Rainfall scenarios from AROME-EPS forecasts using autoencoder and climatological patterns

Arnaud MOUNIER, Laure RAYNAUD, Lucie ROTTNER, Matthieu PLU

CNRM, University of Toulouse, Météo-France, CNRS, Toulouse, France

31 March 2022, ECMWF ML Workshop

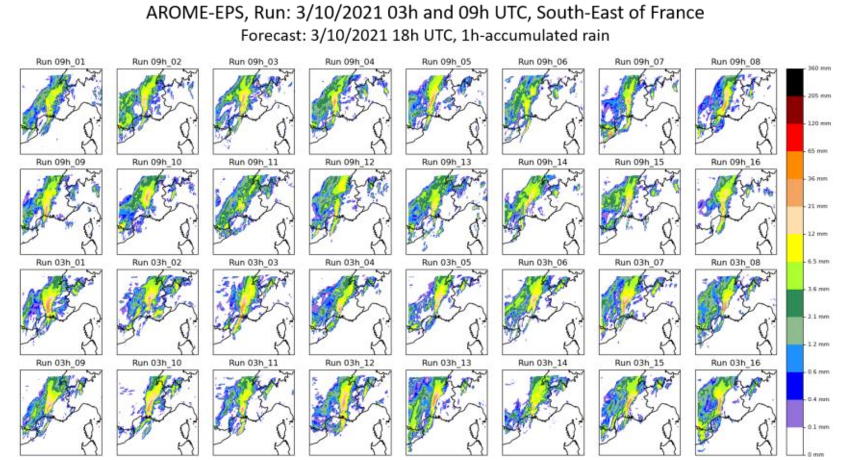
contact : arnaud.mounier@meteo.fr

Introduction *AROME-EPS*

- ❑ AROME-EPS : French convective scale ensemble prediction system. 16 members. Horizontal resolution of 2.5 km (soon 1.3 km). runs 4 times a day and lead times up to 51h. Built to especially improve **severe convective storm forecasts** (MCS, Extreme rainfall in Mediterranean France).
- ❑ AROME-EPS output are most of the time point-based probabilities or percentiles. They provide useful information, but **lack of physical consistency**. We can not recognize meteorological structures. Issue for forecasters in an operational context.
- ❑ Aim: a new approach (scenarios) to summarize AROME-EPS output.
- ❑ Time-lagged ensemble to increase the ensemble size (32 members instead of 16 members)

Introduction

Why weather scenarios ?



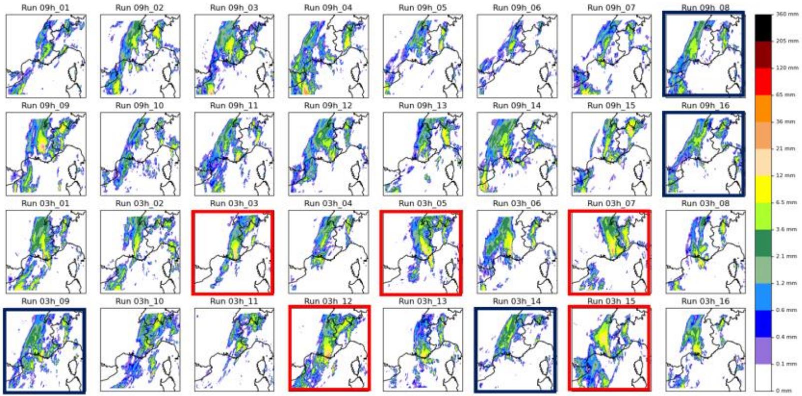
❑ Mainly one scenario (heavy rain over the same area). Relatively high predictability.

Introduction

Why weather scenarios ?

AROME-EPS, Run: 3/10/2021 03h and 09h UTC, South-East of France

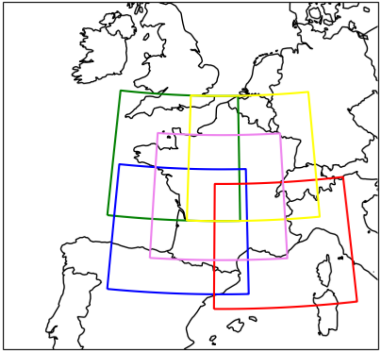
Forecast: 4/10/2021 06h UTC, 1h-accumulated rain



- 12 hours later. Multiple scenarios (red, blue squares + other). Low predictability. How to convey this information?

