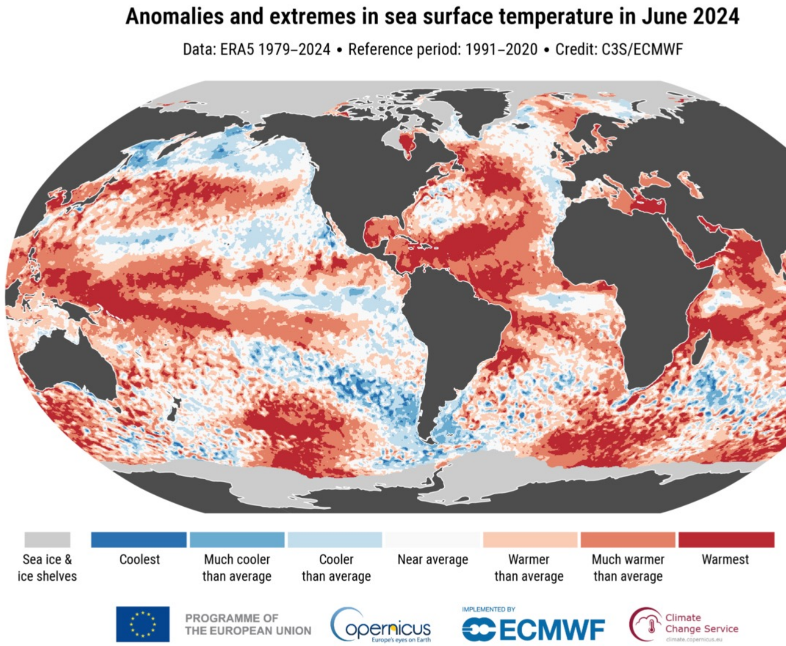
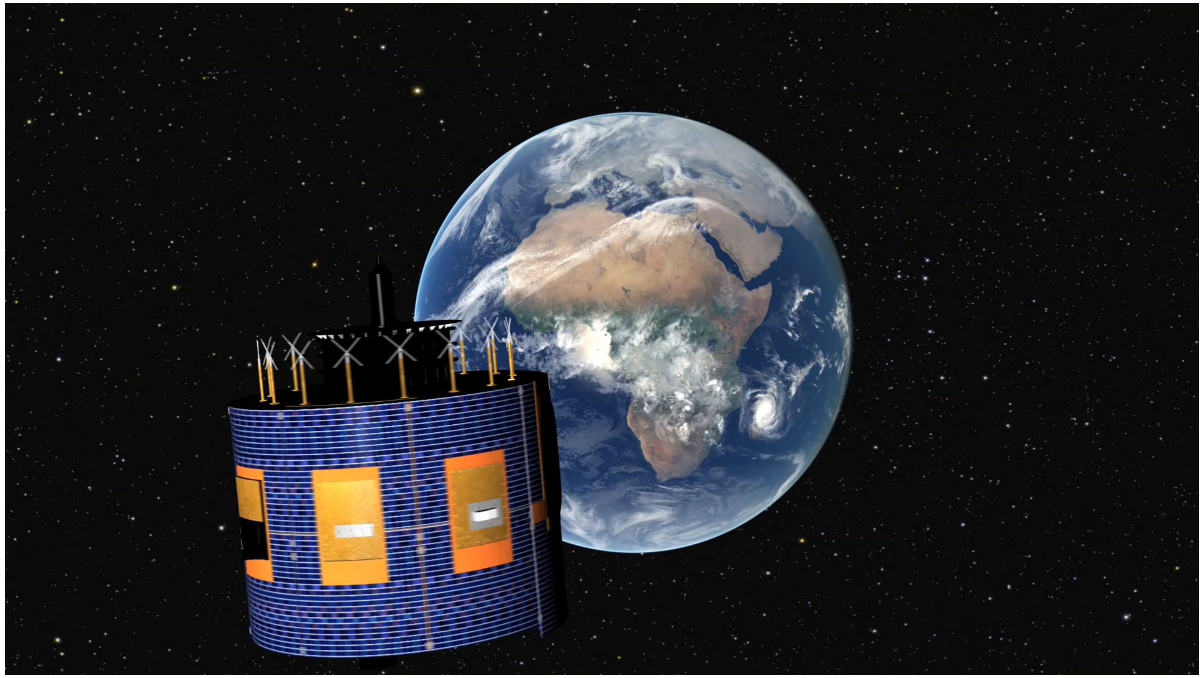
Exploring the relationships between satellite-observed cloud albedo and ERA5 variables to improve data driven short-term irradiance forecast

Nicolas Chea, Sylvain Cros, Arttu Tuomiranta, Jordi Badosa, Amar Meddahi, Martial Haeffelin, and Sébastien Guillon





