A-CDFt
An analytical extension of the CDFt bias-adjustment method adopted in the context of the EU C3S Energy Service

29.06.2023

Letizia Lusito

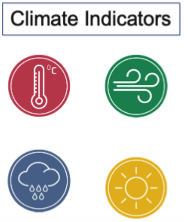
Inside Climate Service

Outline

The EU C3S Energy Service

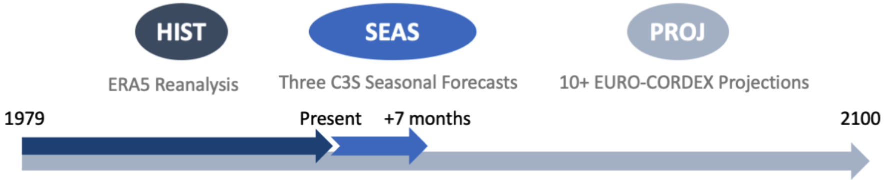
Challenge/Motivation: major transformations in the ENERGY sector:

- increasingly higher share of power supply from variable renewable energy (RE) sources, (wind and solar)
- taking place against a variable and changing climate



C3S Enhanced Energy operational service

to deliver an enhanced operational energy service at the global scale covering data about the **past climate**, multi-model **seasonal forecasts** and multi-model **climate projections**

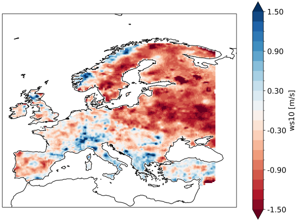


The need for bias adjustment

- Fully physically based climate simulation chain formed by General Circulation Models (GCMs) dynamically downscaled through Regional Climate Models (RCMs), are powerful tools for describing general climate conditions,
- Their direct use in climate change impact or adaptation studies, risk assessment or other analyses requiring climate projections at a regional or local scale is still challenging.
- These data present systematic biases (systematic model errors caused by imperfect conceptualization, discretization, coarse representation of regional features, and spatial averaging within model grid cells) when compared to observations.

Bias Adjustment

the most adopted method to provide 'corrected' climate scenarios; consisting in the application of different post-processing techniques or adjustments that bring the models towards the observed climatology/other observational climate datasets. This statistical post-processing step adjusts selected statistics (mean, variance, distribution) of the so-called "raw" model simulations to better match observed time series over the reference period[1]



[1] Bartòk, Tobin, Vautard, Vrac, Jin, Levavasseur, Denvil, Dubus, Parey, Michelangeli, Troccoli, Saint-Drenan, A climate projection dataset tailored for the European energy sector, Climate Services, 16, 2019, 100138,https://doi.org/10.1016/j.cliser.2019.100138

Delta Change method (DM)

- Adjustment of a "source" dataset based on a reference/target dataset (e.g., observations or long-term (30-year) mean of a climate variable for projection models) simply calculating the change factors (or anomalies or deltas) from the averaged datasets:
- The averages of the source and reference/target dataset are matched by performing a scaling adjustment of the source dataset with the delta factors so that the average of the adjusted source dataset is equal to the average of the reference dataset[2]
- The time series adjustment for temperature is computed as a difference between the prediction/projection and the climatology of the hindcasts/historical run while the time series adjustment for precipitation, solar radiation and wind speed are computed as a ratio (or percentage).

[2] Navarro-Racines, C., Tarapues, J., Thornton, P. et al. High-resolution and bias-corrected CMIP5 projections for climate change impact assessments. Sci Data 7, 7 (2020). https://doi.org/10.1038/s41597-019-0343-8

Quantile Matching

- it is an application of the probability integral transform
- it is designed to adjust the distribution of modelled data, such that it matches observed climatologies by mapping the modelled cumulative distribution function (CDF) of the variable of interest onto the observed CDF
- it attempts to find a transformation h of a modelled variable P_m such that its new distribution equals the distribution of the observed variable P_o: $P_o = h(P_m)$
- If the distribution of the variable of interest is known, the transformation h is defined as: $P_o = F_0^{-1}(F_m(P_m))$

QM can be achieved by[3]:

- using theoretical distributions to represent the CDFs (distribution derived transformations), e.g. mixture of Bernouilli/Gamma etc.
- parametric transformations, e.g. $P_o = bP_m$; $P_o = a + bP_m$
- non parametric transformations, e.g. empirical quantile mapping: the empirical modelled and observed CDFs approximated using tables of the empirical percentiles. Values in between the percentiles are approximated using linear interpolation.

