ENVRI Community International Winter School on DATA FAIRness

https://www.lifewatch.eu/envri-iws-data-fairness-2020

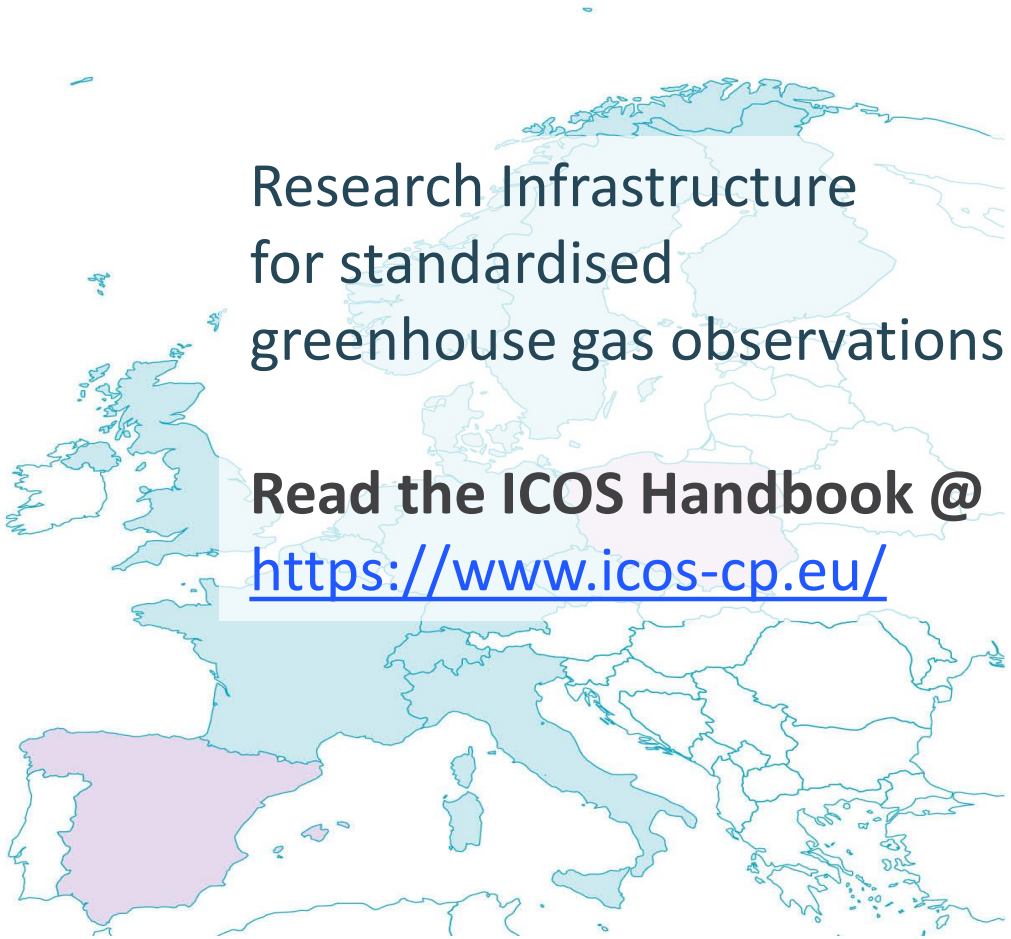
Elaborated Products @ ICOS

Ute Karstens, Karolina Pantazatou, Ida Storm, Claudio D'Onofrio, ICOS Carbon Portal

11-22 January 2021

ENVRI-FAIR has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824068

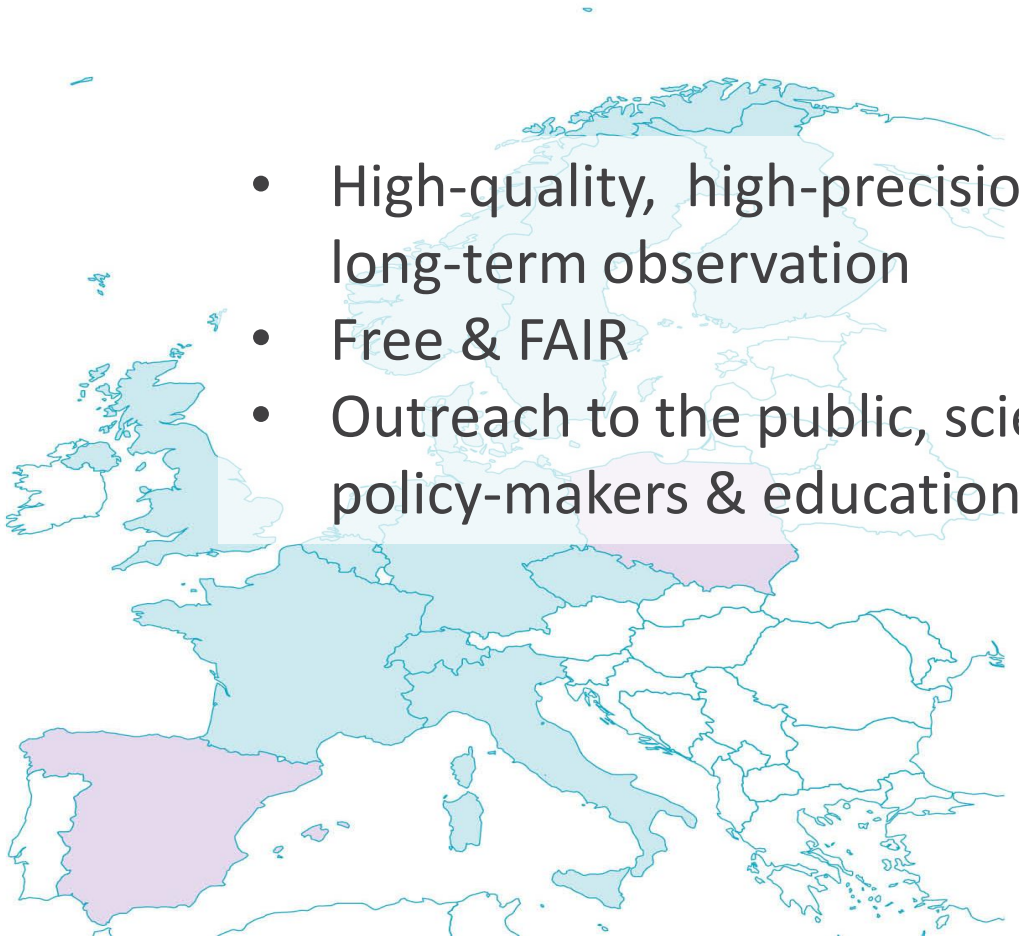
Integrated Carbon Observation System



Research Infrastructure
for standardised
greenhouse gas observations

Read the ICOS Handbook @
https://www.icos-cp.eu/

ICOS Mission



- High-quality, high-precision and long-term observation
- Free & FAIR
- Outreach to the public, science, policy-makers & education

Since this is virtual…..

Ute Karstens
ute.karstens@nateko.lu.se

Claudio D'Onofrio
claudio.donofrio@nateko.lu.se

Karolina Pantazatou
karolina.pantazatou@nateko.lu.se

Ida Storm
ida.storm@nateko.lu.se

ICOS | INTEGRATED CARBON OBSERVATION SYSTEM

GOAL FOR THIS WORKSHOP

Showcase the STILT web application @ ICOS CarbonPortal

Hands on (Jupyter Notebooks) compare STILT result with measurements
- Create a time series plot
- Create an interactive map for the station

Exercise:
Create a timeseries plot (model output vs. observation)
Create a Jupyter notebook with a map (show the station location)
- **Yes**, the same as we do in the workshop
- **No**, you can't use the 'workshop station', you MUST use a different station

Workshop Schedule

STILT On demand calculation Result visualization	Ute	15 min
Station characterization Create a new “data product” based on STILT results	Ida	5 min
Login to VRE	Karolina	5 min
Break		5 min
Hands on – create an interactive map	Claudio	30 min
Break		10 min
Hands on – create a timeseries plot	Karolina	45 min
Conclusion and exercise	Claudio	5 min

STILT Footprint Tool @ ICOS CarbonPortal

STILT is an atmospheric transport model that is used to compute **footprints** and greenhouse gas (GHG) **concentrations** at ICOS atmospheric stations

What is the "footprint" of an atmospheric station?

- *Source area that influences the GHG concentration at a station location*

Why do we need to model the footprint?

- *Analysis of potential impacts of natural and anthropogenic emissions on the atmospheric GHG concentration*

STILT Footprint Tool @ ICOS CarbonPortal

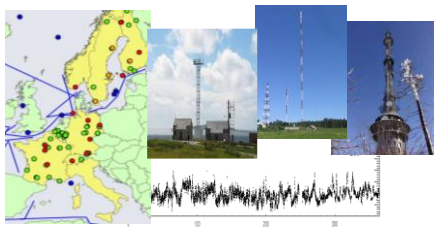
- Web-based service at ICOS Carbon Portal
- On-demand STILT model calculations
  - Model implemented in Virtual Machine to be scalable and transferable
  - Combination of computing resources at ICOS CP and in the cloud
  - Computationally intensive: ≈ 300 CPUs per single footprint
  - ICOS authentication required

  https://stilt.icos-cp.eu/worker/
- Open-access STILT viewer displaying all available computation results

  https://stilt.icos-cp.eu/viewer/

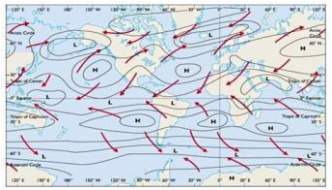
STILT Footprint Tool @ ICOS CarbonPortal

Atmospheric observations



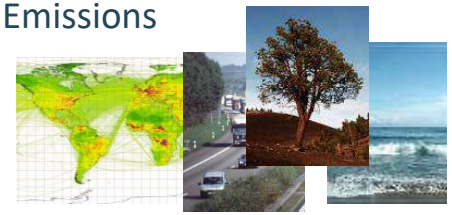
≈ 1 GB

Meteorological driver fields



≈ 2-3 TB

Emissions



≈ 0.5-1 TB

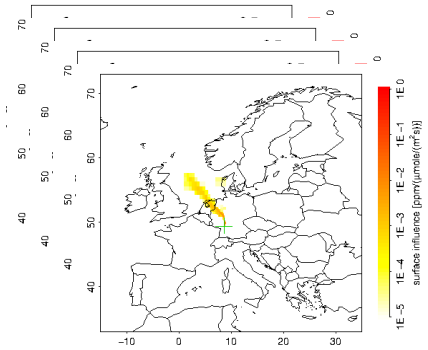
ICOS Carbon Portal

STILT
Atmospheric transport model

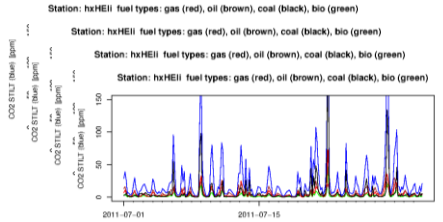
≈ 300 CPUs per footprint
=> 750 CPUh/station/year

