SeaDataCloud

D10.8: DIVA online operational in VRE

C. Troupin, A. Barth & J.-M. Beckers (GHER - ULiège)

7th TTG meeting, Plouzané, France, 14-15 October, 2019

sdn-userdesk@seadatanet.org – www.seadatanet.org

Definitions

DIVA: software tool designed for spatial interpolation
https://github.com/gher-ulg/DIVA

DIVAnd: n-dimensional version of DIVA
https://github.com/gher-ulg/divand.jl

Julia: high-level, high-performance dynamic programming language for numerical computing
https://julialang.org/

Notebook: documents that contain live code, equations, visualizations and narrative text

Jupyter: open-source web application to create and share notebook
http://jupyter.org/

Jupyterhub: multi-user server for Jupyter notebooks
https://jupyterhub.readthedocs.io/en/latest/

1

Context for DIVA in VRE

Installing DIVA was sometimes…painful

Back in 2007

Installing DIVA was sometimes...painful

or in 2013

What have we improved since then?

1. New mathematical formulation — Barth et al. 2014
2. Julia language — instead of Fortran
3. Only 2 (!!!) input files — data & bathymetry
4. Applications as Jupyter notebooks — all in one

What have we improved since then?

Founder Collective
@fcollective

Follow

Congrats to the @JuliaLanguage team on their 1.0 release! We look forward to watching the @JuliaComputing team use it to smash the competition like so much bœuf à la Bourguignonne! github.com/JuliaLang/juli … #ProudInvestor

11:59 PM - 8 Aug 2018

What have we improved since then?

Founder Collective
@fcollective

Follow

Congrats to the @JuliaLanguage team on their 1.0 release! We look forward to watching the @JuliaComputing team use it to smash the competition like so much bœuf à la Bourguignonne! github.com/JuliaLang/juli … #ProudInvestor

11:59 PM - 8 Aug 2018

Jupyter notebooks as guidelines

jupyter 90-full-analysis Last Checkpoint: 02/20/2019 (unsaved changes) Logout

File Edit View Insert Cell Kernel Help Trusted Julia 1.0.3

Markdown

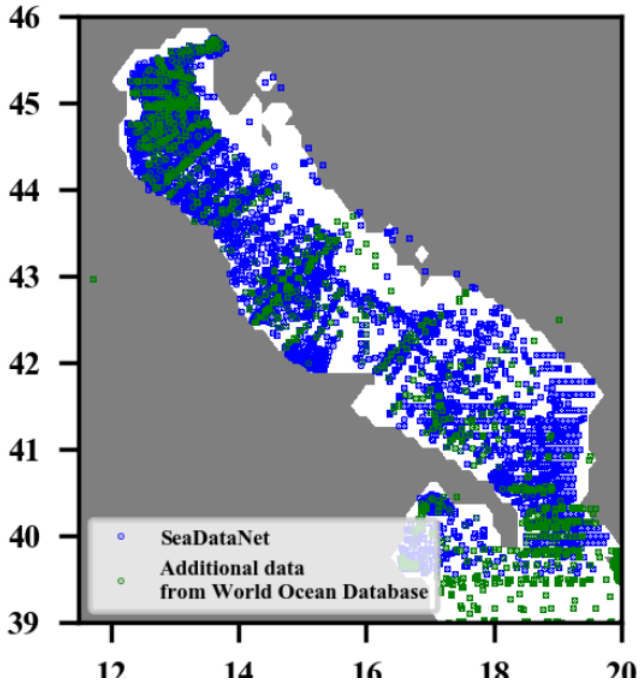
Slide Type Sub-Slide

Create a plot showing the additional data points:

In [27]:

Slide Type Fragment

```
figure("Adriatic-Additional-Data", figsize=(2,2))
ax = subplot(1,1,1)
ax[:tick_params]("both",labelsize=6)
ylim(39.0, 46.0);
xlim(11.5, 20.0);
contourf(bx, by, permutedims(Float64.(mask_edit[:,:,1]),[2,1]),
    levels=[-1e5,0],cmap="binary");
plot(obslon, obslat, "bo", markersize=.2, label="SeaDataNet")
plot(obslonwod[newpoints], obslatwod[newpoints], "go",
    markersize=.2, label="Additional data\nfrom World Ocean Database")
legend(loc=3, fontsize=4)
gca()[:set_aspect](aspect_ratio)
```



