nowcAstIng

Nowcasting of fog using AI (NAI) techniques using satellite observations.

nowcAstIng

A toolbox for developing fog ML models

Llorenç Lliso (AEMET): jllisov@aemet.es

Outline:

- nowcAstIng toolbox
- Current work

nowcAstIng NWC SAF toolbox

The toolbox is proposing NWC SAF cloud products along with METAR reports to develop fog nowcasting ML models. Satellite radiances are also accessible for EUMETSAT registered users.

The users can develop, train and reuse their models for a particular airport.

The objective is to exploit the information coming from satellite, along with METAR reports, to nowcast fog some hours in the future

nowcAstIng

nowcAstIng

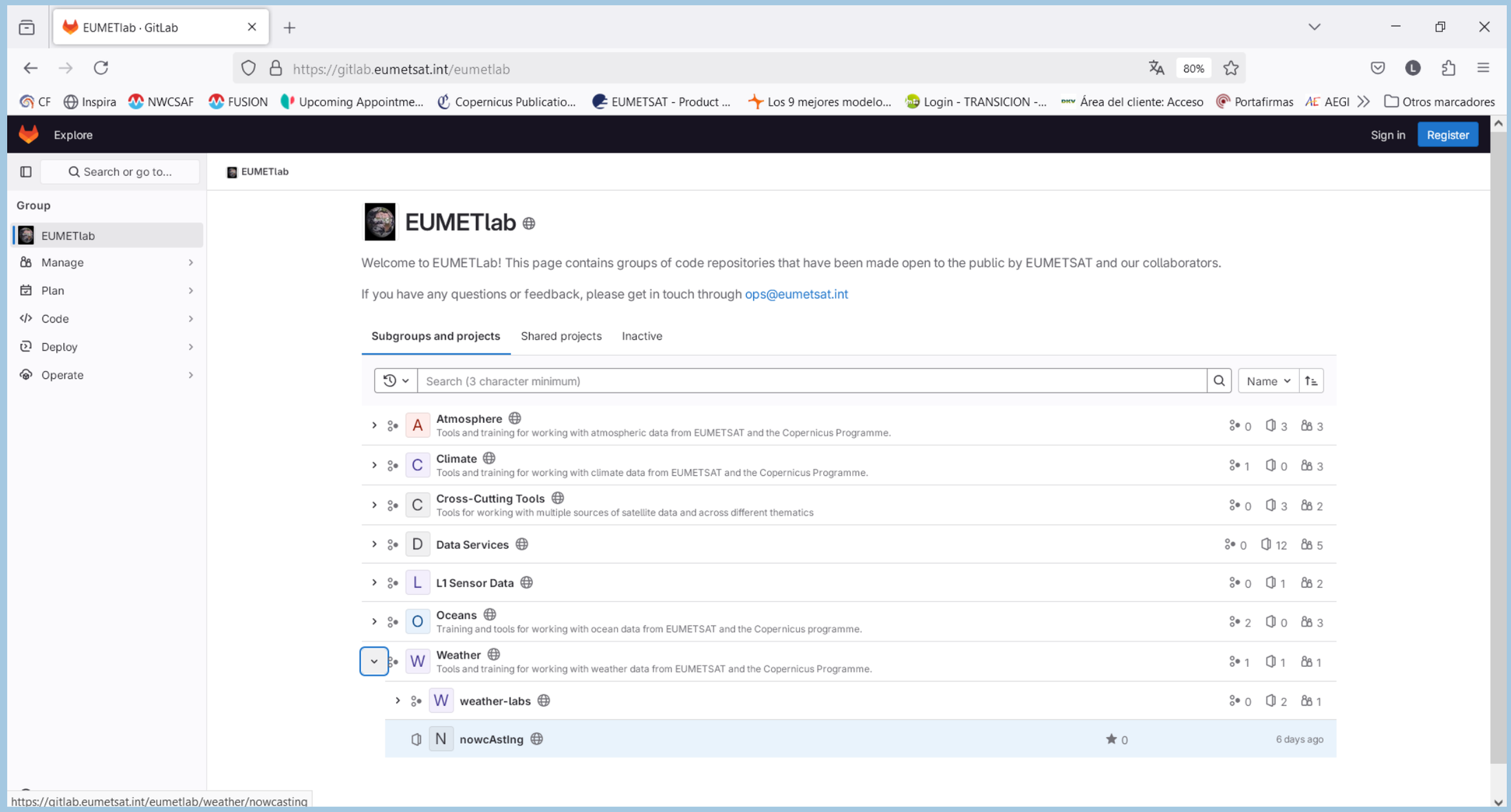
A toolbox for developing fog nowcasting ML models

https://gitlab.eumetsat.int/eumetlab/weather/nowcasting

nowcAstIng is available on EUMETLab.

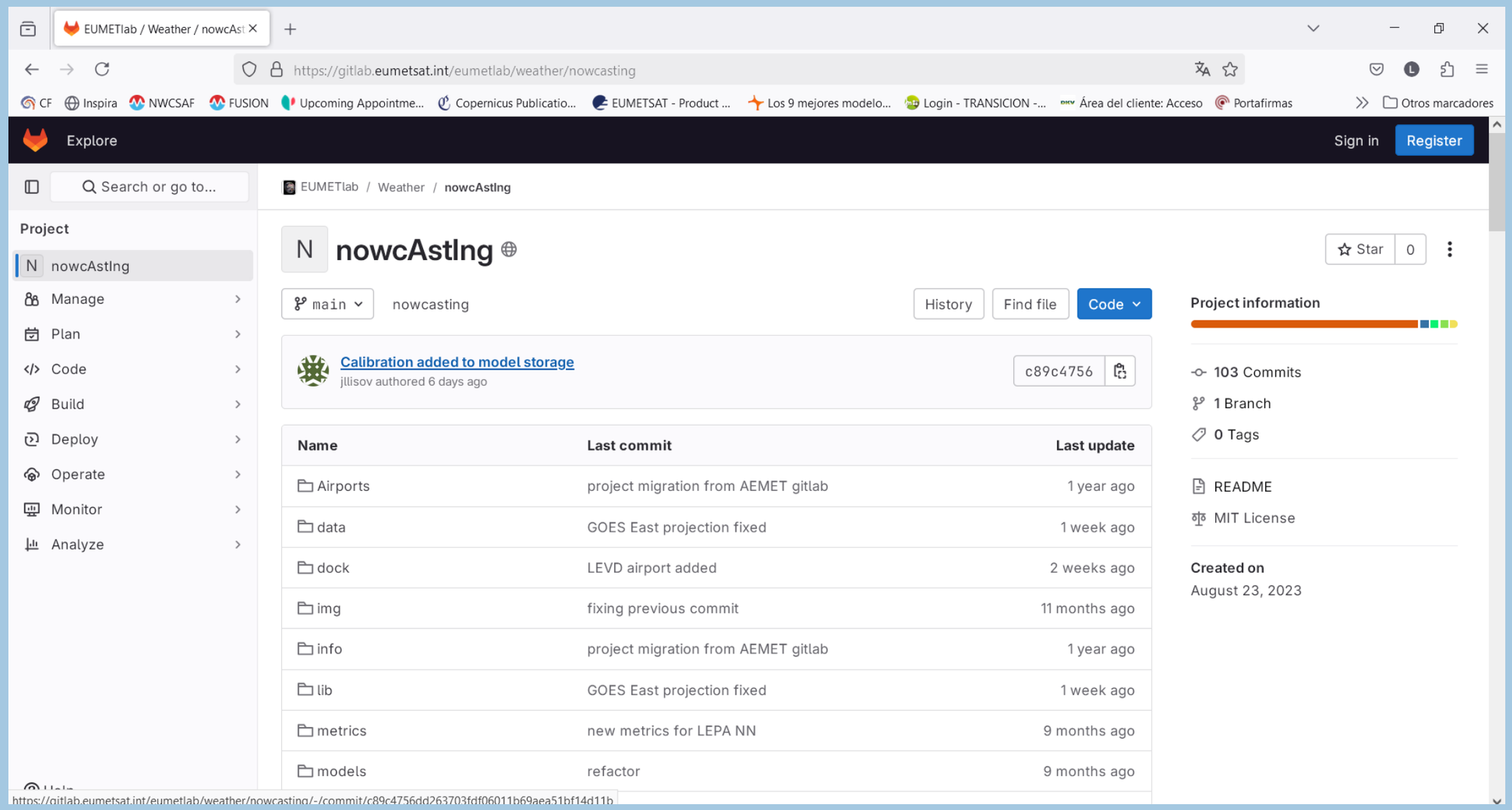
The toolbox is released with a capacity-building approach under the umbrella of the EUMETSATs' EUMETLab. You can clone it from:

https://gitlab.eumetsat.int/eumetlab/weather/nowcasting.git



nowcAstIng is released under MIT License.

The nowcAstIng targeted user are Machine Learning practitioners, Academia and NMS. These communities are encouraged to fork the project, give feedback and contribute.
The nowcAstIng toolbox is in the format of a Jupyter notebook collection.



nowcAstIng targets to co-development.

The nowcAstIng targeted user are Machine Learning practitioners, Academia and NMS. **These communities are encouraged to fork the project, give feedback and contribute.**

The nowcAstIing toolbox is in the format of a Jupyter notebook collection, allowing:
1. To download the NWC SAF clouds (and sat channels under license) and collocate them with METAR reports
2. To get insight on the fog formation ;
3. To establish baseline models ;
4. To develop ML models to forecast the occurrence of fog ;
5. To store their models for operational reuse.

The data proposed to feed the ML models are NWC SAF cloud and METAR data.

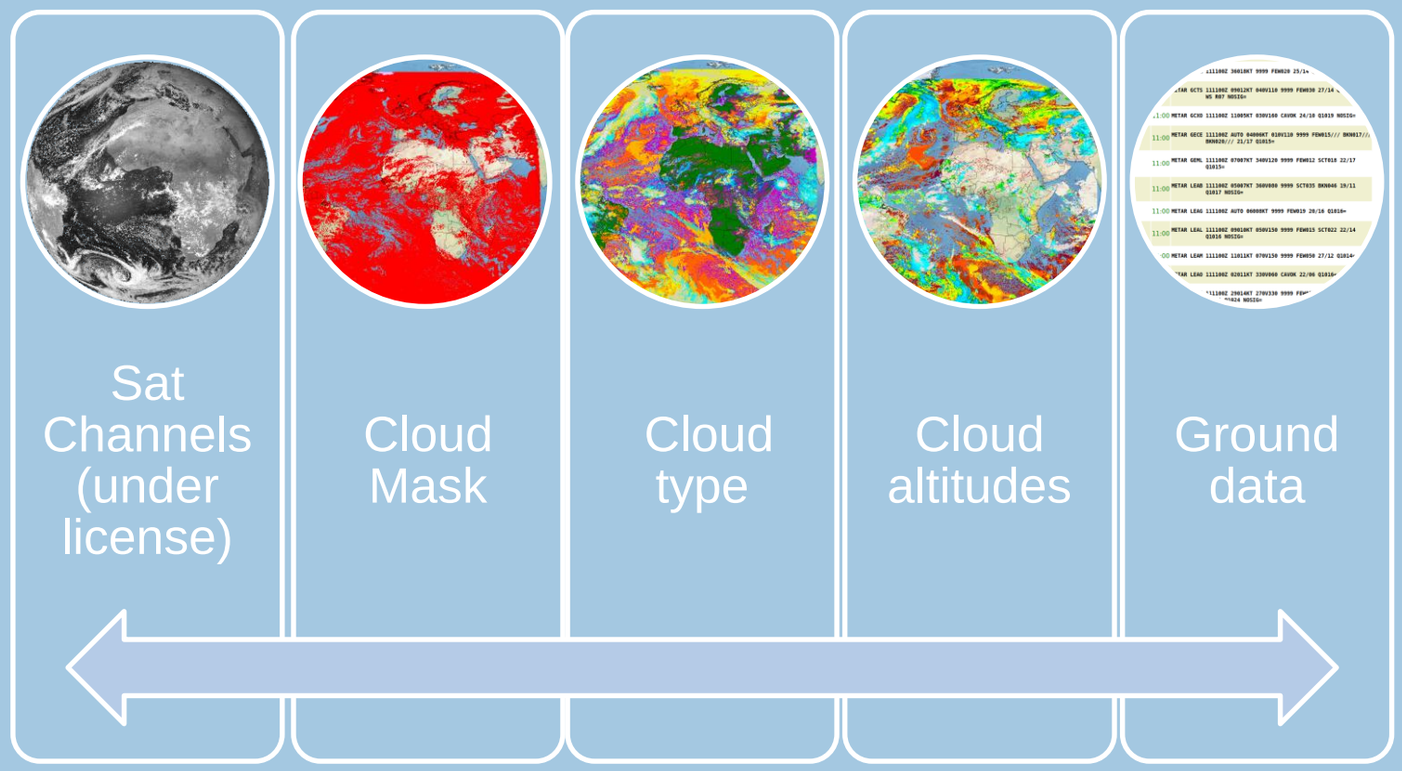


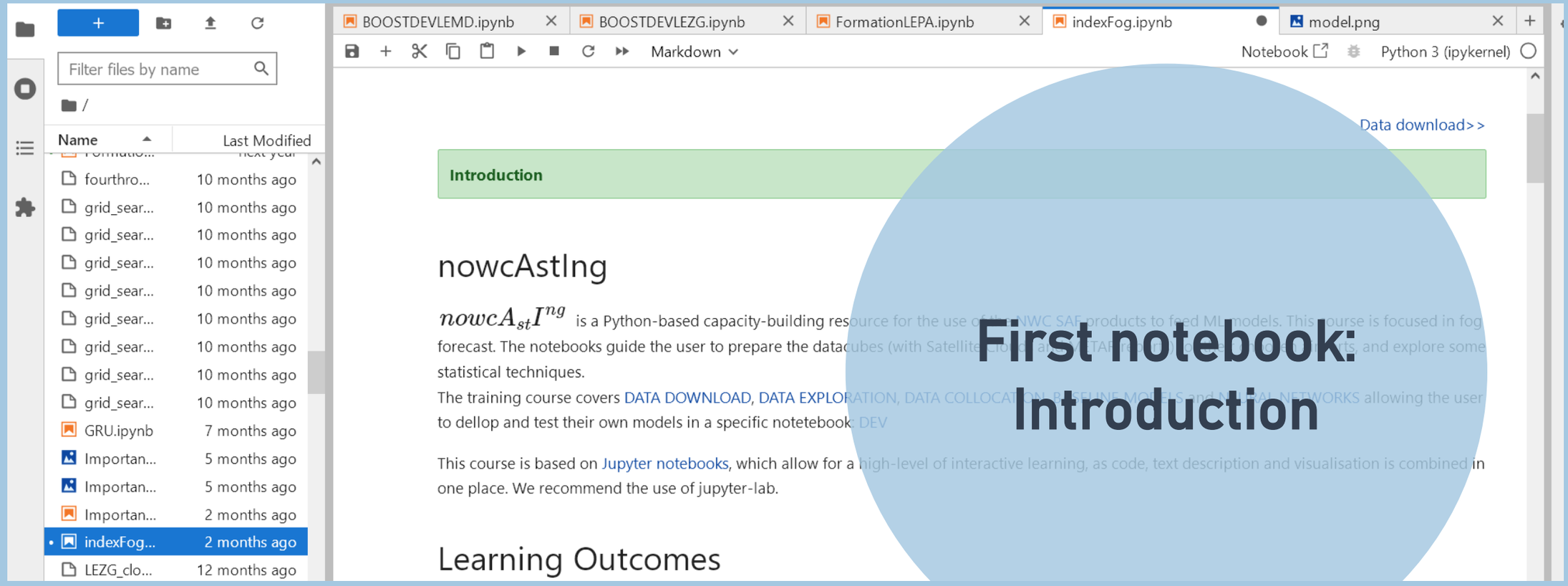
Photo: Andrew Malone via Wikimedia Commons