Federated Cloud

Station Footprints



GHG concentrations



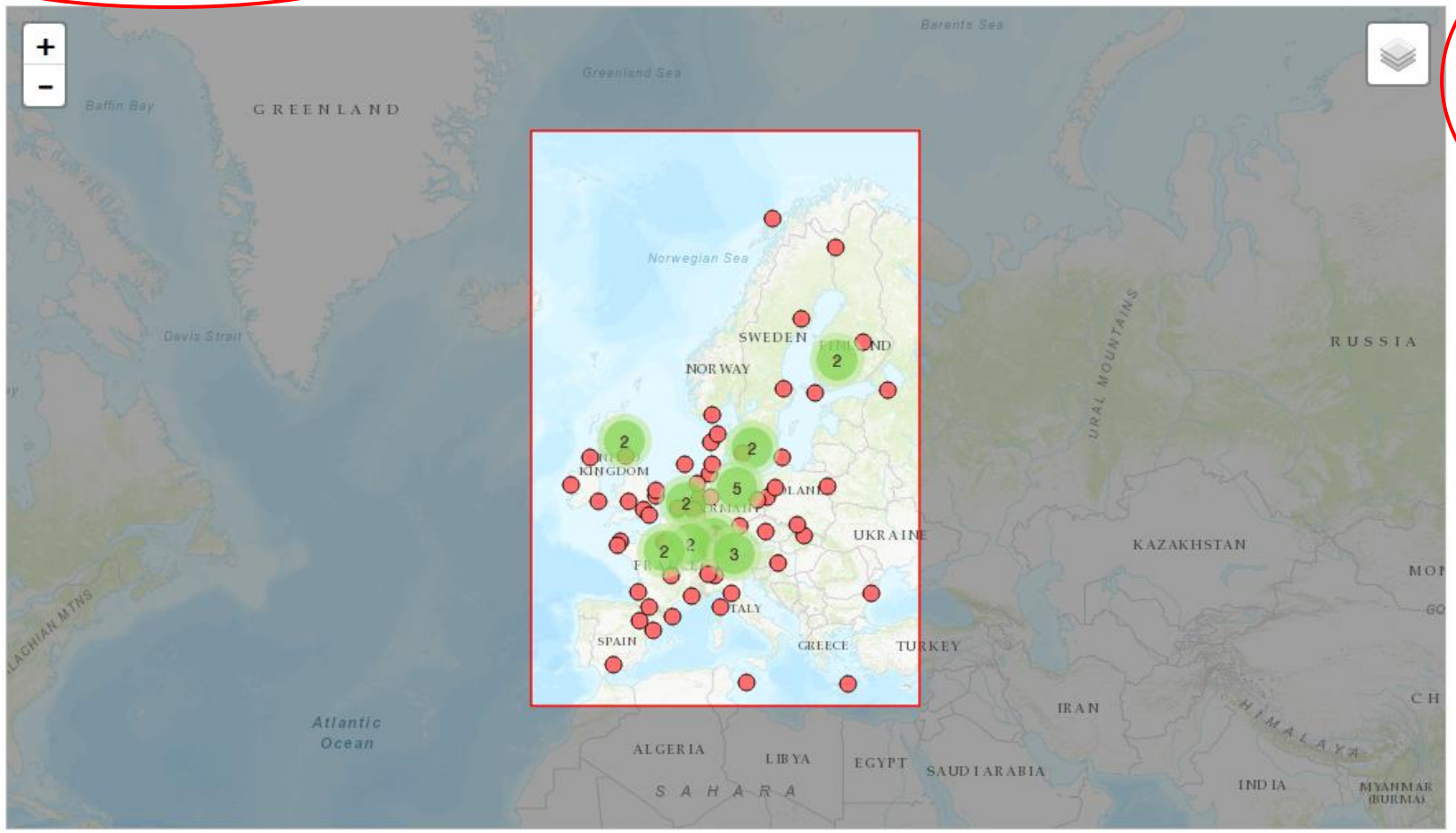
≈ 1-2 TB
per year

STILT calculation service Job starter

Logged in as ute.karstens@nateko.lu.se

Existing STILT footprints

Select station here or on the map



Create new STILT footprint

Latitude (decimal degree)

Longitude (decimal degree)

Altitude above ground (meters)

Site id (usually a 3 letter code)

Load data

Start date (YYYY-MM-DD)

End date (YYYY-MM-DD)

Submit STILT job

Submitted STILT jobs

Show details

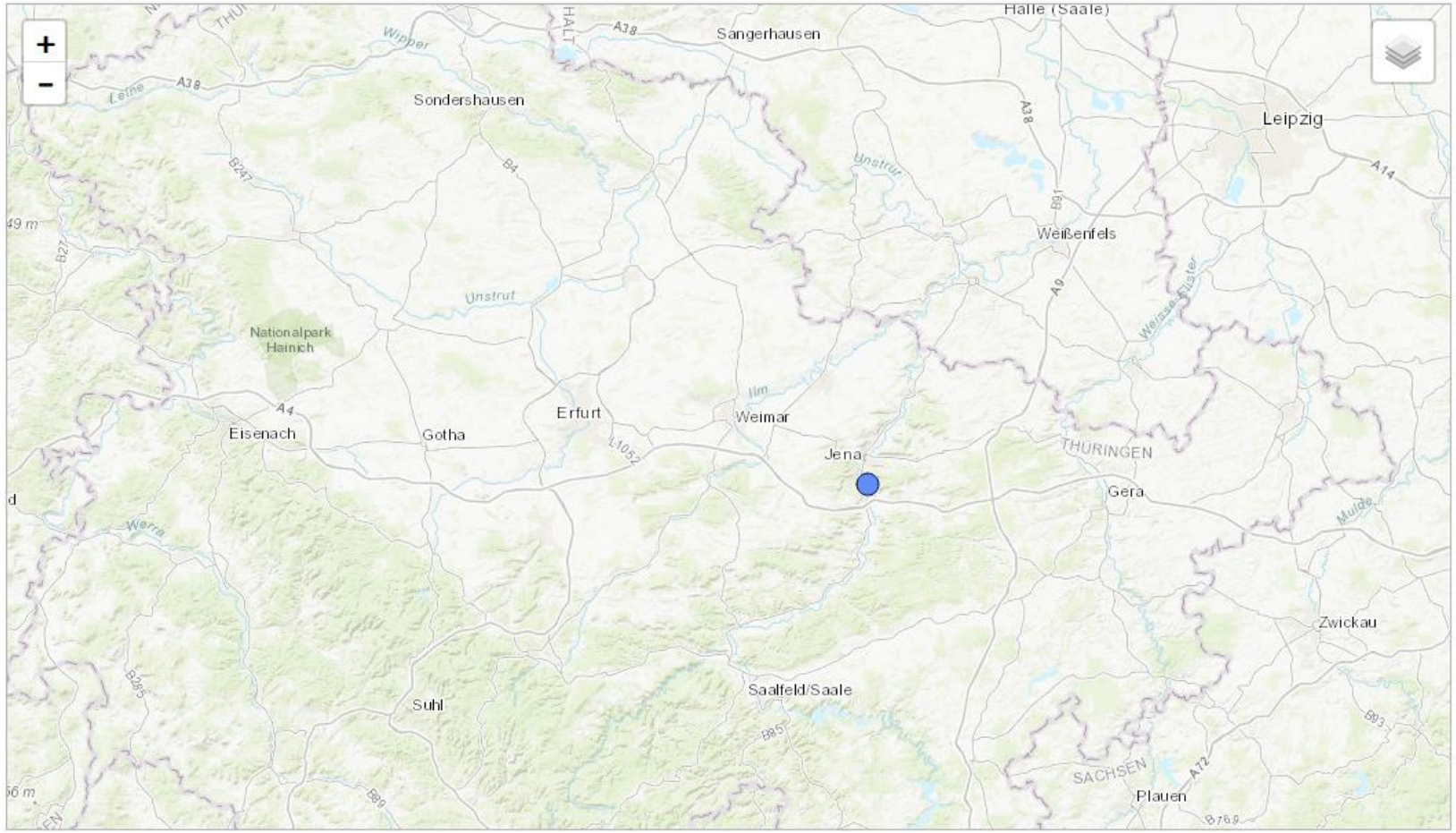
No jobs have been submitted

STILT calculation service Job starter

Logged in as ute.karstens@nateko.lu.se

Existing STILT footprints

Select station here or on the map



Create new STILT footprint

Latitude (decimal degree)

50.9

Longitude (decimal degree)

11.6

Altitude above ground (meters)

100

Site id (usually a 3 letter code)

JENA Load data

Start date (YYYY-MM-DD)

End date (YYYY-MM-DD)

Submit STILT job

Submitted STILT jobs

Show details

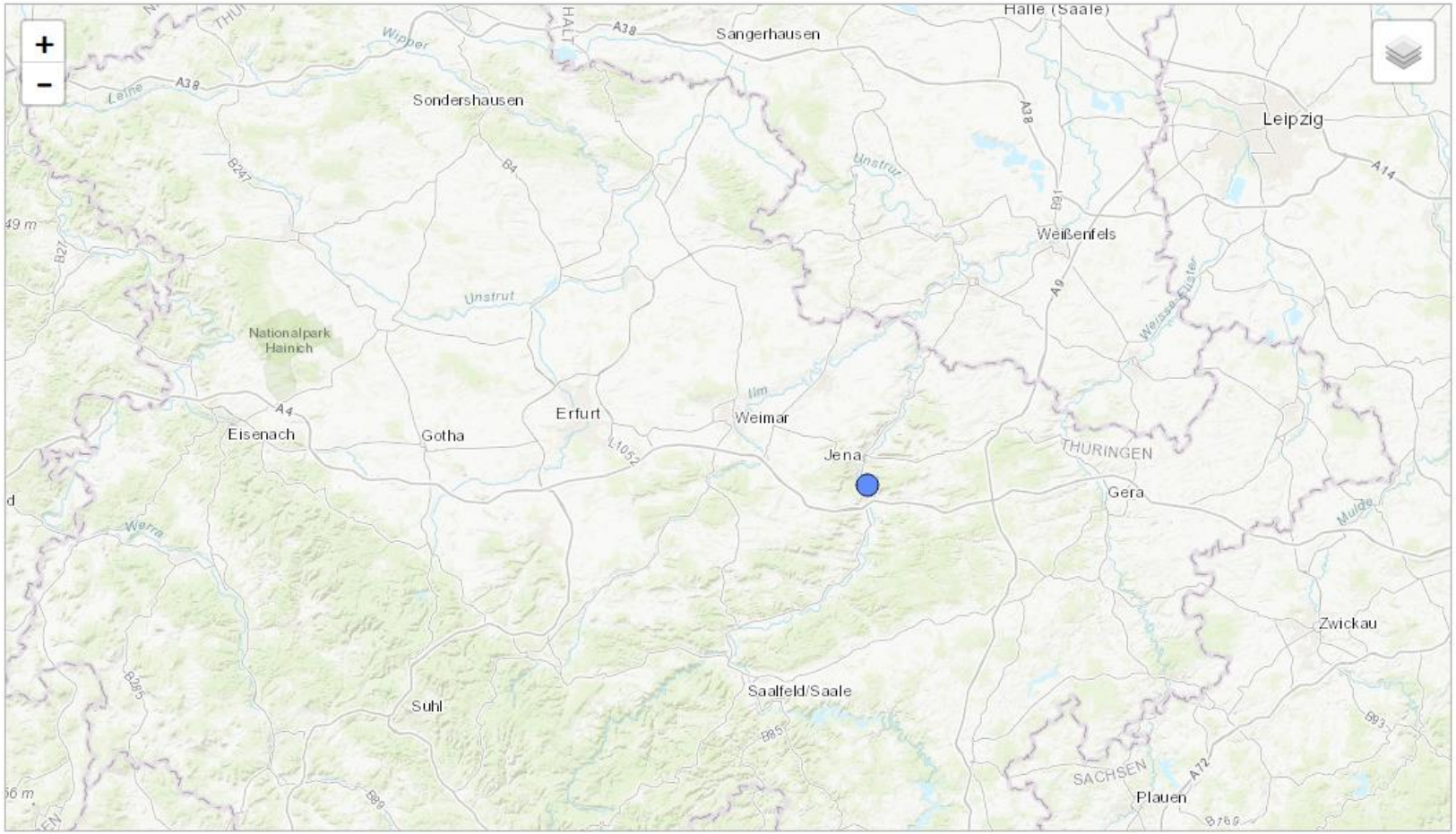
No jobs have been submitted

STILT calculation service Job starter

Logged in as ute.karstens@nateko.lu.se

Existing STILT footprints

Select station here or on the map



Create new STILT footprint

Latitude (decimal degree)