Jupyter notebooks as guidelines

jupyter 90-full-analysis Last Checkpoint: 02/20/2019 (unsaved changes) Logout

File Edit View Insert Cell Kernel Help Trusted Julia 1.0.3

Markdown

Slide Type Sub-Slide

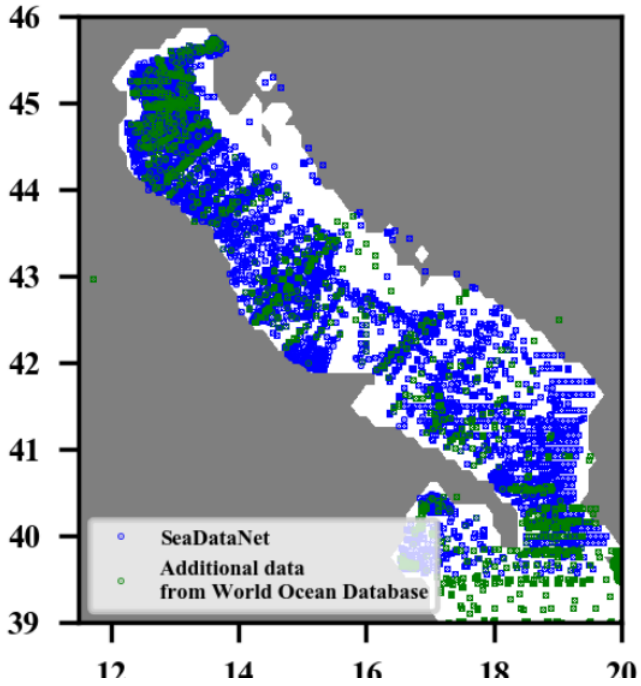
Create a plot showing the additional data points:

Explenatory text

In [27]:

Slide Type Fragment

```
figure("Adriatic-Additional-Data", figsize=(2,2))
ax = subplot(1,1,1)
ax[:tick_params]("both",labelsize=6)
ylim(39.0, 46.0);
xlim(11.5, 20.0);
contourf(bx, by, permutedims(Float64.(mask_edit[:,:,1]),[2,1]),
    levels=[-1e5,0],cmap="binary");
plot(obslon, obslat, "bo", markersize=.2, label="SeaDataNet")
plot(obslonwod[newpoints], obslatwod[newpoints], "go",
    markersize=.2, label="Additional data\nfrom World Ocean Database")
legend(loc=3, fontsize=4)
gca()[:set_aspect](aspect_ratio)
```



Jupyter notebooks as guidelines

jupyter 90-full-analysis Last Checkpoint: 02/20/2019 (unsaved changes) Logout

File Edit View Insert Cell Kernel Help Trusted Julia 1.0.3

Markdown

Slide Type Sub-Slide

Create a plot showing the additional data points:

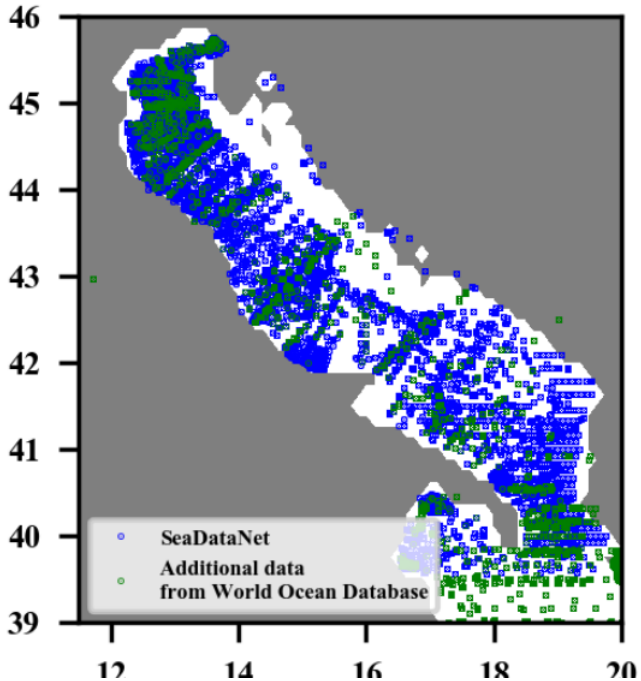
Explenatory text

In [27]:

