ARPEGE cloud cover forecast post-processing with convolutional neural network

Florian Dupuy[1], Olivier Mestre[2], Michaël Zamo[2]

[1]IRT Saint Exupéry, France

[2]Météo-France, France

Deep4cast project

IRT Saint-Exupéry – Météo-France – IRIT – CERFACS

Context and objective

Forecast errors are common in Numerical Weather Predictions (NWP). A post-processing step is generally added to correct them.

Operational cloud cover forecasts post-processing at Météo France are based on random forests (RF) and linear quantile regressions (LQR), yielding a large improvement.

RF and LQR are approved machine learning technics massively used to post-process NWP. However, the algorithms are not especially designed to take advantage of spatial features of the data.

Convolutional Neural Networks (CNNs) are designed to process gridded data, making them a suitable tool to work with NWP data.

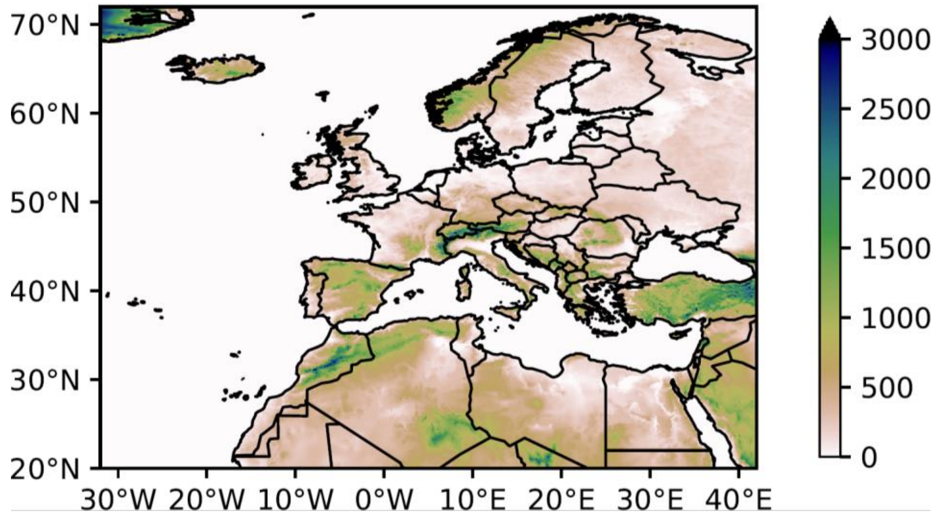
We want to evaluate the usefulness of CNNs to post-process ARPEGE cloud cover forecasts

Data

3 sources of data:

- A satellite based cloud cover analysis, produced by Météo-France (ground truth)
- ARPEGE forecasts: 31 output variables at lead time 15h from the runs starting at 00UTC (predictors)
- Surface variables from SURFEX (predictors)

ARPEGE relief (m) on the studied region

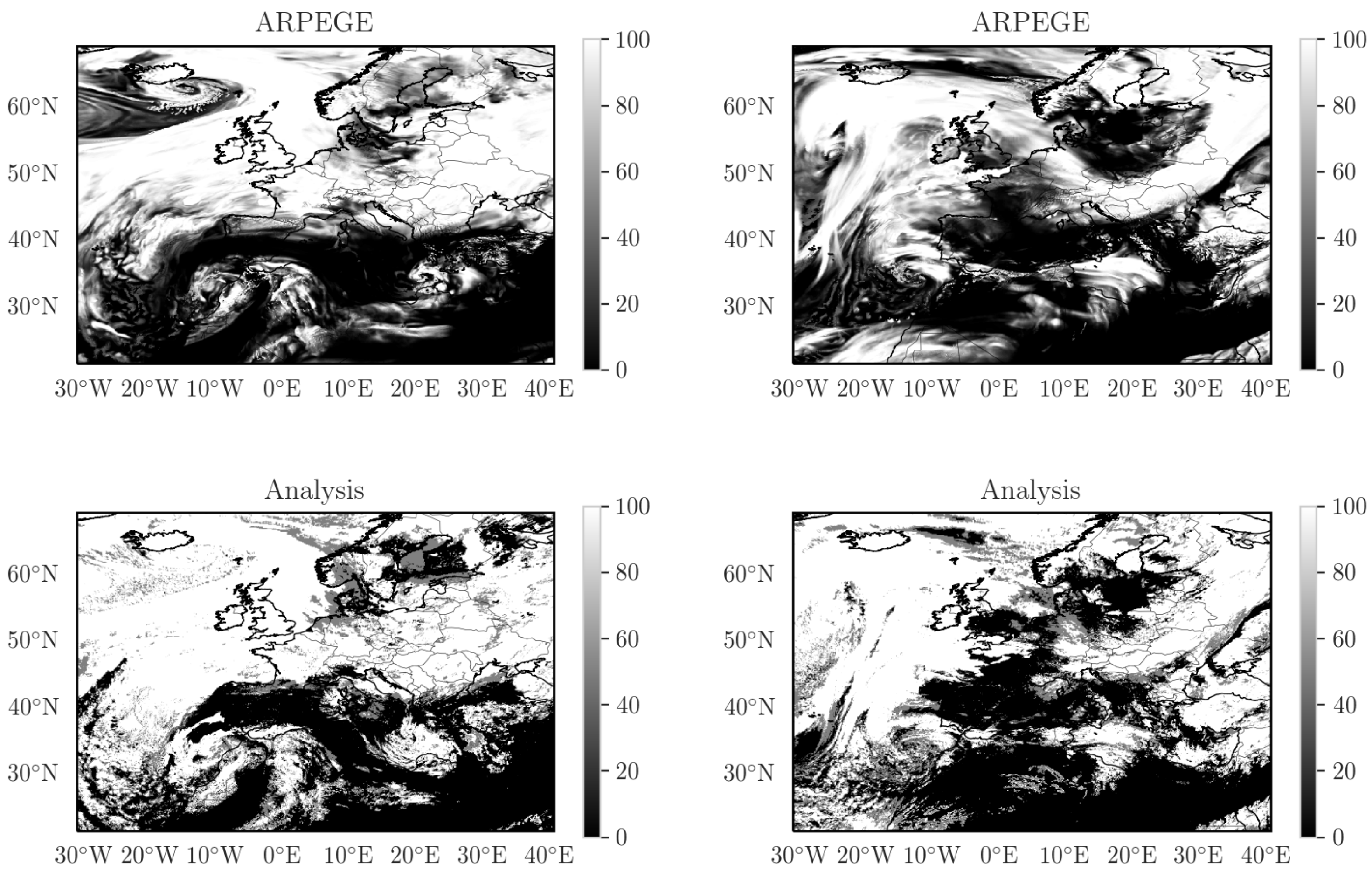


All the data are available on the same:

