

Existing AI tools and services: A Galaxy Platform focus

27 Janvier 2026

Using Galaxy-Ecology as a sustainable Biodiversity common virtual research environment: Focus on integration, access and automation of AI algorithms

Yvan Le Bras
MNHN / PNDB-Data Terra / UAR DoHNÉE
French BON
France



Coline Royaux
MNHN / PNDB – DATA TERRA / UAR DoHNÉE / French BON
France

Marie Jossé
CNRS / – DATA TERRA / French BON
France

Pauline Seguineau
MNHN / PNDB – DATA TERRA / UAR DoHNÉE / French BON
France

Arthur Barreau
SU / PNDB – DATA TERRA / UMR ISYEB / French BON
France

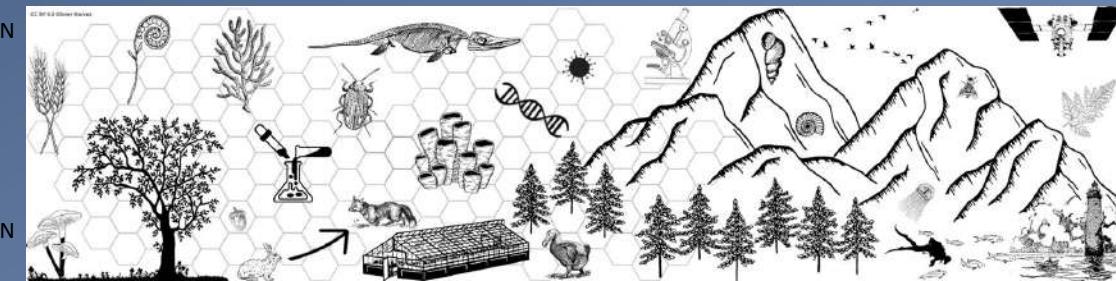
Nadine Le Bris
SU / PNDB – DATA TERRA / UMR ISYEB / French BON
France

Olivier NORVEZ
CNRS / PNDB & Theia – DATA TERRA / French BON
France

Francis Clément
MNHN / PNDB – DATA TERRA / French BON
France

Nicolas Buisine
MNHN / PNDB – DATA TERRA / UAR DoHNÉE / French BON
France

Jian-Sheng Sun
MNHN / PNDB – DATA TERRA / UAR DoHNÉE / French BON
France



Galaxy Ecology: From a platform to best practices in Biodiversity data analysis

Applying “Galaxy-E” best practices to reuse!

REEF LIFE SURVEY fishes data

Atomization & Generalization to reuse & transpose



Guidance framework to apply best practices in ecological data analysis: lessons learned from building Galaxy-Ecology

Coline Royaux^{1,2,*}, Jean-Baptiste Mihoub³, Marie Jossé⁴, Dominique Pellerier⁵, Olivier Norvez⁶, Yves Anne Fouilloux⁹, Helena Rasche⁹, Saskia Hiltemann¹¹, Bérénice Batut^{12,13}, Eléaume Marc^{14,15}, Pauli Guillaume Massé¹⁶, Alan Amosé¹⁷, Claire Bissery^{8,18}, Romain Lorrillière¹³, Alexis Martin¹⁹, Yves Bas^{3,20}, Thimothée Virgoulay^{21,22}, Valentin Chambon¹⁷, Elie Arnaud²¹, Elisa Michon²³, Clara Urfer^{2,24}, Eloïse Trigodet²¹, Grégoire Loïs³, Romain Julliard³, Björn Grünig²⁵, Yvan Le Bras^{3,22}, and The Galaxy-E community

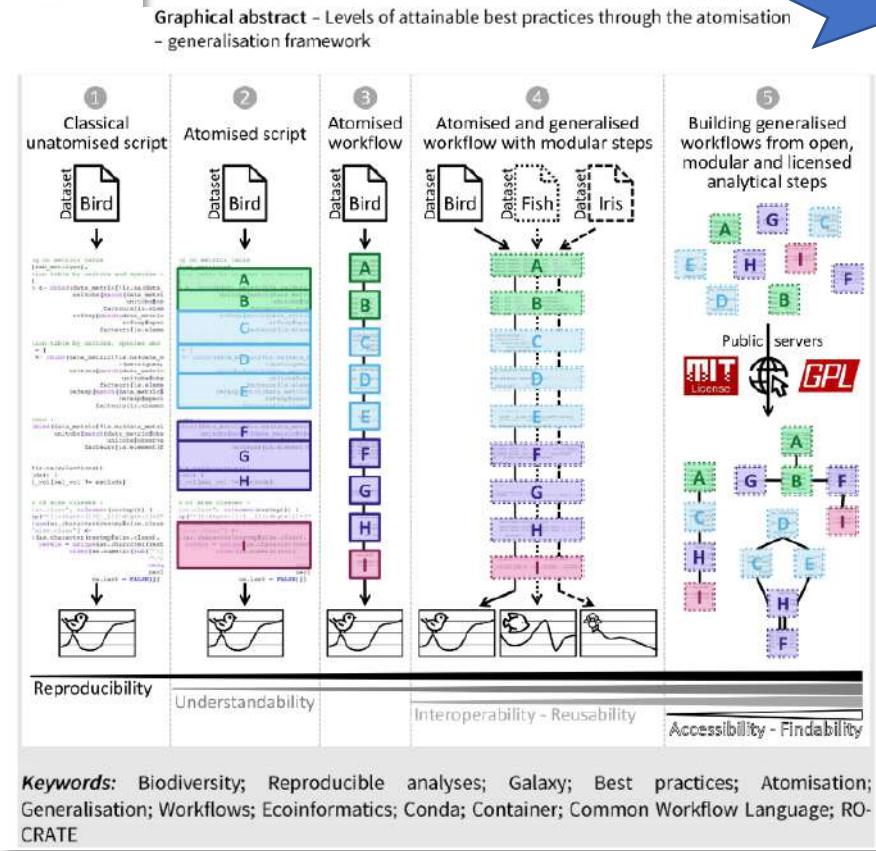
Guidance framework to apply best practices in ecological data analysis: Lessons learned from building Galaxy-Ecology



PCI recommendation: <https://doi.org/10.24072/pci.ecology.100694>

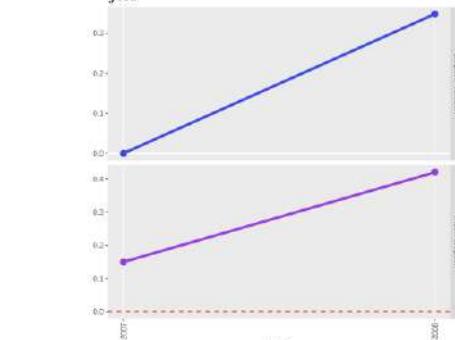
Royaux Coline^{1,2,*}, Mihoub Jean-Baptiste³, Jossé Marie⁴, Pellerier Dominique⁵, Norvez Olivier⁶, Reecht Yves^{7,8}, Fouilloux Anne⁹, Rasche Helena¹⁰, Hiltemann Saskia¹¹, Batut Bérénice^{12,13}, Eléaume Marc^{14,15}, Seguinéau Pauline^{14,15}, Massé Guillaume¹⁶, Amosé Alan¹⁷, Bissery Claire^{8,18}, Lorrillière Romain³, Martin Alexis¹⁹, Bas Yves^{3,20}, Virgoulay Thimothée^{21,22}, Chambon Valentin¹⁷, Arnaud Elie²¹, Michon Elisa²³, Urfer Clara^{2,24}, Trigodet Eloïse^{21,24}, Delannoy Marie³, Loïs Grégoire³, Julliard Romain³, Grünig Björn²⁵, The Galaxy-E community, Le Bras Yvan²

GigaScience, 2025, 14, 1–12
DOI: 10.1093/gigascience/giac122
Review

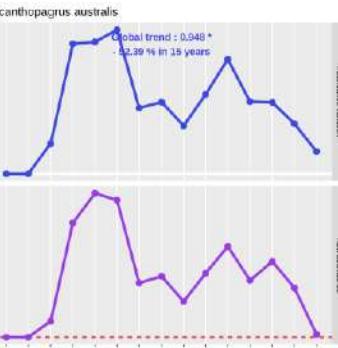
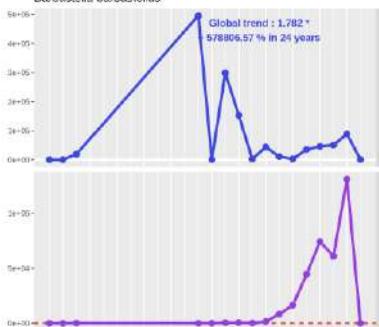


PAMPA

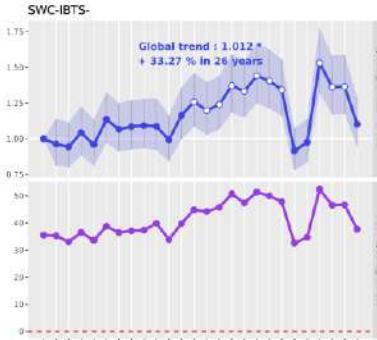
Marine Protected Areas (PAMPA) fishes data



VIGIE CHIRO bats data



DATRAS fishing data



Applying “Galaxy-E” best practices to reuse!

Different data, same primary variables, same workflow

 Location

 Year

 Species

 Occurrence

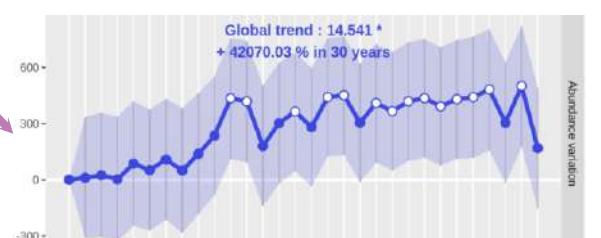
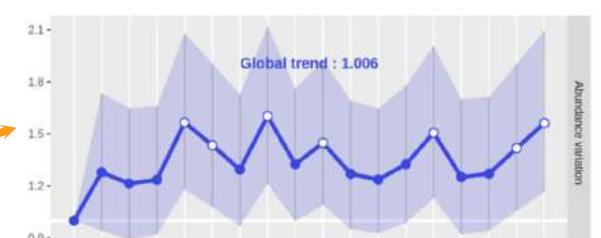
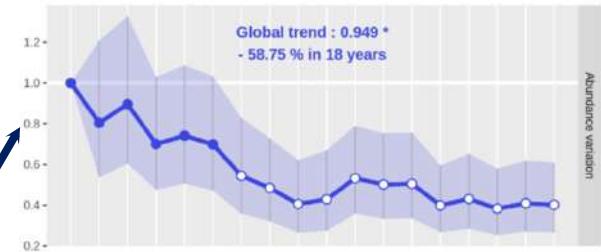
	carre	annee	espece	abond
1	2	2016	ACCGEN	0
2	2	2017	ACCGEN	0
3	2	2018	ACCGEN	0
4	2	2019	ACCGEN	0
5	183	2016	ACCGEN	0
6	183	2017	ACCGEN	0



	Survey	Year	Quarter	Area	AphiaID	Species	LngtClass	CPUE_number_per_hour
1	BITS	1991	1	22	126281	Anguilla anguilla	0	0.000000
2	BITS	1991	1	22	126281	Anguilla anguilla	720	0.009160
3	BITS	1991	1	22	126417	Clupea harengus	0	0.000000
4	BITS	1991	1	22	126417	Clupea harengus	80	0.075785
5	BITS	1991	1	22	126417	Clupea harengus	85	0.075785
6	BITS	1991	1	22	126417	Clupea harengus	95	0.012492
7	BITS	1991	1	22	126417	Clupea harengus	100	0.012492
8	BITS	1991	1	22	126417	Clupea harengus	105	0.012492
9	BITS	1991	1	22	126417	Clupea harengus	110	0.618357



	UnitObs	rotation	codeSp	sexe	taille	classe_taille	poids	nb_ind
1	AS140155	3	Hemifasc	-999	-999	P	-999	1
2	AS140159	1	Nasosp.	-999	-999	P	-999	3
3	AS140159	3	Gompvari	-999	-999	P	-999	1
4	AS140160	3	Gompvari	-999	-999	P	-999	1

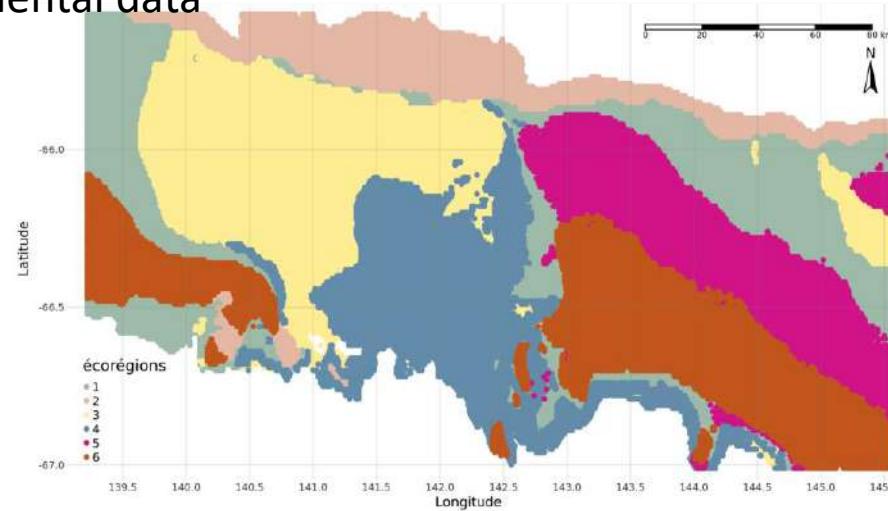


Applying “Galaxy-E” best practices to reuse!