Slide Type Fragment

```
figure("Adriatic-Additional-Data", figsize=(2,2))
ax = subplot(1,1,1)
ax[:tick_params]("both",labelsize=6)
ylim(39.0, 46.0);
xlim(11.5, 20.0);
contourf(bx, by, permutedims(Float64.(mask_edit[:,:,1]),[2,1]),
    levels=[-1e5,0],cmap="binary");
plot(obslon, obslat, "bo", markersize=.2, label="SeaDataNet")
plot(obslonwod[newpoints], obslatwod[newpoints], "go",
    markersize=.2, label="Additional data\nfrom World Ocean Database")
legend(loc=3, fontsize=4)
gca()[:set_aspect](aspect_ratio)
```

Code fragment



Jupyter notebooks as guidelines

jupyter 90-full-analysis Last Checkpoint: 02/20/2019 (unsaved changes) Logout

File Edit View Insert Cell Kernel Help Trusted Julia 1.0.3

Markdown

Slide Type Sub-Slide

Create a plot showing the additional data points:

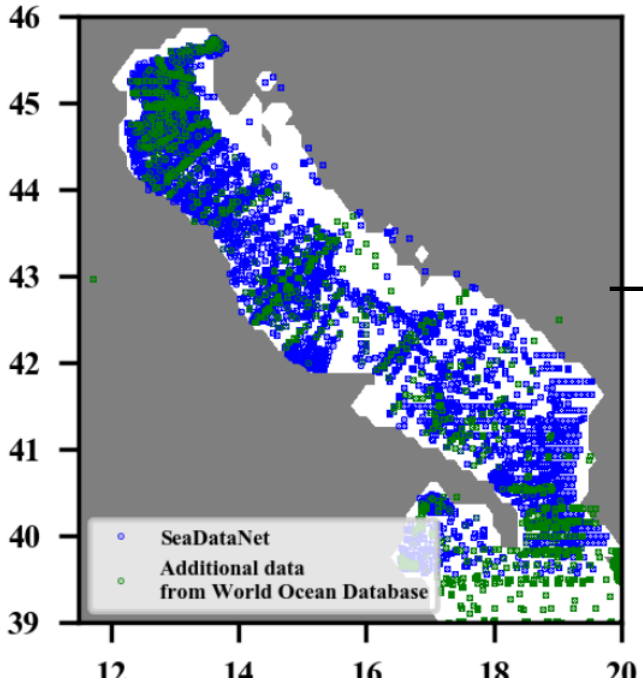
Explenatory text

In [27]:

Slide Type Fragment

```
figure("Adriatic-Additional-Data", figsize=(2,2))
ax = subplot(1,1,1)
ax[:tick_params]("both",labelsize=6)
ylim(39.0, 46.0);
xlim(11.5, 20.0);
contourf(bx, by, permutedims(Float64.(mask_edit[:,:,1]),[2,1]),
    levels=[-1e5,0],cmap="binary");
plot(obslon, obslat, "bo", markersize=.2, label="SeaDataNet")
plot(obslonwod[newpoints], obslatwod[newpoints], "go",
    markersize=.2, label="Additional data\nfrom World Ocean Database")
legend(loc=3, fontsize=4)
gca()[:set_aspect](aspect_ratio)
```

Code fragment



Results or figure

Jupyter notebooks as guidelines

jupyter 90-full-analysis Last Checkpoint: 02/20/2019 (unsaved changes) — Logout

File Edit View Insert Cell Kernel Help — Trusted — Julia 1.0.3

Markdown

Kernel (language)