- 2017 – 2018 period
- grid of 0.1° horizontal resolution

ARPEGE cloud cover forecasts

21/02/2017 15/03/2017

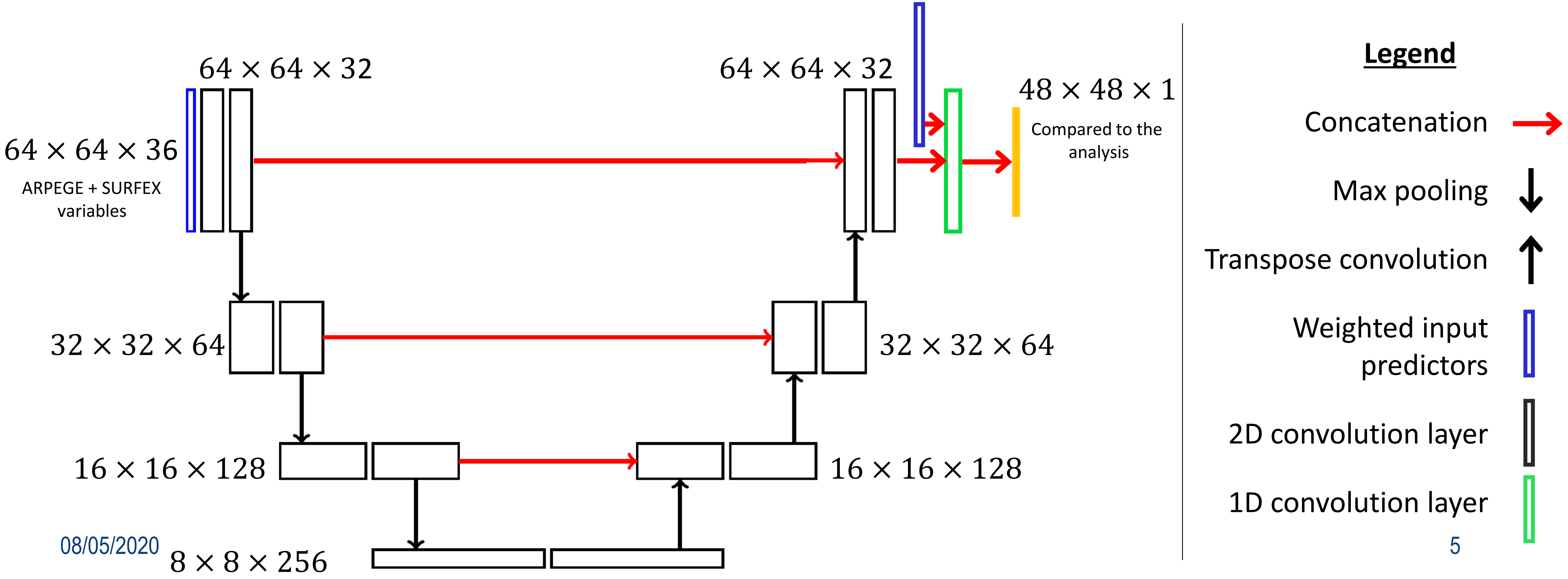


Strengths and weaknesses

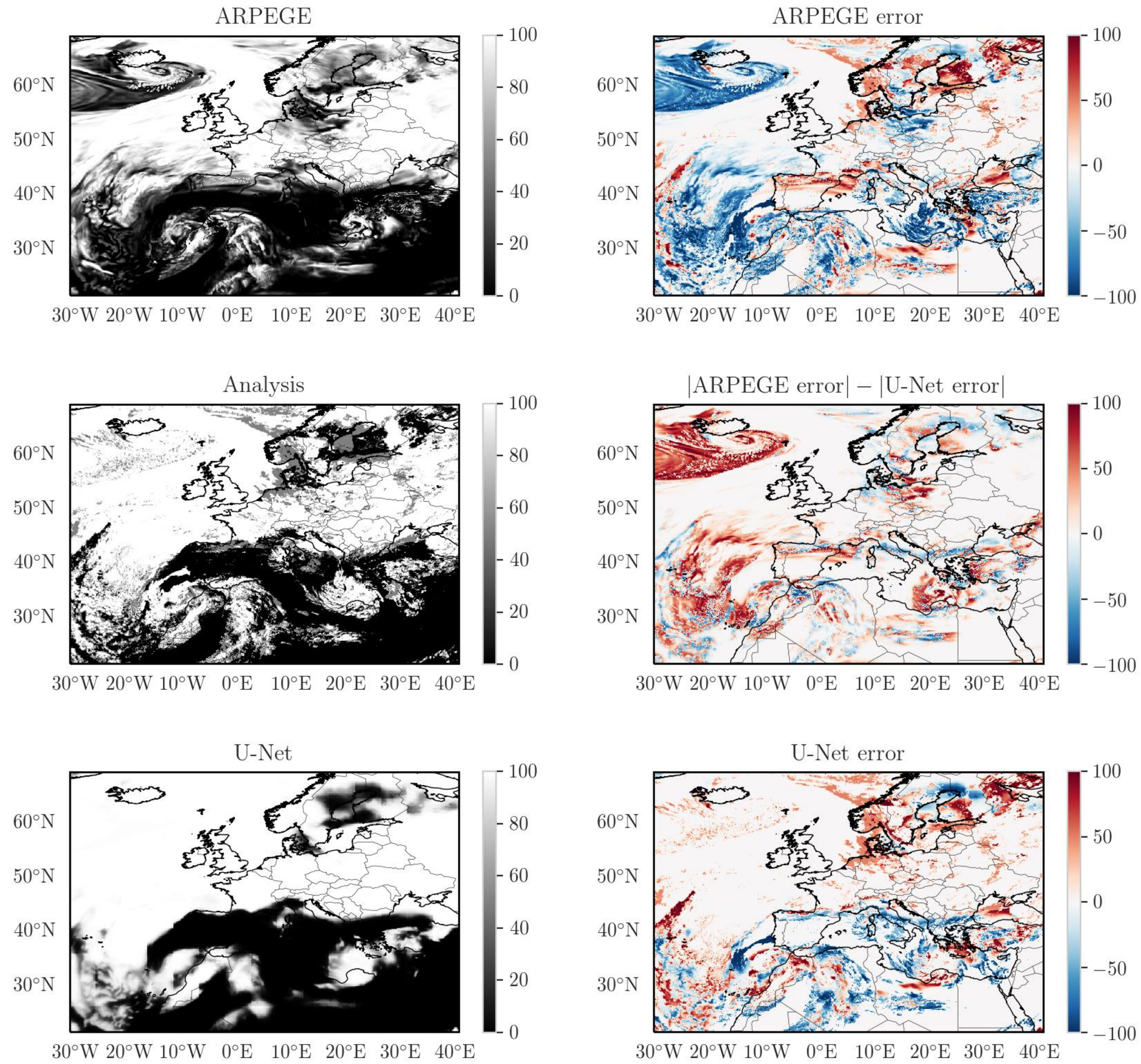
- At the large scale, ARPEGE locates the different cloudy areas…
- … but with problems of spatial extension of these areas…
- … and with a recurrent too clear sky over Lows
- Too many intermediate values causing more errors

CNN architecture: U-Net

Different architectures of U-Net were tested but we only present the results of the best U-Net in the following. It consists on a weighted predictors layer followed by a traditional U-Net architecture.



