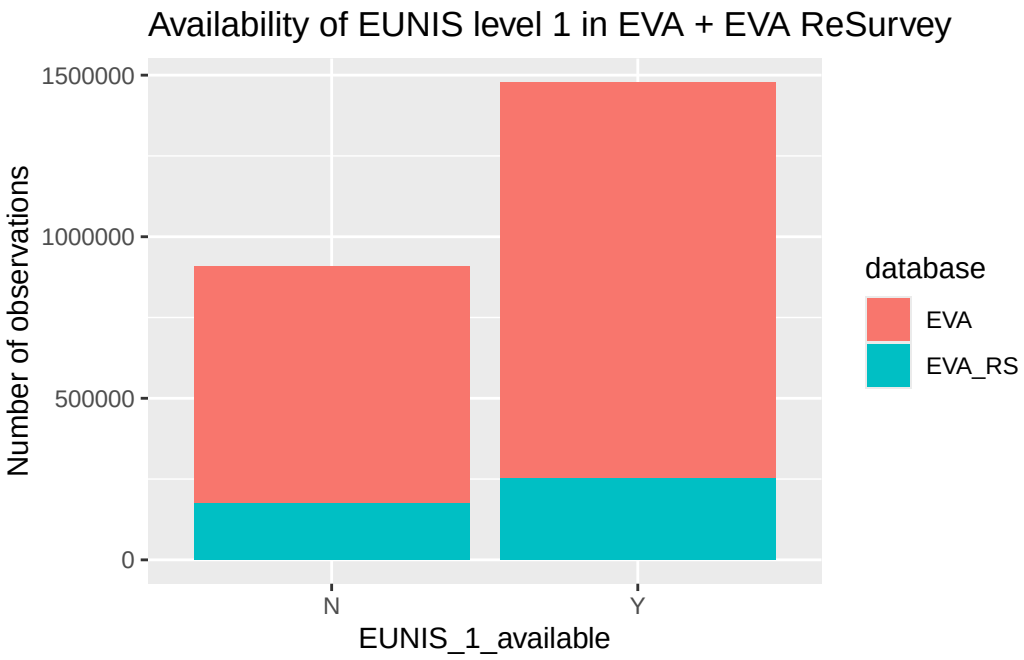
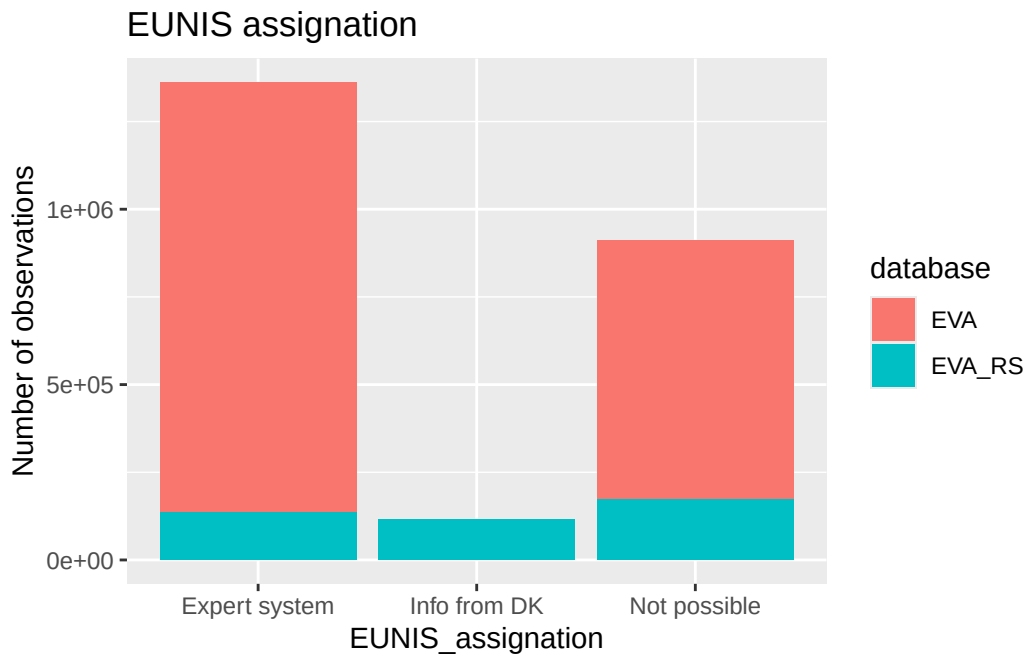
Script to check accuracy and % of missing EUNIS for EVA and EVA Resurvey

