

IceNet@GalaxyClimate

Vanessa Stoeckl, University of Freiburg, Germany

Thanks to Alejandro Coca-Castro, Anne Fouilloux, Bjoern Gruening, Jean Iaquina

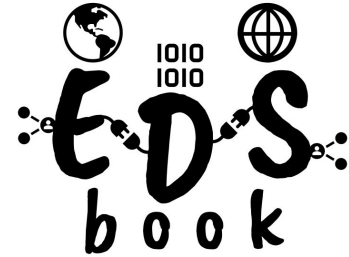


DOI: [10.5281/zenodo.7723371](https://doi.org/10.5281/zenodo.7723371)

Agenda

- Environmental Data Science book
- IceNet@EDS
- Galaxy Project for Climate Science
- IceNet@Galaxy

Environmental Data Science (EDS) book



A Community-driven resource for Environmental Scientists

Reproducible, scalable, & shareable
ENVIRONMENTAL DATA SCIENCE



Scribbles



Contribution



Collaborative
Reviewing



Publication



A screenshot of a digital object metadata page. The page has a dark grey background with white text. At the top, there is a 'Title' section with a tag icon and the text 'Tags (Environment, Theme)'. Below this is a 'RoHub FAIR Executable Research Object' section with 'launch' and 'binder' buttons. The 'Context' section includes the text 'purpose, highlight, contributions'. The 'Data' section is empty. The 'Analysis' section is empty. The 'Citation' section is empty. At the bottom right, there is a 'jupyter' logo.

Open Science and Community of Practice with EDS book

Environmental Data Science book

A Community-driven resource for Environmental Scientists



Reproducible/Reusable



Open-source tools



Shareable
jupyter {book}

Overview Content Quality Activity Life cycle Relations Impact

Reliance

Created: 27.03.2022 (21:35), last modified: 08.04.2022 (10:39)

0.00 / 5 0 0

23 Downloads **3** Views

Hide more details

Resources	10
Annotations	41
Events	71
Forks	0
Snapshots	0
Archives	0
Size	831.45 KB

AGENTS

The Environmental Data Sc... Creator

QUALITY 100%

KEYWORDS: DEEP LEARNING

TOOLBOX

SHARE

CITE AS

Sebastian H. M. Hickman, and Alejandro Coca-Castro. "Tree crown delineation using detectreeRGB (Jupyter Notebook) published in the Environmental Data Science book." ROHub. Mar 27 2022. <https://w3id.org/ro/id/94486a74e04646116bb9334ec7b57040>.

LICENSE MIT

Content:

- tool
- Lock conda file for linux-64
- Lock conda file for osx-64
- Conda environment
- Pip requirements for lock conda environments
- Jupyter notebook
- Online rendered version of the Jupyter notebook

IceNet - Seasonal Arctic Sea Ice Forecasting with Probabilistic Deep Learning



Alejandro Coca-Castro (The Turing Institute, UK) is an early adopter of the RELIANCE services and created a Research Object showing how to use IceNet to make seasonal sea ice forecasts.

nature communications

Explore content ▾ About the journal ▾ Publish with us ▾

nature > nature communications > articles > article

Article | [Open Access](#) | [Published: 26 August 2021](#)

Seasonal Arctic sea ice forecasting with probabilistic deep learning

[Tom R. Andersson](#), [J. Scott Hosking](#), [María Pérez-Ortiz](#), [Brooks Paige](#), [Andrew Elliott](#), [Chris Russell](#), [Stephen Law](#), [Daniel C. Jones](#), [Jeremy Wilkinson](#), [Tony Phillips](#), [James Byrne](#), [Steffen Tietsche](#), [Reena Balan Sarojini](#), [Eduardo Blanchard-Wrigglesworth](#), [Yevgeny Aksenov](#), [Rod Downie](#) & [Emily Shuckburgh](#)