TCC maps for 21/02/2017

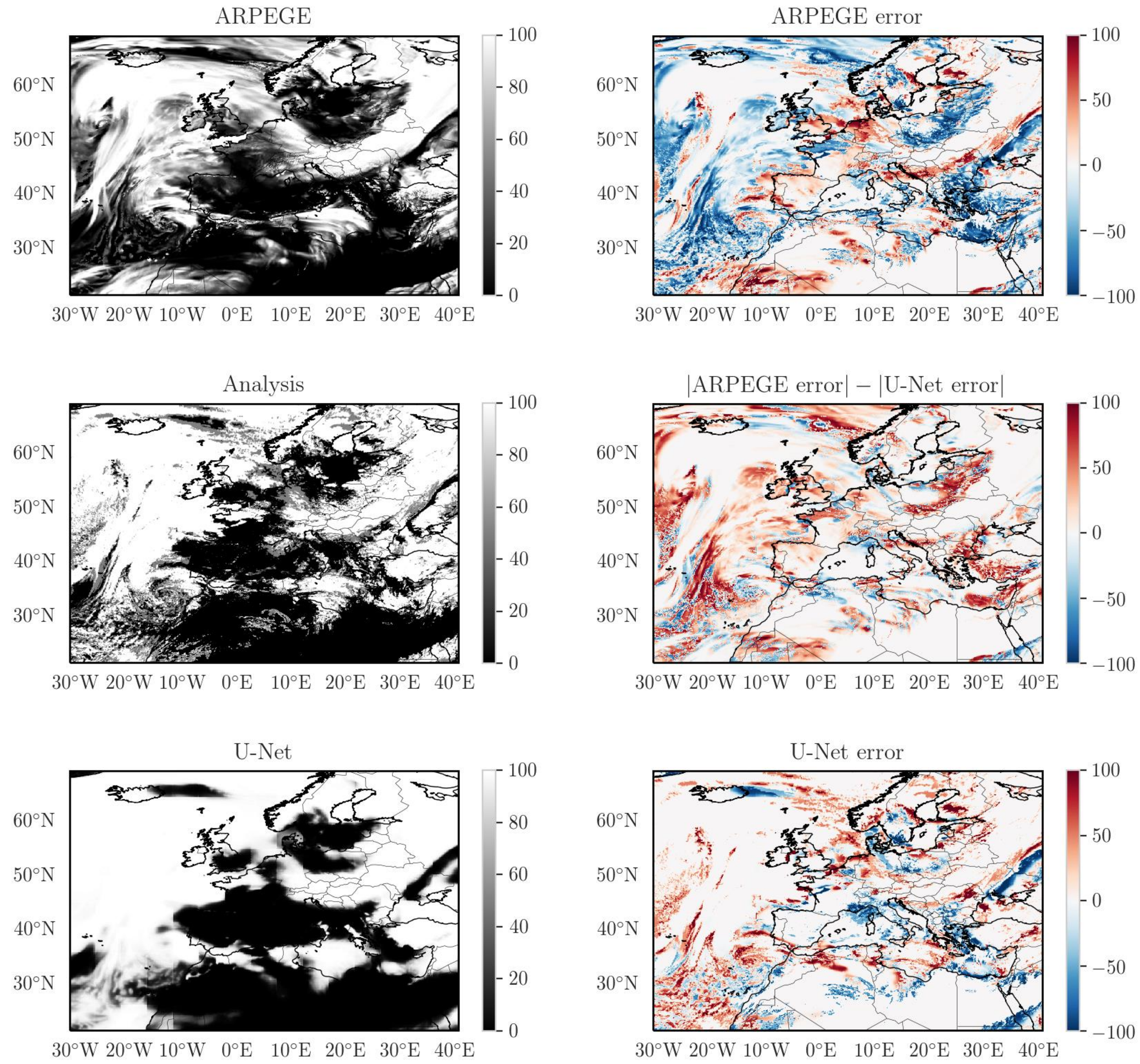


Strengths and weaknesses

- The U-Net corrects most of the spatial extension errors…
- … and the recurrent negative bias over Lows
- Too smoothed
- Recurrent problem with intermediate values

Comparison of cloud covers corresponding to the 21/02/2017 at 15h UTC

TCC maps for 15/03/2017



Strengths and weaknesses

- The U-Net corrects most of the spatial extension errors…
- … and the recurrent negative bias over Lows
- Too smoothed
- Recurrent problem with intermediate values

Comparison of cloud covers corresponding to the 15/03/2017 at 15h UTC

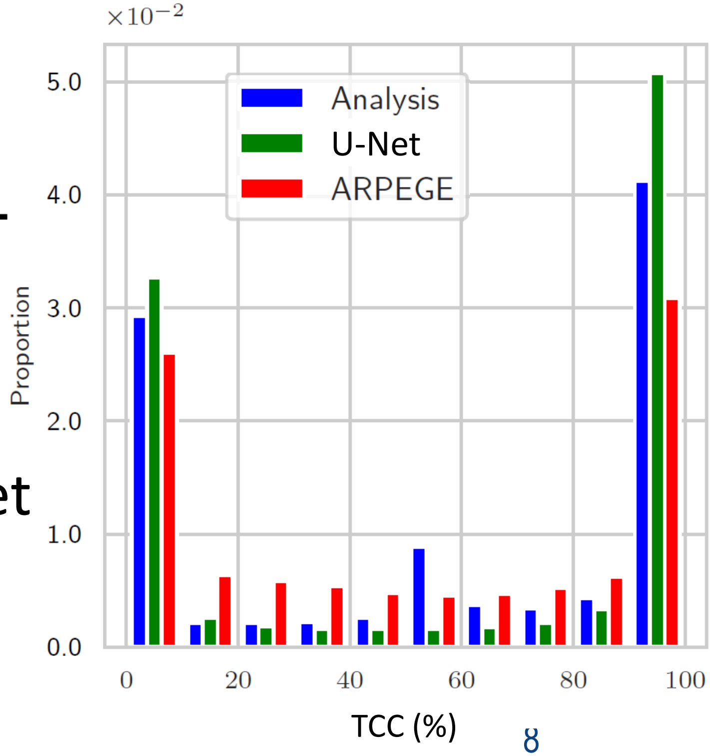
U-Net results – Climatology

- The U-Net improves the cloud cover forecast of ARPEGE and reach better results than a linear quantile regression (LQR) and a random forest (RF)