The set of notebooks is proposing an example of fog forecasting ML model construction.

The notebooks one by one:

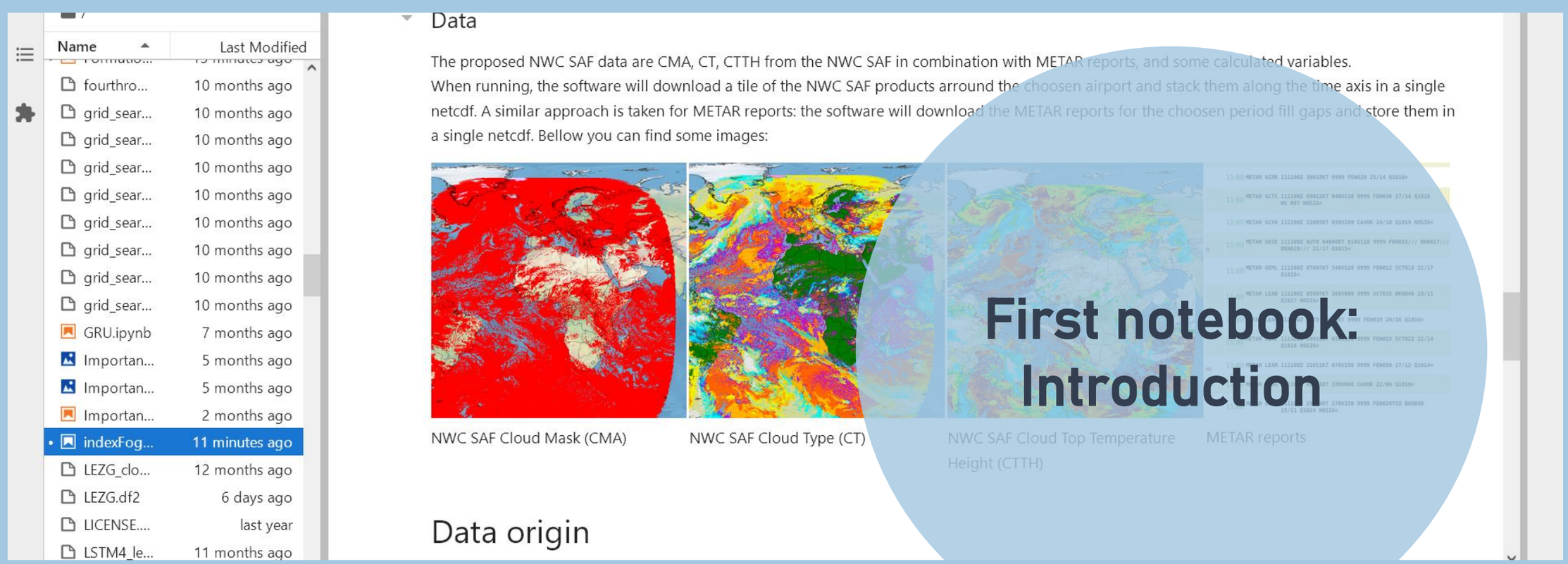
The first notebook gives a general description:
It explains the proposed data to train the model: NWC SAF cloud products, Information extracted from METAR reports and some geophysical calculated variables (as solar elevation...). The first notebook includes an index pointing the user to each chapter of the course.

The toolbox download the data, store them, and make them accessible via xarray objects. The xarray Datasets follow CF convention.

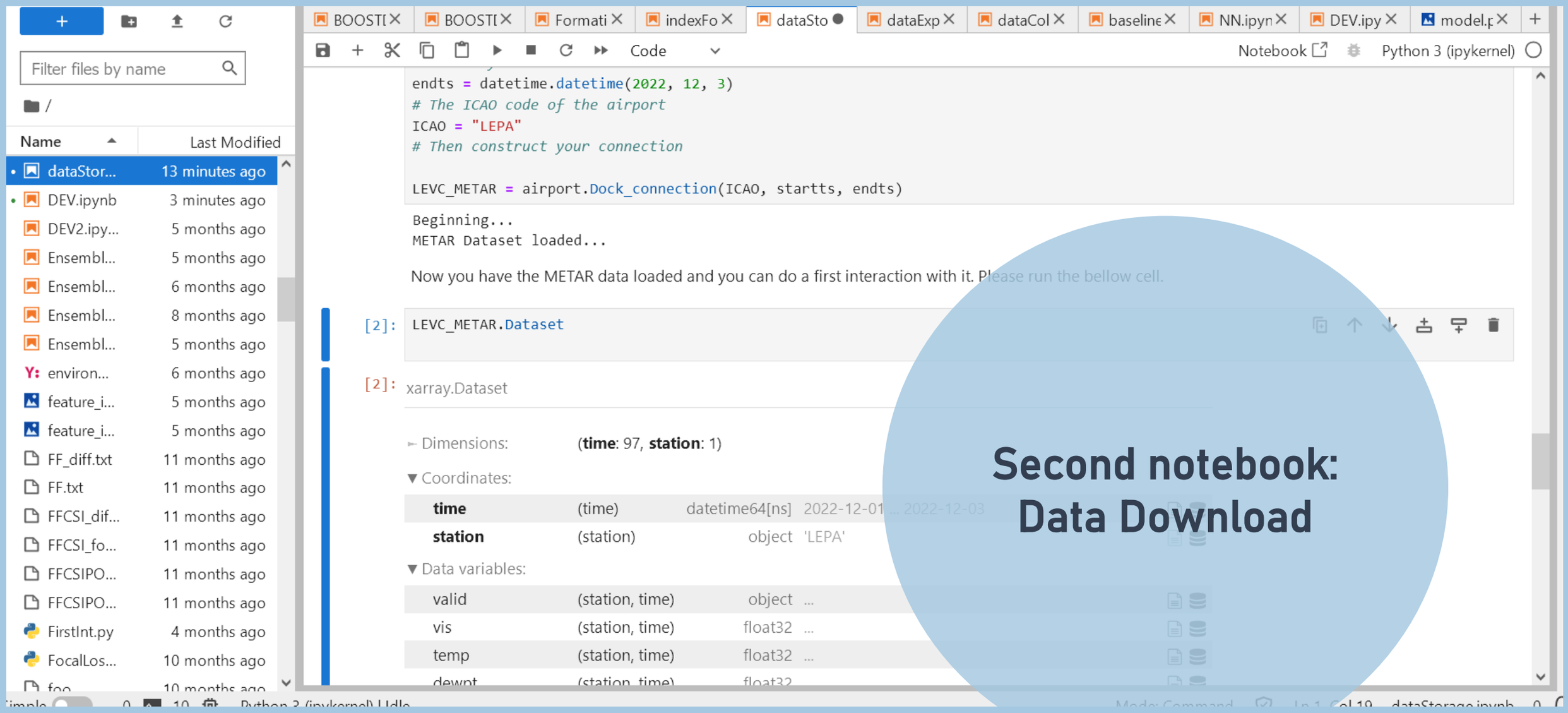


The first notebook also explains the learning outcome: at the end of the series the user:

- Will be familiarized with the content of NWC SAF GEO Cloud Products.
- Will be more familiar with information extracted from METAR reports.
- Will know how to merge both data: clouds and METAR.
- Will see an example of Naïve Bayes implementation.
- Will see an example of a NN basic implementation for a particular airport.
- Will learn how to store and compare its own ML models

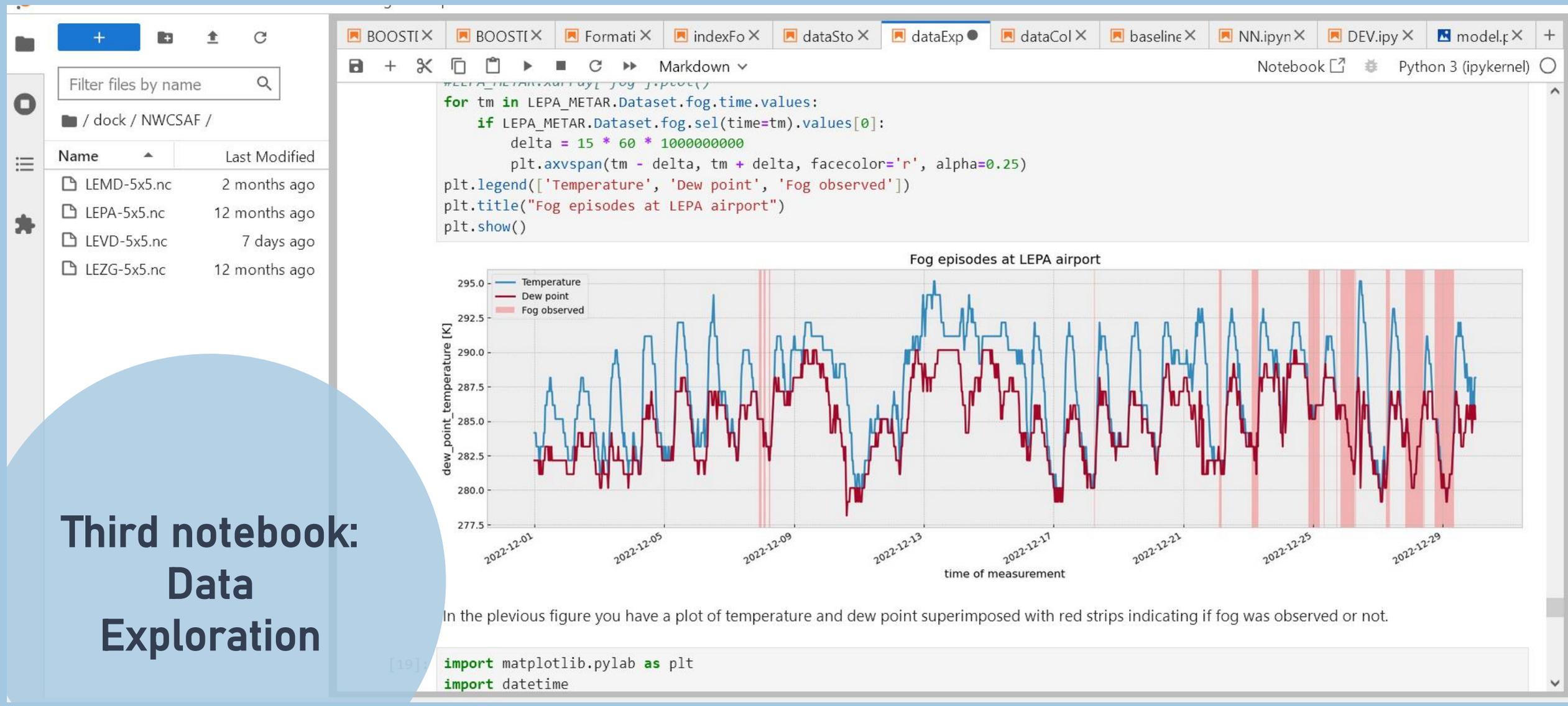


The second notebook gives some general ideas and explains how to download the data for a particular airport. It also makes some considerations on data attribution and where to get credentials to access the data.
The software downloads the data and pack the downloaded in two individual netCDF CF files for reusing. The data are exposed as xarray Datasets.



