Species Distribution Models

Numerical tools that **combine** species occurrences and/or the knowledge about a species with environmental estimates to **predict** species distribution across time and space[1]

[1] Elith et al. 2009

Preparation

Knowledge of species ecology

Assess data properties
- Types of data: Presence-only, Presence-absence, abundances
- Temporal and spatial scale
- Number of observations

Assess data quality
Remove duplicates, check for taxonomy problems (e..g synonyms) and other suspicious records

Evaluate data biases
Use cartograms (sampling) and expert knowledge (taxonomic) to identify biases.

Imperfect detection
Species vary in detectability Include this factor in the process

Species occurrences
Databases (e.g. GBIF)
Local sites
Museums and herbaria

Environmental variables
Bioclimatic
Topographic
Land use and cover

Assess data quality and properties
Resolution, temporal and spatial extent, type of data

Select relevant variables
Select the variables that match the requirements of the species

Assess variables correlations
Remove highly-correlated variables

Match the data
Scale, extent and projection

Data **pre-processing**

Modelling

Examine you data
And select the candidate models that match the objective, expertise and the time and data available

Understand the model selected
The assumptions and modelling process

Model **selection**

Expertise | Data | Knowledge | Time

Process-based
Hybrid
Correlative

Models

Absences

True absences should be verified

creating **Pseudoabsences**
1. Within areas environmentally different from presences.
2. Within geographical distance from presences and in areas environmentally different from presences.

Pseudoabsences are affected
By biases in the species occurrences

Multiple diagnostics
Use multiple metrics to assess and compare outcomes

Model **diagnostics**
- Evaluate predicted vs. real distribution
- Perform analyses of specificity (% of true negatives) and sensitivity (% of true positives).
- For example, True skill statistic (TSS) considers both.

Threshold selection
Be aware that most metrics are affected by threshold selection.

Outcome

Report
- Data sources and properties
- Bias assessment and methods to tackle it.
- Model selection and diagnostics
- Uncertainty at each step
- Spatial/temporal resolution

Map of predictions
Probability of presence

+

Process & diagnostics **report**

+

Uncertainty map
Cartogram

Sources of uncertainty
- Modelling approach
- Data quality & quantity
- Knowledge of the species ecology