Impact of the statistical method, training dataset, and spatial scale of post-processing to adjust ensemble forecasts of the height of new snow

Jari-Pekka Nousu, Matthieu Lafaysse, Guillaume Evin, Matthieu Vernay, Joseph Bellier, Bruno Joly, Maxime Taillardat, Mickaël Zamo

METEO FRANCE

Impact of the statistical method, training dataset, and spatial scale of post-processing to adjust ensemble forecasts of the height of new snow

Jari-Pekka Nousu[1,2], Matthieu Lafaysse[1], Guillaume Evin[3]
Matthieu Vernay[1], Joseph Bellier[4,5], Bruno Joly[6],
Maxime Taillardat[6,7], Mickaël Zamo[6,7]

1 Univ. Grenoble Alpes – Université de Toulouse – Météo-France – CNRS – CNRM, Centre d'Etudes de la Neige, Grenoble, France
2 University of Oulu, Finland ; 3 Univ. Grenoble Alpes – INRAE, UR ETNA, Grenoble, France ; 4 NOAA Earth System Research Lab., Boulder, USA ; 5 Univ. Grenoble Alpes, CNRS, IRD, Grenoble INP, IGE, Grenoble, France ; 6 CNRM – Université de Toulouse – Météo-France – CNRS, GMAP, Toulouse, France, 7 Météo-France, DirOP/COMPAS, Toulouse, France

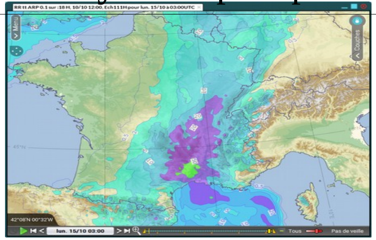
Context

- Forecasting the height of new snow:
  - Safety and economic concerns

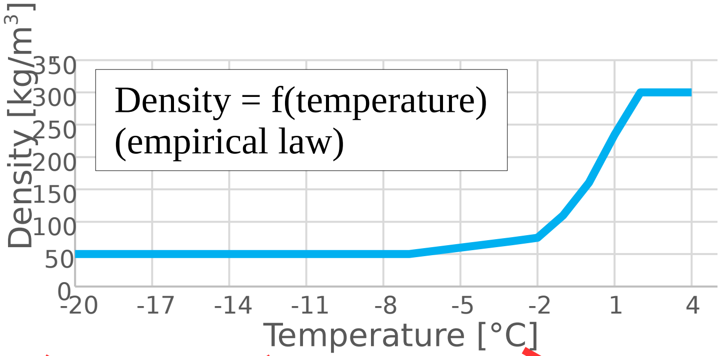
- Meteo-France **automatic forecasts** currently available (website and smartphone apps) :

24h height of new snow =

$\sum_{24h}$ (**Hourly** NWP precipitation



×

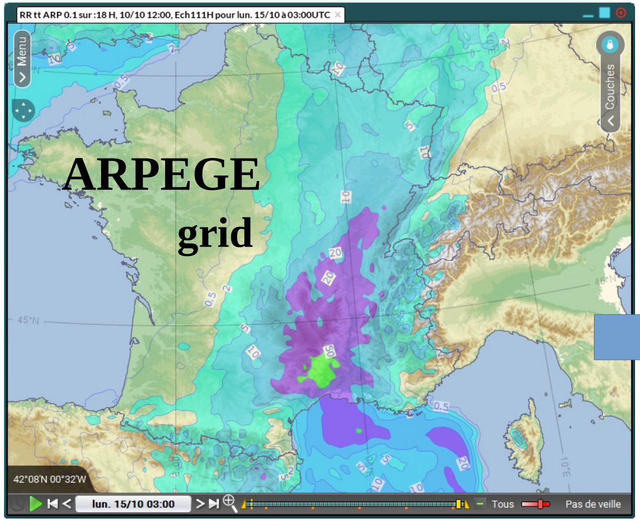


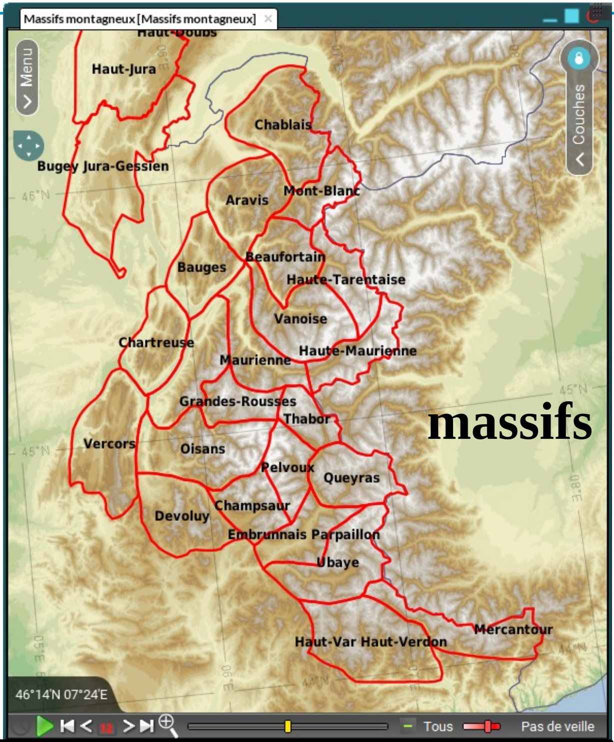
)

