5th ECMWF-ESA Machine Learning Workshop
Bologna, 14 April 2026

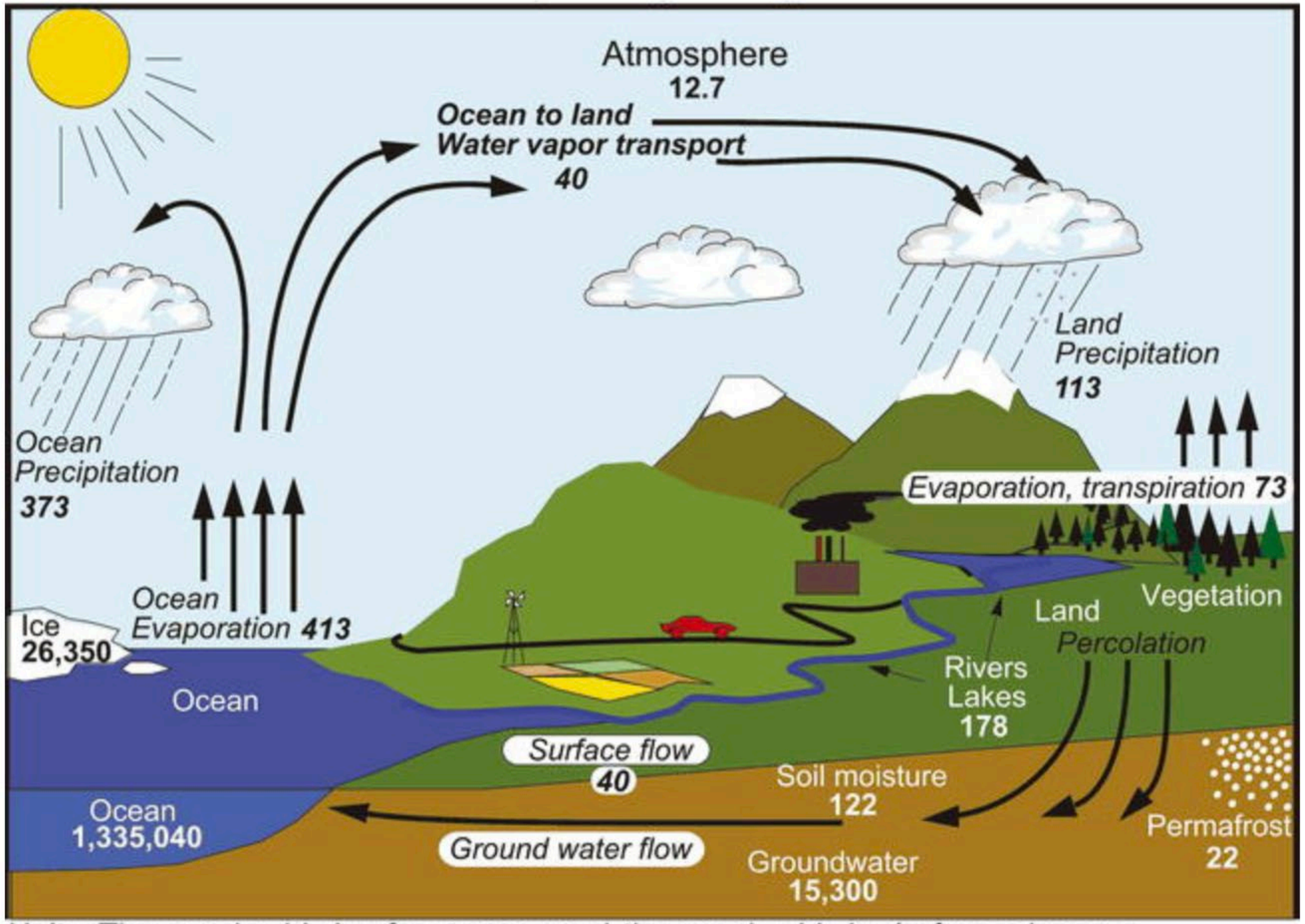
Enhanced land surface data exploitation using machine learning

Patricia de Rosnay, Eulalie Boucher, David Fairbairn, Sébastien Garrigues,
Christoph Herbert, Zdenko Heyvaert, Eleni Kalogeraki, Jana Kolassa, Tony McNally,
Ewan Pinnington, Nina Raoult, Kirsti Salonen, Pete Weston,
and many others

ECMWF

Earth system model

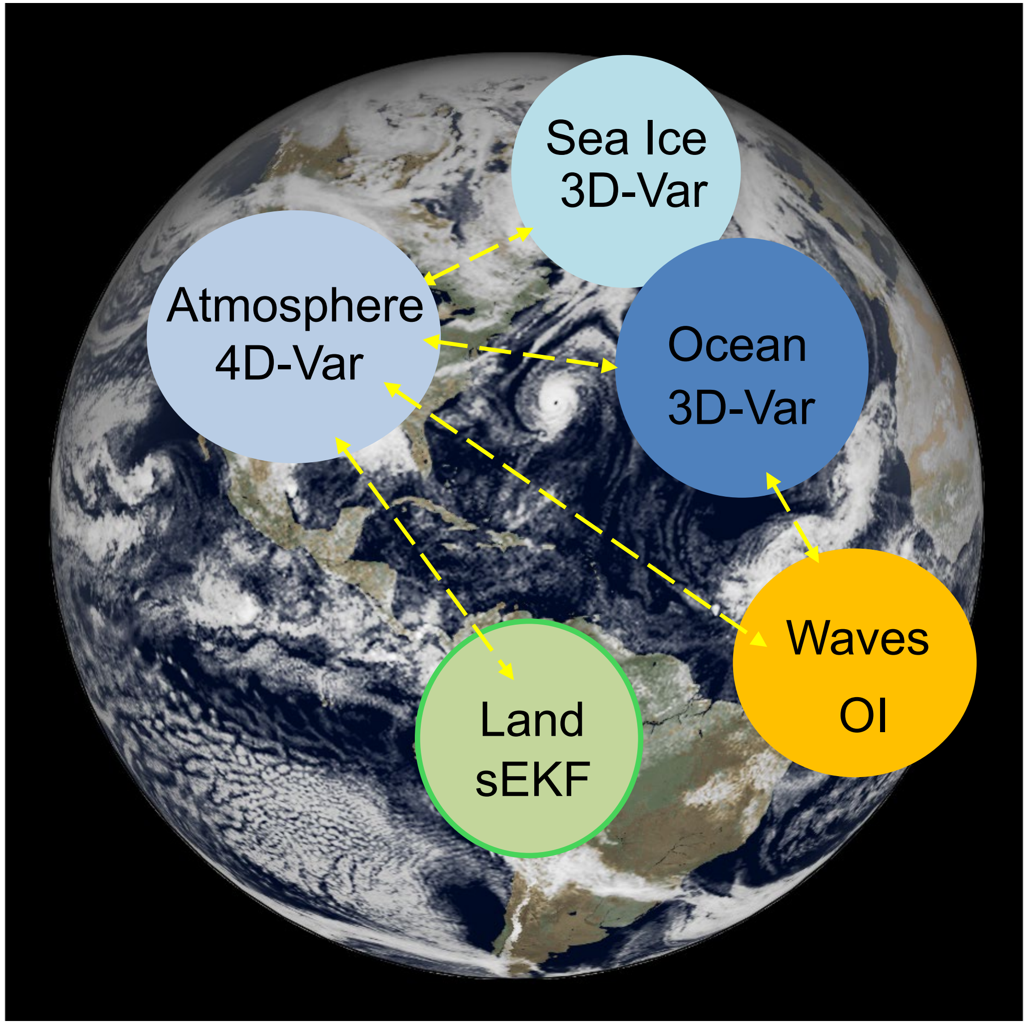
Image: Trenberth et al
J. Hydrometeorol. 2007



