Supplementary Material

Statistical downscaling of seasonal spring-frost predictions

Sebastiano Roncoroni[a] and Panos J. Athanasiadis[a]

[a]CMCC Foundation - Euro-Mediterranean Center on Climate Change, Italy

Contents

S1 Cross-validation **2**

S2 Analysis of trends in the predictand **2**

S3 Study of an individual frost event **3**

S4 Figures **5**

S1 Cross-validation

Overfitting occurs when a statistical model with too many degrees of freedom is used without constraints to learn a relationship $y = f(X)$ between predictand y and predictors X. The model "learns" the noise in the data alongside the signal, and will not generalise well if applied to a fresh dataset. One way to avoid overfitting is to adopt a cross-validation strategy, and Leave-One-Out (LOO) cross-validation is a common choice in the context of seasonal predictions (Cionni et al., 2022, Ramon et al., 2021, Torralba et al., 2017). Within LOO, each sample is set aside once and all remaining data are used to predict it: one benefit of LOO is that it closely mimics an operational setting (Torralba et al., 2017), when all the available information from the past is used to predict ahead in time.

However, following a cross-validation protocol for the learning step does not guarantee that the statistical method used to effect the downscaling is fully protected against overfitting. For example, information can leak from the training to the test dataset through composites and anomalies. For this reason, we include composite and anomaly computations within the LOO cycle whenever making predictions with the statistical downscaling method.

Finally, overfitting can also ensue from excessive tuning of the statistical method's hyper-parameters. In our case, relevant hyper-parameters are the regularisation factor for the logistic regression model, the number of predictors used, and the geographical domain they are defined in. The difficulty is normally addressed by dedicating part of the data for the tuning: this strategy, however, is not viable here due to the scarceness of the available record. For this reason, we do not optimise the statistical method's hyper-parameters (thereby trading sub-optimality for robustness), and rely on physical intuition for the important choices.

S2 Analysis of trends in the predictand

Global warming imprints a predictable signal on climate, e.g. in the form of linear trends. In the context of climate predictions, it has been shown that trends in the target variable can inflate the skill of the prediction systems used to predict it (Wulff et al., 2022). Here, we check for the existence of linear trends in March-mean and March-minimum daily minimum temperature over the 1950–2021 period. The linear trend is defined as the slope of the best-fit linear approximation of the data obtained with an ordinary least-squares method. The significance of trends is assessed through the p-value, which we compute according to the Wald test with t-distribution of the statistic

using the scipy Python library (Virtanen et al., 2020). A p-value close to zero means that the observed trend would be unlikely under the null hypothesis of no trend. Trends are deemed statistically significant when the p-value is $p < 0.05$.

Figure S5 shows the linear trends for March-mean (top) and March-minimum (bottom) daily minimum temperature. Regarding March-mean temperatures, we find a significant trend for approximately 10% of the target grid points in our domain only (top left): in about half of the cases the trend is positive (warming), while in the other half it is negative (cooling). Furthermore, the p-value relative to the trend of spatially-averaged March-mean minimum temperature over the Catalan domain (top right) is $p = 0.73$, well above the 0.05 mark. As for March-minimum temperatures, the trend is significant for approximately 5% of the target grid points, and the p-value for the spatially-averaged time series is $p = 0.97$. We conclude that there is no strong evidence for the existence of significant linear trends in the target domain.

S3 Study of an individual frost event

We analyse the synoptics associated with an individual frost event in Catalonia, cherry-picked for illustration purposes. Rows 1–3 of Figure S6 show the synoptic situation over the Euro-Atlantic domain (ERA5 data) every third day between March 3rd 2010 and March 18th 2010, when Catalonia was affected by widespread and severe frost conditions: in fact, the event is one of the coldest in the entire 1950–2021 record. A time series of daily minimum temperature (E-OBS data) averaged over the target grid points in Catalonia (solid line) is shown in the bottom panel of Figure S6, together with the ± 1 standard deviation spread (shaded interval). The zero degrees line is also shown for ease of visualisation (dashed line). On March 3rd, no blocking structures are detected in the Euro-Atlantic region, a strong anomalous geopotential low flanks Catalonia to the west, and a high influences the central Mediterranean: this large-scale circulation pattern favours northward advection of warm air masses into Catalonia. Consistently, we observe that the region-averaged minimum temperature is well above zero degrees. Then, for approximately a week from March 6th to March 12th, a persistent *Rex*-type blocking located over the British Isles and the North Sea obstructs the westerly mid-tropospheric flow, and determines centres of strongly anomalous high (low) geopotential height over Iceland (the western Mediterranean). Contrary to the previous case, this synoptic situation is conducive to injection of cold air masses into the Iberian peninsula from