~~Compaction during snowfall~~ **~~Possible melting~~** **~~Rain on snow~~**

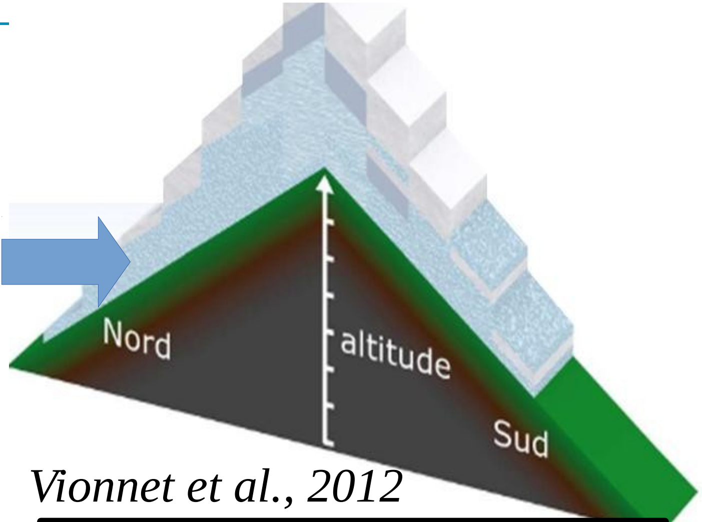
→ **Severe biases**

Alternative : Physical modelling SAFRAN-Crocus





Durand et al., 1998



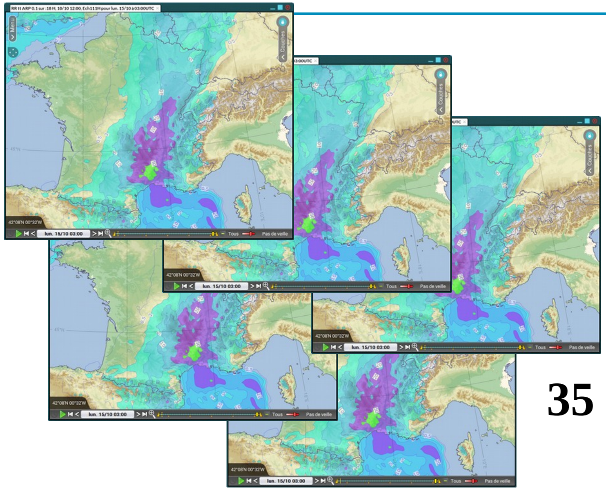
Vionnet et al., 2012

SAFRAN :
- Spatial aggregation of ARPEGE on *massifs* (~1000 km²)
- Adjust meteorological variables at various elevations

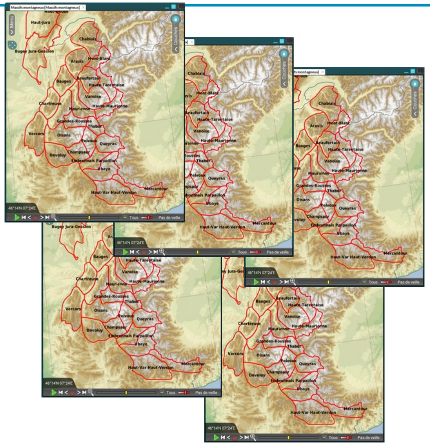
Crocus :
- Falling snow density = f(temperature, wind speed)
- Explicit mechanical compaction
- Melting (energy balance)
- Compaction due to liquid water (rain on snow)