Units: Thousand cubic km for storage, and *thousand cubic km/yr* for exchanges

Earth system data assimilation

Integrated Forecasting system (IFS)



Over land: sEKF (simplified Extended Kalman Filter)

- de Rosnay P. et al QJRMS 2022 → Coupled data assimilation (DA) strategy
- Browne et al. ECMWF Spring Newsletter 2026 → ocean-atmosphere outer loop coupling
- Herbert et al. ECMWF Spring Newsletter 2026 → land-atmosphere outer loop coupling

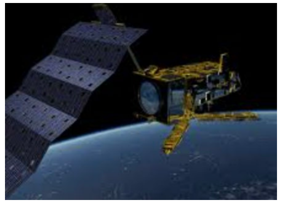
Coupled data assimilation

- Balanced initial conditions across the coupled forecast model components
- Opportunities to enhance the exploitation of surface sensitive satellite data
  → towards an "all surface" approach using interface observations
- Challenges related to complexity of land surfaces and related radiative transfer modelling

→ increasingly rely on ML approaches to link level 1 satellite data to land surface variables

Existing missions | Future missions

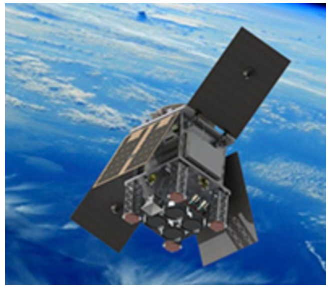
ASCAT



SMOS

Sentinel-5p

HydroGNSS



CRISTAL

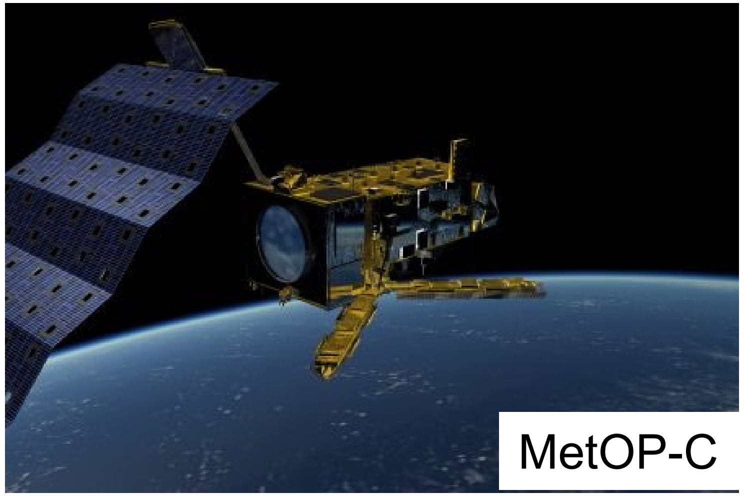
LSTM

CIMR

Soil moisture satellite observations used for NWP
(along with in situ T2m, RH2m screen level observations)

Active microwave data:

ASCAT: Advanced Scatterometer
MetOP-B (2012-), MetOP-C (2018-)
C-band (5.6GHz) **backscattering coefficient**
EUMETSAT Operational mission



Scatterometer soil moisture also used in ERA5 and ERA6 (ERS-SCAT, Metop/ASCAT)

Passive microwave data:

SMOS: Soil Moisture & Ocean Salinity (2009-)
L-band (1.4 GHz) **Brightness Temperature (TB)**
ESA Earth Explorer, dedicated soil moisture mission
(Kerr et al., 2016)