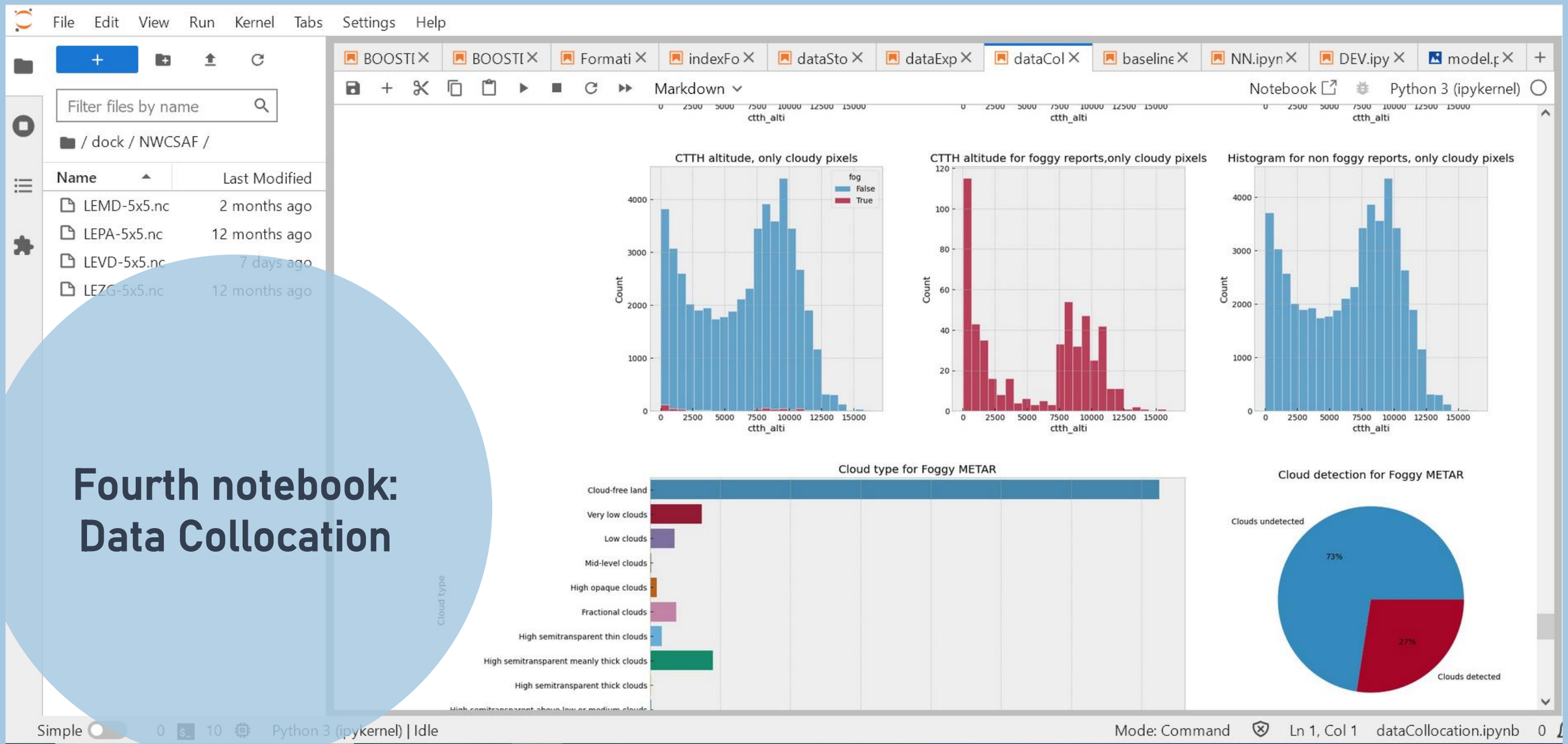
Second notebook: Data Download

The third notebook teaches how to connect with the stored data, how to plot some basic graph. The objective is to gain insight on the data characteristics



Third notebook: Data Exploration

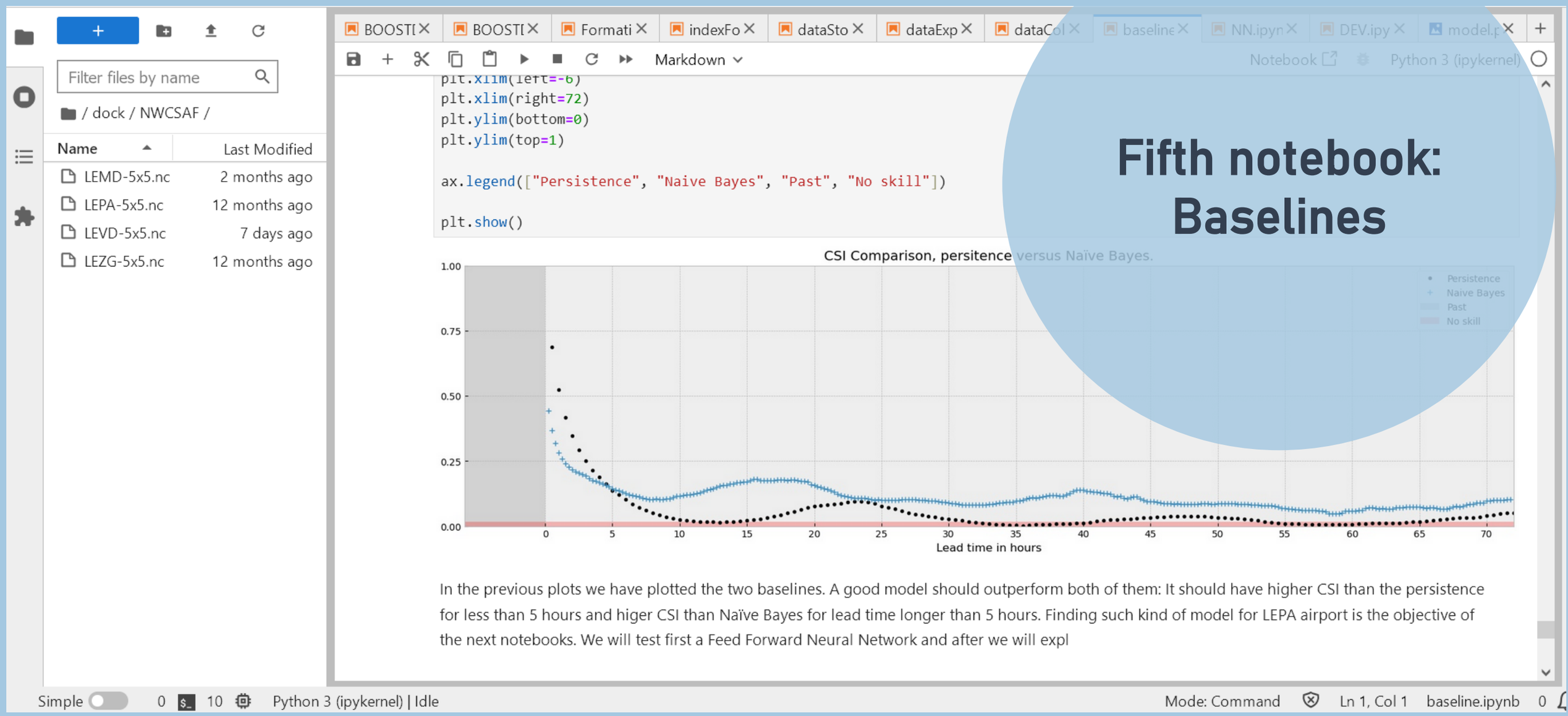
The fourth notebook teaches how to collocate the data for simultaneous exploitation. The satellite cloud data are a cube of small stamps with a very long time axis, on the other hand the METAR data are a long time series with different time steps. The notebook teaches how to select the central pixel of the stamps and how to deal with the METAR data. Two methods are provided: .resample() and .interpolate() for METAR data



Fourth notebook: Data Collocation

In the fifth notebook, how to build some baseline models, including Naïve Bayes, is explained. These models will be the reference for the actual models.