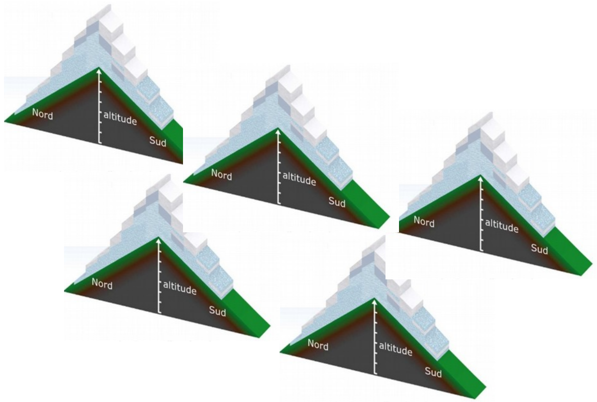
Ensemble forecasts PEARP-S2M

Experimental from 2014
Operational : october 2019



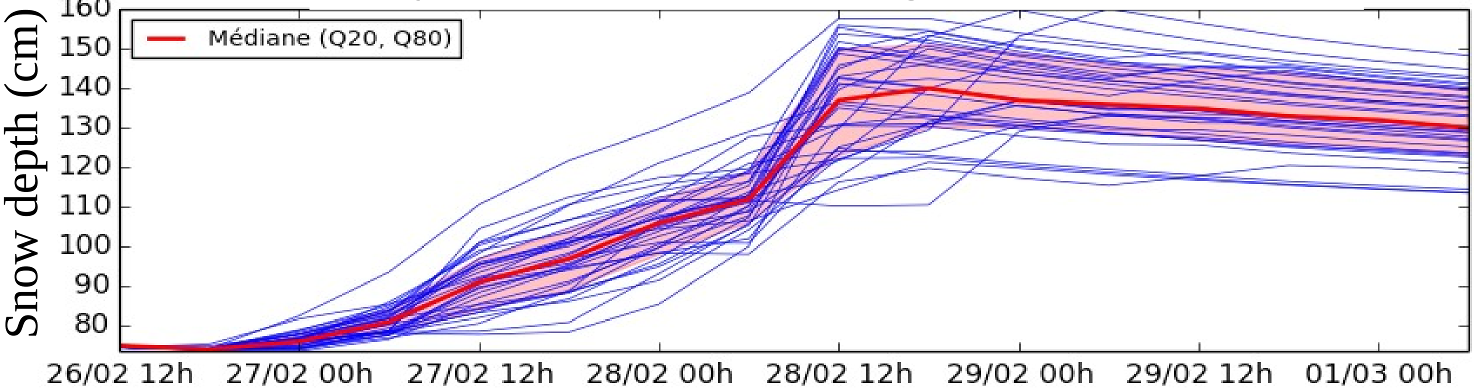
35 members





PEARP **SAFRAN** **Crocus**

4-day forecast, Mercantour region, 2100 m elevation



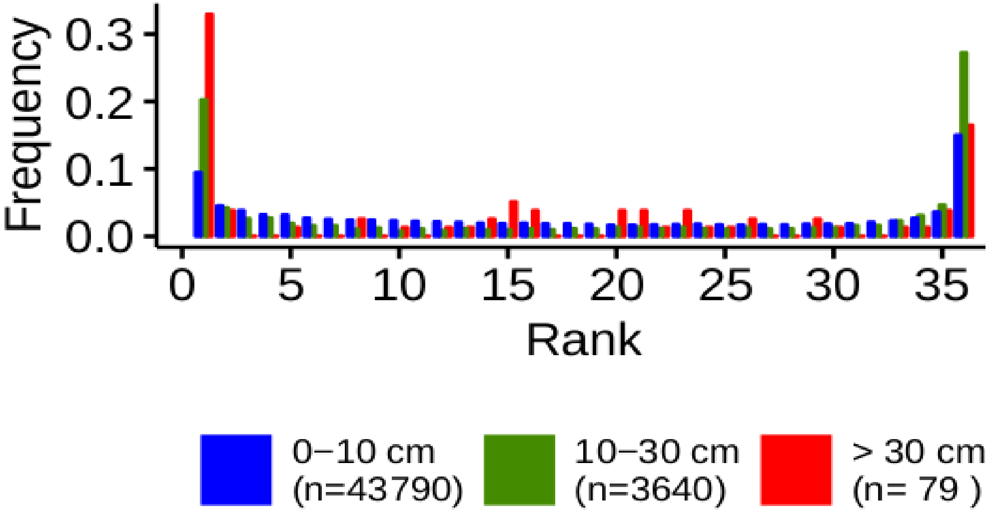
Vernay et al. 2015

METEO FRANCE

Raw ensemble forecasts PEARP-S2M

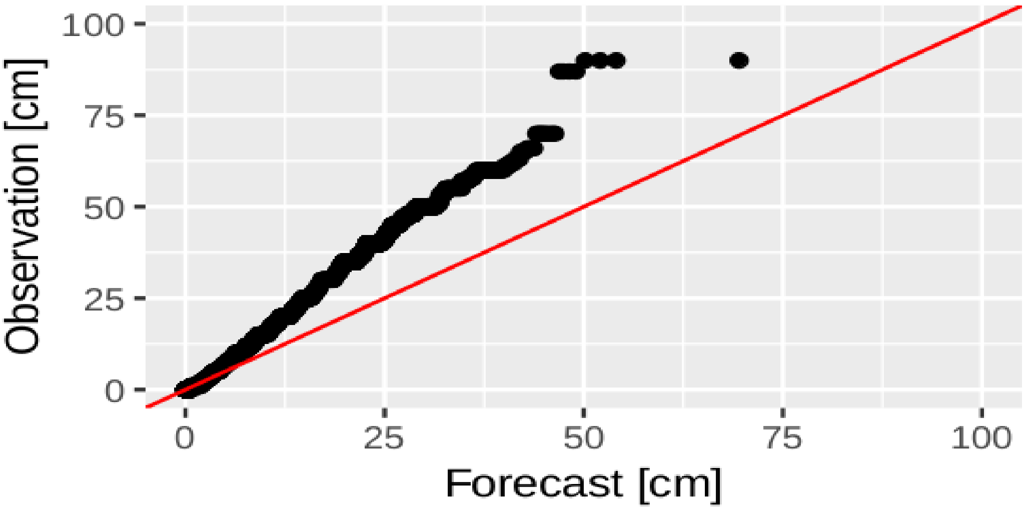
Evaluation over all massifs
Winter 2017-2018

Rank diagram



Underdispersion !

Quantile-Quantile plot



Bias !

State of the art

- Physical ensemble modelling of the snowpack improves the forecast of the height of new snow compared to:
  - **Direct NWP outputs** (*Champavier et al., 2018*)
  - **Deterministic systems** (*Vernay et al., 2015*)

- **Ensemble Model Output Statistics (EMOS)** are useful to forecast the height of new snow from direct ensemble NWP outputs (precipitation and temperature) (*Stauffer et al., 2018 ; Scheuerer and Hamill, 2019*)
- **Quantile Regression Forests (QRF)** can incorporate more predictors and have added value for precipitation forecasts *(Taillardat et al., 2019)*

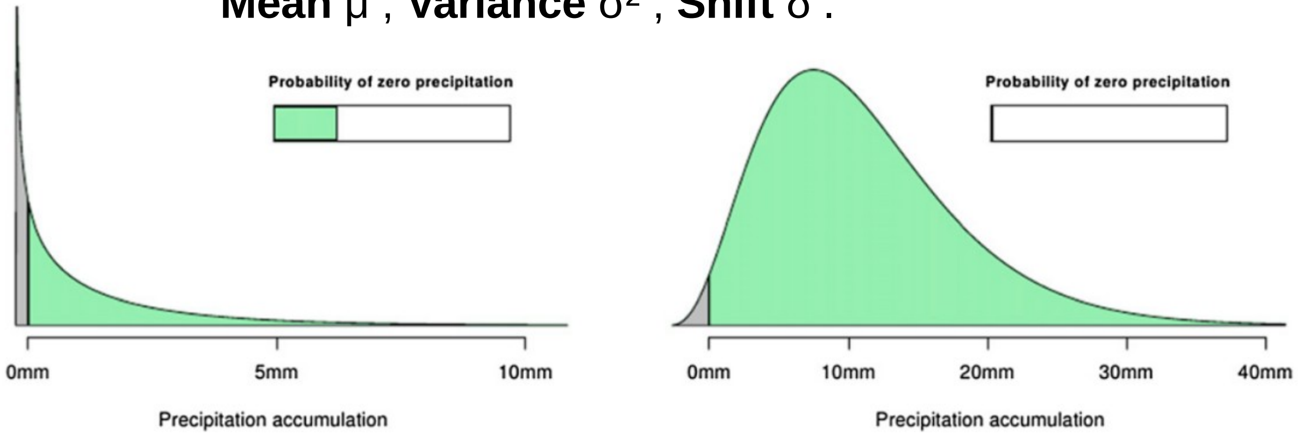
Questions

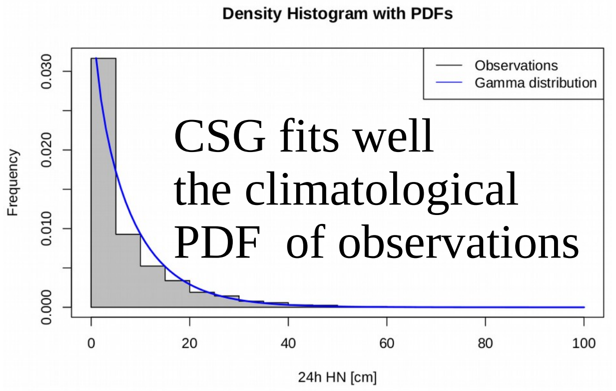
- Can **Ensemble Model Output Statistics (EMOS)** improve the forecasts from physical modelling ?
  - What is the best training dataset ?
  - What is the spatial validity of the post-processing ?
- Can **Quantile Regression Forests (QRF)** improve the skill compared to EMOS ?