Nature Communications 12, Article number: 5124 (2021) | [Cite this article](#)

18k Accesses | 25 Citations | 328 Altmetric | [Metrics](#)

Abstract

Anthropogenic warming has led to an unprecedented year-round reduction in Arctic sea ice extent. This has far-reaching consequences for indigenous and local



The screenshot shows the 'EXECUTABLE RESEARCH OBJECT' interface for 'Sea ice forecasting using IceNet'. It includes a map of the Arctic region, a 'Content' section with a 'tool' and 'conda environment', and a 'Jupyter' logo. The interface is designed for interactive use, allowing users to explore the data and code used in the research.

Open Access paper in *Nature Communications* where IceNet, a probabilistic deep learning method, has been developed for seasonal sea ice forecasts: **data and codes are available and can be reused.**

FAIR Jupyter Notebook

Acknowledge all contributions:
executable RO tracks all the
contributions.

PUBLIC MANUAL LIVE EXECUTABLE RESEARCH OBJECT ENVIRONMENTAL DATA SCIENCE BOOK COMMUNITY JUPYTER NOTEBOOK

CLIMATOLOGY ENVIRONMENTAL RESEARCH

Sea ice forecasting using IceNet (Jupyter Notebook) published in the Environmental Data Science book

Alejandro Coca-Castro

Contributed by Tom Andersson, Nick Barlow
Published by Environmental Data Science Book Community



Overview Content Completeness Enrichment Activity Life cycle Relations Impact

The research object refers to the Sea ice forecasting using IceNet notebook published in the Environmental Data Science book.

☆ 0.00 / 5 🗨️ 0 🍷 0

8 Downloads 12 Views

[Hide more details](#)

📁 Resources	10
📄 Annotations	45
📅 Events	87
🍴 Forks	1
📷 Snapshots	0
🗃️ Archives	0
📏 Size	348.46 KB

AGENTS

Environmental Data Science Book C...
Creator

COMPLETENESS 84%

DISCOVERED METADATA: 0

PUBLISHING EARTH SCIENCES
PHYSICAL GEOGRAPHY AND ENVIRONMENTAL
BOOK INDUSTRY LITERATURE LANGUAGE
GEOSCIENCES GEOPHYSICS

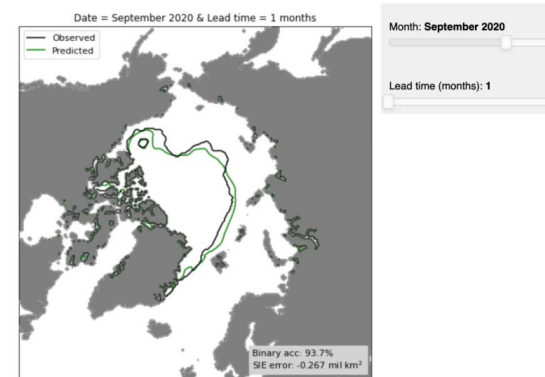
TOOLBOX



SHARE



CITE AS



CONTENT

Image showing interactive plot of IceNet seasonal forecasts ... (337Kb)

tool

- Conda environment
- Lock conda file for linux-64
- Lock conda file for osx-64
- Jupyter notebook
- Online rendered version of the Jupyter notebook