50.9

Longitude (decimal degree)

11.6

Altitude above ground (meters)

100

Site id (usually a 3 letter code)

JENA Load data

Start date (YYYY-MM-DD)

January 2006

Su	Mo	Tu	We	Th	Fr	Sa
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	1	2	3	4
5	6	7	8	9	10	11

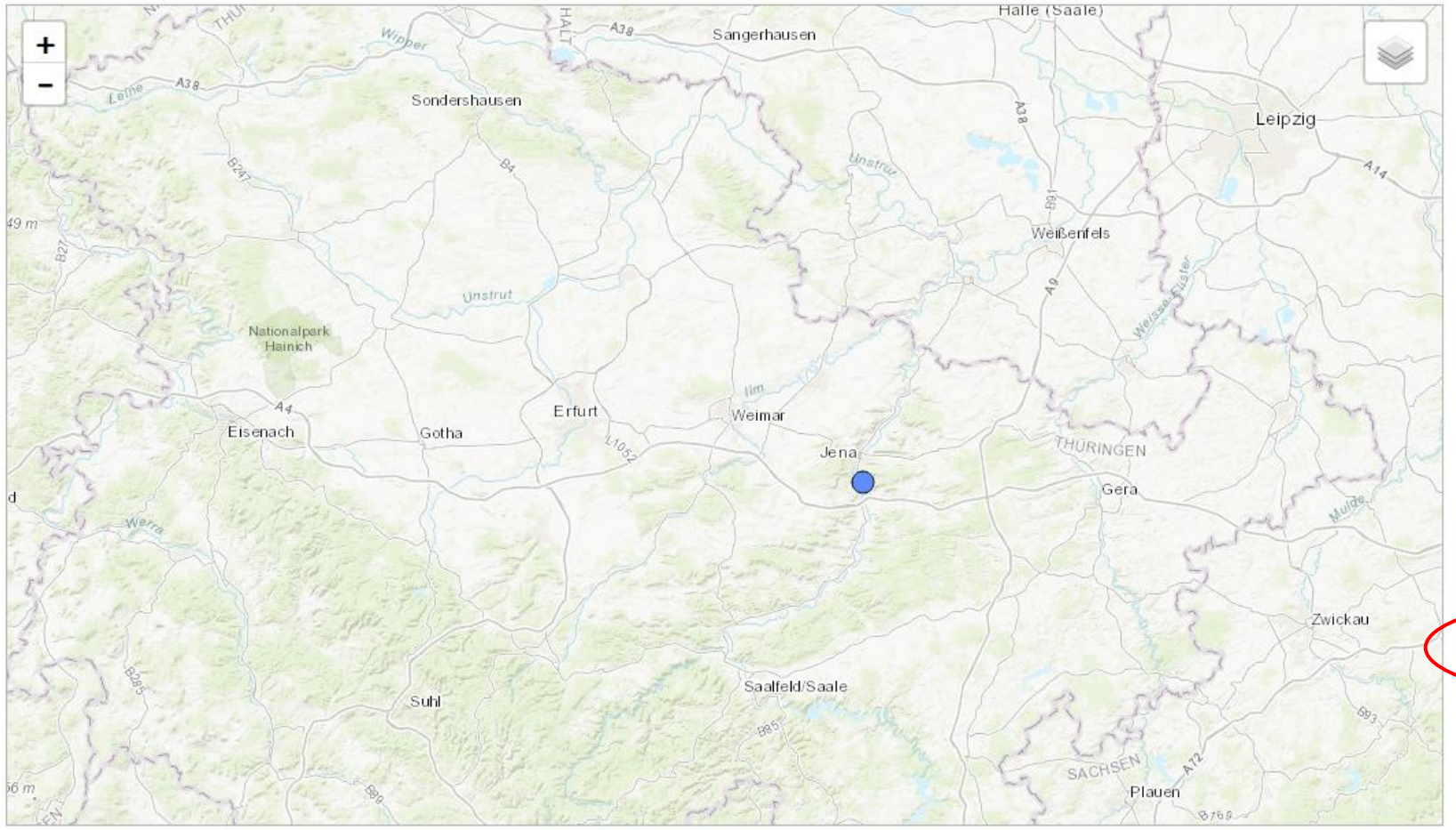
Show details

No jobs have been submitted

STILT calculation service Job starter

Existing STILT footprints

Select station here or on the map



Create new STILT footprint

Latitude (decimal degree)

50.9

Longitude (decimal degree)

11.6

Altitude above ground (meters)

100

Site id (usually a 3 letter code)

JENA Load data

Start date (YYYY-MM-DD)

2016-06-01

End date (YYYY-MM-DD)

2016-06-07

Submit STILT job

Submitted STILT jobs

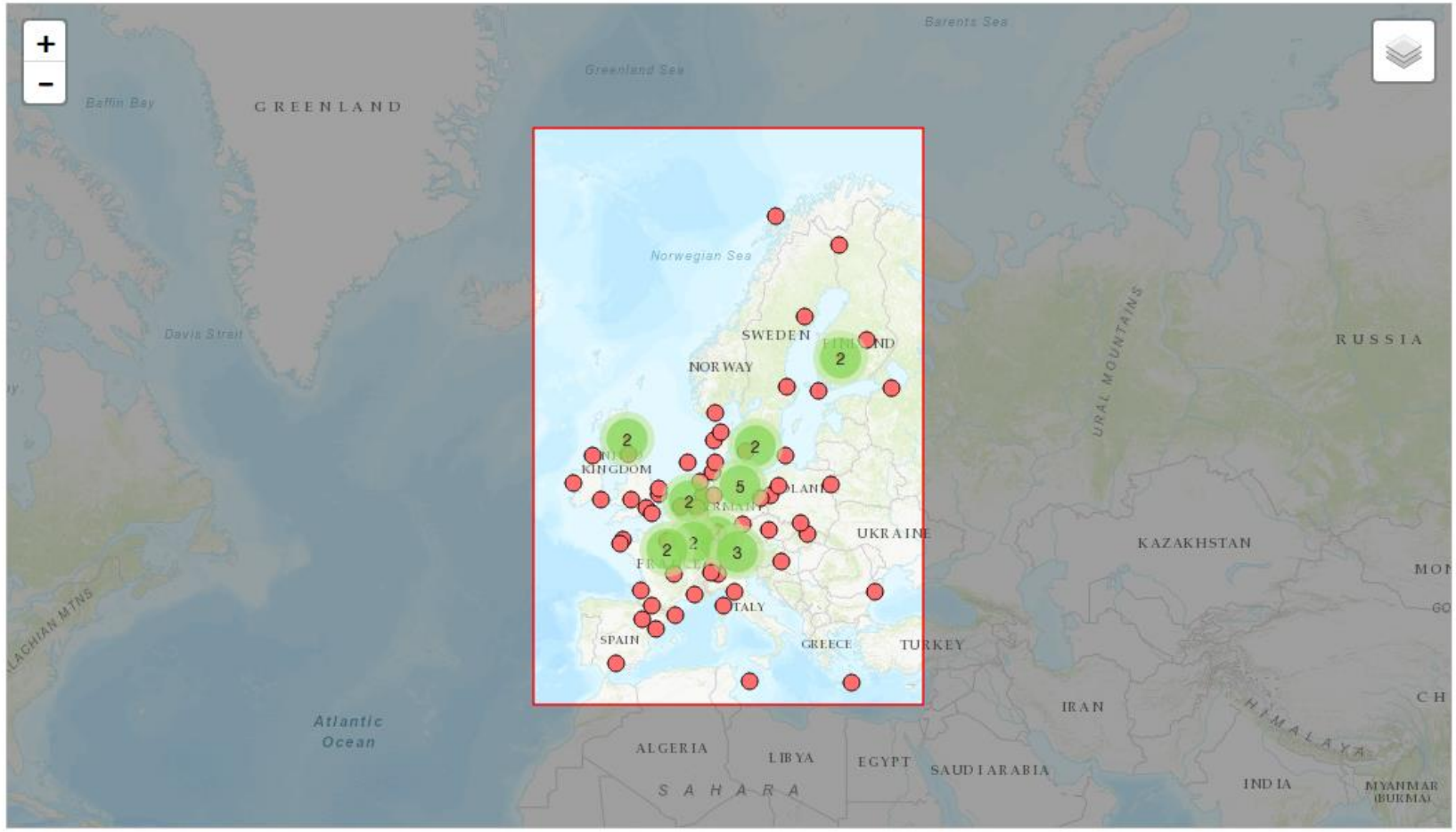
Show details

No jobs have been submitted

STILT calculation service Job starter

Existing STILT footprints

Select station here or on the map



Create new STILT footprint

Latitude (decimal degree)

Longitude (decimal degree)

Altitude above ground (meters)

Site id (usually a 3 letter code)

Load data

Start date (YYYY-MM-DD)

End date (YYYY-MM-DD)

Submit STILT job

Submitted STILT jobs

Show details

Running computations

★ Site 'JENA'

STILT calculation service Dashboard

Logged in as ute.karstens@nateko.lu.se

Computational resources

Node	Free CPUs	Total CPUs
akka.tcp://StiltCluster@localhost:2561	0	40

Running computations

Cancel job **Site id:** ***JENA*** (**lat:** 50.9, **lon:** 11.6), **alt:** 100, **start:** 2016-06-01, **stop:** 2016-06-07, **done:** 1 of 56 - submitted by ute.karstens@nateko.lu.se

To the job starter

STILT calculation service Dashboard

Logged in as ute.karstens@nateko.lu.se

Computational resources

Node	Free CPUs	Total CPUs
akka.tcp://StiltCluster@localhost:2561	31	40

Running computations

Cancel job **Site id:** ***JENA*** (**lat:** 50.9, **lon:** 11.6), **alt:** 100, **start:** 2016-06-01, **stop:** 2016-06-07, **done:** 47 of 56 - submitted by ute.karstens@nateko.lu.se

To the job starter

STILT calculation service Dashboard

Logged in as ute.karstens@nateko.lu.se

Computational resources

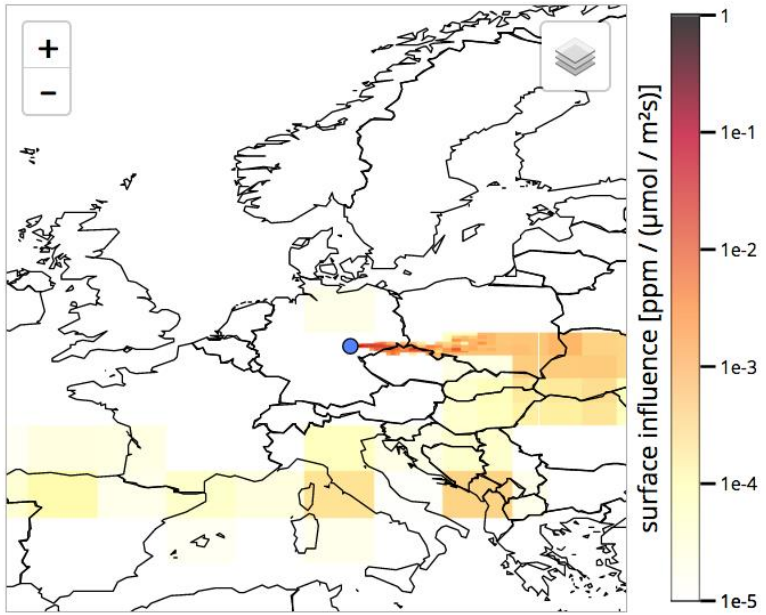
Node	Free CPUs	Total CPUs
akka.tcp://StiltCluster@localhost:2561	40	40

