



Presentation of final solution

– *LifeWatch showcase*

DA P6 Climate Change Data Challenge Application:
Göteborg University, Department of Marine Sciences



Title

Biodiversity Virtual Laboratory – A platform for integrating disparate data and predicting ecological responses to climate change

Authors

Matthias Obst¹, Gunnar Anderson², Alan Williams³, Francisco Quevedo Fernandez⁴, Anders Lindström², Finn Bacall³, Alex Hardisty⁴

Affiliations

¹Göteborg University, Department of Marine Sciences; ²Swedish National Veterinary institute; ³School of Computer Science, University of Manchester, Manchester, UK; ⁴School of Computer Science and Informatics, Cardiff University, Queens Buildings, 5 The Parade, Cardiff CF24 3AA, UK



Demonstration of the BioVeL Infrastructure

In the following presentation we demonstrate the seamless integration of the CORINE land cover dataset into BioVeL's ecological modelling environment.

The BioVeL virtual laboratory (www.biovel.eu) is a platform for executing Taverna workflows. These workflows orchestrate complex and distributed tasks for data access and analysis.

In our demonstration we use workflow technology to connect the CORINE data set with climatic data and carry out scalable and semi-automated analytical cycles to predict the spread of mosquito-borne diseases under various climate scenarios.

The demonstration features 3 principal steps:

Step 1: The CORINE dataset was customized to the needs of the scientists

Step 2: The customized data set was made accessible in the workflow environment

Step 3: A number of predictive modelling experiments were performed using the CORINE and climatic datasets

Step 1: Customizing the dataset

The CORINE land cover dataset was transformed according to the needs of the scientists. This resulted in 44 individual data layers (see below) with continuous variables for the various land types. All transformation was done using ArcGIS.

Layer no.	LABEL1	LABEL2	LABEL3
1	Artificial surfaces	Urban fabric	Continuous urban fabric
2	Artificial surfaces	Urban fabric	Discontinuous urban fabric
3	Artificial surfaces	Industrial, commercial and transport units	Industrial or commercial units
4	Artificial surfaces	Industrial, commercial and transport units	Road and rail networks and associated land
5	Artificial surfaces	Industrial, commercial and transport units	Port areas
6	Artificial surfaces	Industrial, commercial and transport units	Airports
7	Artificial surfaces	Mine, dump and construction sites	Mineral extraction sites
...
44	Water bodies	Marine waters	Sea and ocean

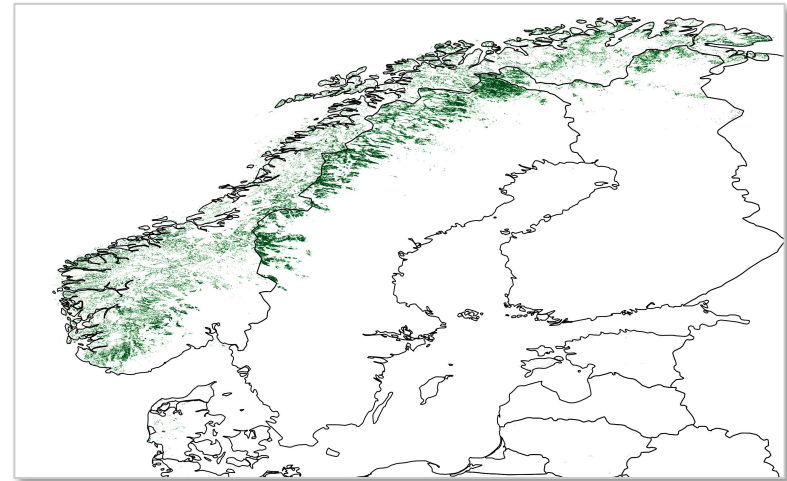
Overview over 44 data layers with information on land use types created from the CORINE land cover data set

Step 1: Customizing the dataset

Mosquitos can be stationary or mobile. Stationary species have an access radius to surrounding habitats of max 300m, while mobile species have an access radius of max 3000 m.