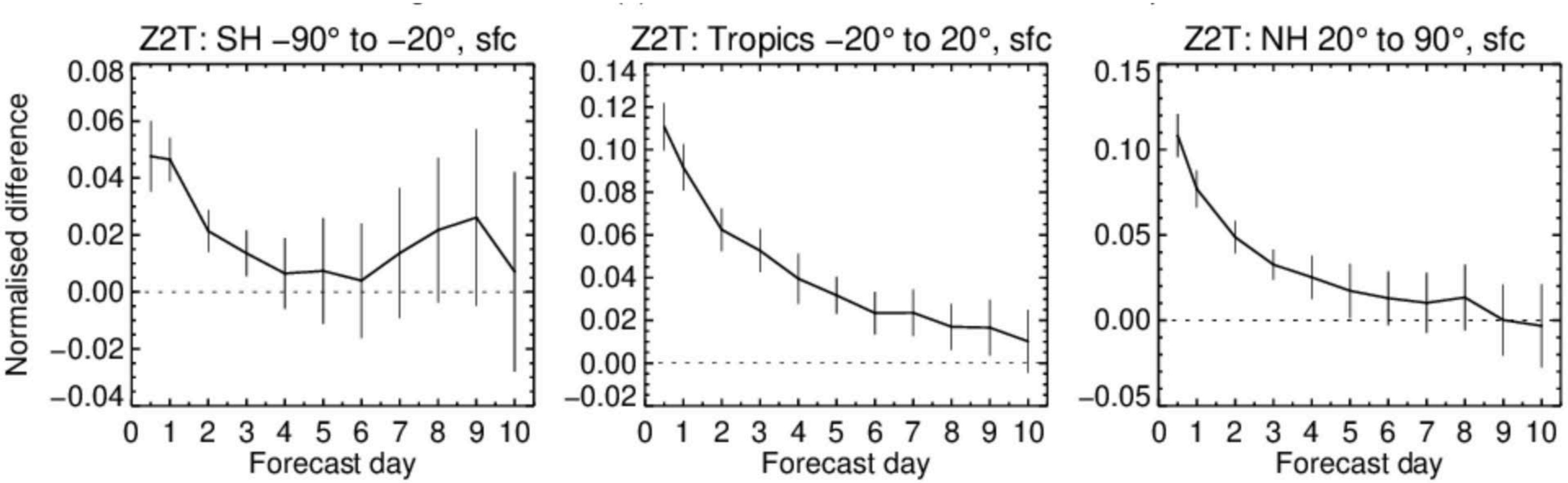
SMOS also used in ERA6

SMAP (monitoring)
L-band TB 2015-
NASA Dedicated
soil moisture mission

Unified soil analysis impact on the atmospheric forecast

T2m RMSE

June-July 2024
IFS cycle 50r1

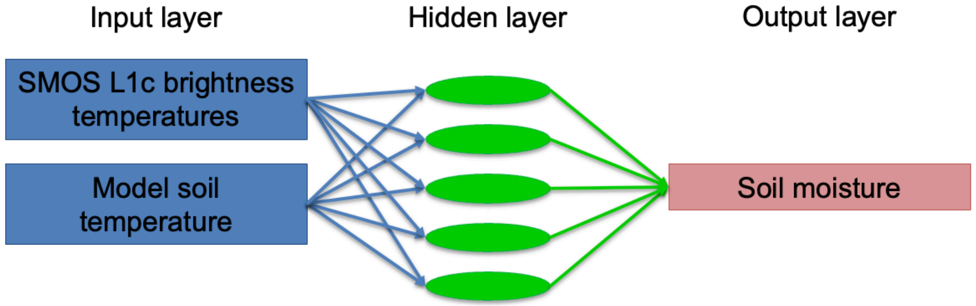


Degradation when soil analysis is off

→ **Significant positive impact of the soil analysis on low level atmospheric temperature forecasts**

The CERISE project (grant agreement No101082139) is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

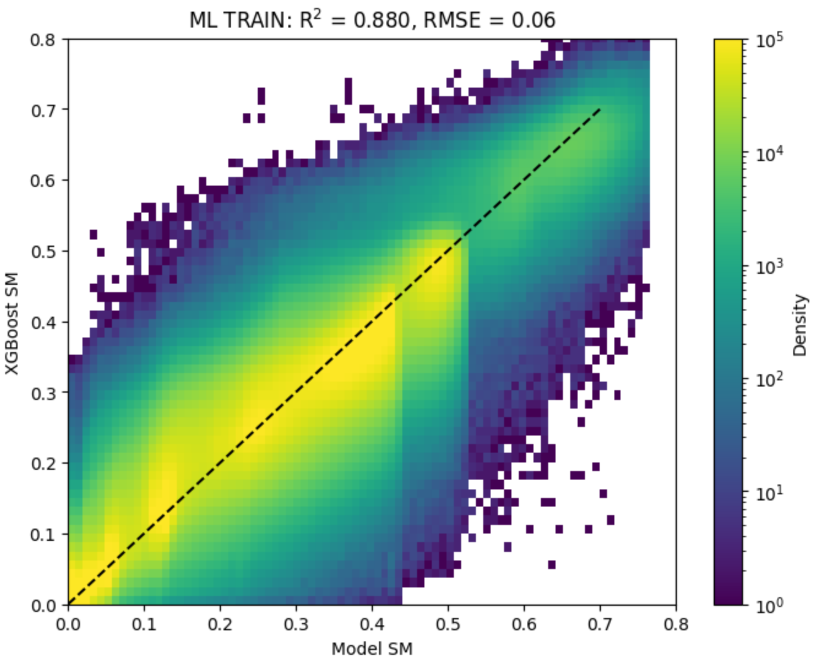
SMOS soil moisture retrieval for data assimilation

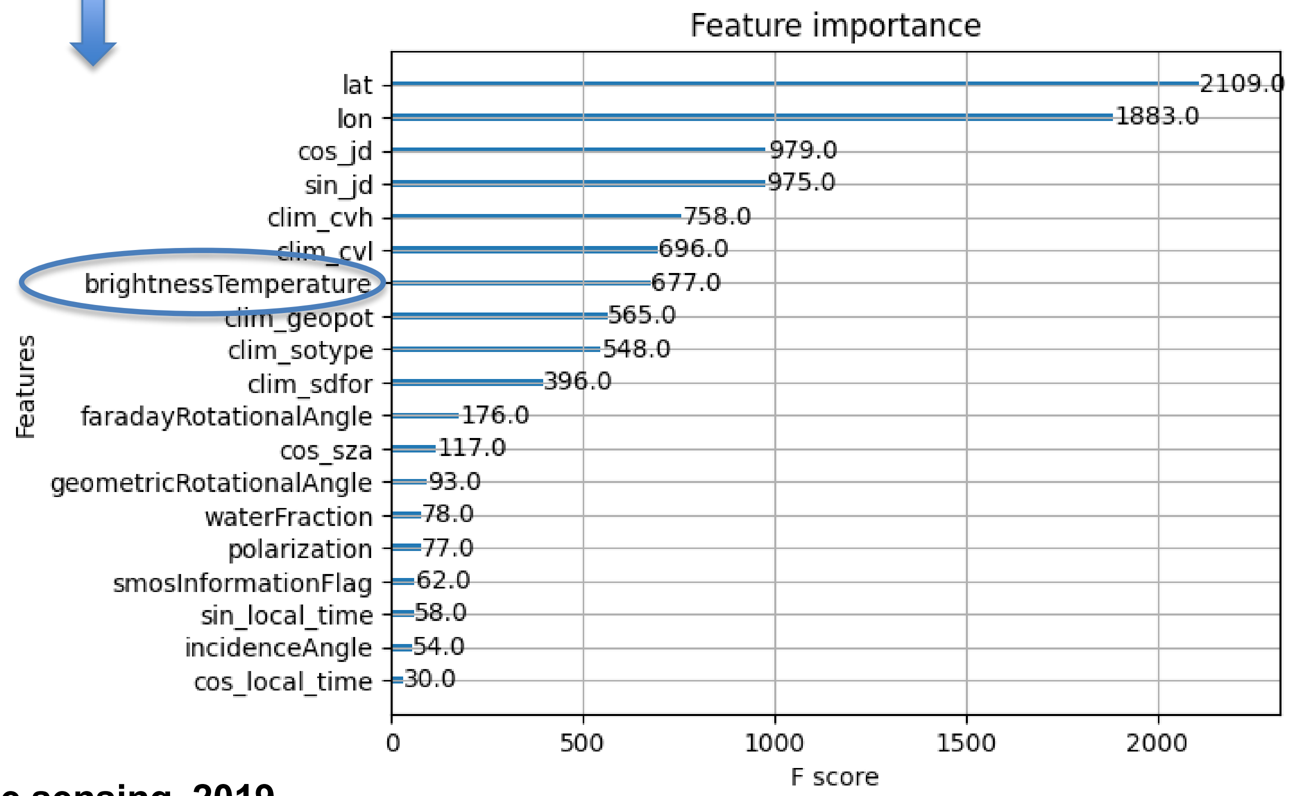


Used for operational NWP since 2019

using a Neural Network (2019-2025)