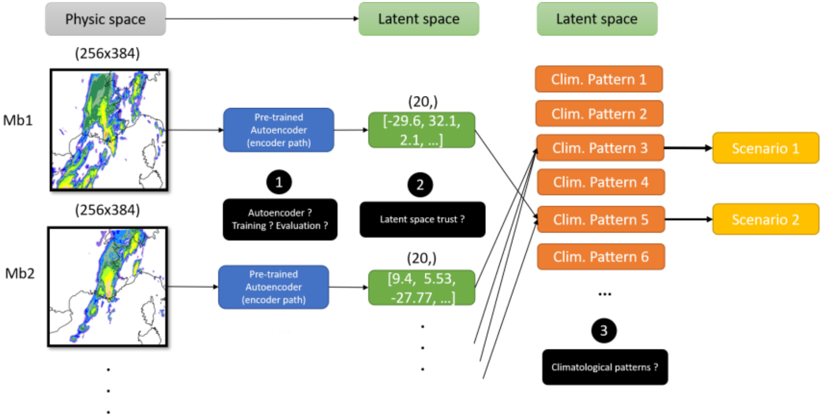
Introduction

Regional scenarios



- If entire domain, not enough members compared with atmospheric degrees of freedom
- In add, risk to take into account multiple meteorological structures simultaneously (cold front + extreme rainfall).
- Regional scenarios are computed. 5 areas of exactly same shape (256×384 grid points)

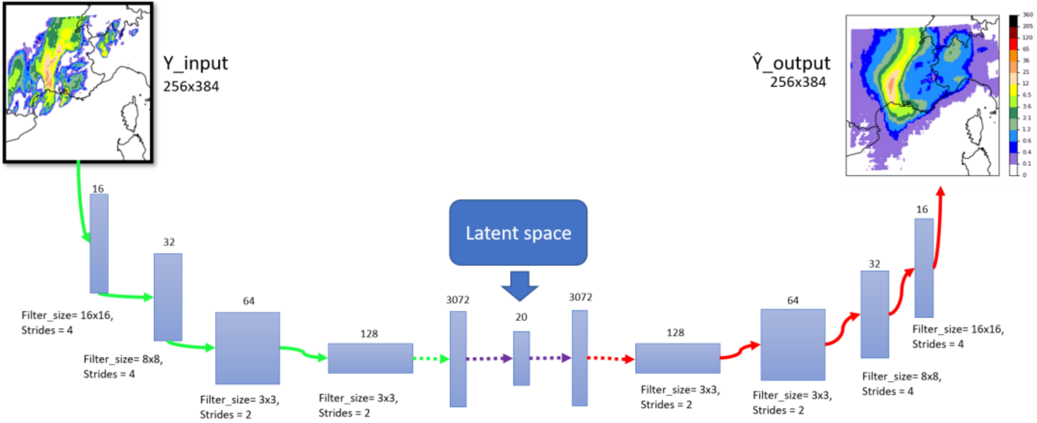
Scenario design *Overview*



- Convolutional autoencoder to define a latent space, dimension reduction (256x384 grid points to few dimensions) to avoid grid point-based metrics
- Each vector in latent space assigned to a pre-defined climatological pattern
- Inspired by Neal et al. (2016) and Karim et al. (2020)

Scenario design

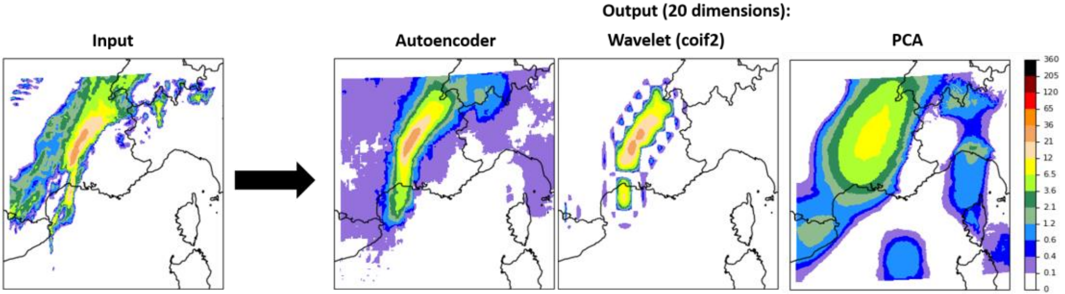
Autoencoder



- ❑ Inspired by Guo et al. (2017)
- ❑ Encoder path : Multiple convolutional layers + 1 dense layer
- ❑ Decoder path : 1 dense layer + Multiple up-convolutinal layers
- ❑ Loss function : MSE