METEO FRANCE

Statistical post-processing: method

- In *Nousu et al., NPG, 2019*, we apply the EMOS method used by *Scheuerer and Hamill (2015 ; 2018)* for precipitation forecasts:
  - We assume that the conditional distribution of the forecast HN to the raw ensemble forecasts follow a Censored Shifted Gamma (CSG) defined by 3 parameters : **Mean** μ ; **Variance** σ² ; **Shift** δ .





  - Regression model between CSG parameters and synthetic properties of the raw ensemble (**mean, dispersion, probability of 0 cm**)

METEO FRANCE

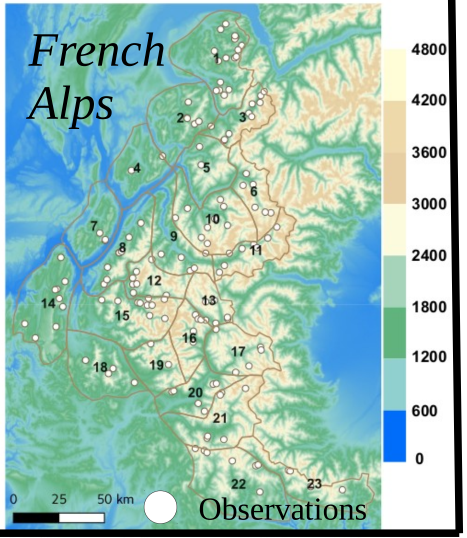
Statistical post-processing: calibration

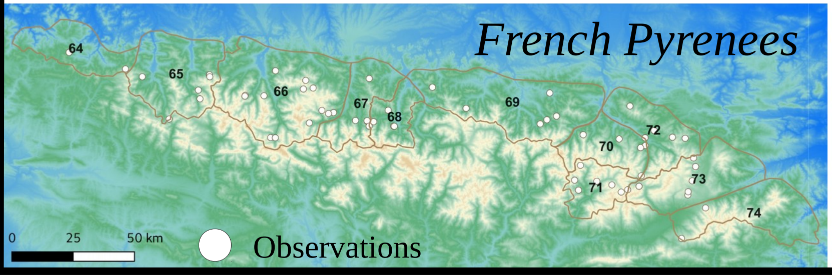
Nousu et al., NPG, 2019

Predictand :
Network of local observations of the 24h height of new snow

2 Predictor datasets: Ensemble forecasts PEARP-S2M

	Period	Members	Initial conditions	Resolution and physics
Reforecast	1994-2016	10	Unperturbed	Homogeneous
Real-time forecasts	2014-2017	35	Perturbed	Heterogeneous





Spatial scale of the calibration:
- Massif scale
- Station scale

Evaluations :

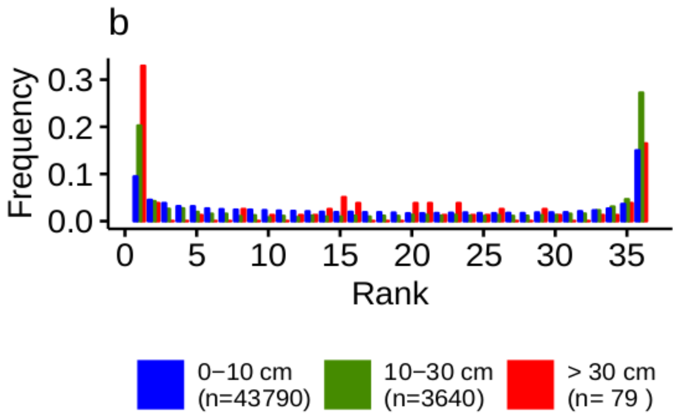
- From real-time forecasts, winter **2017-2018**

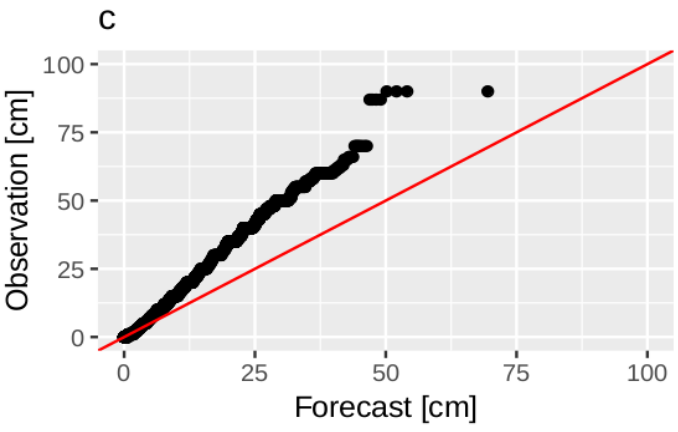
EMOS: results

Nousu et al., NPG, 2019

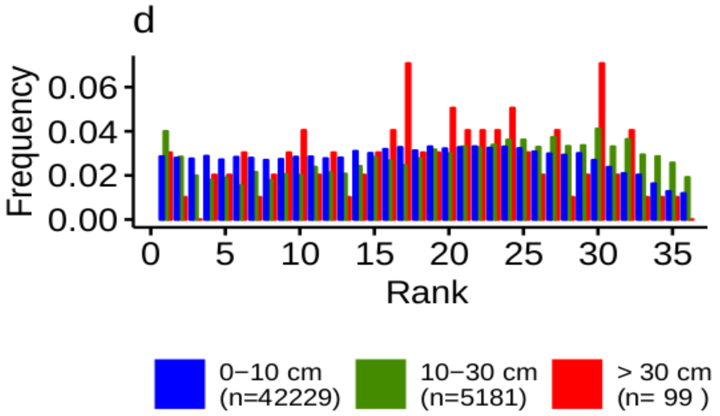
Training: reforecasts 1994-2016
Evaluation: real-time forecasts 2017-2018

