



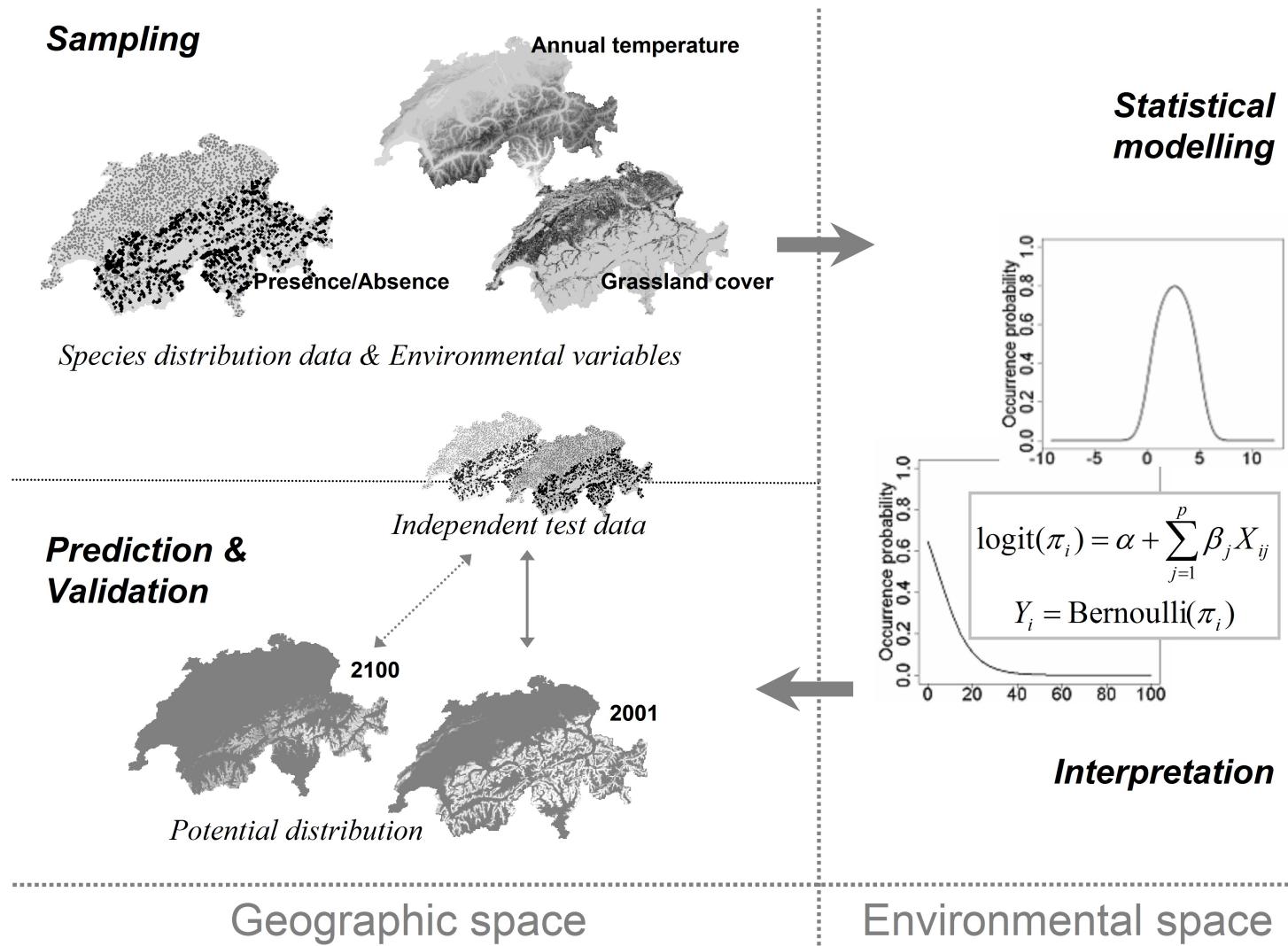
Species distribution modelling: data, data cleaning, pseudo- absences, and more ...

