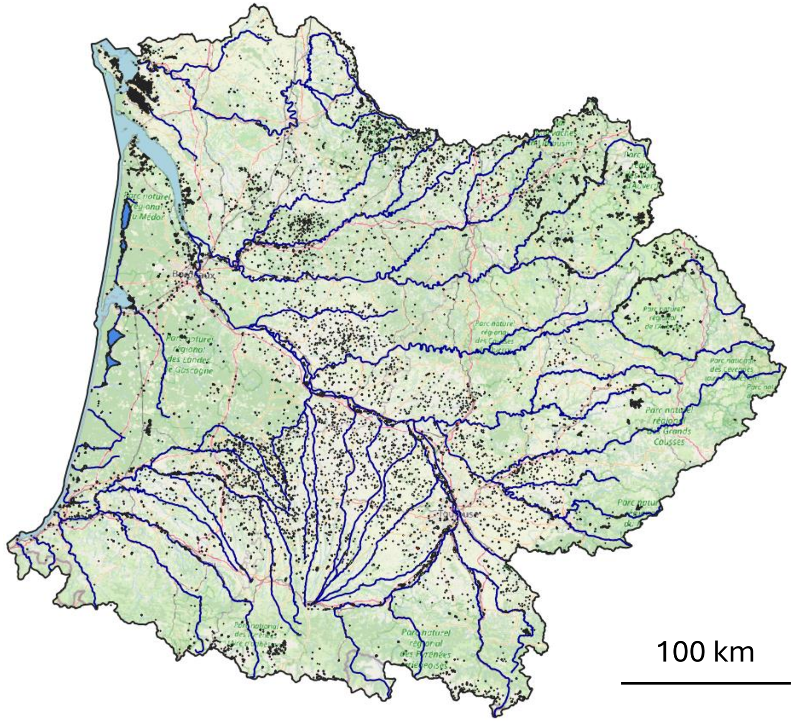
Assessment of eutrophication dynamics of lakes at a large scale by coupling Sentinel-2 remote-sensing, machine-learning and field observations

Mathilde Joffre, Roxelane Cakir, Vanessa Dos Santos, Matheus Tavares, Joana Roussillon, Jean-Michel Martinez, Sabine Sauvage

Study area

Adour-Garonne basin

- ~ 120 000 km^2
- 9672 lakes > 1 ha
- Monitored lakes: 165



GRAND SUD-OUEST
AGENCE DE L'EAU ADOUR-GARONNE

Context

Climate change

↗ frequency and duration of droughts

Impacts on water resources

↘ low flow discharges (July-October)

(LIFE Eau&Climat; Explore2)

↗ Eutrophication risk

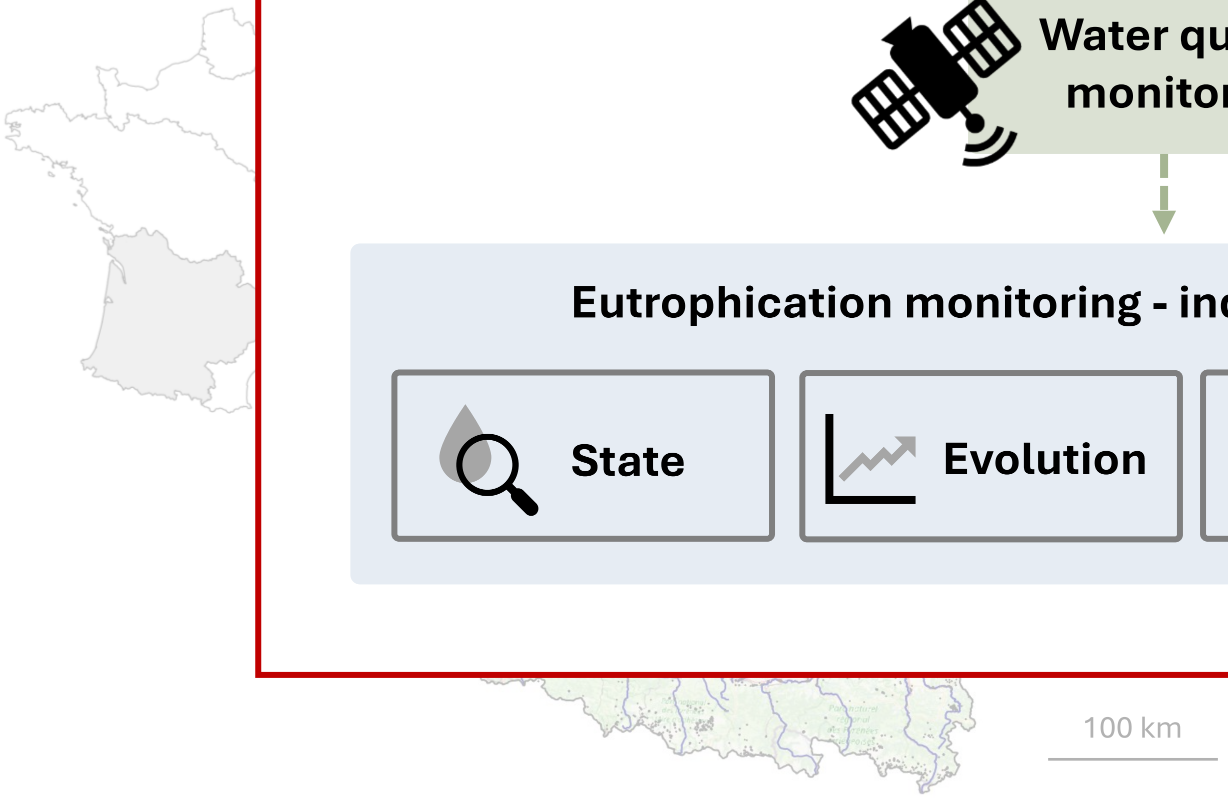
Uses ↔ Water quality ↔ Biodiversity

Introduction

XTREMQUALITY SCO SPACE FOR CLIMATE OBSERVATORY

Study area

Context



hts

er)