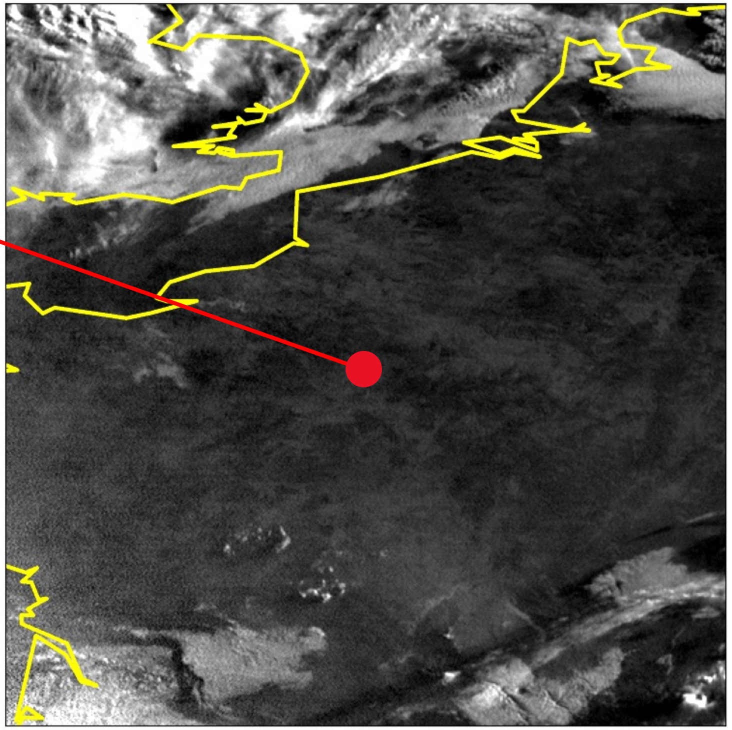
nicolas.chea@polytechnique.edu

September 6th, 2024

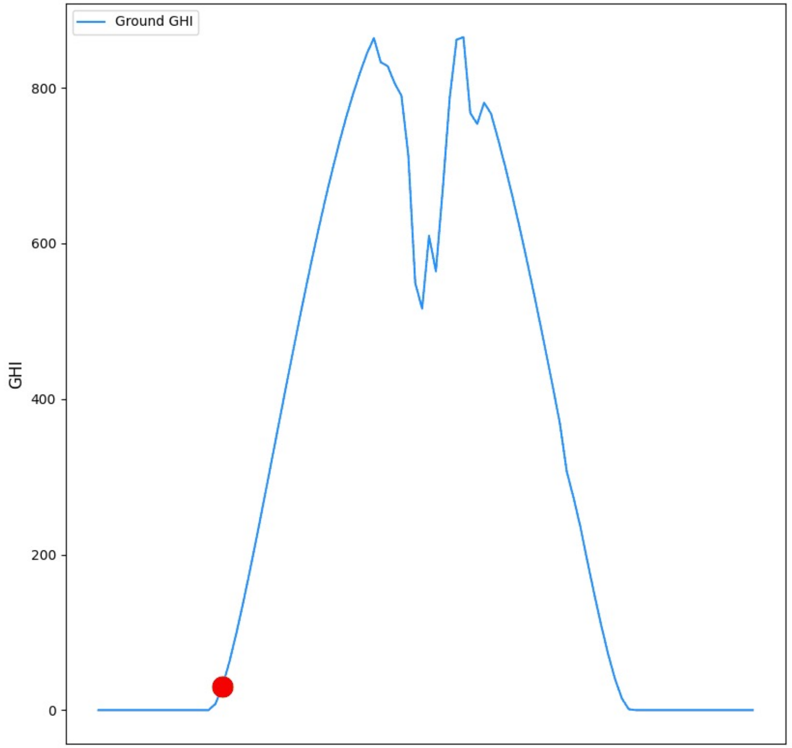
Nicolas Chea, PhD student at Laboratoire de Météorologie Dynamique at École polytechnique in Palaiseau, France (25km from Paris)

Thesis: « Short-term Global Horizontal Irradiance (GHI) forecasting using deep learning techniques applied on meteorological satellite images »

SIRTA



Satellite images



Ground GHI

Motivation

Limitation: satellite data only gives information on the top cloud layer

- Is this sufficient to **anticipate cloud formation and dissipation?**
- Can **meteorological variables** help improve the forecasts?

Can **ERA5 variables** help the forecast reach better performance by including more information on **atmospherical processes** and on the **vertical profile of cloud cover**?