Damaris Zurell

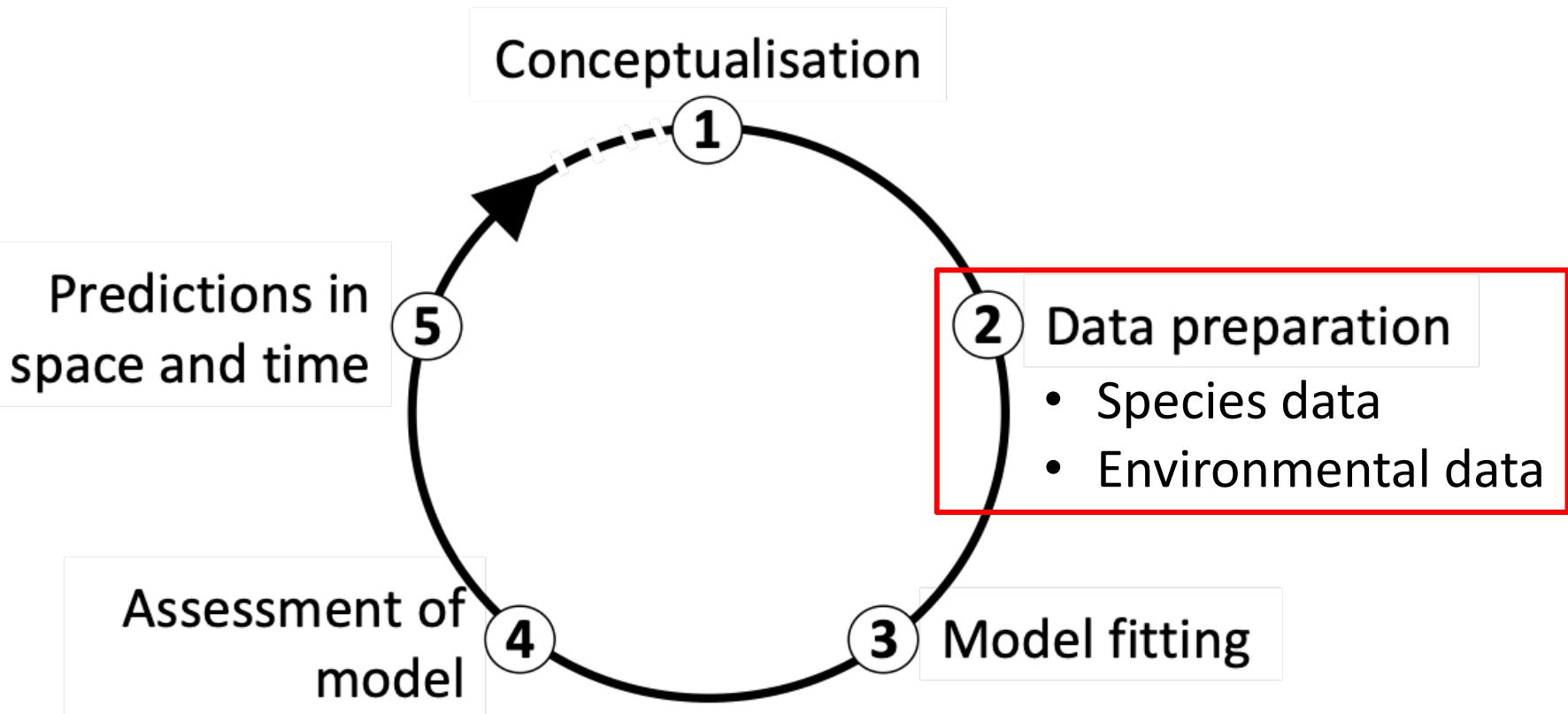
<https://damariszurell.github.io>

 @ZurellLab

Species distribution models



SDMs – model building steps

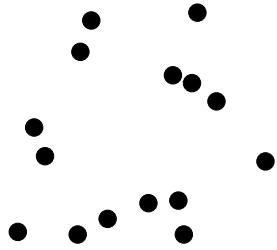


Species data types

Data type	Sampling Method	Taxonomic scope	Spatiotemporal scope	Output	Characteristics	Examples
Incidental records		?			<ul style="list-style-type: none"> • Single species • Spatiotemporally specific • Absences unknown 	<ul style="list-style-type: none"> • Museum records • Many amateur observations • GBIF points
Small-area inventories					<ul style="list-style-type: none"> • Multi-species • Spatiotemporally specific • Absences often reliable 	<ul style="list-style-type: none"> • Surveys with protocol, gridded atlas efforts • Relevé or forest plots • Visual, acoustic sensors • Trap/trawl-based surveys
Large-area inventories					<ul style="list-style-type: none"> • Multi-species • Spatiotemporally unspecific • Absences somewhat reliable 	<ul style="list-style-type: none"> • Regional checklists • National Park inventories
Expert synthesis maps					<ul style="list-style-type: none"> • Single species focus • Spatiotemporally unspecific • Absences somewhat reliable 	<ul style="list-style-type: none"> • Expert ranger maps, following manual or regional delineation

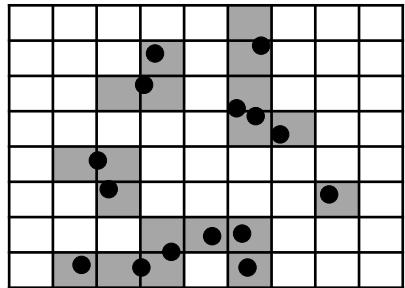
Species data types

Point occurrences



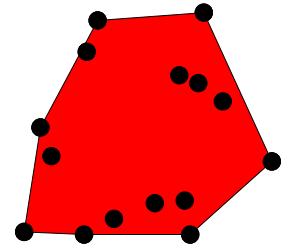
e.g. GBIF, OBIS
(Macroecology
practical 4)

Inventories: grid maps



e.g. UK breeding bird atlas
(Macroecology practical 3)

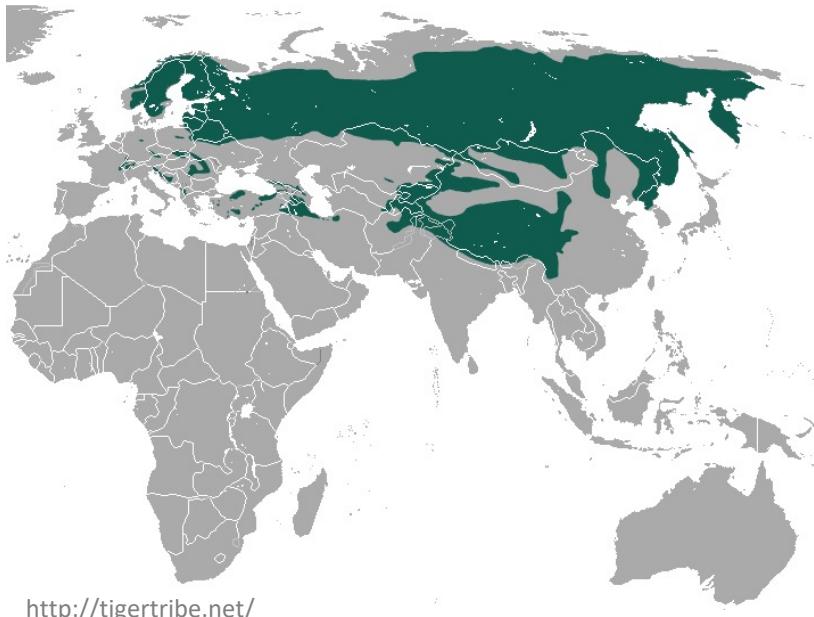
Range map



e.g. IUCN
(Macroecology
practical 4)

