



1st European Soil Observatory Stakeholder Forum

19-21 October 2021, Soil monitoring and contamination session

Challenges in geo-processing of Large scale Soil datasets. A geographer's perspective.

Leonidas LIAKOS (presenter)

(Geographer - Information Systems Agent - Technology expert)

Panos PANAGOS

(Project Officer - Scientific Research)

JRC.D.3-Land Resources
European Soil Data Centre (ESDAC)

For geography the main research field is “geographical space”.

- **Spatial Analysis**
- **Mapping**
- **Interpretation**
- **Prediction**

Spatial phenomena and their analysis requests **Interdisciplinary** approach.



European
Commission

Challenges working in JRC...



Suddenly, **~7900** raster files
need pre-processing !!!

Big data:

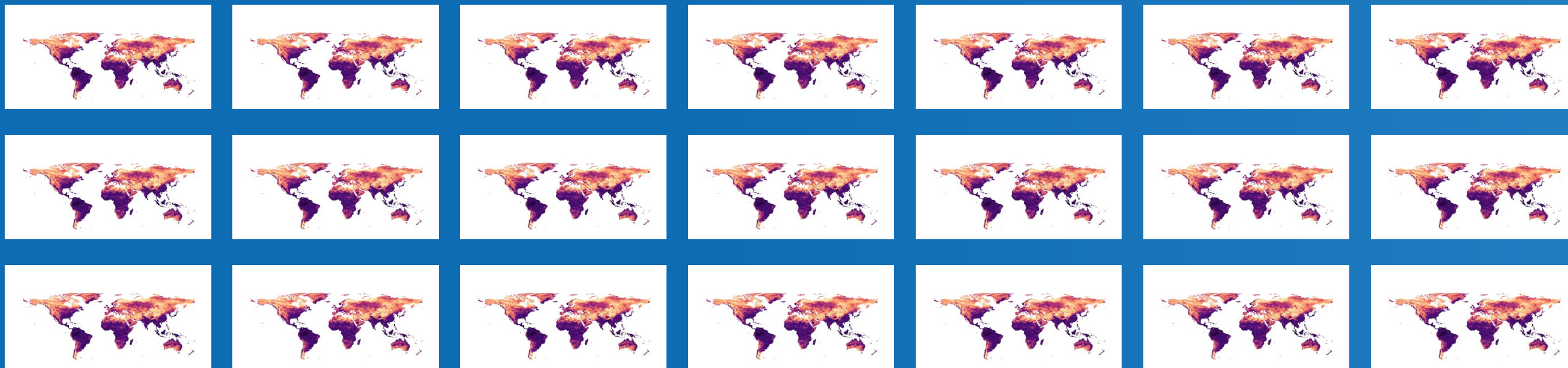
- Large scale (European or global area) datasets
- Long time series



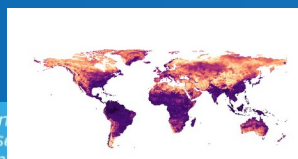
European Commission

Challenges working in JRC...

Dozens of global rasters needs composition.



Pixel based mean



```

LEQNCQ"70"MLZIG(,{ONPQTL)\20"CMIB2"(A99L)"ELOSZITL"(ZCBU9LTO)"(ZIFJOTQEX)"ZT4,"LOIWA(ONPQTL=ONPQTL"LE9L"LE9L" ZCBU9LTO=ZCBU9LTO" ZT
LEQNCQ"ub(p9uq)}[:]=ub'usueqz{9LL" qxts = 0"qfAb=ub' f09fz5 )
9LL{-ub'ZzITIG(9LL)}=ub'usu
9LL = ub'9LONUD(9LL' J2)
9LL = JTW'Z9CQKQ"ub(p9uq)
LOI p9uq ZU LE9D(0" JTW'Z9CQKQ"bLOBEIT9E"uLOIB9UQ()):
LEQNCQ = b)'9LONUD(L"CL0B'J9U(}JTW'Z9CQKQ" 0)
JTW'Z9CQKQ = b)'9LONUD(L"Z9CQKQ'J9U(}JTW'JZ4)
JTWJZ4 = b)'JTWJZ4(}b)'JTW(X'ZIFJOTQEX = ZIFJOTQEX' ZIFJOT9F = ZIFJOT9F) LOI x ZU ZT4(2)
LOI ZIFJOTQEX ZU LE9D(0"ZIFJOT9F):
ZIFJOT9F = 4
  