sity

Framework objectives

Water quality monitoring

Eutrophication monitoring - indicators

State | **Evolution** | **Drivers**

Including small, unmonitored lakes (>1ha)

Sentinel-2 ChlA/turbidity

L1C

Sentinel-2/MSI images → **MAJA** - L2A
Surf. reflectance
Cloud/snow mask

↓ **GRS**
Atmospheric & sunglint correction
Harmel et al. 2018

↓ **WaterDetect**
Water mask
Cordeiro et al. 2021

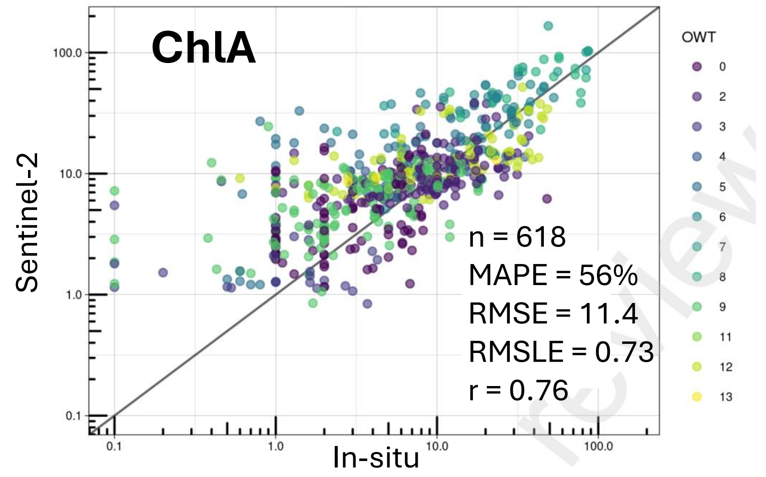
L2A

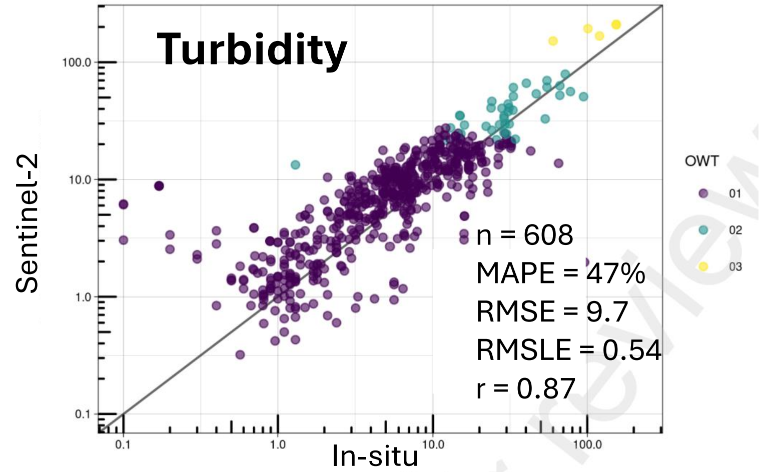
Remote sensing Reflectance (Rrs)

L2B

Filtering steps (poor Rrs, adjacency)

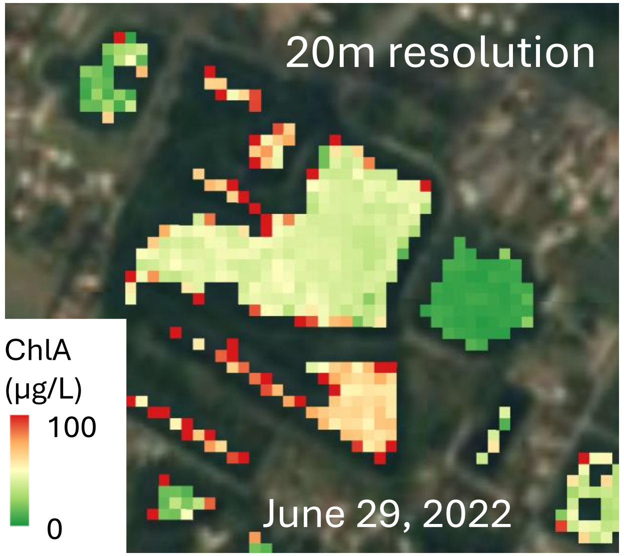
OWT classif.	**Inv. algorithms**	
13 OWT *Spyrakos et al. 2018* →	*Neil et al. 2019; Liu et al. 2021* →	**ChlA**
4 OWT *Cordeiro, 2022* →	*Cordeiro, 2022* →	**Turbidity**