For each grid cell we calculated the proportion of the land type that is accessible in a 300m as well as 3000 m radius. This information was incorporated into the grid cells in each of the 44 data layers.

Although these data layers are now highly customized, they still have generic value. The 44 variables can now be used by other scientists who wish to investigate correlations between land types and abundance of terrestrial species (with similar habitat access range) across the European continent.



Example. Data layer with variable no. 27, as extracted and transformed from the CORINE land use data set. The map shows continuous values describing access to moors and heathland with a mobility range of 300m across the Scandinavian shield.

Step 2: Making the dataset accessible

Because of its generic value, we made the data set publicly available in the BioVeL modelling environment.

Both sets of layers with 300m and 3000m access range to the 44 different land types were deposited at one of BioVeL's service providers in Brazil. For details please visit <https://www.biodiversitycatalogue.org/services/37>

Now the layers are seamlessly accessible through the Ecological Niche Modelling workflow (ENM). For details on the workflow, please visit <http://purl.ox.ac.uk/workflow/myexp-3355.20>

The ENM workflow can now be executed on the BioVeL portal (<http://portal.biovel.eu>) and has access to the new layers

If you want to try yourself, follow the guidelines starting on the next slide...

Step 3: Run modelling experiments with the CORINE data set

The screenshot displays the BioVeL portal interface. The browser address bar shows the URL: https://portal.biovel.eu/workflows?category_id=2&filtered=true&page=all&visibility=public. The navigation bar includes links for Home, Workflows, Runs, Contact, Matthias Obst, and Log out. A red arrow points from the 'Contact' link to a text box containing the instruction: '1. Register and login at <http://portal.biovel.eu>'. Below this, a list of workflows is shown, filtered by 'Ecological Niche Modelling'. The first workflow is 'Bioclim workflow', followed by 'BioVeL ESW DIFF - ENM Statistical Workflow with raster difference computation', and 'Ecological niche modelling workflow'. A second red arrow points from a text box containing the instruction: '2. Choose 'Ecological niche modelling workflow'' to the 'Ecological niche modelling workflow' entry. The left sidebar contains filters for 'Currently showing' (Ecological Niche Modelling, Public workflows), 'Filter by category' (Taxonomic Refinement, Ecological Niche Modelling, Metagenomics, Phylogenetics, Population Modelling, Ecosystem Modelling, Other), 'Filter by visibility' (Your private workflows, Only visible to BioVeL members, Only visible to Swedish LifeWatch members, Only visible to registered users, Public workflows), and 'Filter by uploader' (Matthias Obst (you), Robert Haines, Finn Besell).

Step 3: Run modelling experiments with the CORINE data set

1. Register and login at <http://portal.biovel.eu>

2. Choose 'Ecological niche modelling workflow'

Step 3: Run modelling experiments with the CORINE data set

3. Find the tutorials and documentation

4. Find the original Taverna workflow

5. Start the workflow

The screenshot shows the BioVeL portal interface for the 'Ecological niche modelling workflow'. The browser address bar displays 'https://portal.biovel.eu/workflows/440'. The page has a green header with navigation links: Home, Workflows, Runs, Contact, Matthias Obst, and Log out. The main content area has a green banner with the title 'Ecological niche modelling workflow'. Below the banner, there are three buttons: 'Run workflow', 'Download workflow', and 'Add to Favourites'. To the left of these buttons, there is a 'Version: 5 (latest)' dropdown menu and two buttons: 'Documentation' and 'View on myExperiment'. Below these buttons, there is a text box containing the workflow description and requirements. At the bottom, there are sections for 'Inputs (2)', 'Outputs (18)', 'Interactions (1)', and 'R Scripts (0)'. On the right side, there is a 'Related runs' section listing several workflow runs with their dates and times.