Expert range maps

- Expert-drawn outlines of species distributions
- Coarse-grained, with potentially high false positive rates
- Reliably indicate absence outside range boundaries



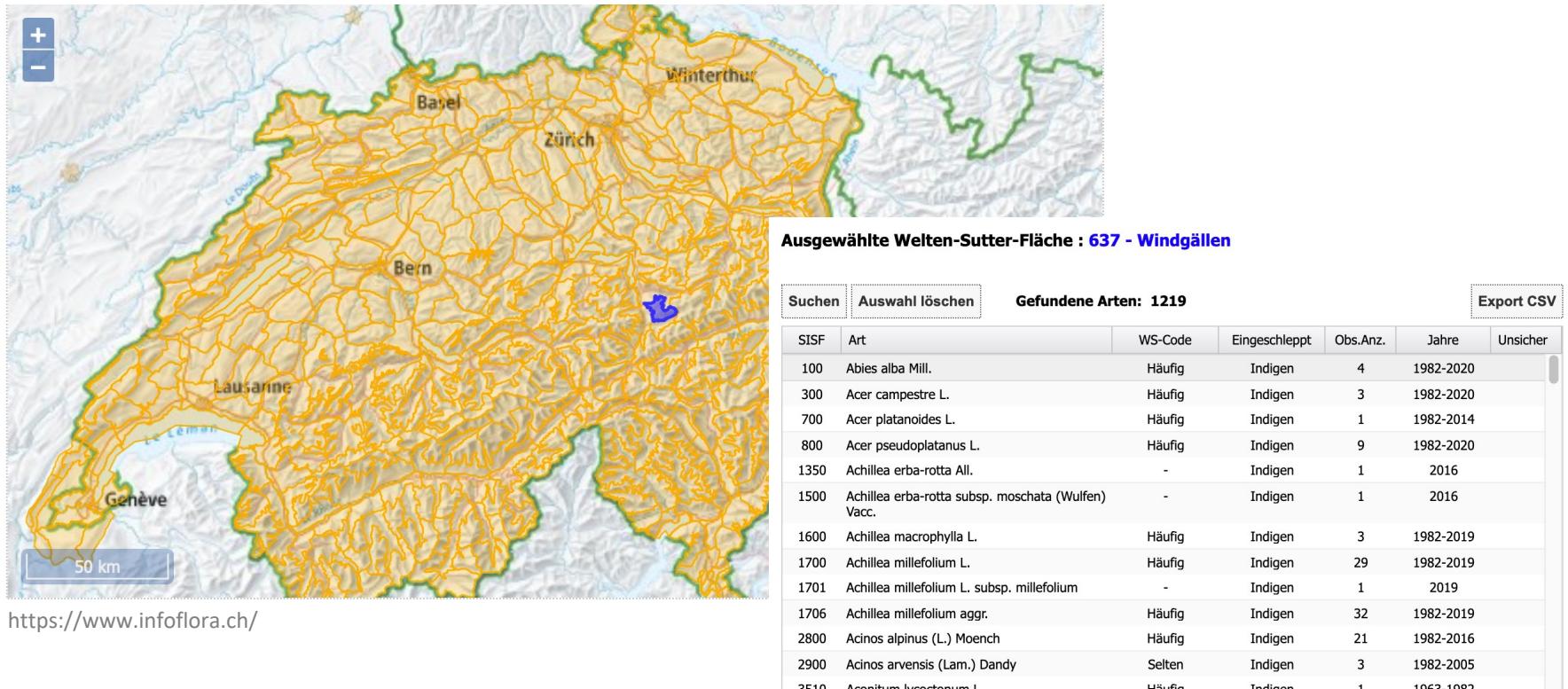
<http://tigertribe.net/>



<https://inhabitat.com/>

Inventories: regional checklists

- Indicate species occurrence within broad geopolitical, geographic, or bioclimatic regions
- E.g. Welten-Sutter plant inventory Switzerland



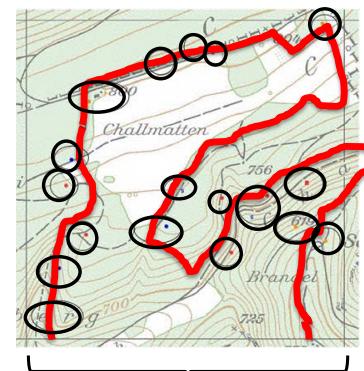
<https://www.infoflora.ch/>

Inventories: survey & atlas data

- comprise whole-taxon inventories of standardized search units (e.g. transects, field plots, atlas grid cells).



vogelwarte.ch



— Transect
○ Individual territory



Incidental records / opportunistic data

- Geographically localised specimen, field observation or tracking data
- Often biased
- E.g. Citizen science data, museum records