	Regression (%)		Classification				
Case	ME	MAE	HR_{10}	F_{10}	PSS_{10}	FAR_{10}	PC_{10}
ARPEGE	−5.2	23.5	0.634	0.111	0.523	0.295	0.814
LQR	+0.8	19.3	0.770	0.131	0.640	0.288	0.840
RF	+2.9	18.8	0.743	0.108	0.635	0.257	0.848
UNet	+2.1	17.3	0.819	0.125	0.694	0.267	0.858

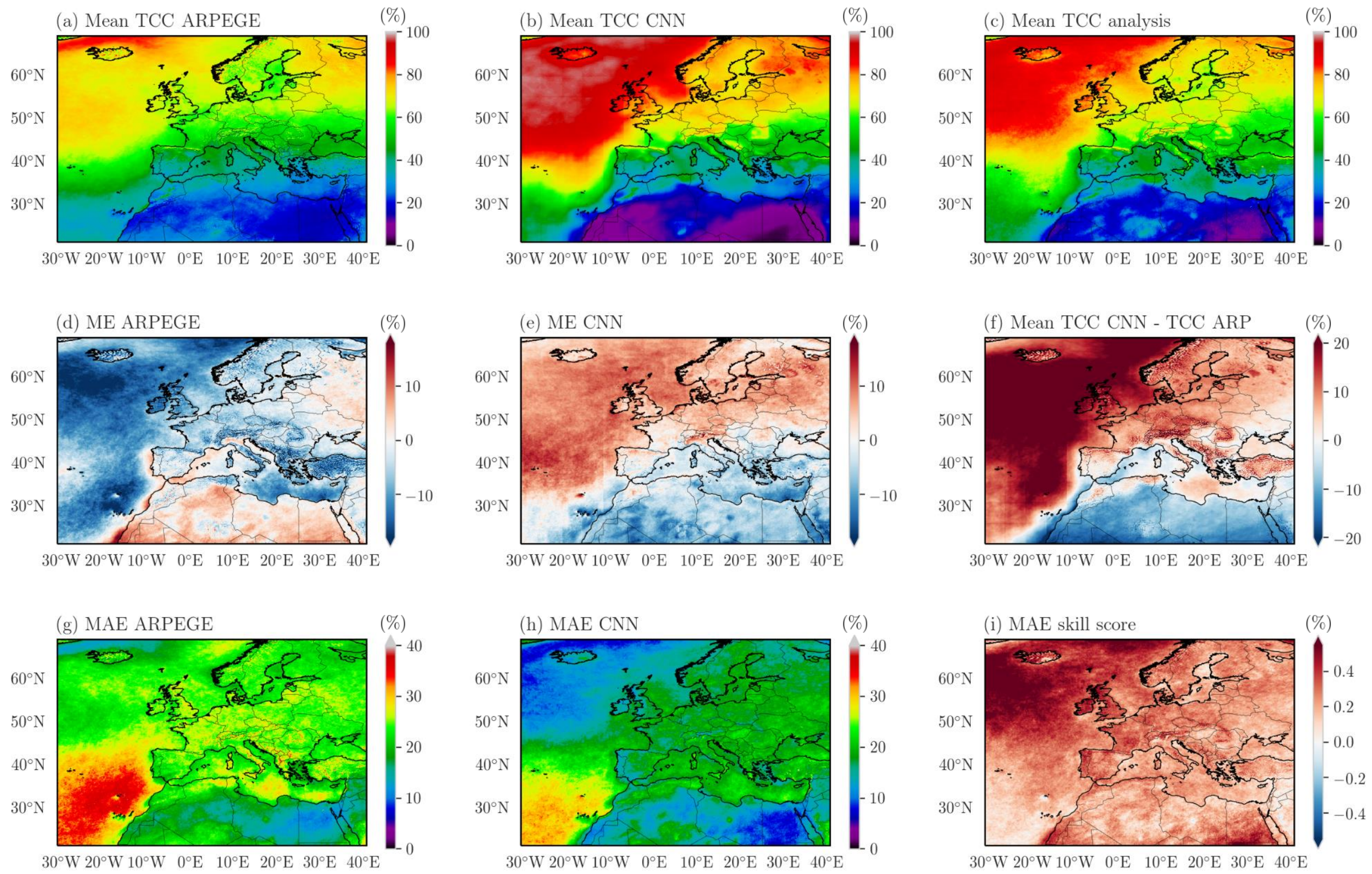
ME: Mean Error; MAE: Mean Absolute Error; HR_{10}: Hit Rate for cloud covers < 10%; F_{10}: False Alarm Rate for cloud covers < 10%; $PSS_{10} = HR_{10} - F_{10}$: Pierce Skill Score; FAR_{10}: False Alarm Ratio for cloud covers < 10%; PC_{10}: Proportion correct for 2 cloud cover classes (above or below 10%)

Distributions of total cloud covers (TCC) over the 2-years period.



