RISKS POSED BY INVASIVE MAMMALS TO THE EMERGENCE AND SPREAD OF ZOONOTIC DISEASES

Supervisor: Belinda Gallardo Armas

Paola Monguilod Brun

Universidad Zaragoza 1542

Facultad de Ciencias
Universidad Zaragoza

IPE CSIC

JAE INTRO
Becas De Introducción A La Investigación

Invasive Alien Species (IAS)

Non-native organisms

Spread extensively

Introduced through human activities (both intentionally and unintentionally)

Negative impact on biodiversity and ecosystems

Invasive Animals and Human Health

Zoonoses originating from wildlife represent >40% of all global EIDs

Invasive Animals:

- Bring new pathogens to the introduced range

- Are introduced and proliferate in human-modified environments where animal-human contact is frequent

- Are reservoirs: can harbor pathogens and be asymptomatic

- Have competitive traits that allow them to rapidly reproduce and spread to other areas

- Invasive mammals in particular are phylogenetically close to humans, which increases the possibilities of zoonotic spillover.
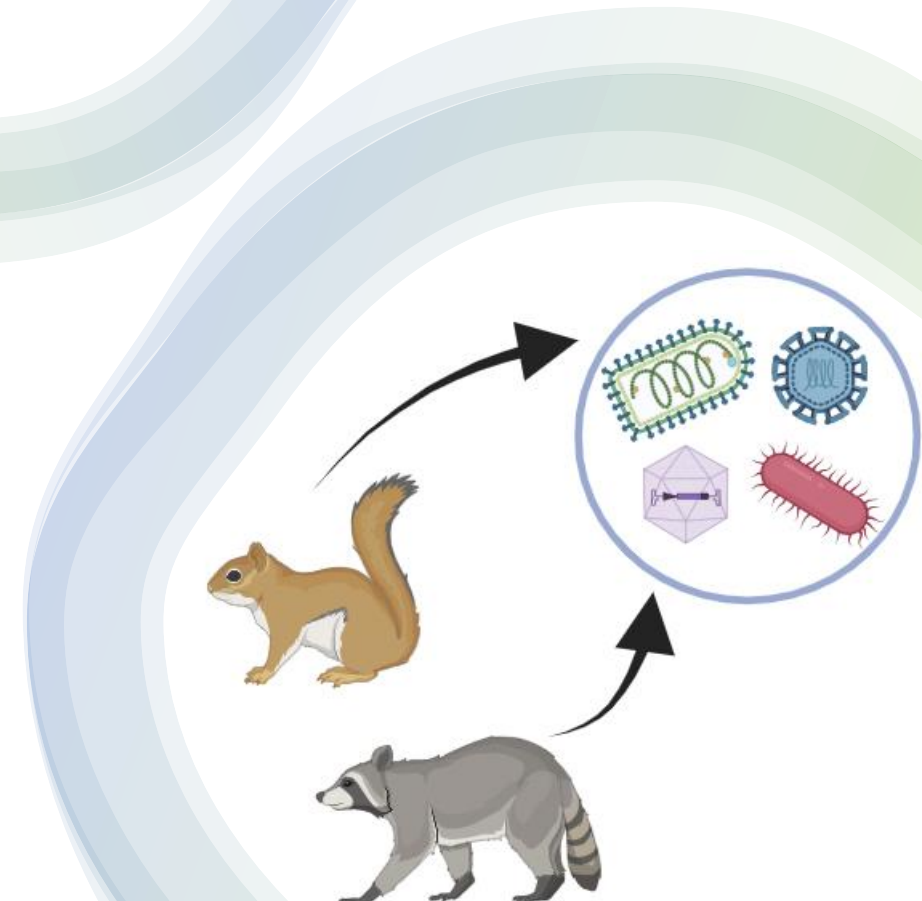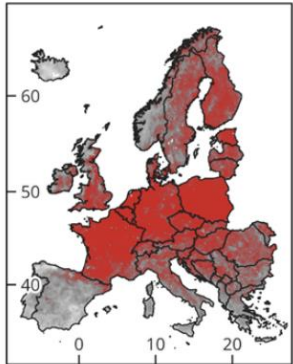
IAS and Climate Change

- The increase in temperature due to climate change facilitates the geographic expansion of IAS

- Climate change will result in increased international trade and travel, contributing to the unintentional spread of IAS to new regions

Epstein, 2001; Walther et al., 2009

Objectives

1. Identify relevant IAS susceptible to zoonotic disease transmission in Europe.

2. Identify the pathogens hosted by the selected IAS, that are relevant for human health.

3. Spatially identify the areas under greatest risk of invasion and therefore EID transmission in Europe under current climate conditions: "disease hotspots".
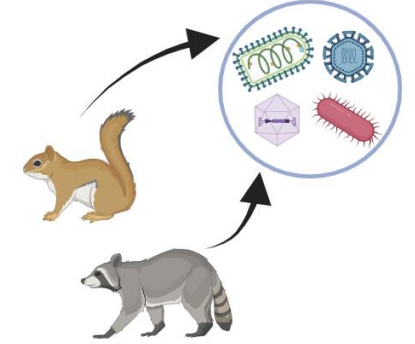
4. Develop a practical assessment tool to prioritize IAS in terms of their risk to human health, that integrates the information generated in this study.