Home > Ecological Niche Modelling > Ecological niche modelling workflow

Ecological niche modelling workflow

Run workflow Download workflow Add to Favourites

Version: 5 (latest)

Documentation View on myExperiment

This workflow takes as input a file containing species occurrence points to create a model with the openModeller Web Service. Algorithm, environmental layers and mask are selected during the workflow. The model is tested (internal test and optional cross validation external test) and then projected one or more times. All points from the input file are used to create a single model, even if there are differences in the scientific names. Cross validation calculates the mean AUC. Model projections can be downloaded from the links in the workflow output. They are geotiff files with suitability values ranging from 0 to 254 (nodata=255).

For more information about the input file format, please check the documentation of the corresponding parameter. The default occurrence points are from a marine species called Gammarus tigrinus, so it is necessary to choose marine environmental layers during the modelling procedure to use it.

Workflow requirements: When running on Taverna workbench, this workflow requires Internet connection and workbench version >= 2.5.

Please note that ecological niche modelling experiments can take a long time to run depending on the parameters - sometimes several hours. This may happen with high resolution environmental layers, thousands of occurrence points and heavy algorithms, such as ANN and GARP BS. Cancelling a workflow run may not cancel the corresponding job on the server side, so if this procedure is repeated the server may get overloaded.

More information and documentation about this workflow can be found here:
<https://wiki.biovel.eu/display/doc/Ecological+Niche+Modelling+%28ENM%29+Workflow>

Inputs (2)

Outputs (18)

Interactions (1)

R Scripts (0)