And XGBoost (since Feb 2025)





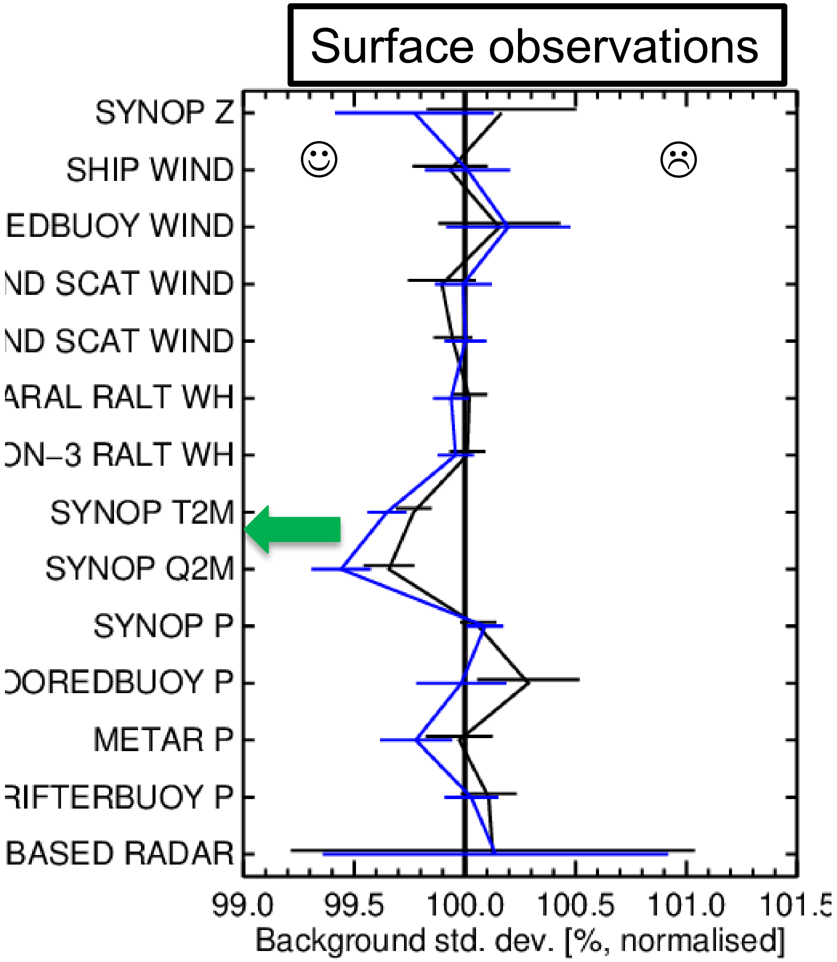
ECMWF

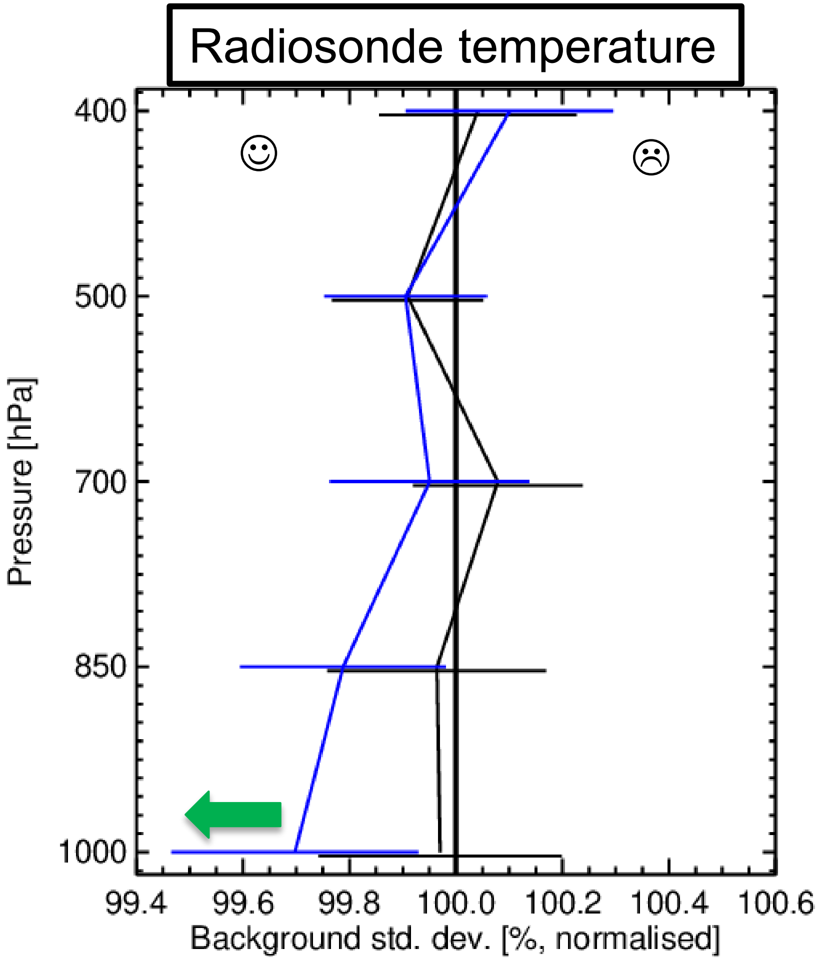
Rodriguez-Fernandez et al, Remote sensing, 2019
Salonen et al., in prep 2026

SMOS data assimilation impact

- Comparing tree-based method (XGBoost) and Neural Network (Rodriguez-Fernandez, RS 2019) SMOS DA
- **SMOS soil moisture DA has positive near surface impact when compared to no SMOS DA (100% line).**
- The positive impact of SMOS DA is stronger with **XGBoost** than **NN** SMOS soil moisture in 49r1.