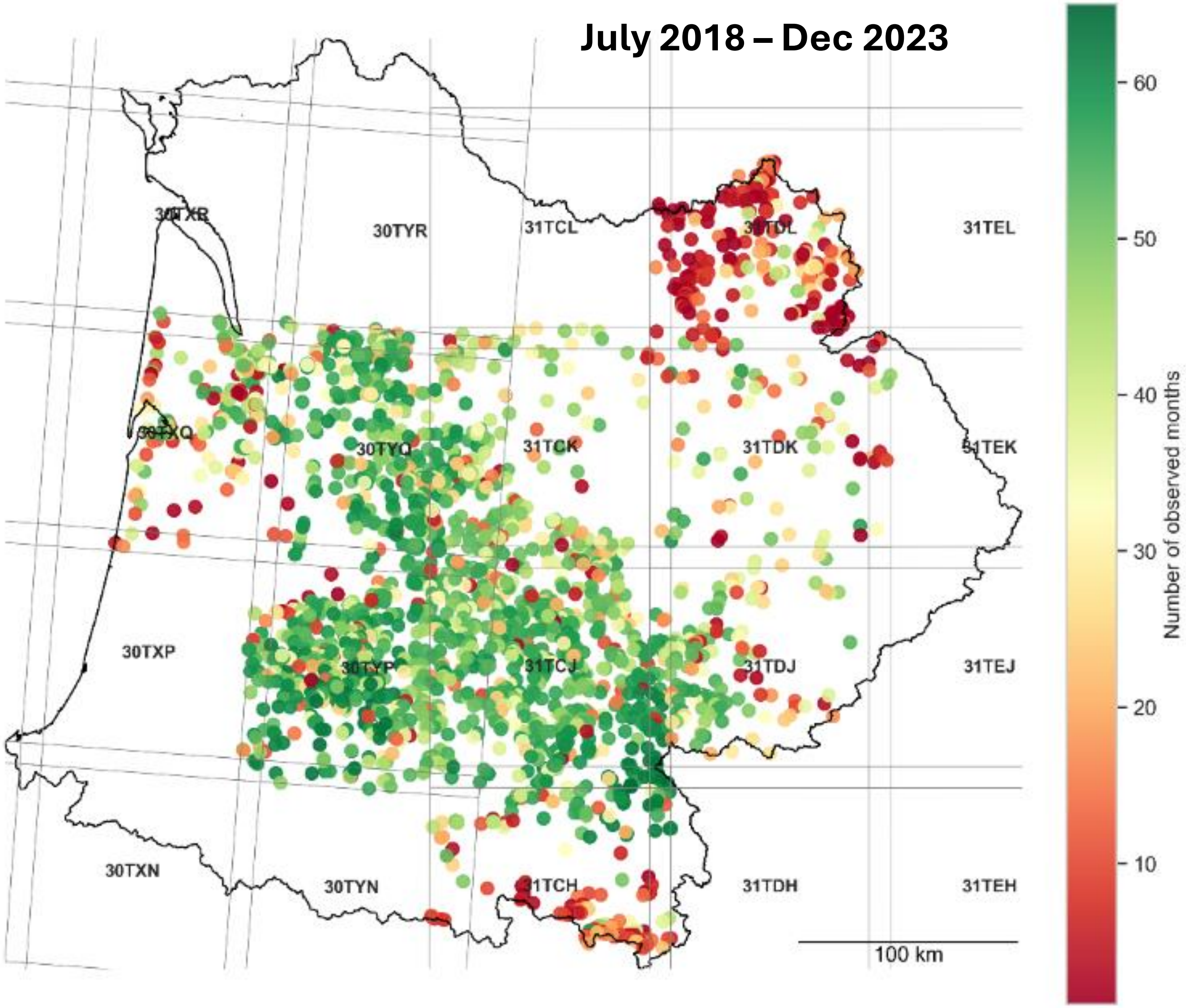
Tavares, M., Guimarães, D., Roussillon, J., Baute, V., Cucherousset, J., Boulêtreau, S., & Martinez, J.-M. (2025). A Framework to Retrieve Water Quality Variables in Small, Optically Diverse Freshwater Ecosystems Using Sentinel-2 Msi Imagery (SSRN Preprint). https://doi.org/10.2139/ssrn.5102644

