Post-traitement des prévisions d'ensemble de précipitations à Météo-France

Maxime Taillardat

maxime.taillardat@meteo.fr

29 novembre 2018

Machine Learning techniques applied to NWP post-processing

Olivier Mestre[1,2], Michaël Zamo[1,2], Maxime Taillardat[1]
Marion Riverain[1], Harold Petithomme[1]

Météo-France [1]DirOP/COMPAS – [2]CNRM/GMAP

Model Output Statistics

Context : Forecast automation

Post-processing evolution

- Demand for high resolution gridded post-processed data
- Calibration of Ensemble Systems
- Aggregation of forecasts provided by multiple systems

Methods

- Machine learning (Random Forest, Boosting…)
- Efficient spatialisation methods

Tools

- HPC + computation optimisation

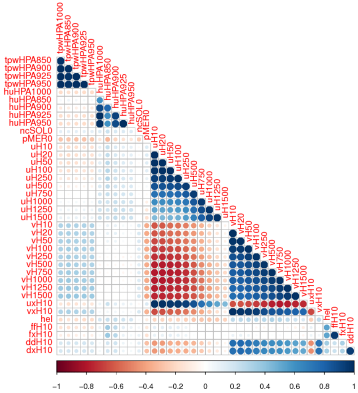
Model Output Statistics

- **Numerical Weather Prediction Model (NWP)**
  - *Pro* : large scale dynamics
  - *Con* : often biased at local scale ; expensive
- **Statistics**
  - *Pro* : unbiased outputs ; cheap
  - *Con* : no dynamics
- **MOS**
  - Modelling statistical relationship between real *in-situ* observations and NWP predictors at the corresponding lead time

    Some information on NWP errors in hidden in NWP outputs

Model Output Statistics

- **NWP Predictors Soup**



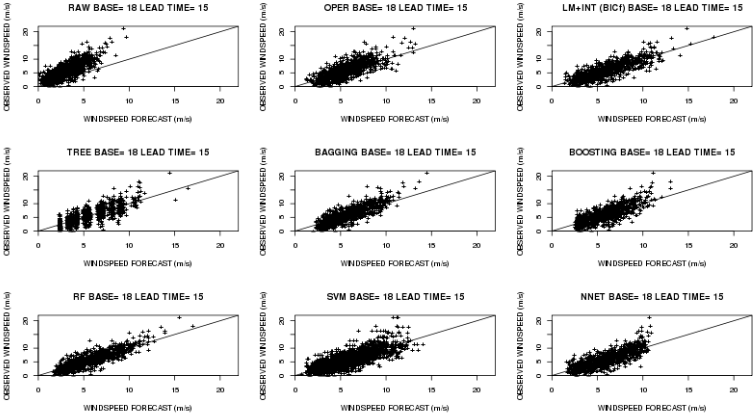
Reduction dimension regularization ?
PCA/PLS/SIR/LASSO ?

Use your knowledge first

building synthetic variables with a physical meaning, vertical gradients, etc.

Application of Machine Learning techniques

- Observed vs Predicted using several machine learning methods
- Avoiding **overfitting** : tuning takes time



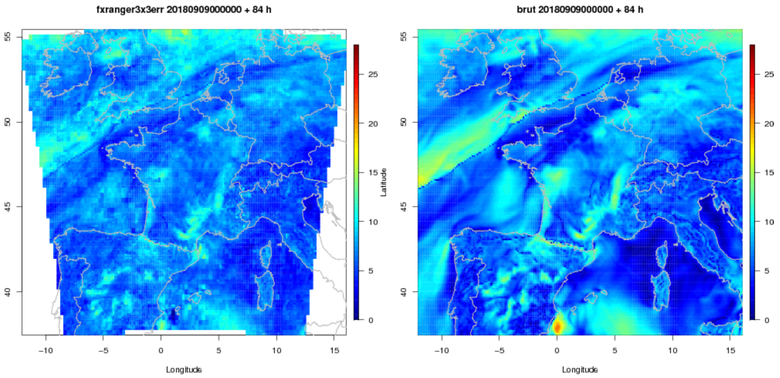
Computing Model Output Statistics

- **Learning**
  - — size of archive ~ several Tbytes
  - — **Avoiding overfitting** : **Tuning** takes time
- **Predicting**
  - — has to be quick (<15')
  - — Limitation : size of I/O
    - ▸ Predictors ~ 0,5 Gbyte
    - ▸ Random Forests ~ 100 Gbyte *after optimization*
- **Parallel computing (R HPC vignette)**
  - — mclapply (openMP) + foreach openMP/MPI framework nested loops

Windspeed/windgusts forecasts

ARPEGE Gridded MOS 01x0.1 grid

- ARPEGE gridded Block-MOS Random Forest
  Zamo *et al.*, 2016, Weather & Forecasting