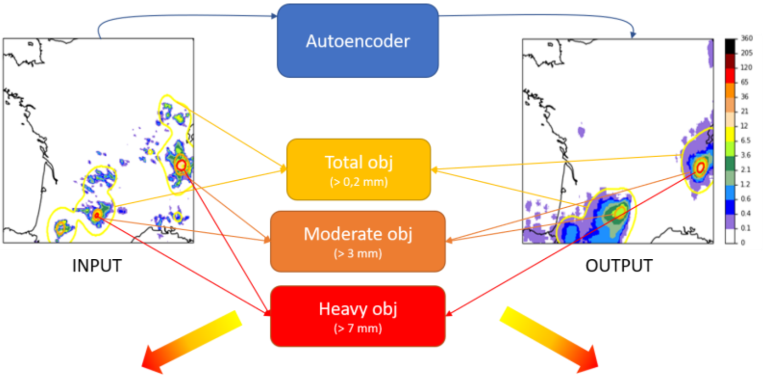
Scenario design *Autoencoder, training & main results*

- One parameter: 1h-accumulated rainfall
- Training database: 4 years of AROME-EPS (2017 to 2020)
- Validation database: 1 year (2021)
- Same autoencoder for the 5 regional areas (architecture & dataset)
- Preprocessing steps: grids without rain are removed + importance sampling method, mainly inspired by Ravuri et al. (2021)
- Comparison Autoencoder vs Wavelet vs PCA:



Scenario design

Autoencoder, training & main results



Global scores (for each kind of object) :

- From contingency table. Rule : detected object in input and/or output ? (Hit rate and False alarm rate)
- Distance between mass centers of input and output objects

Other **specific scores** based on object attributes (not shown) :

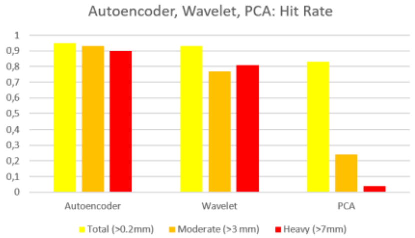
- Q90 inside objects correlation
- Non detection features (small area ? small Q90 ?)
- …

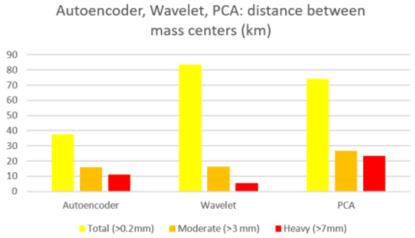
❑ Object-oriented scores from input and output fields (inspired by Davis et al. (2006))

Scenario design

Autoencoder, training & main results

- Autoencoder vs Wavelets vs PCA: example for 20 dimensions (best autoencoder and best wavelet available). Autoencoder clearly better than wavelets and PCA.
- Latent space study (ongoing) based on the validation dataset and idealized patterns (see example in appendix).





Scenario design

Climatological patterns (ongoing)

First idea:

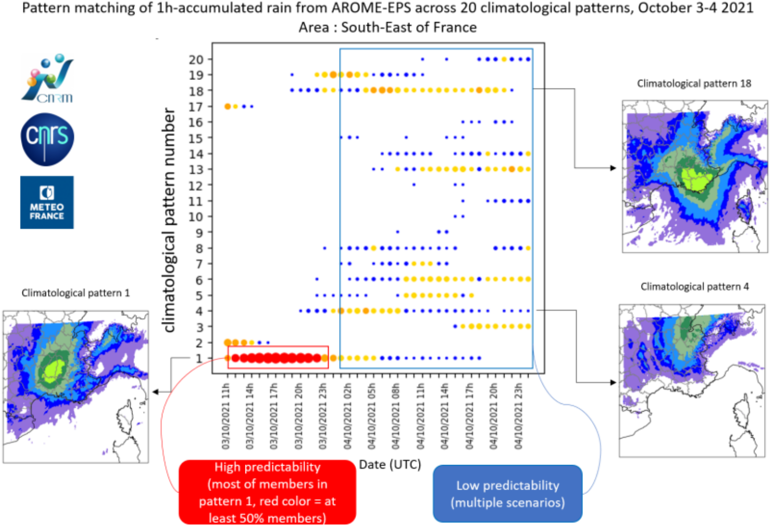
- clustering of the latent space vectors (1 vector = 1 member)
- Unstable results in account of limited points (32 members in around 20 dimensions). Results are highly sensitive to clustering methods, metrics.
- This idea does not seem reliable

Chosen idea (climatological patterns):

- Inspired by Neal et al. (2016)
- Computing a large amount of latent space data.
- Clustering these data to define climatological patterns.
- First results in the next illustrations with KMeans method (5 years of data from 2016 to 2020). 20 climatological patterns are defined (no optimum concerning the number according to classical clustering scores).