Objectives

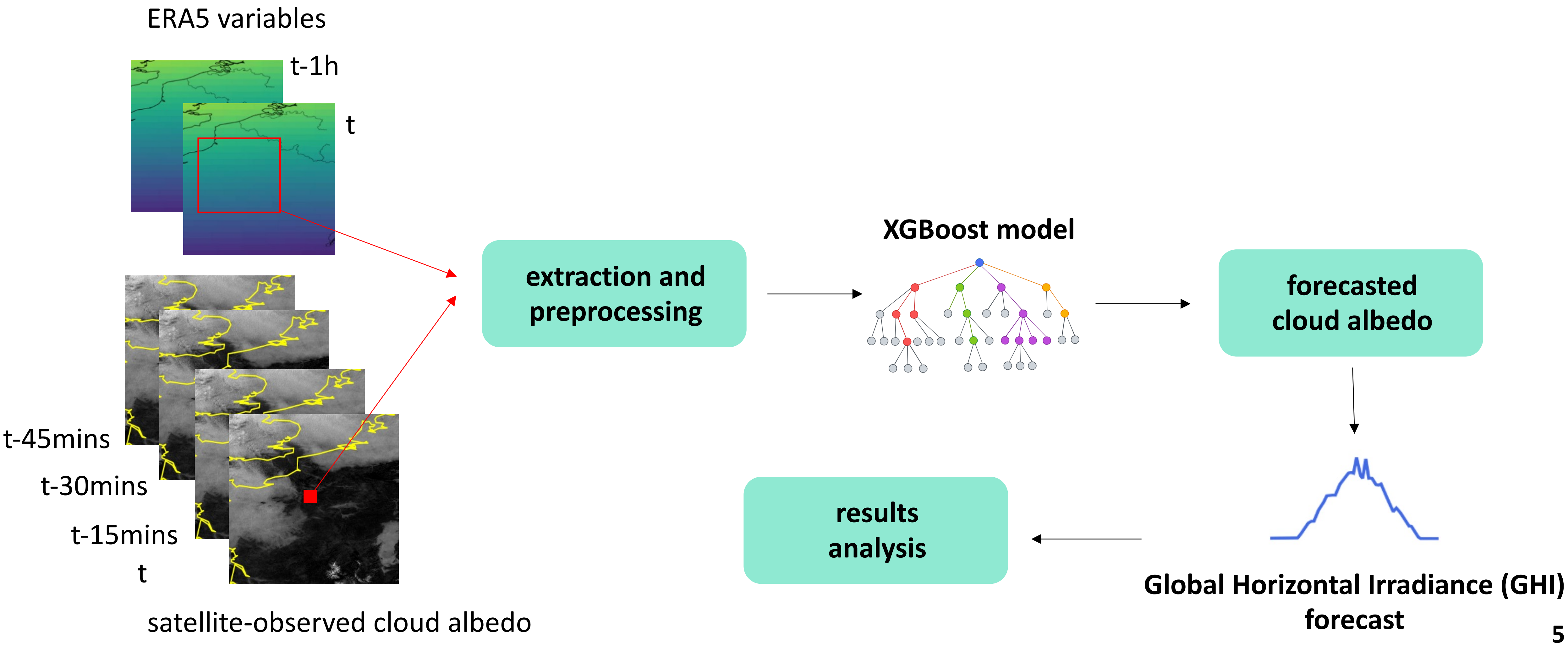
Assess the forecast performance of **satellite-observed cloud** albedo using **ERA5 reanalysis**

Train a model with and without ERA5 variables, and assess the difference in performance

Identify the most relevant features to integrate into a **deep learning forecasting framework**

Determine the key variables the model relies on for generating forecasts

Methodology

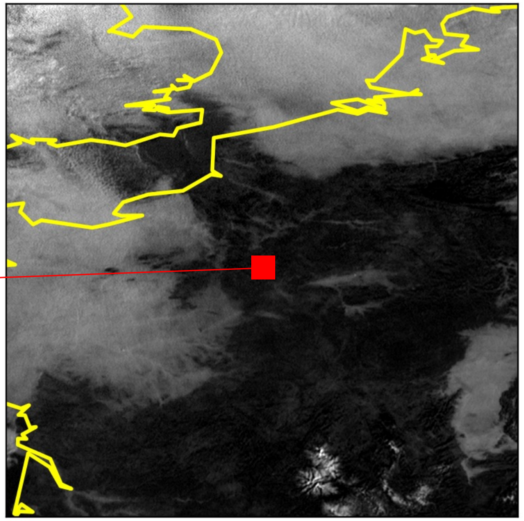


Satellite-observed cloud albedo

Cloud albedo (Heliosat principles, Müller et al., 2012) from **Meteosat Second Generation (MSG)**, High Resolution Visible (HRV) channel

EUMETSAT

SIRTA observatory
48.713° N, 2.208° E



cloud albedo
(value between 0 and 1)

Resolution

Temporal resolution	15 minutes
Spatial resolution	~2*2km

When sun elevation > 5°

ERA5 reanalysis variables

ECMWF Copernicus

Cloud cover state	
Relative humidity	Water vapour pressure as a percentage of the value at which the air becomes saturated
Temperature	Temperature in the atmosphere
Fraction of cloud cover	Proportion of a grid box covered by cloud and varies between zero and one

Horizontal cloud motion	
U wind component	Horizontal speed of air moving towards the east
V wind component	Horizontal speed of air moving towards the north