Slide Type Sub-Slide

Create a plot showing the additional data points:

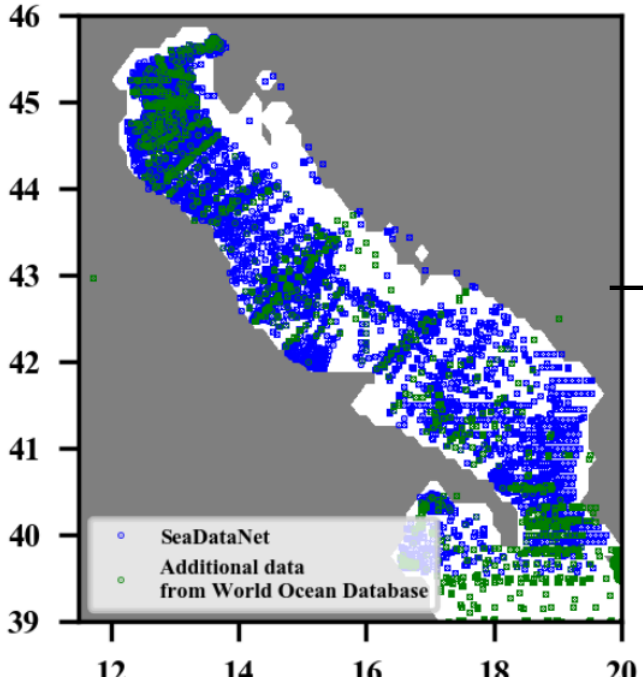
Explenatory text

In [27]:

Slide Type Fragment

```
figure("Adriatic-Additional-Data", figsize=(2,2))
ax = subplot(1,1,1)
ax[:tick_params]("both",labelsize=6)
ylim(39.0, 46.0);
xlim(11.5, 20.0);
contourf(bx, by, permutedims(Float64.(mask_edit[:,:,1]),[2,1]),
    levels=[-1e5,0],cmap="binary");
plot(obslon, obslat, "bo", markersize=.2, label="SeaDataNet")
plot(obslonwod[newpoints], obslatwod[newpoints], "go",
    markersize=.2, label="Additional data\nfrom World Ocean Database")
legend(loc=3, fontsize=4)
gca()[:set_aspect](aspect_ratio)
```

Code fragment



Results or figure

What should we improve?

What should we improve?

+ Access to computing power
+ Data availability
+ Documentation

What should we improve?

+ Access to computing power
+ Data availability
+ Documentation

= Virtual Research Environment!

2

Deployment in the VRE

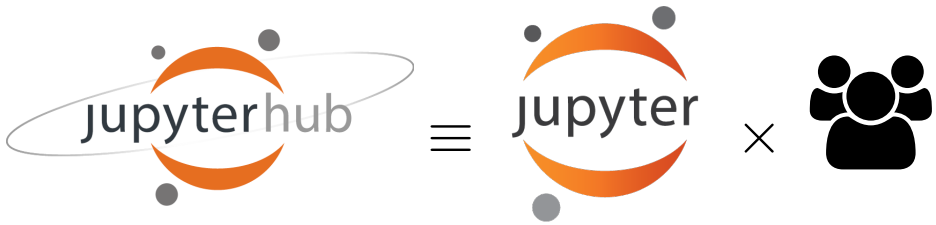
Goals: seamless generation of DIVA products

Data

Tool

Computing power

Providing users access to notebooks



2017 1st deployment @CINECA

2019 Deployment transferred to DKRZ
for the first training workshop

Multiple copies of a Docker container are run
(https://hub.docker.com/r/abarth/divand-jupyterhub)

What's the recipe for the container?

- ☑ Libraries: netCDF, unzip, git, …

What's the recipe for the container?

- Libraries: netCDF, unzip, git, …
- Julia language (V1.2.0)

What's the recipe for the container?

- Libraries: netCDF, unzip, git, ...
- Julia language (V1.2.0)
- Julia packages: PyPlot, NCDatasets, DataStructures, ...