Different data, same primary variables, same workflow

Transpose to other taxonomic and/or geographical and/or temporal context

Project data (invertebrates)
Project environmental data



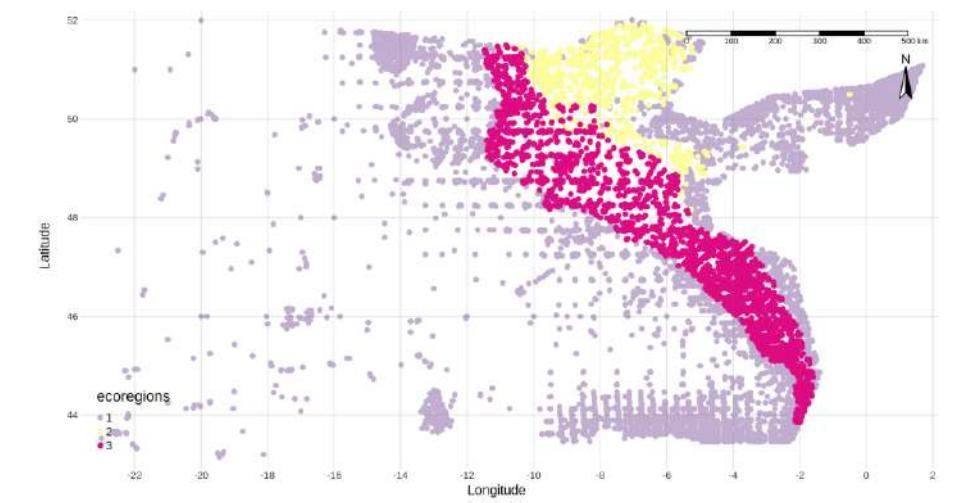
From Antarctica to North East Atlantic

From Invertebrates to Fishes

North Atlantic

OBIS marine occurrence data (fishes)
OBIS marine environmental data (all)

c1: decimallongitude ✕ c2: decimallatitude ✕ c5: shoredistance ✕ c6: bathymetry ✕ c7: sst ✕ c8: sss ✕



Applying “Galaxy-E” best practices to reuse!

Different data, same primary variables, same workflow

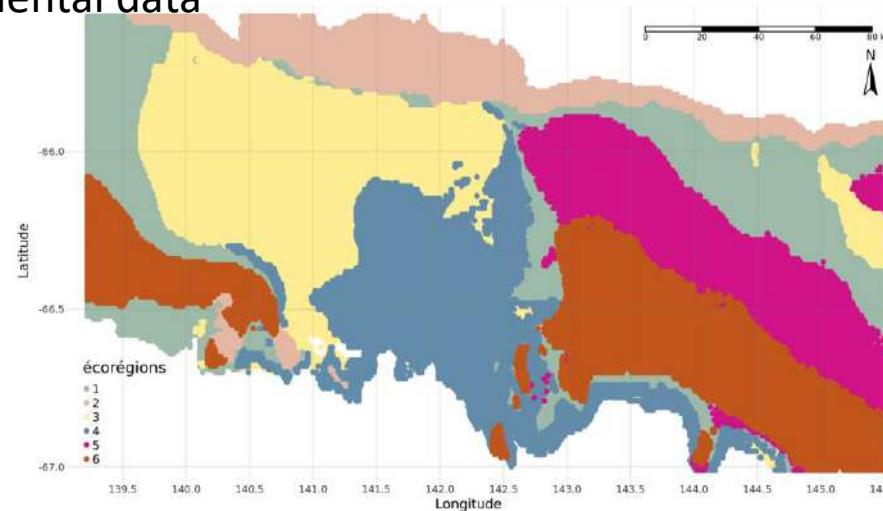
Test several scenarios

From Antarctica to North East Atlantic

From Invertebrates to Fishes

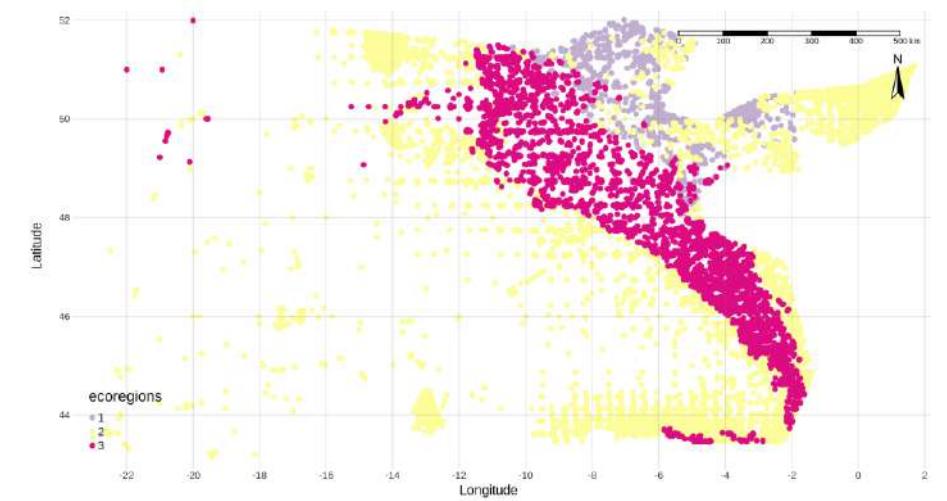
Modifying modelling parameters

Project data (invertebrates)
Project environmental data



North Atlantic
OBIS marine occurrence data (fishes)
OBIS marine environmental data (sst / bath)

c1: long ✕ c2: lat ✕ c4: bathymetry ✕ c5: sst ✕ c6: sss ✕



Training tutorials for Ecology



Data and Metadata Management

These tutorials are focusing on data and metadata management in Ecology.

Lesson	Slides	Hands-on	Recordings	Input dataset	Workflows
Cleaning GBIF data using OpenRefine biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Creating FAIR Quality assessment reports and draft of Data Papers from EML metadata with MetaSHIMPS Metadata EML FAIR Data Paper biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Creating metadata using Ecological Metadata Language (EML) standard with EML Assembly Line functionalities Metadata EML Ecology Biodiversity FAIR Data Paper	Slides	Hands-on	Recordings	Input dataset	Workflows
Data submission using ENA upload Tool	Slides	Hands-on	Recordings	Input dataset	Workflows

Data access

These lessons focus on ways to access data classically used in Ecology

Lesson	Slides	Hands-on	Recordings	Input dataset	Workflows
QGIS Web Feature Services earth-system GIS Geographical Information System WFS Spatial data Maps OGC biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows

Data preprocessing

These lessons focus on manners to preprocess data used in Ecology

Lesson	Slides	Hands-on	Recordings	Input dataset	Workflows
Biodiversity data exploration taxonomic data data quality biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Checking expected species and contamination in bacterial isolate illumina bacteria microgalaxy	Slides	Hands-on	Recordings	Input dataset	Workflows
Cleaning GBIF data for the use in Ecology gbif data management data cleaning biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows

Data visualization

These tutorials showcase data visualization in Ecology

Lesson	Slides	Hands-on	Recordings	Input dataset	Workflows
Visualization of Climate Data using NetCDF xarray Map Plotting pangeo	Slides	Hands-on	Recordings	Input dataset	Workflows
Visualize EBV cube data with Panoply netCDF viewer interactive-tools EBV cube Data visualization	Slides	Hands-on	Recordings	Input dataset	Workflows

Data analysis

These lessons focus on ways to analyse data in Ecology

Lesson	Slides	Hands-on	Recordings	Input dataset	Workflows
Champs bioc indicators Completeness EBV class EBV dataset EBV workflow Marine ecosystems biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Compute and analyze biodiversity metrics with PAMPA toolsuite Species population EBV class Community composition EBV class EBV dataset EBV workflow modeling biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Ecoregionalization workflow tutorial ecology taxonomic data EBV workflow modeling gbif ocean earth-system interactive-tools biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
From NDVI data with OpenEO to time series visualisation with Holoviews earth-system land degradation NDVI Copernicus Holoviews	Slides	Hands-on	Recordings	Input dataset	Workflows
Life Traits Ecoregionalization workflow biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Marine Omics identifying biosynthetic gene clusters earth-system ocean marine omics biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Metabarcoding/edNA through Obitools Genetic composition EBV class Community composition EBV class EBV dataset EBV workflow eDNA Metabarcoding biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Obis marine indicators earth-system ocean marine omics biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Phylogenetic diversity analysis quick tutorial ecology taxonomic data	Slides	Hands-on	Recordings	Input dataset	Workflows
Preparing genomic data for phylogeny reconstruction phylogeny data handling functional annotation biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
RAD-Seq Reference-based data analysis RAD-seq Genetic composition EBV class Species population EBV class EBV dataset EBV workflow biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
RAD-Seq de-novo data analysis RAD-seq Genetic composition EBV class Species populations EBV class EBV dataset EBV workflow biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
RAD-Seq to construct genetic maps RAD-seq Genetic composition EBV class EBV dataset EBV workflow biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Regional GAM Species populations EBV class Species traits EBV class EBV dataset EBV workflow biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Sentinel 2 biodiversity Remote sensing biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Species distribution modeling interactive-tools modeling gbif species populations EBV class biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows
Taxonomic Analysis of eDNA biodiversity	Slides	Hands-on	Recordings	Input dataset	Workflows

Training tutorials for Ecology



New tutorial using
AI model (*in
review*)

Link to the PR

Data and Metadata Management

These tutorials are focusing on data and metadata management in Ecology.

github.com/mjoudy/training-material/blob/yolo_predict_tutorial_deepsea/topics/imaging/tutorials/yolo_predict/tutorial.md

Input

training-material / topics / imaging / tutorials / yolo_prediction / tutorial.md

Preview Code Blame 206 lines (207 loc) • 12.9 KB

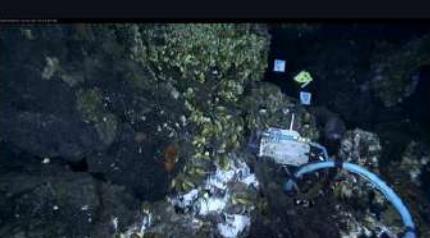
Raw       

Part 1: Detection of marine species

🔗 Dataset

👉 We will use selected images from the SEANOE dataset (% cite lebeaud2024deepsea %).

The [SEANOE](#) collection features real underwater images captured by deep-sea observatories as part of a citizen science initiative called DeepSeaSpy. These non-destructive imaging stations continuously monitor marine ecosystems and provide snapshots of various fauna. In this dataset, multiple annotators—including trained scientists and enthusiastic citizen scientists—have manually labeled images with polygons, lines, or points highlighting marine organisms. These annotations were then cleaned and converted into bounding boxes to create a training-ready dataset for object detection with YOLOv8. Though the exact species vary, images often include deep-sea fish, species, making this dataset well-suited for practicing detection tasks.