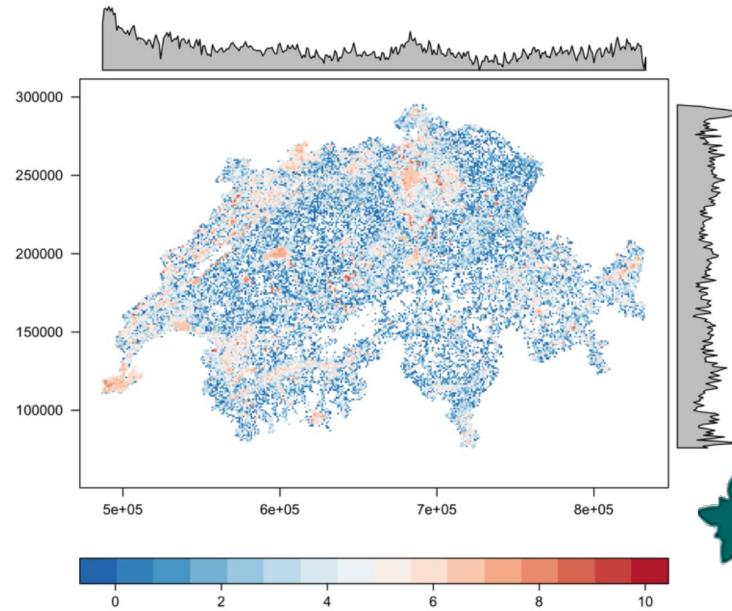
eButterfly



NaturaList



eBird



info flora

GBIF – global biodiversity information facility



Get data How-to Tools Community About

GBIF | Global Biodiversity Information Facility

Free and open access to biodiversity data

OCCURRENCES SPECIES DATASETS PUBLISHERS RESOURCES

Search

WHAT IS GBIF? ABOUT GBIF GERMANY

Torenia crustacea observed in Ngarchelong, Palau by Lin Scott. Photo via iNaturalist (CC BY)

Occurrence records 1,582,173,895 Datasets 53,924 Publishing institutions 1,637 Peer-reviewed papers using data 4,639

 News

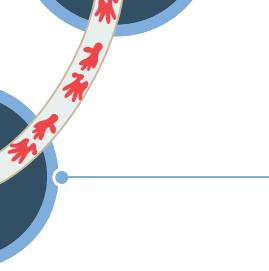
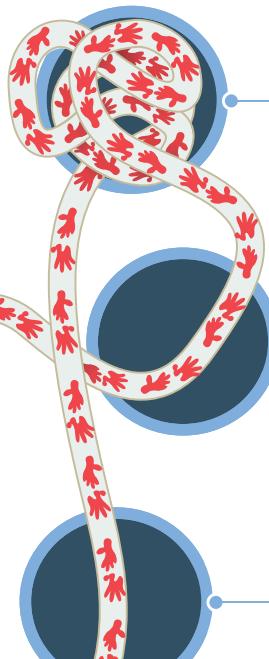
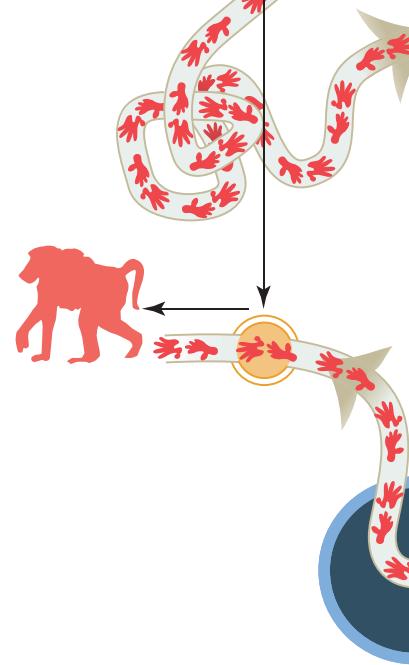
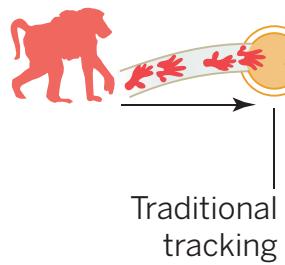
 News

 Data use

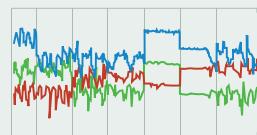
 News

Guide to publishing sequence- New data-clustering feature aims Protecting frogs of the future in BIFA programme awards funding

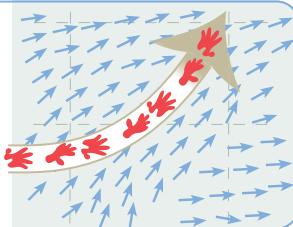
Animal tracking data



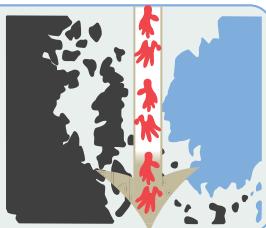
3D Accelerometer showing behavior and energy use



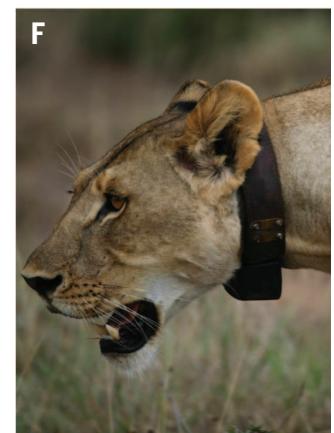
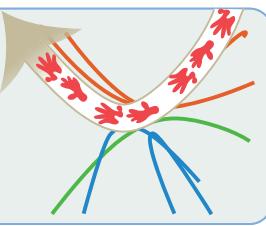
Remote sensing weather data