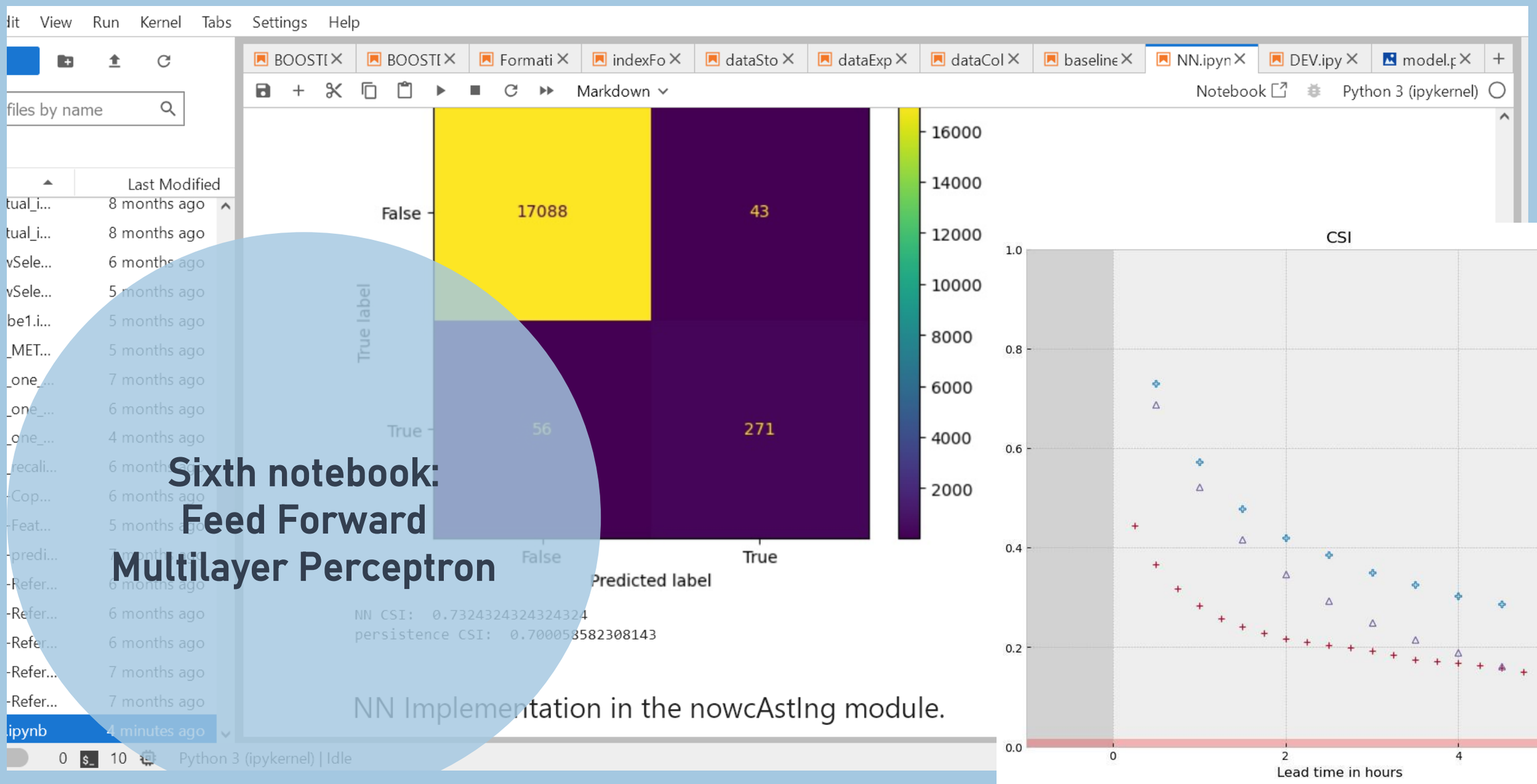
Through the toolbox the users models and their performance are stored on disk, you don't need to retrain the model each time you need to make a prediction.



Fifth notebook: Baselines

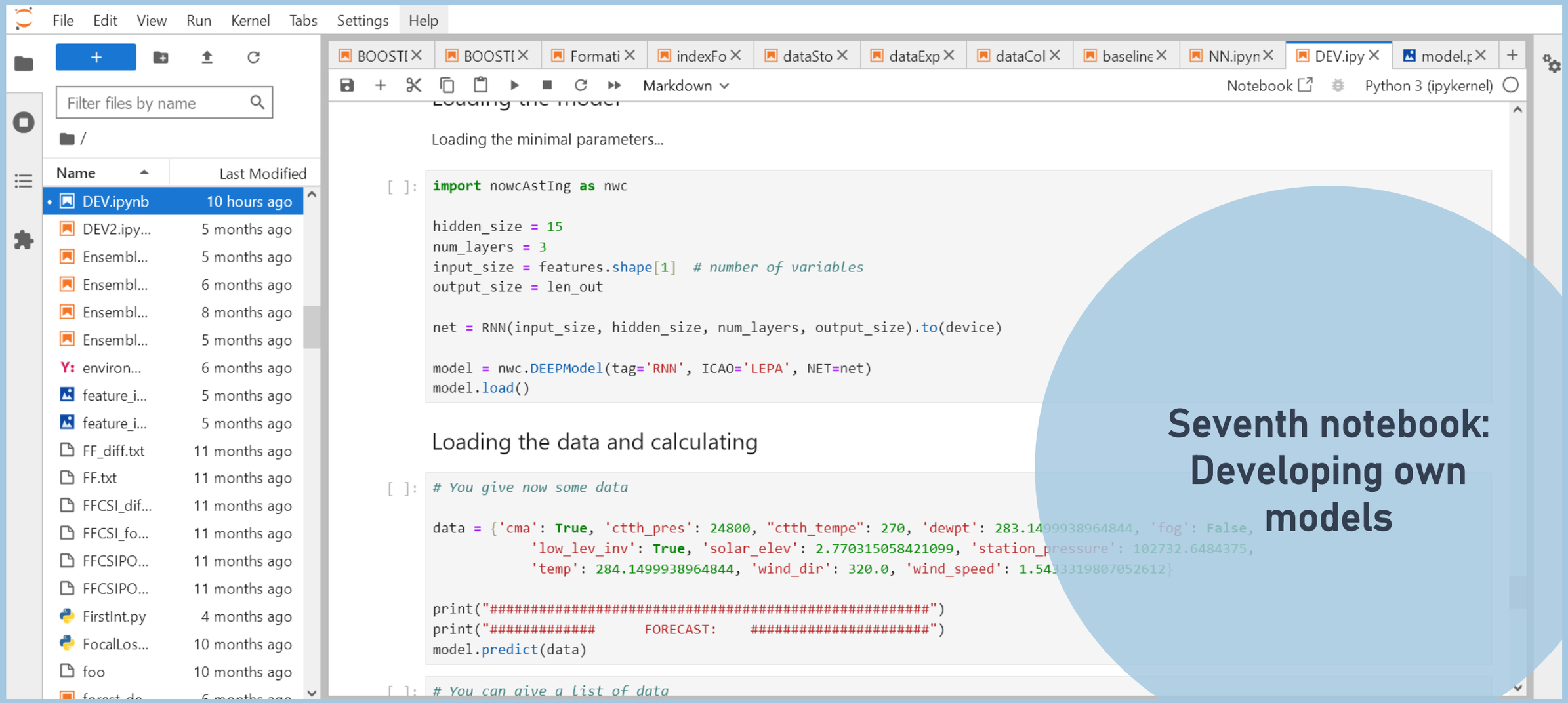
In the sixth notebook, Deep Learning basis are explained. We use a Feed Forward Multilayer perceptron (stored also as a class of the toolbox). Special attention is given to the loss function.

This model has proved to outperform the baselines for LEPA airport .