north-eastern Europe along the south-eastern flank of the high [e.g., Brunner et al. (2017)]. Concurrently, the average minimum temperature drops to a minimum of nearly −5 °C. Eventually, the *Rex* block dissipates and on March 15th–18th a subtropical ridge exerts its influence over the western Mediterranean. Concomitantly, temperature starts to increase and is well above zero by the end of the period considered.

S4 Figures

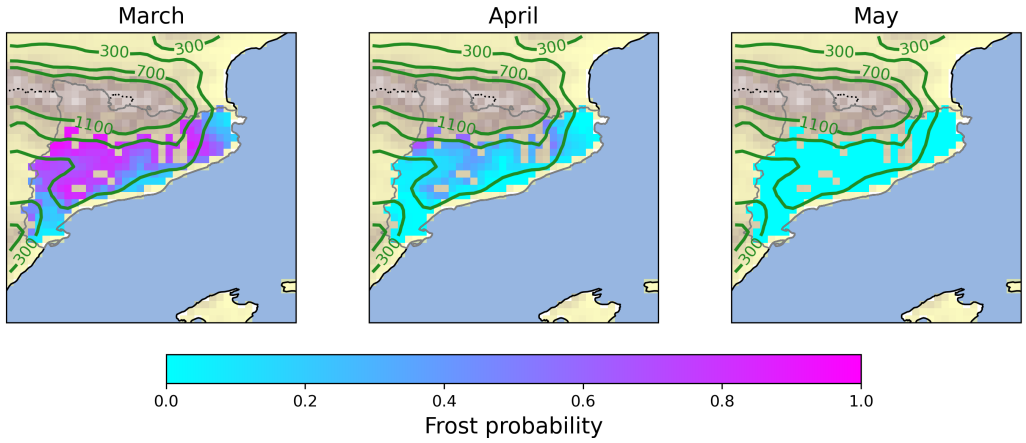


Figure S1: Climatological probability of frost occurrence in Catalonia below 700 m surface elevation over the 1950–2021 period in March (left), April (middle), and May (right). The 300, 700, and 1100 m elevation contours are denoted by thick green lines (contours have been smoothed with a Gaussian filter for ease of visualisation). Surface elevation in represented in the background by colour shading.