Remote sensing habitat data



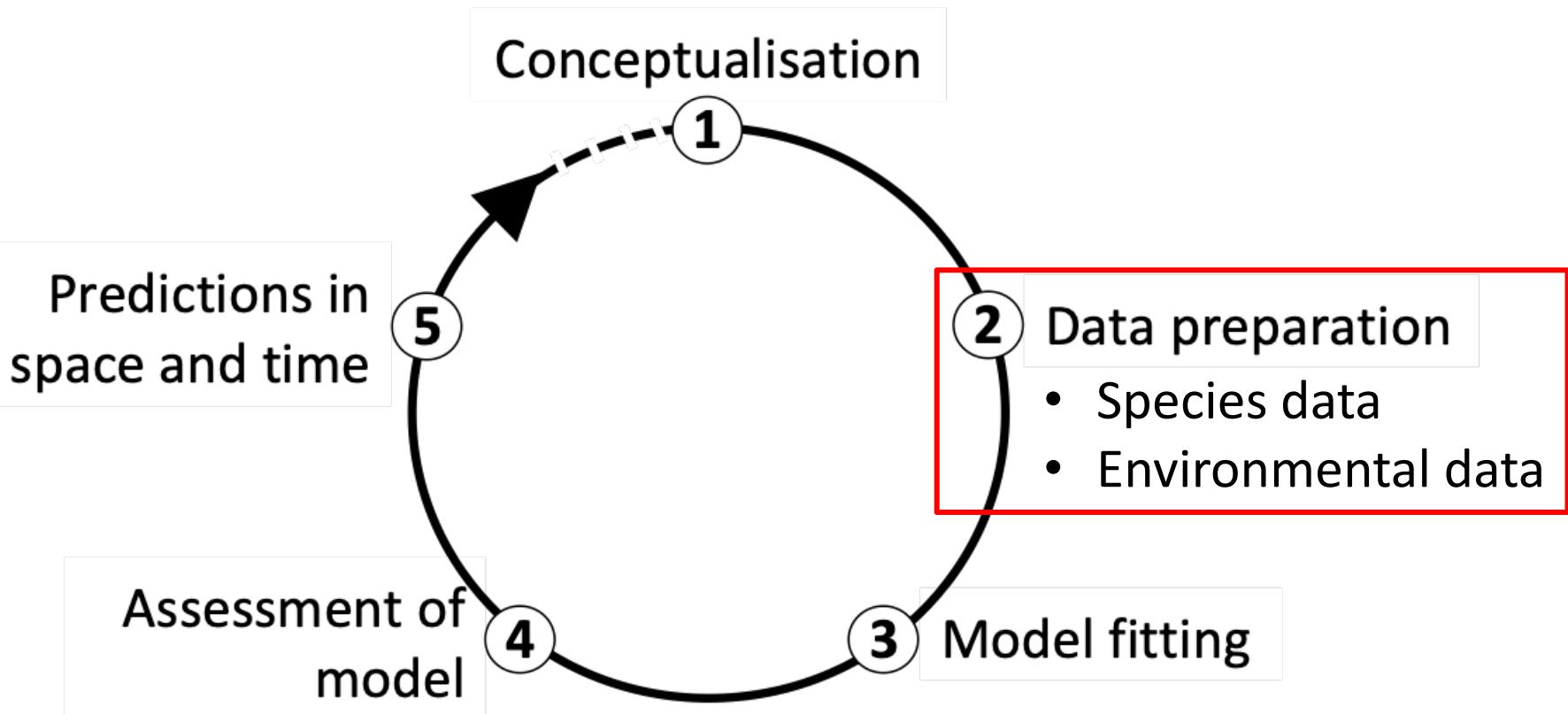
Interactions with other tagged animals



SDM and species data sources

- All of these data sources come with specific assumptions and biases
- All species data need to be checked for potential spatial biases, errors in coordinate specification, taxonomy, etc.
- If absences are unknown, adequate methods for deriving pseudo-absences or background data are required
- SDM methods and workflow need to be tailored to the data and the research question, e.g.
 - Simpler models (envelopes) for paleodata
 - Mixed models or step-selection functions for GPS data
 - Occupancy modelling for repeated surveys
 - ...

SDMs – model building steps

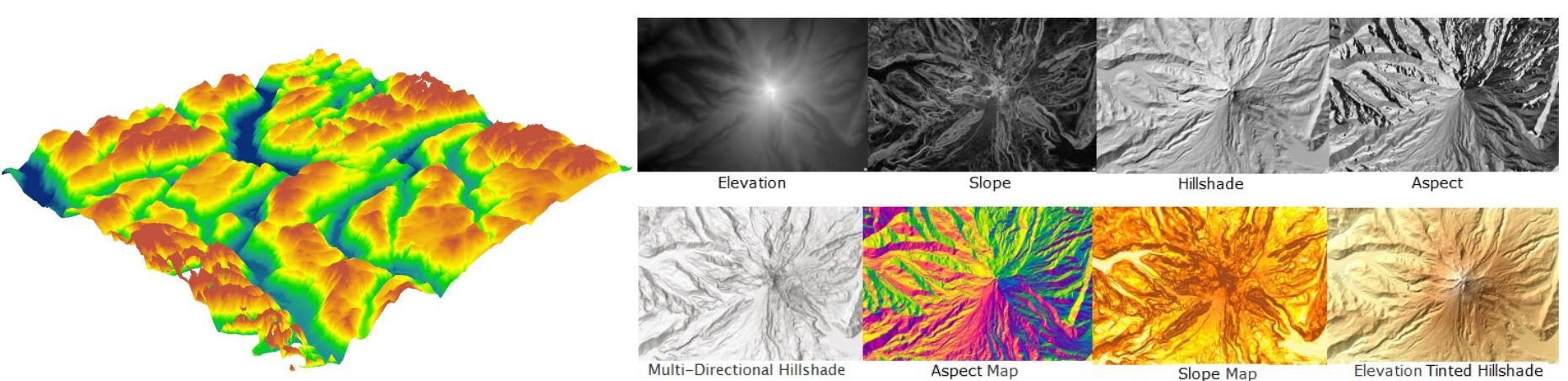


Typical environmental data in SDMs

- Digital elevation data and derivatives
- Climate data (incl. climate scenarios)
- Land use data