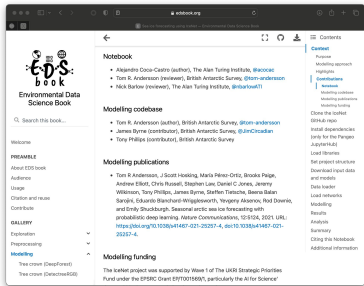
Reusable & Reproducible

Binder

Open-source tools



Rendered online notebook



The Galaxy Project and the Galaxy Climate Workbench

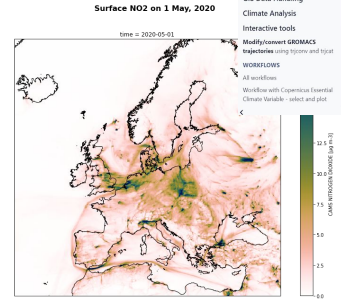
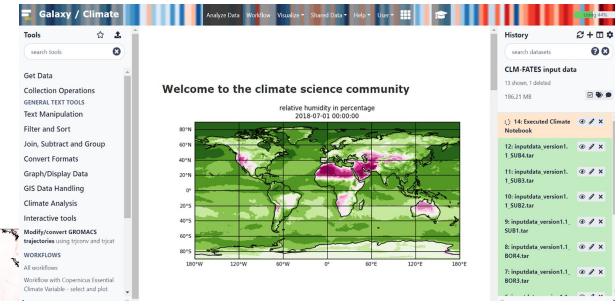


Galaxy open-source platform for FAIR data analysis offers:

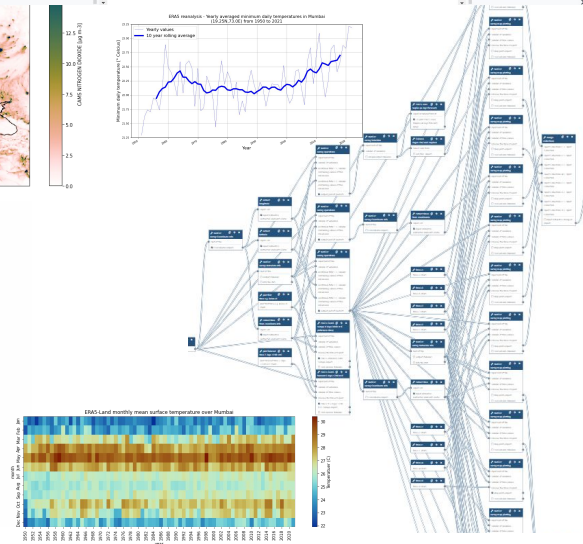
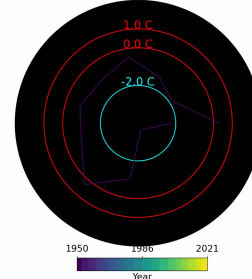
- Pangeo and Climate Science jupyter notebook deployment (dask local) **available to everyone (free registration);**
- Galaxy Tools for fully **automated climate science workflows;**
- **GUI** for users with **no programming skills;**
- Self-Paced **Learning material** and organisation of online training events with the Galaxy Training Network;
- **Training Infrastructure as a Service** is a free and ready to use with private queues where only training's jobs run.



<https://climate.usegalaxy.eu/>

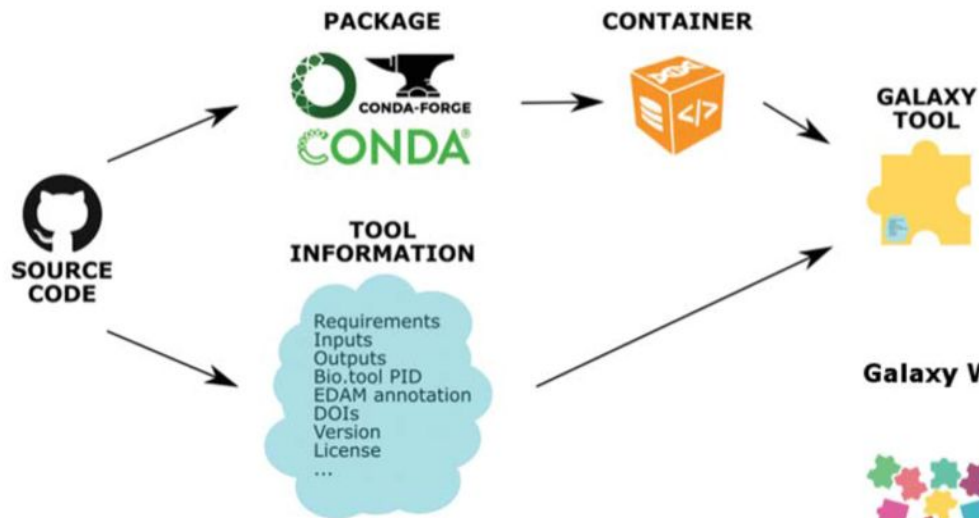


Temperature anomalies in Mumbai (cf. 1980-2010)