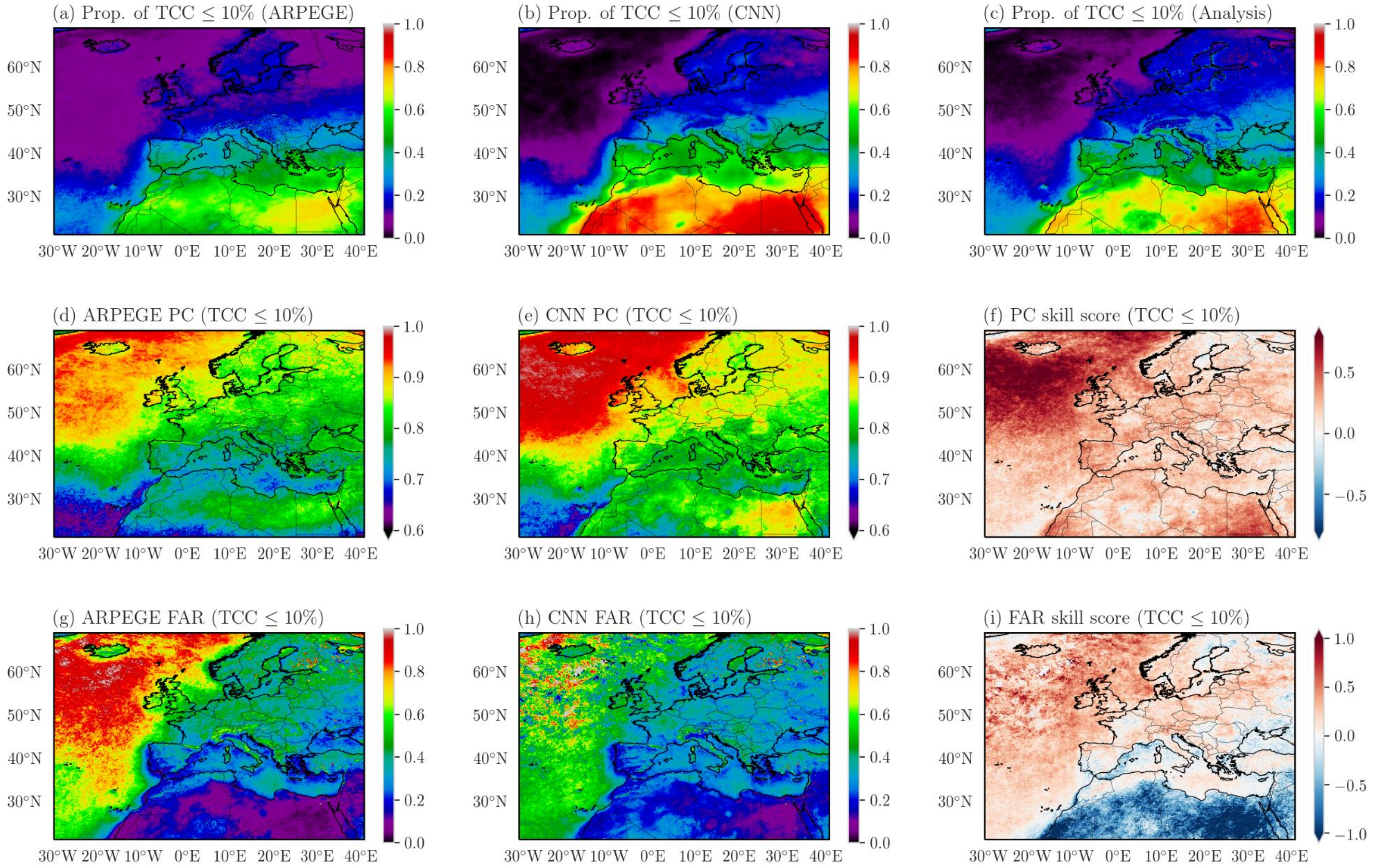
- The ARPEGE distribution was too flat. On the opposite, the U-Net overestimates the occurrence of clear sky and overcast
- Spike at 50% represented neither by ARPEGE nor by the U-Net

U-Net results – Climatology



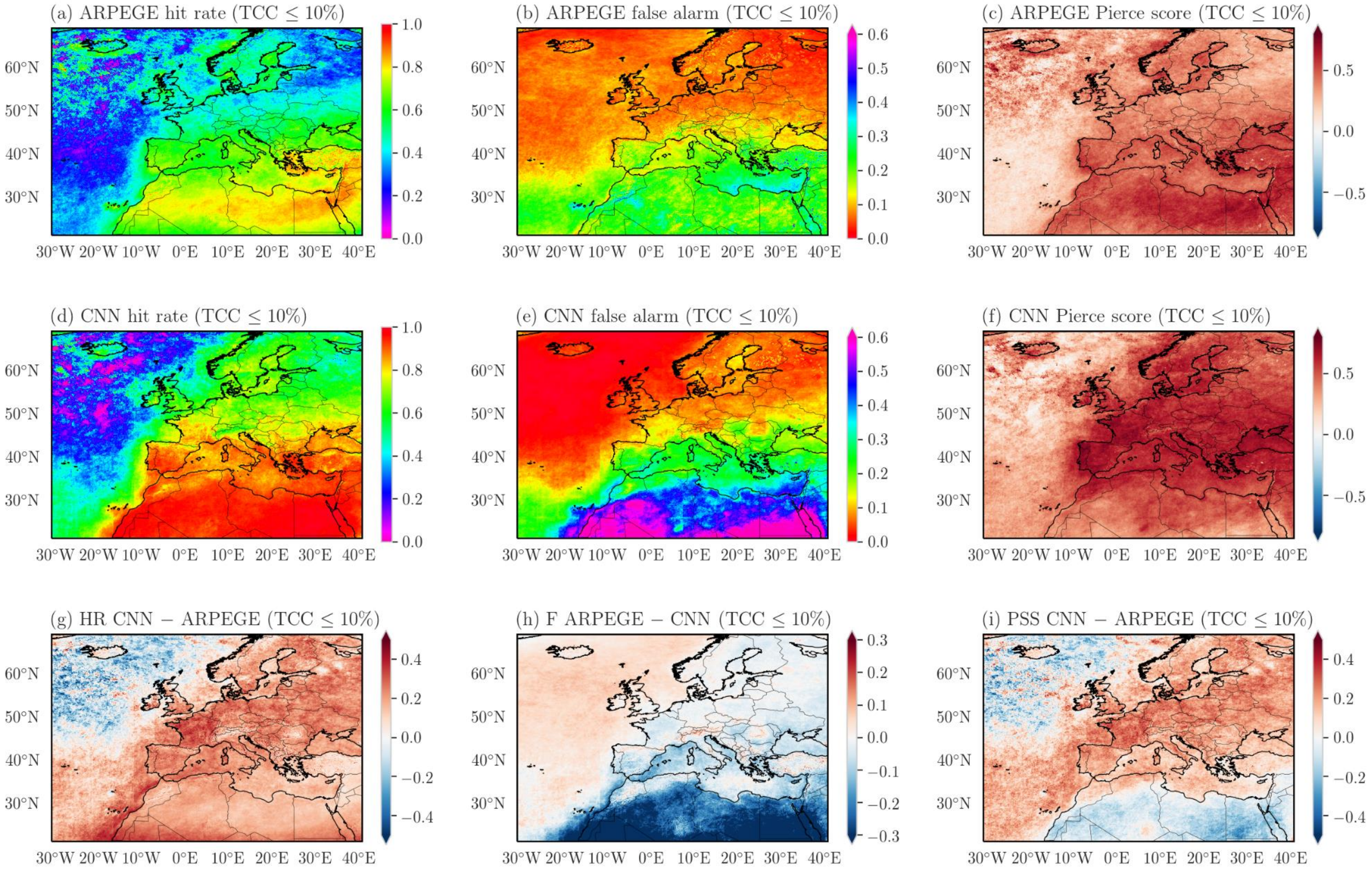
- Large improvement of mean TCC especially over mountains
- Generalized decrease of MAE
- Exaggeration of the climatology: positive bias over regions with high mean TCC and negative bias where mean TCC are low

