## Spatial datasets online validation tool

**ReMAP project**

21.02.2025

Andžej Miloš, Deborah Shinoda

1. **Validation tool**

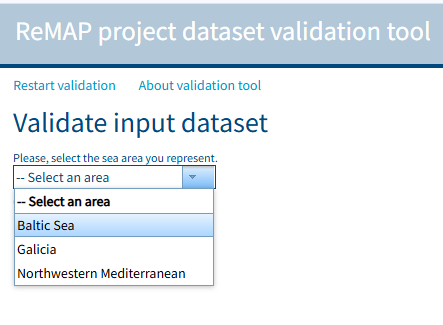
Spatial datasets online validation tool is a secure web application. Registered users can log in, choose a dataset to validate, upload the dataset for validation, see the validation report. If dataset is valid (according to validation rules) – upload dataset to the database. If there are errors – no upload is possible.

1. **User interface**

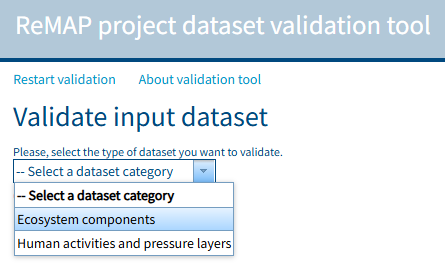
The tool was developed for the ReMAP project, in which there are 3 geographic study areas for which spatial datasets are validated. The tool can be adjusted for any geographic areas.

The tool is implemented as one page web application. User performs the following steps to validate a dataset:

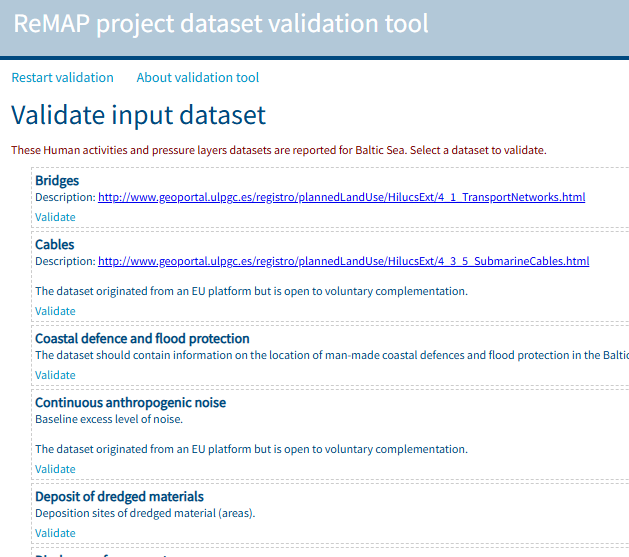
* Select a study area from the predefined list



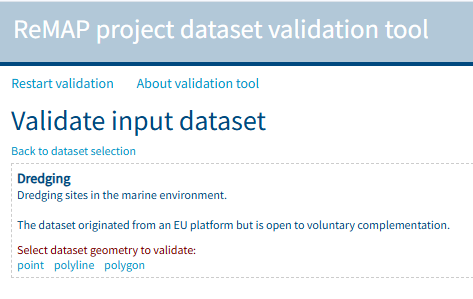
* Select a dataset category from the predefined list