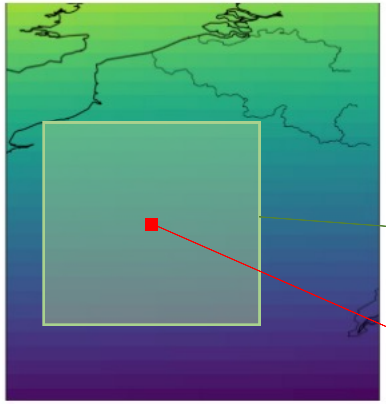
Cloud formation and dissipation	
Convective inhibition	Measure of the amount of energy required for convection to commence
K index	Measure of the potential for a thunderstorm to develop, calculated from the temperature and dew point temperature in the lower part of the atmosphere
Convective Available Potential Energy (CAPE)	Indication of the instability (or stability) of the atmosphere and can be used to assess the potential for the development of convection
Vertical velocity	Speed of air motion in the upward or downward direction
Divergence	Horizontal divergence of velocity. It is the rate at which air is spreading out horizontally from a point

Resolution

Temporal resolution	1 hour
Spatial resolution	0.25° (~25*25km)

ERA5 data preprocessing

Horizontal preprocessing



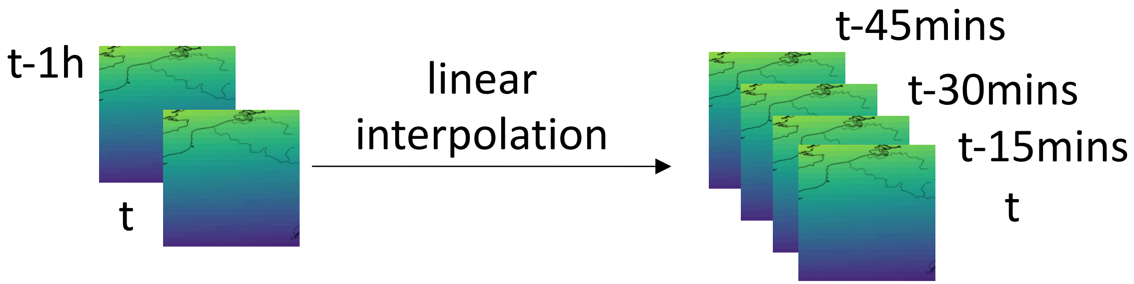
Grid points centered on **SIRTA observatory**
48.713°, 2.208°

For each variable:

mean, std of a 5*5 grid points (1.25°) centered on SIRTA

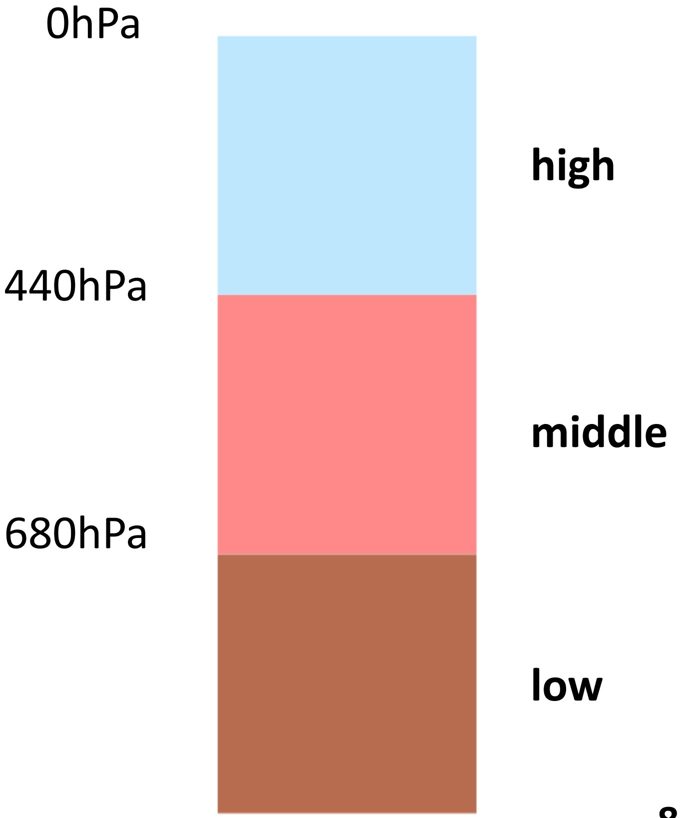
extraction of the closest grid point

Temporal preprocessing



Pressure level layers aggregation

Aggregation in **3 standard atmospheric layers** with pressure thresholds of **440 and 680 hPa** (Luo, 2023)



Forecasting model: XGBoost

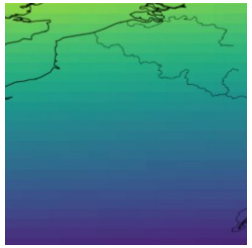
- **Tree-based gradient boosting model**: can capture non-linear relationships between cloud albedo and ERA5 variables
- **Feature importance**: identifies key variables through feature importance (gain-based)
- **Performance:** good tradeoff between performance and fast iteration