What's the recipe for the container?

- Libraries: netCDF, unzip, git, ...
- Julia language (V1.2.0)
- Julia packages: PyPlot, NCDatasets, DataStructures, ...
- DIVAnd.jl (V2.4.0, August 2019)

What's the recipe for the container?

- Libraries: netCDF, unzip, git, ...
- Julia language (V1.2.0)
- Julia packages: PyPlot, NCDatasets, DataStructures, ...
- DIVAnd.jl (V2.4.0, August 2019)
- DIVAnd notebooks (latest version)

What's the recipe for the container?

- Libraries: netCDF, unzip, git, ...
- Julia language (V1.2.0)
- Julia packages: PyPlot, NCDatasets, DataStructures, ...
- DIVAnd.jl (V2.4.0, August 2019)
- DIVAnd notebooks (latest version)

Note: no need to edit the docker file when the DIVAnd code or the notebooks are modified

```
RUN julia --eval 'using Pkg;
pkg"add https://github.com/gher-ulg/DIVAnd.jl#master"'
```

Julia is becoming more famous!

Julia is becoming more famous!

TOOLBOX · 30 JULY 2019

Julia: come for the syntax, stay for the speed

Researchers often find themselves coding algorithms in one programming language, only to have to rewrite them in a faster one. An up-and-coming language could be the answer.

Jeffrey M. Perkel

🔗 https://www.nature.com/articles/d41586-019-02310-3

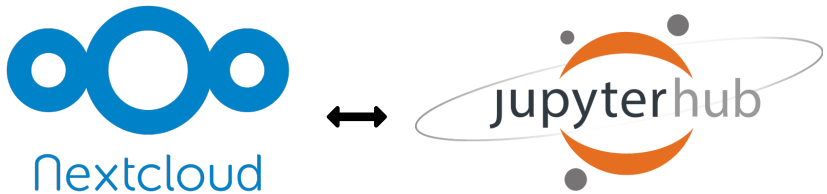
3

Files I/O

DIVAnd file I/O

Pre-processing	Input	Output	Visualisation
webODV	**Observations** SeaDataNet WOD CMEMS **Bathymetry** EMODnet GEBCO	Gridded fields (netCDF) Metadata (XML)	Deltares VIZ

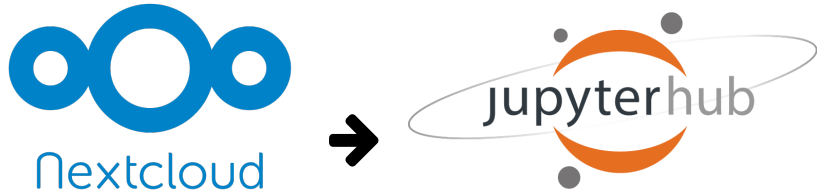
DIVAnd file I/O



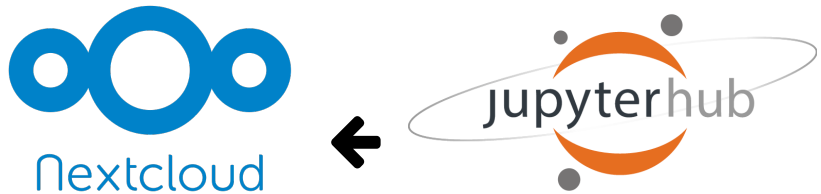
Protocol: web  (Web Distributed Authoring and Versioning)

WEBDAV_USERNAME WEBDAV_PASSWORD WEBDAV_URL	defined at user login

DIVAnd file I/O




```
get("nextcloud_file","jupyterhub_file")
```




```
put("jupyterhub_file","nextcloud_file")
```

4

Multi-user performance during the SDN training course

Oostende, 19–26 June, 2019

1. Introduction to Julia
2. Extraction of a regional bathymetry
3. Data import using ODV
4. Climatology generation

Oostende, 19–26 June, 2019

- ☑ Almost all the users managed to create a small climatology
- ☑ Despite the computation load, the VRE stayed alive

Thanks
for your attention