```



Projects

European or global scale, long time series

- 1) Bezak, Nejc, Matjaž Mikoš, Pasquale Borrelli, Leonidas Liakos, and Panos Panagos. 2021. “**An In-Depth Statistical Analysis of the Rainstorms Erosivity in Europe.**” CATENA 206 (November): 105577. <https://doi.org/10.1016/j.catena.2021.105577>.
- 2) Panagos, Panos, Martin Jiskra, Pasquale Borrelli, Leonidas Liakos, and Cristiano Ballabio. 2021. “**Mercury in European Topsoils: Anthropogenic Sources, Stocks and Fluxes.**” Environmental Research, June, 111556. <https://doi.org/10.1016/j.envres.2021.111556>.
- 3) Panagos, Panos, Cristiano Ballabio, Mihaly Himics, Simone Scarpa, Francis Matthews, et al. 2021. “**Projections of Soil Loss by Water Erosion in Europe by 2050.**” Environmental Science & Policy 124 (October): 380–92. <https://doi.org/10.1016/j.envsci.2021.07.012>.
- 4) Köninger, Julia, Emanuele Lugato, Panos Panagos, Mrinalini Kochupillai, Alberto Orgiazzi, et al. 2021. “**Manure Management and Soil Biodiversity: Towards More Sustainable Food Systems in the EU.**” Agricultural Systems 194: 103251. <https://doi.org/10.1016/j.agsy.2021.103251>.
- 5) Panagos, Panos, Pasquale Borrelli, Francis Matthews, Leonidas Liakos, Nejc Bezak, Nazzareno Diodato, and Cristiano Ballabio. 2021. “**Global Rainfall Erosivity Projections for 2050 and 2070.**” Journal of Hydrology (*submitted 13/09/2021*).
- 6) **Phosphorus budget in European agricultural topsoils with an empirical model** (*under development*).
- 7) **Estimating water table depth from Sentinel-1 C-band at large-scale** (*Natural Capital Project*).



European Commission

Solution provided by JRC...

JRC Big Data Platform (JEODPP):

Petabyte scale data hub

JEO-desk: Remote data science desktop (based on Ubuntu)

JEO-batch: High-Performance Computing mechanism (based on htcondor)

JEO-lab: JupyterLab/Python kernel

Distributed computing

Secure and encrypted communication

Multi-factor authentication

Secure file transfer

Critical Geospatial Data Collections already available (Copernicus and USGS, Base and Project data)

Issue tracking and wiki

Standard CPU processing servers:

~2,000 cores, 12-19 GB RAM per core (JEO-batch/desk/lab) • ~500 cores (Other services, like download, DBs, dev, ...)



JEODPP

The JRC Earth Observation Data and Processing Platform (Soille et al.,2017)

Free and open-source software (FOSS) for geospatial analysis

JEODESK:

- GRASS (geoprocessing)
- QGIS (mapping)
- snap/gpt (image processing)
- Htcondor (batch processing)
- R/Rstudio, Python/Spyder (geoprocessing, analysis, plots)
- GDAL libraries (geodata conversion)





Reproducibility



Scripting: R, Python, bash

pyjeo:

Python package for image processing for geospatial data implemented in JRC Ispra



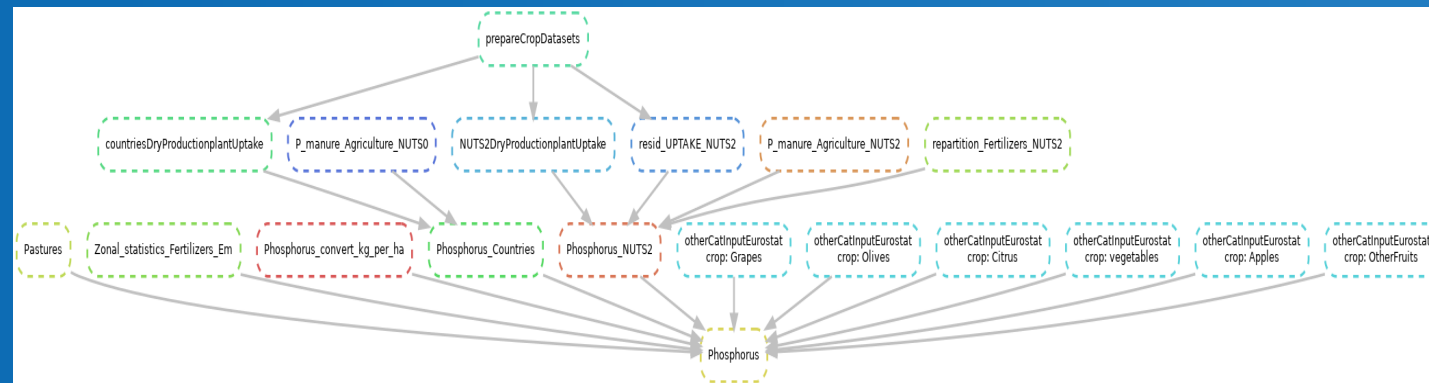
Version Control System: git



Source: Yasemin Turkyilmaz-van der Velden et al 2020



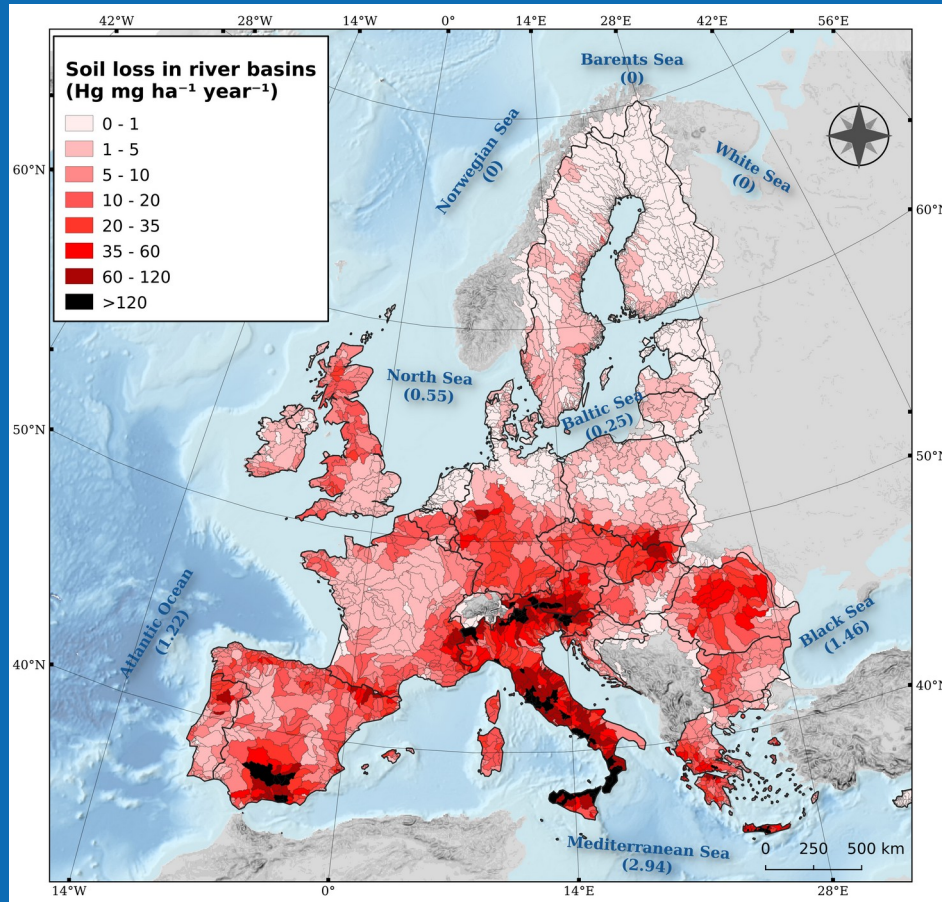
Workflow management:
Snakemake



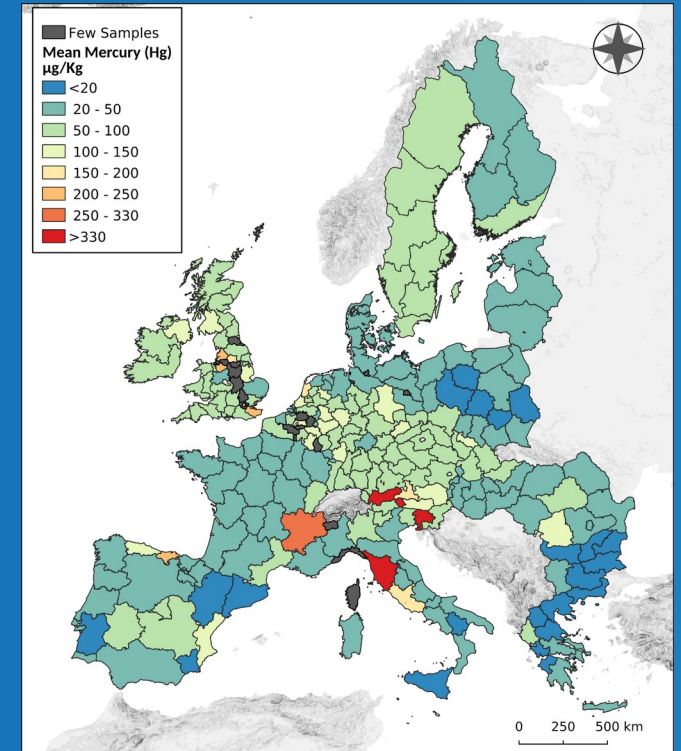
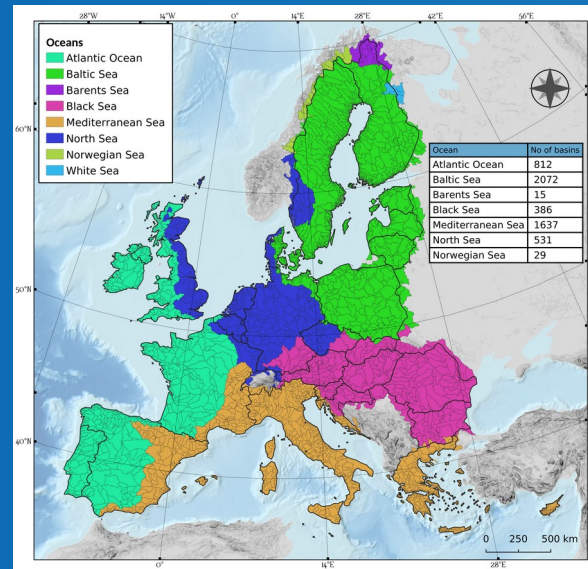
Maps-Plots-Techniques

Exploratory Data Analysis, Descriptive Statistics, Geoprocessing

- Zonal statistics
- Composites
- Data harmonization
- Projections
- Plots
- Maps