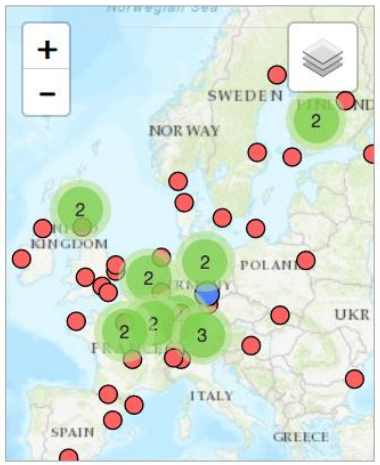
Finished computations

View results **Site id:** ***JENA*** **(lat:** 50.9, **lon:** 11.6), **alt:** 100, **start:** 2016-06-01, **stop:** 2016-06-06, **done:** 48 of 48 - submitted by ute.karstens@nateko.lu.se

To the job starter

STILT single-site scoped viewer





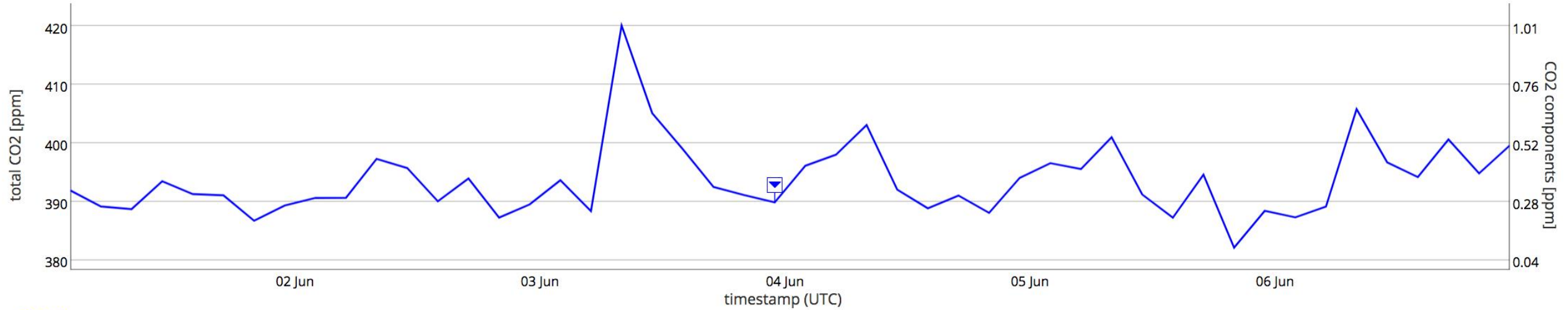
Pre-selected site: JENA **Start date:** 2016-06-01 **End date:** 2016-06-06

Footprint: 2016-06-03 21:00

Primary Y-axis: ☑ co2.stilt ☐ co2.background

Secondary Y-axis: ☐ co2.bio ☐ co2.bio.gee ☐ co2.bio.resp ☐ co2.fuel ☐ co2.fuel.oil ☐ co2.fuel.coal ☐ co2.fuel.gas ☐ co2.fuel.bio ☐ co2.energy ☐ co2.transport ☐ co2.industry ☐ co2.others

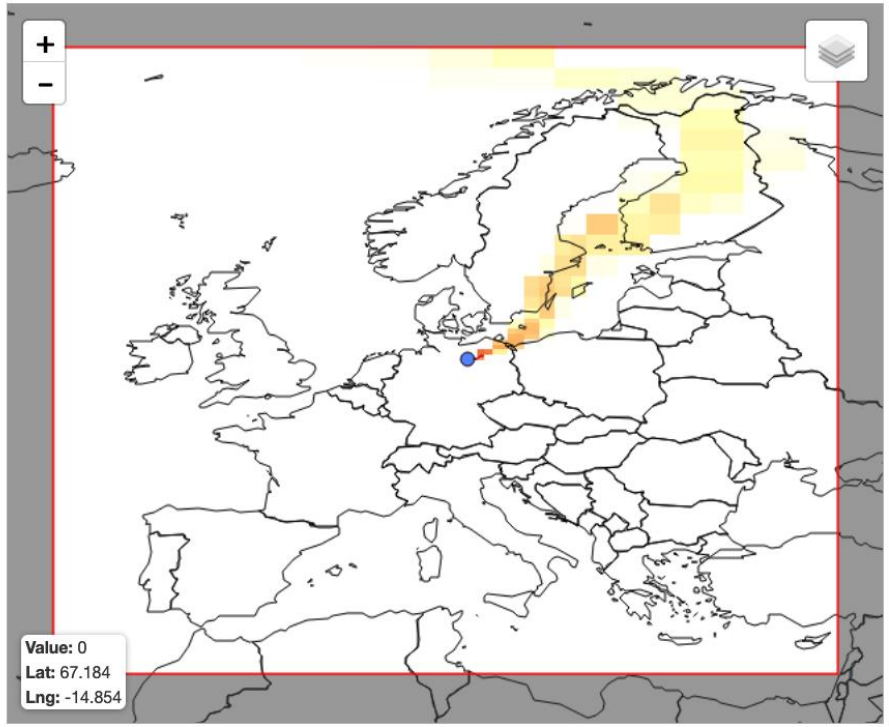
Playback ◀ ▶ ▶ **Playback speed** Fast (up to 10 fps)

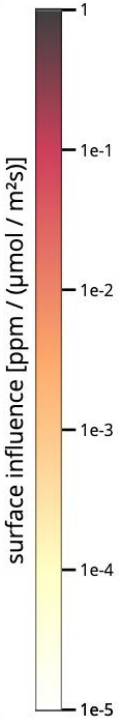


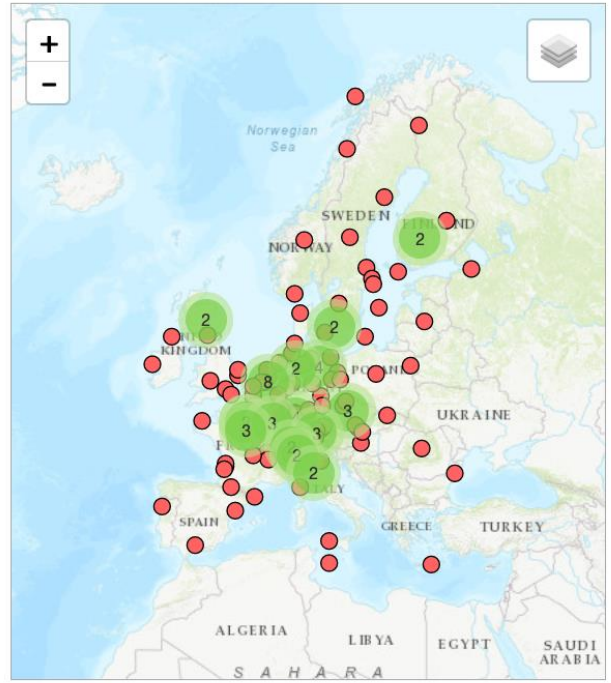
— co2.stilt

STILT results viewer

Help







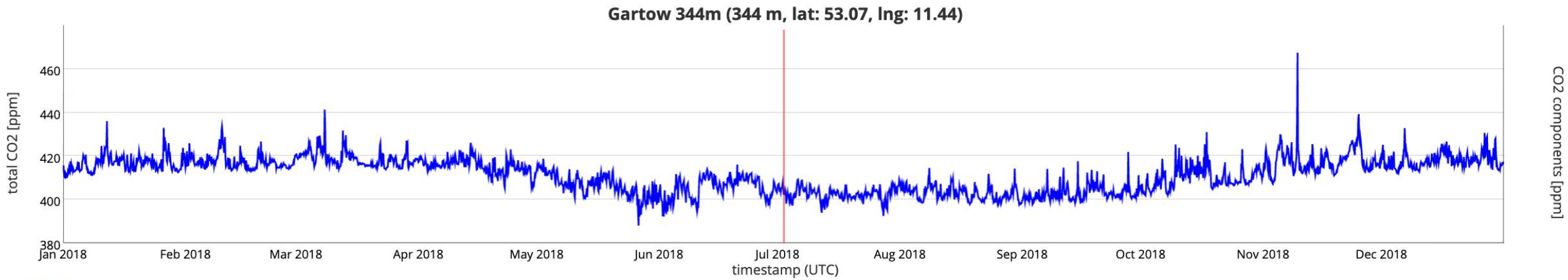
GAT344 (Gartow 344m, 344 m) | 2018

Footprint: 2018-07-02 09:00

Primary Y-axis:
☐ co2.observed ☑ co2.stilt ☐ co2.background

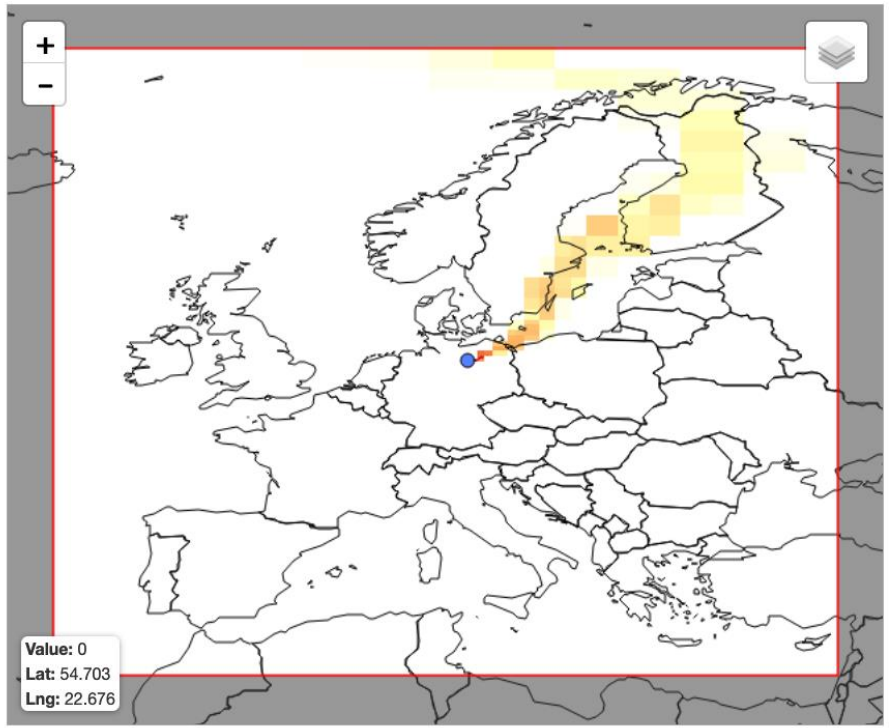
Secondary Y-axis:
Biosperic CO2
☐ co2.bio
Biosperic CO2 split into Photosynthetic uptake and respiration
☐ co2.bio.gee ☐ co2.bio.resp
Anthropogenic CO2
☐ co2.fuel
Anthropogenic CO2 split into Fuel types
☐ co2.fuel.coal ☐ co2.fuel.oil ☐ co2.fuel.gas ☐ co2.fuel.bio
Anthropogenic CO2 split into Source categories
☐ co2.energy ☐ co2.transport ☐ co2.industry ☐ co2.others