Illustrations

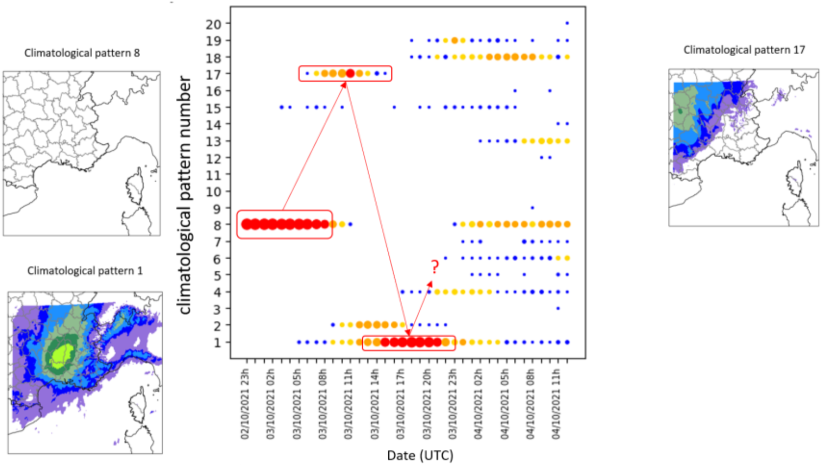
South-East of France, October 3-4 2021



Illustrations

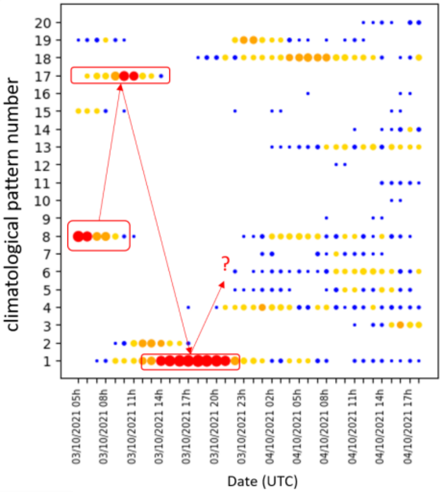
South-East of France, October 3-4 2021

Pattern matching of 1h-accumulated rain from AROME-EPS across 20 climatological patterns, October 3-4 2021
AROME-EPS runs: **2 October 2021 15h and 21h UTC**, Area : South-East of France



Illustrations *South-East of France, October 3-4 2021*

Pattern matching of 1h-accumulated rain from AROME-EPS across 20 climatological patterns, October 3-4 2021
AROME-EPS runs: **2 October 2021 21h and 3 October 03h UTC**, Area : South-East of France

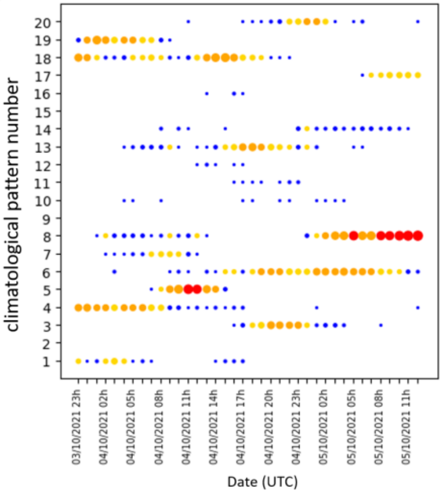


Illustrations

South-East of France, October 3-4 2021

Pattern matching of 1h-accumulated rain from AROME-EPS across 20 climatological patterns, October 3-4 2021
AROME-EPS runs: **3 October 2021 15h and 21h UTC**, Area : South-East of France

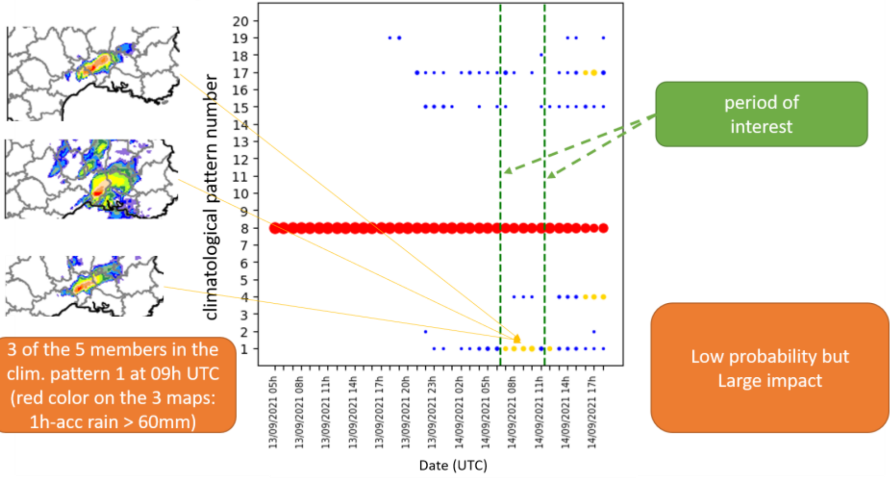
Really low predictability ! Consistent with the forecaster feedback.