SMOS XGBoost DA
SMOS NN DA





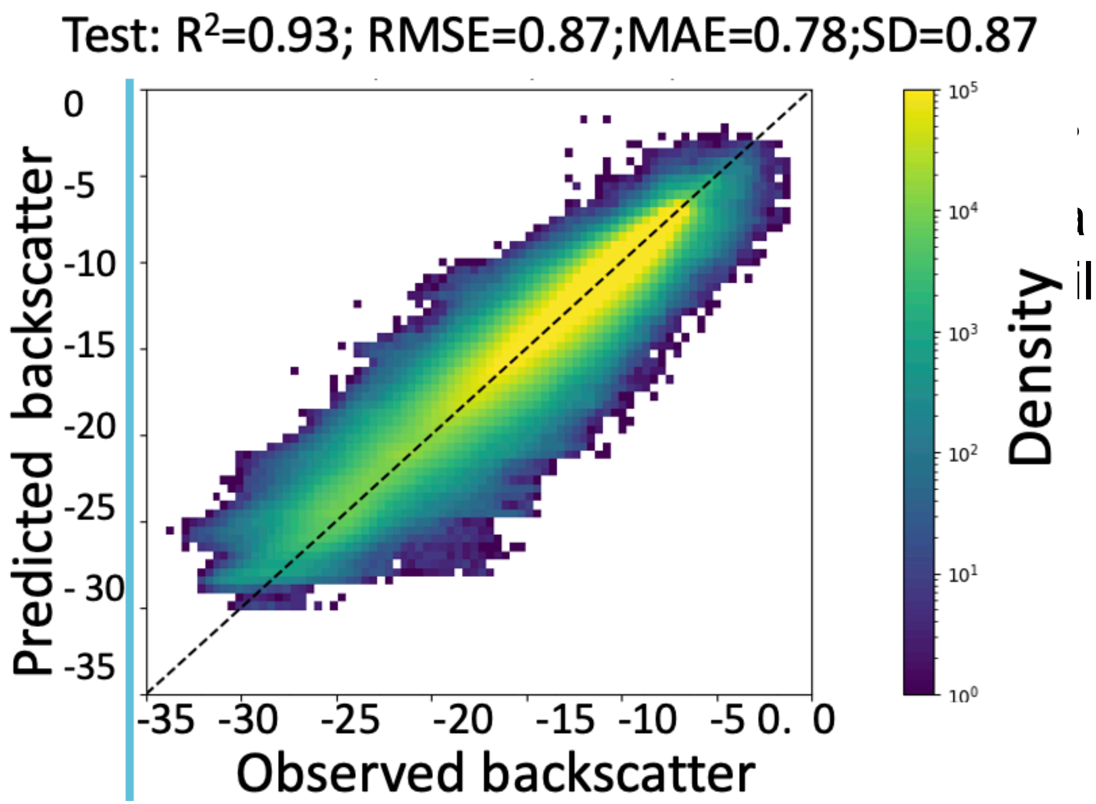
ECMWF

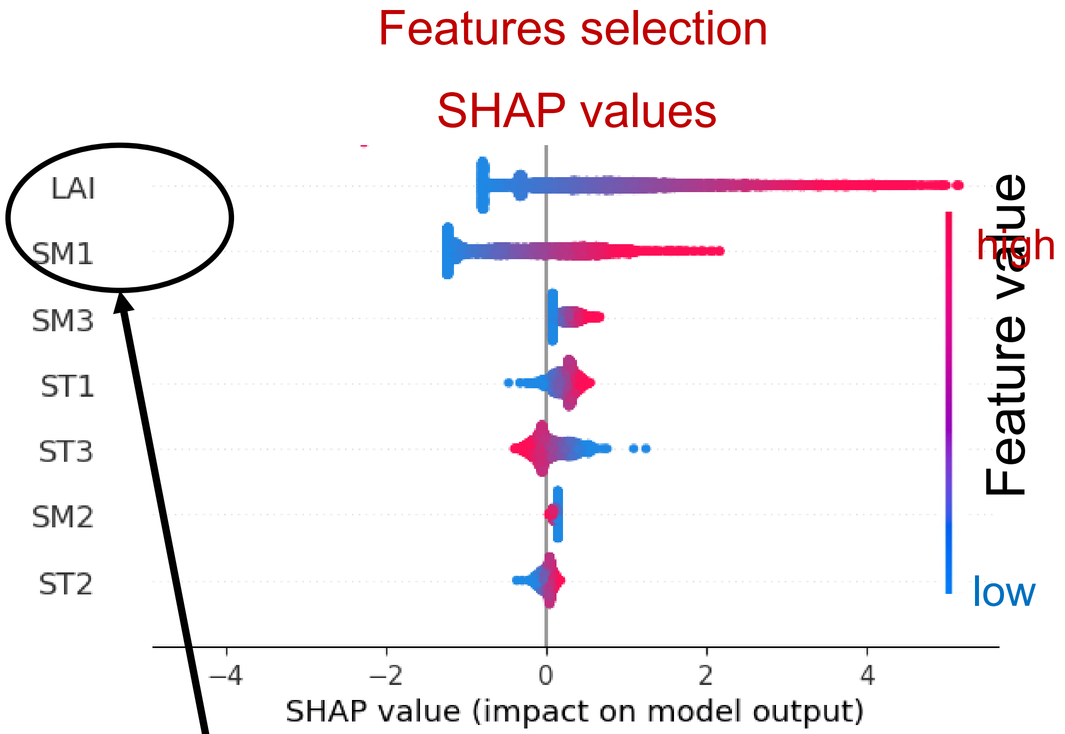
Microwave (MW) observation operators

Enhance the exploitation of satellite observations in coupled land-atmosphere assimilation to constrain vegetation water and carbon cycle variables.

→ Development of Machine Learning (ML)-based observation operators. Example for ASCAT:

Funded by the European Union





Vegetation (LAI) and surface soil moisture (SM1) are the most influent variables

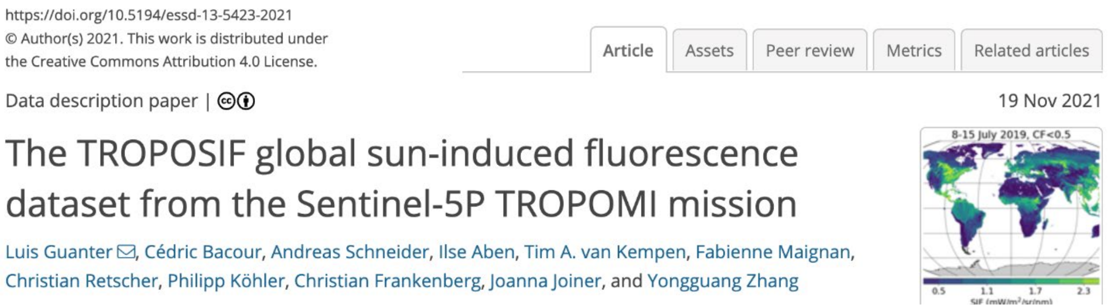
IFS DA experiments show that ASCAT Backscatter DA improves surface variables compared to ASCAT SM DA

ECMWF

→ Poster from Garrigues et al.