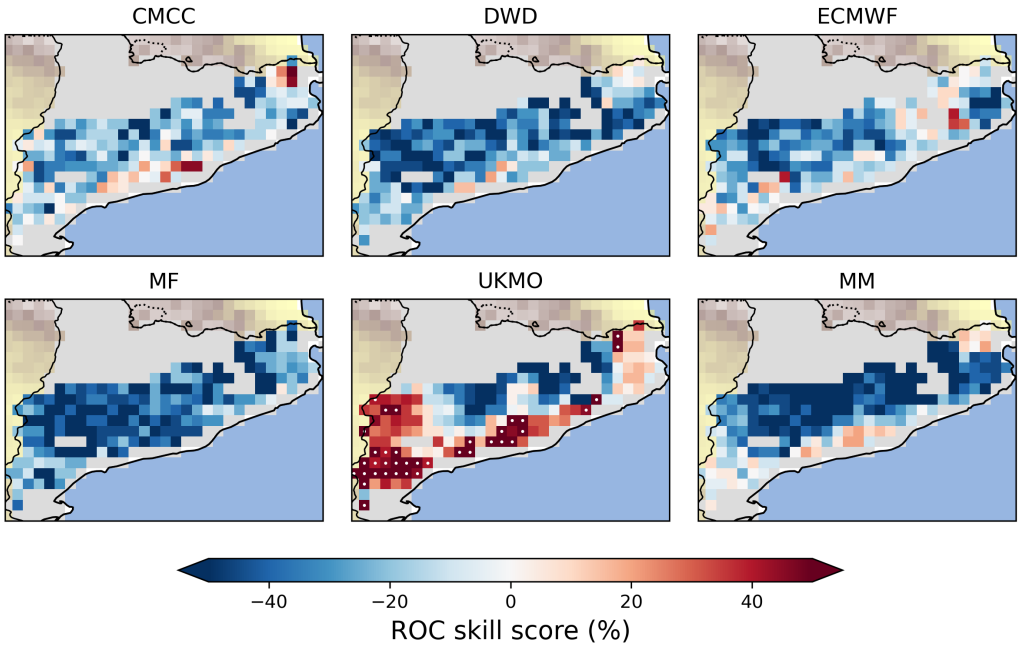


Figure S2: ROC skill score of the dynamical predictions with lead-dependent bias adjustment against the climatological reference for the individual seasonal prediction models (CMCC, DWD, ECMWF, MF, UKMO) and for the multi-model ensemble (MM). Dots mark grid points where the skill score is statistically significant at the 95% significance level. Bias adjustment was effected in Leave-One-Out cross-validation mode.

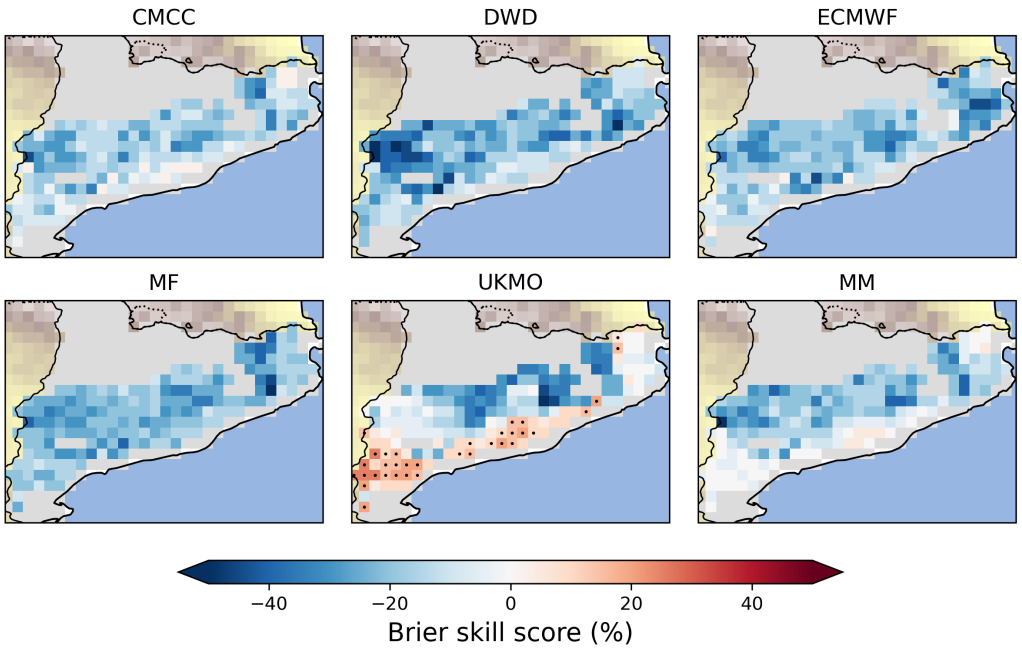


Figure S3: As in Figure S2, but for the Brier skill score.

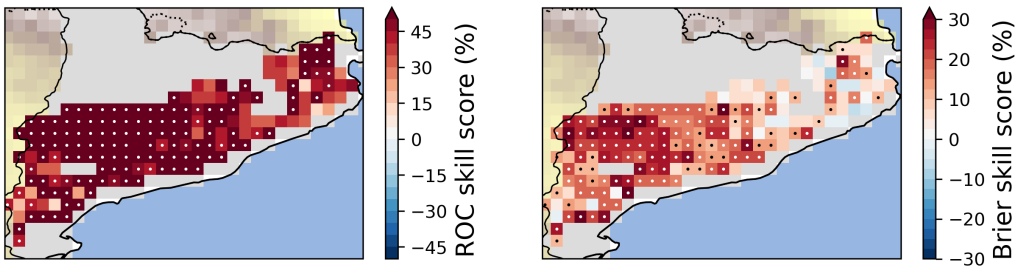


Figure S4: ROC (left) and Brier (right) skill scores for the perfect-prognosis predictions obtained with the statistical downscaling model (note the different scale of the colorbars). The training period is 1950–2021, and the verification period is 1993–2016. Points where the skill score is statistically significant at the 95% significance level are marked by dots. LOO cross-validation covers training of the logistic regression model, but not computation of anomalies and composites.