[3] Gudmundsson, L., Bremnes, J.B., Haugen, J.E., Engen Skaugen, T., 2012. Technical note: downscaling RCM precipitation to the station scale using quantile mapping—a comparison of methods. Hydrol. Earth Syst. Discuss. 9, 6185–6201, http://dx.doi.org/10.5194/hessd-9-6185-2012.

CDF-transform (CDFt)

- A probabilistic downscaling method where the relationship between statistical properties (i.e. probabilistic, CDFs) between a target (or reference such as observations or RCMs or historical data) and a source (biased projection model or GCMs) datasets are modelled.
- No assumption on the shape of the relationship to be modelled, or on the family of the CDFs, but rather non-parametric correspondences between the predictor and predictand CDFs are used
- $\exists$ a transformation T allowing to "translate" the CDF of a source variable into the CDF of the reference $\ni T(F_{Gh}(x)) = F_{Sh}(x)$ with $F_{Sh}(x)$ being the CDF of the reference data for the calibration (h, historical) period and $F_{Gh}(x)$ the CDF of the source/to-be-adjusted for the calibration period.
- T can be used to model the relationship between $F_{Sf}(x)$ and $F_{Gf}(x)$, i.e., same CDFs but for a future (different) period, assuming that $F_{Gf}(x)$ is known[4,5]:

$$T(u) = F_{Sh}(F_{Gh}^{-1}(u)) \qquad F_{Sf}(x) = T(F_{Gf}(x)) \qquad F_{Sf}(x) = F_{Sh}(F_{Gh}^{-1}(F_{Gf}(x)))$$

[4] Michelangeli, Vrac, Loukos (2009) Probabilistic downscaling approaches: Application to wind cumulative distribution functions. Geophysical Research Letters, 36, L11708, <doi:10.1029/2009GL038401>

[5] Vrac, Drobinski, Merlo, Herrmann, Lavaysse, Li, Somot (2012) Dynamical and statistical downscaling of the French Mediterranean climate: uncertainty assessment. Nat. Hazards Earth Syst. Sci., 12, 2769-2784, www.nat-hazards-earth-syst-sci.net/12/2769/2012/, <doi:10.5194/nhess-12-2769-2012>.

CDF-transform (CDFt)

The estimates of F_{Sh} , F_{Gh}^{-1}, F_{Gf}, are empirically modelled respectively from the historical observations and the historical and future large-scale simulated data.

$$F_{Sf}(x) = F_{Sh}(F_{Gh}^{-1}(F_{Gf}(x)))$$

the estimator of F_{Sf} can be calculated.

Difference between QM and CDFt

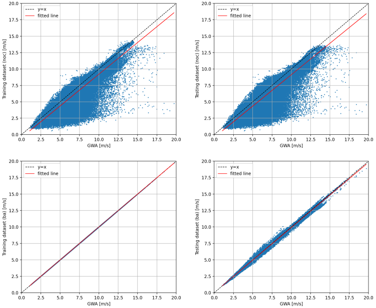
- CDFt takes into account the change in the large-scale CDF from the historical to the future time period, QM only projects the simulated large-scale values onto the historical CDF to compute and match quantiles.
- QM cannot provide local-scale quantiles outside the range of the historical observations (limitation for a changing climate context), whereas CDFt allows one to overcome this problem by taking advantage of the simulated future large-scale CDF.