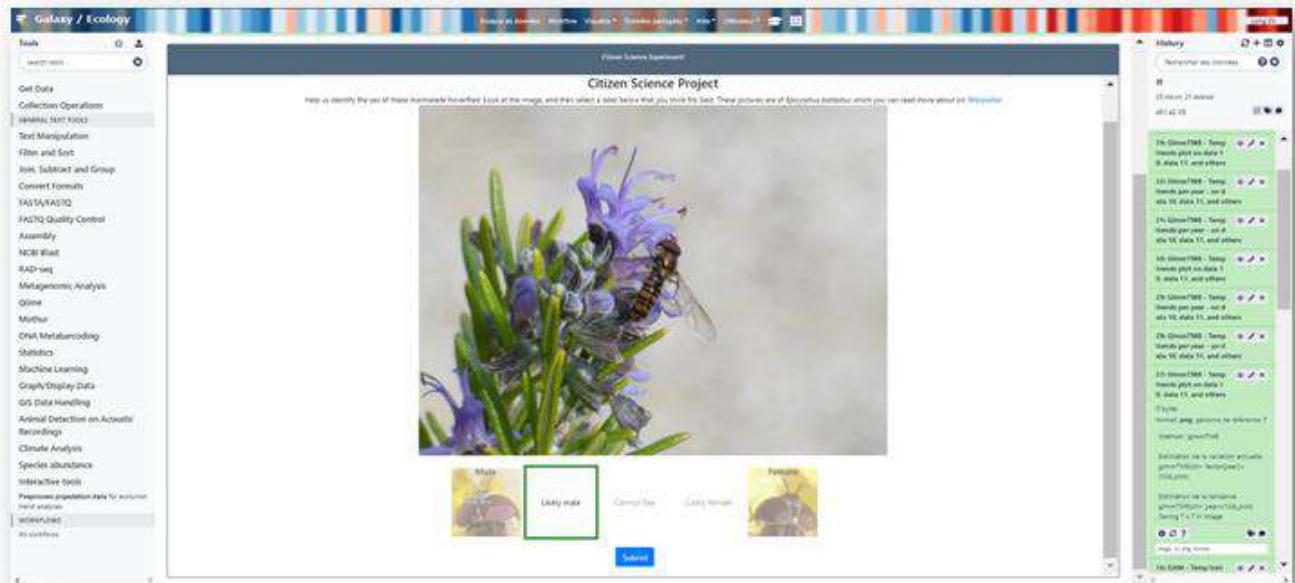




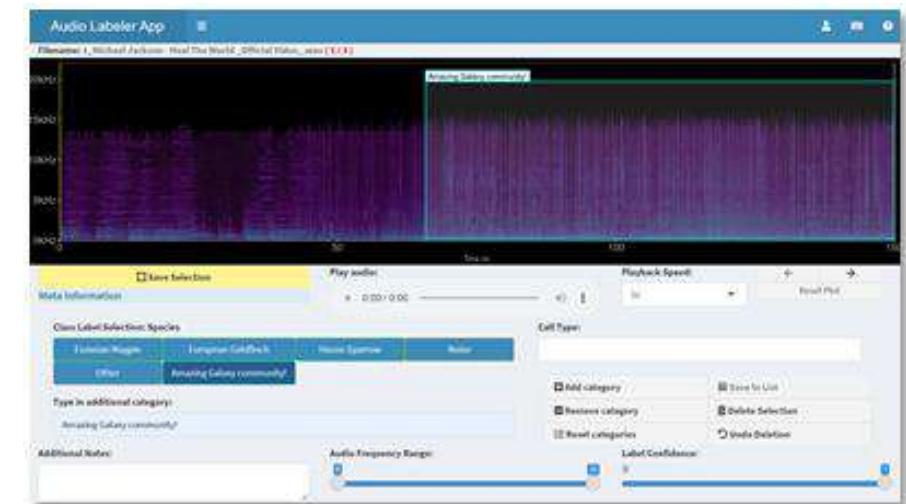
Galaxy Ecology: From citizen science to AI

Citizen science & crowdsourcing

Crowdsourcing image



Annotation of sounds



VIGIE NATURE

Un réseau de citoyens qui fait avancer la science



>3400 classifications per month / >110 per day

Annotation of images



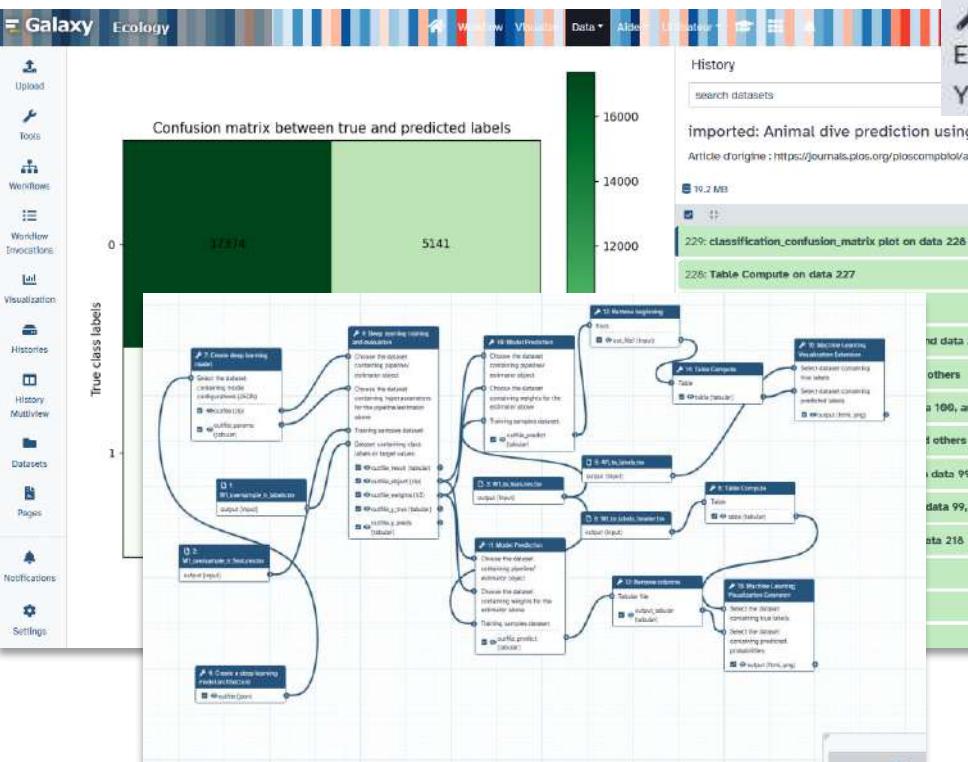
MOOREV



SPIPOL

Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS

RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

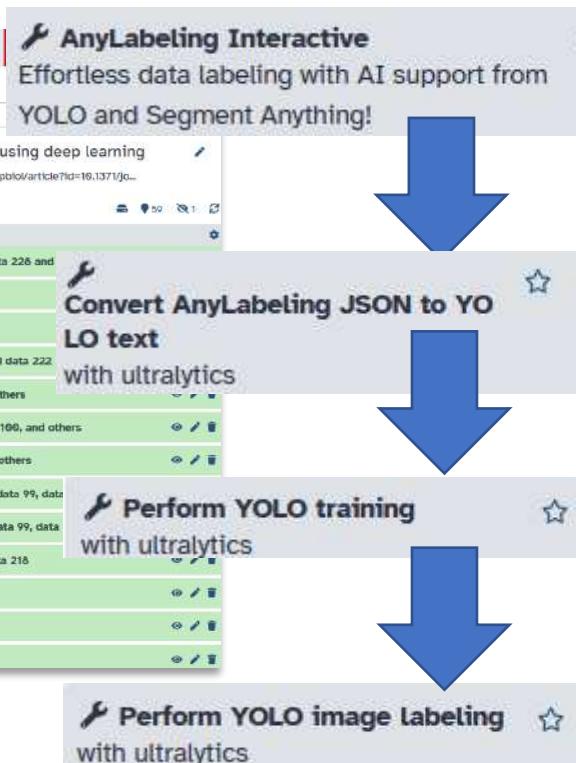
Amédée Roy , Sophie Lanco Bertrand, Ronan Fablet