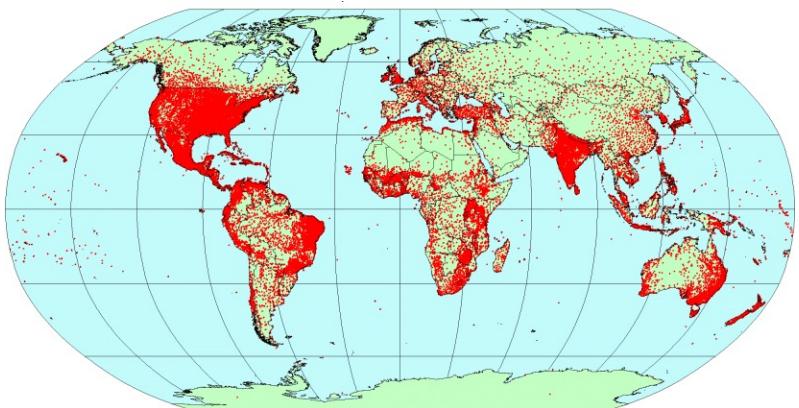
Digital elevation model (DEM)

- A three-dimensional representation of a terrains's surface
- Useful for deriving topographic predictors like altitude, slope, and aspect etc.
- Free global sources, e.g. SRTM (Space Shuttle Radar Topography Mission; 30 m)

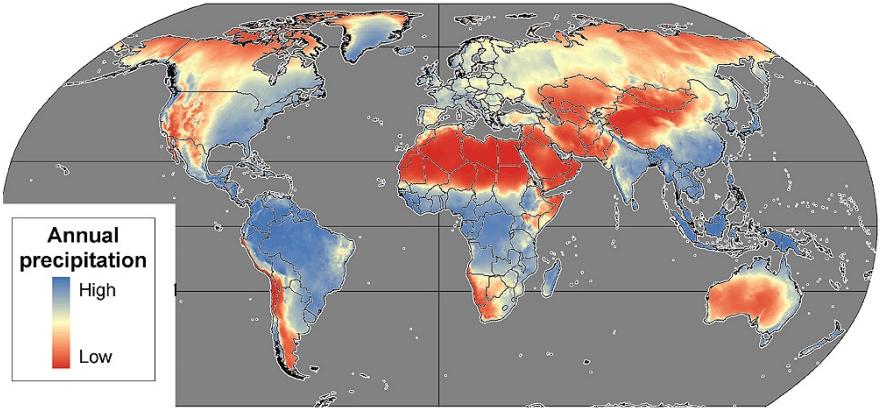


Climate data

- Spatial interpolation of climate values from long-term weather station data
- E.g. Worldclim, CHELSA
- Temperature and precipitation, and derived bioclimatic variables