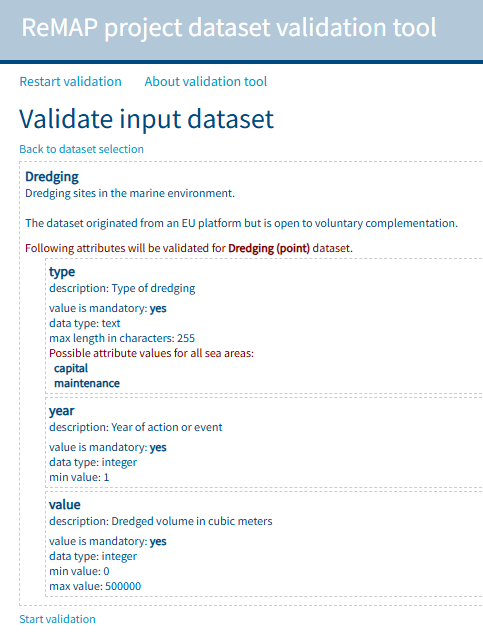
* Select a dataset to validate



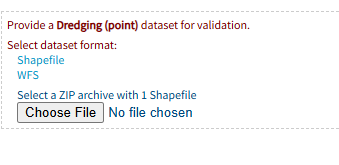
* Select a geometry of the dataset



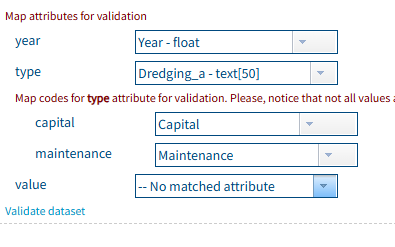
* The list of attributes that will be validated is displayed at this step. Each attribute has a set of validation rules. Users can start validation process.



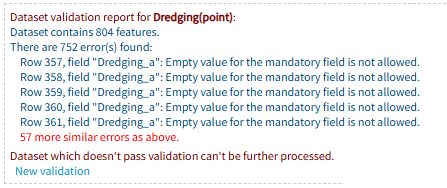
* Select input dataset format and upload a dataset



* Perform attribute and value mapping



* Start validation and check the validation report. If no errors – dataset can be uploaded to the database, otherwise data errors has to fixed and validation process repeated.

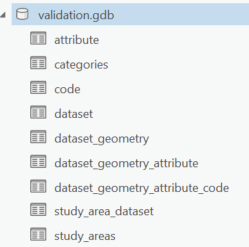


The validation tool was created with Javascript and ArcGIS API for Javascript 3. The source code is available in the *web-source-code* folder.

1. **Validation rules**

In the ReMAP project it was defined beforehand which datasets will be collected and validated for each study area. It was also defined what spatial representations are possible for each dataset. There were defined attributes, their constraints and possible values if any for each dataset.

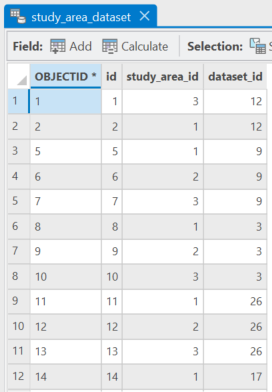
Having in mind above a database (*validation.gdb*) with validation rules was created. Validation database contains a few tables with study area, dataset, attribute, code name and their relations:



The relations define for example:

* what dataset is available for the study area;
* what geometries (point, polyline, polygon or raster) are possible for dataset;
* what attributes are present for dataset
* what possible values (codes) are possible for attribute.

Below is a “study area – dataset” relation example:



For the ReMAP project all these tables and relations were created and filled manually.

1. **Backend**

All tables and relations were published with ArcGIS Server and made available for the validation tool through ArcGIS Server REST interface. When user selects options in the validation tool interface, queries are generated and sent to the tables and appropriate results are returned and displayed in the validation tool.

Additionally, some Python tools were developed and published with ArcGIS Server. Tools are also available for the validation tool through ArcGIS Server REST interface. These Python tools are:

* Get attribute name from the uploaded Shapefile;
* Get attribute name from the WFS;
* Get unique attribute values;
* Validate dataset;
* Upload dataset.

Python source code for the tools is available in the *python-source-code* folder.

The URLs to all tables and tools available in ArcGIS Server REST are listed in the *web-source-code/1.0/config/config.json* file. The services are secured.

For each study area and data category an output file geodatabase was created. Each output file geodatabase contains empty datasets with the required attributes for that dataset. The empty datasets with attributes were created with the Python script (*python-source-code/createEmptyDatasets.py*).

Anytime the data upload is performed with the validation tool, valid data are appended to appropriate dataset.

All Python scripts can be run from the ArcGIS Pro Toolbox *tools.atbx*.

1. **Software requirements**

The online validation tool is a Javascript web application. Any web server is suitable to deploy the web application. Currently it is deployed on the IIS web server.

The ArcGIS Server is required for publishing validation tables and Python tools.

Python tools are created using arcpy and a few other Python site packages. Toolbox was tested in ArcGIS Pro 3.1.x and 3.3.x environments on Windows OS.