Version 2

Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>

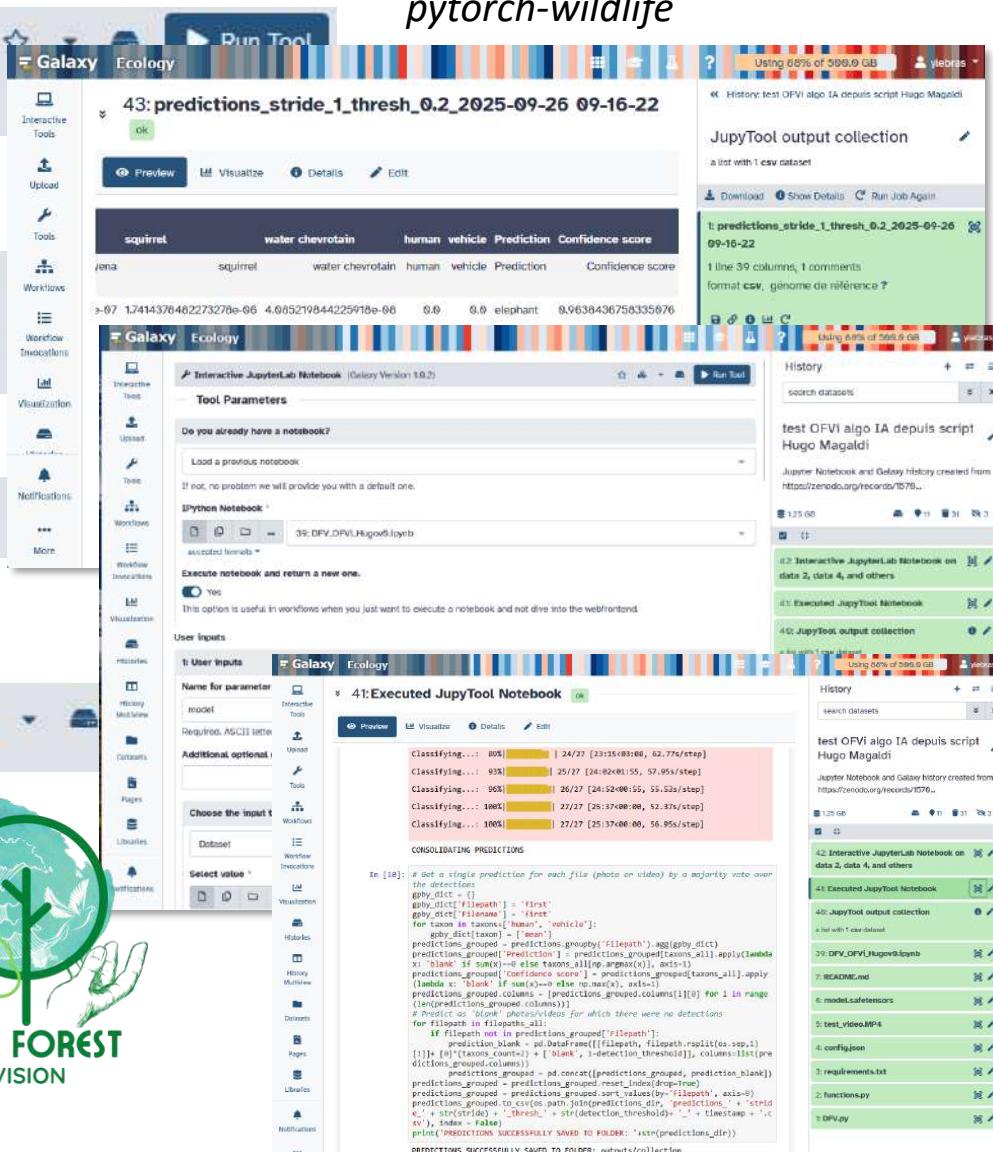
Adapt existing tools to use cases

From Interactive annotation to training to labeling



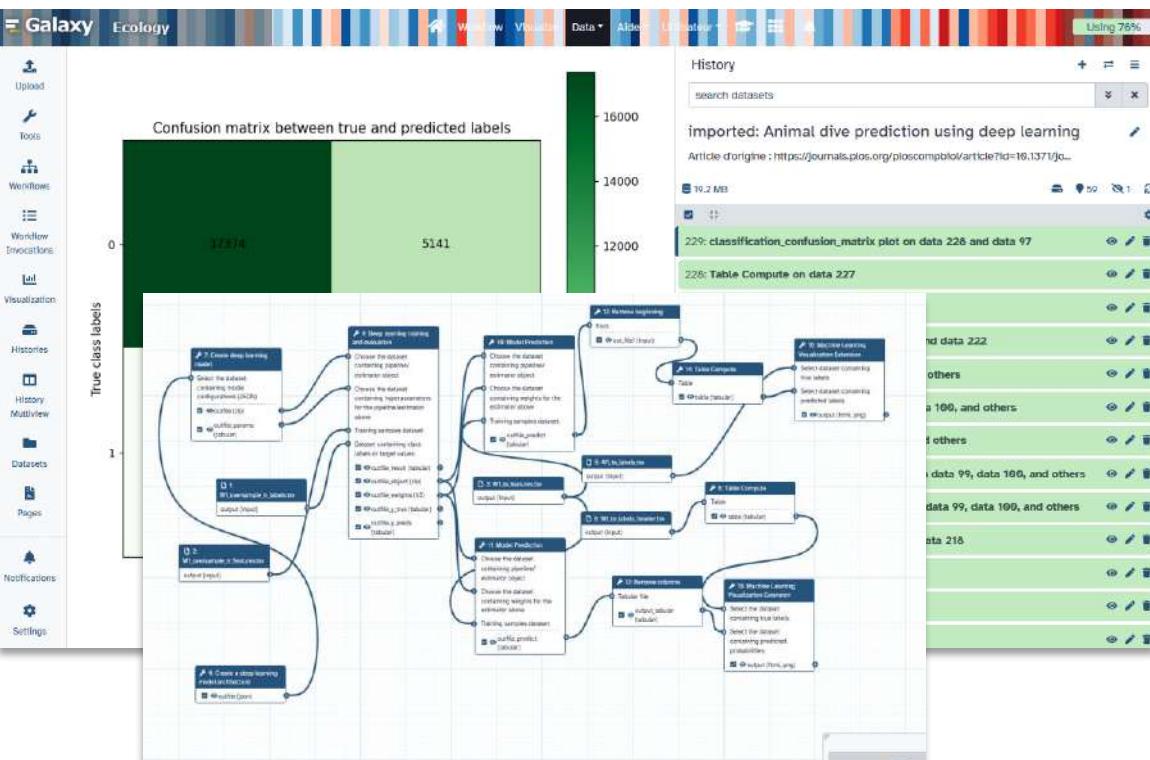
Using Jupyter notebook

From camera trap video to species detection with megadetector & pytorch-wildlife



Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS

RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

Amédée Roy  Sophie Lanco Bertrand, Ronan Fablet

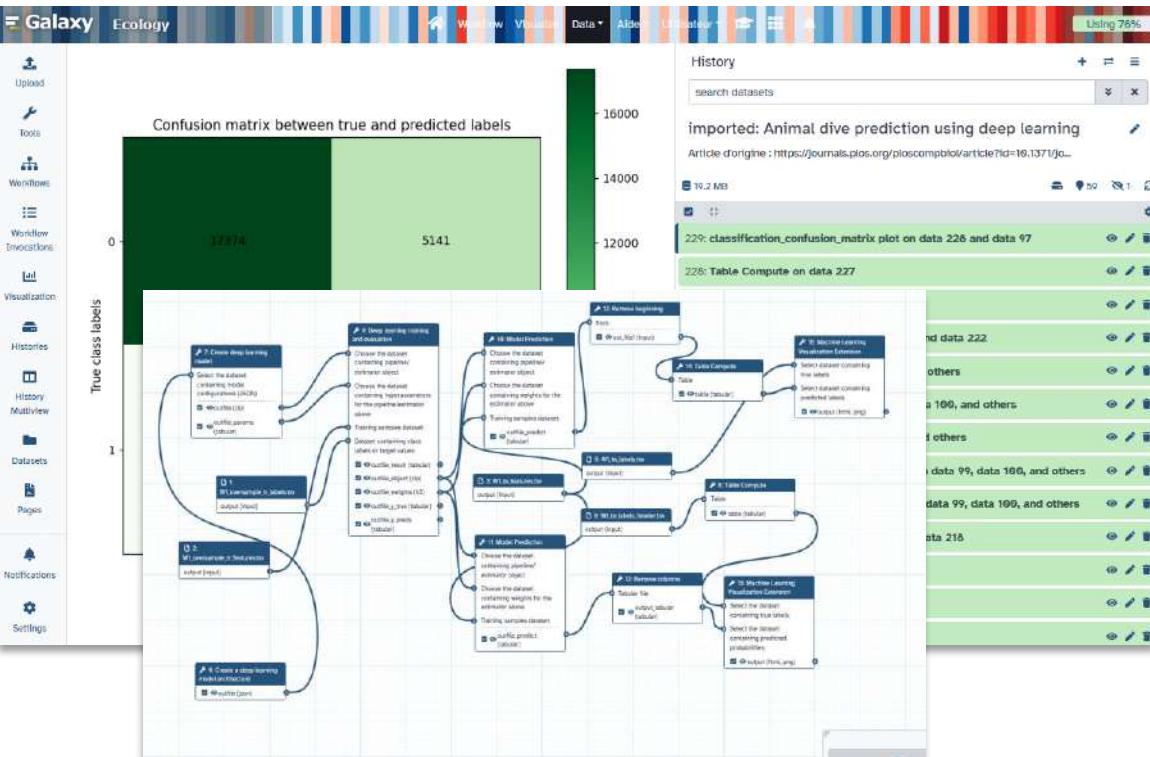
Version 2

Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>

The screenshot shows the Galaxy web interface with the 'Ecology' tool selected. The main area is a 'Create a deep learning model architecture' tool, version 6.5.9. The left sidebar contains links for Interactive Tools, Upload, Tools, Workflows, Workflow Invocations, Visualization, Histories, History Multiview, Datasets, Pages, Libraries, and Notifications. The top right shows system status: 'Using 91% of 569.6 GB' and '1 users'. The main tool parameters are set for a 'Sequential' model with an 'input_shape' of '(3,)'. The 'LAYER' section is expanded, showing the configuration for the first layer: 'Core -- Dense' with 'units' set to '8'. The 'Activation function' is set to 'relu'. A note at the bottom of the layer section says: 'Type in key words arguments if different from the default' (optional), with an example provided: 'use_bias=True, kernel_initializer='glorot_uniform', bias_initializer='zeros', kernel_regularizer=None, bias_regularizer=None, activity_regularizer=None, kernel_constraint=None, bias_constraint=None'. The right side of the interface shows a 'History' panel with a list of recent datasets and files, including '159: Keras Model Config' (selected), '159: Keras Model Config', '159: W1_oversample_tr_features.tsv', '99: W1_oversample_tr_labels.tsv', '98: W1_te_features.tsv', '97: W1_te_labels.tsv', '96: W1_val_features.tsv', and '95: W1_val_labels.tsv'. The '159: Keras Model Config' entry shows its content as a JSON object: { "class_name": "Sequential", "config": { "name": "sequential_1", "layers": [] } }.

Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS

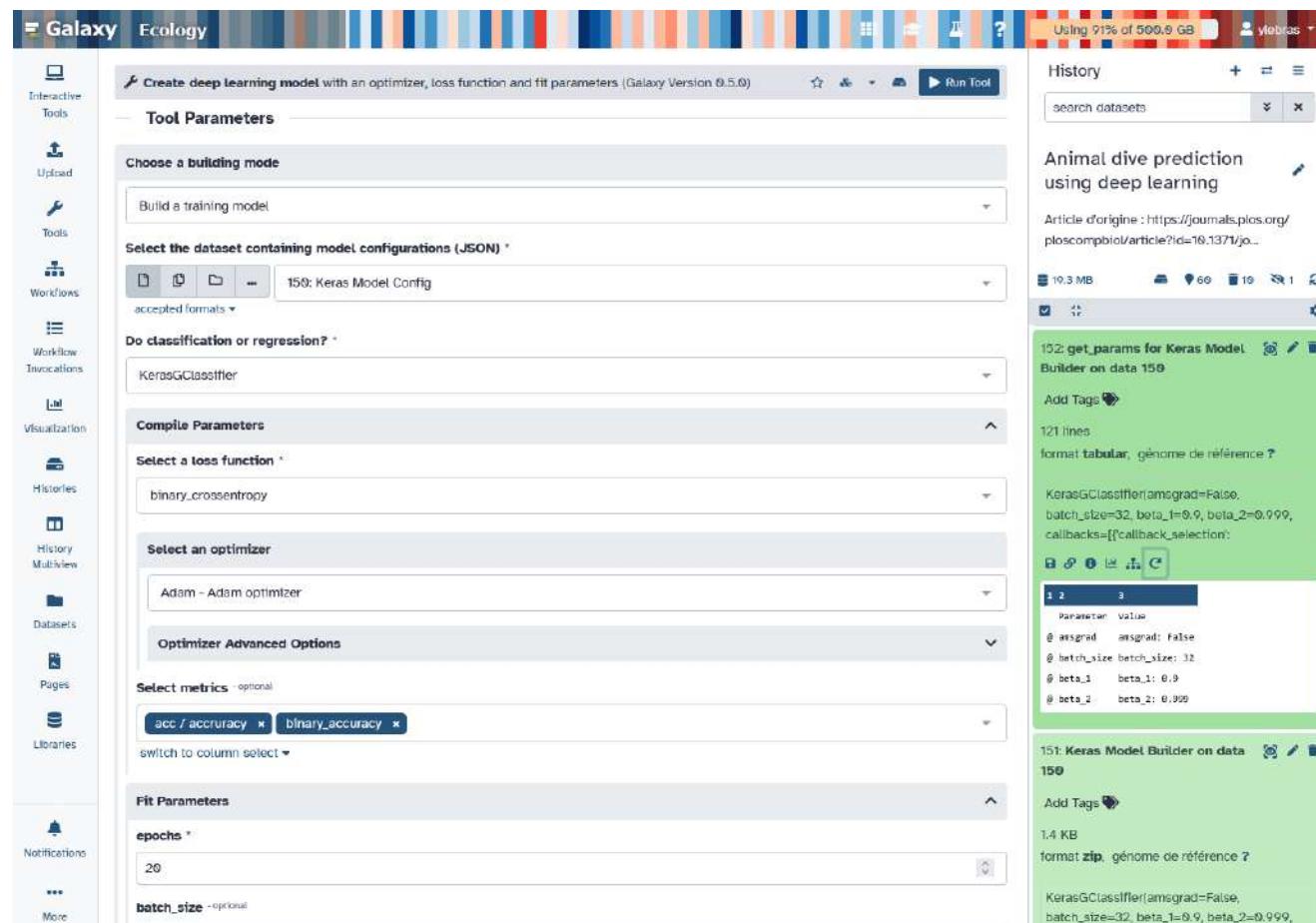
RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