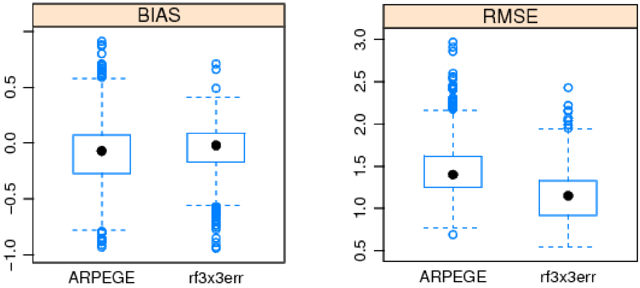


Windspeed/windgusts forecasts

ARPEGE Gridded MOS 01x0.1 grid

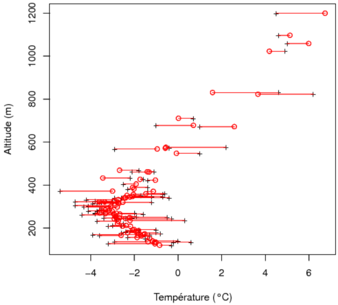
- ARPEGE gridded Block-MOS Random Forest

  Zamo *et al.*, 2016, Weather & Forecasting



Real time spatialization of temperature MOS

- ~1600 MOS for temperature forecasts at station locations
- Spatialization on a 1kmx1km grid (~4e6 gridpoints)
- Principle : similar to regression kriging
  - Multiple linear regression+change-point model (altitude)
  - Kriging replaced by multilevel B-Splines

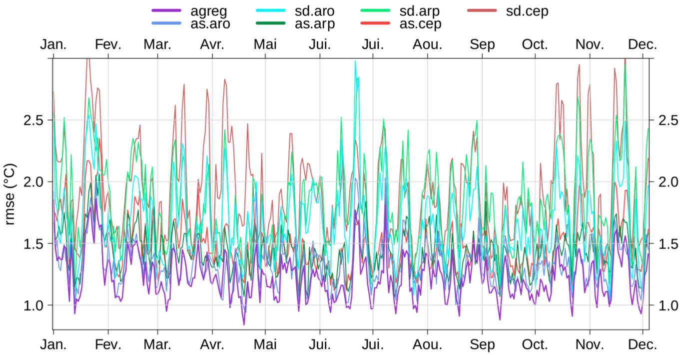


Temperature deterministic forecasts

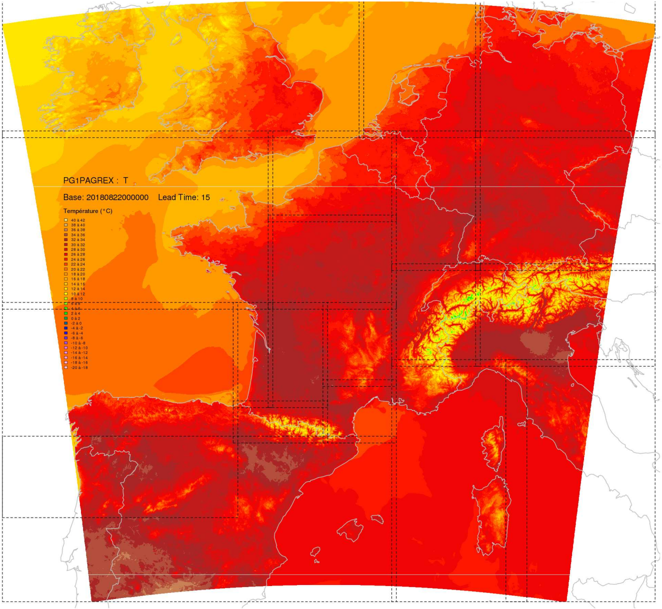
Sequential aggregation of forecasting systems

Spatialisation on AROME 0.01x0.01 grid

- 6 experts : raw and MOS AROME/ARPEGE/IFS
- AROME MOS : Boosting Aggregation algorithm : BOA



0


PG1PAGREX : T
Base: 20180822000000 Lead Time: 15
Température (°C)
40 à 42
38 à 40
36 à 38
34 à 36
32 à 34
30 à 32
28 à 30
26 à 28
24 à 26
22 à 24
20 à 22
18 à 20
16 à 18
14 à 16
12 à 14
10 à 12
8 à 10
6 à 8
4 à 6
2 à 4
0 à 2
-2 à 0
-4 à -2
-6 à -4
-8 à -6
-10 à -8
-12 à -10
-14 à -12
-16 à -14
-18 à -16
-20 à -18

Conclusion – Current research

- **Usefulness of NWP post-processing**

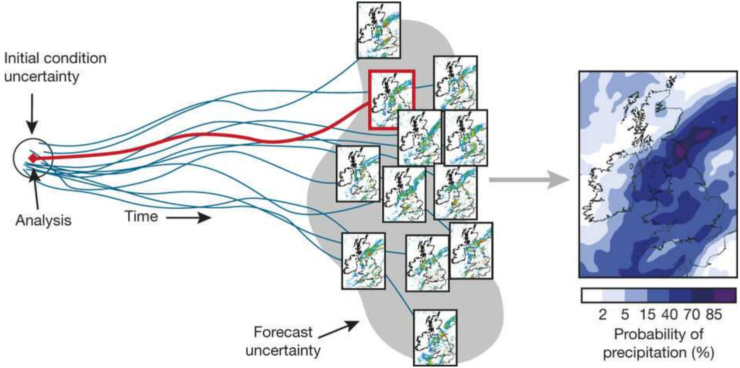
- **Current research**
  - — Calibration of PEAROME ensemble rainfall amounts + reconstruction of ensemble members
  - — Calibration of other elements (visibility,…)
  - — Usefulness of Deep-Learning on NWP post-processing and Nowcasting