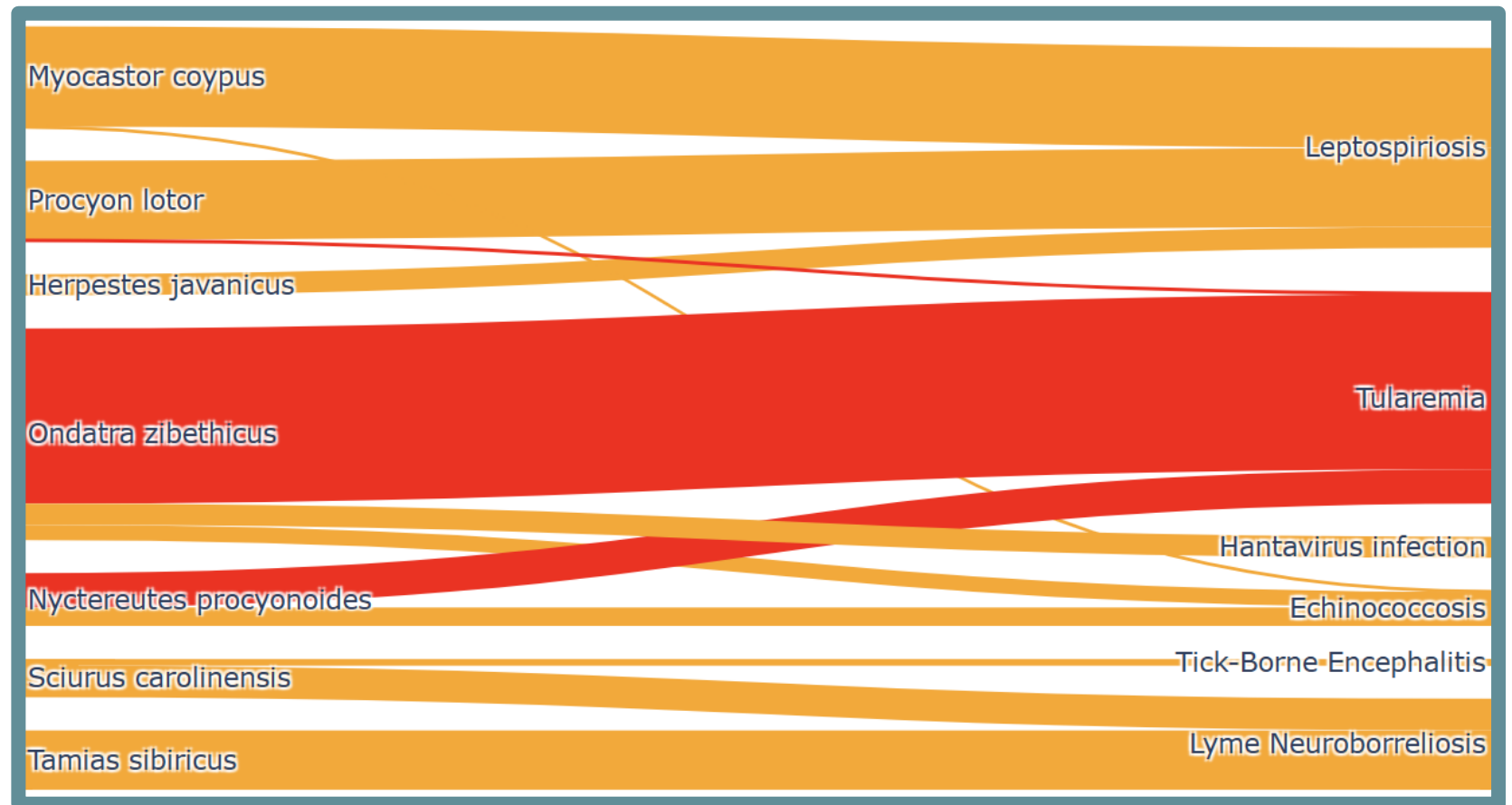# 1. Which invasive mammals pose a threat to human health in Europe?

Carnivora

Herpestes javanicus

Nasua nasua

Nyctereutes procyonoides

Procyon lotor

Rodentia

Sciurus carolinensis

Tamias sibiricus

Myocastor coypus

Ondatra zibethicus

2. Which pathogens do the selected mammals harbor that are of relevance for human health?

97 distinct pathogens

Average of 16 pathogens per invasive mammal studied

Medium-High risk | High risk

Myocastor coypus · Procyon lotor · Herpestes javanicus · Ondatra zibethicus · Nyctereutes procyonoides · Sciurus carolinensis · Tamias sibiricus

Leptospiriosis · Tularemia · Hantavirus infection · Echinococcosis · Tick-Borne Encephalitis · Lyme Neuroborreliosis

3. Which regions are suitable for invasive mammals that pose a threat to human health?

GBIF

Geo-referenced data: Presence of the chosen invasive mammals

↓

Species Distribution models (SDMs)

↓

Predict the potential distribution of the chosen invasive mammals under current and future conditions

Identify regions that have conditions suitable for species establishment based on the predictors used:
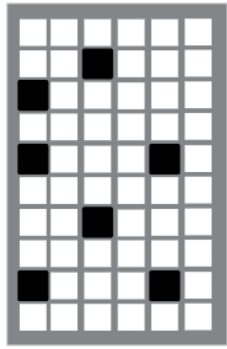
- Accessibility
- Elevation
- Temperature-related variables
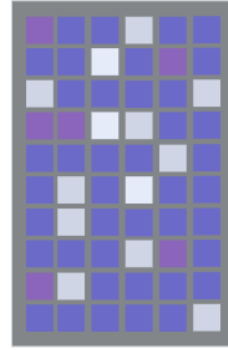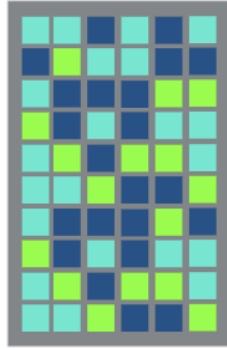- Precipitation-related variables

Temperature and Accessibility most important predictors in distribution of the chosen invasive mammals

(Gallardo & Aldridge, 2020; Guisan et al., 2017; Pérez et al., 2022; Srivastava et al., 2019)

