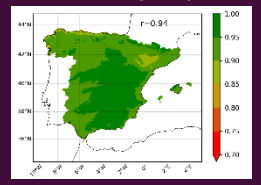
A critical view on the suitability of Machine Learning techniques to downscale climate change projections

Alfonso Hernanz[1], Juan Andrés García-Valero[2], Marta Domínguez[1], Ernesto Rodríguez-Camino[1]
1 Spanish Meteorological Agency (AEMET), Madrid, 28040, Spain
2 AEMET, Murcia, 30107, Spain
Correspondence to: Alfonso Hernanz (ahernanzl@aemet.es)

Data.

Two variables stronghly correlated (see Fig. 1) have been used, daily máximum surface temperature (TMAX) and temperature at 850 hPa (T850), in a historical period 1980-2016.
TMAX has been taken from an observational grid (Peral *et al.*, 2017) and T850 comes from the ERA-Interim reanalysis (Dee *et al.*, 2011)



Correlation between T850 and TMAX.

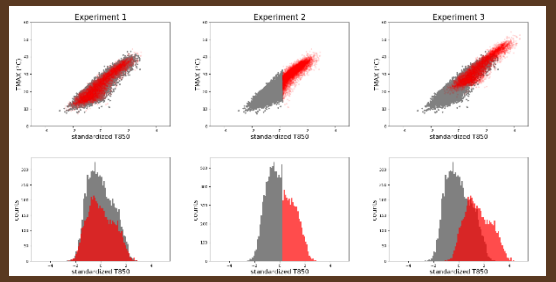
Methods.

- MLR: Multiple Linear Regression
- ANN-RELU: Artificial Neural Network with a rectified linear activation function.
- ANN-LOG: Artificial Neural Network with a logistic activation function.
- SVR: Support Vector Machine.

pyClim-SDM software (Hernanz *et al.*, 2022a, https://github.com/ahernanzl/pyClim-SDM) has been used for the downscaling.

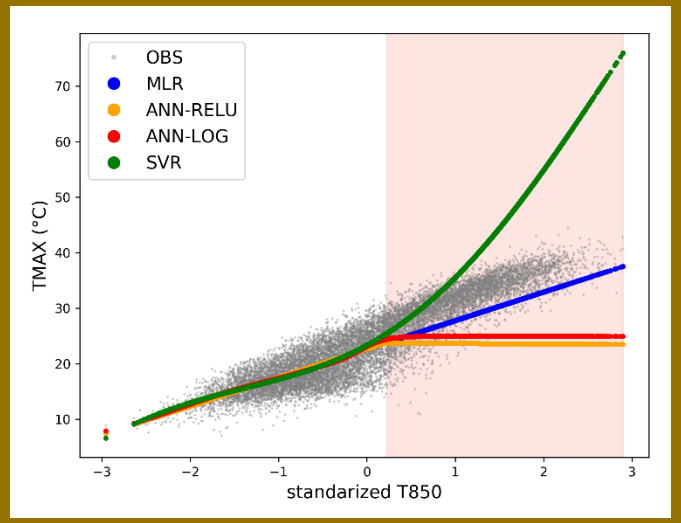
Experiments.

The training/testing split has been performed in three different ways:
1. Total overlapping
2. No overlapping
3. Partial overlapping

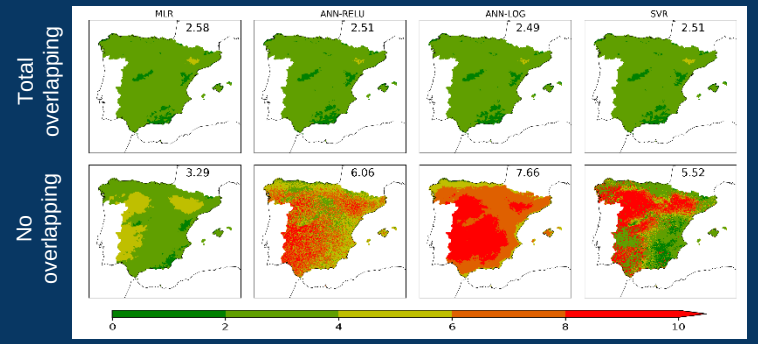


T850 training (grey) and testing (red) distributions used for each experiment.

Can we trust Machine Learning algorithms under extrapolation?



T850 vs. observed (grey) and downscaled TMAX (ºC) by MLR (blue), ANN-RELU (orange), ANN-LOG (red) and SVR (green). Extrapolation in red background.



RMSE (ºC) by each method (in columns: MLR, ANN-RELU, ANN-LOG, SVR) for total and no overlapping (in rows, respectively)

Each ML method presents a different sensitivity to the degree of extrapolation

Partial overlapping

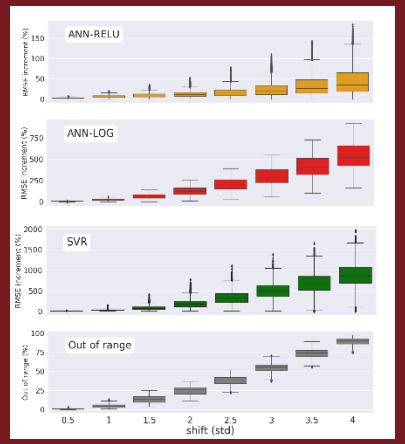


Fig. 4. Increment of RMSE (%) for ANN-RELU (first panel, orange), ANN-LOG (second panel, red) and SVR (third panel, green) for different degrees of training/testing overlapping

References

Dee, D.P., Uppala, S.M., Simmons, et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137, 553–597. https://doi.org/10.1002/qj.828

Hernanz, A., García-Valero, J. A., Domínguez, M., Correa, C. and Rodríguez-Camino, E. (2022a). pyClim-SDM 1.0: a software for statistical downscaling of climate change projections with a graphical user interface. Geoscientific Model Development. (Submitted)

Hernanz, A., García-Valero, J. A., Domínguez, M., & Rodríguez-Camino, E. (2022b). A critical view on the suitability of Machine Learning techniques to downscale climate change projections: illustration for temperature with a toy experiment. Atmospheric Science Letters, e1087. https://doi.org/10.1002/asl.1087

Peral, C., Navascués, B., and Ramos, P. (2017). Serie de precipitación diaria en rejilla con fines climáticos. Nota Técnica nº 24 AEMET. http://www.aemet.es/documentos/es/conocermas/recursos_en_linea/publicaciones_y_estudios/publicaciones/NT_24_AEMET/NT_24_AEMET.pdf