U-Net results – Climatology of TCC ≤ 10%



- Better representation of the proportion of clear sky
- Improvement of the proportion of correct classification
- Increase of false alarm over Africa

U-Net results – Climatology of TCC ≤ 10%



- Global improvement

except for:

- Africa, where the false alarm increases
- Northern part of the Atlantic Ocean where the hit rate decreases

U-Net results – Predictors importance

A first step toward model interpretation is to find which predictors are the most important for the model. Traditional methods to perform such task are permutation importance or forward (or backward) feature selection. However, their computational time is huge.

We introduced a weighted predictors layer prior to the traditional U-Net architecture that performs a predictor importance ranking. Contrary to the traditional methods, the ranking is done during the training phase by fitting these weights.

The additional trainable parameters equals the number of predictors (36 in our case), which is negligible comparing to the U-Net itself (millions of trainable parameters). Therefore, there is no impact on the computational time.

U-Net results – Predictors importance

Weights of the weighted predictors layer. The 4 values per variable correspond to the 4 models consecutive to a cross-validation using 6-months periods for the test set.

Fundamental meteorological variables*. Ts: surface temperature; T 2 m: 2-m temperature; MSLP: mean sea level pressure; U and V 100 m: zonal and meridional wind components at 100 m a.g.l.*

Cloud-related variables*. LOW LV CC: low level cloud cover; MID LV CC: middle level cloud cover; HIGH LV CC: high level cloud cover; CONV CC: convective cloud cover; TCC: total cloud cover; CLD FRACT: cloud fraction.*

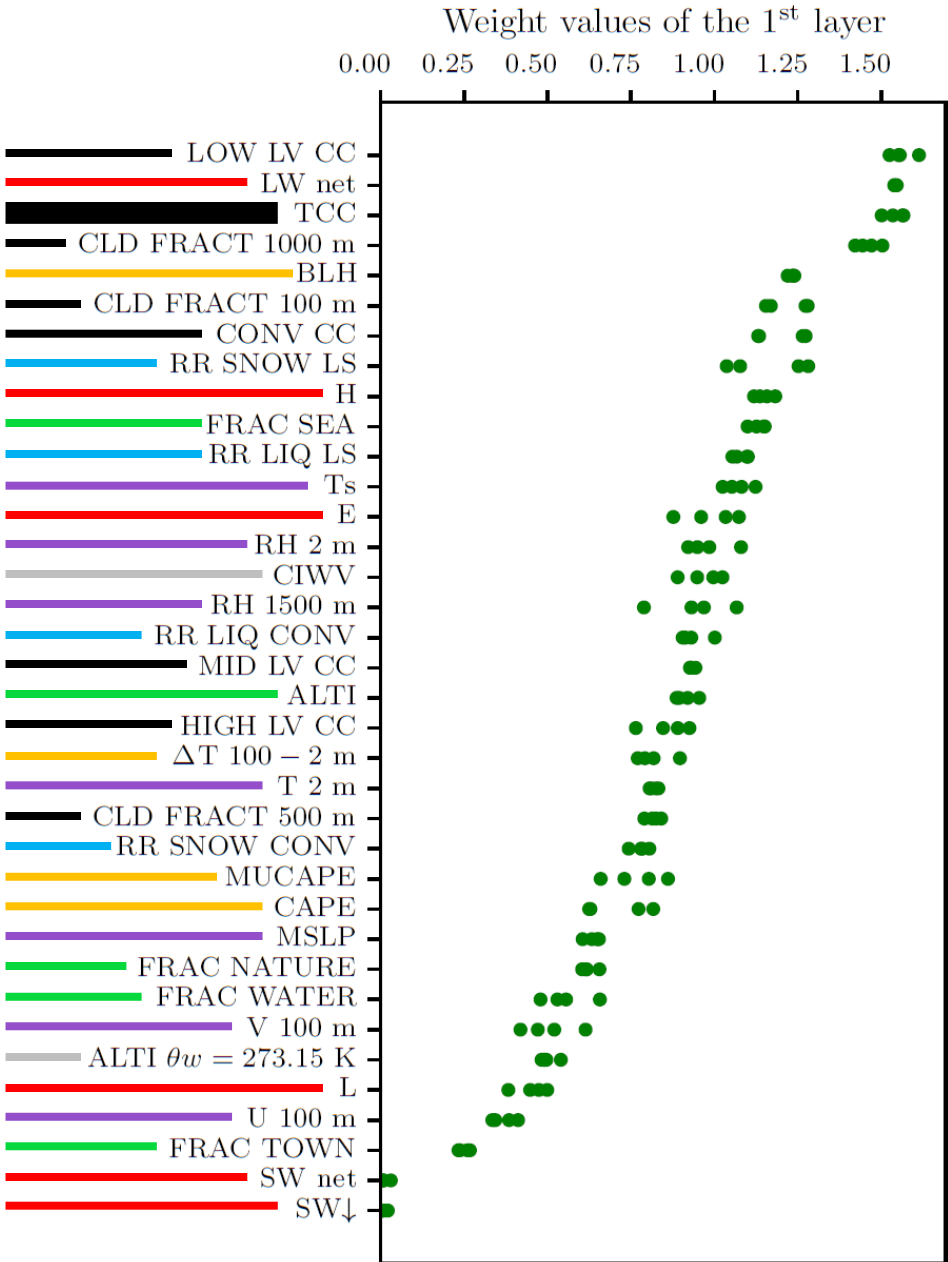
Precipitation variables*. RR corresponds to 3-hours rainfall accumulation, SNOW and LIQ distinguish snow and liquid precipitations while LS and CONV means large scale and convective precipitations.*

Flux variables*. LW net: net longwave radiation at the surface; H: sensible heat flux; E: evaporation flux; L: latent heat flux; SW net: net shortwave radiation at the surface; SW↓: ongoing shortwave radiation at the surface.*

Atmospheric stability*. BLH: boundary layer height; ΔT100 – 2 m: vertical difference of temperature between 100 and 2 m; CAPE: convective available potential energy in the model; MUCAPE: most unstable CAPE.*