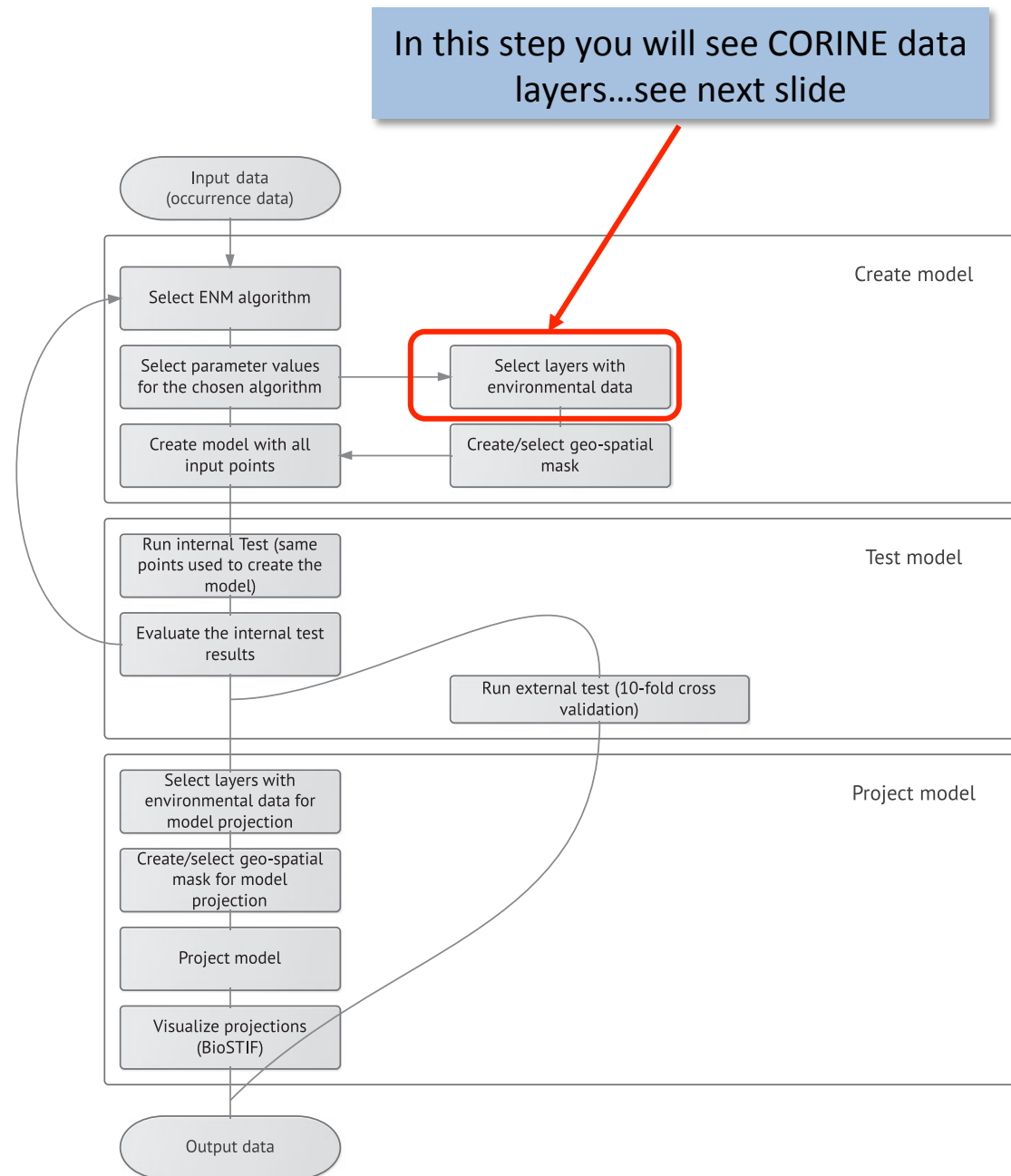
Related runs

- Ecological niche modelling workflow sweep
- Ecological niche modelling workflow sweep
- 5 layers (meanTemp, MeanPrec, 3 forest) - low res - PILOT
- 5 layers (low res: Temp, Prec, highres: 3 forest layers)
- Ecological niche modelling workflow sweep
- Ecological niche modelling workflow (v5) run 30 Aug 2015 07:52:37 UTC
- Ecological niche modelling workflow (v5) run 30 Aug 2015 06:57:12 UTC
- Ecological niche modelling workflow (v5) run 30 Aug 2015 06:53:28 UTC
- Ecological niche modelling workflow (v5) run 30 Aug 2015 06:48:28 UTC
- Ecological niche modelling workflow (v5) run 29 Aug 2015

Step 3: Run modelling experiments with the CORINE data set

The ecological niche modelling (ENM) will guide you through a series of interactions where the algorithm, parameters, environmental layers, and geographical masks are selected. An example input file is provided by default. Your model is tested and projected under the climate scenarios that you select. The model projection outputs are geotiff files with habitat suitability values.

Try yourself following the tutorial at <http://tinyurl.com/ng59nv6>



Ecological niche modelling workflow as developed for climate studies. For details see <http://dx.doi.org/10.1111/jbi.12395>

Step 3: Run modelling experiments with the CORINE data set

In step 10 of the tutorial (<http://tinyurl.com/ng59nv6>) you will be able to select the CORINE data (a) together with climatic and environmental data (b).

Various climate scenarios can be selected by choosing climate layers for 2050, 2100, etc.

(b)

(a*)

**the layer names are currently replaced to identify the variables, e.g. 'corine_land_cover_original_r300m_code_14' will be named 'Corine14_Rice fields' by 1st of September*

The screenshot displays the 'Run Interaction' interface with a tree view of data layers. A red arrow labeled '(b)' points to the 'climate' layer under the 'terrestrial' category. Another red arrow labeled '(a*)' points to a list of CORINE land cover codes at the bottom. The interface includes 'Submit selected layers' buttons for both sections.

Run Interaction

Select layers for model creation

- oM Server
 - layers
 - marine
 - global
 - bio-oracle
 - incofish
 - terrestrial
 - climate (b)
 - soil
 - global
 - hwsd
 - subsoil
 - 10arc-minutes
 - Base Saturation - average value from all soil categories (%)
 - Base Saturation - value from the dominant soil category (%)
 - Bulk Density - average value from all soil categories (kg/dm3)*100
 - Bulk Density - value from the dominant soil category (kg/dm3)*100