Amédée Roy  Sophie Lanco Bertrand, Ronan Fablet

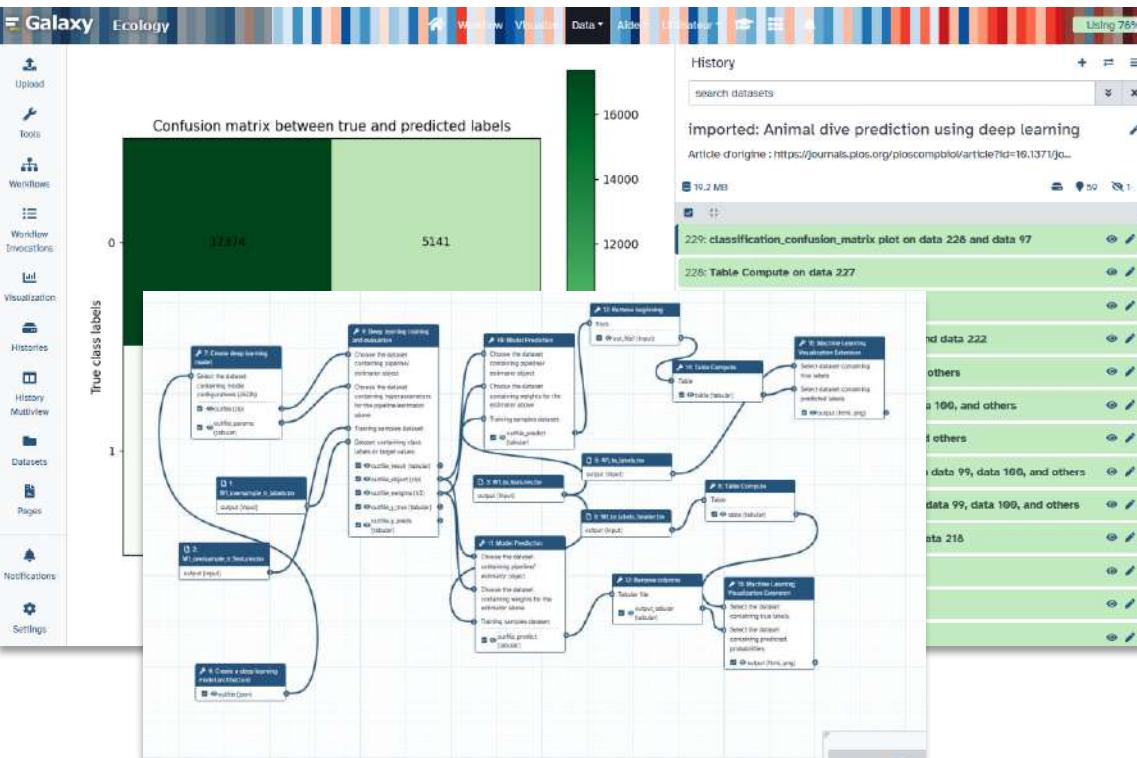
Version 2

Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>



Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

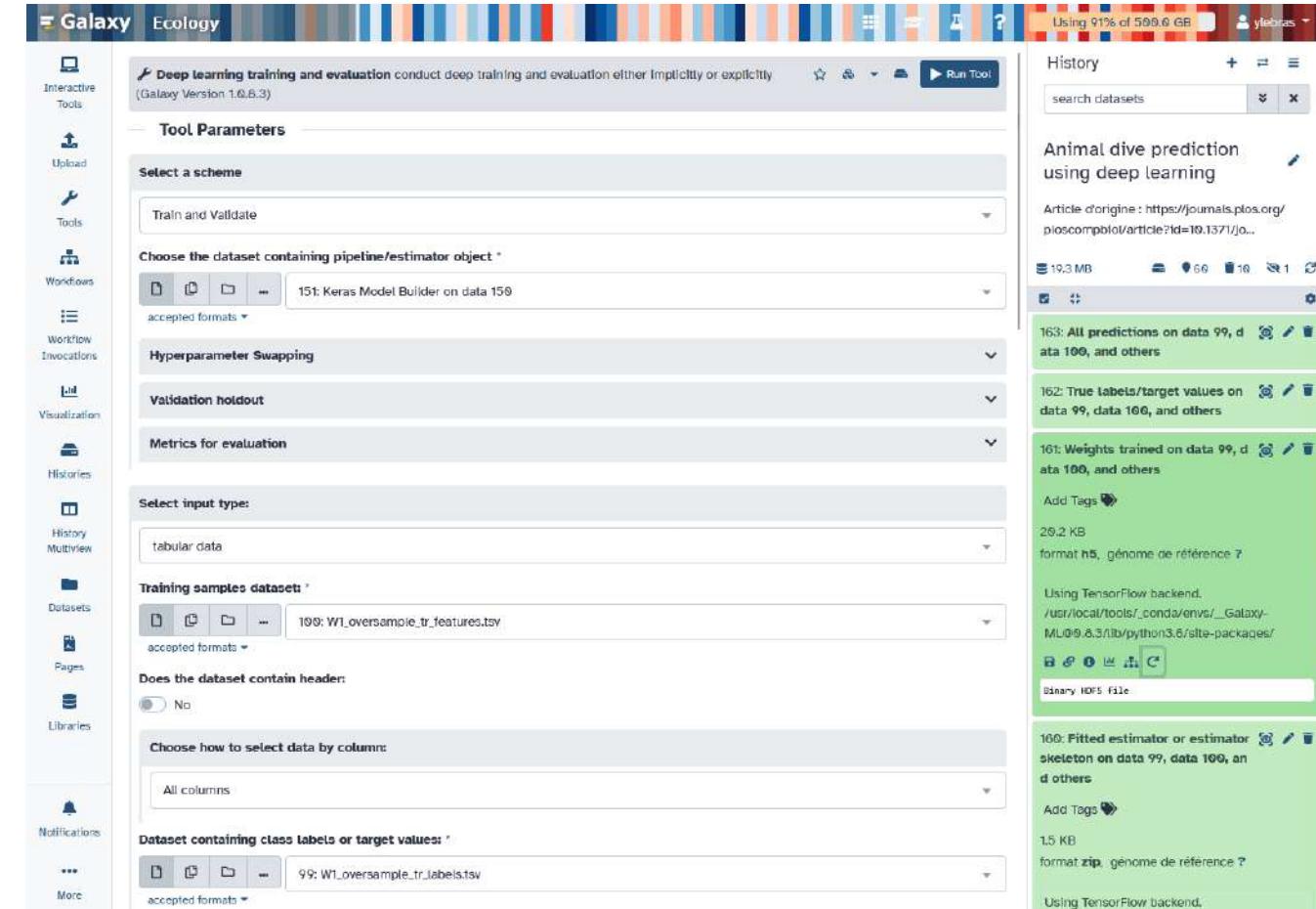
RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

Amédée Roy, Sophie Lanco Bertrand, Ronan Fablet

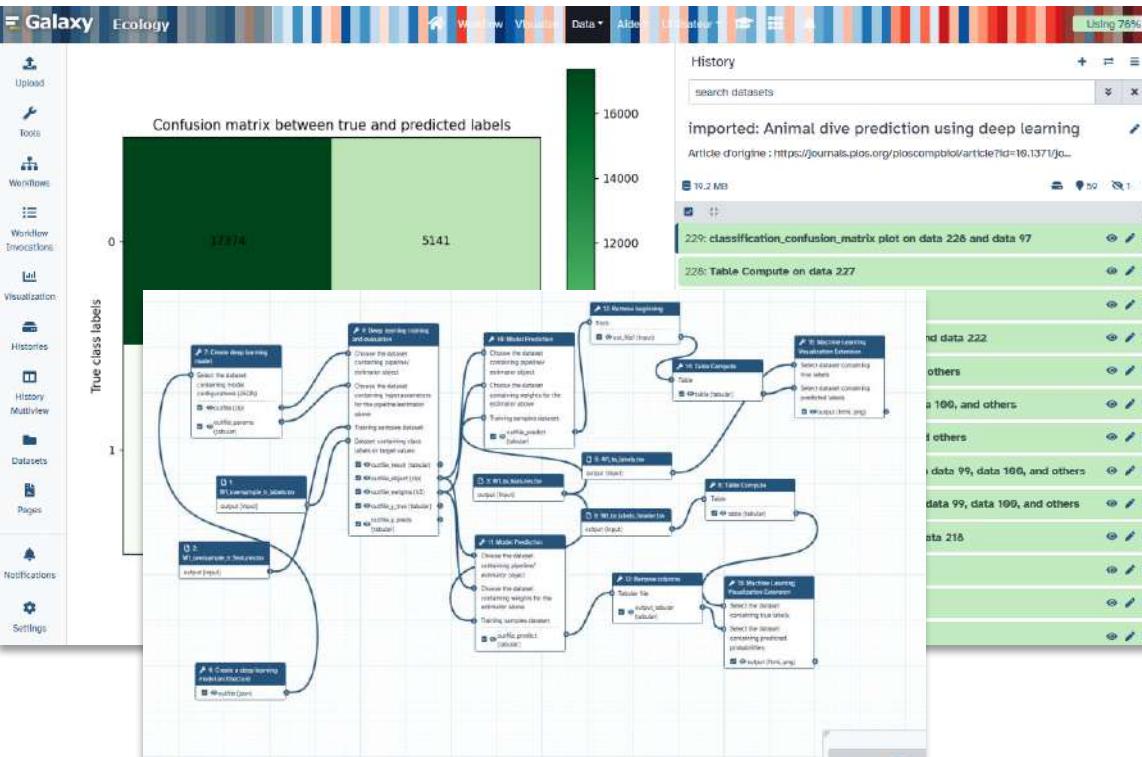
Version 2

Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>



Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

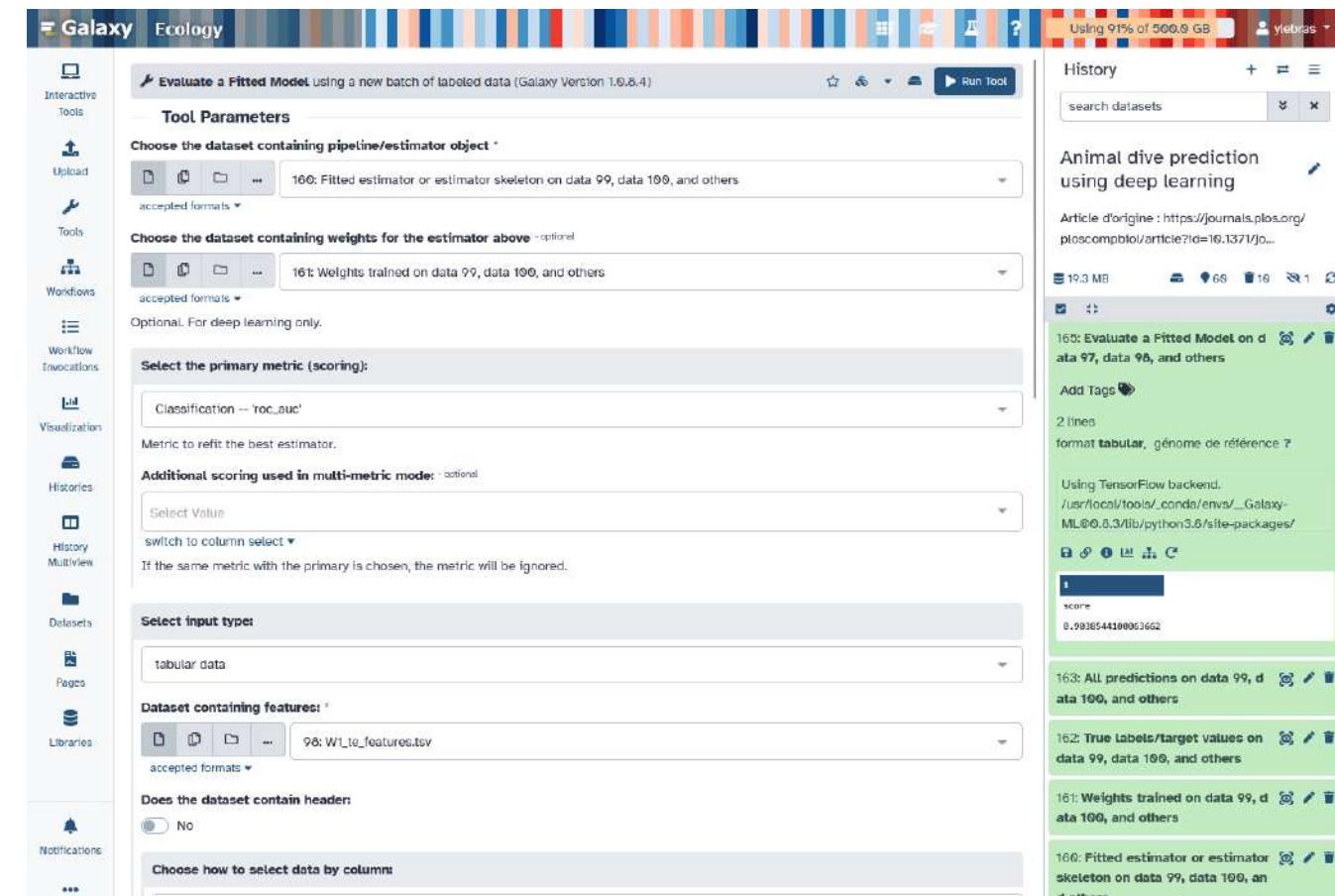
RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