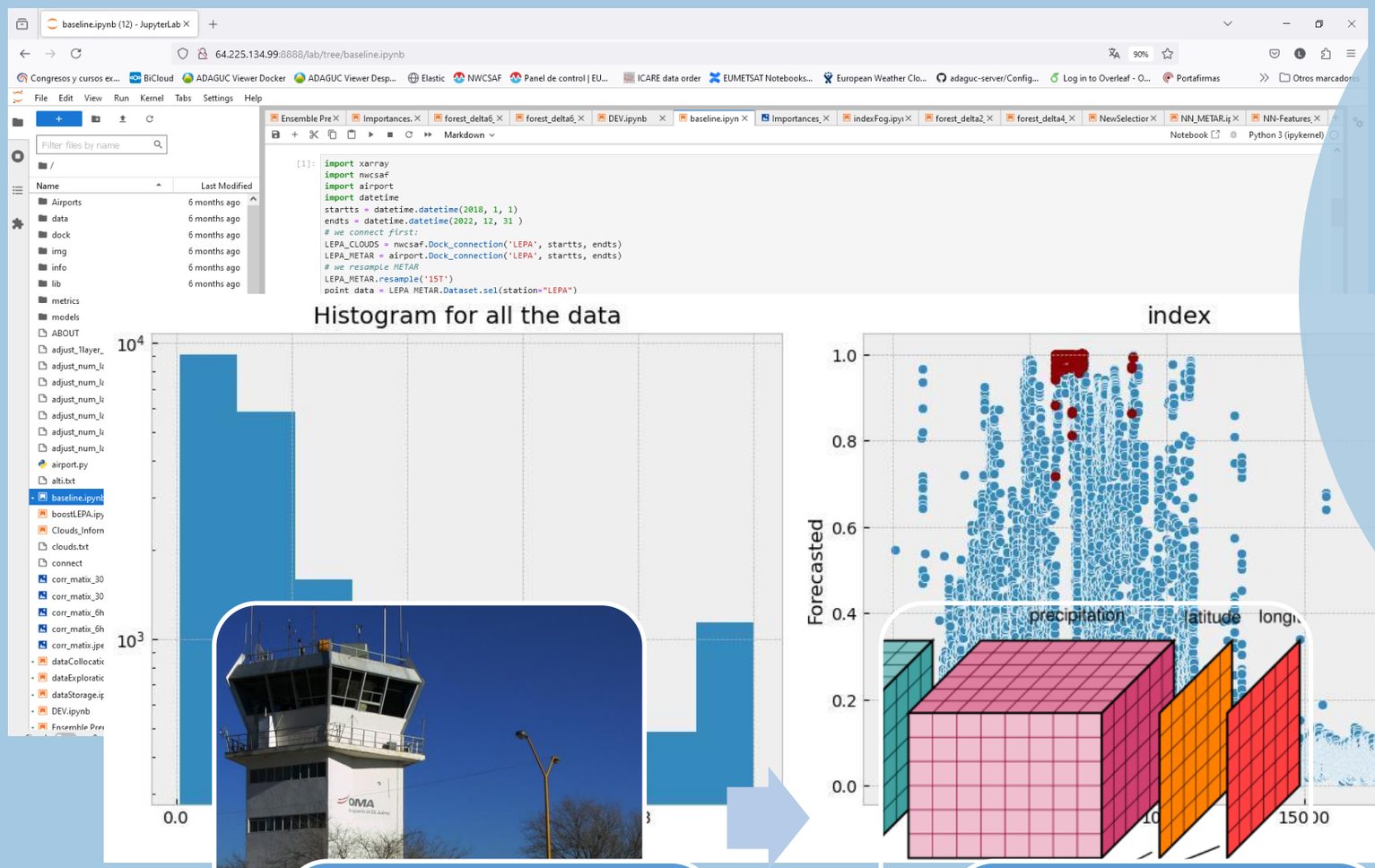
The seventh notebook provides a host class to test own models.
The name of the class is DEEPmodel, it takes as parameters a long list of "configurable" parameters: a tag, the ICAO code of the airport, the data, the model architecture, the list of predictors, the learning rate…

This allow to develop and intercompare different models. The code is prepared to run either on CPU or GPU.



Seventh notebook: Developing own models

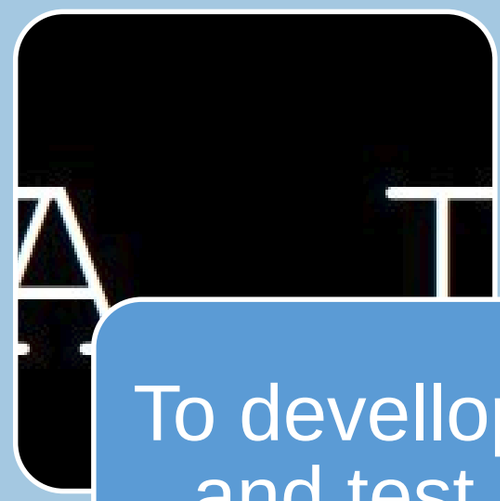
Summary



We provide the toolbox to the community through a jupyter notebook collection, allowing:

To choose a location

To download the data



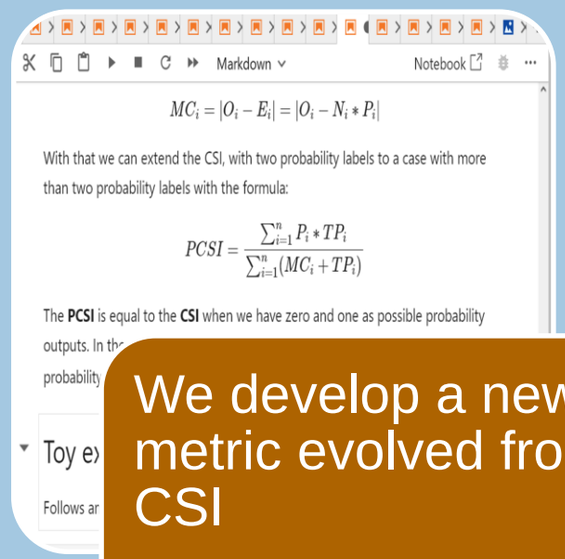
To devellop and test their own models

Dev>Ops

AEMET is in the process of testing in operations .

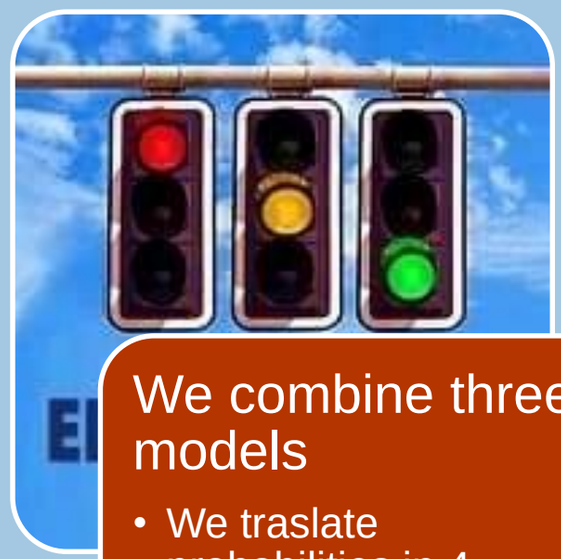
We talk with forecasters

- They forecast four categories
- **Not fog**
- **P30**
- **P40**
- **Fog**



We develop a new metric evolved from CSI

- PCSI is an extensión of CSI to multicategories
- The metric is insensitive to true negatives

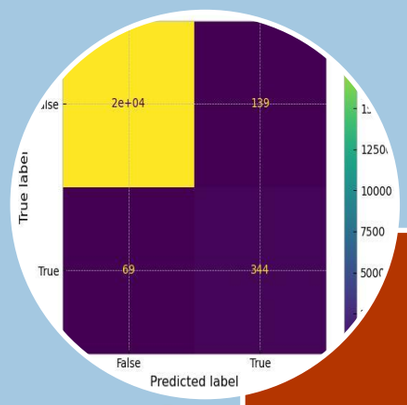


We combine three models

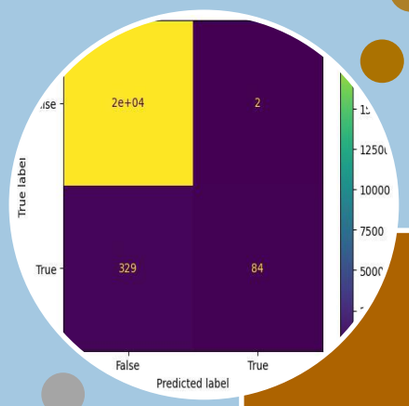
- We traslate probabilities in 4 warning levels
- Still not tested o the forecast rooms

