Suomen ympäristökeskus
Finlands miljöcentral
Finnish Environment Institute

METSÄHALLITUS
FORSTSTYRELSEN
MEAHCIRÁÐĐEHUS

A1.2 Deliverable - Data management plan for updating information from field inventories to model development and deliverable - Plan for model operationalization and model development

19.12.2022

Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.

BIODIVERSEA
for marine nature

LIFE20 IPE/FI/000020 LIFE-IP BIODIVERSEA

Life

NATURA 2000

Louise Forsblom & Elina Virtanen

Finnish Environment Institute (Syke)

with contributions from Metsähallitus Parks and Wildlife Finland

Data management plan

Current status

This section describes the flow of underwater diversity data from the field all the way to modeling as related to the data management plan described in section A1.2 (Deliverable - Data management plan for updating information from field inventories to model development).

The underwater biodiversity inventory program, Velmu, has been mapping underwater biodiversity since 2004. The collection of data is described in more detail in the Velmu Method Guide 2022. Parks and Wildlife Finland adds all the collected data to the LajiGIS database that they maintain.

In 2021, when Syke updated species distribution models using Velmu data and described in Virtanen et al. 2018, Syke's internally maintained version of the Velmu data. In addition to the 170 000 observations gathered in Velmu's mapping efforts, additional data was used from Sykes's bottom fauna database POHJE, endangered species monitoring sites stored in LajiGIS database, as well as additional data contributed by Åbo Akademi and Ålands Landskapsregering, and separately retrieved data on alien species retrieved from the Finnish Biodiversity Information Facility (FinBIF) (Figure 1). The quality of the data retrieved from different data sources varies; for example, the data from Velmu and POHJE also contain data on instances where species have not been observed (verified zero observation), whereas citizen observations only provide data on observed species.

Data flow

Currently, the availability data is affected by restrictions on data retrieval in the user interface of LajiGIS, which limits the search to a certain number of rows. The data for each species is printed on its own row (long format) when searching for the material, which means multiple rows for each unique spatial location if there are multiple species. Velmu's data alone includes more than 170,000 observations, so the entire data cannot be obtained with a single search and there is a need for a separate request for information. Considering the plan for operationalizing the modelling effort, it would be sensible to design a solution for this, but due to the changes expected to the user interface, it is not effective to tackle this at the moment. In the future, we aim for the most agile and automated arrangement possible, where both Parks and Wildlife Finland and Syke use the same version of the species observations, by doing a joint data requested at least once a year from the main user of LajiGIS.