Illustrations

South-East of France, September 13-14 2021

Pattern matching of 1h-accumulated rain from AROME-EPS across 20 climatological patterns, October 3-4 2021
AROME-EPS runs: **12 September 2021 21h and 13 September 03h UTC**, Area : South-East of France

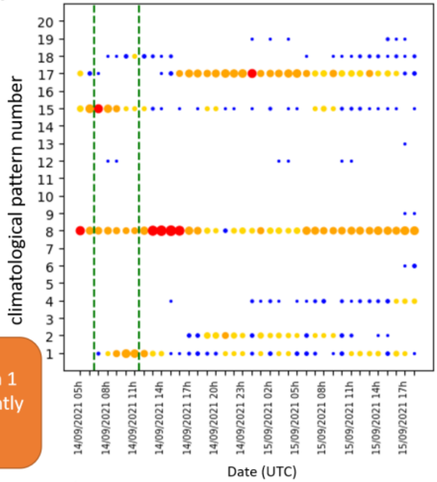


3 of the 5 members in the clim. pattern 1 at 09h UTC (red color on the 3 maps: 1h-acc rain > 60mm)

Illustrations

South-East of France, September 13-14 2021

Pattern matching of 1h-accumulated rain from AROME-EPS across 20 climatological patterns, October 3-4 2021
AROME-EPS runs: **13 September 21h and 14 September 03h UTC**, Area : South-East of France



Probability of pattern 1 (and 15) has significantly increased!

Observation:
280mm locally in 3 hours.

Conclusion

- ❑ Aim : AROME-EPS member classification
- ❑ New methodology to define weather scenarios (autoencodeur + climatological patterns)

Future work:

- ❑ Computing and study of latent spaces + climatological patterns.
- ❑ Climatological study of scenarios (number according to lead time, comparison with observation, ...)
- ❑ The unperturbed EPS member could be added also in these synthesis plots
- ❑ Severity index for each scenario
- ❑ Selection of the 'good' number of climatological patterns (+'good' latent space size) in cooperation with the forecasting department.

References

C. Davis, B. Brown, and R. Bullock. Object-based verification of precipitation forecasts. part i: Methodology and application to mesoscale rain areas. *Monthly Weather Review*, 134, 2006. doi: 10.1175/MWR3145.1.

X. Guo, X. Liu, E. Zhu, and J. Yin. Deep clustering with convolutional autoencoders. In *Neural Information Processing*, 2017.

M. Karim, O. Beyan, A. Zappa, I. Costa, D. Rebholz-Schuhmann, M. Cochez, and S. Decker. Deep learning-based clustering approaches for bioinformatics. *Briefings in Bioinformatics*, 22(1):393-415, 2020. doi: 10.1093/bib/bbz170.

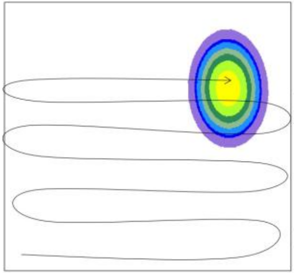
R. Neal, D. Fereday, R. Crocker, and R. E. Comer. A flexible approach to defining weather patterns and their application in weather forecasting over Europe. *Meteorological Applications*, 23(3): 389-400, 2016. doi: https://doi.org/10.1002/met.1563.

S. Ravuri, K. Lenc, M. Willson, D. Kangin, R. Lam, P. Mirowski, M. Fitzsimons, M. Athanassiadou, S. Kashem, S. Madge, R. Prudden, A. Mandhane, A. Clark, A. Brock, K. Simonyan, R. Hadsell, N. Robinson, E. Clancy, A. Arribas, and S. Mohamed. Skilful precipitation nowcasting using deep generative models of radar. *Nature*, 597:672-677, 2021. doi: 10.1038/s41586-021-03854-z.

Scenario design

Autoencoder, Latent space study example (20 dimensions)

Multiple idealized pattern positions, evolution of each dimension in latent space + Norms



The norms of the latent space vectors are relatively stable according to the pattern positions (especially the 2-Norm).