    Deep4Cast project
    - Involving DirOP/CERFACS/IRT/UPS (coord. Michaël Zamo)
    - 4 post-doctoral positions
    - Deep-Learning applied to various forecasting problems

Principle of ensemble forecasting

Ensemble forecasting

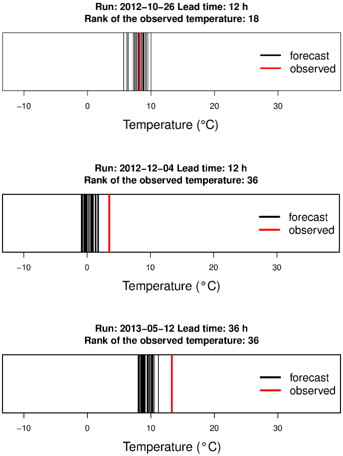
- Sampling the uncertainty in both initial conditions and physics representations
- Running several ensemble **members**
- Producing an estimate of the forecast uncertainty



Motivation of post-processing ensemble forecasts

Météo-France's 35-members global ensemble system (PEARP), 10 km resolution over France.

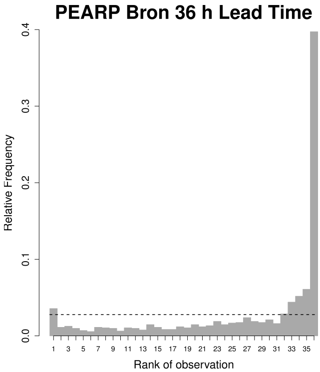
Observations and forecasts of 2-m temperature (T2m) at Lyon-Bron for the run of 1800UTC (different lead times)



Motivation of post-processing ensemble forecasts

Météo-France's 35-members global ensemble system (PEARP), 10 km resolution over France.

Observations and forecasts of 2-m temperature (T2m) at Lyon-Bron for the run of 1800UTC (different lead times)



- Need of a **simultaneous** correction of bias and dispersion
- Whatever the quality of the raw ensemble, post-processing improves forecast attributes (Hemri, 2014)

The state of the art

- Existing methods: Analogs, Perfect Prog., Rank-based matching, CDF-matching, Member dressing, Bayesian Model Averaging, SAMOS...
- A review of techniques: Gneiting, 2014

The most (famous // widely-used // efficient) post-processing method :

- **Ensemble model output statistics (EMOS)** (Gneiting, 2005) also called Non-homogeneous Regression (NR)
  - fitting parameters linearly on predictors on some training period:

$$f(y|x_1,\ldots,x_N) = \mathcal{N}(\mu = a_0 + \sum_{i=1}^{N} a_i x_i, \sigma^2 = b + cs^2)$$

y response variable, $x_1,\ldots,x_N$ ensemble member values or any other predictor, s^2 ensemble variance

Benefits of non-parametric post-processing

No assumptions on the weather variable you deal with

- Raw ensemble (not post-processed) :

EMOS

Non-parametric

Benefits of non-parametric post-processing

No assumptions on the weather variable you deal with

- Raw ensemble (not post-processed) :

EMOS

- EMOS post-processing :

Non-parametric

Benefits of non-parametric post-processing

No assumptions on the weather variable you deal with

- Raw ensemble (not post-processed) :

EMOS

- EMOS post-processing :

Non-parametric

- Post-processing possible results :

Benefits of non-parametric post-processing

No assumptions on the weather variable you deal with

- Raw ensemble (not post-processed) :

EMOS

- EMOS post-processing :

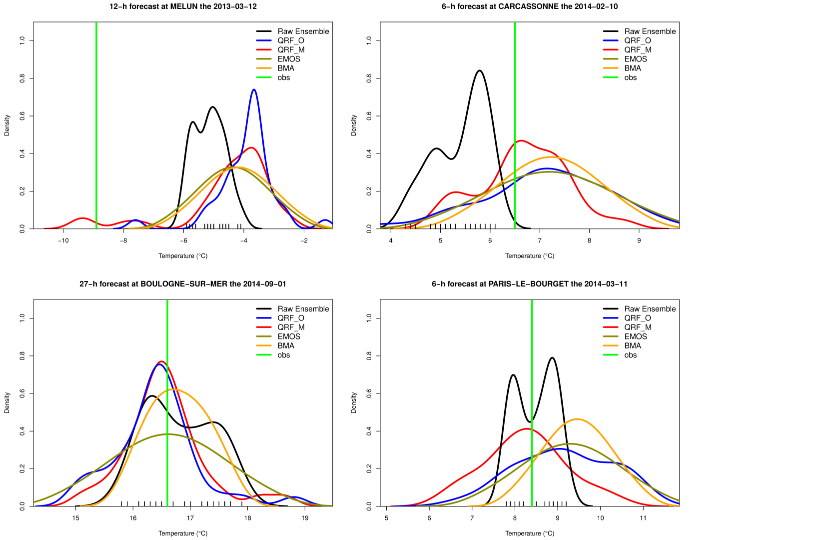
Non-parametric

- Post-processing possible results :