train / validation / test split

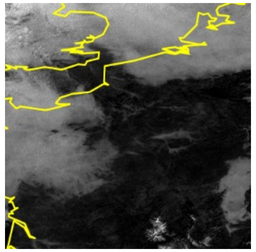
Train/validation	2021-2022
Test	2023

Cross-validation is performed to identify the optimal set of hyperparameters

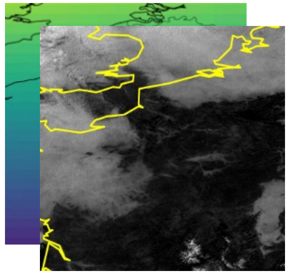
3 model configurations



ERA5 only



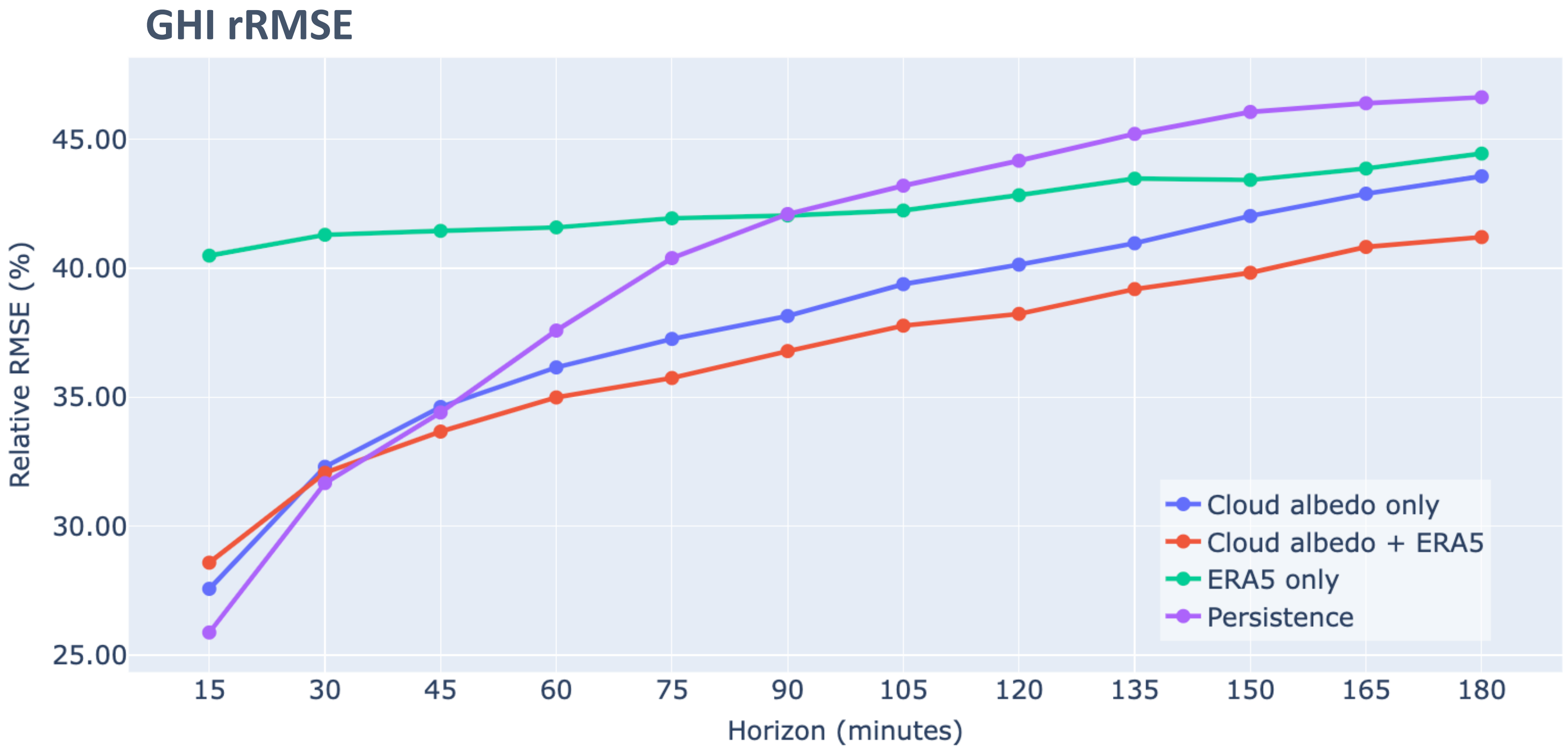
Cloud albedo only



Cloud albedo + ERA5

1 model by configuration and by time horizon from **t+15mins to t+180mins** (15 mins step) to **forecast satellite-observed cloud albedo**