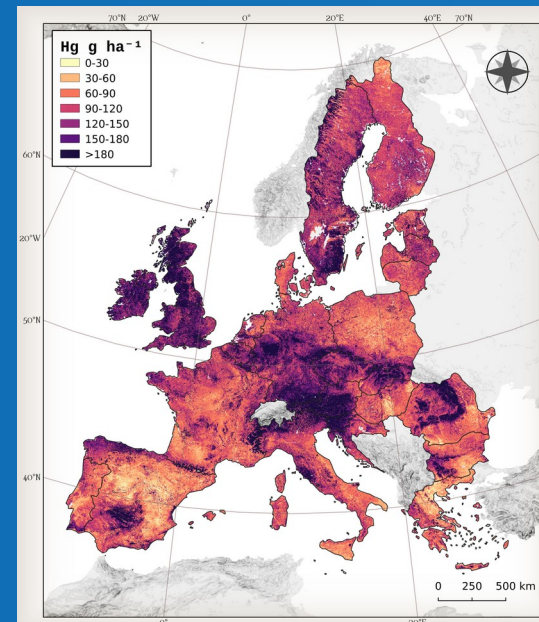
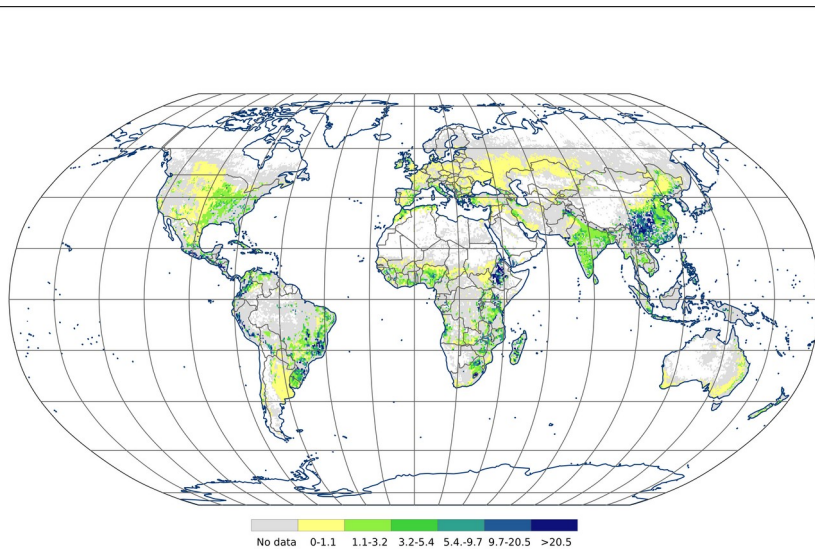
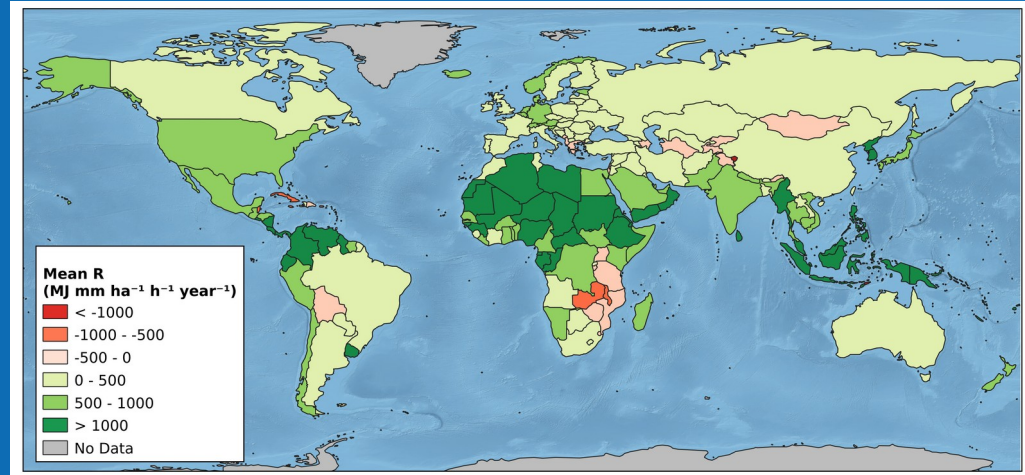
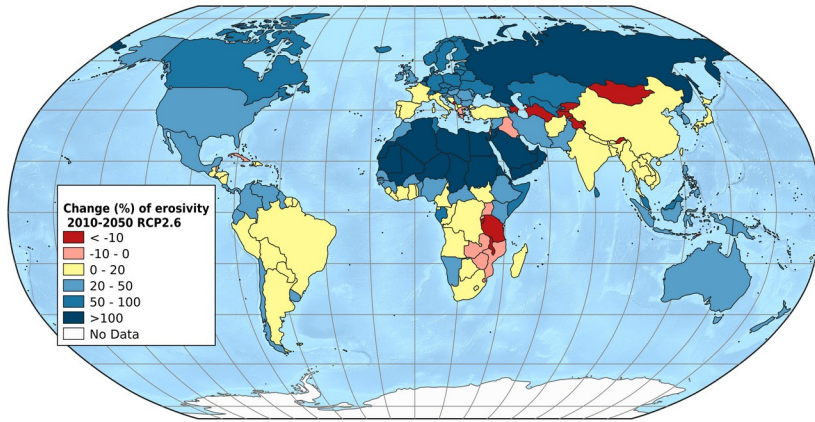