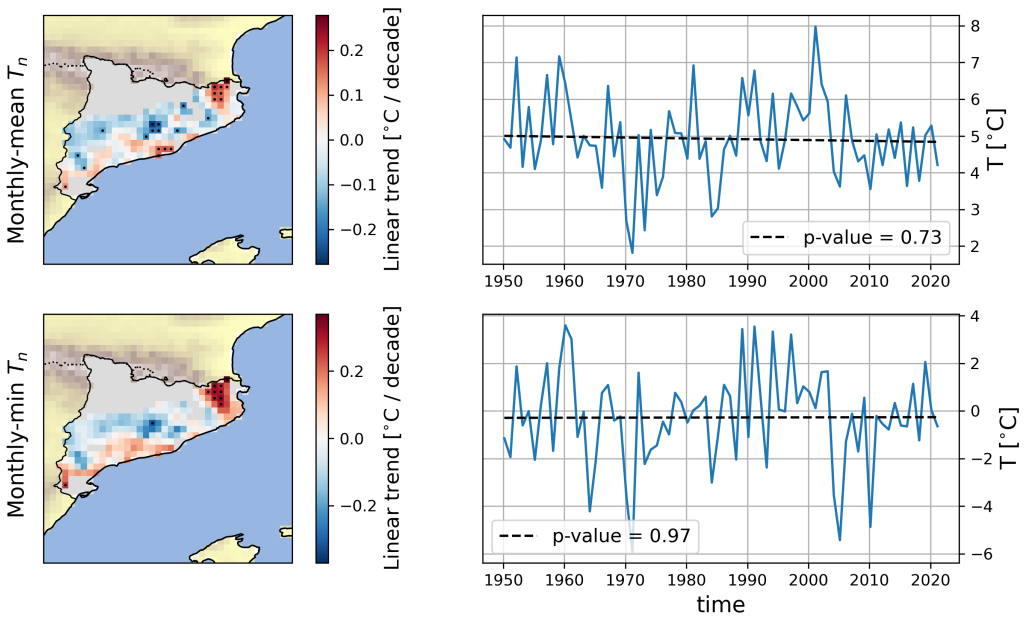


Figure S5: Left: linear trend of March-mean daily minimum temperature (top) and March-minimum daily minimum temperature (bottom). Grid points significant at the 95% significance level according to a two-tailed test are marked by dots. Right: time series of March-mean daily minimum temperature (top) and March-minimum daily minimum temperature (bottom) averaged over the target grid points in Catalonia. The corresponding least-squares linear fits are shown by the dashed black lines, and the associated p-values are reported in the legend.