Global Horizontal Irradiance (GHI) relative RMSE (rRMSE) model comparison



Target	Single pixel: SIRTA's location
Dates	2023

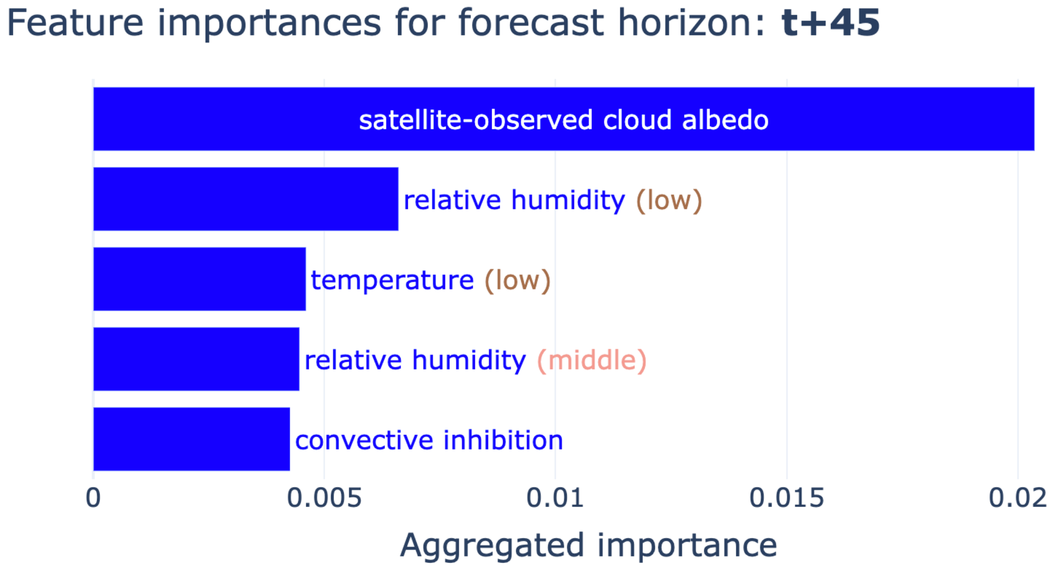
When sun elevation > 5°

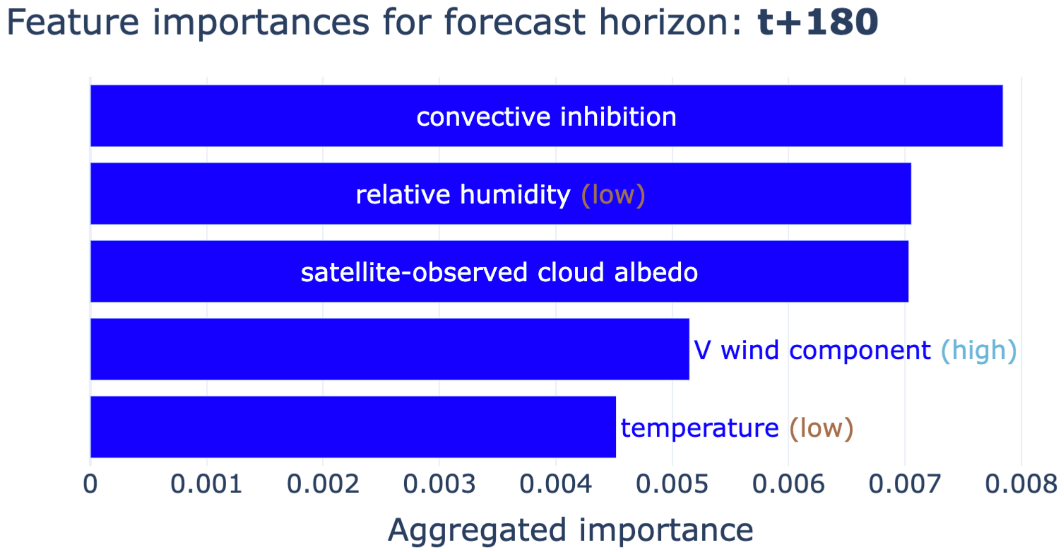
For rRMSE: lower is better

GHI is computed using the **extracted value of cloud albedo forecasted** and the **simplified Solis clear-sky model** (Ineichen, 2007)