Time series reconstruction



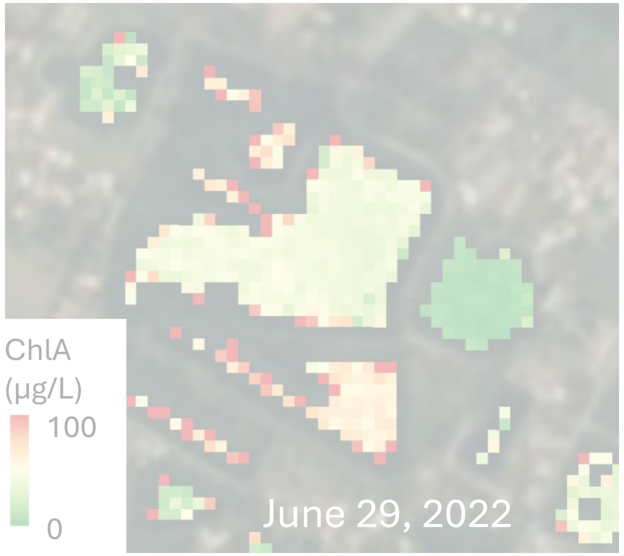
Median value

Monthly aggregation



Clouds, atmospheric correction (altitude), small lakes

Time series reconstruction



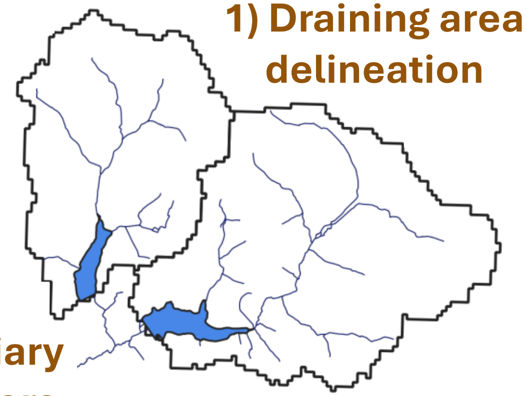
Median value

Monthly aggregation

Missing value prediction

MissForest algorithm

(Stekhoven & Bühlmann., 2012)

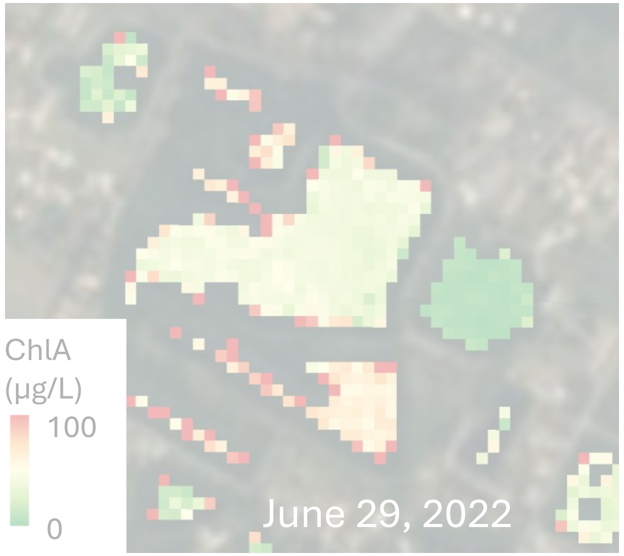


1) Draining area delineation

2) Auxiliary predictors

Predictors	Temporal resolution	*Sources*
Meteorological variables		
Temperature Visible radiation Rainfall	Current month + past 6, 12, 18 and 24 months	*SAFRAN (Meteo France)*
Wind speed	Current month	
Land-use		
% of land use categories	Current year	*IGN*

Time series reconstruction



Median value

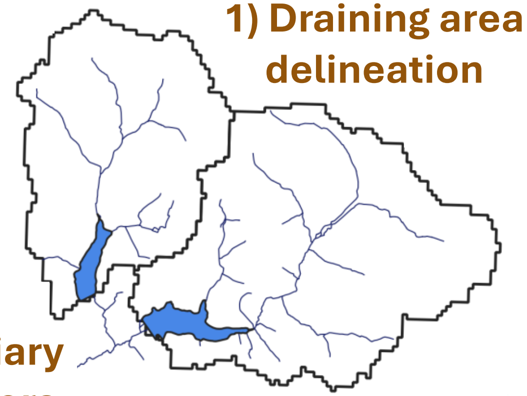
Monthly aggregation

Missing value prediction

MissForest algorithm

(Stekhoven & Bühlmann., 2012)