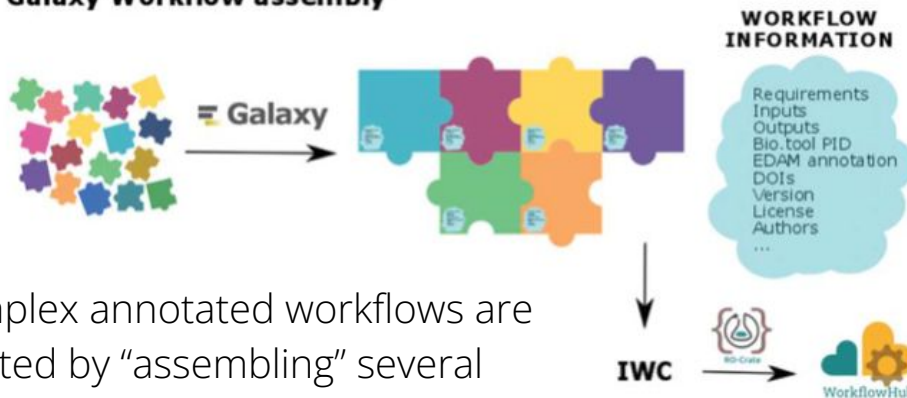
Accessible FAIR and Open Science

Create **FAIR** tools and workflows



For each release, a new version of the corresponding **Galaxy Tool is automatically created** and published in the Galaxy Tool Repository (Galaxy AppStore).

Galaxy Workflow assembly



Each step is reviewed



Complex annotated workflows are created by “assembling” several Galaxy Tools.

From FAIR Jupyter Notebook to Annotated Galaxy Tools & Workflows



One Jupyter Notebook is turned into one script

EDS book Jupyter Python/R/Julia Notebook

Set project structure

Let's follow the structure of the iceNet paper as it is indicated in the source code config.py file. The structure allows conveniently using iceNet's custom data loader.

```

# data_folders
data_folders = {
    'data': './data',
    'sourcecode_folders' : './sourcecode_folders'
}

def get_data_folders():
    """
    Return the list of data folders.
    """
    return data_folders

def get_sourcecode_folders():
    """
    Return the list of sourcecode folders.
    """
    return sourcecode_folders
    
```

Download input data and models

iceNet consists of 25 ensemble members (i.e. models). For this demonstrator, we only download three of them to reduce computational cost (note that this will reduce performance compared with the full ensemble). We also fetch analysis-ready (i.e. preprocessed data of stream observations, ground truth sea ice concentration (SIC) and a iceNet's project configuration file from a Zenodo repository. Finally, we call a script from the iceNet paper repo to generate masks required for computing metrics and visualization.