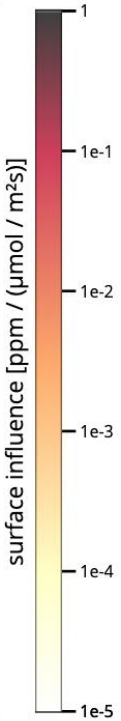
Playback ◀ ▶ ▶ **Playback speed** Fast (up to 10 fps)

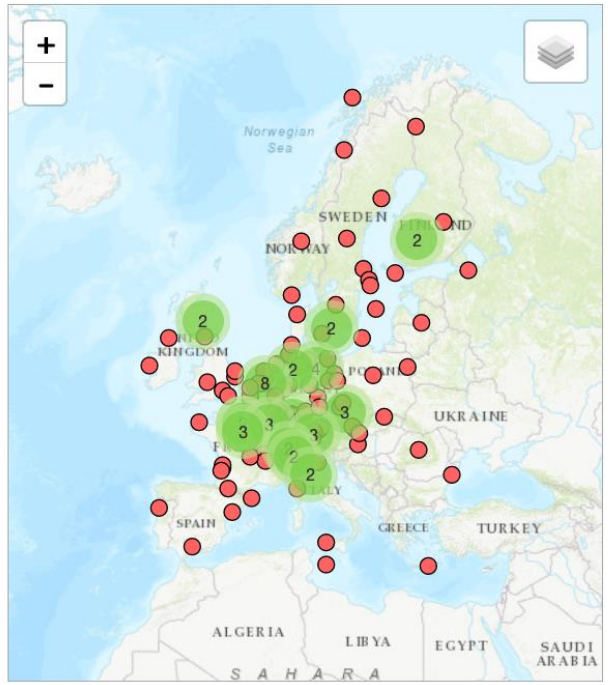


STILT results viewer

Help







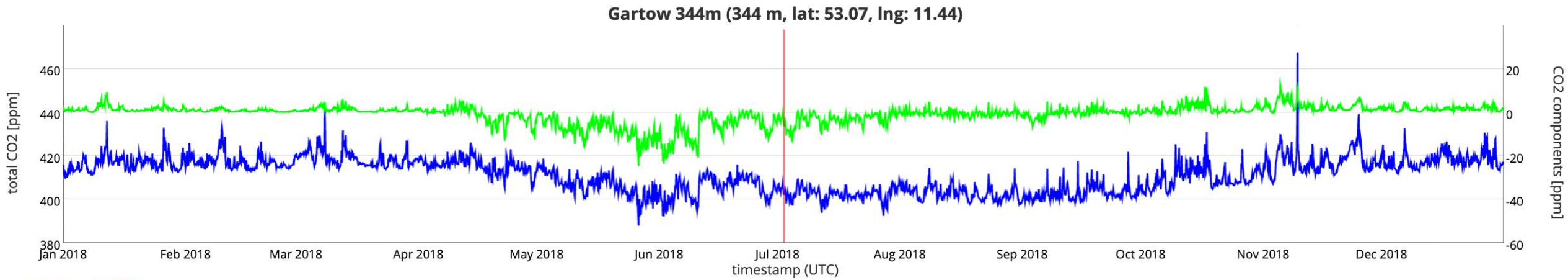
GAT344 (Gartow 344m, 344 m) | 2018

Footprint: 2018-07-02 09:00

Primary Y-axis:
☐ co2.observed ☑ co2.stilt ☐ co2.background

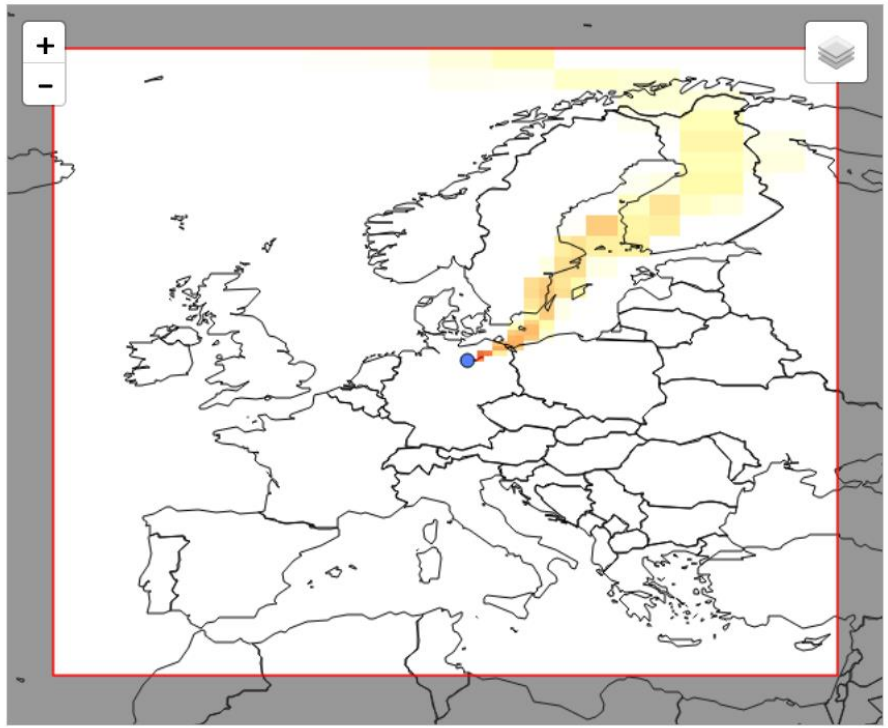
Secondary Y-axis:
Biosperic CO2
☑ co2.bio
Biosperic CO2 split into Photosynthetic uptake and respiration
☐ co2.bio.gee ☐ co2.bio.resp
Anthropogenic CO2
☐ co2.fuel
Anthropogenic CO2 split into Fuel types
☐ co2.fuel.coal ☐ co2.fuel.oil ☐ co2.fuel.gas ☐ co2.fuel.bio
Anthropogenic CO2 split into Source categories
☐ co2.energy ☐ co2.transport ☐ co2.industry ☐ co2.others

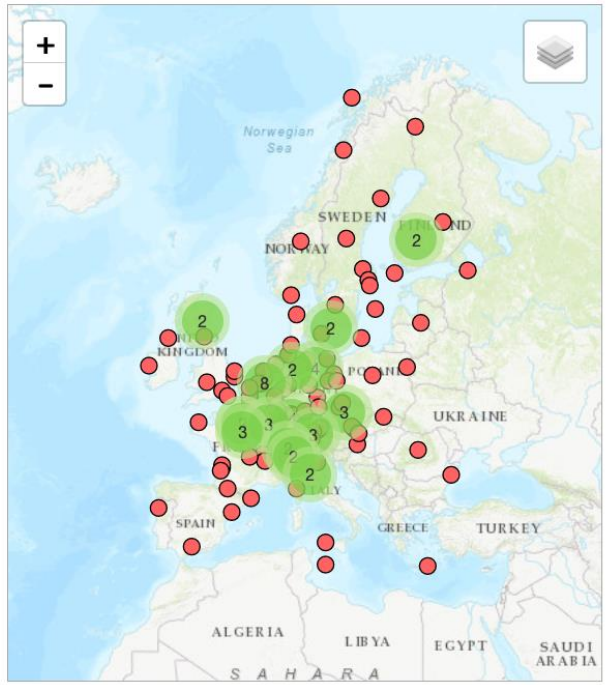
Playback ◀ ▶ ▶ **Playback speed** Fast (up to 10 fps)



STILT results viewer

Help





GAT344 (Gartow 344m, 344 m) | 2018

Footprint: 2018-07-02 09:00

Primary Y-axis:
- [ ] co2.observed
- [x] co2.stilt
- [ ] co2.background

Secondary Y-axis:

Biosperic CO2
- [x] co2.bio

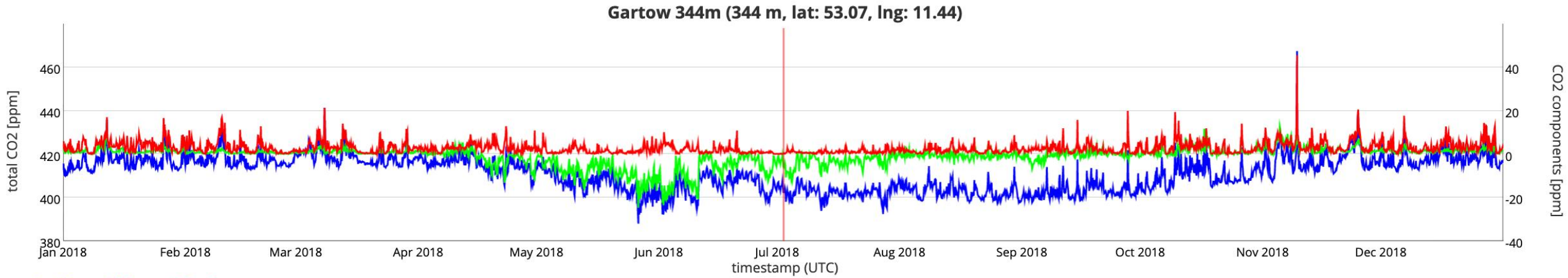
Biosperic CO2 split into Photosynthetic uptake and respiration
- [ ] co2.bio.gee
- [ ] co2.bio.resp

Anthropogenic CO2
- [x] co2.fuel

Anthropogenic CO2 split into Fuel types
- [ ] co2.fuel.coal
- [ ] co2.fuel.oil
- [ ] co2.fuel.gas
- [ ] co2.fuel.bio

Anthropogenic CO2 split into Source categories
- [ ] co2.energy
- [ ] co2.transport
- [ ] co2.industry
- [ ] co2.others

Playback ◀ ▶ ▶ **Playback speed** Fast (up to 10 fps)



STILT Footprint Tool @ ICOS CarbonPortal

All footprints and concentration results from the on-demand calculations are

- instantly available at ICOS CP
- directly accessible in Jupyter notebooks for further analysis, e.g.
  - ➢ combination with additional datasets
  - ➢ comparison with ICOS observations