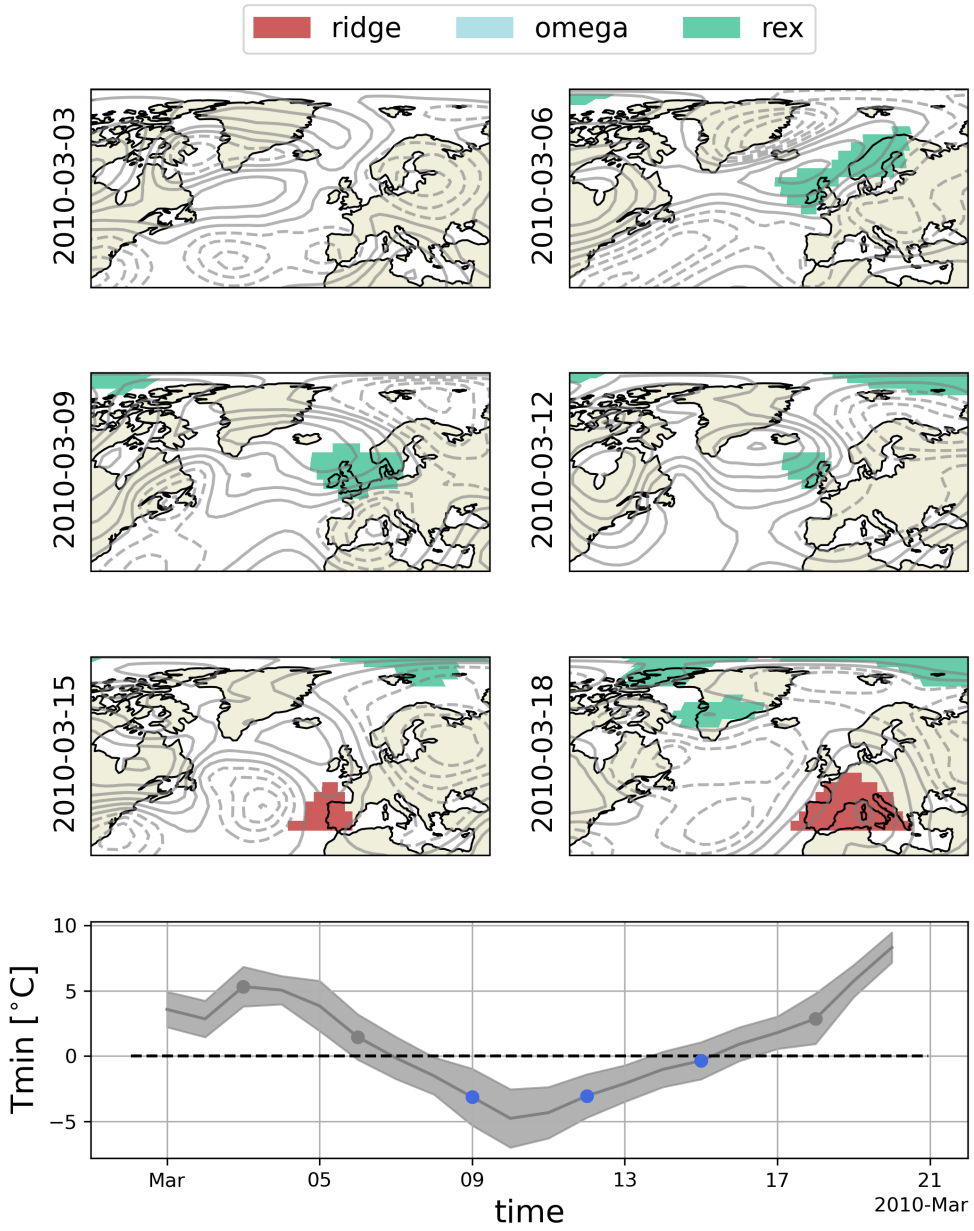


Figure S6: Rows 1–3: blocking events (see legend) and anomalous 500 hPa geopotential height (contours; the anomaly is with respect to a smoothed daily climatology. Contour spacing is 60 hPa) over the Euro-Atlantic region every third day between March 3rd 2010 and March 18th 2010. Bottom: time series of daily minimum temperature averaged over the target grid points in Catalonia. Dots mark the days when the snapshots shown in the top panels are taken (in blue if temperature is below zero degrees).

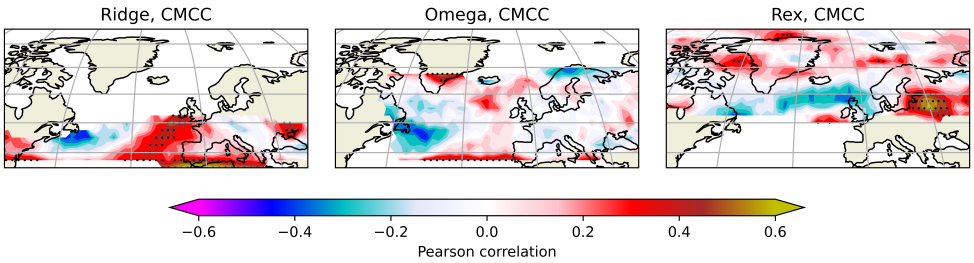


Figure S7: As in Figure 7 in the main text, but for the CMCC system only.

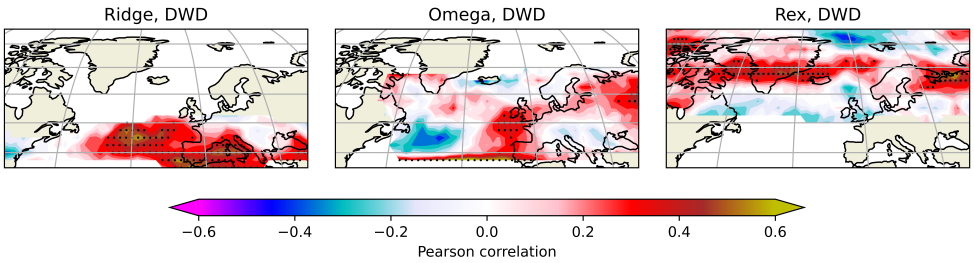


Figure S8: As in Figure 7 in the main text, but for the DWD system only.