Download pretrained iceNet models

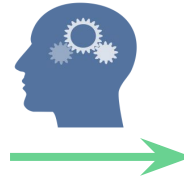
We downloaded 3 out of 25 ensemble members (i.e. models). For this demonstrator, the models are numbered from 36 to 60. For this example we use the networks 36, 42 and 53. We wish to mention other pre-computed results from the Nature Communications paper can be downloaded including output results table, uncertainty, netCDF forecast of the 25 ensemble members, among others.

```

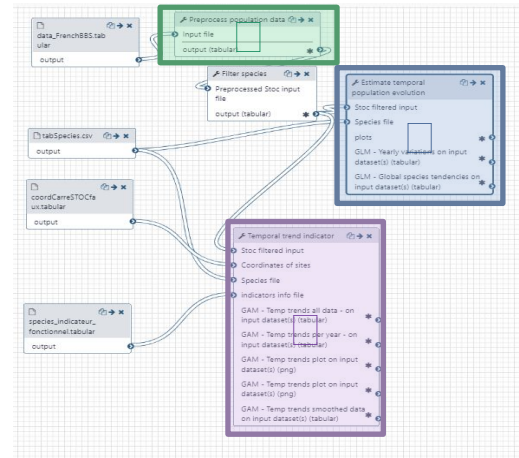
tab[is.na(tab)] <- 0
# filename <- "trouverUnNom"
chemin <- paste(rep, filename, sep = "/")
write.table(tab, chemin)
colnames(tab) <- sub("nombre.", "", colnames(tab))
return(tab)

## sous jeux de donnees si choix d'espece d'annee ou d'un pourcentage de carres
makeSourceTab <- function(tab, vecsp, methodeEchantillon, "carre", vecannees=NULL) {
  flush.console()
  # reduction de la table a certaines especes
  if(!is.null(vecsp)) {
    cat(" selection", length(vecsp), "espece(s):\n -> ")
    cat(vecsp)
    cat("\n")
    tab <- data.frame(carre = tab$carre, annee = tab$annee, tab[,vecsp])
    colnames(tab) <- c("carre", "annee", vecsp)
  }
  ## reduction de la table pour certaines annees
  if(!is.null(vecannees)) {
    tab <- subset(tab, annee ~ vecannees[1] & annee ~ vecannees[2])
  }

  if(echantillon != 1) {
    if(echantillon < 1 & echantillon > 0) {
      nbinit <- nrow(tab)
      if(methodeEchantillon == "global") {
        nb <- round(nrow(tab) * echantillon)
        cat(" echantillonage", echantillon)
        cat("\n")
        chat <- conservation de "nb" lignes sur "nbinit."
        flush.console()
        tab <- tab[sample(1:nrow(tab))[(1:nb)]]
      } else {
        (methodeEchantillon == "carre") {
          cat(" echantillonage", echantillon)
          cat("\n")
          % des carres par la methode, methodeEchantillon, "\n")
          nbcarrinit <- length(unique(tab$carre))
          chat <- sample(unique(tab$carre),
            length(unique(tab$carre)) * echantillon, replace=F)
          cat(" conservation de", length(chat), "carres sur",
            nbcarrinit)
          tab <- subset(tab, subset = carre %in% chat)
          cat(" (" , nrow(tab), " lignes sur ", nbinit, ")\n", sep="")
        } else {
          stop("Methode d'echantillonage non reconnue")
        }
      }
    }
  }
}
    
```



Several atomic scripts from which generic and fully annotated Galaxy Tools are created



From FAIR Jupyter Notebook to Annotated Galaxy Tools & Workflows



One Jupyter Notebook is turned into one script

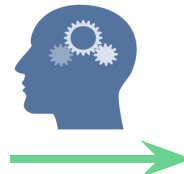
EDS book Jupyter
Python/R/Julia
Notebook

Title
Tags (Environment, Theme)
RoHub FAIR Executable Research Object
launch binder
Context
purpose, highlight, contributions
Data
Analysis
Citation