Mercury in European topsoils: Anthropogenic sources, stocks and fluxes (Panagos et al. 2021)





European Commission



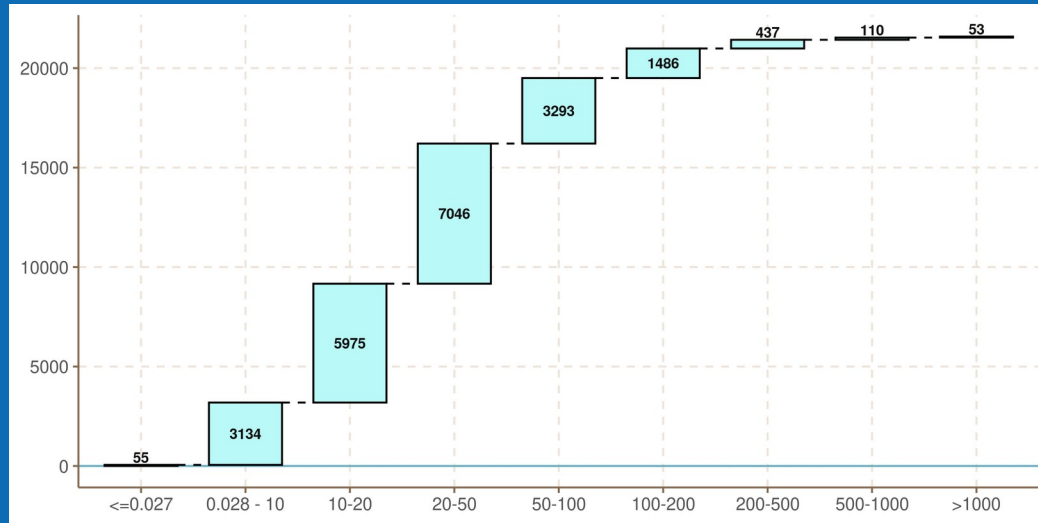
Projections:

- WGS '84 (world maps)
- Robinson (world maps)
- ETRS89-extended/LAEA Europe

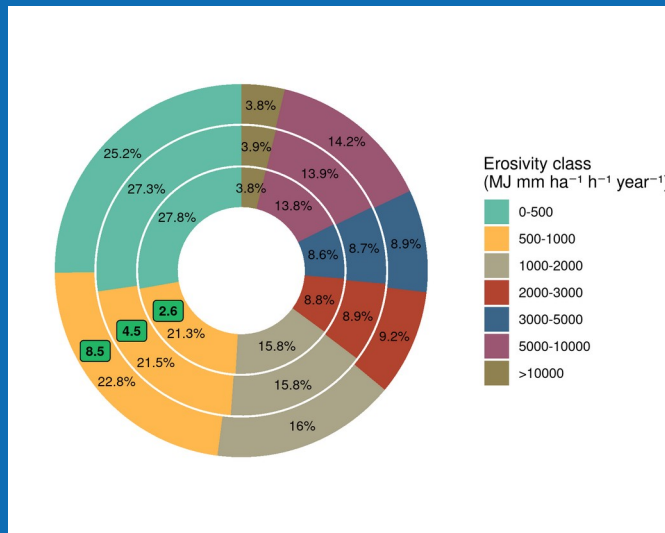
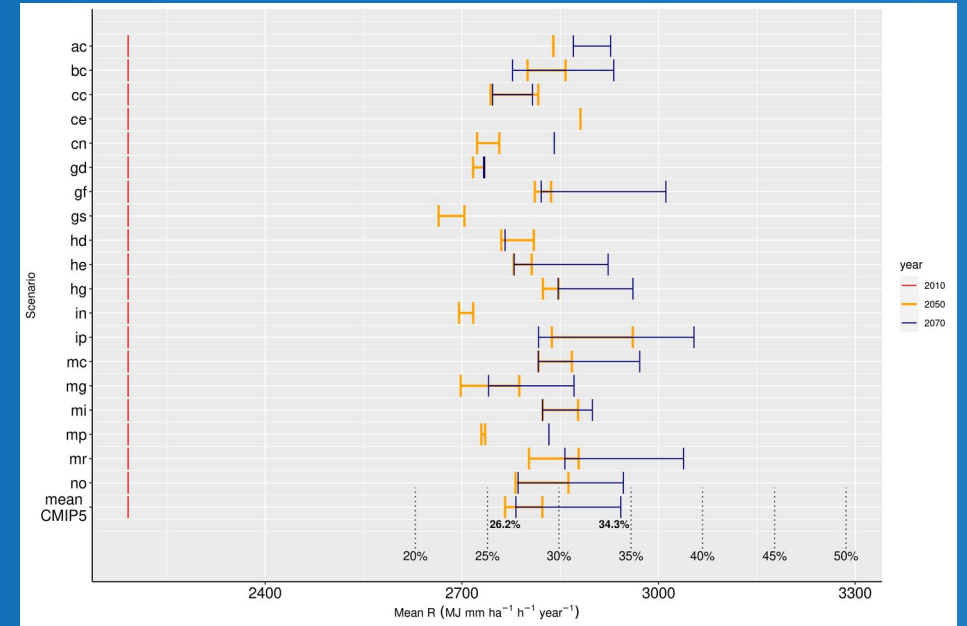


