Post-processing ensemble forecasts using Ensemble Model Output Statistics (EMOS)

2m temperature v.s. observations

Ivana Aleksovska
ivana.aleksovska@ecmwf.int

Thomas Haiden (ECMWF), Zied Ben Bouallègue (ECMWF), Sándor Baran (University of Debrecen)

European Centre for Medium-Range Weather Forecasts

ECMWF

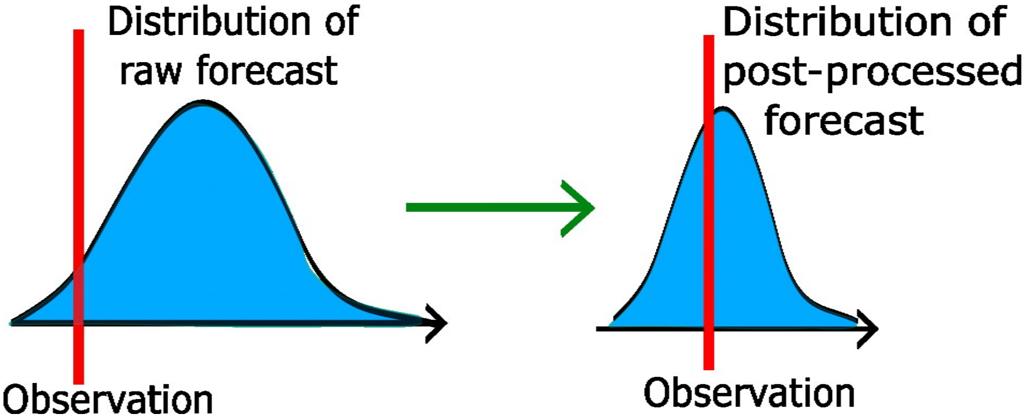
Why do we need post-processing and how do we do it?

- Systematic errors in Numerical Weather Prediction (NWP) Models;
- Statistical correction methods:
  - Establish a statistical relationship between the forecast and the corresponding observations on the training period;
  - Use it to correct future forecasts -> post-processed forecast (pp).

Post-processing methods

Two main groups: Parametric and non-parametric.

-> Parametric : Ensemble Model Output Statistics (EMOS) (Gneiting, 2014)



Ensemble Model Output Statistics (EMOS)

- Estimate the conditional distribution of y, the weather variable of interest, given the forecasts of the different members $x_1, ..., x_N$;

$$Y \mid x_1, ..., x_N \sim p(Y \mid x_1, ..., x_N).$$

- Temperature: Gaussian underlying distribution (Gneiting et al., 2005)

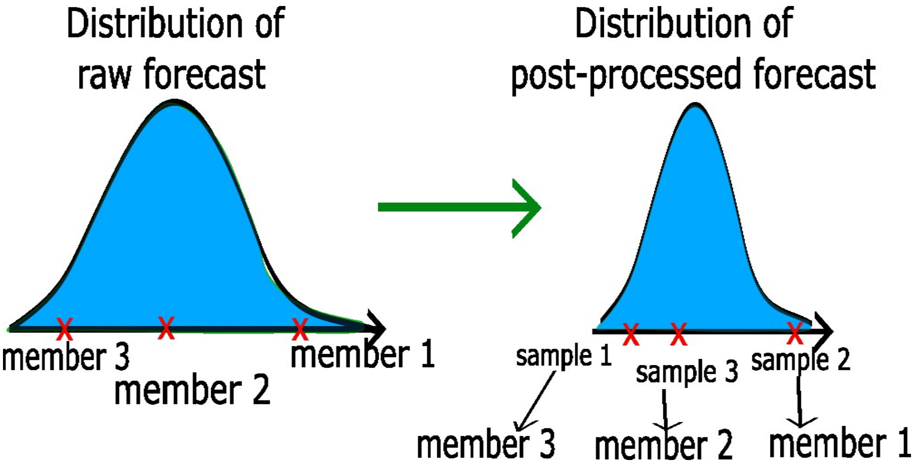
$$Y \mid x_1, ..., x_N \sim \mathcal{N}\left(\alpha_0 + \alpha_1 \mu_e,\ \beta_0 + \beta_1 \sigma_e^2\right),$$

  where μ_e and σ_e^2 are the mean and variance of the ensemble;
  -> α_0, α_1 correct for systematic biases in the ensemble mean,
  -> β_0, β_1 adjust the dispersion of the ensemble.