Submit selected layers

 - ☐ corine_land_cover_original__r300m_code_05
 - ☐ corine_land_cover_original__r300m_code_06
 - ☐ corine_land_cover_original__r300m_code_07
 - ☐ corine_land_cover_original__r300m_code_08
 - ☐ corine_land_cover_original__r300m_code_09
 - ☐ corine_land_cover_original__r300m_code_10
 - ☐ corine_land_cover_original__r300m_code_11
 - ☐ corine_land_cover_original__r300m_code_12
 - ☐ corine_land_cover_original__r300m_code_13
 - ☐ corine_land_cover_original__r300m_code_14
 - ☐ corine_land_cover_original__r300m_code_15
 - ☐ corine_land_cover_original__r300m_code_16
 - ☐ corine_land_cover_original__r300m_code_17

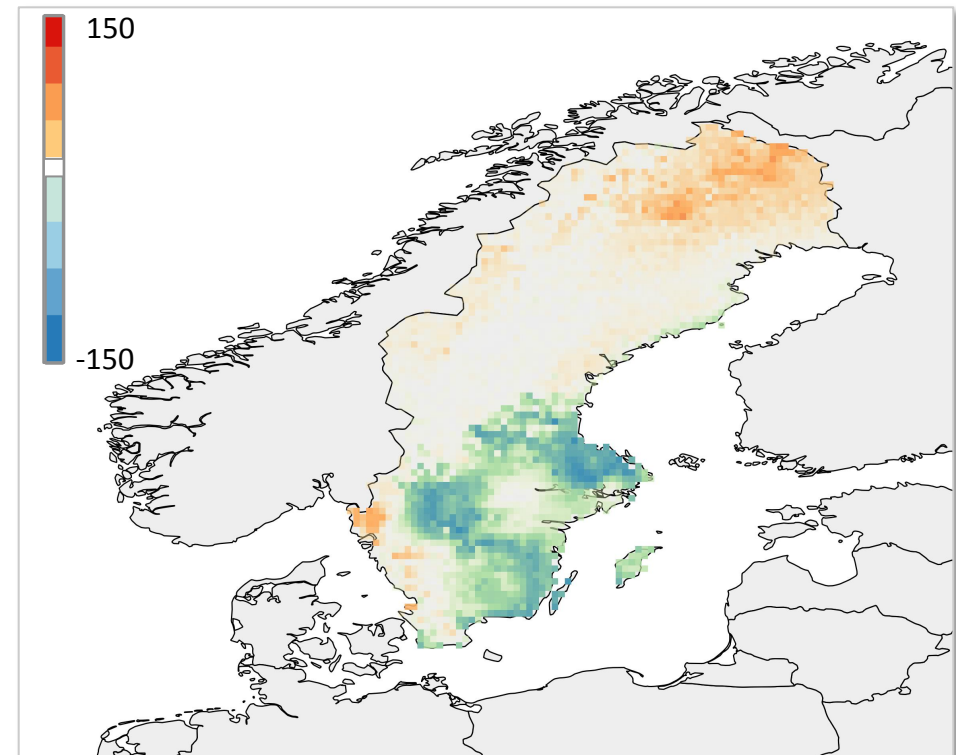
Submit selected layers

Step 3: Run modelling experiments with the CORINE data set

As an example we created and projected a simple model using CORINE and climate data for one mosquito species (*Aedes cantans*). The experiment is documented here:

<http://portal.biovel.eu/runs/9868>

To see the model projections, click on 'BioSTIF_link' and activate the projections on the left hand panel in the upcoming GIS interface.