1) Draining area delineation

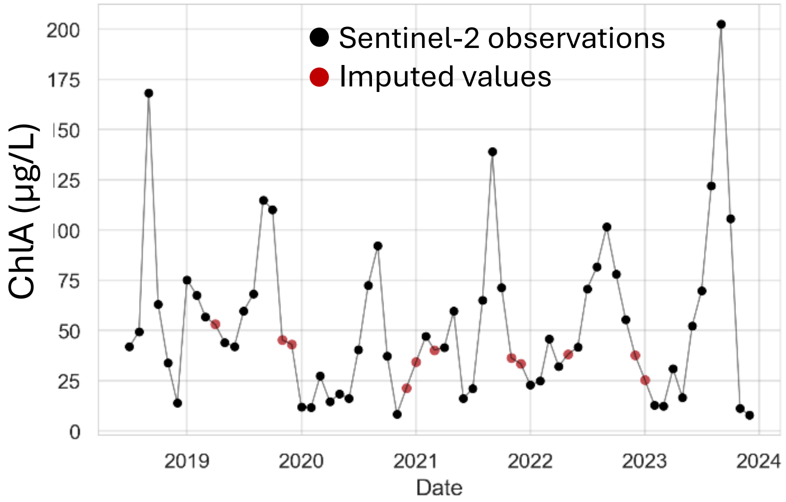


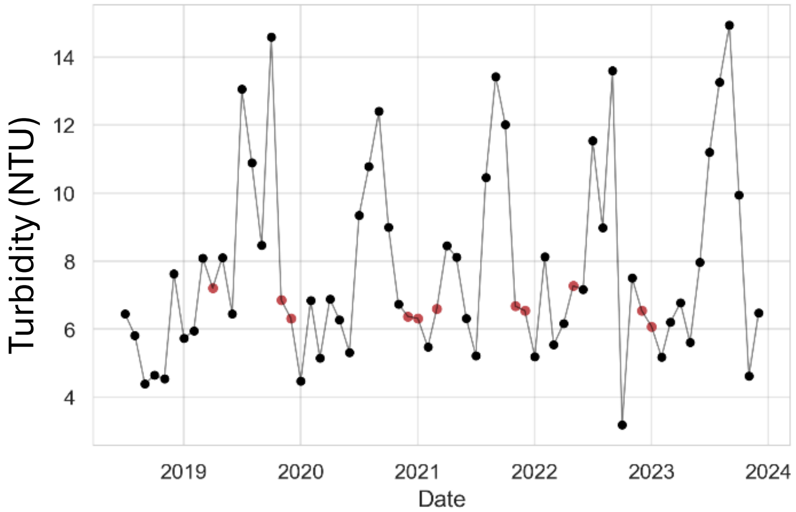
2) Auxiliary predictors

Predictors	Temporal resolution	*Sources*
Meteorological variables		
Temperature	Current month + past 6, 12, 18 and 24 months	*SAFRAN (Meteo France)*
Visible radiation		
Rainfall		
Wind speed	Current month	
Land-use		
% of land use categories	Current year	*IGN*

3) Time series imputation (July 2018 – 2023)

2125 lakes (min obs nb = 20)



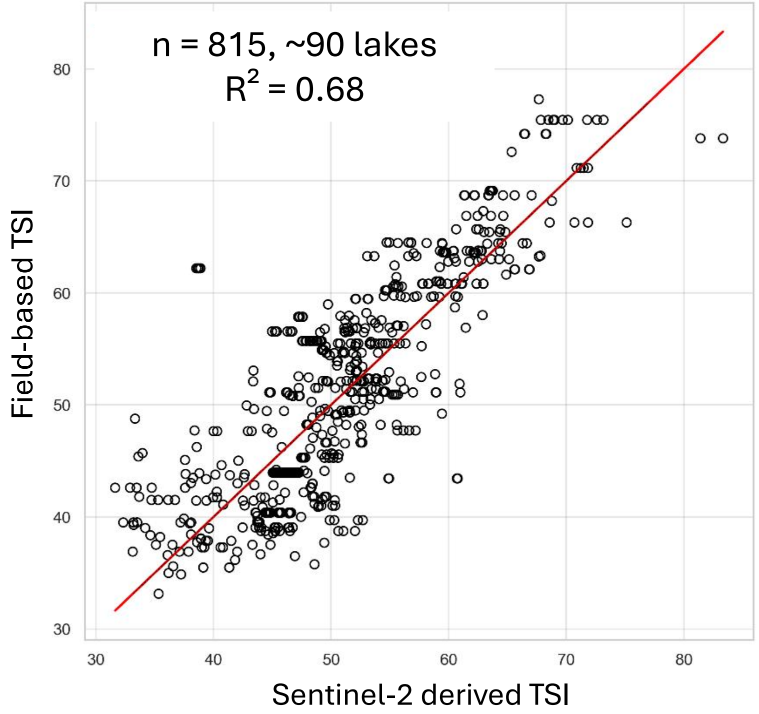


Trophic state model

TSI model based on field data

Field TSI based on ChlA, Secchi, TP, TN *(Carlson, 1977; Kratzed & Brezonik, 1981)*