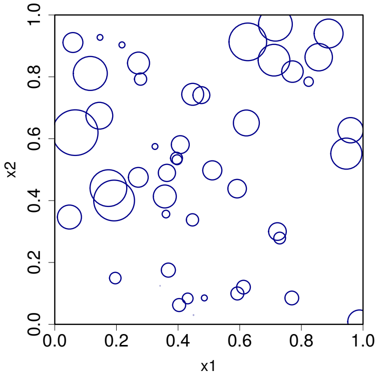




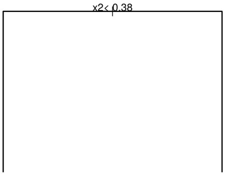
Best-of post-processing possible outputs

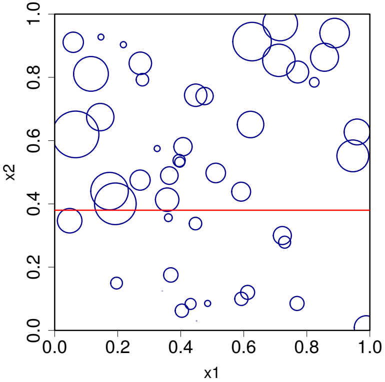


- Example: CART algorithm (Breiman, 1984)



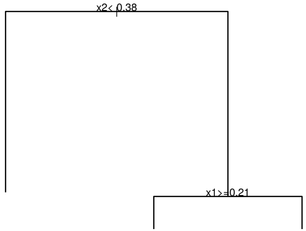
- Example: CART algorithm (Breiman, 1984)

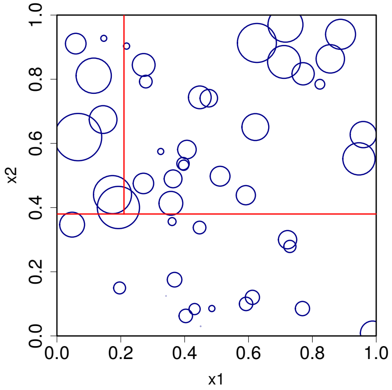




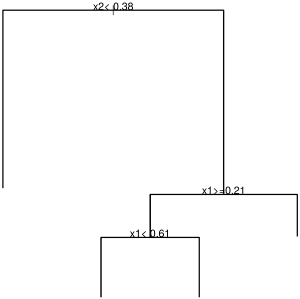
From binary regression trees...

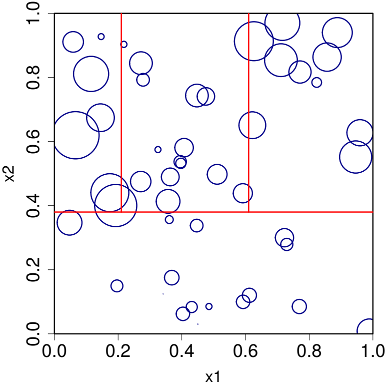
- Example: CART algorithm (Breiman, 1984)





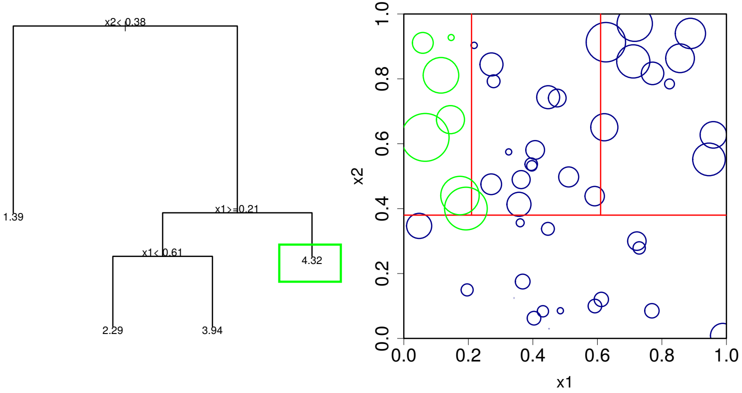
- Example: CART algorithm (Breiman, 1984)





From binary regression trees...

- Example: CART algorithm (Breiman, 1984)



...to Random Forest...and Quantile Regression Forests

- Bagged trees are too much correlated
- **Random Forest** (RF): Add randomness in tree node's selection (Breiman, 2001)

Example:

- In CART: at each node, split trying all predictors
- In CART-RF: at each node, split trying only a random subset of predictors

- Each tree is suboptimal, but trees are less correlated.