Other variables*. CIWV: column integrated water vapor; ALTI $\theta w = 273.15\ K$: altitude of the 0°C wet-bulb potential temperature level.*

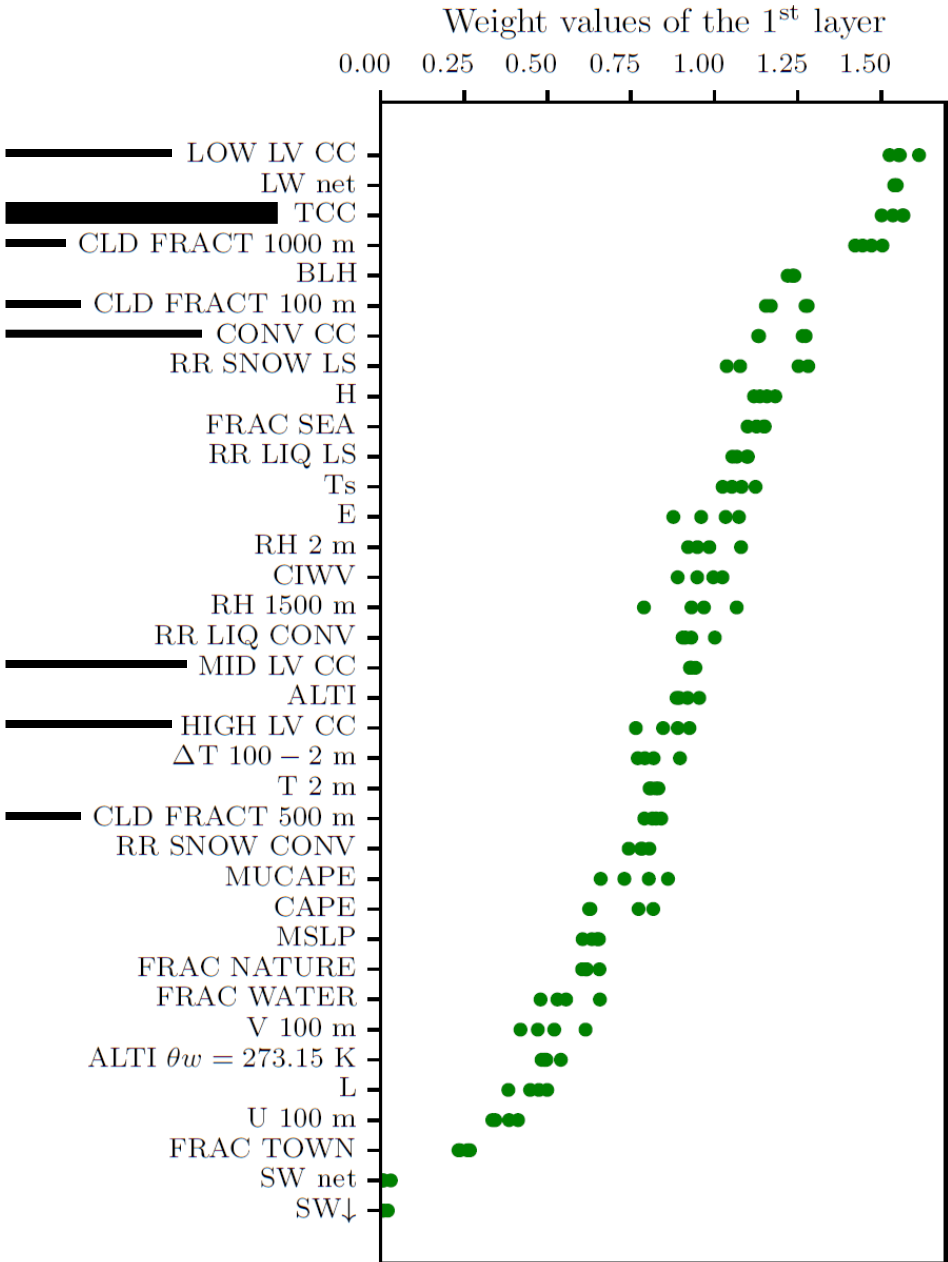
Surface variables*. ALTI: altitude; FRAC SEA, NATURE, WATER and TOWN: grid cell fraction occupied by seas and oceans, natural surfaces, continental water bodies and artificial surfaces (from SURFEX).*



U-Net results – Predictors importance

Cloud-related variables:

- They are mostly on the top of the ranking
- TCC is one of the most important variables which makes sense since it is the variable we want to correct
- Low level clouds forecast is known to be very challenging in NWP. They are responsible of repeatable errors in the ARPEGE TCC forecasts. This makes this variable useful for the U-Net to correct these errors.
- Cloud fraction at 500 m is less important:
  - redundancy of informations with cloud fractions at 100 m and 1000 m $R_{100/1000} = 0{,}28$; $R_{100/500} = 0{,}49$; $R_{500/1000} = 0{,}59$
  - more clouds at 100 m (fog) and 1000 m (closer to the top of the BL) than at 500 m