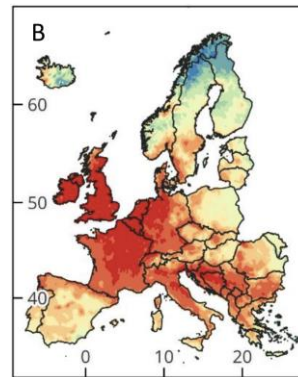# Species Distribution Models (SDMs)

Species Occurence Data

GBIF

Predictors

Model Calibration
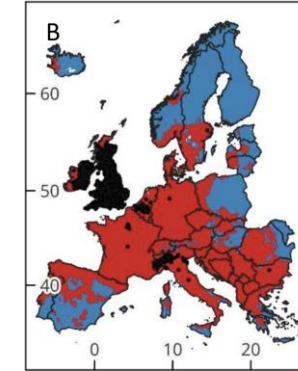Algorithms: GLM, GAM, RF, GBM (Ensemble modelling)
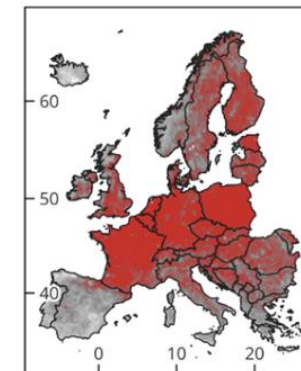
Projection of ensemble models onto Europe

Optimize TSS

Binary species suitability maps

Continuous species suitability maps

Add continuous maps of IAS resevoirs for each EID, weighted by their prevalence

Disease hotspot maps

Load species presence data

```r
188  Ozibe_file <- read.csv("Species/Ozibe.csv")
189  x <- Ozibe_file[,"x"]
190  y <- Ozibe_file[,"y"]
191  presence <- Ozibe_file[,"presence"]
192  OzibeReor <- cbind(x, y, presence,  Mcoyp = NA, Nnasu = NA, Hjava=NA,  Nproc= NA, Ozibe=1, Ploto=NA, Scaro=NA, Tsibi=NA)
193
194  Ploto_file <- read.csv("Species/Ploto.csv")
195  x <- Ploto_file[,"x"]
196  y <- Ploto_file[,"y"]
197  presence <- Ploto_file[,"presence"]
198  PlotoReor <- cbind(x, y, presence,  Mcoyp = NA, Nnasu = NA, Hjava=NA,  Nproc= NA, Ozibe=NA, Ploto=1, Scaro=NA, Tsibi=NA)
199
200  Scaro_file <- read.csv("Species/Scaro.csv")
201  x <- Scaro_file[,"x"]
202  y <- Scaro_file[,"y"]
203  presence <- Scaro_file[,"presence"]
204  ScaroReor <- cbind(x, y, presence,  Mcoyp = NA, Nnasu = NA, Hjava=NA,  Nproc= NA, Ozibe=NA, Ploto=NA, Scaro=1, Tsibi=NA)
205
206  Tsibi_file <- read.csv("Species/Tsibi.csv")
207  x <- Tsibi_file[,"x"]
208  y <- Tsibi_file[,"y"]
209  presence <- Tsibi_file[,"presence"]
210  TsibiReor <- cbind(x, y, presence,  Mcoyp = NA, Nnasu = NA, Hjava=NA,  Nproc= NA, Ozibe=NA, Ploto=NA, Scaro=NA, Tsibi=1)
211
212  species <- rbind(MycopReor, NnasuReor, HjavaReor, NprocReor, OzibeReor, PlotoReor, ScaroReor, TsibiReor)
213  write.csv(species, file = "species.csv", row.names = FALSE)
```

	x	y	presence	Mcoyp	Nnasu	Hjava	Nproc	Ozibe	Ploto	Scaro	Tsibi
1	0.91667	43.58	1	1	NA	NA	NA	NA	NA	NA	NA
2	-95.58333	29.42	1	1	NA	NA	NA	NA	NA	NA	NA
3	4.41667	43.42	1	1	NA	NA	NA	NA	NA	NA	NA
4	-46.91667	-23.58	1	1	NA	NA	NA	NA	NA	NA	NA
5	-58.41667	-34.58	1	1	NA	NA	NA	NA	NA	NA	NA
6	14.41667	50.08	1	1	NA	NA	NA	NA	NA	NA	NA
7	-123.75000	46.25	1	1	NA	NA	NA	NA	NA	NA	NA
8	7.25000	45.08	1	1	NA	NA	NA	NA	NA	NA	NA

Calibration of the models: Selection of pseudo absences

- Package BIOMOD2 version 4.2-3 in R environment version 4.2.2

- Calibration of models at global scale (including information about the native and invasive ranges)

- Ensemble modelling approach combining 4 algorithms

- The algorithms used require explicit data of presence and absence of the species

- Random selection of 10,000 pseudo-absences, repeated 3 times

```
251     ### Initialization
252     # Select pseudo absences
253     myBiomodData <- BIOMOD_FormatingData(resp.var = myResp,        # Response variable: the presence data
254                                          expl.var = myExpl,          # Explanatory variables: environmental predictors at global scale
255                                          resp.xy = myRespXY,         # Spatial coordinates (latitude and longitude) of the species presence data
256                                          resp.name = myRespName,     # Define a name for the response variable
257                                          PA.nb.rep = 3,              # Number of times the 10.000 selection is repeated (independently)
258                                          PA.nb.absences = 10000,     # Number of pseudo-absence points to be randomly selected  over the whole
        study area (global)
259                                          PA.strategy = 'random',     # Strategy for selecting pseudo-absences. Set to 'random', meaning the
        pseudo-absence points are chosen randomly from the study area.
260                                          na.rm= T)                   # Removes missing values (NA) from the data set
```

Calibration of the models: Modelling step

- Calibration of models using 4 algorithms: GLM and GAM (regression models), RF and GBM (machine learning).