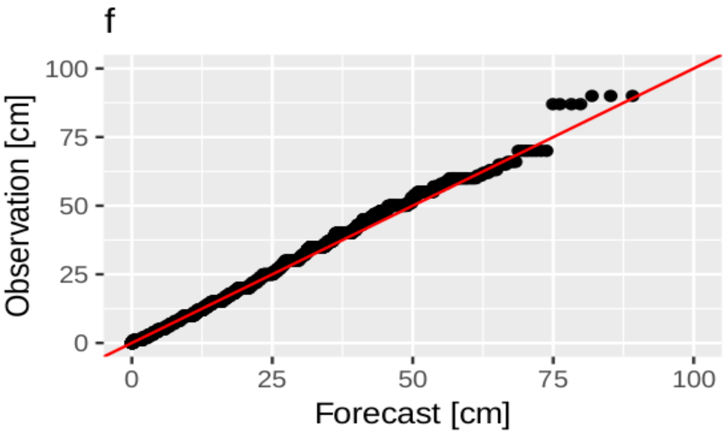
Raw forecasts



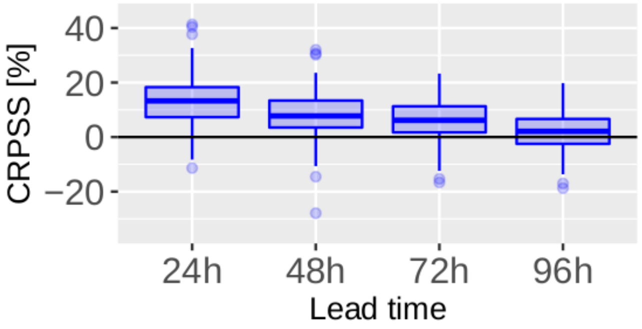


Corrected forecasts





CRPSS (Reference : raw forecasts)



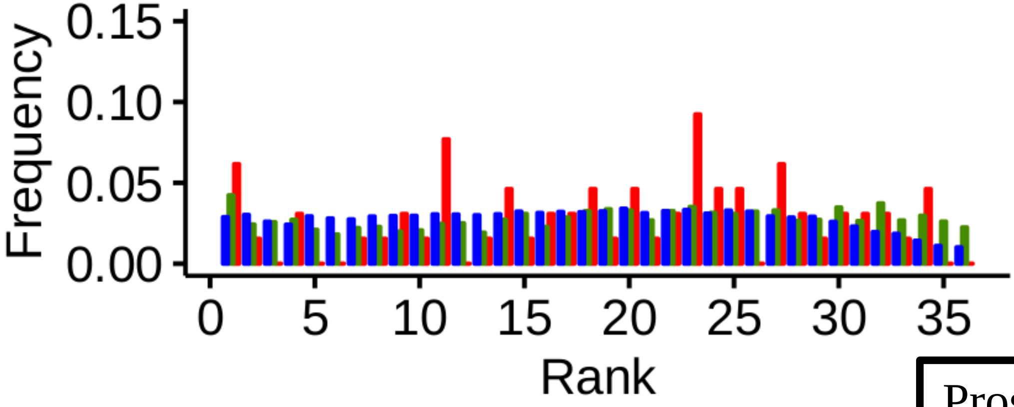
- **Remove bias and underdispersion**
- **Improvement of CRPS** on most stations.
- Larger improvement at short lead times

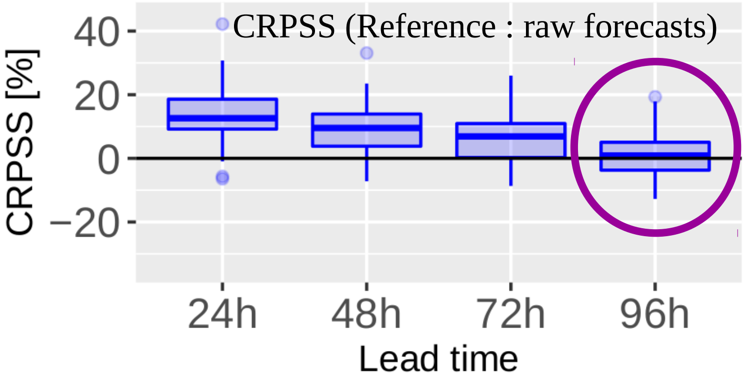
Sensitivity to training dataset

Nousu et al., NPG, 2019

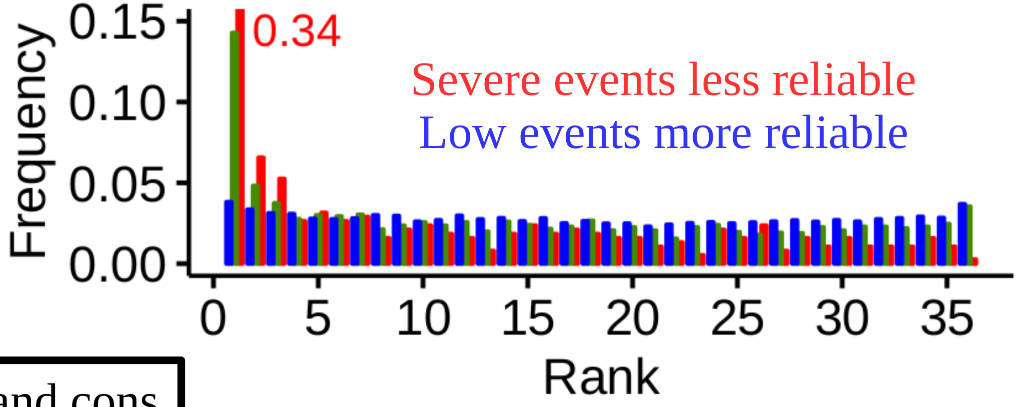
0–10 cm (n=18728) | 10–30 cm (n=2763) | > 30 cm (n= 65)

Training: reforecasts 1994-2016
Evaluation: real-time forecasts 2017-2018



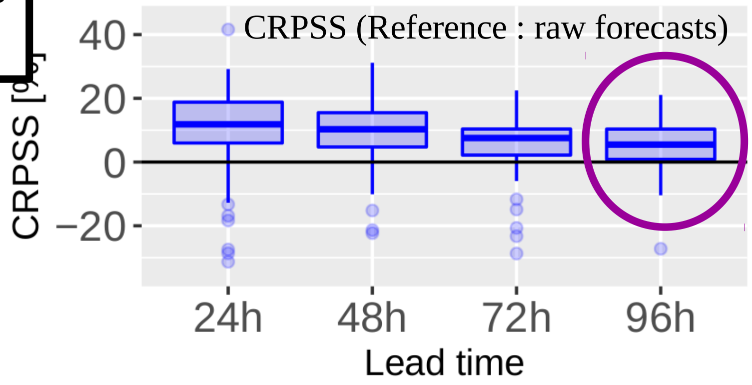


Training: real-time forecasts 2014-2017
Evaluation: real-time forecasts 2017-2018



