**Simple algorithms to display ignorance maps of raw distributional data accessed from species observation databases: the Swedish LifeWatch implementation**

Author: Alejandro Ruete, for the Swedish LifeWatch

Swedish Species Information Centre, Swedish University of Agricultural Sciences (SLU). P.O. 7007, SE-750 07. Uppsala, Sweden.

E-mail: [aleruete@gmail.com](mailto:aleruete@gmail.com)

*Keywords*: citizen-science data, open-access biodiversity database, presence-only data, sampling effort, spatial bias, species distribution model, temporal bias

This interactive application exemplifies the use of three different algorithms to map the ignorance (i.e. bias and lack of sampling effort) found in the observations accessed through the Swedish LifeWatch biodiversity analysis portal [www.swedishlifewatch.se](http://www.swedishlifewatch.se).

Examples are provided for seven reference taxonomic groups (i.e. Amphibians, Birds, Butterflies, Land Mammals, Harvestmen, Dragonflies, and Vascular Plants) as groups with different sampling effort intensity and extension as well as different number of amateur observers reporting observations to the database. For each group we show examples of a common (or widely distributed) and a rare (or locally distributed) species. Note that in many cases common species are not reported as enthusiastically as more “interesting” species are (in Sweden, a moose is not as frequently reported as hedgehogs are). Therefore, one can expect that for common species the observations are very scattered and they do not cover the whole expected range of the species.

**Tabs**

You are now reading under the “Read Me” tab. The data is shown the form of “Maps” and “Data plots”. Four maps of Sweden with a resolution of 10 x 10 km are shown under the tab “Maps” displaying:

1. the number of observations per pixel for the reference taxonomic group selected,
2. the ignorance scores per pixel,
3. an estimate of pseudo-absence (henceforth PA) for the selected target species. PAs are estimated with the same algorithms explained to produce the ignorance maps. However, the results have slightly different interpretation. For individual focal species, the lack of observations could represent a true absence or lack of observers. For a reference group the number of observations (or lack of them) are assumed to represent the ignorance we have for that grid cell.
4. and the presence map (1-PA) . This is an example of the use of ignorance maps to weight and mask different estimates for target species. In this case, the black to white scale represents the knowledge on the species presence. On top, a layer multiplying 1-ignorance (i.e. certainty) by the PA (red) or P (green) shows areas where a minimum certainty for PA or P is achieved after weighting. Then in red are shown areas where there is low ignorance but the target species has not been found, while in green are areas where the target species has been found but a minimum certainty is required.

The outline in those maps is a 10km buffer around the Swedish land surface.

Under the “Data plots” tab you find

1. a density plot of the number of observations per grid cell,
2. a species discovery plot and
3. curves of the transformations of the number of observation into ignorance scores (note that this later plot is reactive to the value set for *O*0.5).

**Options**

On the left panel you find options for the Reference taxonomic group and Target species to be displayed. First, you can opt to use the raw observations as a measure of sampling effort or the Observation Index, a sampling effort relative to the number of species sampled in a particular grid cell. On each subpanel you can set the algorithm can be set and O0.5 (the number of observations that are enough to reduce the ignorance score to 0.5) for the Inversed algorithm. For target species, a Step algorithm is added where any observation number below O0.5 sets the ignorance score to 0. Else the Step algorithm is equal to the Inversed algorithm.

Finally, on the subpanel for Visualization options you can set the minimum certainty required for pseudo-absences and presence scores to be highlighted in red and green, respectively, and the transparency of these layers.