- Evaluate the predictive performance of the models: presence of species was randomly divided into training data (70%) and test data (30%).

- 12 model replicas per species (4 algorithms x 3 data partitions)

- Metrics for assessing the calibrated models: TSS and ROC

- Permutations to assess the importance of the variables

```
262    #myBiomodData
263    myBiomodOptions <- BIOMOD_ModelingOptions()
264
265    ### Modelling
266    myBiomodModelOut <- BIOMOD_Modeling(bm.format = myBiomodData,          # Uses the formatted data object created earlier
267                                        bm.options = myBiomodOptions,        # Specifies the modeling options configured for BIOMOD
268                                        models = c("GLM", "GBM","RF", "GAM"), # These are the algorithms I'll be using
269                                        CV.nb.rep = 3,                        # Number of repetitions for cross-validation
270                                        CV.perc = 0.7,                        # Percentage of data used for training during data splitting (70%
       training, 30% testing)
271                                        var.import = 3,                       # Calculates variable importance three times for increased
       reliability
272                                        metric.eval = c('TSS','ROC'),         # Evaluation metrics used
273                                        CV.do.full.models = FALSE             # Indicates whether to create full models using all available
       data (FALSE means only cross-validation models are created)
274                                        )
```

Calibration of the models: Ensemble modelling

- **Ensemble models** were created **for each species**

- **Each** of the 12 **model replicas** was **assigned a weight** based on its predictive capacity, measured by the **TSS**

- Replicas with **TSS < 0.7** were **discarded**

```
285    ### Building ensemble-models (Put the options to do the ensemble)
286    myBiomodEM <- BIOMOD_EnsembleModeling(bm.mod = myBiomodModelOut,      # Use the output from the modeling step as input for the ensemble
287                                          models.chosen = 'all',          # Include all model replicas from the previous step in the ensemble modeling
288                                          em.by = 'all',                  # Create ensemble models by considering all available algorithms and scenarios
289                                          em.algo = c('EMmean'),          # Define the ensemble modeling algorithm as 'EMmean', which calculates the mean probability
       across models
290                                          metric.select = c('TSS'),       # Define the selection metric for including models in the ensemble as 'TSS'
291                                          metric.select.thresh = c(0.7),  # Define the threshold of the selecting metric as 0.7; only models with TSS ≥ 0.7 are
       included
292                                          metric.eval = c('TSS', 'ROC', 'SR'), # Define the metrics used to evaluate the ensemble models: ' 'TSS', 'ROC' and 'SR'
293                                          var.import = 3,                 # Specify that variable importance should be calculated three times for the ensemble
294                                          EMci.alpha = 0.05,              # Define the confidence level for the confidence interval as 0.05
295                                          EMwmean.decay = 'proportional', # Define the average weighted decay form as 'proportional'
296                                          prob.mean.weight = T) # Use the average weight of the probabilities for creating the ensemble models
297
```

Model projections onto Europe for the current scenario

- myBiomodProj:

  - Continuous and binary maps for individual models (12 model replicas per species)

- myBiomodEF:

  - Continuous and binary maps for ensemble models (1 per species)

- Continuous maps (suitability 0–1000 scale) were converted into binary maps (0: unsuitable, 1:suitable)

  optimizing the TSS of the model

```
303    # Make projection (Europe present)
304  myBiomodProj <- BIOMOD_Projection(
305                    bm.mod = myBiomodModelOut,  # Use the output from the modeling step as input for the projection
306                    proj.name = 'Current_',     # Define the name of the projection as 'Current_'
307                    new.env = myExplEUR,        # Sets the environmental variables for the projection as myExplEUR (Europe's current variables)
308                    models.chosen = 'all',      # Selects all available models from the previous step for the projection
309                    metric.binary = 'all',      # Treats all metrics  (TSS, ROC, SR) as binary in the projection
310                    metric.filter = 'all',      # Includes all available metrics without applying any filter
311                    compress = T,               # Compresses the output files of the projection to save space
312                    build.clamping.mask = F,    # Does not build a clamping mask during the projection
313                    output.format = '.img'      # Sets the output format of the projected files as '.img'
314                    )
315
316
317  # Make ensemble-model projections based on the models and environmental variables
318  myBiomodEF <- BIOMOD_EnsembleForecasting(bm.em = myBiomodEM,         # Uses the ensemble-model output as input
319                            bm.proj = myBiomodProj,                     # Uses the projection output for generating ensemble forecasts
320                            models.chosen = 'all',                      # Includes all models for ensemble forecasting
321                            metric.binary = "TSS",                      # Generates binary maps based on the TSS metric (threshold applied)
322                            metric.filter = NULL,                       # Does not filter the metrics (all included)
323                            output.format = '.img'                      # Specifies the output file format as '.img'
324  )
```