European Commission

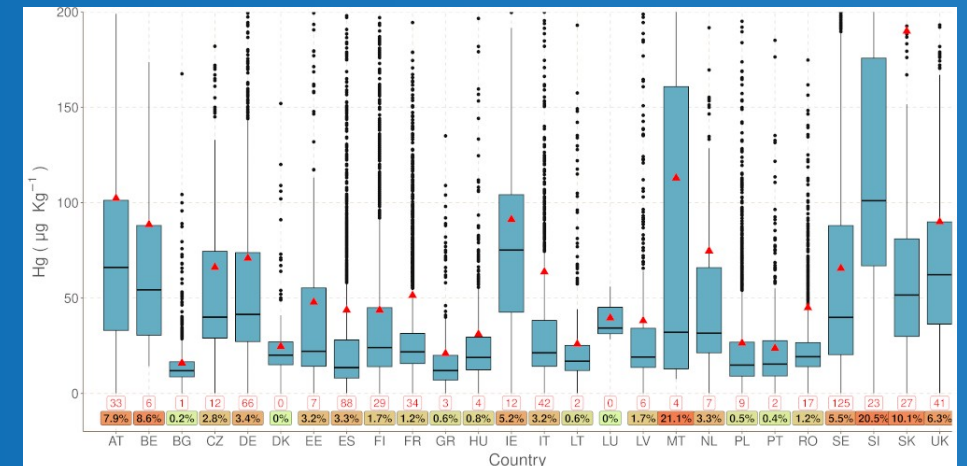
Mercury in European topsoils: Anthropogenic sources, stocks and fluxes (Panagos et al. 2021)



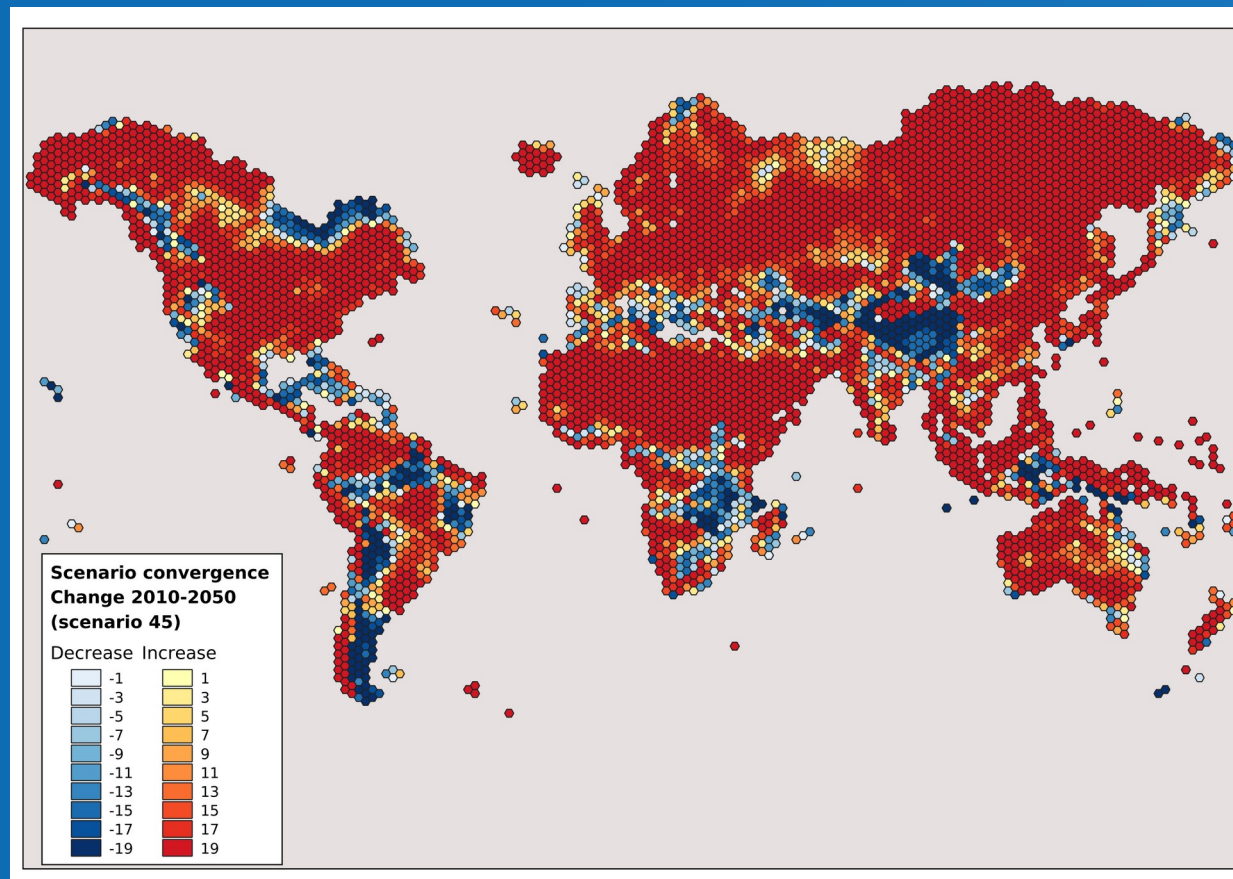
Projections of soil loss by water erosion in Europe by 2050 (Panagos et al. 2021)