- Unknown parameters are estimated using the CRPS of forecasts and past observations over a training period (Gebetsberger et al., 2018);

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

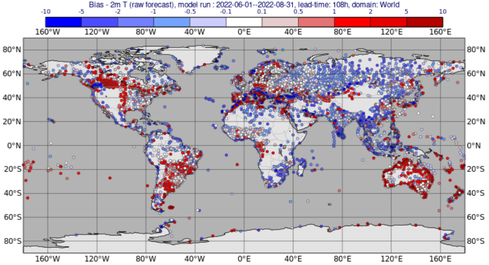
Ensemble Copula Coupling (ECC)

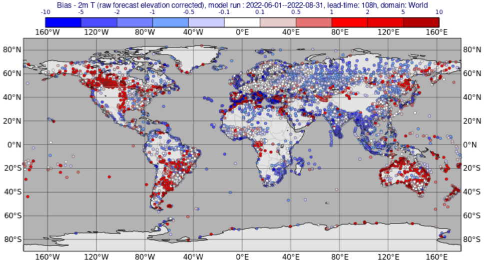


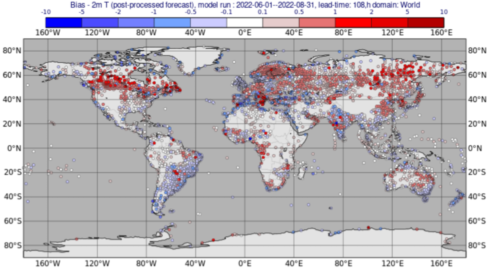
Experimental design and verification score metrics

1. Type of data:
   - SYNOP observations (all over the world);
   - Raw ensemble forecast (ECMWF);
   - Raw corrected ensemble (heigh elevation):
     $T_{\text{corr}} = T + 0.0065 *$ height elevation;
     -> Height elevation: difference between model orography and station elevation;
     -> 0.0065: 'Standard atmosphere' gradient (widely used for height correction);
2. Period:
   - October 2020 - February 2023;
   - Training period: sliding window of 60 previous days.
3. Scores:
   - Bias (ensemble mean - observation);
   - Continuous Ranked Probability Score (CRPS);
   - Continuous Ranked Probability Skill Score (CRPSS).

Bias: raw, corrected raw and pp ensemble forecast (2m T)







Aleksovska et . al.

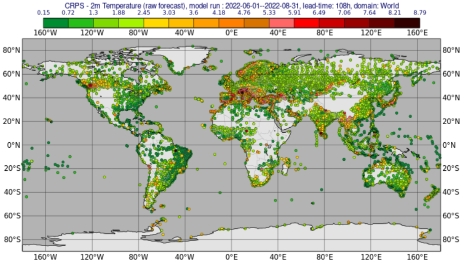
Continuous Ranked Probability Score (CRPS)

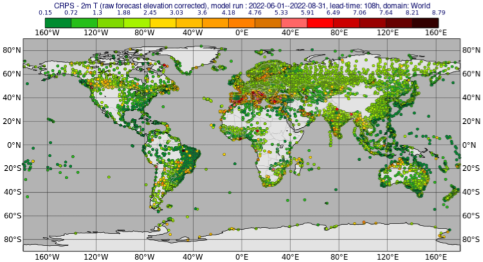
- Probabilistic forecast verification to assess quality and performance;
- CRPS measures global performance (Candille et Talagrand, 2005);