GHI persistence: *cloud albedo(t+h) = cloud albedo(t)* but clear sky GHI varies

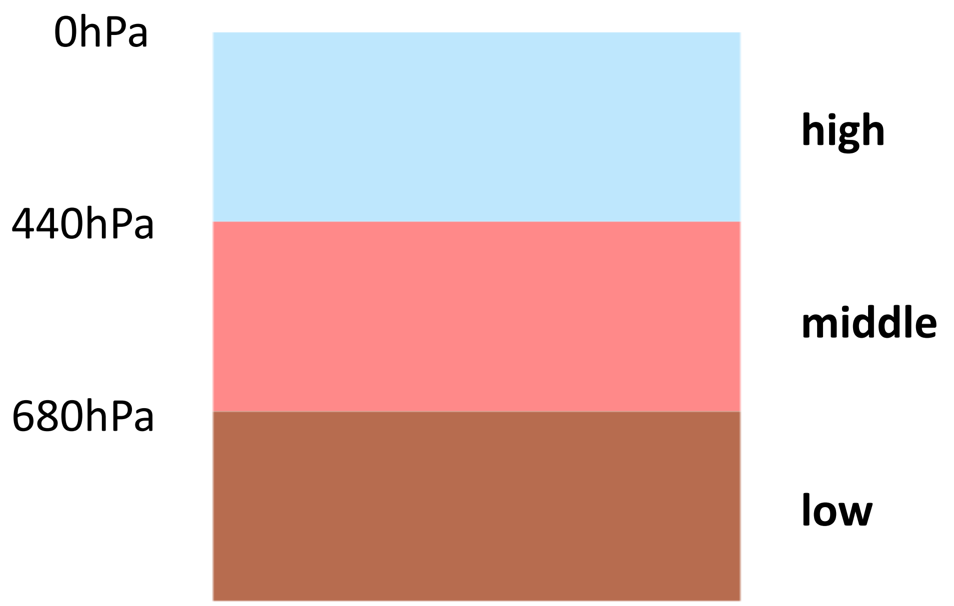
Feature importance analysis



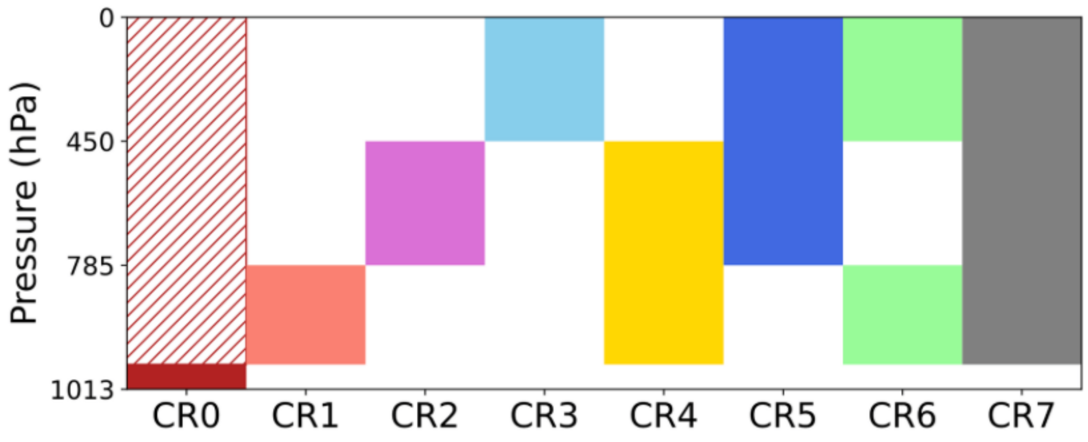


- **Convective inhibition:** negative energy that must be overcome by a parcel of air for convection to start
- **Relative humidity:** water vapour pressure as a percentage of the value at which the air becomes saturated
- **V component:** northward component of the wind