Model projections onto Europe for the future scenarios

Continuous suitability maps
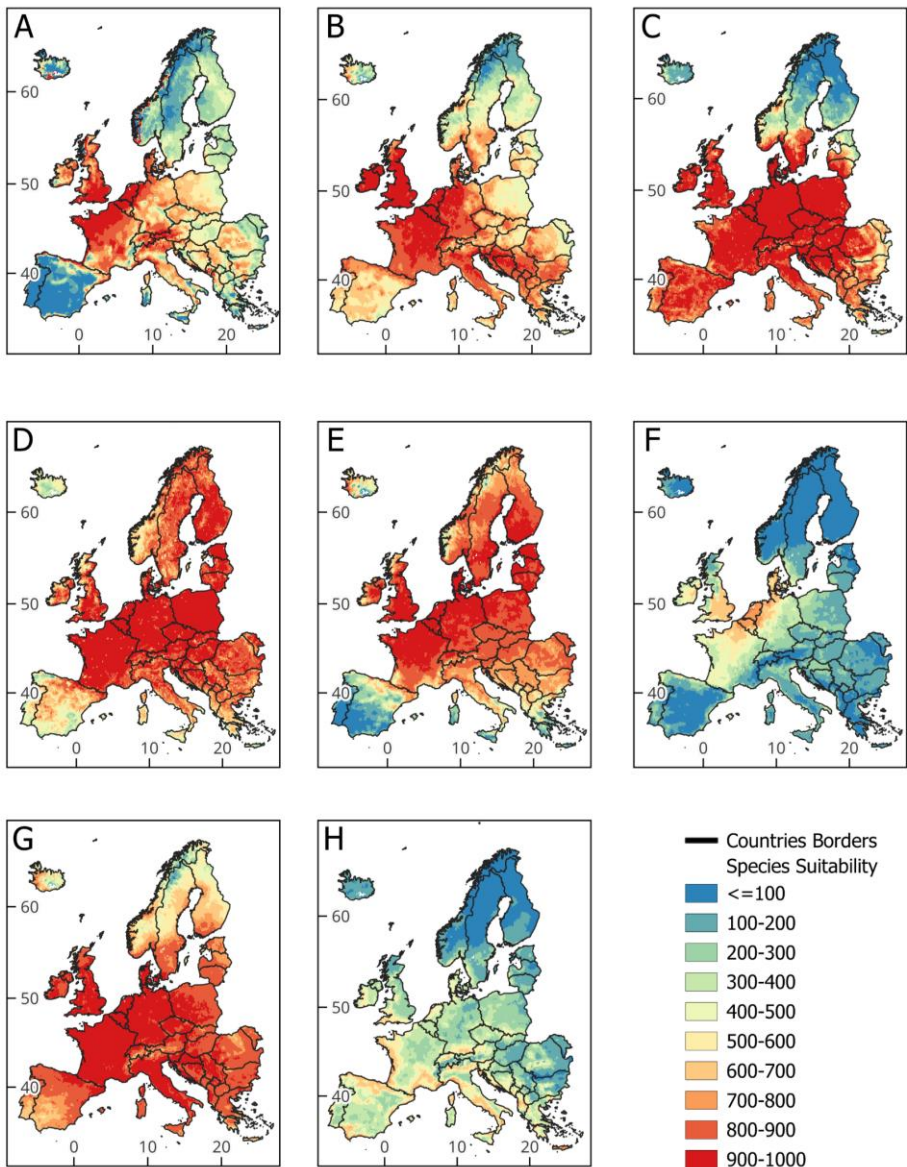
Current IAS Distribution