*Predicted changes (present-2050) in habitat suitability for the disease carrying mosquito *Aedes cantans* as calculated from the ENM workflow using CORINE and climatic data*

Step 3: Run modelling experiments with the CORINE data set

portal.biovel.eu/runs/9868

Home Workflows **Runs** Contact Matthias Obst Log out

Home > Ecological Niche Modelling > Ecological niche modelling workflow > RDA demonstration 1: Modelling suitable habitat for one mosquito species (Aedes cantans) under 2012 and 2050 climate scenario

RDA demonstration 1: Modelling suitable habitat for one mosquito species (Aedes cantans) under 2012 and 2050 climate scenario

Download all results Data sweep based on this run Manage run Delete

Name: RDA demonstration 1: Modelling suitable habitat for one mosquito species (Aedes c) Save

Visibility: Only visible to members of BioVeL, BioVeL Portal Guests, BioVeL, and Swedish LifeWatch

Workflow: Ecological niche modelling workflow Created at: 31 Aug 2015 17:38:46 UTC

Category: Ecological Niche Modelling Started at: 31 /

Status: Finished Finished at: 31 /

Outputs

Jump to:

Results: answer area_statistics BioSTIF_csv_data_url BioSTIF_link create_final_model_log external_auc_list external_omission_list internal_test_model_log internal_test_model_statistics mean_auc mean_omission projection_url project_model_output_log serialized_final_model xval_create_model_log xval_test_model_log xval_test_model_statistics xval_threshold

Results (18)

answer (text/plain)

portal.biovel.eu/sweeps/new?run_id=9868

Download value

Once the parameters are fixed, the analytical cycles can be executed for a large number of species using the 'Data sweep' function.

Step 3: Run modelling experiments with the CORINE data set

As an example we created a simple Sweep analysis (i.e. batch run execution) for 4 species (*Aedes cantans*, *A. cinereus*, *A. communis*, and *Coquillettidia richiardii*). The experiment is documented here: <https://portal.biovel.eu/sweeps/238>

The screenshot shows the BioVeL portal interface. The browser address bar displays <https://portal.biovel.eu/sweeps/238>. The portal header includes navigation links (Home, Workflows, Runs), user information (Contact, Matthias Obst), and a Log out button. The breadcrumb trail indicates the current page is under 'Home > Ecological Niche Modelling > Ecological niche modelling workflow > RDA demonstration 2: Modelling suitable habitat for 4 mosquito species (Aedes cantans, A. cinereus, A. communis, Coquillettidia richiardii) under 2012 and 2050 climate scenario'. The main title of the sweep is 'Sweep: RDA demonstration 2: Modelling suitable habitat for 4 mosquito species (Aedes cantans, A. cinereus, A. communis, Coquillettidia richiardii) under 2012 and 2050 climate scenario'. Below the title are buttons for 'Manage sweep' and 'Delete sweep'. A 'Visibility' box states: 'Only visible to members of BioVeL, BioVeL Portal Guests, and Swedish LifeWatch'. A 'Runs' section header is followed by a search bar. A table lists four runs, all completed, showing details for each species and the workflow used.

Home > Ecological Niche Modelling > Ecological niche modelling workflow > RDA demonstration 2: Modelling suitable habitat for 4 mosquito species (Aedes cantans, A. cinereus, A. communis, Coquillettidia richiardii) under 2012 and 2050 climate scenario

Sweep: RDA demonstration 2: Modelling suitable habitat for 4 mosquito species (Aedes cantans, A. cinereus, A. communis, Coquillettidia richiardii) under 2012 and 2050 climate scenario

[Manage sweep](#) [Delete sweep](#)

Visibility: Only visible to members of BioVeL, BioVeL Portal Guests, and Swedish LifeWatch