Severe events less reliable
Low events more reliable

Pros and cons on both sides



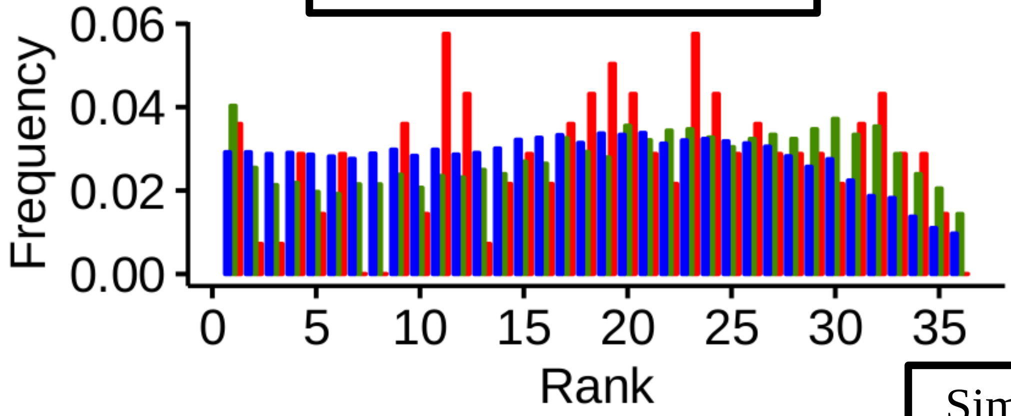
Better skill at the longest lead time

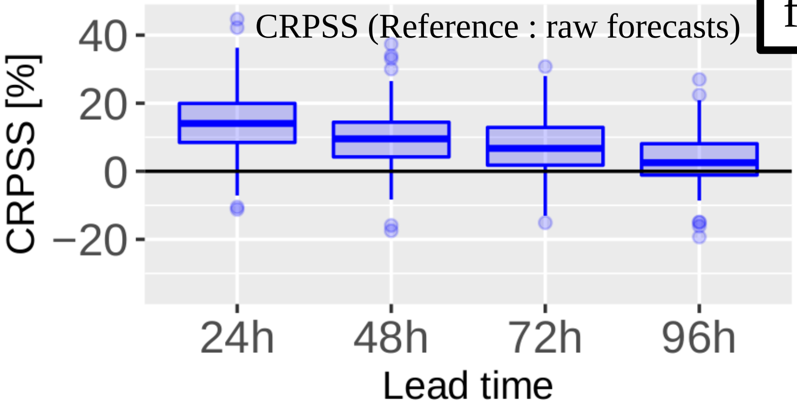
Sensitivity to spatial scale

Nousu et al., NPG, 2019

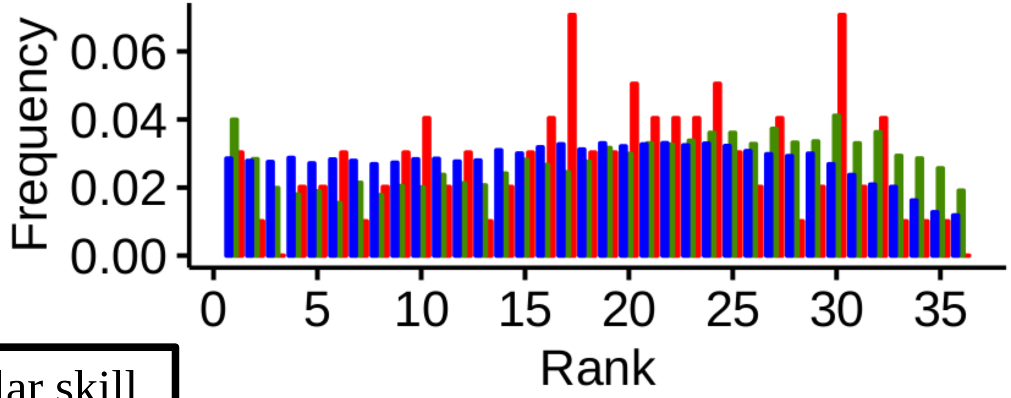
0–10 cm (n=18728)
10–30 cm (n=2763)
> 30 cm (n= 65)

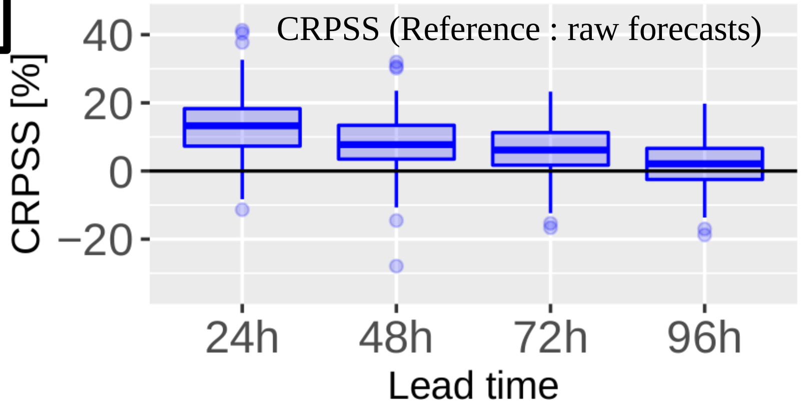
Training: <u>local</u> scale
Evaluation: local scale