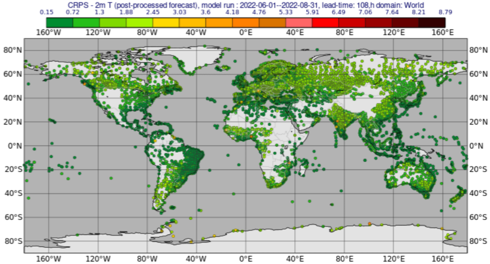
$$\text{CRPS}(\mathbb{F}, \mathbb{F}_o) = \int_{\mathbb{R}} (\mathbb{F}(x) - \mathbb{F}_o)^2 \, dx \, .$$

- $\mathbb{F} = \mathbb{P}[\mathbb{X} \leq x]$ cumulative distribution function of ensemble;
- $\mathbb{F}_o$ cumulative distribution function associated with the observation;
- The lower the score, the better the forecast.

CRPS raw, corrected raw and pp ensemble forecast (2m T)







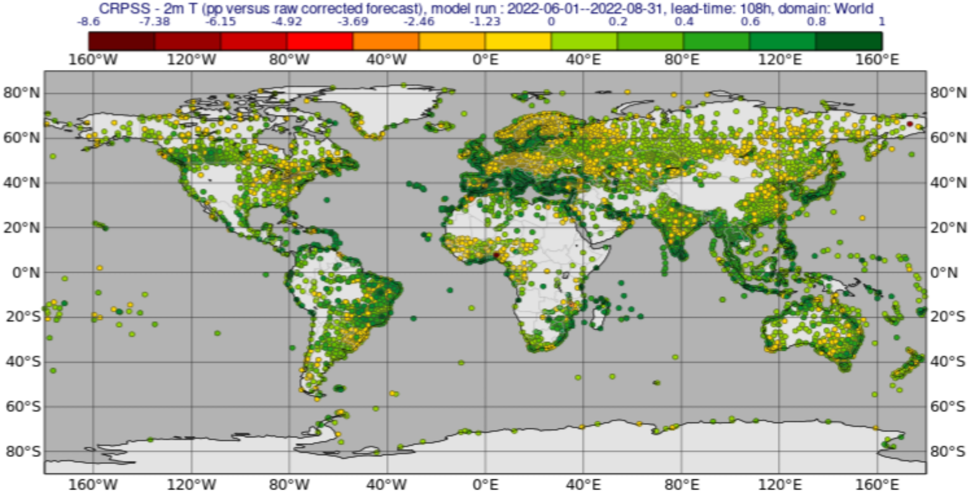
Continuous Ranked Probability Skill Score (CRPSS)

- CRPSS is a measure of how good two forecasts are in matching observed outcomes (Hersbach, H., 2000);

$$\text{CRPSS}(\mathbb{F}, \mathbb{F}_{\text{ref}}) = 1 - \frac{\text{CRPS}(\mathbb{F}, \mathbb{F}_o)}{\text{CRPS}(\mathbb{F}_{\text{ref}}, \mathbb{F}_o)}.$$

- $\text{CRPS}(\mathbb{F}, \mathbb{F}_o)$ - CRPS score of the ensemble forecast;
- $\text{CRPS}(\mathbb{F}_{\text{ref}}, \mathbb{F}_o)$ - CRPS score of the reference ensemble forecast.
- Interpretation:
  1. CRPSS = 1 the forecast has perfect skill compared to the reference forecast;
  2. CRPSS = 0 the forecast has no skill compared to the reference forecst;
  3. CRPSS = negative value the forecast is less accurate than the reference forecast.

CRPSS pp v.s. raw corrected ensemble forecast (2m T)



Conclusion

- Good performance of EMOS over the entire globe;

  –> Verification with scores: Bias, CRPS and CRPSS.

- *Next steps*: Use EMOS as a benchmark to compare against other methods for post-processing (e.g. generative neural networks).

Thank you for your attention!