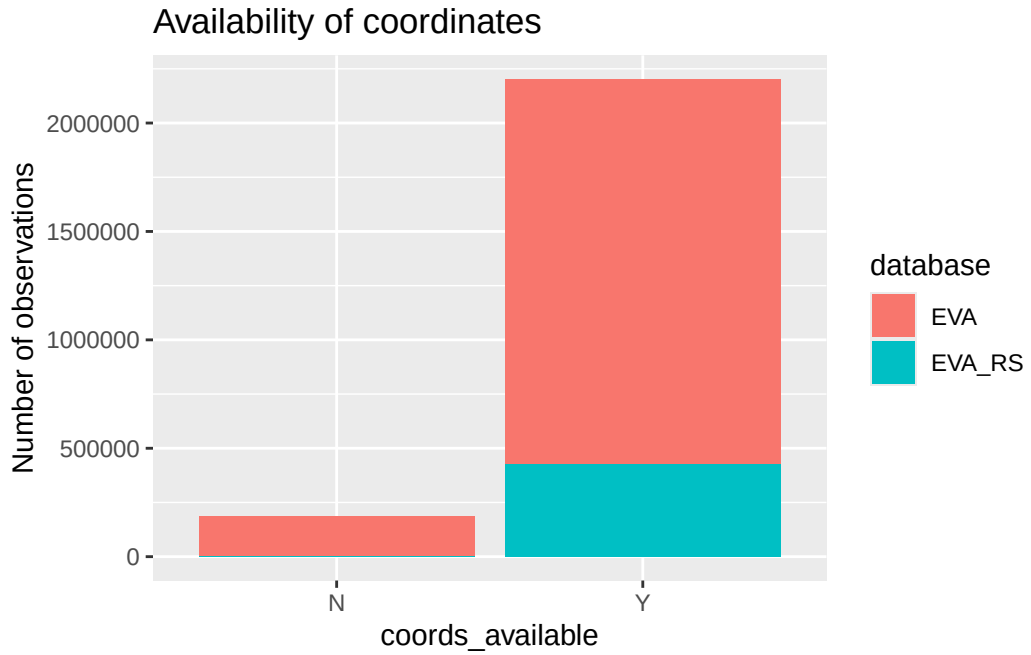
EUNIS level 1 availability



EUNIS assignation

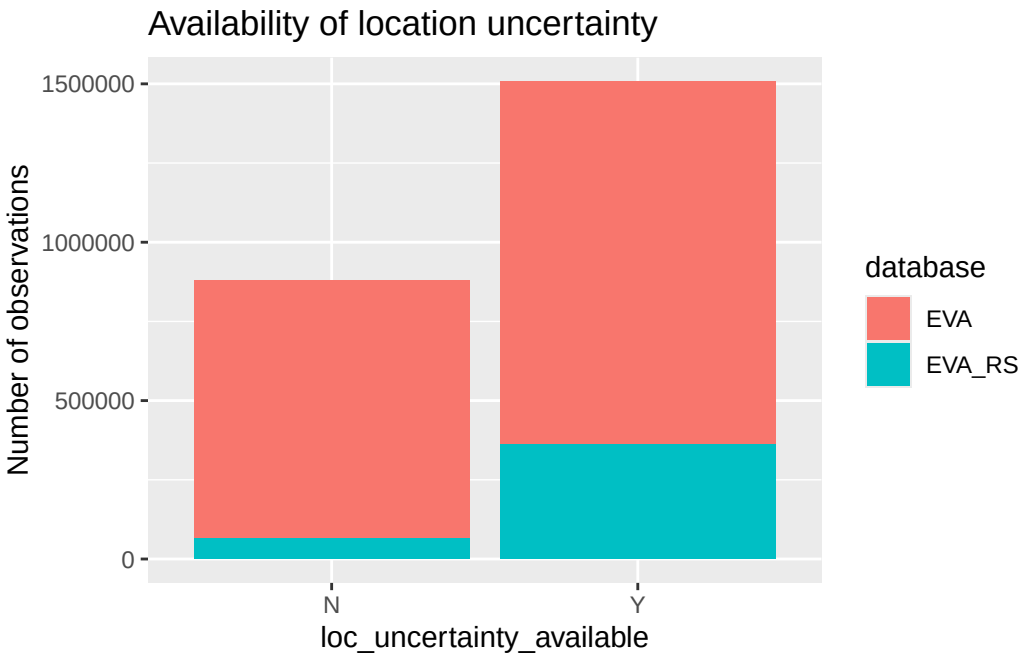


Coordinates available

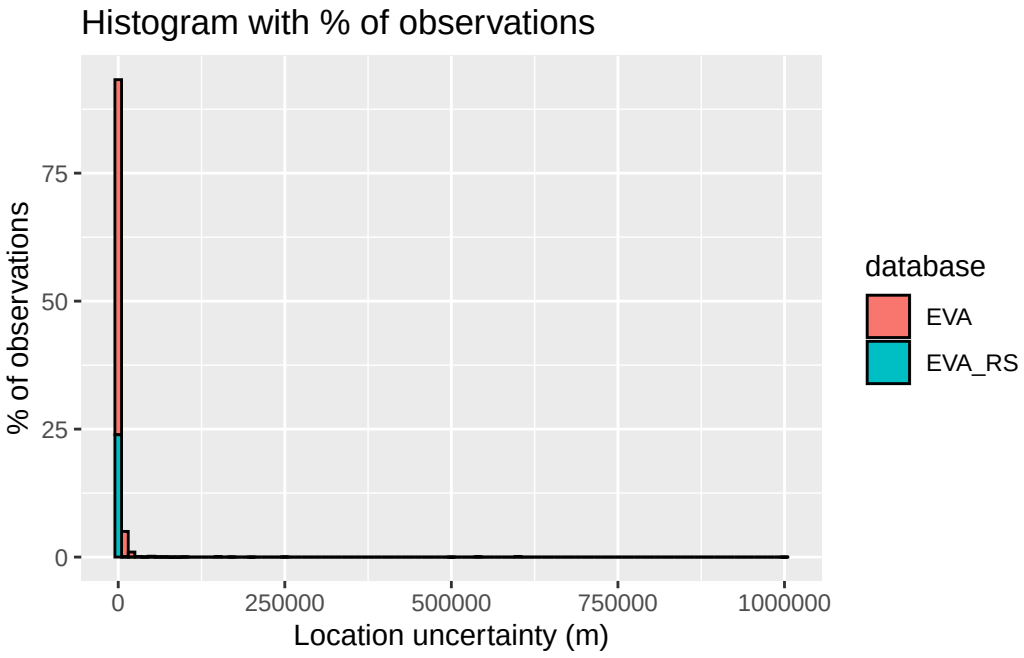