Training: <u>massif</u> scale
Evaluation: local scale



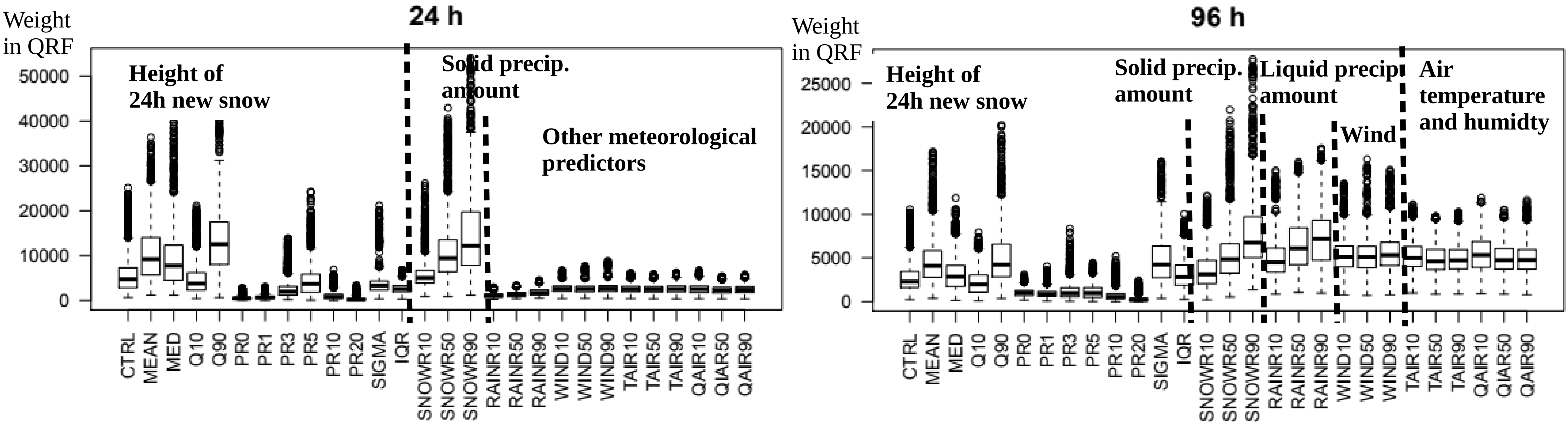


Similar skill for all criteria

Added value of Quantile Regression Forests (QRF)

Evin et al., in prep.

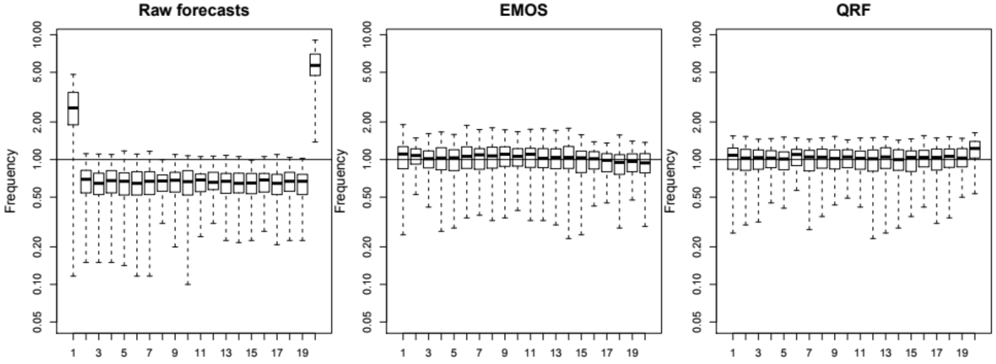
- Limitation of EMOS :
  - When all raw members expect 0 cm of snow but some rainfall, EMOS always forecast 0 cm (it does not account for potential errors in the rain-snow limit elevation)
- **QRF** has been tested with a large set of variables as predictors
  - It is shown that rainfall amount and temperature are useful predictors to be associated with the simulated new snow depth, **<u>especially at the longest lead times</u>**



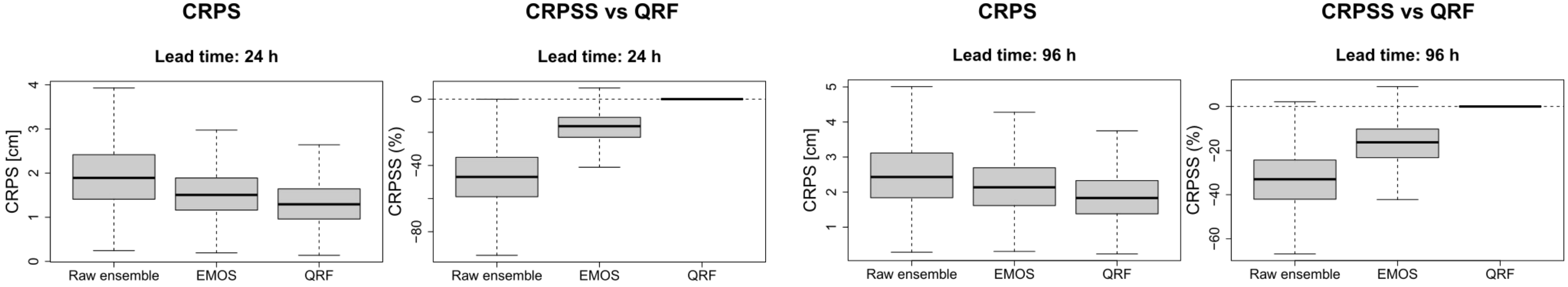
Added value of Quantile Regression Forests (QRF)

Evin et al., in prep.

- The **statistical properties** of the post-processed are **satisfactory** in both cases (flat rank histograms for both EMOS and QRF)