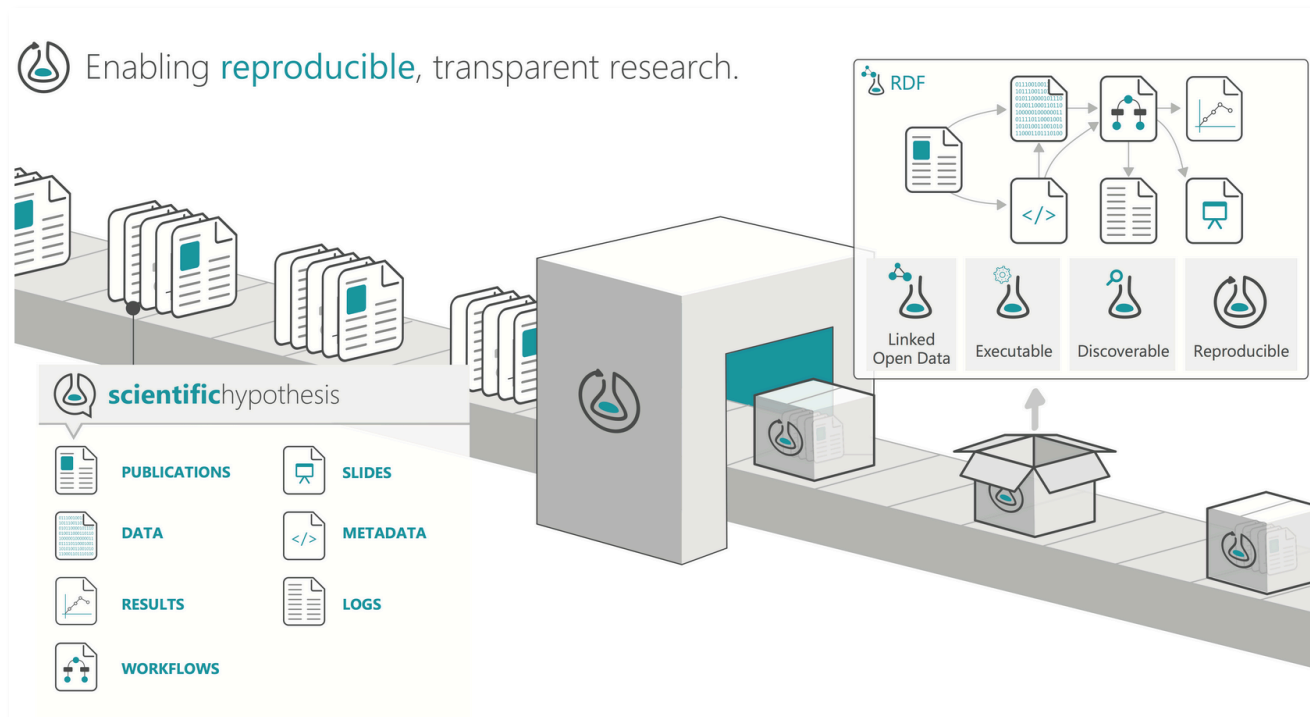
Runs

Search:

| Run | Workflow | Category | State | Created | Finished | Actions |
|---|-------------------------------------|----------------------------|----------|--------------------|--------------------|------------------------|
| Suitable habitat for Aedes cantans under 2010 and 2050 climate scenario | Ecological niche modelling workflow | Ecological Niche Modelling | Finished | about 15 hours ago | about 15 hours ago | Delete |
| Suitable habitat for Coquillettidia richiardii under 2010 and 2050 climate scenario | Ecological niche modelling workflow | Ecological Niche Modelling | Finished | about 15 hours ago | about 15 hours ago | Delete |
| TableSuitable habitat for Aedes cinereus under 2010 and 2050 climate scenario | Ecological niche modelling workflow | Ecological Niche Modelling | Finished | about 15 hours ago | about 15 hours ago | Delete |
| TableSuitable habitat for Aedes communis under 2010 and 2050 climate scenario | Ecological niche modelling workflow | Ecological Niche Modelling | Finished | about 15 hours ago | about 15 hours ago | Delete |

Research object

A fully documented analysis will be published as a Research Object (www.researchobject.org) in the Zenodo online repository (<https://zenodo.org>) once the scientific study is completed. Here is an example of an earlier LifeWatch use case using a marine data set: <https://zenodo.org/record/14839#.VeQ4Fc5FGbg>



Acknowledgements and financial disclosure

The analysis as well as underlying tool development was supported by the EU's Seventh Framework Program project BioVeL (www.biovel.eu, grant no. 283359) and by the Swedish Research Council through the Swedish LifeWatch project (<http://swedishlifewatch.se>, grant no. 829-2009- 6278)