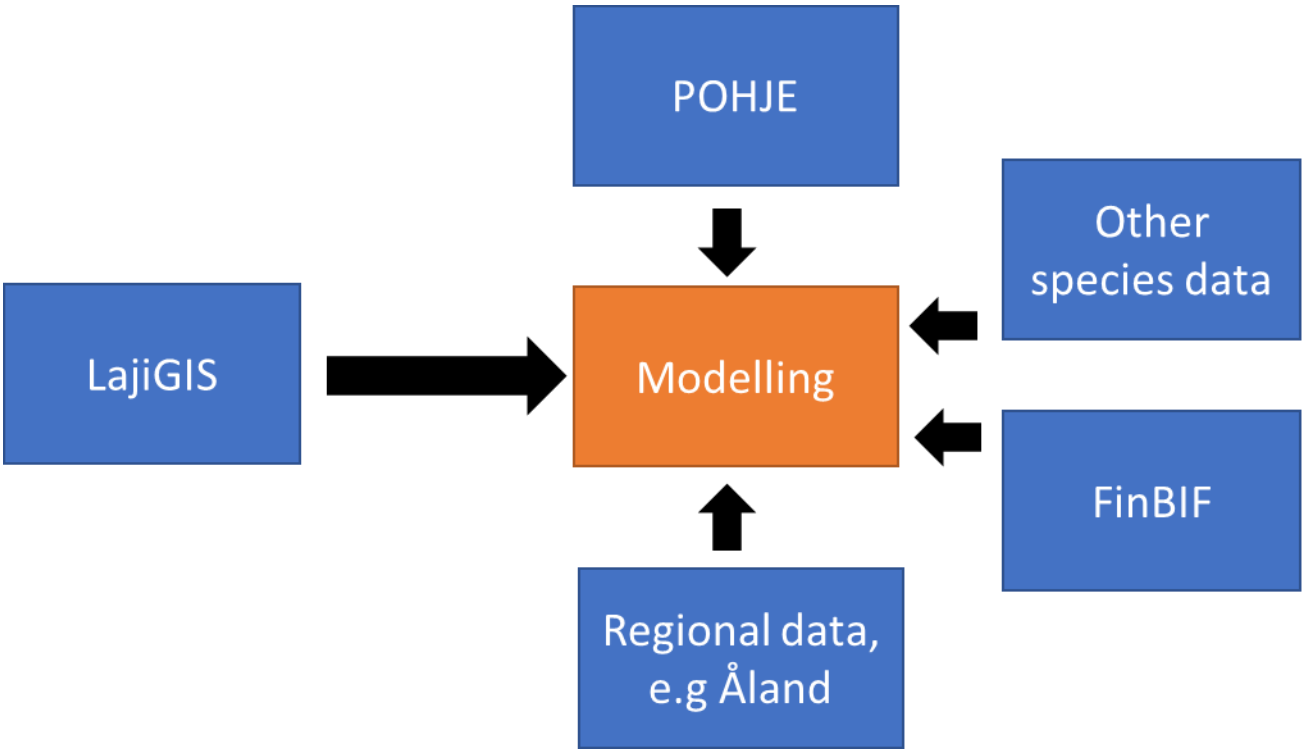

Figure 1. The flow of species data for modelling from different sources of information.

Operationalization

This section describes the preliminary plan for the operationalization of modelling in relation to the objectives described in section A1.2 (Deliverable - Plan for model operationalization and model development).

Within the framework of the Biodiversea project, the operationalization of modelling is carried forward with the aim of achieving a higher degree of operationalization by the end of 2025. This is a continuation of the work that has been going on in Velmu. Figure 2 describes the modelling workflow from the main types of datasets, through modelling to the final products. In order to keep the work agile, the operationalization of the coming years will be carried out in parts so that all sub-areas (Figure 2 boxes) will be operationalized separately if possible.

The goal is to significantly reduce the number of steps to be done by hand, but because the chain from field observations to final model output is long, it is not possible to automate everything. Reducing the amount of manual work helps combat the amount of errors. The automation also aims to make the modelling repeatable, which also increases the transparency of the whole process. The development must also take into account changes in different databases that can be hard to anticipate exactly, for example, the future renewal of the LajjGIS database. Ideally, by the end of 2025, all the points described in Figure 2 will be at least in such a format that each separate block will have its own custom script that, with a specific data input, that produces the desired result; be it, for example, a pressure layer or a species model.

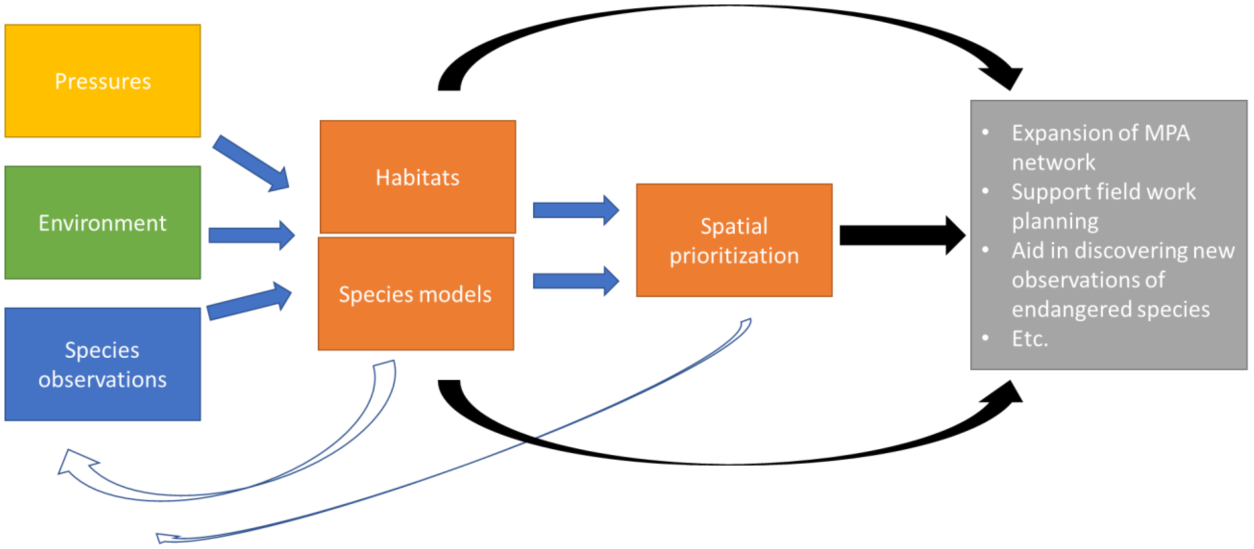

Figure 2. Different stages of modeling. The left-hand panes reflect the material needed for modeling, and the modeling itself is depicted in the middle (orange). The final products are depicted on the right (grey).

