<https://www.dropbox.com/s/46s6x2tg16be3k1/SwissBreedingBirdAtlas_2013_2016.pdf?dl=0>

tutorial

<https://damariszurell.github.io/SDM-Intro/>

birds

* one species list per km^1 (no repetition
* about 100-2000 1km^2 squares
* Some species will only have presence, others will have counts
* Absence will not be explicit, but if a species doesn't appear in a list, then we can treat is as observed absence (but we have to add the info the way we need it ourselves).
* Data has coordinates
* A few dozen species
* Excel- oder Text-File kommen als "Positiv-Liste"

Environmental data

* 16 they have, but wel’ll see
* might have to get it from elsewhere
* tutorial – download data

I asked about receiving the finest resolution, too, and we could have that, but then we should aggregate ourselves. Marc said that the squares are pretty spread across CH and that we could probably even use that for a model based only on coordinates. I asked for the 1km^2 data, but I can ask to change that if you prefer. What do you think?

The person responsible for environmental data might be on holidays atm, so we might not receive them before next week. It would be easier if we were happy with covs for the 1km^2 areas of bird data, but I'm negotiating to see if they would be happy to share some variables for the whole of CH, or at least for other areas, so we can extrapolate. Otherwise, I guess we should be able to find something on worldclim, official swiss land cover, and elevation, and that should get us somewhere. In that case, it might have to be part of the deal that people spend a bit more time on data gathering an preparation, but that's generally not a bad skill to train I think, and we can help.

I think the tutorial is great for what we're doing, especially if we need to get data from elsewhere. I think different types of model algorithms available for SDMs would benefit from more input, but I assume that will be well covered in your or Mollies lecture? Marc will cover GLMs, too. (They mention machine learning at some point in the tutorial, which I found a bit unnecessary, but that’s not a problem.) I can tell you more about the data that was used there and the one we’re getting when we meet.

Best wishes,

Leila

Liebe Leila,  
  
habe mit Kollege Hans Schmid geredet, und Ihr werdet da sicher ein paar schutzmaessig unproblematische Vogeldaten bekommen. Hans wird bereits eine Auswahl machen, sodass pro 1km2-Quadrat nur eine Liste kommt (also keine Wiederholungen), die seltenen Arten sind draussen (also eben keine Schutzprobleme), und Ihr werdet noch immer **mehrere Dutzend Arten** haben, von denen die meisten nur Praesenz/Absenz haben, einige Arten aber auch Zaehlungen haben. Also Moeglichkeit fuer Poisson- oder aehnliche Modelle, neben den ueblichen Bernoulli-Modellen.  
  
Die Daten werden als Excel- oder Text-File kommen als "Positiv-Liste" an jedem Ort (= km2-Quadrat), d.h., es werden nur die beobachteten Arten drin sein. Die Methodik (wozu wir Euch auch einen kleinen Text liefern werden) laesst aber den klaren Umkehrschluss zu, dass eine Art, die nicht auf einer Liste an einem Ort ist, dort nicht beobachtet worden ist. Somit koennt Ihr aus dem Gesamt-Set der Orte die Nuller-Daten eruieren; das muesst Ihr dann aber selber machen.  
  
Zu den Habitatsdaten habe ich noch keine wirklichen Neuigkeiten. Der Kollege, der damit vor allem zu tun hat (und auch ueber legale Aspekte mehr Bescheid weiss), ist evl diese Woche noch in den Ferien. Es ist aber sicher so, dass es unproblematischer waere, wenn Ihr Habitatsdaten nur von denjenigen Orten mit Vogeldaten kriegen wuerdet (das werden wohl 1000-2000 Orte verteilt ueber die Schweiz sein). Fuers Modell-fitten ist das dann OK, einzig fuer die Extrapolation (= map-making) geht das dann nicht, weil Ihr dazu ja auch die Umweltdaten fuer die nicht besuchten Orte benoetigt. Wie seht Ihr das: wollt Ihr Schweizer Karten machen ?  
  
Viele Gruesse  --- Marc

Hi Leila,

It's probably simplest to aggregate the data to the same spatial resolution as the covariates, so 1km^2 would be great. That being said, if we were interested in detailing a spatial model based only on coordinates, we could use the highest resolution available. Will coordinates be provided with the data?

I came across a useful tutorial on SDMs over the weekend. Here is the [link](https://damariszurell.github.io/SDM-Intro/" \t "_blank). Interestingly, it uses an example from the Swiss breeding bird atlas for years 1993-1996. It provides a nice conceptual overview of SDMs and then runs through a simple example using both a glm and random forest. The tutorial also mentions key SDM references throughout. Would this tutorial or something similar be useful to provide to the Workshop participants?