Mercury in European topsoils: Anthropogenic sources, stocks and fluxes (Panagos et al. 2021)

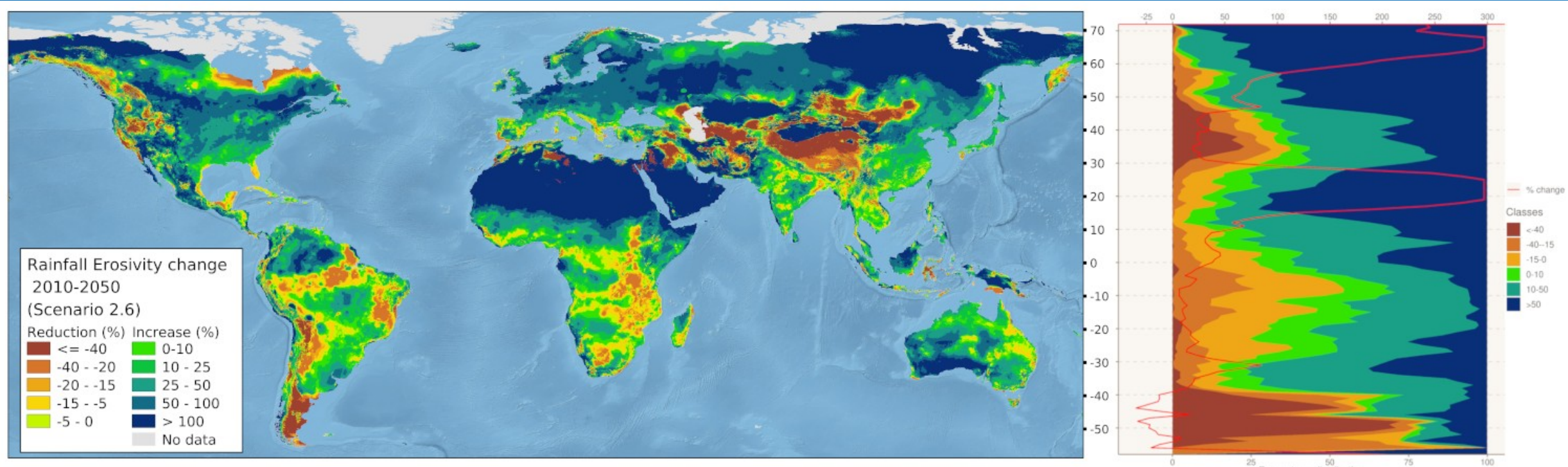


Global Rainfall Erosivity projections for 2050 and 2070 (Panagos et al., submitted)



Innovative maps

Global Rainfall Erosivity projections for 2050 and 2070 (Panagos et al. 2021)

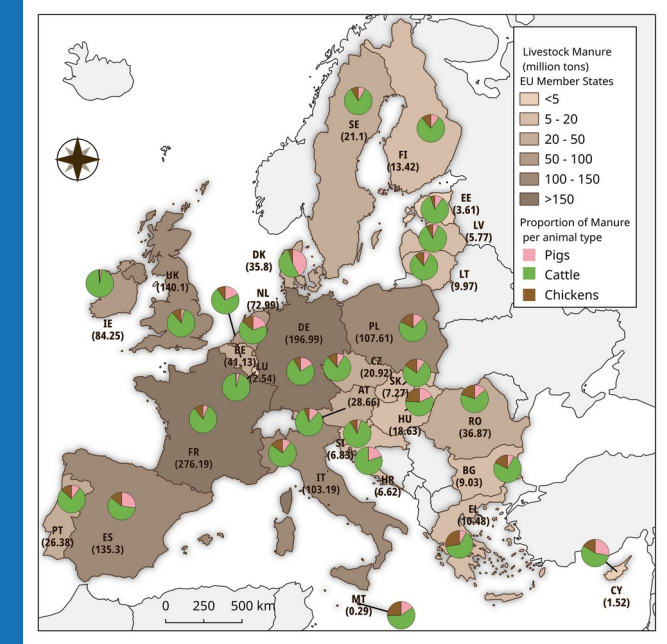
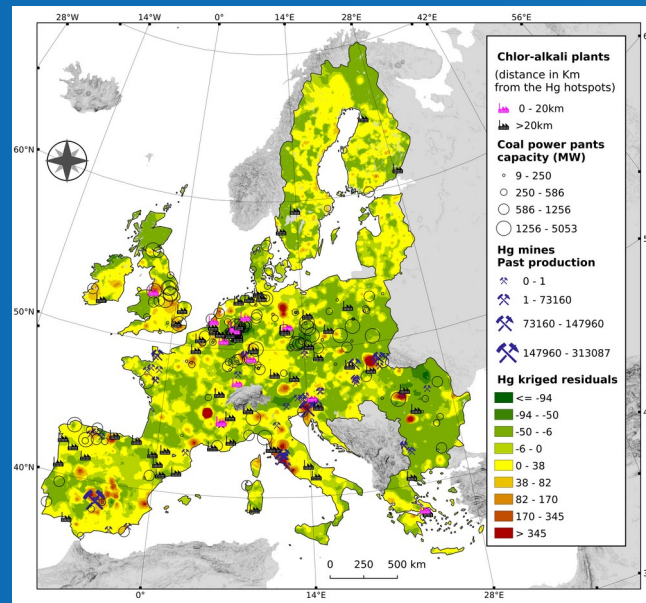
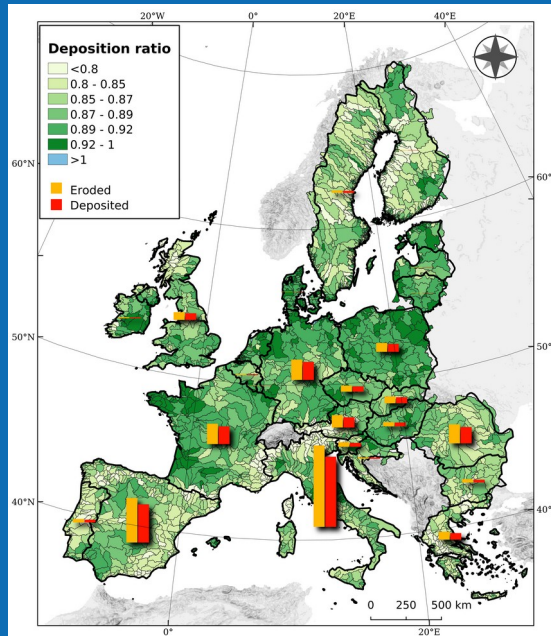