Amédée Roy, Sophie Lanco Bertrand, Ronan Fablet

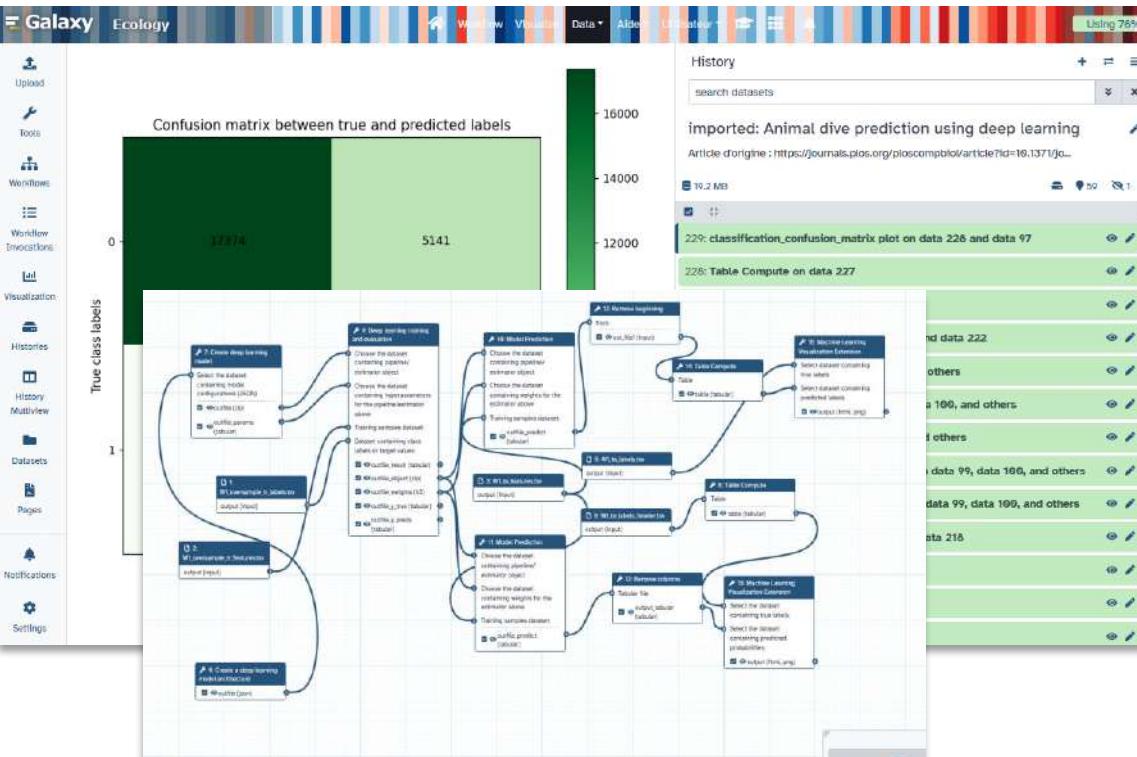
Version 2

Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>



Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

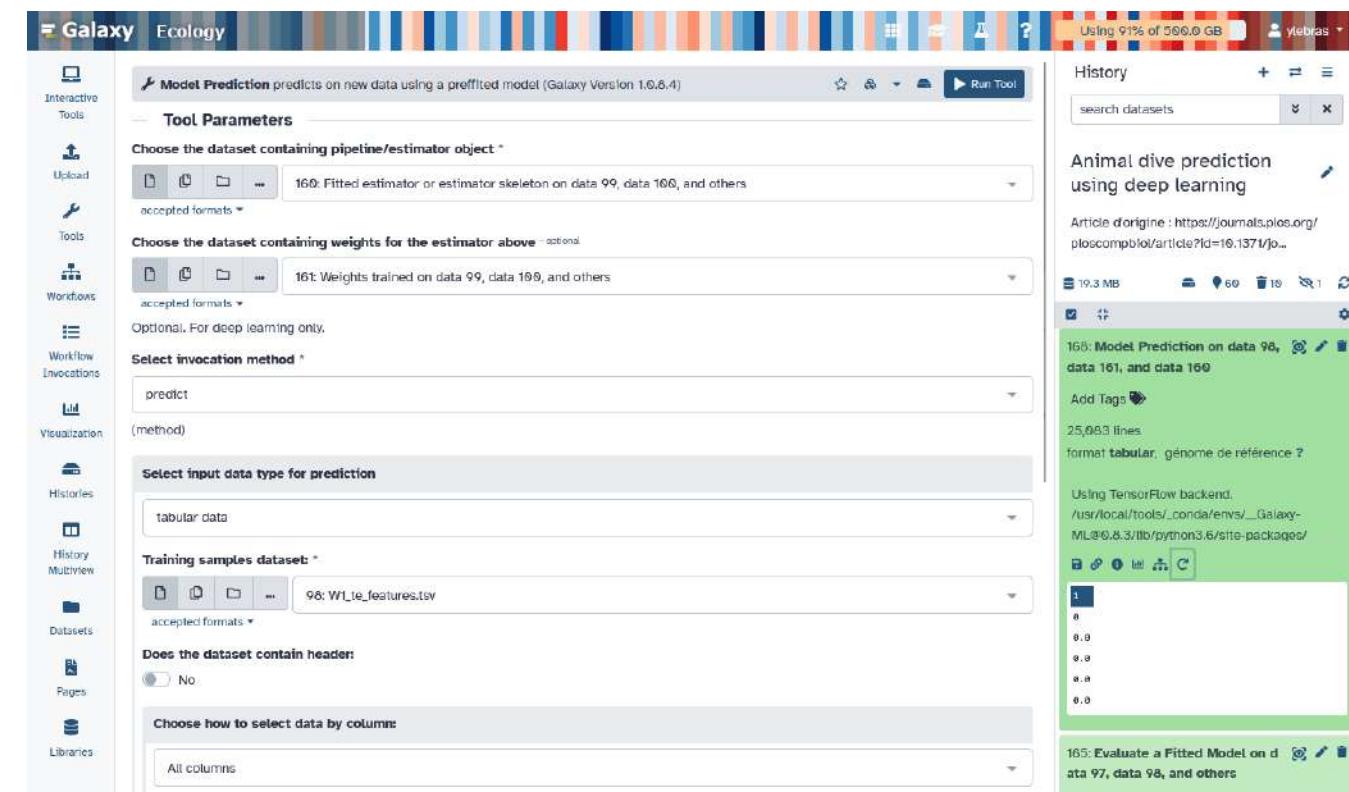
RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

Amédée Roy, Sophie Lanco Bertrand, Ronan Fablet

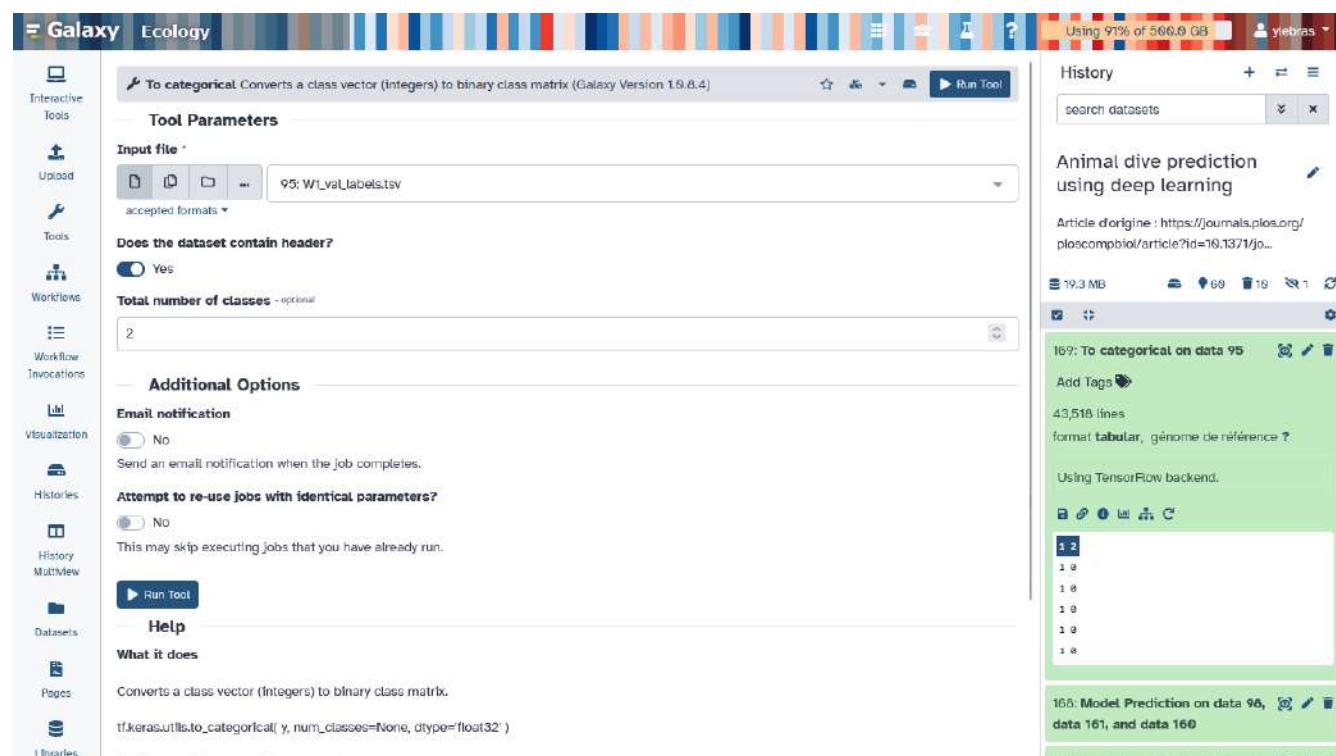
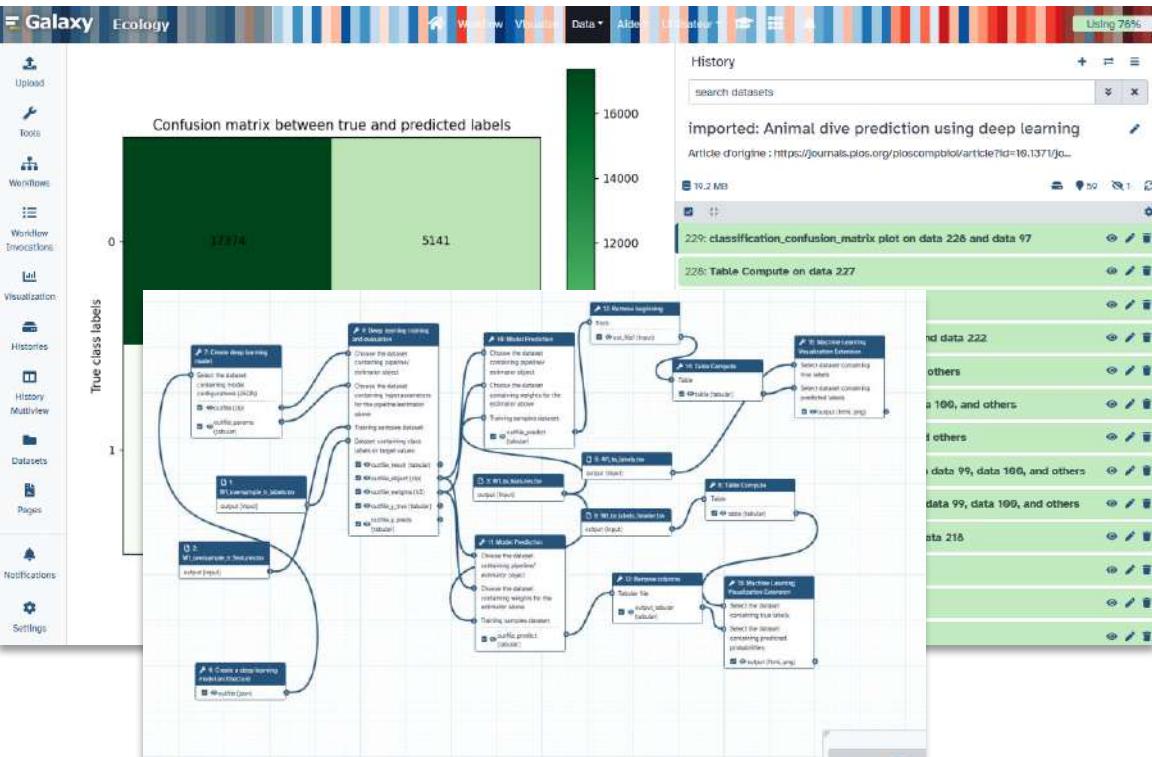
Version 2

Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>



Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

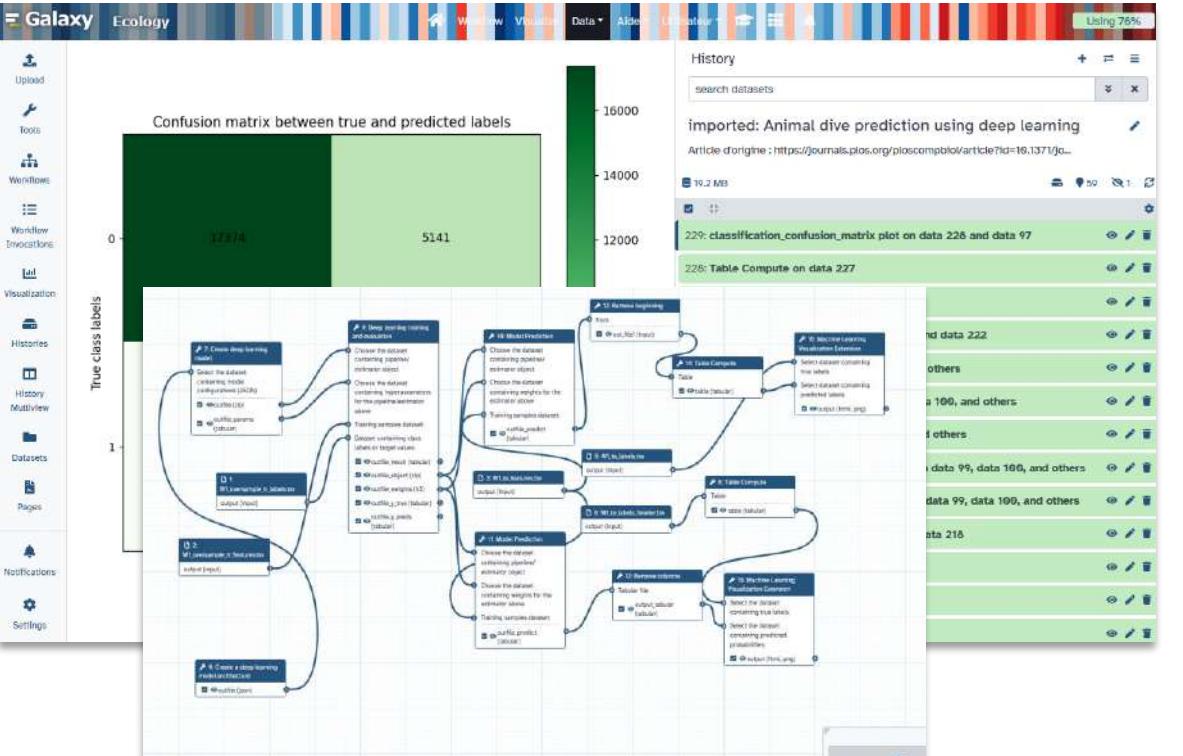
Amédée Roy, Sophie Lanco Bertrand, Ronan Fablet

Version 2

Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>

Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

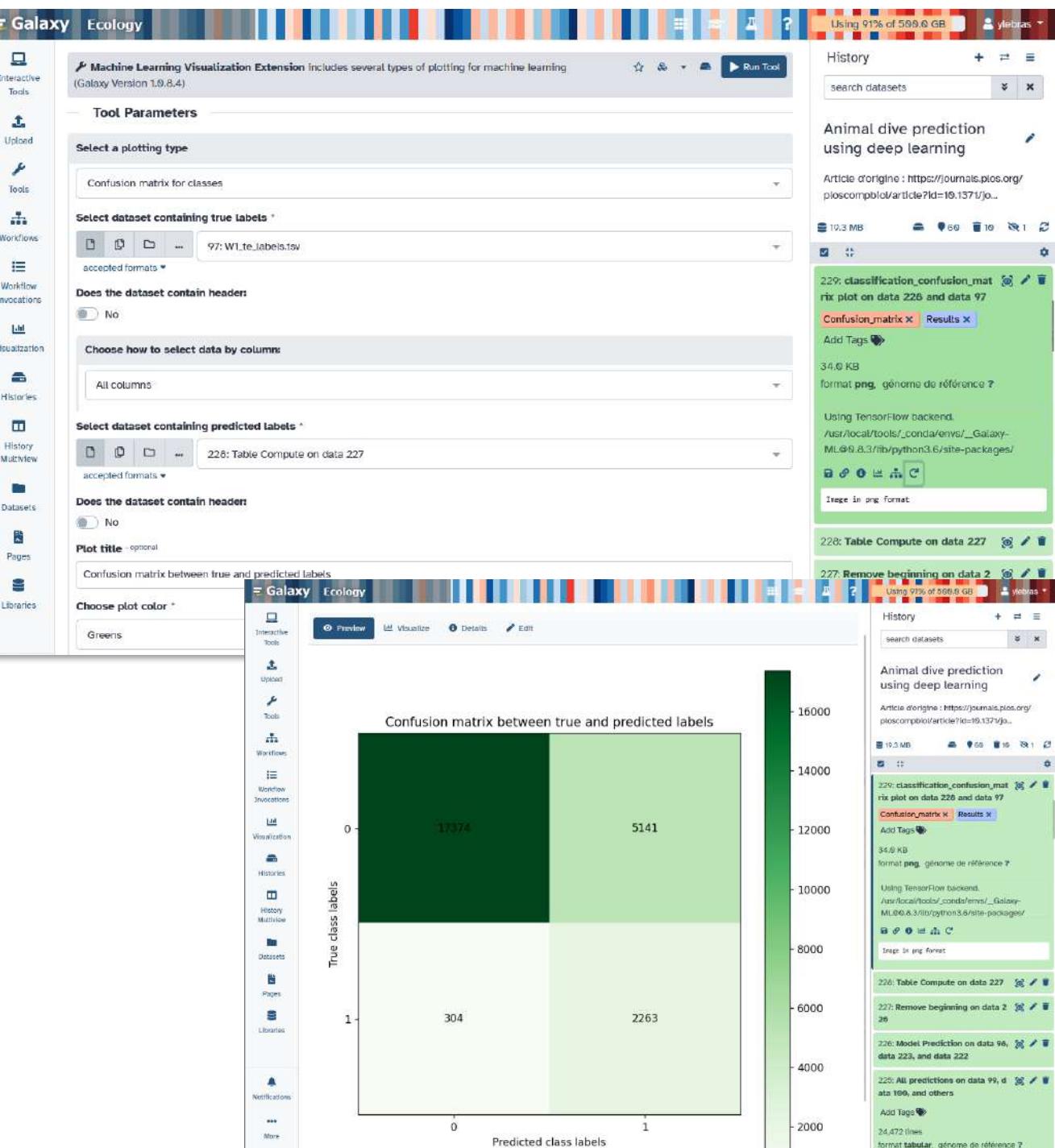
RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

Amédée Roy, Sophie Lanco Bertrand, Ronan Fablet

Version 2

Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>



Annotations & AI

Adapt existing tools to use cases
*From Interactive annotation to training
to labeling*



Annotations & AI



Adapt existing tools to use cases
From Interactive annotation to training to labeling



Annotations & AI

Adapt existing tools to use cases
From Interactive annotation to training to labeling

Galaxy Ecology

Perform YOLO training with ultralytics (Galaxy Version 6.3.0+galaxy2)

Tool Parameters

Input Images

Search for options

Accepted formats

Unselected (14) Selected (28)

131: MED_Annotation-Crisp2022_DSC_6494-copie.jpg.tif4af03db13fe790d5eb9af9e1cb65e69f.jpg
 130: MED_Annotation_DSC03718-laser-Zoanthid-lasers.jpg.tif.b22da45a29beaecd63c3117a7ed162872.jpg
 129: DSC09761.JPG.tif49da3cbb6d22fffd6649f1f7971.jpg
 128: DSC06117-massif-madrepore-lopelia-en-copie.jpg.tif.9873b3bca9d9a65e157b87f14f1935.jpg
 127: DSC06736-copie.JPG.tif.2101c79af7e7e547bbae5269a33d4672.jpg
 126: DSC08291-copie.JPG.tif.aae038be44fb0f1395e1f45ebfc32bcea.jpg
 125: DSC06114-plastique-Lophelia-madrepore.jpg.tif.a920de36f761bb5cf07ae987b15ee98.jpg
 124: Dicotylusum-rose_DSC06228-laser.jpg.tif.fec7f01eaaa3c83b0767b90113e3b63f.jpg

Shift to highlight range. Ctrl to highlight multiple.
 switch to simple select ▾

Input YOLO txt files

Search for options

Accepted formats

Unselected (21) Selected (28)

220: AnyLabeling Interactive on data 17: version.txt
 213: Training Metrics
 209: AnyLabeling Interactive on data 17: version.txt
 208: AnyLabeling Interactive on data 17: version.txt
 166: AnyLabeling Interactive on data 17: version.txt
 149: MED_Annotation-Crisp2022_DSC_6494-copie.jpg.tif4af03db13fe790d5eb9af9e1cb65e69f.txt
 139: Dicotylusum-rose_DSC06228-laser.jpg.tif.fec7f01eaaa3c83b0767b90113e3b63f.txt
 138: MED_Annotation_DSC03718-laser-Zoanthid-lasers.jpg.tif.b22da45a29beaecd63c3117a7ed162872.txt

Shift to highlight range. Ctrl to highlight multiple.
 switch to simple select ▾

The YOLO text files, each text file must correspond to one input image (same name different extension).

model_url

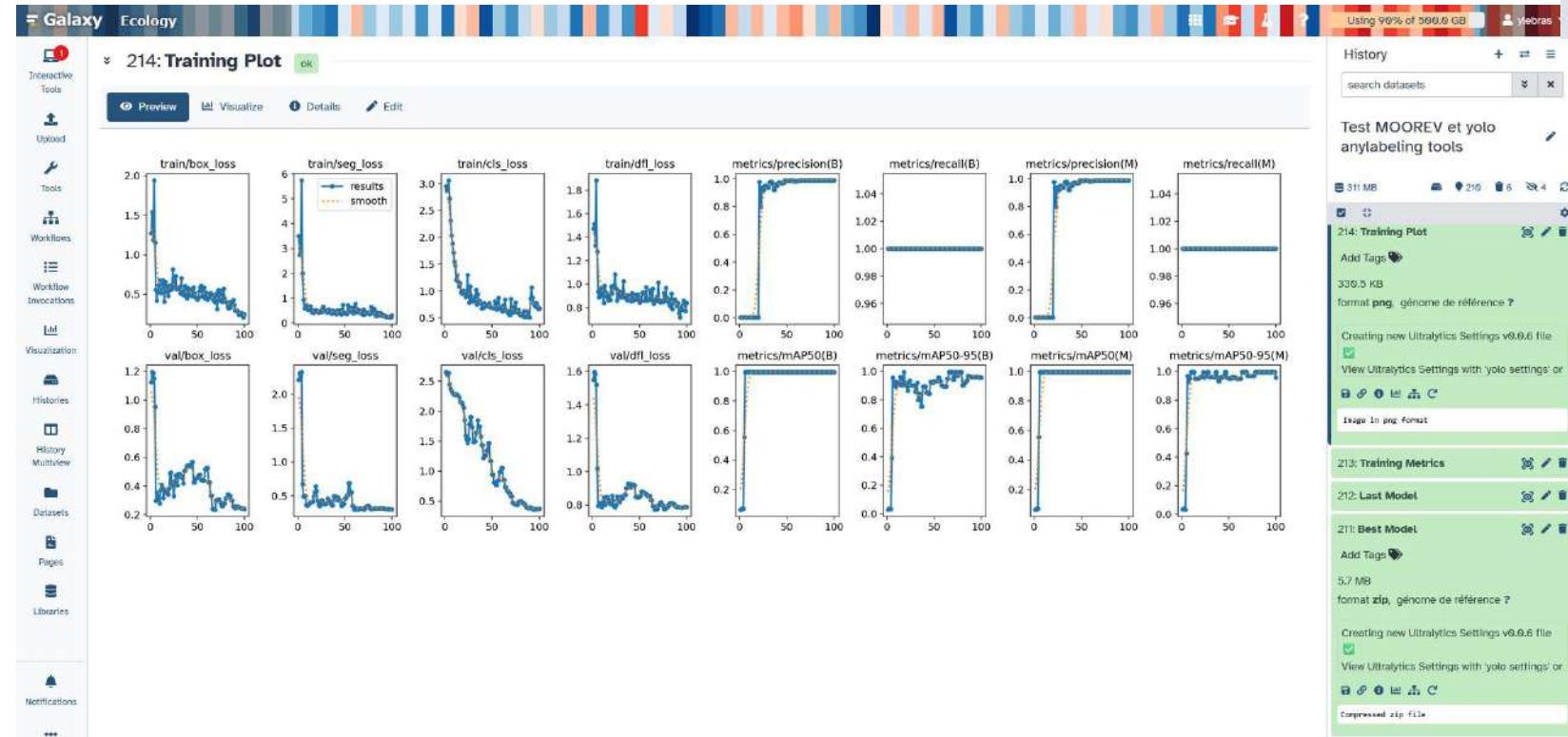
YOLOIn-seg

Training Parameters



Annotations & AI

Adapt existing tools to use cases
From Interactive annotation to training to labeling



Annotations & AI



Using Jupyter notebook From camera trap video to species detection with megadetector & pytorch-wildlife

Using 68% of 590.0 GB

History: test OFVi algo IA depuis script Hugo Magaldi

Jupyter output collection
a list with 1 csv dataset

Download Show Details Run Job Again

43: predictions_stride_1_thresh_0.2_2025-09-26 09-16-22

ok

Preview Visualize Details Edit

squirrel	water chevrotain	human	vehicle	Prediction	Confidence score
0.0	0.0	0.0	0.0	elephant	0.9638436758335076

1 line 39 columns, 1 comments
format csv, génome de référence ?