DM application

- Used to adjust variables in projection models (2015-2100) where trend is not a key feature (e.g., ghi) or variables where the target data is available as reference values and not time series (e.g. Global Wind Atlas)
  - in C3S, DM used to adjust historical (ERA5) wind speed, using as a reference the Global Wind Atlas[6,7] with $< ERA5 >$ the mean (2008-2017) wind speed:

$$F = \frac{GWA}{<ERA5>} \quad WS_{i,b.a.} = WS_{i,orig} * \mathrm{F}$$

- train_ba: bias-adjusted mean wind speed over the period 2008-2017
- test_ba: bias-adjusted mean wind speed over the period 1971-2000
- train_noc: original mean wind speed over the period 2008-2017
- test_noc: original mean wind speed over the period 1971-2000



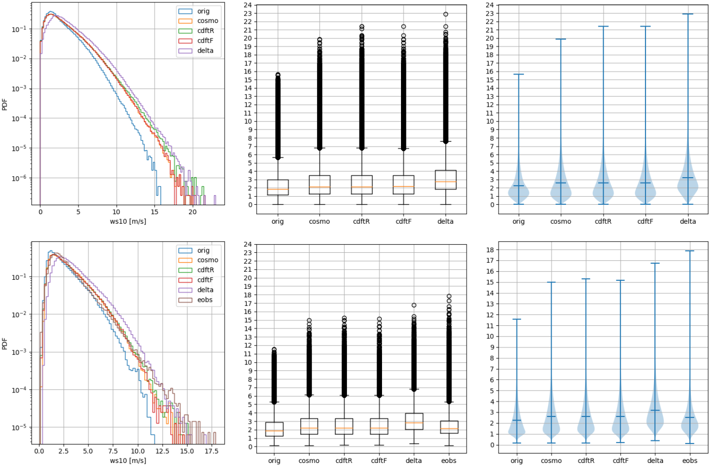
Height	ERA5 Dataset	Slope (95%)	Intercept (95%)	R^2	RMSE
100m	Testing (ba)	0.99016 +/- 0.00023	0.01936 +/- 0.00159	0.99448	0.1613
	Testing (noc)	0.95719 +/- 0.00142	-0.43792 +/- 0.01005	0.80920	1.2107
	Training (ba)	0.99941 +/- 0.00001	0.00280 +/- 0.00005	0.99999	0.0050
	Training (noc)	0.96585 +/- 0.00143	-0.45543 +/- 0.01011	0.81012	1.1901

[6] https://globalwindatlas.info

[7] Murcia, J. P., Koivisto, M. J., Luzia, G., Olsen, B. T., Hahmann, A. N., Sørensen, P. E., Als, M. “Validation of European-scale simulated wind speed and wind generation time series”, Applied Energy, vol. 305, 117794 (2022) https://doi.org/10.1016/j.apenergy.2021.117794

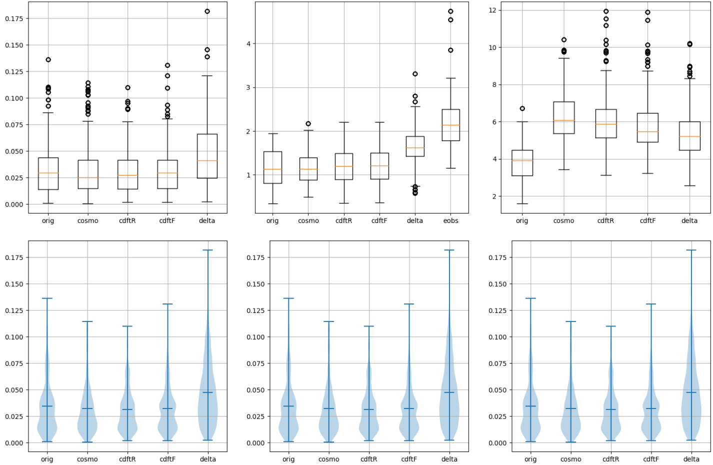
CDFt application

Adjustment of variables in projection models when climate change trends are non negligible; the target and source/predictand PDFs are very different (as for the wind speed) and an averaged-adjustment (as in DM) is not appropriate