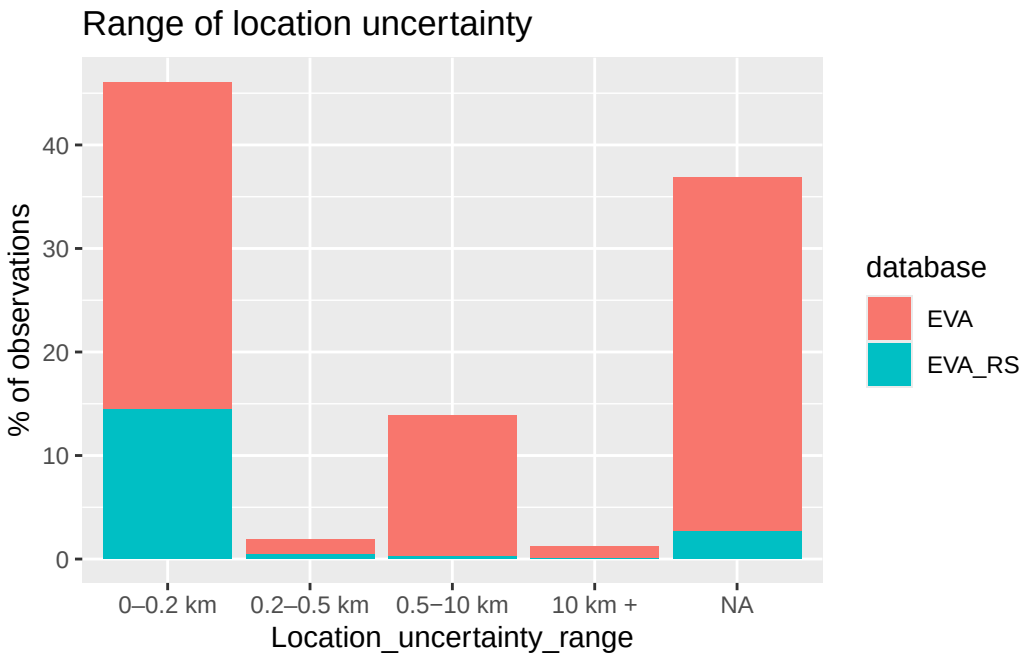
Location uncertainty

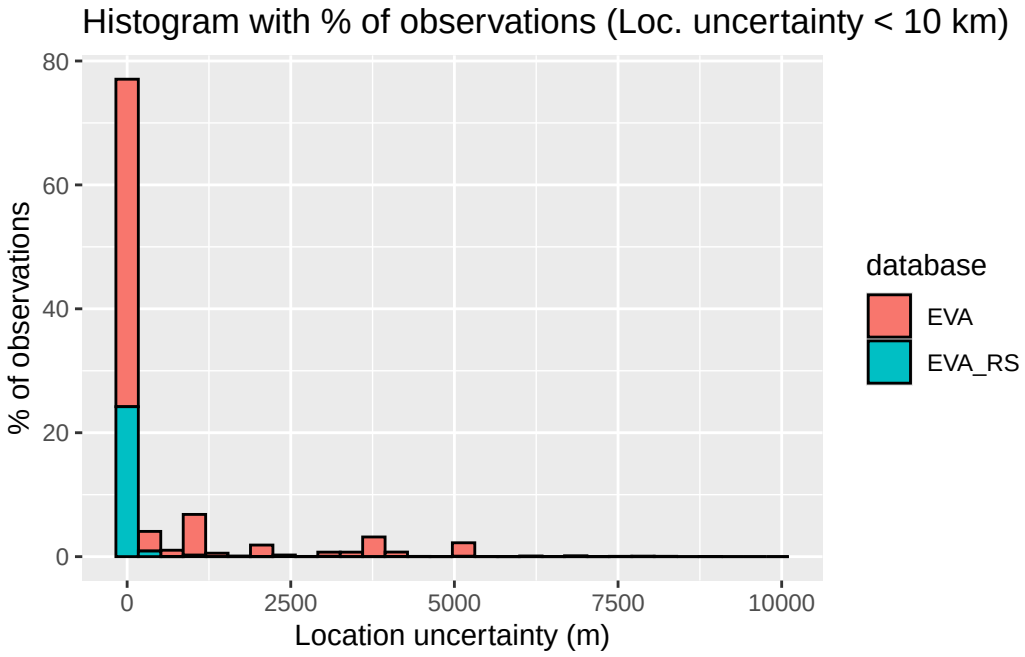
Availabilty

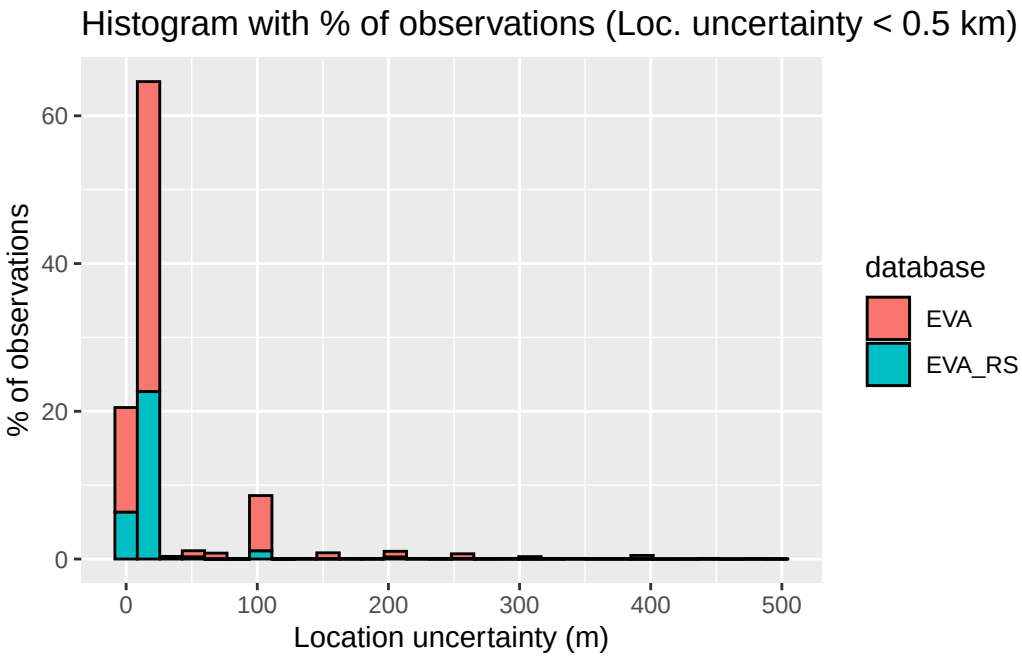


Distribution when available




Range of location uncertainty
% of observations
40
30
20
10
0
0–0.2 km
0.2–0.5 km
0.5−10 km
10 km +
NA
Location_uncertainty_range
database
EVA
EVA_RS