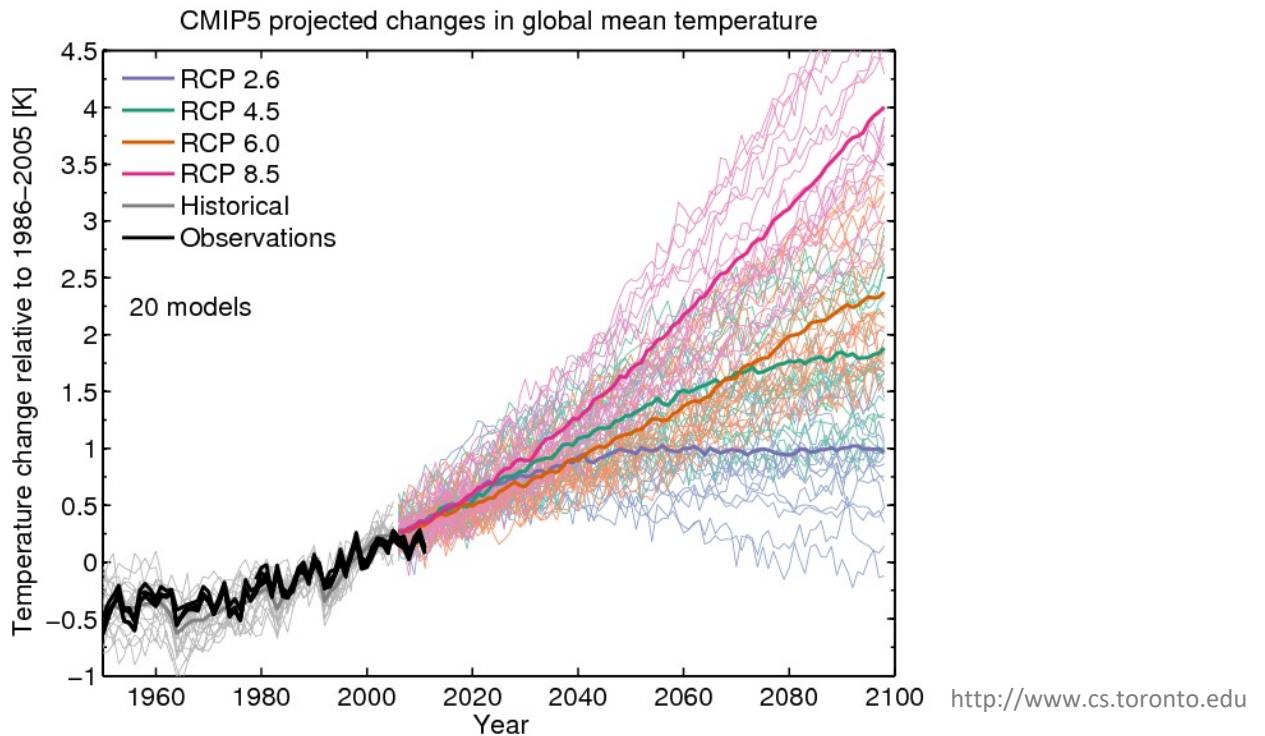


<https://databasin.org/>



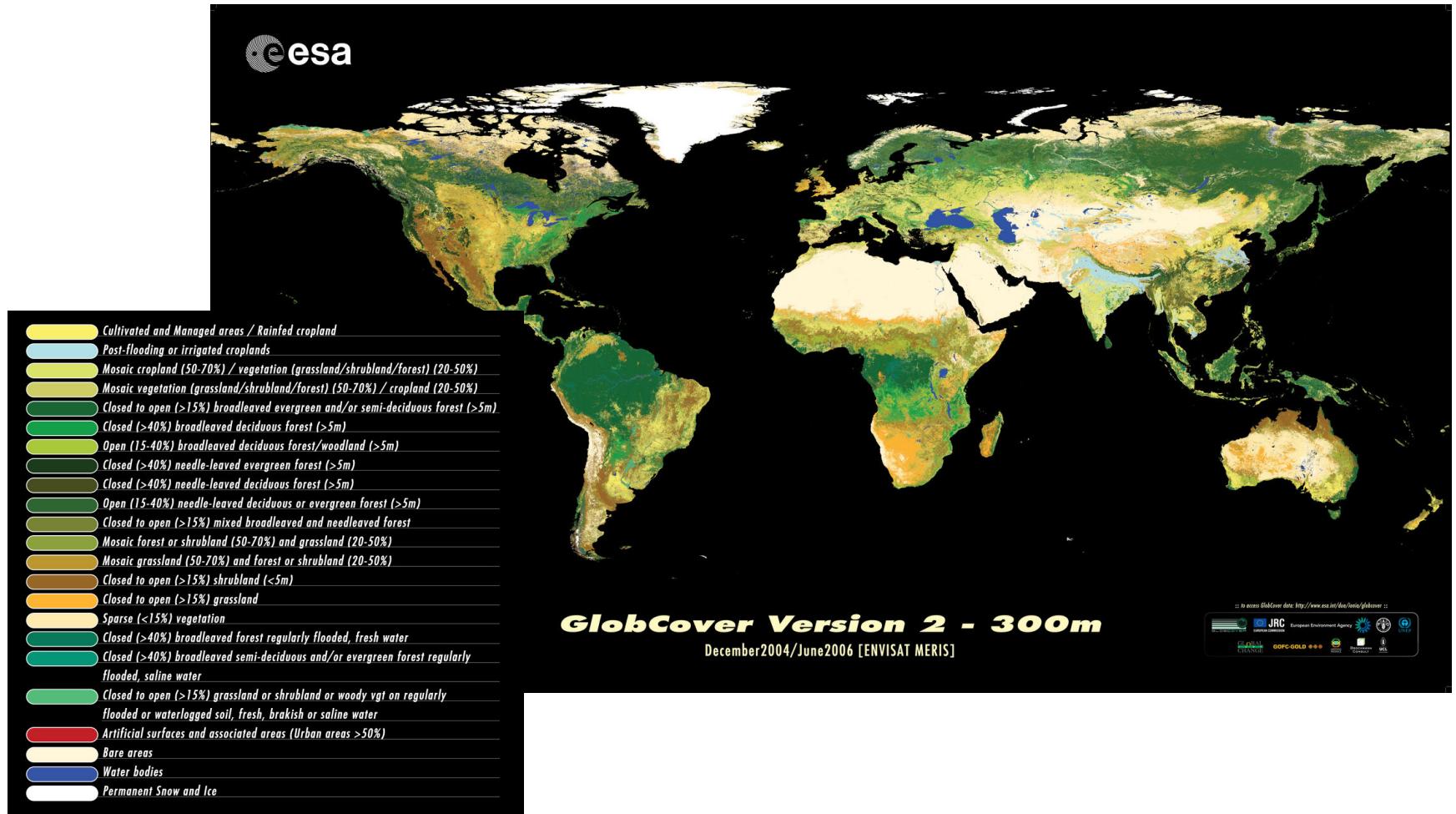
Climate scenarios

- Derived from global circulation models (GCMs)
- Based on different emission assumptions:
 - Representative concentration pathways (RCPs) describe the increase in radiative forcing 1850-2100 [W/m²]