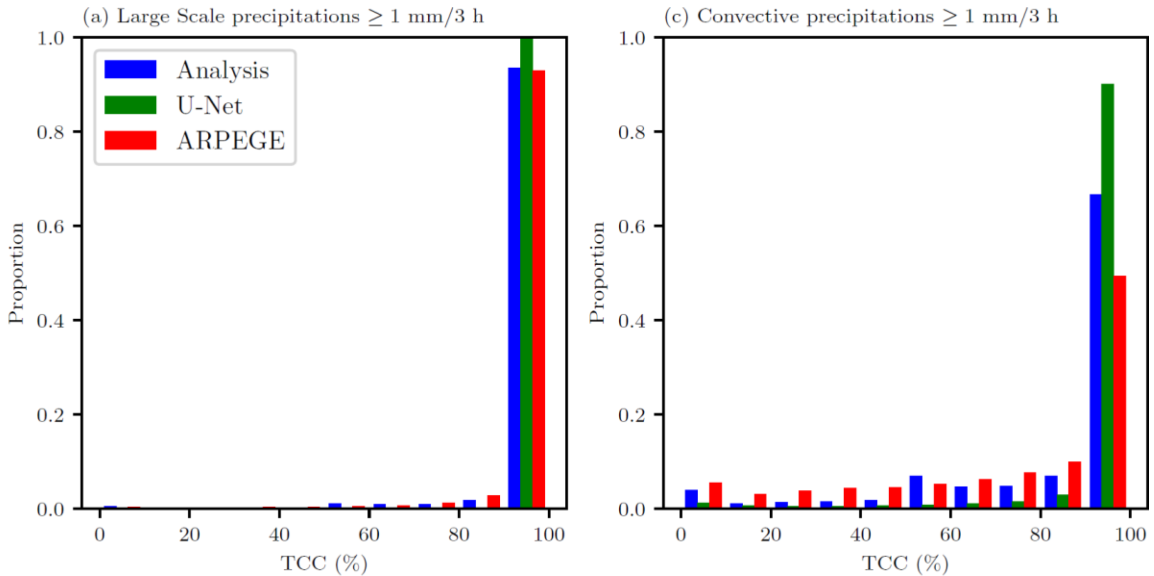


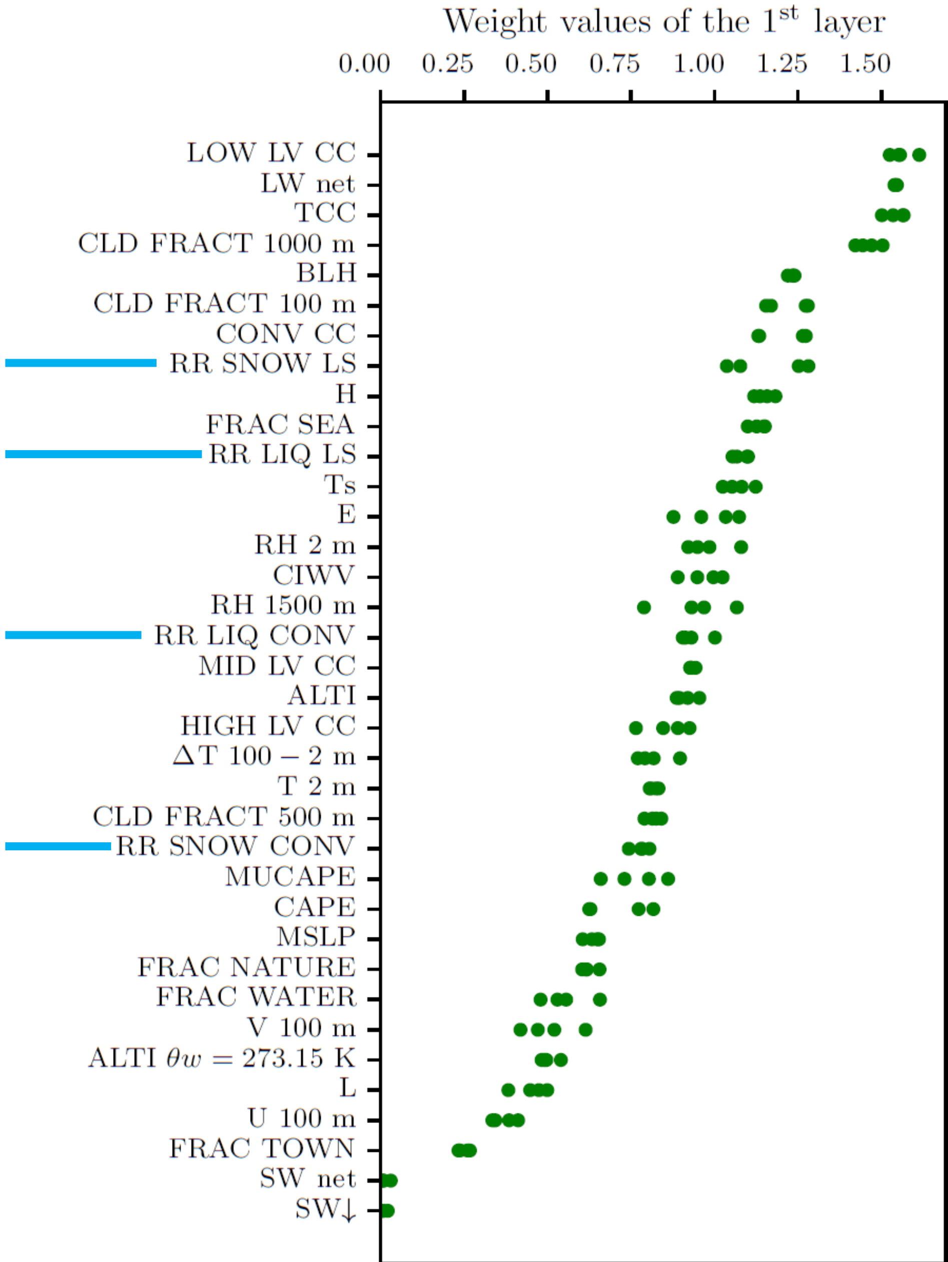
U-Net results – Predictors importance

Precipitation variables:

Large scale precipitations (LSP) are more important

- LSP are mostly associated to overcast conditions, which the U-Net reproduces (and exaggerates)
- TCC associated to convective precipitations are more variable
  - ↪ LSP are a good predictor to diagnose overcast conditions





Conclusion – post-processing

We applied a deep learning algorithm, a convolutional neural network, to a weather forecast post-processing task: the ARPEGE cloud cover forecasts correction

The U-Net, a particular CNN, proved its efficiency to solve this problem in comparison to random forests and a linear regression.

- Better localization of cloudy and clear sky areas, especially concerning their spatial extend
- Improvement of all the metrics over a large part of the domain: Europe and the southern part of the Atlantic Ocean
- The improvement over mountains is remarkable

- Slight exaggeration of the climatology: too few clouds over Africa, (↗ false alarm), and too much clouds over the northern part of the Atlantic Ocean (↘ hit rate)
- TCC fields are too smooth
- No improvement concerning intermediate values

Conclusion – interpretation

The modified U-Net performs a predictor importance ranking during its training phase, meaning there is no impact on the computational time. This represents an improvement compared to traditional methods such as permutation importance or forward (or backward) feature selection.

- Cloud-related variables are very important which seems logic.
  - The ARPEGE TCC, which is the variable we want to correct, is one of the most important variables.
  - Low level CC is also important, probably because it is responsible of numerous errors on the TCC.
- Large scale precipitations are more important than convective precipitations. They are a good variable to diagnose overcast conditions
- Further analysis have to be done concerning the other variables

The deep4cast project

This work is part of the deep4cast project, which is a French collaboration aiming at exploring some possibilities of **deep learning** applied to **meteorological** problems (https://www.researchgate.net/project/Deep4Cast).

Additional work done in deep4cast concern:

- Precipitation nowcasting using deep neural network (EGU – session HS3.4)
- Efficient POD-kriging surrogate models for rainfall forecasting (EGU – session HS7.2)
- NetCDF: Performance and storage optimization of Meteorological data (EGU – session ITS4.1/NP4.2)

The deep4cast project is funded by the STAE foundation

Merci de votre attention