Example: adjusting ERA5 wind speed (1980-2018) with COSMO (2006-2018) as reference; time series: 2009-2018; lat: 45-49, lon: 5-8

CDFt application



Evaluation of extremes:

- Wind generation minima:
  - annual min of hourly data
  - 5 days
- Wind generation maxima:
  - RAMP annual max on 3 hours

Issues

Issues encountered:

- Very high computational resources needed
- Difficulties in evaluating the adjustments in the tails of the PDFs
- Difficulties in adjusting the variables if shape/extend of the distrubution will change due to an indipended factor (e.g. climate change trends)

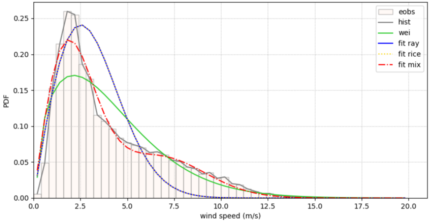
The idea

- Find the most appropriate PDF that describes reference/biased data (analytical description)
- Find "delta factors" based on the ratio of the two PDF’s and scale the biased data accordingly (such as in the delta method)
- Find the most appropriate PDF that describes the corrected data: the parameters of the corrected data PDF should be much closer to the reference data: their distance is also a way to quantify goodness-of-adjustment/the uncertainties of adjustment.

Fitting PDFs with a continuous distribution

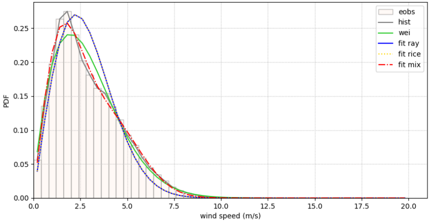
Example: Rayleigh-Rice mixture PDF for a particular location in France (Orange, lon: 4.82°, lat: 44.13°)

E-OBS



- α_{mix}: 0.36
- μ_{rice}: 6.16
- σ_{rice}: 2.57
- σ_{ray}: 1.84

ERA5

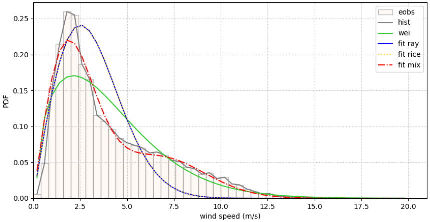


- α_{mix}: 0.45
- μ_{rice}: 3.71
- σ_{rice}: 1.72
- σ_{ray}: 1.50

Normalisation/scaling

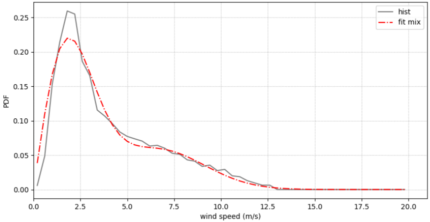
Example: Rayleigh-Rice mixture PDF for a particular location in France (Orange, lon: 4.82°, lat: 44.13°)

E-OBS



- α_{mix}: 0.36
- μ_{rice}: 6.16
- σ_{rice}: 2.57
- σ_{ray}: 1.84

Corr. ERA5



- α_{mix}: 0.35
- μ_{rice}: 6.19
- σ_{rice}: 2.54
- σ_{ray}: 1.84

Conclusion

- The need of reliable climate service (high-resolution, and including far future climate projections) is rising rapidly in the Energy Sector
- Very big efforts from the scientific communities to cope with such large demands
- Strategies need to be put in place to allow fast applications/implementations without forgetting proper uncertainties estimations
- Evaluation of extremes in A(djusted)-CDF/PDF datasets
- Evaluation of computational performances

Thank you!

letizia.lusito@inclimateservice.com

https://www.inclimateservice.com