LAI analysis using Solar Induced Fluorescence (SIF)

CORSO
CO2MVS Research on Supplementary Observations


https://doi.org/10.5194/essd-13-5423-2021

Article | Assets | Peer review | Metrics | Related articles

Data description paper | 19 Nov 2021

The TROPOSIF global sun-induced fluorescence dataset from the Sentinel-5P TROPOMI mission

Luis Guanter, Cédric Bacour, Andreas Schneider, Ilse Aben, Tim A. van Kempen, Fabienne Maignan, Christian Retscher, Philipp Köhler, Christian Frankenberg, Joanna Joiner, and Yongguang Zhang

Funded by the European Union

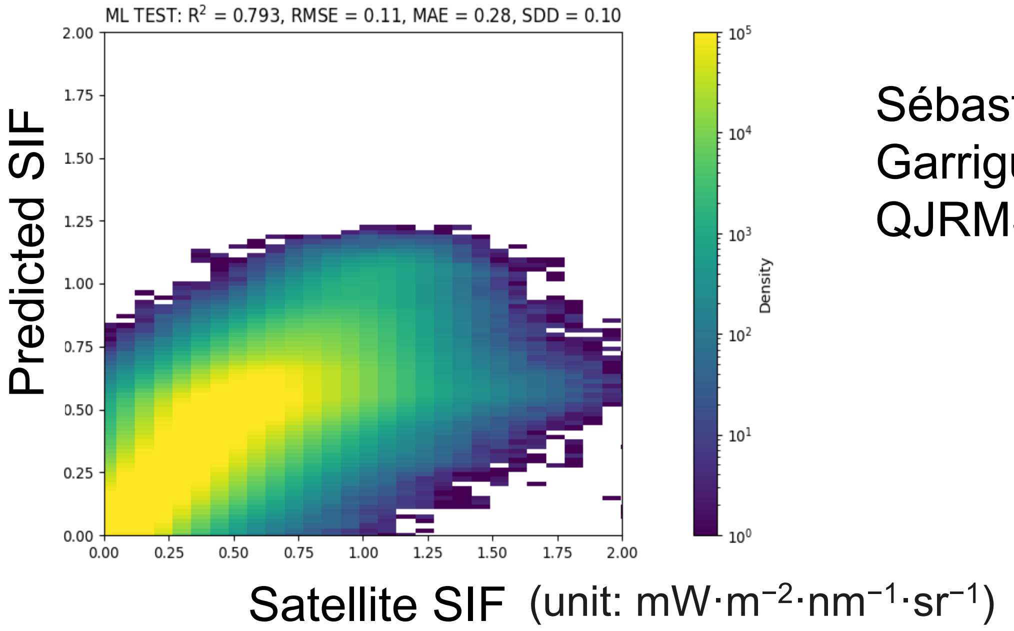
SIF: electromagnetic signal emitted by the chlorophyll of assimilating plants

→ part of the energy absorbed by chlorophyll is not used for photosynthesis but emitted at longer wavelengths in the near infrared

→ Relevant to analyse vegetation LAI and Gross Primary Production (GPP)

TROPOMI SIF 743-758 nm

Exploratory work to use SIF at ECMWF.
Observation operator development



Sébastien Garrigues et al., QJRMS 2026

ECMWF

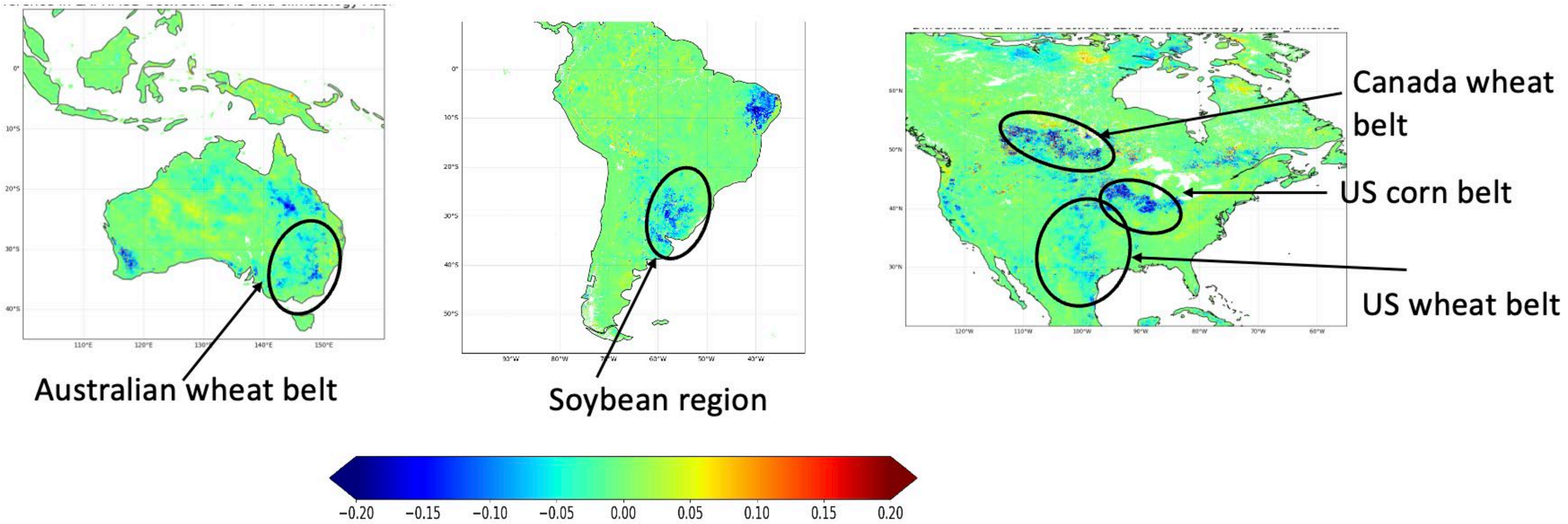
LAI analysis using Solar Induced Fluorescence (SIF)

CORSO
CO2MVS Research on Supplementary Observations

Funded by the European Union

Impact of SIF DA shown as LAI RMSE differences with vs without SIF data assimilation against satellite (Copernicus Land) LAI for 2022