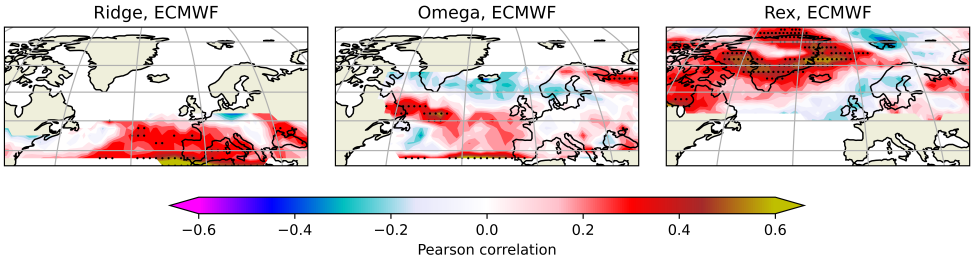


Figure S9: As in Figure 7 in the main text, but for the ECMWF system only.

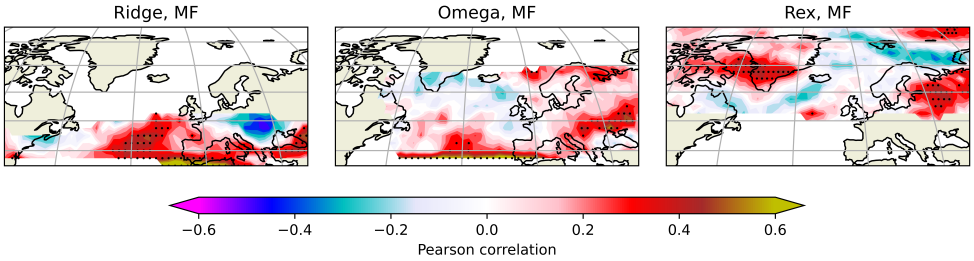


Figure S10: As in Figure 7 in the main text, but for the MF system only.

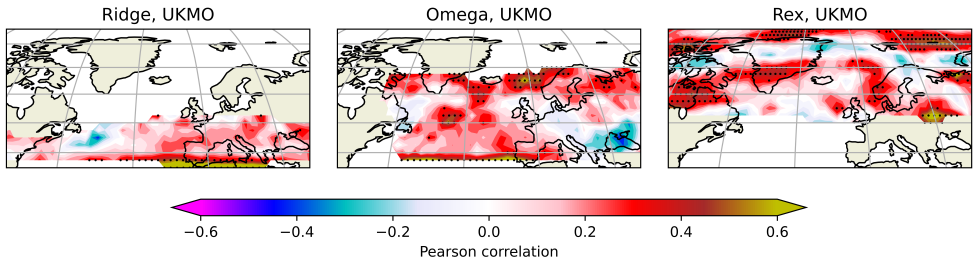


Figure S11: As in Figure 7 in the main text, but for the UKMO system only.

References

Brunner, L., Hegerl, G. C., and Steiner, A. K. (2017). Connecting atmospheric blocking to european temperature extremes in spring. *J. Clim.*, 30(2):585–594.

Cionni, I., Lledó, L., Torralba, V., and Dell'Aquila, A. (2022). Seasonal predictions of energy-relevant climate variables through Euro-Atlantic Teleconnections. *Clim. Serv.*, 26:100294.

Ramon, J., Lledó, L., Bretonnière, P.-A., Samsó, M., and Doblas-Reyes, F. J. (2021). A perfect prognosis downscaling methodology for seasonal prediction of local-scale wind speeds. *Environ. Res. Lett.*, 16:054010.

Torralba, V., Doblas-Reyes, F. J., MacLeod, D., Christel, I., and Davis, M. (2017). Seasonal Climate Prediction: A New Source of Information for the Management of Wind Energy Resources. *J. Appl. Meteor. Clim.*, 56(5):1231 – 1247.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods*, 17:261–272.

Wulff, C. O., Vitart, F., and Domeisen, D. I. V. (2022). Influence of trends on subseasonal temperature prediction skill. *Q.J.R. Meteorol. Soc.*, 148(744):1280–1299.