Remote-sensing TSI based on ChlA and turbidity

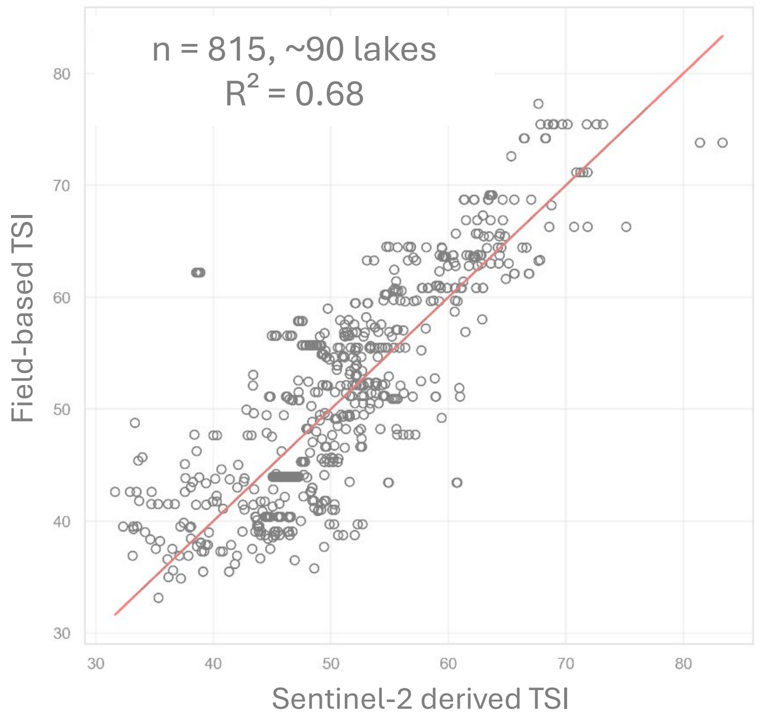


Trophic state evaluation

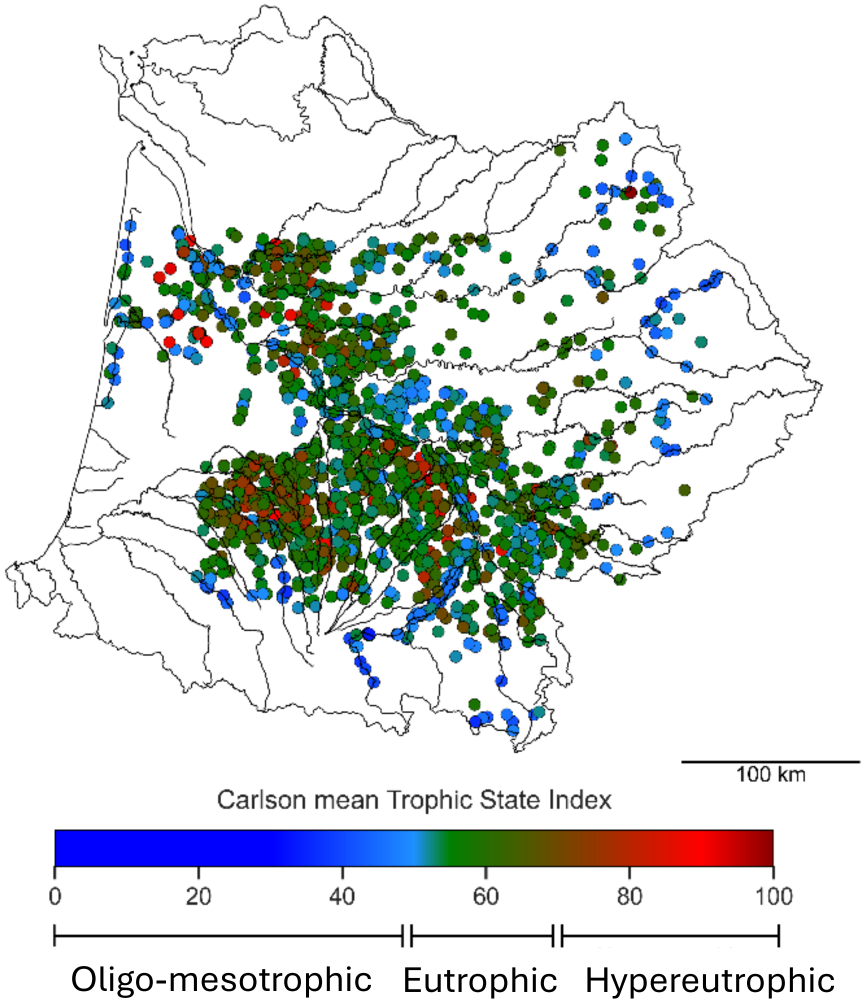
TSI model based on field data

Field TSI based on ChlA, Secchi, TP, TN *(Carlson, 1977; Kratzed & Brezonik, 1981)*