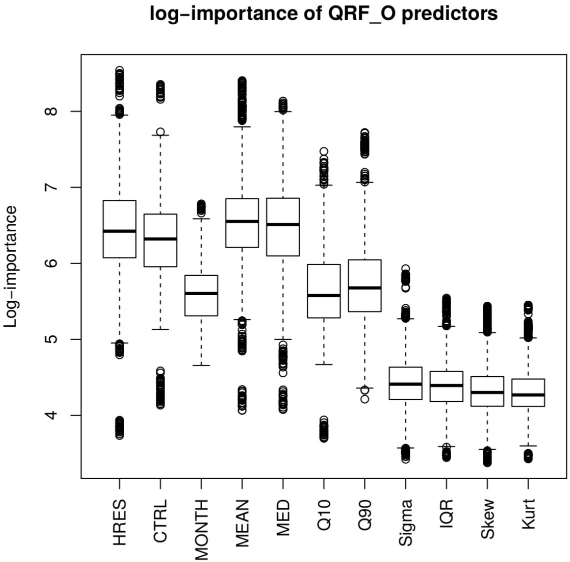
In RF:

Take the mean of final leaves

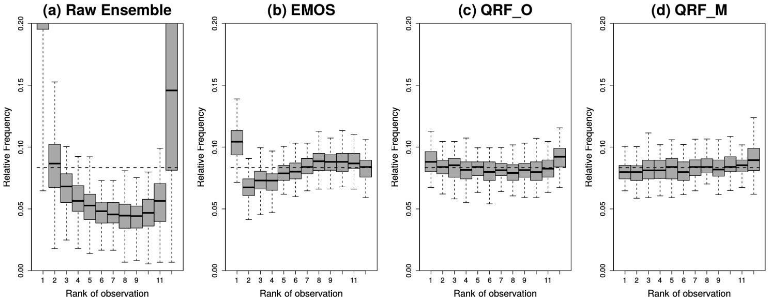
In QRF:

Take the empirical distribution of final leaves

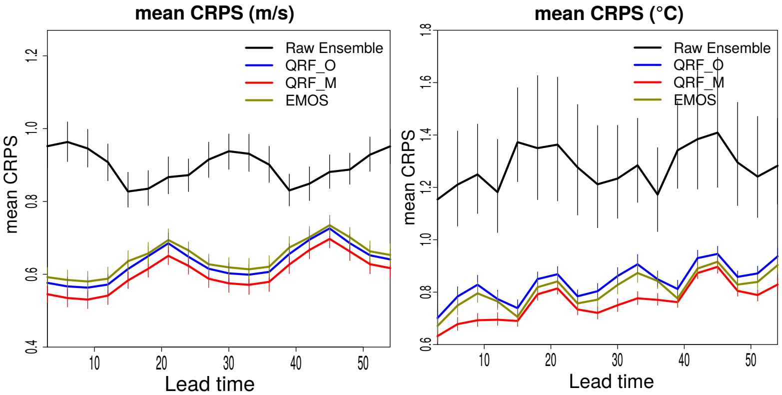
Importance of variables for wind speed, 24-h lead time



Box-Rank histograms for wind speed, 36-h lead time



CRPS for wind speed (left) and temperature (right)



To sum up

- For T2m and FF10m, QRF (and especially QRF with multi-variable predictors) is better (in terms of CRPS) than the "classic" EMOS method
- QRF takes into account some non-linear phenomenas, multimodal distributions, skewed distributions etc.
- QRF performs well also for variables like RR24 (daily rainfall) or TCC (Total Cloud Cover)

- Taillardat, Maxime, Olivier Mestre, Michaël Zamo, and Philippe Naveau. "Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics." *Monthly Weather Review* 144, no. 6 (2016): 2375-2393.

Interest in post-processing RR6 (6-h rainfall) but:

- This variable is (by far !) the most difficult to calibrate whatever the method, a really hard nut to crack for statisticians
- A lot of zeros, a lot of "extremes"

What we have tested

For EMOS:

- Try different distributions "tailored" for extremes (Gamma, GEV...)
- Try different sets of predictors

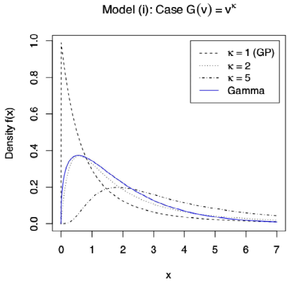
For QRF:

- Improve QRF for "quantiles": Gradient Forests (GF) (Athey et al., 2017)
- **Extrapolation in distributions tails: semi-parametric approach**

Goal: Be skillful for extremes without degrading overall performance
Use QRF/GF outputs to fit a distribution which would:

- Model jointly low, moderate and heavy rainfall
- Be flexible
- Be simple

- Use of an Extended GP distribution (EGP3) (Papastathopoulos and Tawn, 2013 ; Naveau et al., 2016)



A semi-parametric approach

Our final distribution is:

$$G(x) = f_0 + (1 - f_0)\left[1 - \left(1 + \frac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}}\right]^{\kappa}$$

Strategy

1. Run QRF/GF to get $\widehat{F}(y|X = x) = \widehat{\mathbb{P}}(Y \leq y|X = x)$
2. Keep the probability of no rain $\widehat{f_0} = \widehat{\mathbb{P}}(Y = 0|X = x)$ from QRF/GF outputs
3. Estimate $(\widehat{\kappa}, \widehat{\sigma}, \widehat{\xi})$ from non-zero QRF/GF quantiles

What we expect:

A semi-parametric approach

Our final distribution is:

$$G(x) = f_0 + (1 - f_0)\left[1 - \left(1 + \frac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}}\right]^{\kappa}$$