Binary suitability maps
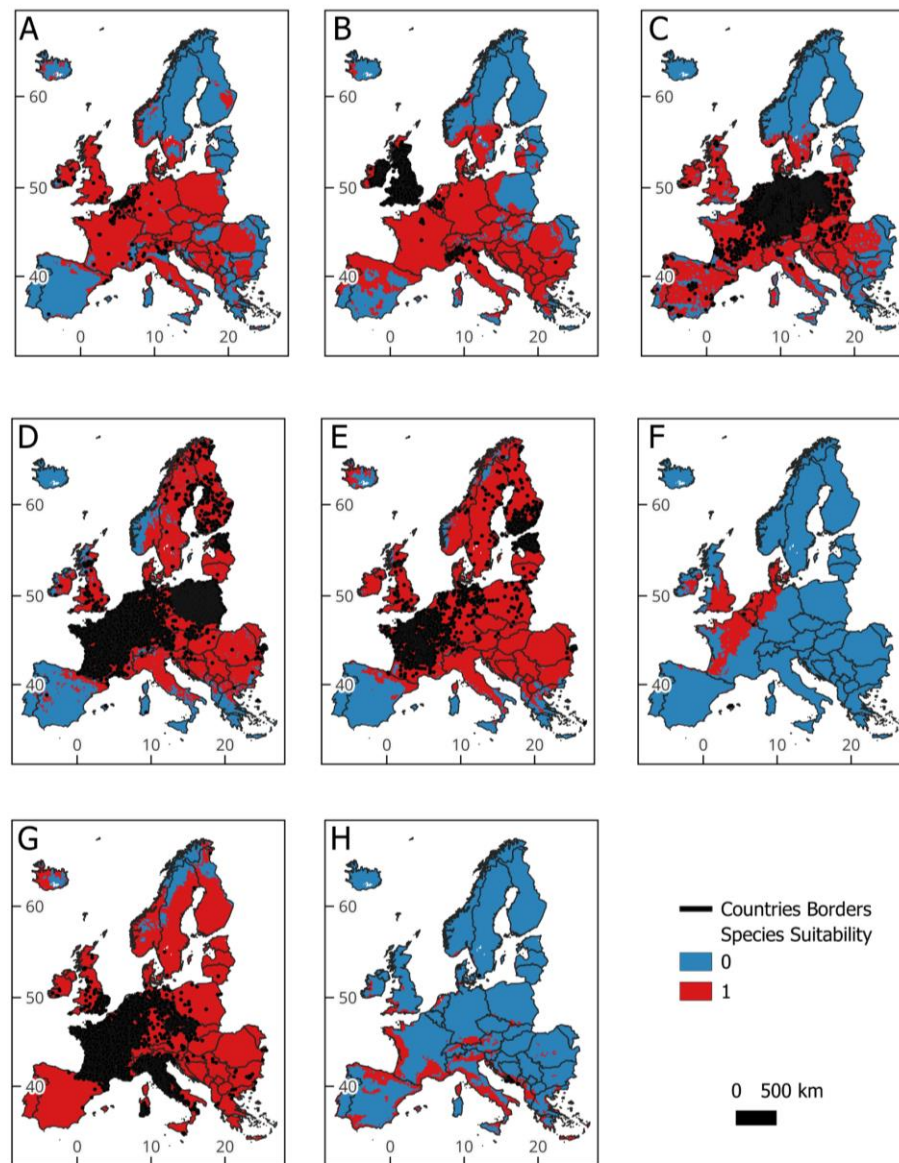

A. *Tamias sibiricus*

B. *Sciurus carolinensis*

C. *Procyon lotor*

D. *Ondatra zibethicus*

E. *Nyctereutes procyonoides*

F. *Nasua nasua*

G. *Myocastor coypus*

H. *Herpestes javanicus*

Countries Borders
Species Suitability

<=100
100-200
200-300
300-400
400-500
500-600
600-700
700-800
800-900
900-1000

0   500 km

Countries Borders
Species Suitability

0
1

0   500 km

15

Species Range Change (SRC) maps
Future IAS distribution

SRC maps = Future binary maps – Current binary maps

Binary maps
0: unsuitable
1: suitable

SRC (future–present)
Gain = 1 – 0 = 1
No change = 1 – 1 / 0 – 0 = 0
Loss = 0 – 1 = –1

QGIS

A. *Tamias sibiricus*
B. *Sciurus carolinensis*
C. *Procyon lotor*
D. *Ondatra zibethicus*
E. *Nyctereutes procyonoides*
F. *Nasua nasua*
G. *Myocastor coypus*
H. *Herpestes javanicus*

Colour Code
— Countries Borders
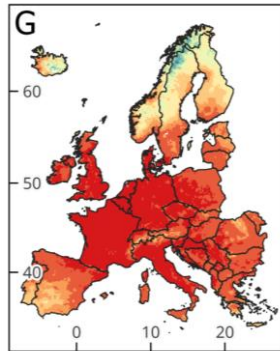■ Loss
□ No change
■ Gain

0 500 km

16
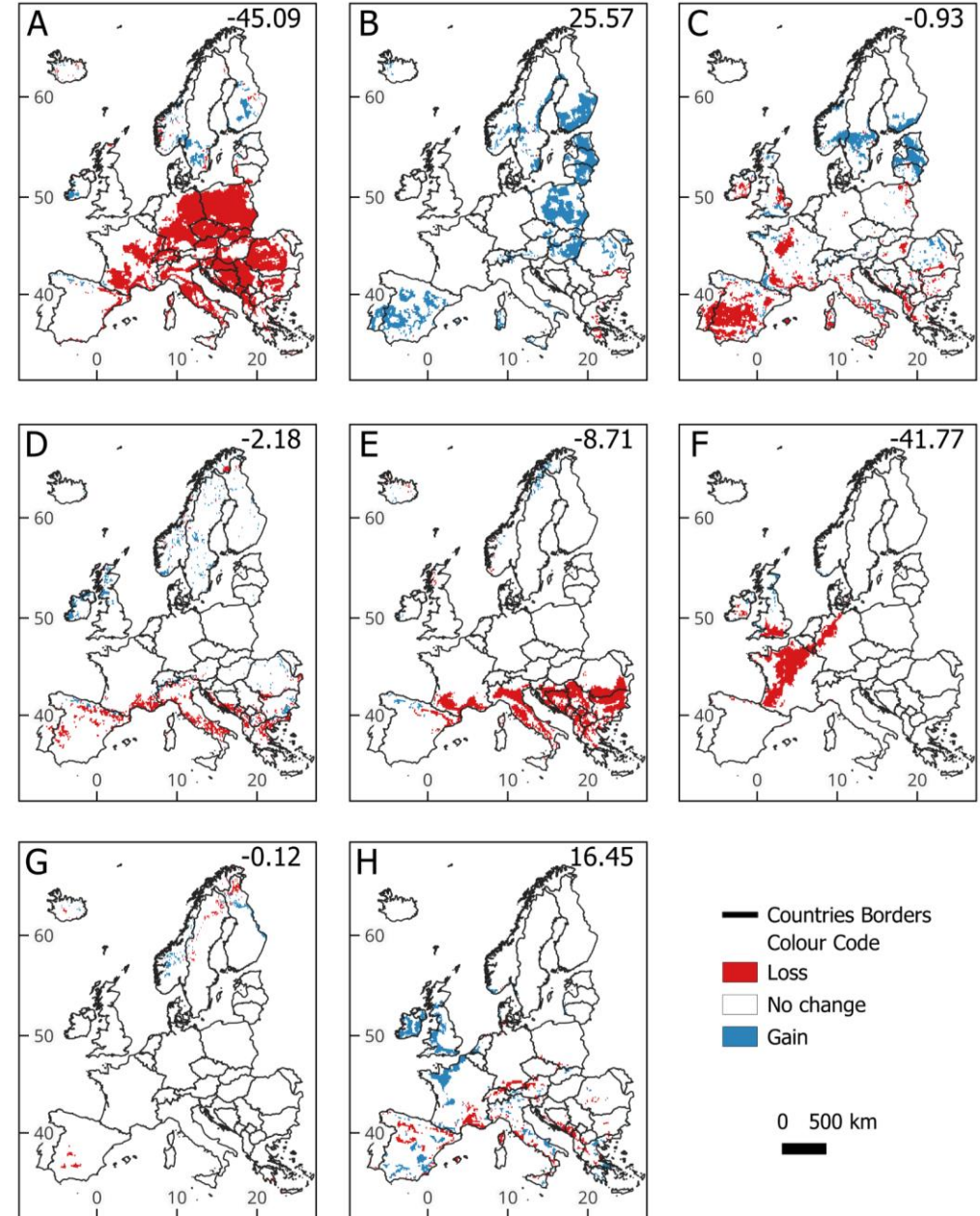
Disease hotspot maps

Example: Echinococcosis

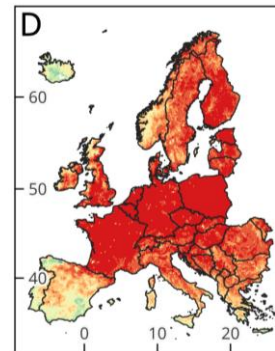| Reservoir Species	Prevalence
Myocastor coypus	0.4%
Nyctereutes procyonoides	12%
Ondatra zibethicus	11%