ENVRI-FAIR has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824068

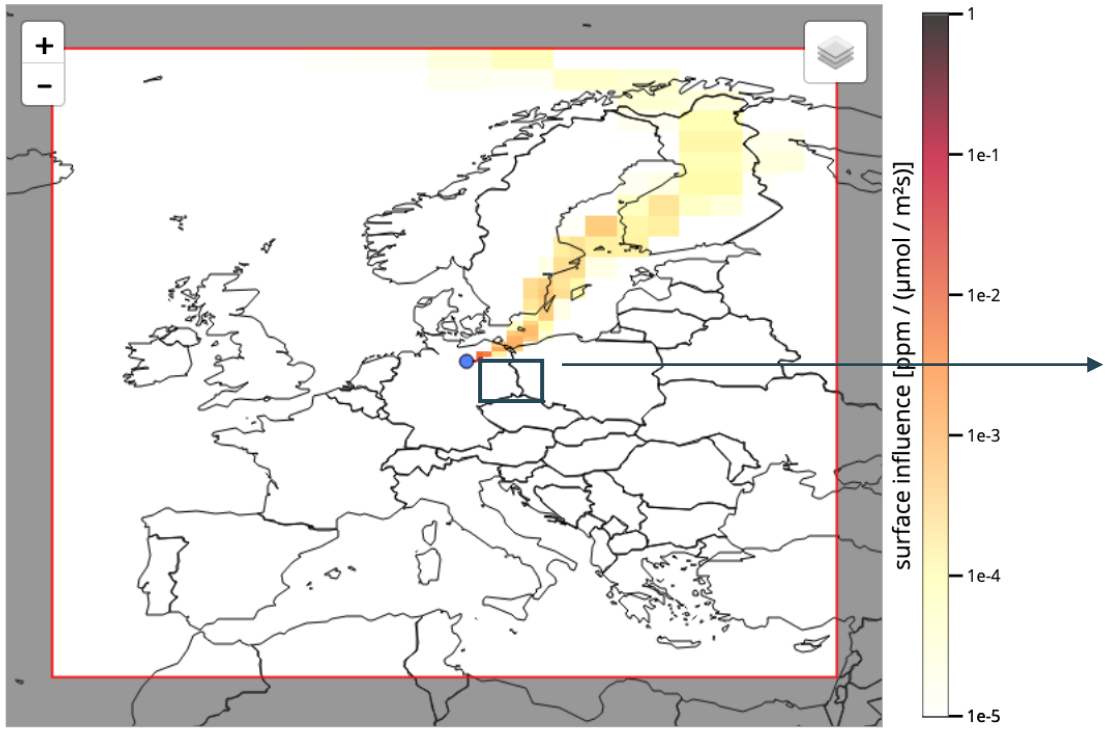
Workshop Schedule

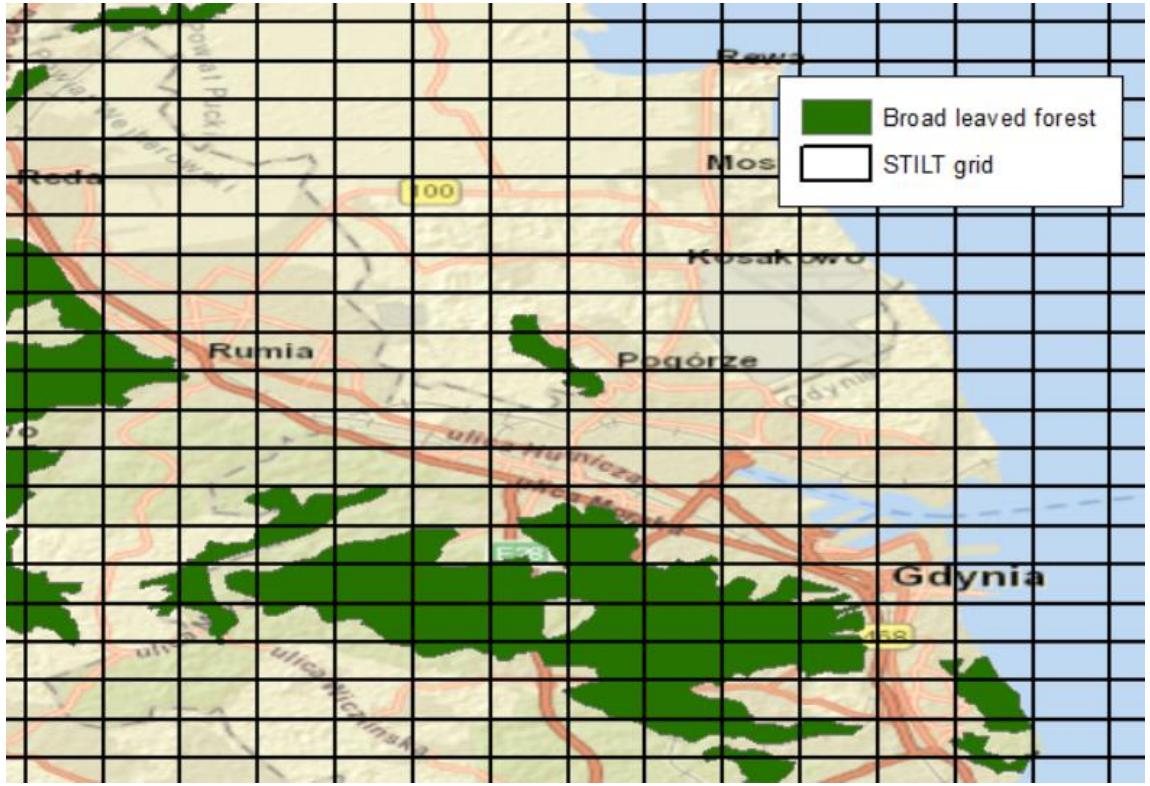
STILT - On demand calculation - Result visualization	Ute	15 min
Station characterization - Create a new “data product” based on STILT results	Ida	5 min
Login to VRE	Karolina	5 min
Break		5 min
Hands on – create an interactive map	Claudio	30 min
Break		10 min
Hands on – create a timeseries plot	Karolina	45 min
Conclusion and exercise	Claudio	5 min

Station Characterization

Characteristics of Gartow's footprint area from combination with additional datasets such as land cover (e.g. broad leaved forests):

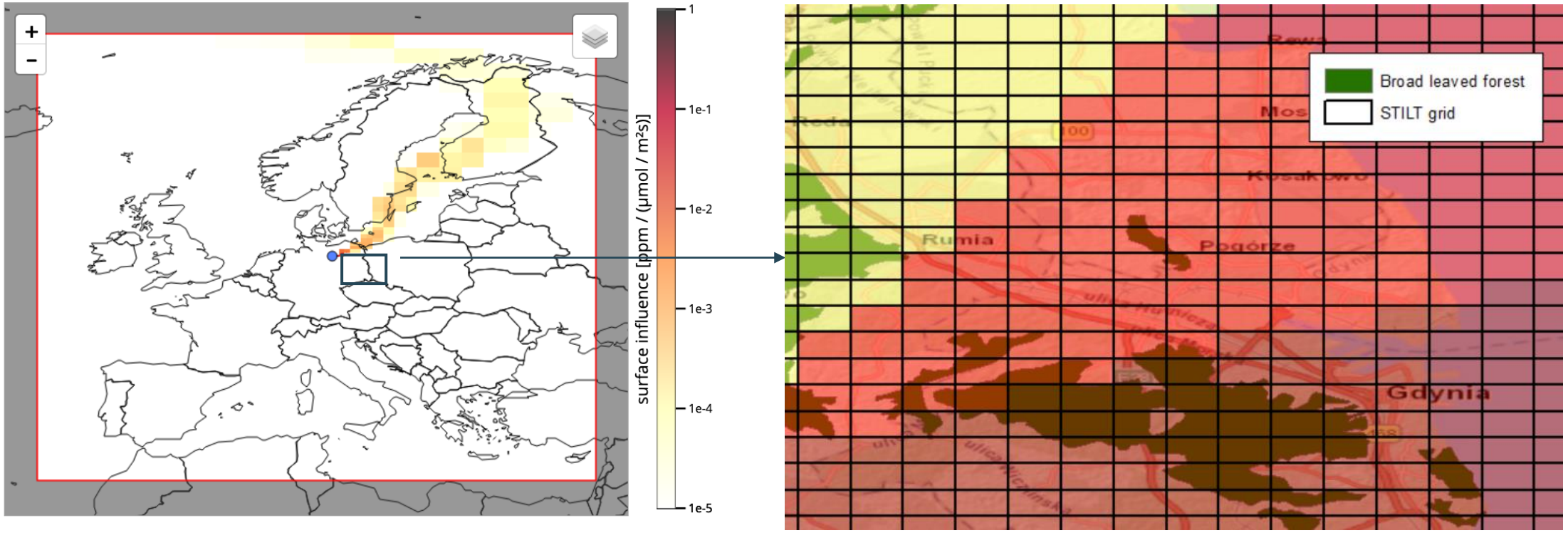
What is "seen" by the station?





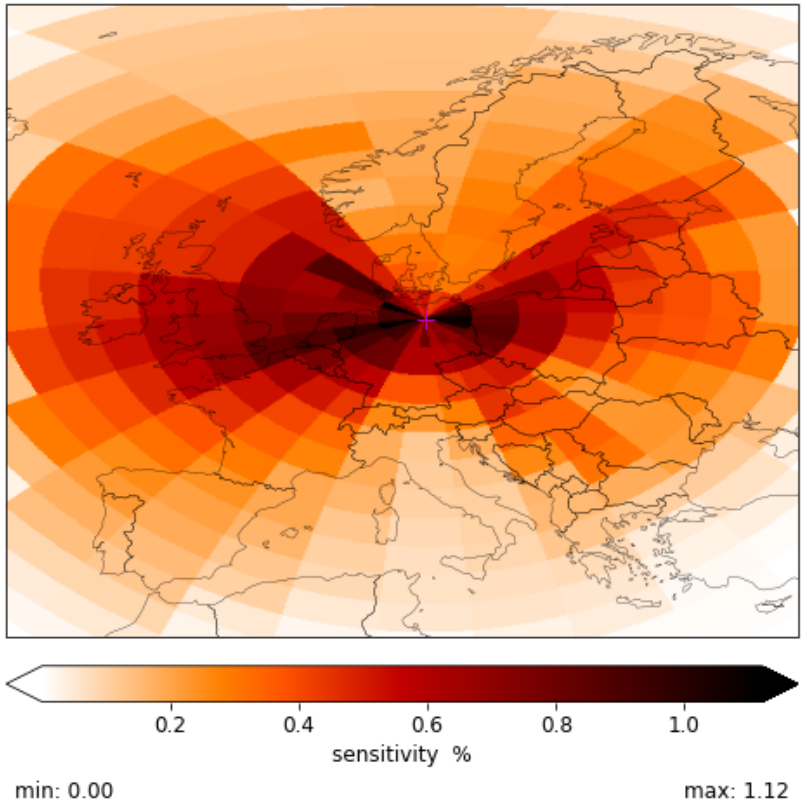
Station Characterization

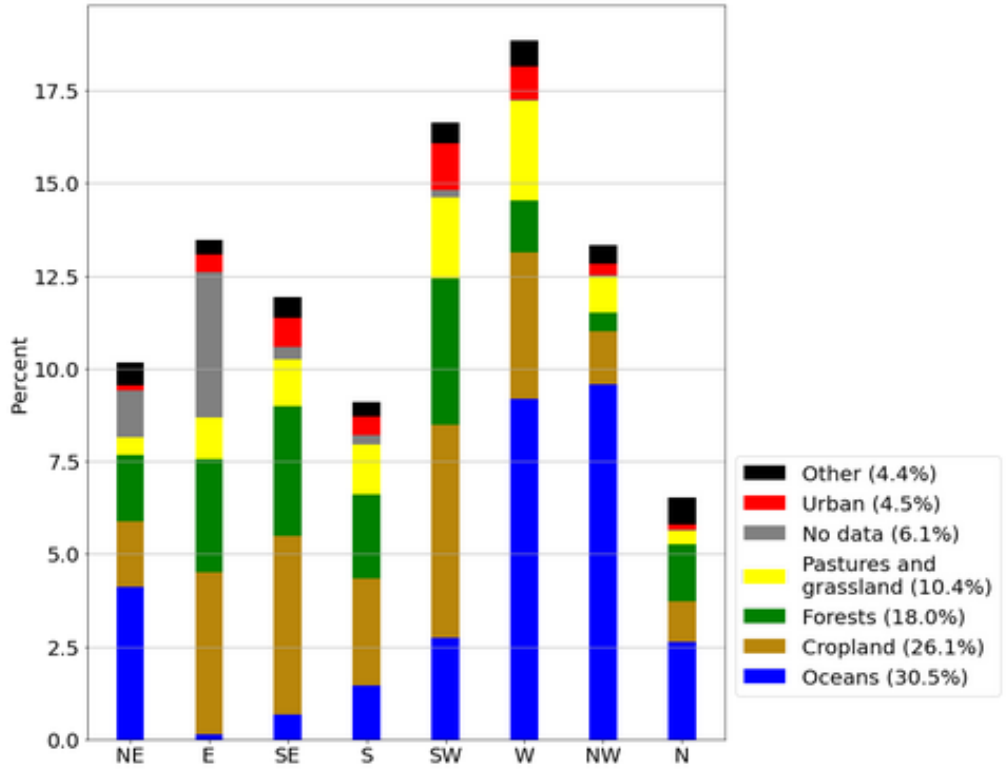
The color of each footprint ”cell” indicate how ”seen” it is by the station. Its value is multiplied by underlying data (e.g. area of broad leaved forest within the cell).