Remote-sensing TSI based on ChlA and turbidity



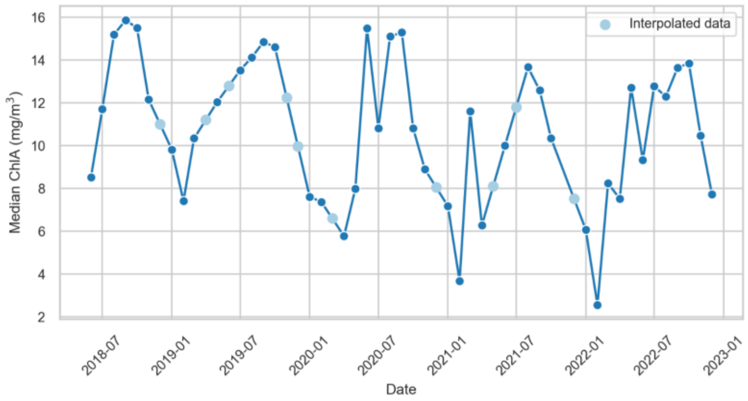
Application to Adour-Garonne lakes

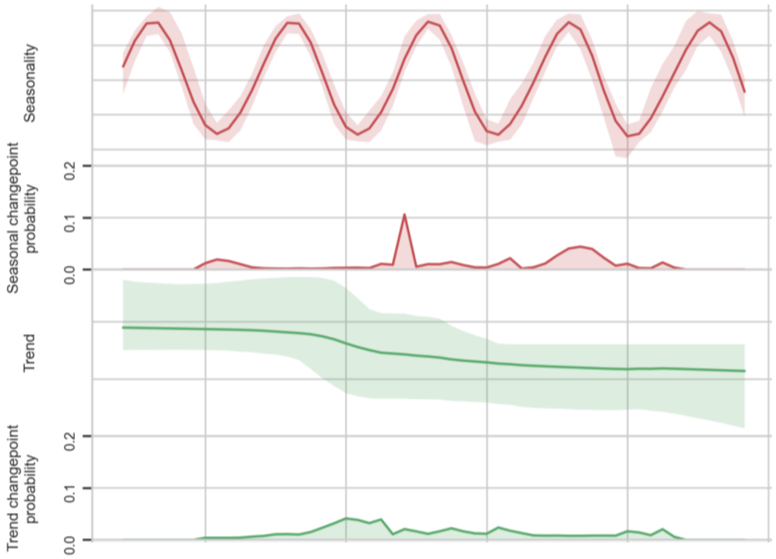


Trend decomposition

Bayesian trend decomposition

BEAST model *(Zhao et al. 2019)*



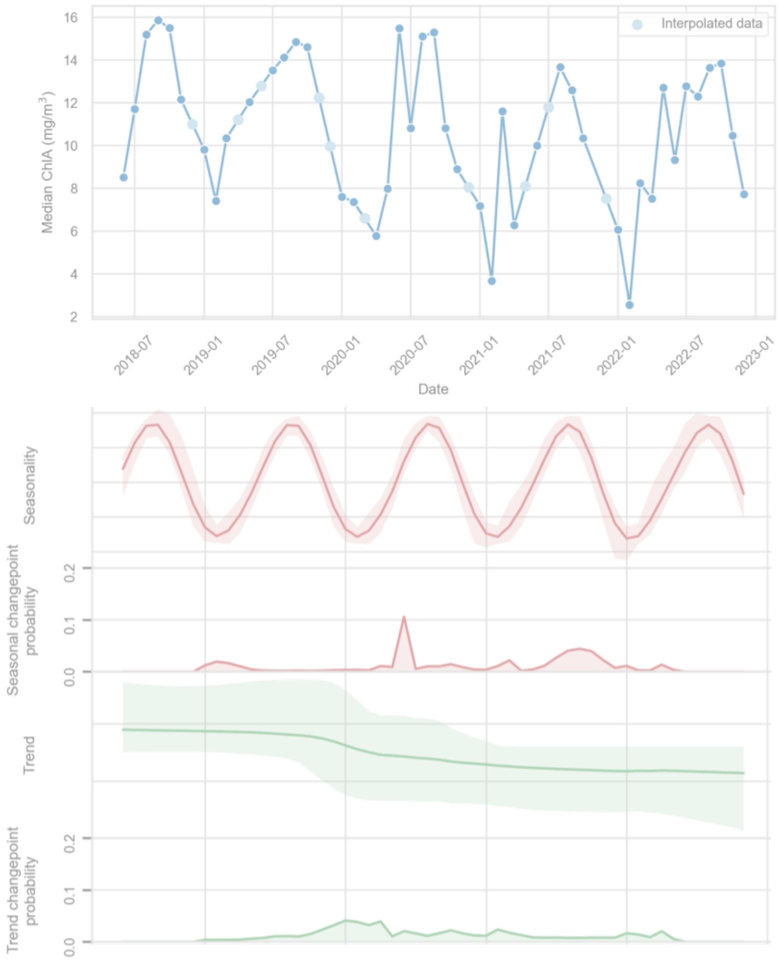


Trend analysis

XTREMQUALITY SCO SPACE FOR CLIMATE OBSERVATORY

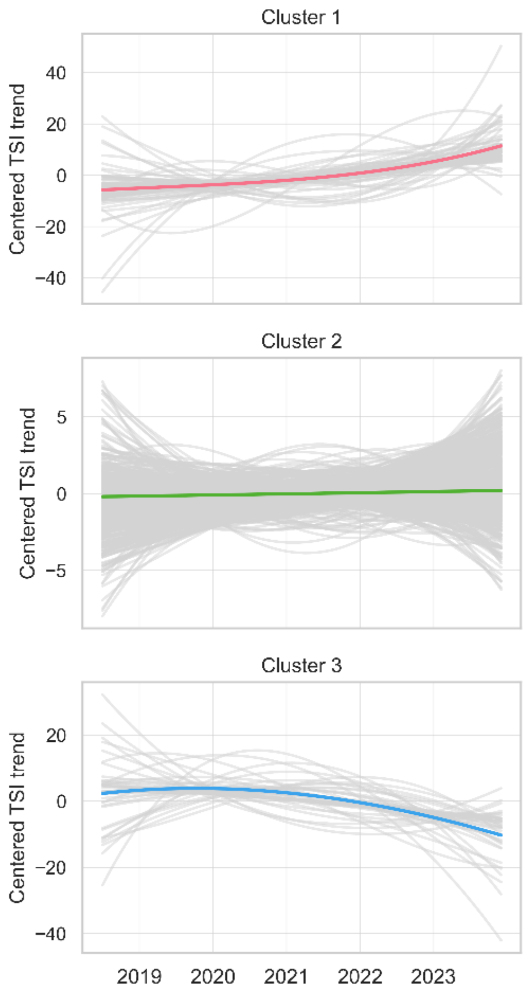
Bayesian trend decomposition

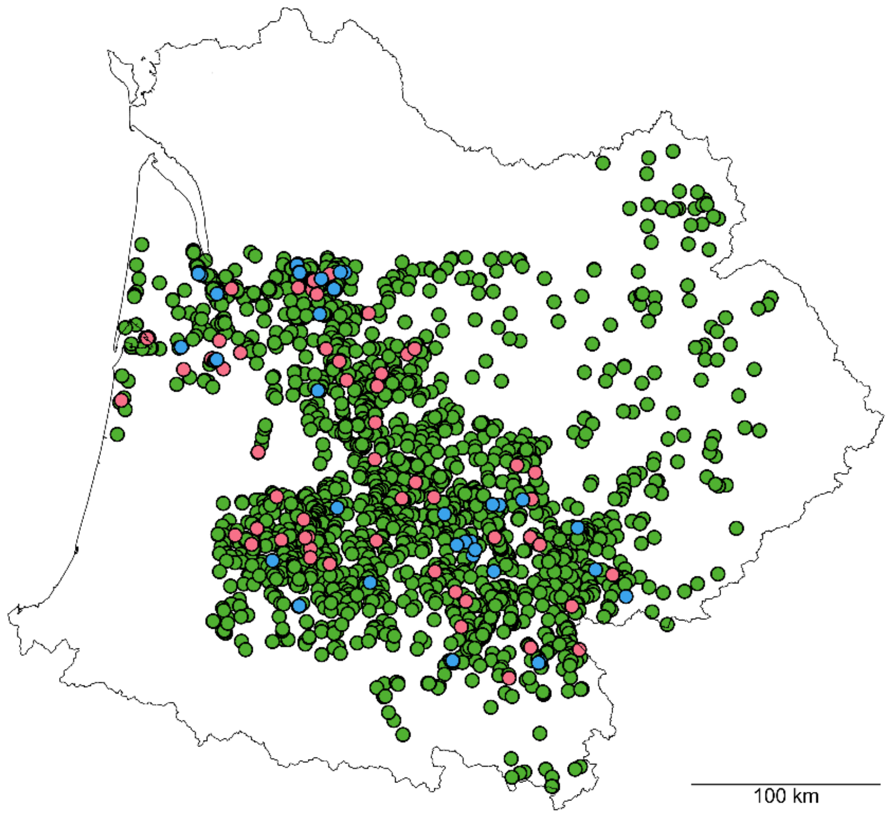
BEAST model *(Zhao et al. 2019)*



Medium-term trajectories classification

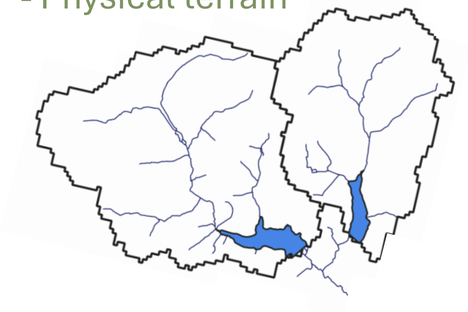
K-means clustering with DTW distance





Machine-learning models

- **Lake characteristics**
- **Draining areas characteristics**
  - Climatic indices
  - Meteorological indices
  - Land use and management
  - Physical terrain



Environmental variables

Quality data

River: NO_3^-/TP (in situ)

Lakes: ChlA/Turb (Sentinel-2)

Nutrient model

N/P inflows to the lakes

Lake model

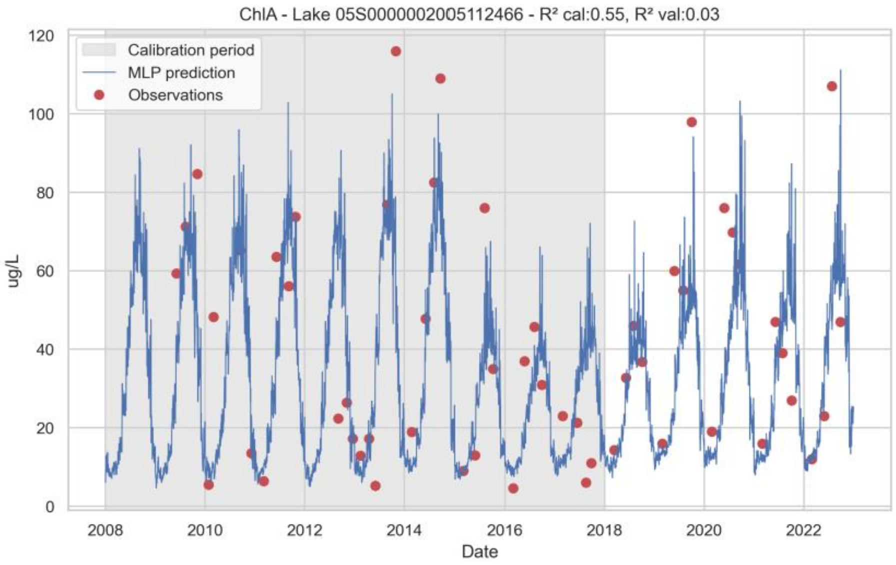
ChlA / Turb /TSI

Descriptor variables

Training variables

ML model (RF, XGBoost, MLP)

1) Improvement of time series predictions
2) Driving factors analysis