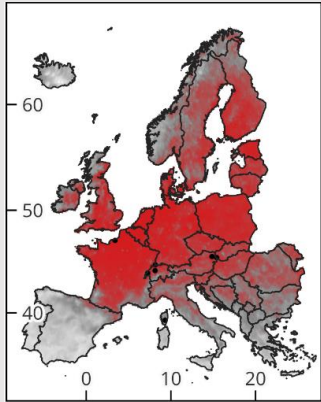*Myocastor coypus* *0,004 + *Nyctereutes procyonoides* *0,12 + *Ondatra zibethicus* *0,11 = A. Echinococcosis
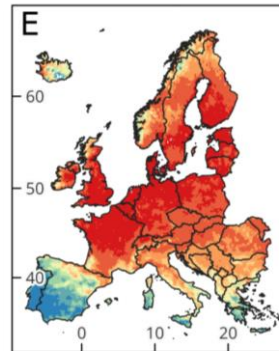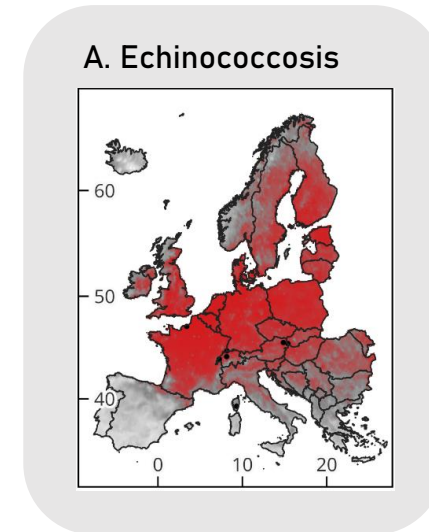
Disease hotspot maps
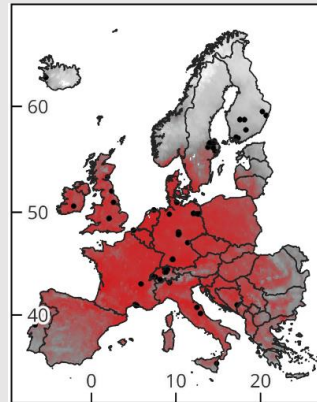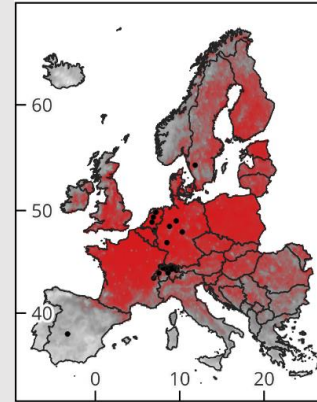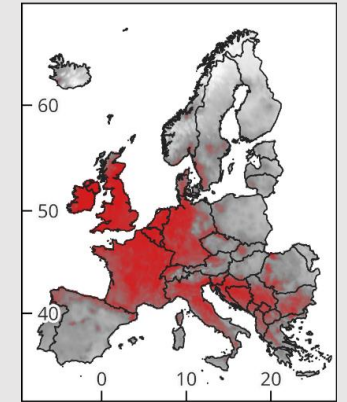

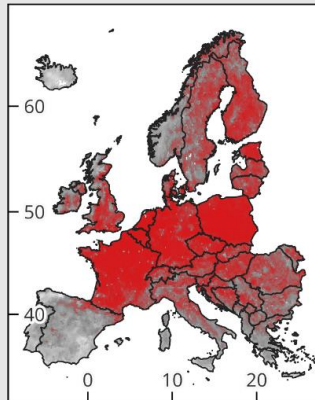A. Echinococcosis
1122

B. Leptospiriosis
1544

C. Tularemia
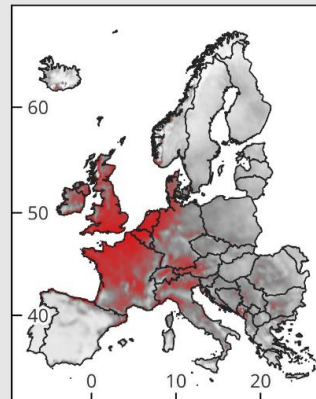853

D. Tick-Borne Encephalitis
1098

E. Hantavirus infection
719

F. Lyme Neuroborreliosis
317

Countries Borders
Disease Hotspots
Max
Min

0 500 km

Disease incidence in Europe (cumulative number of cases)

● Occurrences of the pathogens that cause the disease

GBIF

18

Human Exposure

Percentage of European population residing in areas considered suitable for the species

Calculated using only European raster layers

Example: *Herpestes javanicus* in the current scenario

$$\frac{\text{Sum } (\textit{Herpestes javanicus} \text{ binary Current} * \text{population Current})}{\text{sum (population Current)}}$$

Human exposure provides insights into the likelihood of human populations interacting with the specific IAS under study

To be considered: Files exported from QGIS have elements separated by a comma and decimals indicated with a dot.

To ensure they can be properly read in Excel, the elements should be separated by a semicolon, and decimals should use a comma.