Dev>Ops

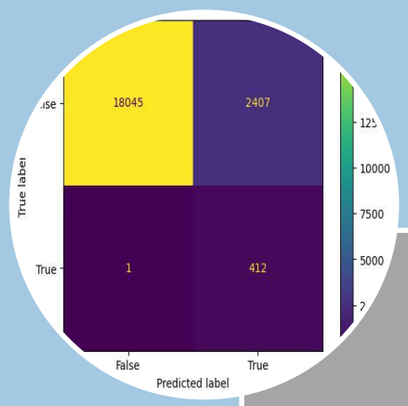
We combine three models with different behaviour and at the end we recalibrate the outputs



We train a model with neutral behaviour



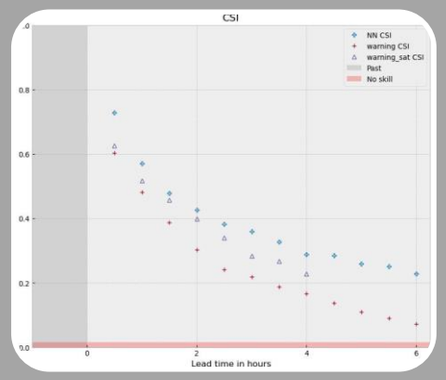
Second we train a very precise model it will give the red labels



First we train a very sensitive model, it will give the green labels

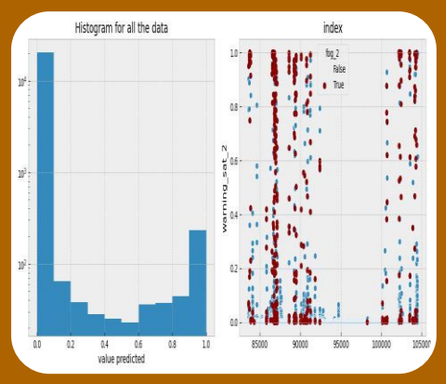
We are facing the problem with a boosting approach.

Dev>Ops



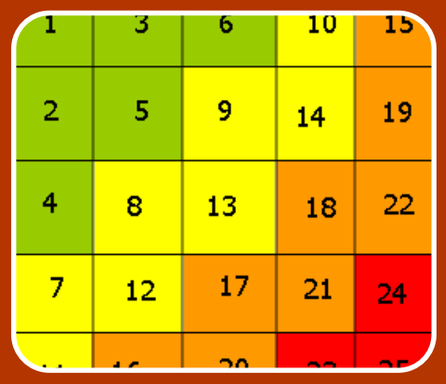
We are nowcating all time step at once

- The overall performance increases with radiances
- The model is more performant nowcasting time steps one by one



The model goes beyond persistence

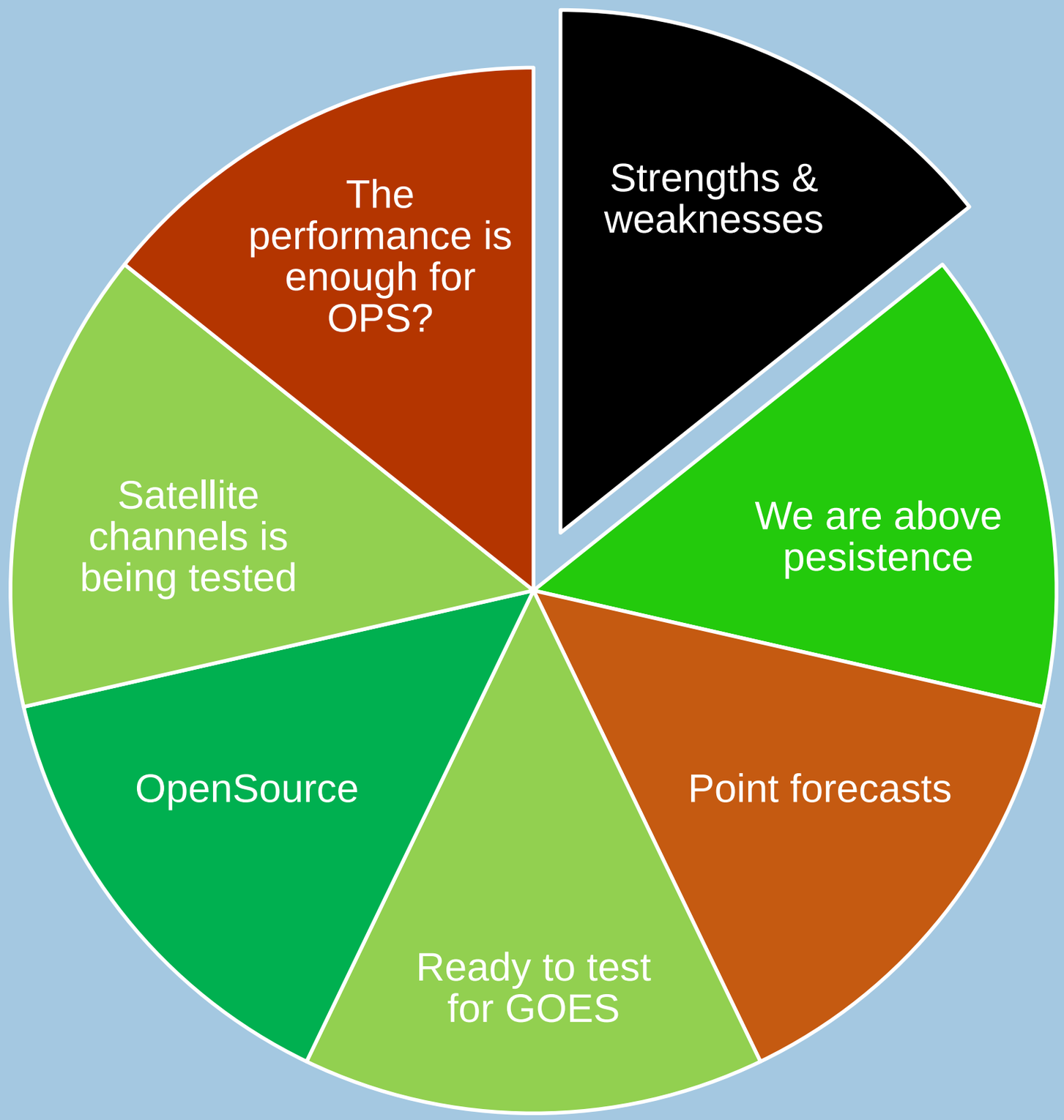
- The model is able to nowcast fog with clear vibility
- The model is able to nowcast notfog in foggy situations



Spatial and temporal info

- We are including the 4 previous time steps of the variables
- For spatial we are just including means and standard deviations of the spatial variables

Final valoration:



Thank you for the attention!

Foto Bernd 📷 Dittrich at Unsplash

jllisov@aemet.es

Lliso, Llorenç(1). Peinado-Galan, Niobe (1). Pellicer-Valero, Oscar J (2). Camps-Valls, Gustau (2). Lopez, Román (1) . Calbet, Xavier (1). Rípodas, Pilar (1) Gultepe, Ismail (3) (4)

(1) Agencia Estatal de Meteorología (AEMET/NWC SAF), Spain

(2) Image Processing Laboratory (IPL), Universitat de València, Spain

(3) Ontario Tech University, Oshawa, Canada

(4) University of Notre Dame, South Bend, Indiana, USA.

Special thanks to Carla Barroso from EUMETSAT

NowcastIng of fog using AI (NAI) techniques using satellite observations

NWC SAF home page: http://nwc-saf.eumetsat.int