Strategy

1. Run QRF/GF to get $\widehat{F}(y|X = x) = \widehat{\mathbb{P}}(Y \leq y|X = x)$
2. Keep the probability of no rain $\widehat{f_0} = \widehat{\mathbb{P}}(Y = 0|X = x)$ from QRF/GF outputs
3. Estimate $(\widehat{\kappa}, \widehat{\sigma}, \widehat{\xi})$ from non-zero QRF/GF quantiles

What we expect:

- QRF/GF possible output

A semi-parametric approach

Our final distribution is:

$$G(x) = f_0 + (1 - f_0)\left[1 - \left(1 + \frac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}}\right]^{\kappa}$$

Strategy

1. Run QRF/GF to get $\widehat{F}(y|X = x) = \widehat{\mathbb{P}}(Y \leq y|X = x)$
2. Keep the probability of no rain $\widehat{f_0} = \widehat{\mathbb{P}}(Y = 0|X = x)$ from QRF/GF outputs
3. Estimate $(\widehat{\kappa}, \widehat{\sigma}, \widehat{\xi})$ from non-zero QRF/GF quantiles

What we expect:

- QRF/GF possible output

- After "EGP TAIL"

Data and model fitting for RR6

- PEARP (ARPEGE Ensemble Prediction System)
- 4 years from 2012 to 2015
- 87 French SYNOP stations
- Lead time: 51 hours
- initialization: 1800UTC
- For EMOS: GEV and Censored/Shifted Gamma (CSG) and EGP3 distributions.
  Predictors selection: AIC and VSURF (Genuer, Poggi and Tuleau, 2014)
- For Analog Method: Delle-Monache metric:

$$d_t = \sum_{p=1}^{N_p} \frac{w_p}{\sigma_{f^p}} \sqrt{\sum_{k=-t_l}^{+t_l} \left(f^p_{t+k} - g^p_{t+k}\right)^2}$$

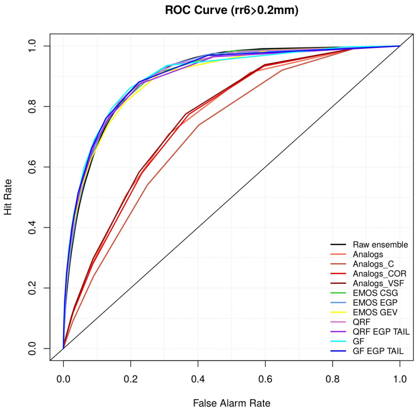
  Different strategies in the weightning of predictors w_p.
- QRF and GF techniques with and without EGP3 fit.

Results on RR6

Table : Mean CRPS. Bootstrap estimation error under 6.1×10^{-3} for all methods.

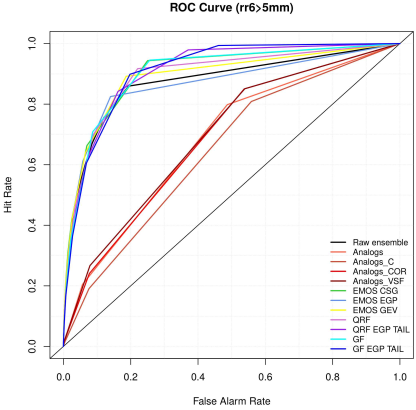
Types	Methods	pdf	CRPS
	Raw ensemble		0.4694
Non-parametric	Analogs		0.5277
	Analogs_C		0.5376
	Analogs_COR		0.5276
	Analogs_VSF		0.5247
	QRF		**0.4212**
	GF		**0.4134**
Parametric with covariates	EMOS	CSG	**0.4224**
	EMOS	GEV	0.4228
	EMOS	EGP	0.4292
Hybrid	QRF	EGP TAIL	**0.4138**
	GF	EGP TAIL	**0.4127**

ROC curves



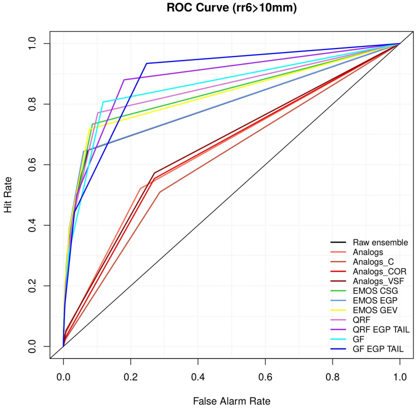
- Value: Focus on the upper left corner
- Semi-parametric methods value increases with RR6 threshold.

ROC curves



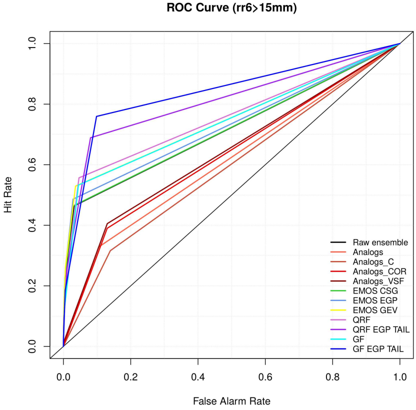
- Value: Focus on the upper left corner
- Semi-parametric methods value increases with RR6 threshold.

ROC curves



- Value: Focus on the upper left corner
- Semi-parametric methods value increases with RR6 threshold.

ROC curves



- Value: Focus on the upper left corner
- Semi-parametric methods value increases with RR6 threshold.