Human Exposure

QGIS GBIF

	Current (%)	Low Emissions (%)	Business As Usual (%)	High Emissions (%)
Herpestes javanicus	22,50	31,03	27,28	29,93
Myocastor coypus	99,43	99,42	99,12	99,58
Nasua nasua	28,80	30,21	24,88	20,67
Nyctereutes procyonoides	87,49	81,04	76,74	75,75
Ondatra zibethicus	85,2	81,33	80,10	80,06
Procyon lotor	81,80	82,95	74,71	76,28
Sciurus carolinensis	82,43	90,35	90,83	92,65
Tamias sibiricus	78,16	60,09	49,14	47,17

→ Highest

→ Lowest

4. Integrate the information into a new tool to assess the risk posed by IAS to human health

Considering:

- **Prevalence** of the disease-causing pathogens in IAS populations

- **Severity of the diseases** caused by the pathogens harbored by IAS.

- **Geographic Distribution** of the IAS

- **Human population exposure** to the IAS

Score of Impact	Average Prevalence of Disease-Causing Pathogens	Average Severity of Diseases the pathogens IAS harbor cause	Human Exposure	Geographic Distribution	Final Score Description
1: Minimal	1: Very low (less than 1%)	1: Mild (BSL-1)	1: Low (0 - 20%)	1: Local (0 - 20 %)	1: IAS with low prevalence of disease-causing pathogens, causing mild diseases, with minimal human exposure. The impact is limited to a local area.
2: Minor	2: Low (1 - 10%)	2: Moderate (BSL-2)	2: Low to Moderate (20 -40%)	2: Local (20 - 40 %)	2: IAS with low prevalence of disease-causing pathogens, causing moderate diseases, with minor to moderate human exposure. Impact remains localized.
3: Moderate	3: Moderate (10 - 30%)	3: Moderate to Severe (BSL-3)	3: Moderate (40 – 60%)	3: Moderate to Widespread (40 - 60 %)	3: IAS with moderate prevalence of disease-causing pathogens, causing moderate to severe diseases, with moderate human exposure. Impact can be localized or extended to multiple regions.
4: Major	4: High (30 - 70%)	4: Severe (BSL-3)	4: High (60 – 80%)	4: Widespread (60 - 80 %)	4: IAS with high prevalence of disease-causing pathogens, causing severe diseases, with significant human exposure. Impact is widespread, affecting multiple regions.
5: Massive	5: Very High (greater than 70%)	5: Very Severe (BSL-4)	5: Very High (80 – 100%)	5: Widespread (80 – 100%)	5: IAS with a very high prevalence of disease-causing pathogens, causing very severe diseases, with massive human exposure. Impact is extensive, potentially affecting ecosystems and regions extensively.

Invasive Species Assessment Tool

Average Prevalence of Pathogens

[1: Very Low (<1%) ▼]

Average Prevalence of Disease-Causing Pathogens that IAS harbor

Severity of Diseases

[1: Mild (BSL-1) ▼]

Average Severity of the diseases the pathogens IAS harbor cause (according to BSL Classification)

Human Exposure

[1: Low (0-20%) ▼]

Percentage of European population residing in areas considered suitable for the IAS

Geographic Distribution

[1: Local (0-20%) ▼]

Percentage of European territory suitable for the establishment of the IAS

Final Score Description

IAS with low prevalence of disease-causing pathogens, causing mild diseases, with minimal human exposure. The impact is limited to a local area.
Exact Final Score: 1
Final Score (rounded): 1 out of 5

Using library
Shiny

23

4. Integrate the information into a new tool to assess the risk posed by IAS to human health

Considering:

- **Prevalence** of the disease-causing pathogens in IAS populations

- **Severity of the diseases** caused by the pathogens harbored by IAS.

- **Geographic Distribution** of the IAS

- **Human population exposure** to the IAS

Species	Human health impact score in our study	Biodiversity impact score in the Union List Risk Assessment
Herpestes javanicus	3	4
Myocastor coypus	4	4
Nasua nasua	3	3
Nyctereutes procyonoides	4	4
Ondatra zibethicus	4	4
Procyon lotor	4	3
Sciurus carolinensis	4	4
Tamias sibiricus	4	3

Main Findings

- 8 invasive mammals regulated in Europe as potential zoonotic hosts and an average of 16 pathogens in them were identified.

- Disease hotspots found in Western and Central Europe

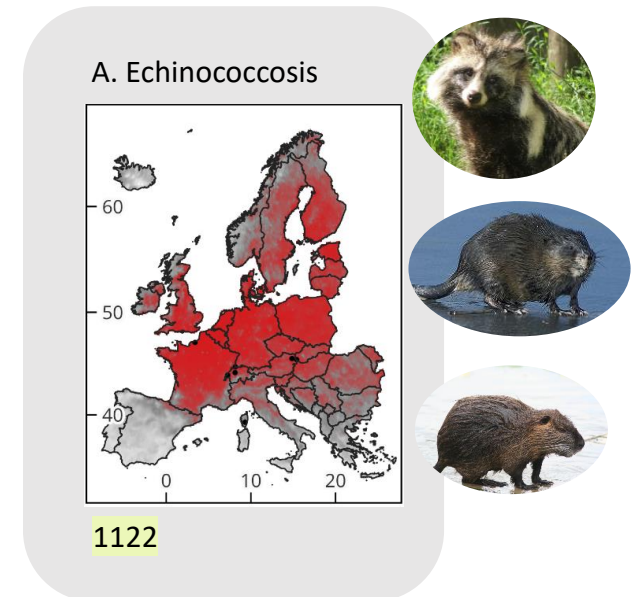- Several of the IAS fell into the Major threat category based on our risk assessment tool

Summary use of GBIF data in this study

Parameters of Risk Assessment tool

- Prevalence of the disease–causing pathogens in IAS populations

- Severity of the diseases caused by the pathogens harbored by IAS.

- Geographic Distribution of the IAS

- Human population exposure to the IAS

Creation of Disease Hotspot maps

A. Echinococcosis

1122

GBIF

Thank you very much for your attention

Paola Monguilod Brun