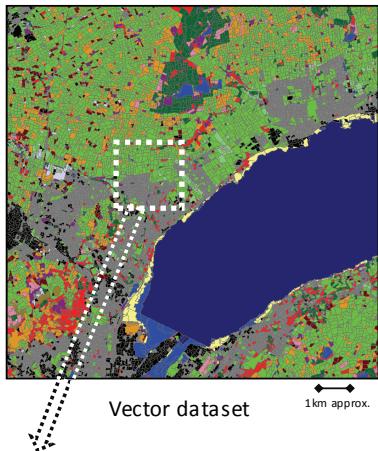
Land cover / land use data

- Global land cover maps derived from remote sensing



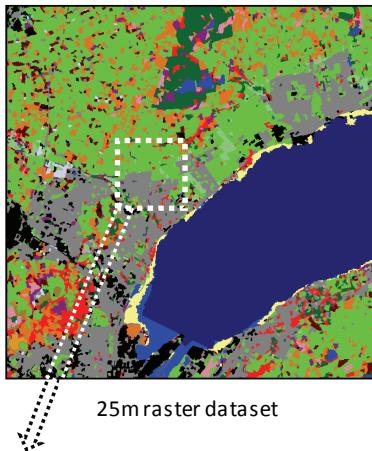
Land cover / land use data

- Local land cover maps derived from aerial images and habitat mapping

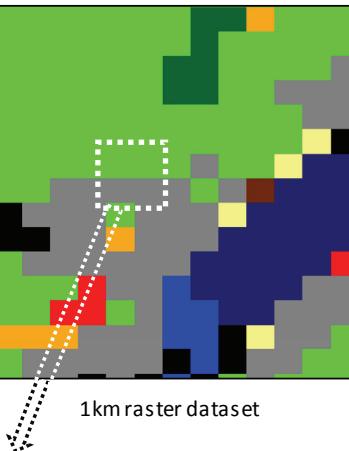


Vector dataset

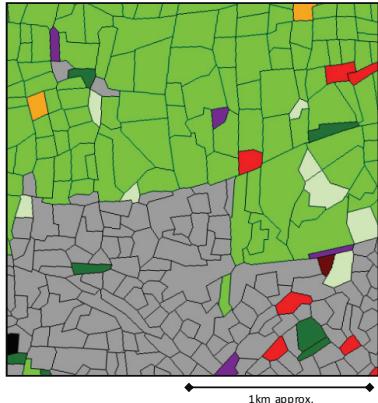
1km approx.



25m raster dataset

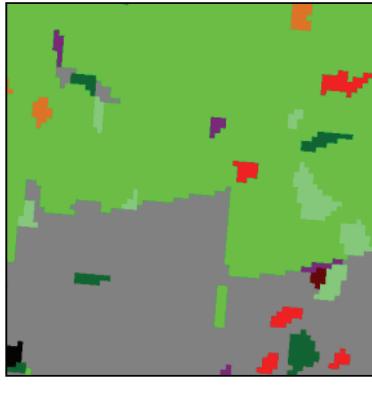


1km raster dataset

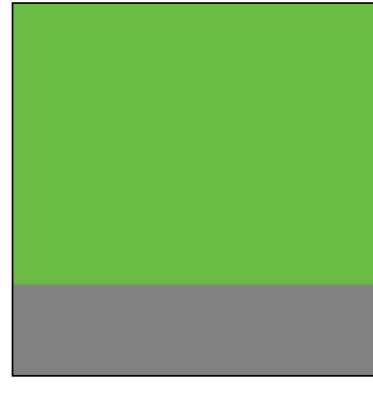


Vector product detail

1km approx.



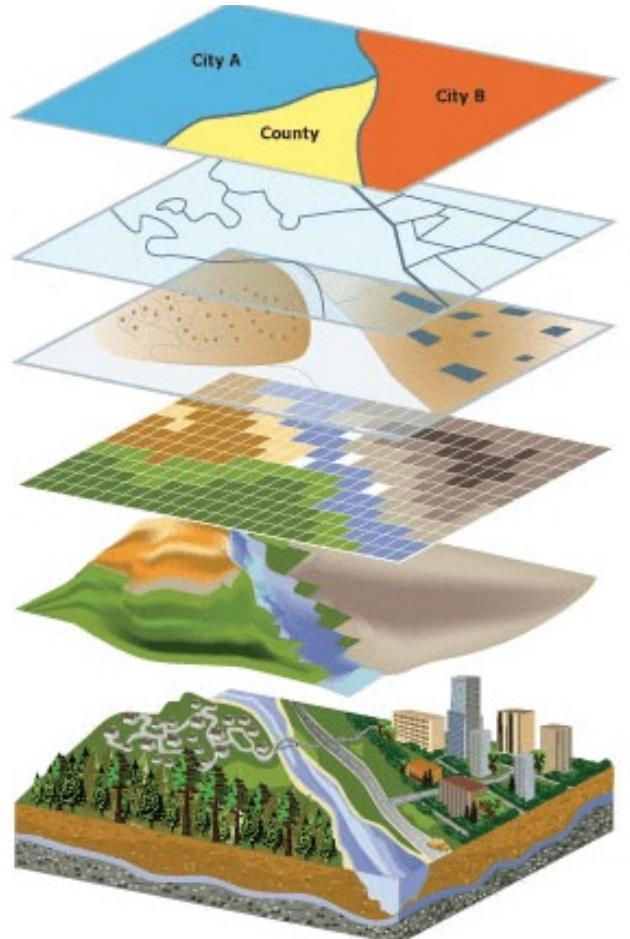
25m raster product detail



1km raster product detail

Environmental data in SDMs

- Obtaining and preparing an optimal set of environmental data requires GIS knowledge or collaborations



Additional resources for practicing

Practicals:

- Introduction to SDMs: <https://damariszurell.github.io/SDM-Intro/>
- Teaching resources (Master level) with more details on single SDM steps: <https://damariszurell.github.io/EEC-MGC/>

Lectures:

- International lecture series (recorded) from experts on ecological niche modelling:
<https://docs.google.com/spreadsheets/d/1RQu1XRKyYfrnFI2V1g677d0sf8tFx2xUvb96cbP02s/edit#gid=0>

Thank you for your interest



Contact:

Damaris Zurell

Ecology & Macroecology
University of Potsdam

<https://damariszurell.github.io>

Email: damaris.zurell@uni-potsdam.de

 @ZurellLab