Histogram with % of observations (Loc. uncertainty < 10 km)
% of observations
80
60
40
20
0
0
2500
5000
7500
10000
Location uncertainty (m)
database
EVA
EVA_RS



TO DO: Location uncertainty (last resurveys for EVA_RS)

Extract last resurvey for each each ReSurvey plot (defined by combination of RS_CODE, ReSurvey site and ReSurvey plot). So far, discard 111 resurvey observations corresponding to ReSurvey project "Swiss arable flora since the 1920s" (RS_CODE = CH_0002) where ReSurvey plot is NA.

Careful! There are plots with muliple records for the last year (and also for other years).

Solve this first and check location uncertainty for last resurveys later on.

Check for plots with multiple records for the same year

For 2100 unique combinations of RS_CODE, ReSurvey site, ReSurvey plot and year, there are multiple records, sometimes with different dates within the year (OK - we choose one), **but also with the same exact date**.

Example below:

```
##   PlotObservationID     RS_CODE ReSurvey.site ReSurvey.plot year       date
## 1            748986 CZ_0019_001         N_Cep         N_Cep 1952 1952-09-12
## 2            748990 CZ_0019_001         N_Cep         N_Cep 1952 1952-09-12
## 3            748991 CZ_0019_001         N_Cep         N_Cep 1952 1952-09-12
## 4            748992 CZ_0019_001         N_Cep         N_Cep 1952 1952-09-12
## 5            748993 CZ_0019_001         N_Cep         N_Cep 1952 1952-09-12
## 6            749003 CZ_0019_001         N_Cep         N_Cep 1952 1952-09-22
## 7            749004 CZ_0019_001         N_Cep         N_Cep 1952 1952-09-22
## 8            749018 CZ_0019_001         N_Cep         N_Cep 1952 1952-09-12
## 9            749028 CZ_0019_001         N_Cep         N_Cep 1952 1952-09-12
##   count        Country EUNISa ReSurvey.project ReSurvey.observation Lon_updated
## 1     9 Czech Republic   <NA>       TrebonFens         N_Cep_1_1952    14.83889
```

```
## 2     9 Czech Republic    Q22       TrebonFens         N_Cep_2_1952    14.83472
## 3     9 Czech Republic    Q22       TrebonFens         N_Cep_3_1952    14.83472
## 4     9 Czech Republic    Q22       TrebonFens         N_Cep_4_1952    14.83472
## 5     9 Czech Republic    Q22       TrebonFens         N_Cep_5_1952    14.82917
## 6     9 Czech Republic    Q24       TrebonFens         N_Cep_6_1952    14.83472
## 7     9 Czech Republic    Q24       TrebonFens         N_Cep_7_1952    14.83472
## 8     9 Czech Republic    Q22       TrebonFens         N_Cep_8_1952    14.83472
## 9     9 Czech Republic    Q22       TrebonFens         N_Cep_9_1952    14.83750
##   Lat_updated
## 1    48.91944
## 2    48.92083
## 3    48.92083
## 4    48.92083
## 5    48.92222
## 6    48.92083
## 7    48.92083
## 8    48.92083
## 9    48.92083
```

I have sent a list to Ilona to see what to do with this issue.

Keep GOOD observations WP3

Defining "good" observations as those that **have an EUNIS assigned** (any level) and where **location uncertainty is available and < 0.5 km** (decide if we use 0.2 or other value?).

From **EVA + EVA ReSurvey**, the **proportion of "good" observations** according to these criteria is **0.231**.

From **EVA ReSurvey only**, the **proportion of "good" observations** according to these criteria is **0.484**.

From **EVA + EVA ReSurvey**, the **proportion of observations with EUNIS assigned, irrespective of location uncertainty** is **0.619**.

From **EVA ReSurvey only**, the **proportion of observations with EUNIS assigned, irrespective of location uncertainty** is **0.591**.

Maybe include those without location uncertainty but where location method is GPS, etc.?

Check for plots with multiple records for the same year in "good" observations

Keep GOOD observations that do not have multiple records for the same year

From **EVA + EVA ReSurvey**, the **proportion of "good" observations** according to the above criteria **that do not have multiple records for the same year** is **0.229**.

From **EVA ReSurvey only**, the **proportion of "good" observations** according to the above criteria **that do not have multiple records for the same year** is **0.473**.