Improvement for cropland



LAI RMSE differences

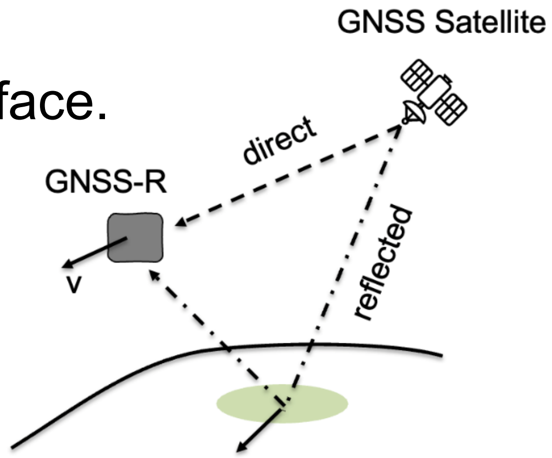
☺ ☹

ECMWF

The CORSO project (grant agreement No101082194) is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

Global Navigation Satellite System - Reflectometry (GNSS-R)

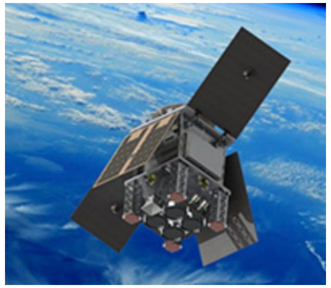
Concept:
GNSS-R gets information from the signal reflected at the surface.
Wet soils reflect more strongly than dry soils



Strengths:

- Uses existing GPS signals → opportunity information
- Frequent observations (Daily and Sub-daily)
- Potentially high spatial resolution
- Low cost compared to other dedicated SM missions

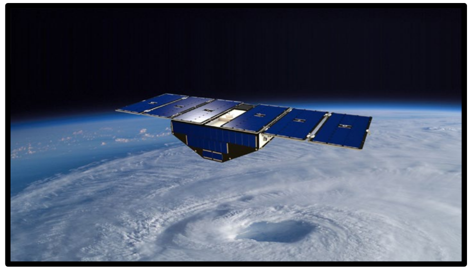
HydroGNSS
Launched Nov 2025



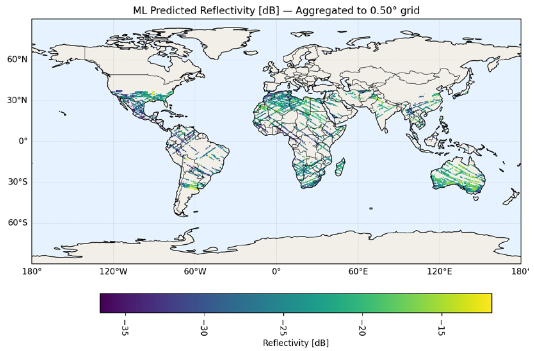
Using CYGNSS (Cyclone GNSS) in preparation of hydroGNSS

Eleni Kalogeraki
Jana Kolassa

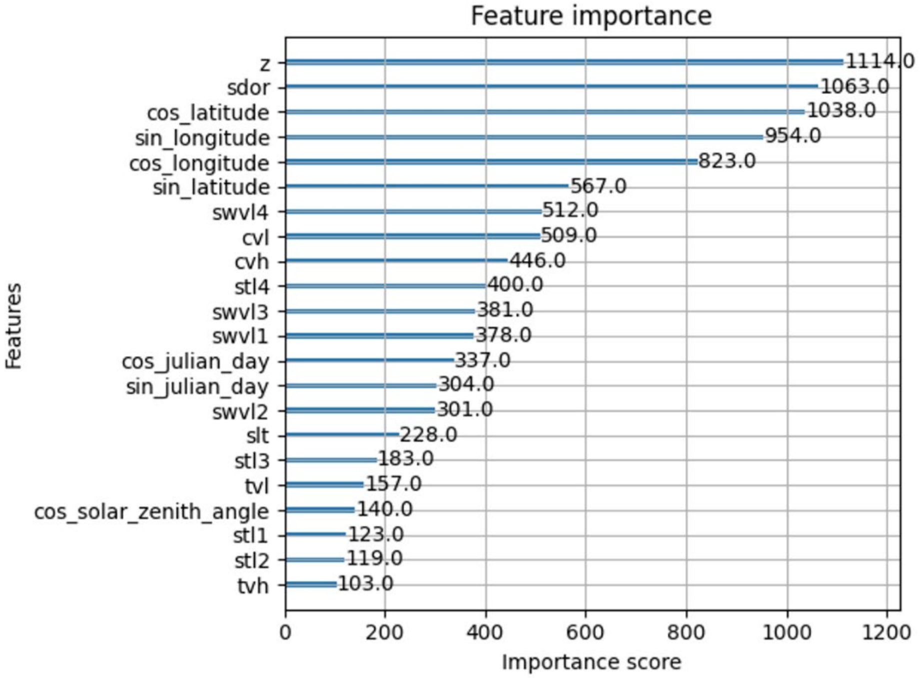
GNSS-R ML-based observation operator



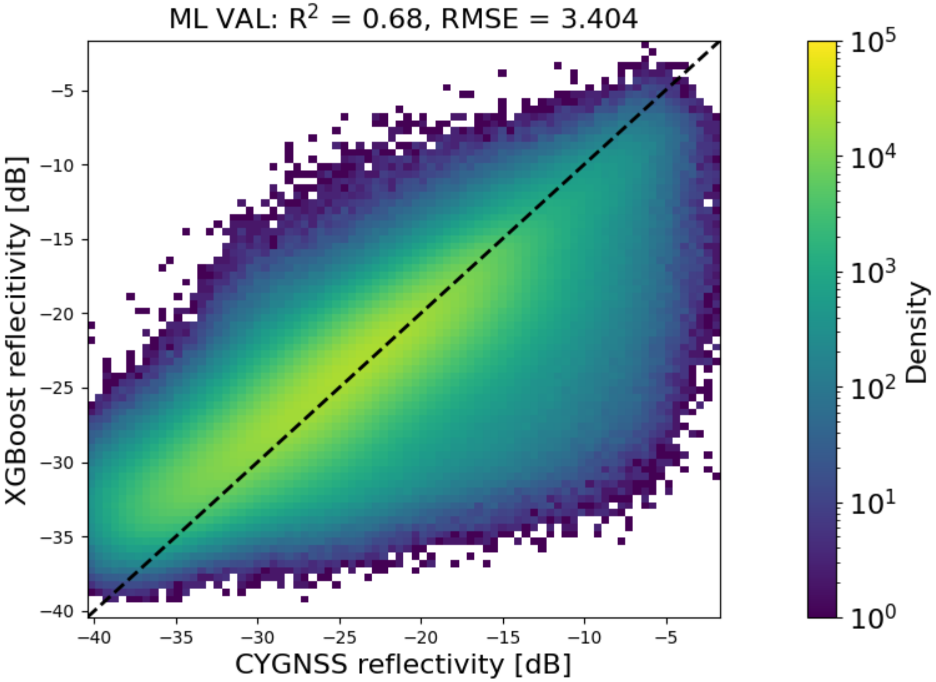
CYGNSS
Constellation of 8 satellites
Orbit between ±38° latitude