Station Characterization

Average footprint calculated from all 2018 footprints (every three hours)
Average footprint multiplied by all underlying different types of land cover.





Station Characterization Tool

It is possible to combine other data layers with the footprints. Visit exploredata.icos-cp.eu to see more examples of this and to characterize different stations!

- Footprints directly accessed in a user-friendly notebook
- Output in accordance with user-specified parameters
- Option to generate a PDF with the results

How to get there:

- https://exploredata.icos-cp.eu/user/jupyter/notebooks/icos_jupyter_notebooks/station_characterization/station_characterization_one_cell.ipynb

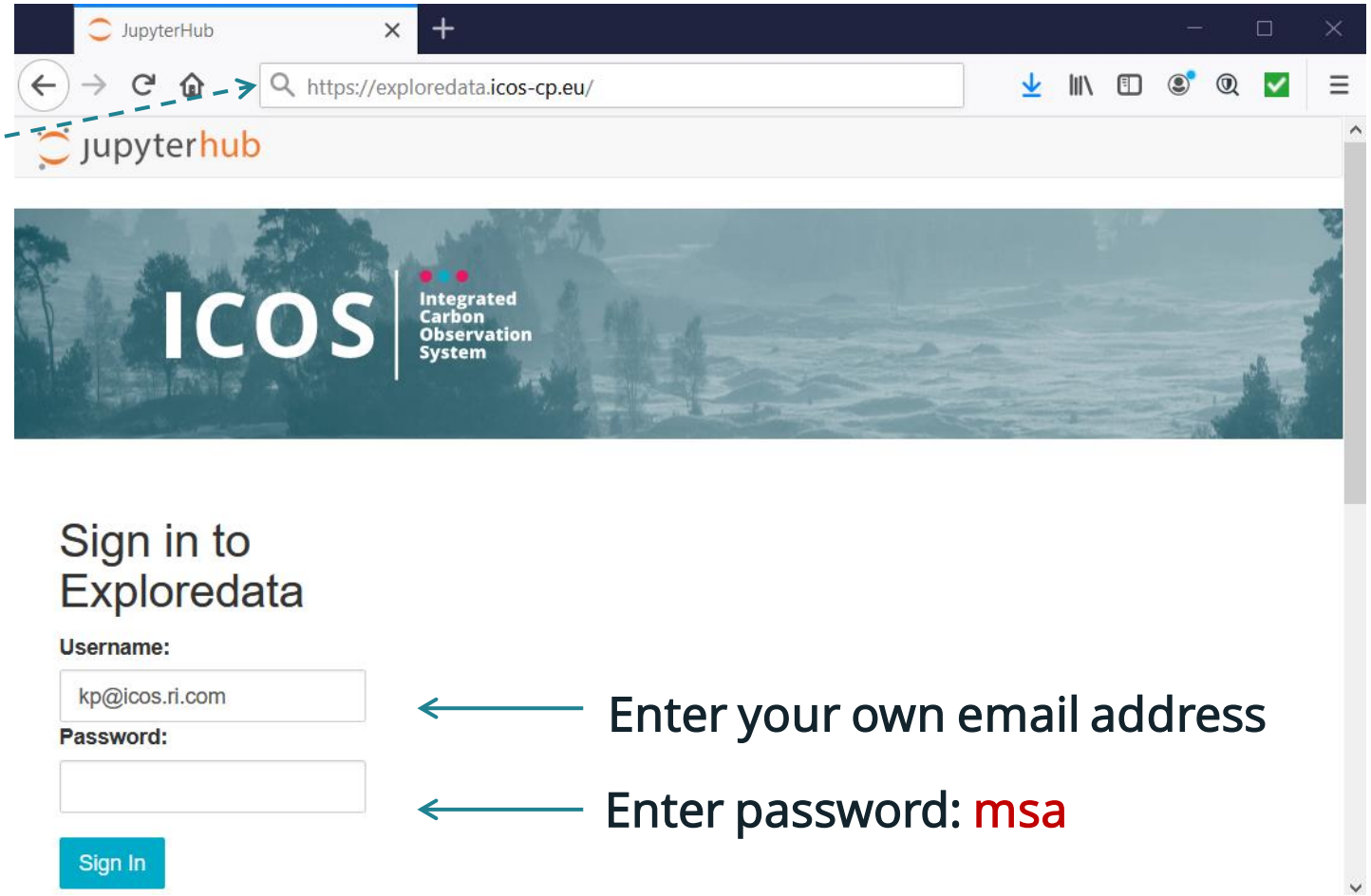
Workshop Schedule

STILT - On demand calculation - Result visualization	Ute	15 min
Station characterization - Create a new “data product” based on STILT results	Ida	5 min
Login to VRE	Karolina	5 min
Break		5 min
Hands on – create an interactive map	Claudio	30 min
Break		10 min
Hands on – create a timeseries plot	Karolina	45 min
Conclusion and exercise	Claudio	5 min

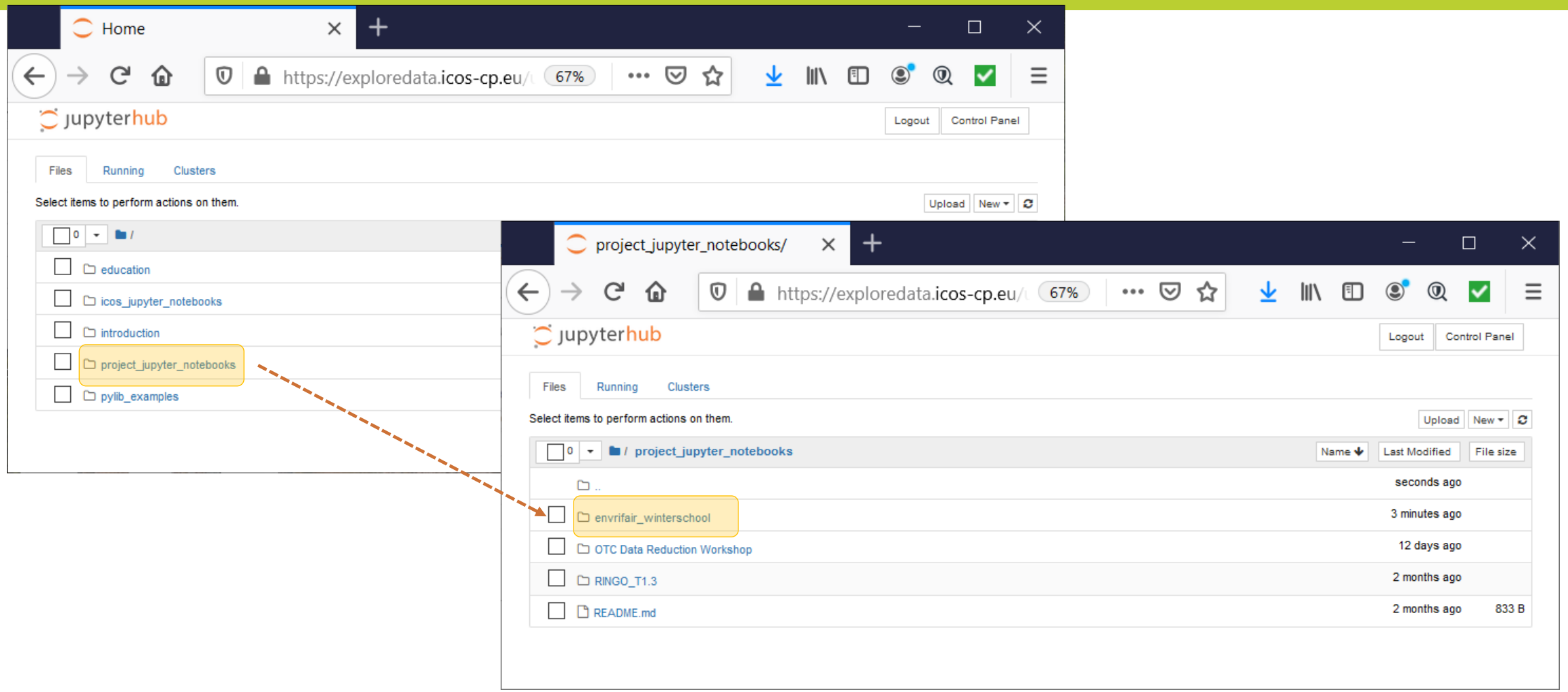
Login to VRE

Enter the following URL in your browser:

https://exploredata.icos-cp.eu

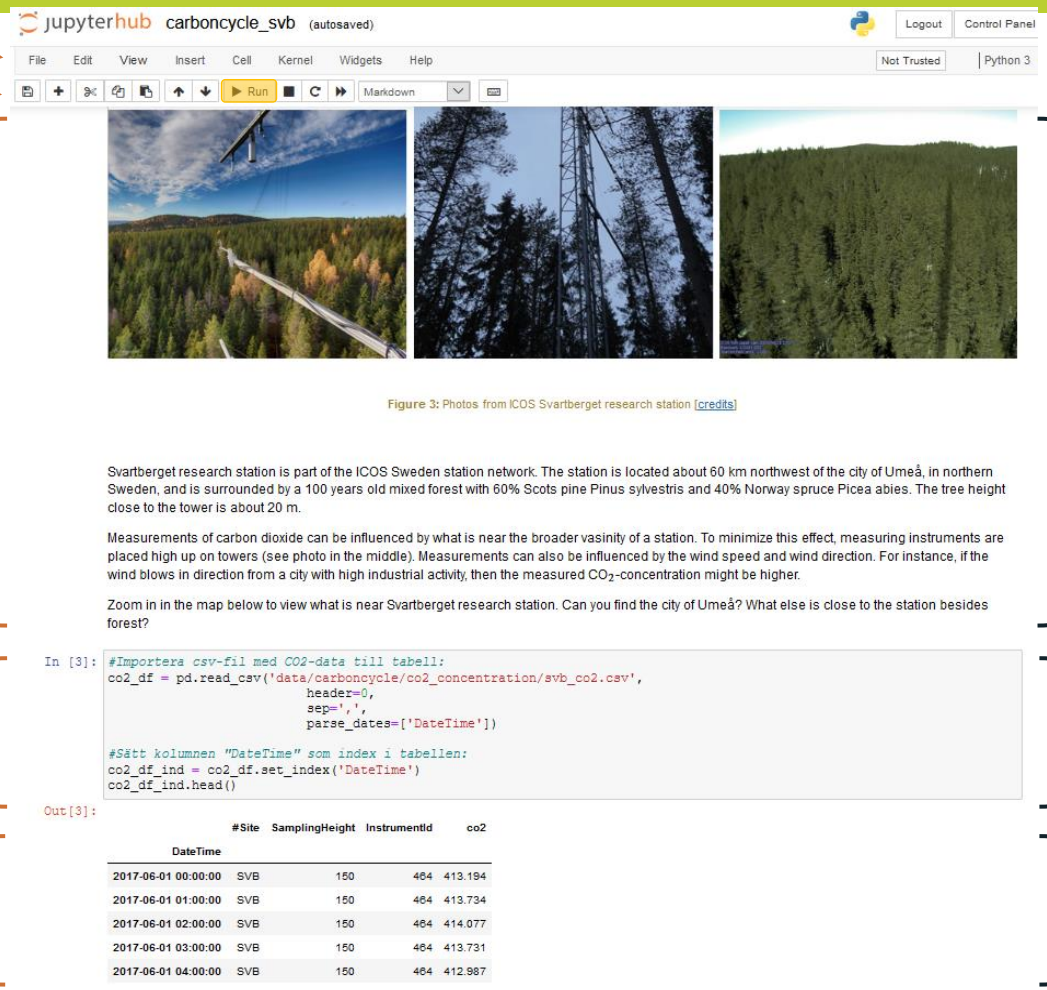


Access working folder



Jupyter Notebook

Menu Bar
Tool Bar
Markdown Cell
Code Cell
Output



Documentation
- ✓ Text
  (regular text, HTML, LateX, etc.)
- ✓ Images
  (png, jpeg, etc.)
- ✓ Links
  (internal & external)