Pressure level aggregation

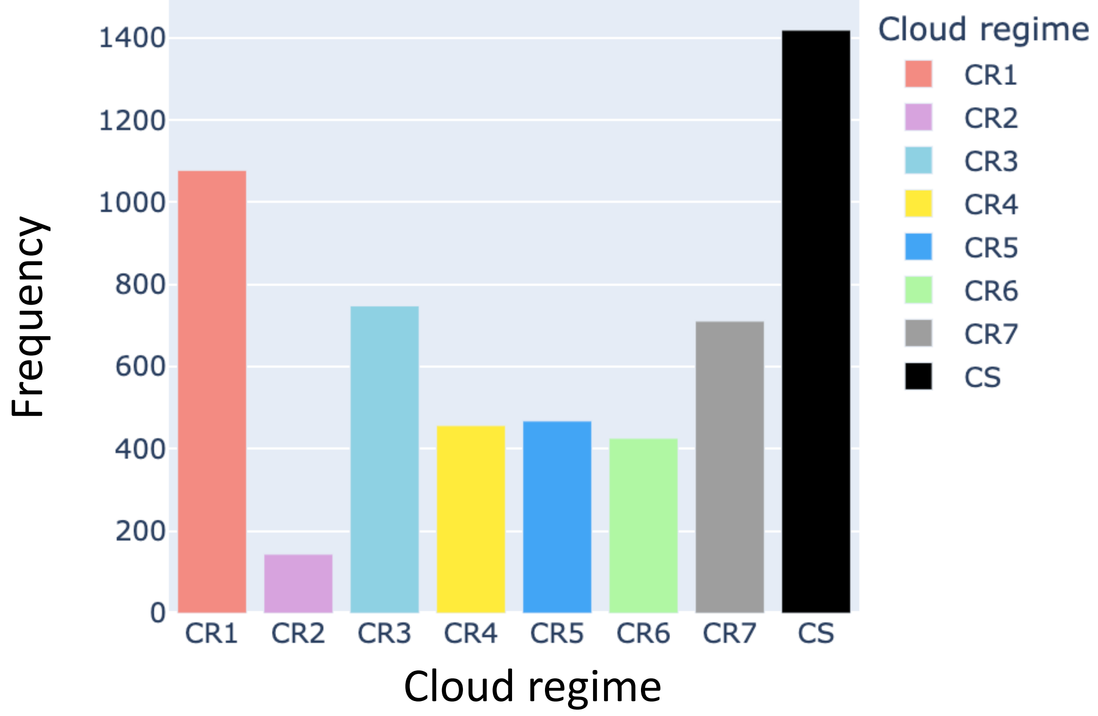


Cloud regime analysis



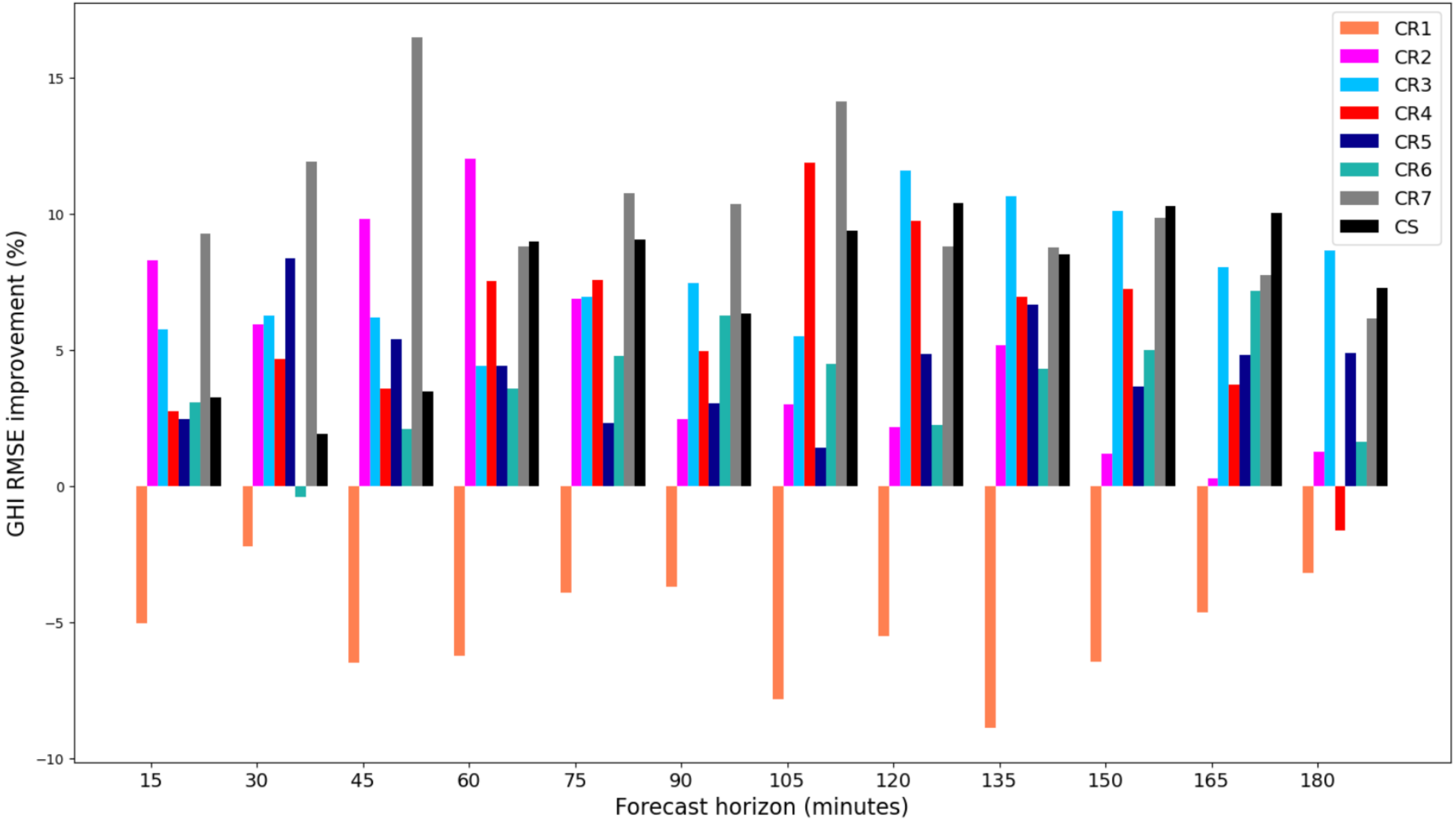
Cloud regimes definition (Magnaldo, 2023)

Cloud regime frequency (SIRTA, 2023)



Cloud regime = 30% threshold of fraction cloud cover (FCC-ERA5) between different pressure levels
*CR0 (fog) not considered in this study

GHI RMSE improvment (%) by cloud regime
ERA5 + cloud albedo model over cloud albedo only model



Conclusion

- **ERA5 variables help** the cloud albedo forecasting model achieve **better performance**
- **Cloud albedo + ERA5 model performs worse** (GHI rRMSE) **than the cloud albedo only model for CR1** (low clouds) but **outperforms it in most other cloud regimes** across all time horizons
- **U and V components** (wind speed) **are not optimally used** (could be better with more spatial information on cloud cover)

Future work

- **Generalization assessment** across various sites with differing cloud conditions and topography
- **Link feature importance to cloud regimes:** which features are the most relevant for the model to forecast a given cloud regime
- Integrate ERA5 variables in a **deep learning short-term irradiance forecasting framework**
- More focus on **spatiotemporal resolution matching** between satellite and ERA5

Thank you for your attention

The authors gratefully acknowledge the financial support from
TotalEnergies for this research project

nicolas.chea@polytechnique.edu

TotalEnergies

ÉCOLE POLYTECHNIQUE IP PARIS

LMD