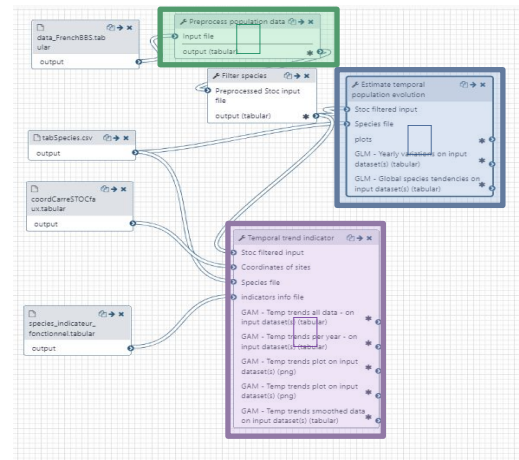
```
direction=wide )
tab[is.na(tab)] <- 0
# filename <- "trouverUnNom"
# chemin <- paste(rep(filename,sep="/")
# write.table(tab, chemin)
colnames(tab) <- sub("nombre.", "", colnames(tab))
return(tab)

## sous jeux de donnees si choix d'espece d'annee ou d'un pourcentage de carres
makeSubstTab <- function(tab, vecspp, methodeEchantillon=1, vecannees=NULL) {
  cat("-- Fabrication du sous jeu de donnees --\n")
  flush.console()
  ## reduction de la table a certaines especes
  if(!is.null(vecspp)) {
    cat("selection", length(vecspp), "espece(s):\n -> ")
    cat(vecspp)
    cat("\n")
    tab <- data.frame(carre = tabs$carre, annee = tabs$annee, tab[, vecspp])
    colnames(tab) <- c("carre", "annee", vecspp)
  }
  ## reduction de la table pour certaines annees
  if(!is.null(vecannees)) {
    tab <- subset(tab, annee ~ vecannees[1] & annee <= vecannees[2])
  }

  if(echantillon != 1) {
    if(echantillon < 1 & echantillon > 0) {
      nbinit <- nrow(tab)
      if(methodeEchantillon == "global") {
        nb <- round(nrow(tab)*echantillon)
        cat("echantillonage", echantillon, "\n")
        "% des donnees par la methode", methodeEchantillon, "\n")
        chat <- conservation de "nb", "lignes sur", nbinit, "\n")
        flush.console()
        tab <- tab[sample(1:nrow(tab))[1:nb], ]
      } else {
        if(methodeEchantillon == "carre") {
          cat("echantillonage", echantillon, "\n")
          "% des carrees par la methode", methodeEchantillon, "\n")
          nbcarreinit <- length(unique(tab$carre))
          chat <- sample(unique(tab$carre),
            length(unique(tab$carre))*echantillon, replace=F)
          cat("conservation de", length(chat), "carrees sur",
            nbcarreinit, "\n")
          tab <- subset(tab, carre %in% chat)
          cat("(", nrow(tab), " lignes sur ", nbinit, ")", "\n", sep="")
        } else {
          stop("Methode d'echantillonage non reconnue")
        }
      }
    }
  }
}
```



Several atomic scripts
from which generic and
fully annotated Galaxy
Tools are created



EDS book notebooks follow best practices: it makes it easier to “modularize” and create Galaxy Tools.

From FAIR jupyter notebook to annotated Galaxy Tools & workflows

IceNet Jupyter Notebook from EDS book

Title

Tags (Environment, Theme)

RoHub FAIR Executable Research Object

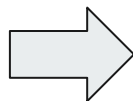
launch binder

Context
purpose, highlight, contributions

Data

Analysis

Citation



Galaxy workflow

Data

[Pretrained IceNet models](#)

[CMIP6 input data](#)

[ERA5 input data](#)

Analysis

Pre-processing for ERA5

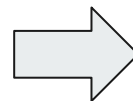
Pre-processing for CMIP6

DataLoader with [dataconfig](#)

IceNet ModelLoader

Generate ensemble forecast

Aggregate results



Copernicus Climate Data Store (C3S) Galaxy Tool; already available

Interactive dashboard & Galaxy Tools

Visualization

Create Interactive plot to show forecast, etc.

IceNet Forecast Galaxy Tools to develop.

IceNet@Galaxy Tool



Tools can be run from the command line or from a GUI using the Galaxy web portal.



Thank you!



DOI: [10.5281/zenodo.7723371](https://doi.org/10.5281/zenodo.7723371)