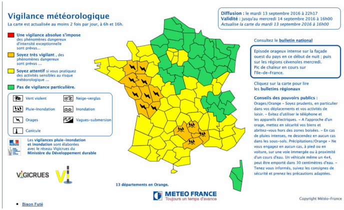
To sum up

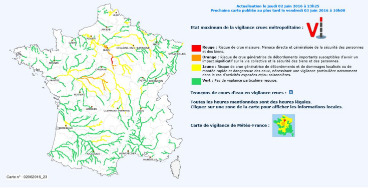
- Even without extrapolations in distributions, QRF and GF competes favourably with EMOS techniques
- Semi-parametric approaches beat EMOS and improve forecasts, especially forecast value for high thresholds

The flood forecasting system in France

- 1 Support Service (SCHAPI) managing 21.000 km watercourse
- 22 Flood Forecasting Services (SPCs)

Production of vigilance maps for flood twice/day (Vigicrues), relayed by Météo-France's warning system





The forecast chain and the need of post-processing

Rainfall forecast $\longrightarrow$ Hydrological forecast $\longrightarrow$ waterflow forecast $\Longrightarrow$ Threshold

This chain is currently **deterministic**

An on-going PhD work (A-L Tibéri-Wadier, CEREMA) has to convert the chain into an **ensemblist** one.

- Need of hourly rainfall scenarios : ensemble forecasts (with members information)
- Feasability study : Do classical post-processing methods can be applied to such a variable ?

Post-processing strategy

We have to provide a post-processing strategy which would be :

1. Simple to set up
2. Fast to run and tailored for the hydrological model

- Work with short datasets and predictors
- Work on a watersheds' scale (kind of block-EMOS)

Case study

- Work on 3 watersheds in Brittany (global size : 720 km^2)
- Météo-France's LAM-EPS PEAROME (12 members, grid scale : 2.5 km), base time : 21UTC, lead times from 1 to 45 hours
- Use of hourly calibrated (over rain gauges) radar precipitation observations ANTILOPE (grid scale : 1 km)
- Data spans from 12/2015 to 03/2016 and from 05/2016 to 06/2016
- Predictors : statistics on the hourly rainfall ensemble, and on surface humidity and temperature

Methods involved

Non homogeneous regression (NGR/EMOS) (Gneiting et al., 2005)

- Use of censored Gamma (CGAMMA) and censored Normal (CNORM) distributions
- Each lead time is treated separately

Quantile Regression Forests (QRF) (Taillardat et al., 2016 ; Whan and Schmeits, 2018)

- Add of a predictor ("MORNING","AFTERNOON","EVENING","NIGHT") : all lead times are treated by the same forest.

Semi-parametric Quantile Regression Forests (QRF_TAIL) (Taillardat et al., 2017)

- Like QRF, but the tail of the distribution is fitted from QRF output

Ensemble Copula Coupling (ECC) (Schefzik et al., 2013)

- Restores physical consistency between grid points and time steps

For each time step, each watershed :

Result of calibration : 1 distribution for the watershed $\xrightarrow{ECC}$ 12 members for each grid point

ECC in a nutshell

Lyon

Membres bruts :

2°C	7°C	6°C	3°C

Rang des membres :

1er	4eme	3eme	2eme

Valeurs calibrées :

1°C	5°C	7°C	9°C

Membres calibrés :

1°C	9°C	7°C	5°C

Saint-Etienne

Membres bruts :

3°C	8°C	7°C	9°C

Rang des membres :

1er	3eme	2eme	4eme

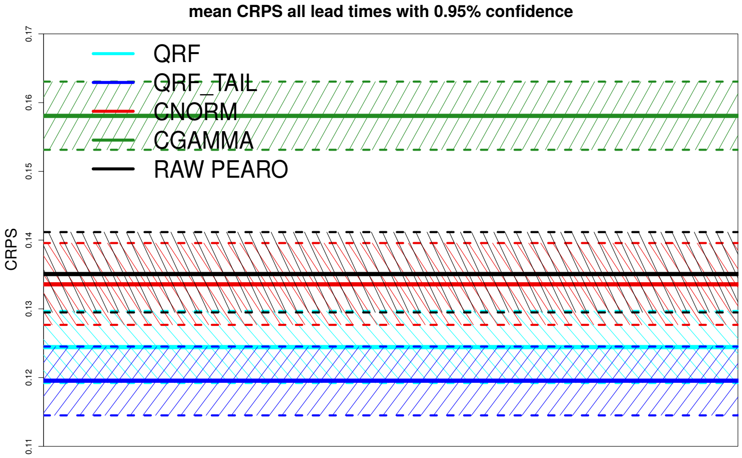
Valeurs calibrées :

2°C	4°C	8°C	10°C

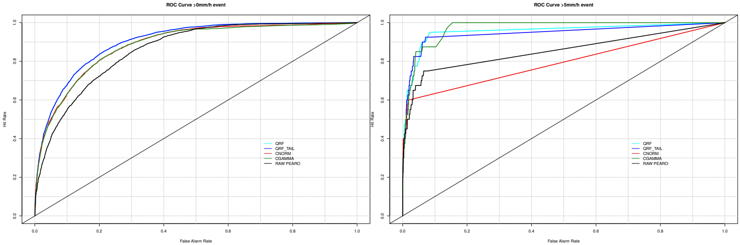
Membres calibrés :