Training data ~36 million samples (across 2022/2023)
Validation data ~9 million samples
(20% retention, 80% rejection)



Importance of orography, geolocation, vegetation, soil moisture and temperature



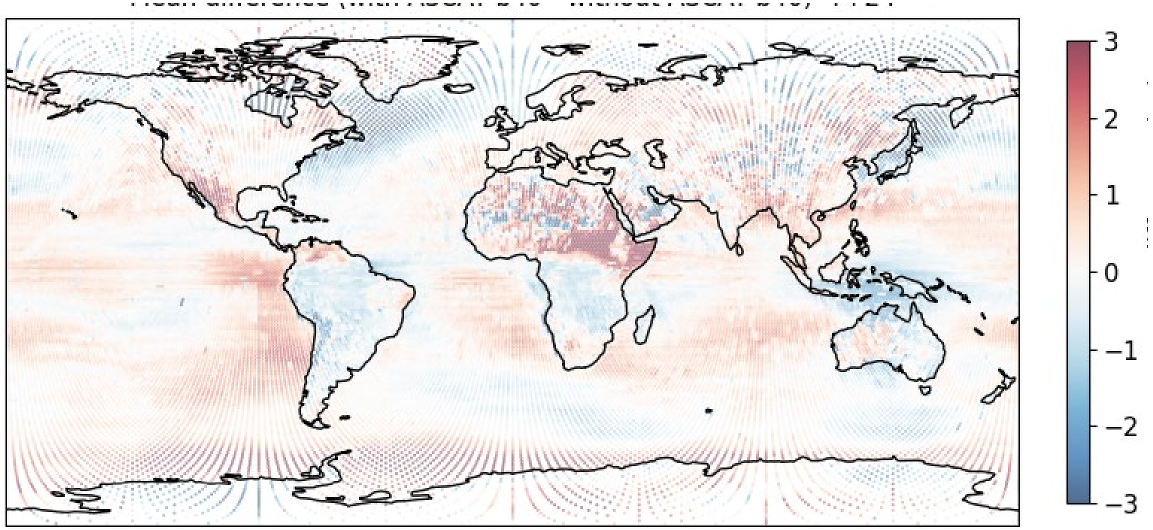
→ Ongoing implementation in the ECMWF Land Data Assimilation System

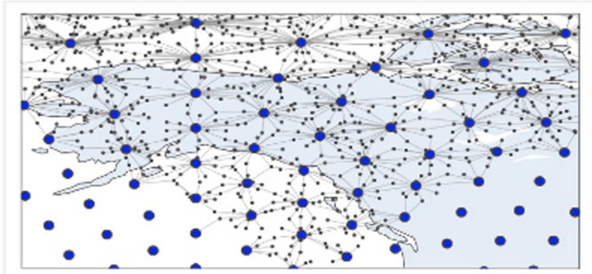
Land surface in data driven machine learning forecasts models

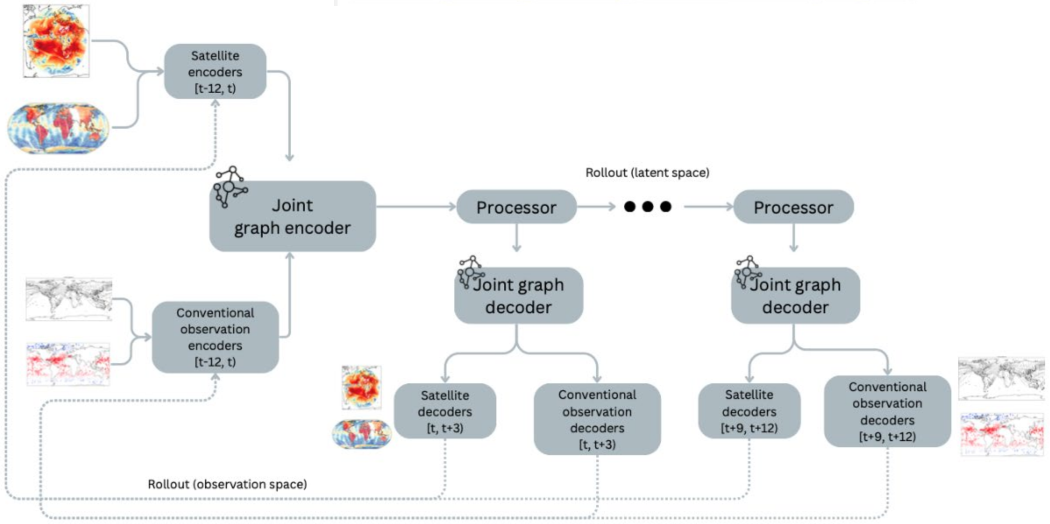
- AIFS, trained on ERA5 and IFS analyses: Adding land surface variables: currently Soil Moisture, temperature, runoff (Moldovan et al., 2026), next version (v2) with snow (Raoult et al. ECMWF NL#186, 2026) → Nina Raoult

- AI-DOP (Direct Observation Prediction → Eulalie Boucher), using observations to learn a physical model of the Earth system: Adding satellite data sensitive to land surfaces → Sébastien Garrigues

Difference in T2m (K) with vs without ASCAT backscatter 40º
In GraphDOP







ECMWF

GrapDOP: Alexe et al., 2024, Boucher et al., 2025

Summary

- Strong and significant impact of land data assimilation analyses on NWP
- Coupled data assimilation → possibilities to enhance the exploitation of surface observations
- Transition towards using raw observations at level 1 to simultaneously analyse soil, vegetation and atmospheric variables
- Machine-learning forward models allow to enhance exploitation of current satellites (e.g. ASCAT), as well as passive microwave data (Pete Weston's presentation), and to explore new land surface observations (e.g. SIF, GNSS-R),
- Integration of land surface variables in AIFS
- Potential of surface observations into end-to-end Machine learning systems such as Graph-DOP
  → unprecedented capability to exploit surface sensitive observations for weather prediction & climate reanalyses

2026 Workshop on Machine Learning for Land and Hydrology

3-5 November

Reading, UK

Prof Pierre Gentine

Columbia University

Rotem Mayo

Google

More speakers to be confirmed!

ECMWF