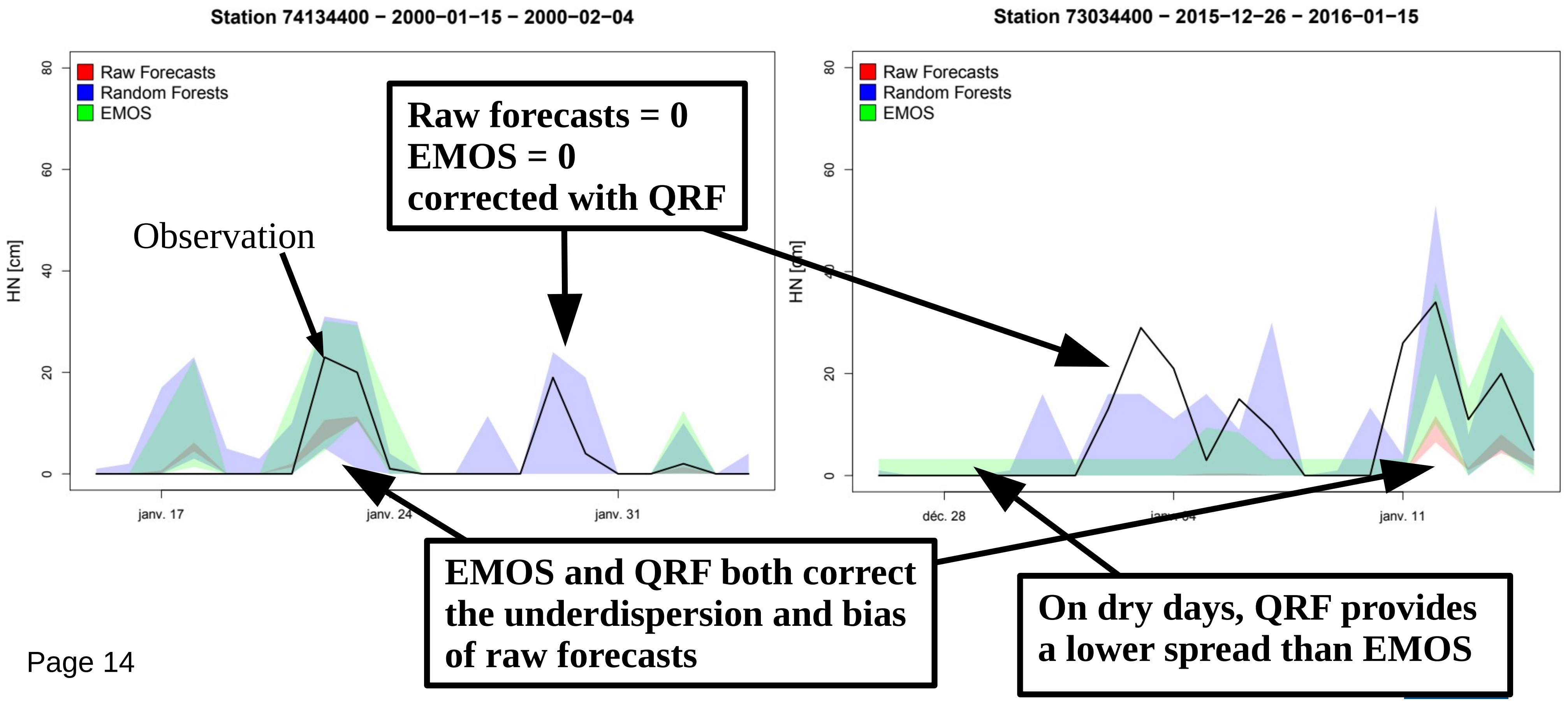
- **A significant improvement of CRPS is obtained with QRF** in theoretical experiments based on the 22-year reforecast dataset (22* [21-year training, 1-year validation])
  → Better predictive power



Added value of Quantile Regression Forests (QRF)

Evin et al., in prep.

- Illustrations on specific cases (24h lead time forecasts):



Conclusions

- Raw ensemble forecasts + snowpack modelling provide predictive but ***biased*** *and* ***underdispersive*** forecasts not well suited for **automated products**.

- **Ensemble Model Output Statistics (EMOS)** improve the forecasts from physical modelling.
  - What is the **best training dataset** ?

    → ***Long reforecasts*** *improve the* ***reliability*** *of the post-processed forecasts for the severe and* ***unusual events***

    → *But they should me* ***more homogeneous*** *with the operational system (initial perturbations)*
  - What is the spatial validity of the post-processing ?

    → ***Spatial consistence of biases*** *allows to apply corrections at the massif scale (1000 km²)*

- **Quantile Regression Forecasts (QRF)**
  - Better predictive skill in theoretical experiments thanks to other predictors
  - Further work required to test the robustness when transfered to real time forecasts

METEO FRANCE

References

More details for the EMOS results in our main reference:

Nousu, J.-P., Lafaysse, M., Vernay, M., Bellier, J., Evin, G., and Joly, B.: Statistical post-processing of ensemble forecasts of the height of new snow, Nonlin. Processes Geophys., 26, 339–357, https://doi.org/10.5194/npg-26-339-2019, 2019.

Other references

Champavier, R., Lafaysse, M., Vernay, M., and Coléou, C.: Comparison of various forecast products of height of new snow in 24 hours on French ski resorts at different lead times, in: Proceedings of the International Snow Science Workshop - Innsbruck, Austria, pp. 1150–1155, http://arc.lib.montana.edu/snow-science/objects/ISSW2018_O12.11.pdf, 2018.

Scheuerer, M. and Hamill, T. M.: Statistical Postprocessing of Ensemble Precipitation Forecasts by Fitting Censored, Shifted Gamma Distributions, Mon. Weather Rev., 143, 4578–4596, https://doi.org/10.1175/MWR-D-15-0061.1, 2015.

Scheuerer, M. and Hamill, T. M.: Generating Calibrated Ensembles of Physically Realistic, High-Resolution Precipitation Forecast Fields Based on GEFS Model Output, J. Hydrometeorol., 19, 1651–1670, https://doi.org/10.1175/JHM-D-18-0067.1, 2018.

Scheuerer, M. and Hamill, T. M.: Probabilistic Forecasting of Snowfall Amounts Using a Hybrid between a Parametric and an Analog Approach, Mon. Weather Rev., 147, 1047–1064, https://doi.org/10.1175/MWR-D-18-0273.1, 2019.

Stauffer, R., Mayr, G. J., Messner, J. W., and Zeileis, A.: Hourly probabilistic snow forecasts over complex terrain: a hybrid ensemble postprocessing approach, Adv. Stat. Climatol. Meteorol. Oceanogr., 4, 65–86, https://doi.org/10.5194/ascmo-4-65-2018, https://www.adv-stat-clim-meteorol-oceanogr.net/4/65/2018/, 2018.

Taillardat, M., Fougères, A., Naveau, P., and Mestre, O.: Forest-based and semi-parametric methods for the postprocessing of rainfall ensemble forecasting, Weather Forecast., in press, https://doi.org/10.1175/WAF-D-18-0149.1, 2019.

Vernay, M., Lafaysse, M., Merindol, L., Giraud, G., and Morin, S.: Ensemble Forecasting of snowpack conditions and avalanche hazard, Cold. Reg. Sci. Technol., 120, 251–262, https://doi.org/10.1016/j.coldregions.2015.04.010, 2015.