1.1.1filepath 2.1.1filename 3.1.1
filepath filename send
galaxy_inputs/test/videos/test_video.M4 test_video.M4 3.1.1

Using 68% of 590.0 GB

History

search datasets

test OFVi algo IA depuis script Hugo Magaldi

Using 68% of 590.0 GB

History

search datasets

test OFVi algo IA depuis script Hugo Magaldi

Interactive JupyterLab Notebook (Galaxy Version 10.2)

Tool Parameters

Do you already have a notebook?

41: Executed Jupyter Notebook

ok

Preview Visualize Details Edit

Classifying... 80% | 24/27 [23:15<03:00, 62.77s/step]
Classifying... 93% | 25/27 [24:02<01:55, 57.99s/step]
Classifying... 98% | 26/27 [24:52<00:55, 55.53s/step]
Classifying... 100% | 27/27 [25:37<00:08, 52.37s/step]
Classifying... 100% | 27/27 [25:37<00:00, 56.95s/step]

CONOLIDATING PREDICTIONS

In [18]: # Get a single prediction for each file (photo or video) by a majority vote over the detections
gbpy_dict = {}
gbpy_dict['filepath'] = 'first'
gbpy_dict['filename'] = 'first'
for taxon in taxons: ['human', 'vehicle']:
 gbpy_dict[taxon] = predictions.groupby('filepath').agg(gbpy_dict)
 predictions_grouped['Prediction'] = predictions_grouped[taxon_all].apply(lambda x: 'blank' if sum(x)==0 else taxons_all[x].argmax(x), axis=1)
 predictions_grouped['Confidence score'] = predictions_grouped[taxon_all].apply(lambda x: 'blank' if sum(x)==0 else np.max(x), axis=1)
 predictions_grouped.columns = [predictions_grouped.columns[i][0] for i in range(len(predictions_grouped.columns))]
 # Predict as 'blank' photos/videos for which there were no detections
 for filepath in taxons_all:
 if filepath not in predictions_grouped['filepath']:
 prediction_blank = pd.DataFrame([[filepath, filepath.rsplit(os.sep, 1)[1], 0]])(taxons_count) + 'blank', 1-detection_threshold), columns=list(predictions_grouped.columns))
 prediction_grouped = pd.concat([predictions_grouped, prediction_blank])
 predictions_grouped = predictions_grouped.sort_values(by='filepath', axis=0)
 predictions_grouped.to_csv(os.path.join(predictions_dir, 'predictions' + '_stride' + str(stride) + '_thresh' + str(detection_threshold) + '_' + timestamp + '.csv'), index=False)
print('PREDICTIONS SUCCESSFULLY SAVED TO FOLDER: ' + str(predictions_dir))

PREDICTIONS SUCCESSFULLY SAVED TO FOLDER: outputs/collection

42: Interactive JupyterLab Notebook on data 2, data 4, and others

41: Executed Jupyter Notebook

40: Jupyter output collection

a list with 1 csv dataset

39: OFVi_HugoMagaldi.ipynb

7: README.md

6: modelsetensors

5: test_video.M4

4: config.json

3: requirements.txt

2: functions.py

1: OFVi.py

Annotations & AI

Using Jupyter notebook From camera trap video to species detection with megadetector & pytorch-wildlife



```

In [18]: # Get a single prediction for each file (photo or video) by a majority vote over the detections
gby_dict = {}
gby_dict['filepath'] = 'first'
gby_dict['filename'] = 'first'
for taxon in taxons: ['human', 'vehicle']:
    gby_dict[taxon] = 'first'
predictions_grouped = predictions.groupby('filepath').agg(gby_dict)
predictions_grouped['Prediction'] = predictions_grouped[taxon_all].apply(lambda x: 'blank' if sum(x)==0 else taxons_all[x].argmax(), axis=1)
predictions_grouped['Confidence score'] = predictions_grouped[taxon_all].apply(lambda x: 'blank' if sum(x)==0 else np.max(x), axis=1)
predictions_grouped.columns = [predictions_grouped.columns[i] for i in range(len(predictions_grouped.columns))]
# Predict as 'blank' photos/videos for which there were no detections
for filepath in filepaths_all:
    prediction_blank = pd.DataFrame([{'filepath': filepath, 'filepath': filepath, 'taxon': 'blank', 'taxon': 'blank', 'confidence': 0}], columns=list(predictions_grouped.columns))
    prediction_grouped = pd.concat([predictions_grouped, prediction_blank])
predictions_grouped = predictions_grouped.reset_index(drop=True)
predictions_grouped = predictions_grouped.sort_values(by='filepath', axis=0)
predictions_grouped.to_csv(os.path.join(predictions_dir, 'predictions' + str(stride) + str(stride) + '_threshold' + str(detection_threshold) + '_'+ timestamp + '.csv'), index=False)

```

PREDICTIONS SUCCESSFULLY SAVED TO FOLDER: 'predictions_dir'

Annotations & AI

New: Using Hugging Face models in Galaxy Ecology

User Preferences / My Repositories / Create New / Hugging Face Hub 😊

Create a Hugging Face Hub 😊 File Source

Name *

Label this new file source with a name.

Description - optional

Provide some notes to yourself about this file source - perhaps to remind you how it is configured, where it stores the data, etc..

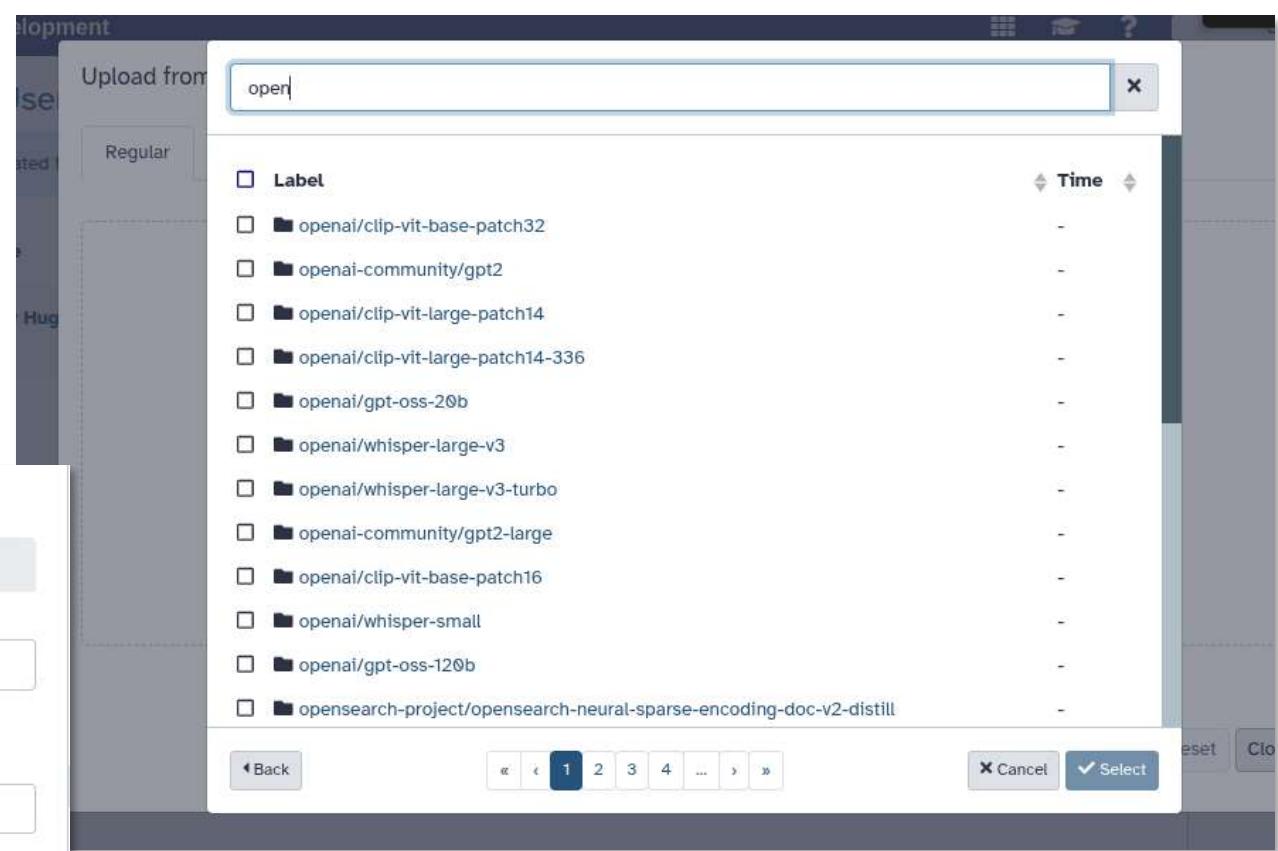
Hugging Face Hub Endpoint

Custom endpoint of the Hugging Face Hub you are connecting to. This should be the full URL including the protocol (http or https) and the domain name. You can leave this blank to use the default Hugging Face Hub endpoint (https://huggingface.co).

Hugging Face Access Token

The personal access token to use to connect to the Hugging Face Hub. You can generate a new token in your Hugging Face account settings. This will allow Galaxy to access private models if you have the necessary permissions.

Create



Galaxy Ecology initiative: A lot of young contributors!



Clara Urfer

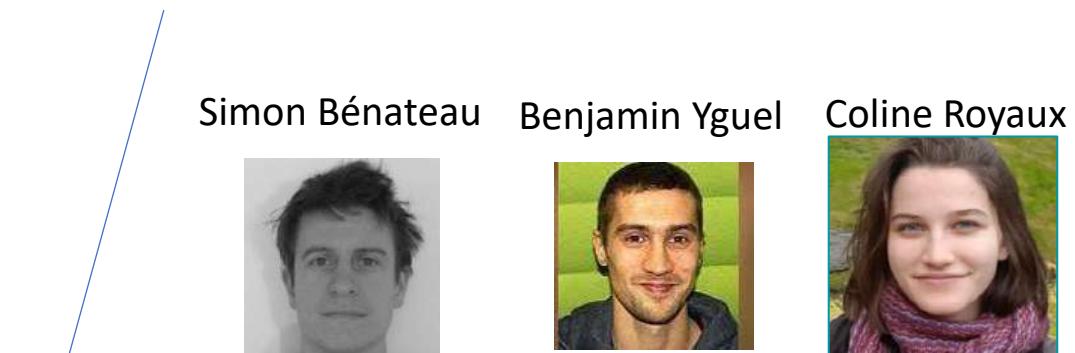
Elisa Michon

Timothée Virgoulay

Eloïse Trigodet

Valentin Chambon

Alan Amossé



2018

2019

2020

2021



Claire Dussin Elouan Le Mestric Jean Le Cras

2021

2022

2023

2024

2025

Molène Mahé

Kévin Payet

Yassine Ankerl



Najat Amoukou



Arthur Barreau

Triskell Cumunel