Conclusion

CONCLUSIONS

- **Global and complete framework for eutrophication assessment**
- **At a large scale and for lakes > 1 ha**

Average TSI: 58
80% lakes with TSI>50

Mostly stable,
↗49 lakes

On-going

LIMITS

- **Image processing** (adjacency, bottom effect, atmospheric correction)
- **Models limitation** (imputation accuracy, driving factors analysis)
- **Trophic state interpretation** (simplification, temporal scale, loss of spatial heterogeneity)

} Need of uncertainty indicators

Conclusion

XTREMQUALITY SCO SPACE FOR CLIMATE OBSERVATORY

ON-GOING and FUTURE WORK

- **Improving Sentinel-2 ChlA and turbidity data**
  - **New atmospheric correction algorithm**
  - **Extending time series**
- **Modeling**
  - **Machine-learning models (RF, XGBoost, MLP)**
  - **Coupling ML with physical models**
- **Results open-source by mid-2025**
- **Tests on other basins**

R&D RECOMMENDATIONS

→ Enhance monitoring of key *in-situ* variables (EO retrievable parameters and auxiliary predictors)

→ Work on uncertainty indicators for satellite-derived indicators

Thank you for your attention!

Contact information:
Mathilde Joffre ; mathilde.joffre@hetwa.fr