A Stochastic Generator of Global Monthly Wind Energy

with Tukey $g$-and-$h$ Autoregressive Processes

Jaehong Jeong$^1$, Yuan Yan$^2$, Stefano Castruccio$^3$, and Marc G. Genton$^2$

Supplementary Material

Section 2

For each spatial location of wind speed after removing the trend, we test significances of skewness and kurtosis over time (Bai and Ng, 2005). The p-values of a significance test for skewness and kurtosis in many spatial locations are smaller than 0.05 as in Figure 1(a) and (b). This indicates a rejection of the Gaussianity null in favor of the non-Gaussian distribution at the usual 5% level. Thus, the first two moments are not sufficient for temporal dependence at least in the modeling of monthly wind speed, and we need to consider the skewness and kurtosis as well.
Figure 1: p-values of a significance test for (a) skewness (equals to 0) and (b) kurtosis (equals to 3) over time for one selected ensemble member.

For Gaussianity, the Lilliefors and Jarque-Bera tests are used for testing a null hypothesis that the data comes from the normal distribution [Lilliefors, 1967; Jarque and Bera, 1987]. For these tests (Figure 2), there are 15,347 (39.77%) and 23,923 (61.99%) out of 38,592 points where p-value < 0.05, respectively. Many locations show clear evidence of significance for non-Gaussian assumption.

Figure 2: p-values for (a) Lilliefors and (b) Jarque-Bera normality tests.
Section 3.2

Figure 3 shows the p-values for the significance of the first temporal lag of the cross-correlation between innovations one lag apart. Figure 4 maps the estimates of the first three autoregressive coefficients from one randomly selected ensemble member. Figure 5 shows the corresponding p-value for the significance. Figure 6 shows the p-values for the significance of the estimated $g$ and $h$ in the Tukey random field.

Figure 3: (a) p-values of the first temporal lag of the cross-correlation for the innovations between locations that are one longitudinal lag apart. (b) Plot of the selected order for the TGH-AR($p$).

Figure 4: Plot of the first three estimated autoregressive parameters, (a) $\hat{\phi}^1_{L_m,\ell_n}$, (b) $\hat{\phi}^2_{L_m,\ell_n}$, and (c) $\hat{\phi}^3_{L_m,\ell_n}$, from one ensemble member.
Figure 5: p-values of the first three autoregressive parameters, (a) $\phi^1_{L_m,\ell_n}$, (b) $\phi^2_{L_m,\ell_n}$, and (c) $\phi^3_{L_m,\ell_n}$, from one ensemble member.

Figure 6: p-values of the Tukey $g$-and-$h$ transformation parameters, (a) $g(L_m,\ell_n)$ and (b) $h(L_m,\ell_n)$, from one ensemble member.
Section 3.4

\( \varphi_{Lm} \) can be written in the following form

\[
\varphi_{Lm} = \begin{pmatrix}
\varphi_{Lm}(c_1) & \varphi_{Lm}(c_2) & \varphi_{Lm}(c_3) & \ldots & 0 & 0 & 0 \\
\varphi_{Lm}(c_2) & \varphi_{Lm}(c_3) & \varphi_{Lm}(c_4) & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & \varphi_{Lm}(c_{N-1}) & \varphi_{Lm}(c_N) & \varphi_{Lm}(c_N) \\
0 & 0 & 0 & \ldots & \varphi_{Lm}(c_N) & \varphi_{Lm}(c_N) & \varphi_{Lm}(c_N)
\end{pmatrix}
\] (S0.1)

Figure 7 shows the parameter estimate for step 3 in the model across the ten sequential time subsamples consisting of the different decades considered. Figure 8 shows similar results for the two-band parameters.
Figure 8: Parameter estimates of (a) $a_{L_m}$ and (b) $b_{L_m}$ for 10 sequential sub-samples. Boxplots of estimates of (c) $a_{L_m}$ and (d) $b_{L_m}$ across latitudes at each sub-sample.
Section 4

For the comparison with our model (T), with a model with no spatial dependence (I), we first focus on a single simulation for May 2020, see Figure 9. While the marginal values simulated by I are still physically coherent (i.e., there are lower winds generally over land), the lack of spatial dependence creates maps which are clearly unphysical, as apparent from the panel (b).

Figure 9: Maps of one realization of wind speeds from (a) T and (b) I runs

For the comparison of T with a model with Gaussian dependence (G), in Figure 10 we showed the empirical estimates of skewness and kurtosis (over time) for one LENS run in panels (a–b) (same as in Figure 1 (e–f) in the main manuscript), compared with the same estimates according to a simulation from the T model (c–d) and G model (e–f). It is apparent that the G model is able to capture neither third nor fourth moment of the distribution, while the T model captures the magnitude and spatial
variability.

Figure 10: The empirical skewness and kurtosis of residuals after subtracting the trend from one LENS, T, and G runs, respectively. Only for the locations where p-values of a significance test are less than 0.05.

Figure 11 shows a comparison in terms of the fitted contrast variances in a region including South Africa.
A Stochastic Generator of Global Monthly Wind Energy

Section 5

We can measure image quality by computing the structural similarity index (SSIM, (Wang et al., 2004)). Figure 13(a) gives a representation of the computed SSIMs, from Jan 2006 to Dec 2100, for an average of reproduced wind speed across 40 runs based on the Tukey g-and-h and
Gaussian cases using the reference of average wind speed across the 40 LENS runs. In Figure 13(b), we show computed SSIMs for one selected, reproduced wind speed based on different cases using the reference of one selected wind speed from the non-training set of the remaining 35 LENS runs. In terms of structural similarity, the Tukey $g$-and-$h$ case has slightly larger values and is more similar to the LENS runs than the Gaussian case. These results are in agreement with the temporal patterns in Figure 16. All LENS and SG runs show similar values of wind speed near the Gulf of Aden as time evolves.

We then compare the empirical skewness and kurtosis from 40 LENS runs with those from each model’s 40 SG runs. Figure 14 represents averages of skewness and kurtosis across longitudes and runs from each model. We find that the skewness values obtained from the Tukey $g$-and-$h$ model are
similar to the empirical values from the climate model than the Gaussian model, while, for several latitudes, it indicates larger kurtosis values than the empirical skewness. Although the empirical kurtosis values are better matched by the Tukey $g$-and-$h$ case at latitudes $30 \sim 60^\circ$N, we observe that the kurtosis values are much larger than the empirical values near several locations, such as western Europe, western China, Central America, and West Africa.

Figure 15: Averages of kurtosis across runs from (a) the LENS and (b) the SG.
Figure 16: Wind speeds near the Gulf of Aden, from Jan 2006 to Jan 2021, from the training set of LENS (blue lines) and five SG runs (red lines, offset by 2 ms$^{-1}$).

Figure 17 shows a comparison with T and I as in the previous Section, same location and months as in Figure 4. The T model clearly follows more closely the ideal 45 degrees line.

Figure 17: QQ-plots between the WPD values from the LENS runs versus T and I runs in (a) March and (b) September 2020.
Figure 18: An example of the GUI for the generation of surrogate runs.

Bibliography