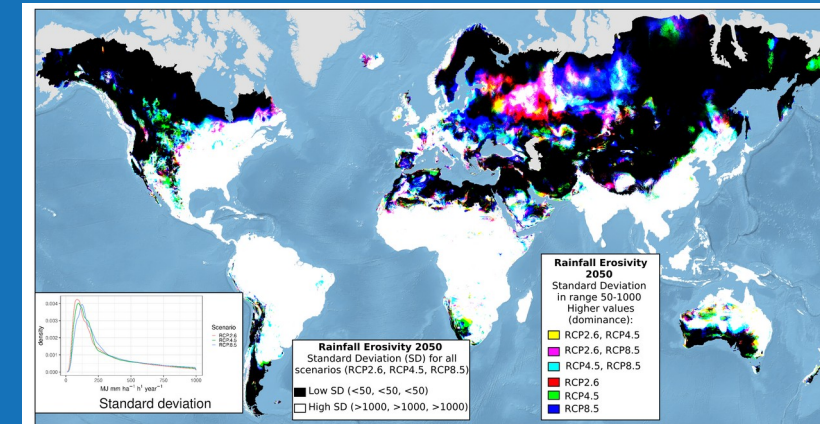
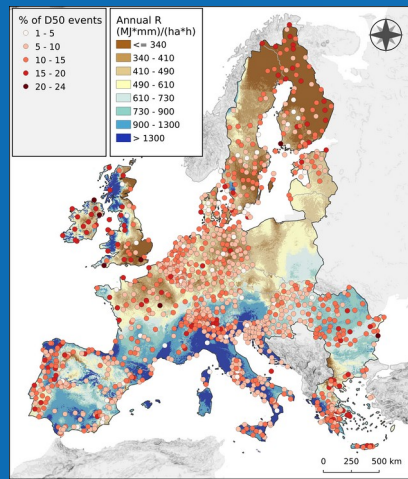
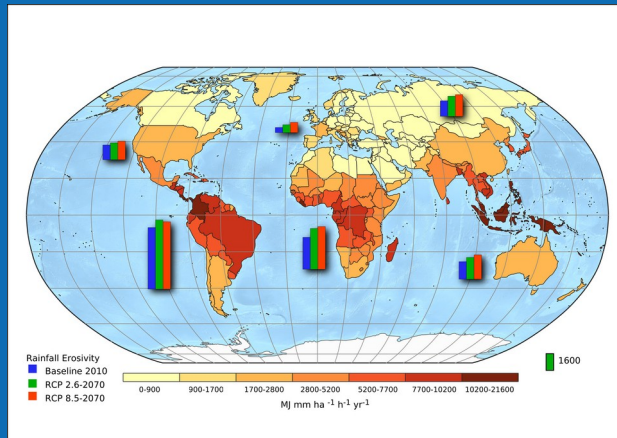
Maps and plots combination and thematic layers

Manure management and soil biodiversity: Towards more sustainable food systems in the EU (Köninger et al., 2021)

Mercury in European topsoils: Anthropogenic sources, stocks and fluxes (Panagos et al., 2021)



An in-depth statistical analysis of the rainstorms erosivity in Europe (Bezák et al., 2021)



Global Rainfall Erosivity projections for 2050 and 2070 (Panagos et al., submitted)



Conclusions

- **Big data** is the new era in geospatial world.
- **JEODPP** provides adequate resources and new opportunities in research.
- **Combination** of **geospatial** (maps) and **descriptive** information (plots) provides Intuitive, Analytic and Inferential insights.
- **Reproducibility** is a critical point in research and science and it should be widely adopted. **Interdisciplinary approach** for the analysis of phenomena with spatial dimensions.