Specific objectives and concrete measures for the separate sections are set out in the following section. In general, all sections aim for modeling, which allows the current 20 x 20 m modeling resolution.

Species observations

Data sources for species observations and the flow of data from the field to modelling are presented in the data management plan (Figure 1). The management and editing of species data is currently implemented using R-scripts, which makes data management repeatable. Operationalization is furthered by creating a script for each source of species information, so that the observations are modified to the necessary format for modelling. New sources of observations are also being considered so that observations reported by citizens are better included.

Action: I) update existing species scripts to support the LajiGIS data format, II) develop a script-based solution for information retrieval from the FinBIF service, III) update other scripts as needed.

Environment Variables

Information on the state and characteristics of the environment is needed as explanatory variables for modelling the distribution of species. Particularly important variables for many aquatic species are, depth, bottom substrate, salinity and exposure.

The needs for updating environmental variables most commonly used for species models have been assessed and the related plans are presented in the table.

Table 1. Variables used for species modeling, the way they are updated and the need for development, and the time interval for the update. Layers under development (x).

Variable	Update procedure and development needs	Update interval	Dev.
Surface and bottom water salinity	Script	At the same time as species models	
Water temperature	Script	At the same time as species models	
Sea surface temperature	Evaluate satellite-based methods		x
Bathymetry	Sentinel2 based bathymetry model	As needed	
Bathymetric position index	Produced using ArcGIS	When bathymetry is updated	
Depth exposure	Based on depth, Bekkby et al. 2008	When bathymetry is updated	
Slope	Script	When bathymetry is updated	
Topographical Shelter Index		When bathymetry is updated	
Ruggedness of the seafloor		When bathymetry is updated	
Seafloor fetch	Script	When bathymetry is updated	
Surface exposition		When bathymetry is updated	
Total phosphorous and nitrogen	ArcGIS, has manual steps		
Turbidity	Satellite based		x
CDOM	Satellite based		x
Secchi	Satellite based		x
Oxygen conditions	Script	At the same time as species models	
Substrate	Script	At the same time as species models	
Sand	Evaluate satellite-based methods	At the same time as species models	x

Pressures

This section takes into account all direct or indirect effects of human activities on the species. As a rule, new pressure layers are produced in the A7 action and the pressure layers produced there aim at the highest possible level of automation. Automation of previously produced levels will also be advanced.

Pressure	Plan	Action
Climate change scenarios	Climate change scenarios are used to when modelling species distributions. Scripts used for species modeling are modified so that the alignment of scenarios is as straightforward as possible.	A7

Identification of human pressures	The purpose is to operationalize the identification of human pressures, especially dredging, identified manually from aerial photographs so that their detection and the identification of their extent would take place automatically.	A7
The distribution of reeds	Remote sensing methods are used to automate the identification of the distributions of reeds.	A7
Mapping of small scale boating	The identification of small boats by remote sensing methods is automated.	A7
User Generated Materials and Social Media	Piloting new pressure layers based on user generated materials and social media.	A7
Pressure models	Pressure modeling is automated (where possible) to be script-based	A7

Species and Habitat Models

All species and habitat models are adapted using data covering the whole of Finland and their results are predicted for the entire Finnish sea area as 20 x 20 m rasters. The fitting of the species models and the subsequent predictions of the distributions are largely automated.

Spatial prioritization

Zonation 5 (Moilanen et al. 2022) scripts for the CSC Puhti environment will be produced for each spatial prioritization product needed (e.g. biodiversity concentrations, degraded areas, sites to be restored). The aim is that the prioritizations can be run almost automatically as the input data are updated. As the number of input data increase or changes, the weighting needs to be updated and that requires expert work.

References

The Finnish Inventory Programme for the Underwater Marine Environment (VELMU) Methodology guide 2022 Version 14.2.2022

Virtanen, E. A., Viitasalo, M., Lappalainen, J., & Moilanen, A. (2018). Evaluation, gap analysis, and potential expansion of the Finnish marine protected area network. *Frontiers in Marine Science*, *5*, 402.

Moilanen, A., Lehtinen, P., Kohonen, I., Jalkanen, J., Virtanen, E. A., & Kujala, H. (2022). Novel methods for spatial prioritization with applications in conservation, land use planning and ecological impact avoidance. Methods in Ecology and Evolution, 13(5), 1062-1072.