2°C	8°C	4°C	10°C

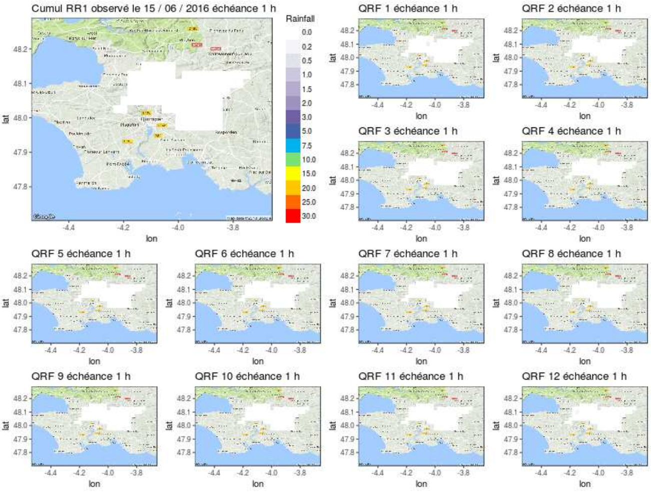
Leave-one-out cross validation results



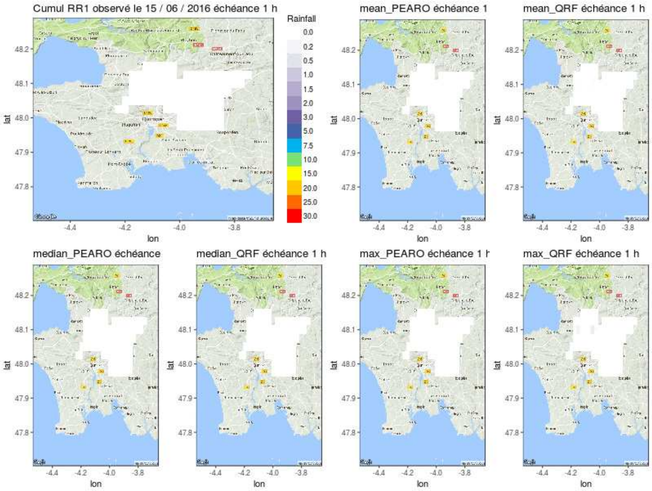
ROC curves


ROC Curve >0mm/h event
Hit Rate
False Alarm Rate
QRF
QRF_TAIL
CNORM
CGAMMA
RAW PEARO
ROC Curve >5mm/h event
Hit Rate
False Alarm Rate
QRF
QRF_TAIL
CNORM
CGAMMA
RAW PEARO

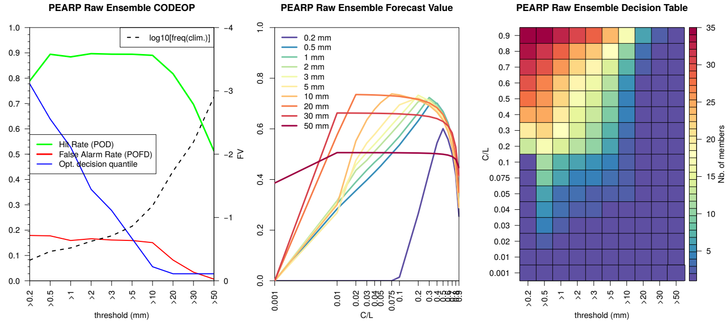
ECC + post-processing visualization



Summary QRF vs. raw ensemble



The CODOP



PEARP Raw Ensemble CODEOP
log10[freq(clim.)]
Hit Rate (POD)
False Alarm Rate (POFD)
Opt. decision quantile
threshold (mm)
FV
PEARP Raw Ensemble Forecast Value
0.2 mm
0.5 mm
1 mm
2 mm
3 mm
5 mm
10 mm
20 mm
30 mm
50 mm
C/L
PEARP Raw Ensemble Decision Table
C/L
threshold (mm)
Nb. of members

References

Taillardat, Maxime, Olivier Mestre, Michaël Zamo, and Philippe Naveau. "Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics." *Monthly Weather Review* 144, no. 6 (2016): 2375-2393.

Taillardat, Maxime, Anne-Laure Fougères, Philippe Naveau, and Olivier Mestre. "Forest-based and semi-parametric methods for the postprocessing of rainfall ensemble forecasts." *Submitted to Weather and Forecasting* (2019). HAL preprint <hal-01643954>. ArXiv preprint 1711.10937

METEO FRANCE

References

Gneiting, T., A. E. Raftery, A. H. Westveld III, and T. Goldman, 2005: Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. *Monthly Weather Review*, **133 (5)**, 1098–1118.

Schefzik, R., T. L. Thorarinsdottir, T. Gneiting, and Coauthors, 2013: Uncertainty quantification in complex simulation models using ensemble copula coupling. *Statistical Science*, **28 (4)**, 616–640.

Whan, K., and M. Schmeits, 2018: Comparing area-probability forecasts of (extreme) local precipitation using parametric and machine learning statistical post-processing methods. *Monthly Weather Review*, **(2018)**.