Python Code

Dataframes, lists, arrays, strings, plots, maps, videos…

Please fill out the form

Could we please ask you to fill in your details. If you don't feel comfortable giving us this information, you can create a 'fake' one (with sensible values)

Open this link, and add your information

https://fileshare.icos-cp.eu/apps/onlyoffice/s/ByCDKPH8LC5Jqz4

If you don't know yor location, checkout this site: https://mylocation.org/

Workshop Schedule

STILT - On demand calculation - Result visualization	Ute	15 min
Station characterization - Create a new “data product” based on STILT results	Ida	5 min
Login to VRE	Karolina	5 min
Break		5 min
Hands on – create an interactive map	Claudio	30 min
Break		10 min
Hands on – create a timeseries plot	Karolina	45 min
Conclusion and exercise	Claudio	5 min

back in 5

Workshop Schedule

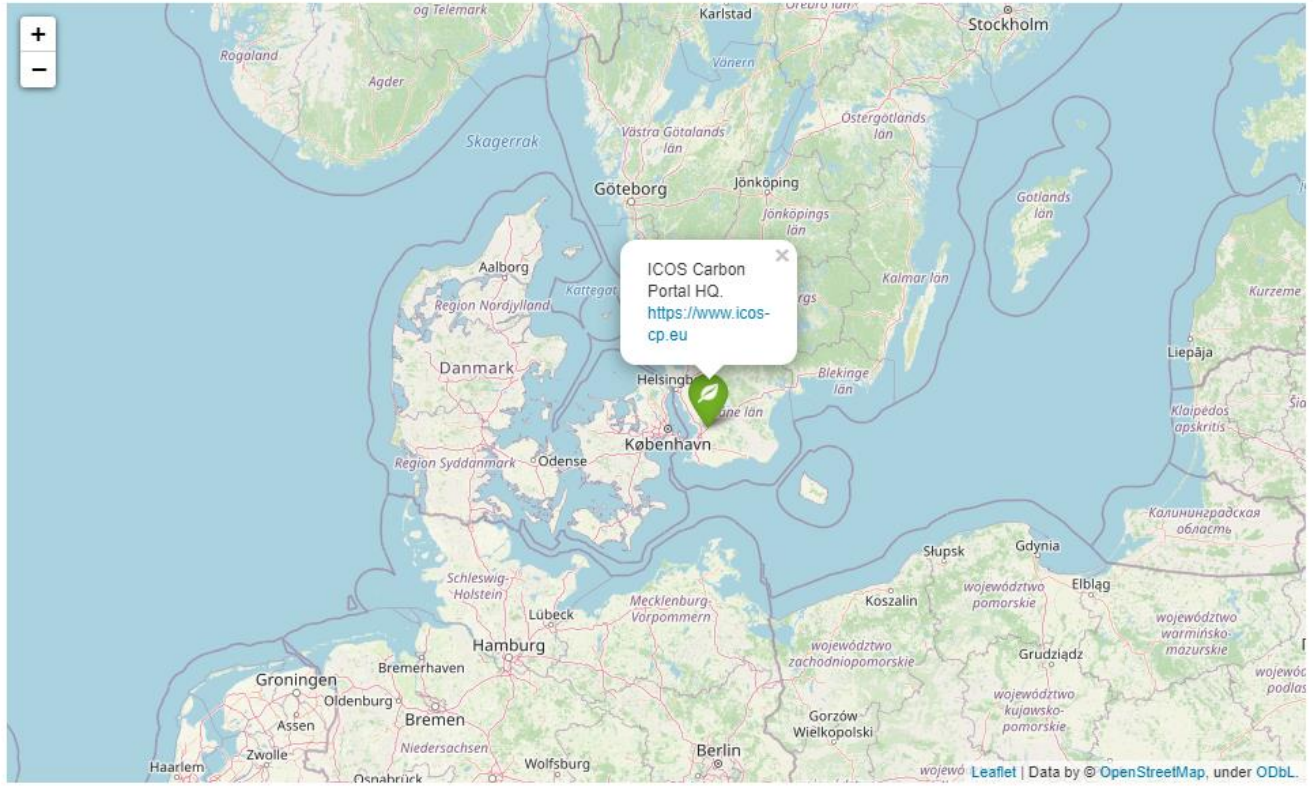
STILT - On demand calculation - Result visualization	Ute	15 min
Station characterization - Create a new “data product” based on STILT results	Ida	5 min
Login to VRE	Karolina	5 min
Break		5 min
Hands on – create an interactive map	Claudio	30 min
Break		10 min
Hands on – create a timeseries plot	Karolina	45 min
Conclusion and exercise	Claudio	5 min

Interactive Map

Show map

```
In [6]:  1 myMap
```

Out[6]:



Workshop Schedule

STILT - On demand calculation - Result visualization	Ute	15 min
Station characterization - Create a new “data product” based on STILT results	Ida	5 min
Login to VRE	Karolina	5 min
Break		5 min
Hands on – create an interactive map	Claudio	30 min
Break		10 min
Hands on – create a timeseries plot	Karolina	45 min
Conclusion and exercise	Claudio	5 min

back in 10

Workshop Schedule

STILT	Ute	15 min
On demand calculation		
Result visualization		
Station characterization	Ida	5 min
Create a new “data product” based on STILT results		
Login to VRE	Karolina	5 min
Break		5 min
Hands on – create an interactive map	Claudio	30 min
Break		10 min
Hands on – create a timeseries plot	Karolina	45 min
Conclusion and exercise	Claudio	5 min

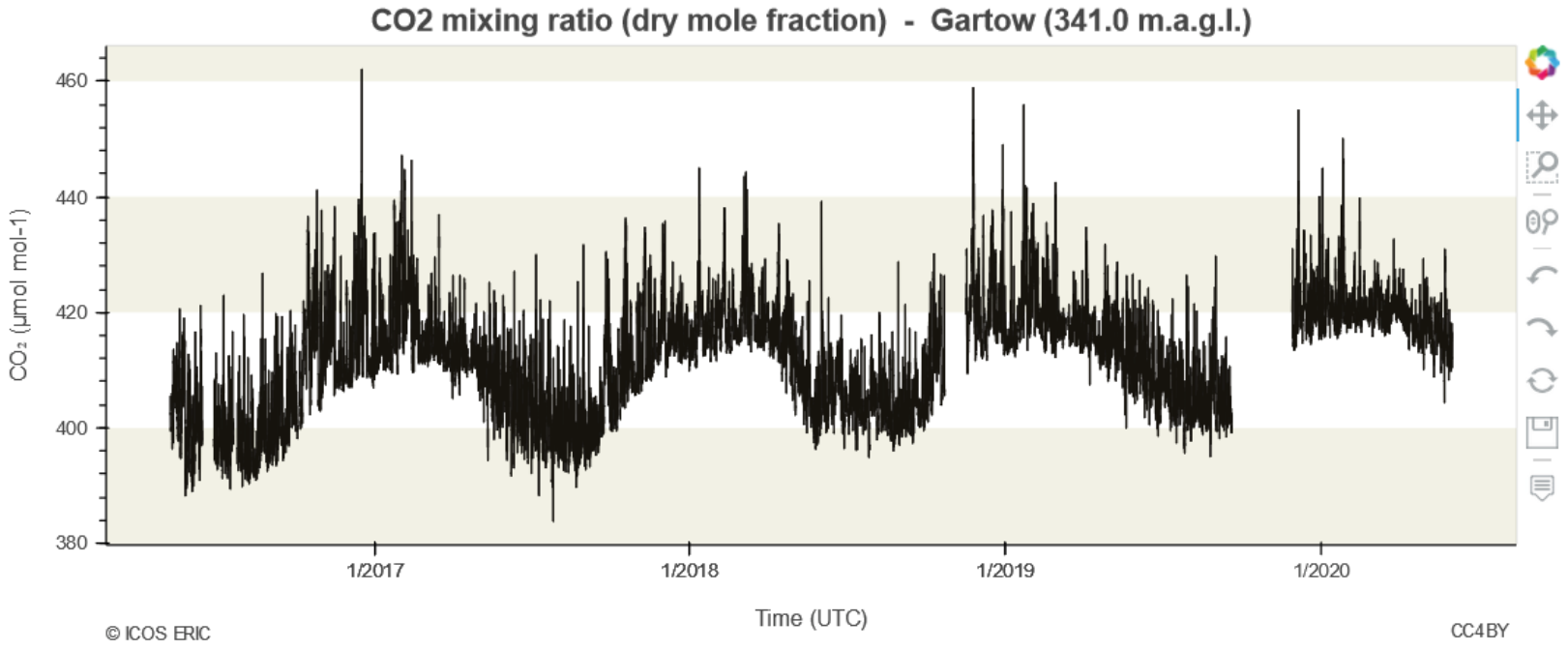
Timeseries Plot

4.5. Interactive plot - using ICOS tools

In [16]:

```
#Call function to get plot object:
p = plot_icos_single_station_binary(dobj, 'co2', 'line', color='#0F0C08') #line/point

#Show interactive plot:
show(p)
```



Workshop Schedule

STILT - On demand calculation - Result visualization	Ute	15 min
Station characterization - Create a new “data product” based on STILT results	Ida	5 min
Login to VRE	Karolina	5 min
Break		5 min
Hands on – create an interactive map	Claudio	30 min
Break		10 min
Hands on – create a timeseries plot	Karolina	45 min
Conclusion and exercise	Claudio	5 min

Exercise

Create a new notebook to include your solutions to the following tasks:

I. Write code to produce an interactive map with a marker for a station of your choice. Add functionality so that a pop-up message displaying station metadata appears, when the marker is clicked.
(select a station from the station list included in the *"1. Explore ICOS atmosphere stations"* part

of the *time series* notebook).

II. Repeat all steps in the *time series* notebook for the station you selected in *task I*.
Note that you are meant to pick a different station than the one used in the notebook example!

Submit exercise

- Remember to enter your **name** and **affiliation** in the bottom part of the notebook.
- Export your notebook as a **ipynb**, **pdf** or **html** file and change the filename to include you name.
  (e.g. filename_kim_andersson.ipynb)
- Please send your answers to: jupyter-info@icos-cp.eu
- Observe that exploredata **deletes** your instance after **15 min of inactivity**, so make sure to download your notebook if you do not plan to complete the exercise in one go.

Conclusion

ICOS has opted for Jupyter Hub and Python Notebooks because ...:

- userfriendly
- easy to learn
- easy to share
- enable and empower users to interact and explore data
- write interactive documentation
- create transparency

We think that all of the above are essential building blocks for the users to pave the way for FAIR content

Conclusion

ICOS has opted for Jupyter Hub and Python Notebooks because ...:

- easy to deploy
- out of the box
- customize and adjust
- low maintenance

ENVRI FAIR

envri.eu/envri-fair

@envri_fair

company/envri-fair

facebook.com/ENVRIcomm