This poster participates in OSP
Outstanding Student Poster Contest

European Geosciences Union General Assembly 2016
Vienna, Austria, 17-22 April 2016
Session HS7.4: Change in climate, hydrology and society,
Vol. 18, EGU2016-14946.

Stochastic investigation of wind process for climatic variability identification

Ilias Deligiannis, Vasileios Tyrogiannis, Olympia Daskalou, Panayiotis Dimitriadis, Yannis Markonis, Theano Iliopoulou and Demetris Koutsoyiannis

The poster can be downloaded at: http://www.itia.ntua.gr/

Department of Water Resources
and Environmental Engineering
National Technical University of Athens

1. Introduction

The difference in atmospheric pressure between two locations can affect the wind speed [1] and, as a result, precipitation. According to researchers, the air pressure has been globally affected by human activity during the last half century. This change is of high importance for the climatic processes [2].

In general, a change in wind's behaviour is commonly attributed to anthropogenic climatic change. However, most of the studies have not taken into consideration the "Hurst phenomenon", also known as "long-term persistence" for the analysis of hydro-climatic processes and particularly wind speed. Usually, high (low) values of wind speed are followed by high (low) ones, meaning that observations appear in groups [3]. In other words, the autocorrelation coefficient remains quite high as the scale increases due to this clustering effect.

In this study, the wind speed is analyzed in terms of its climacogram (i.e., plot of variance or standard deviation of the mean-aggregated random variable versus scale) in order to determine whether it exhibits behaviour of long-term persistence. The justification for the use of the climacogram as a measure of statistical uncertainty can be seen in [4]. In this analysis, we use hourly wind speed data from over 7,000 wind stations from around the globe (https://www.ncdc.noaa.gov/cdo-web/) and we also estimate the Hurst coefficient (or equivalently, we calculate the slope of the decay of the climacogram) for various time periods. Finally, we estimate the prediction interval (or error) and we comment on the results:

- If the prediction intervals of the wind speed are large (close to unity) for all examined periods and for each station, then the model can describe adequately the climatic variability of wind and so, it is possible that the changes observed during the last decades can be well described by the Hurst phenomenon.
- In contrast, in case a significant variation of the prediction interval is observed for various time periods, then the model used cannot effectively describe the climatic variability of wind.

Aim: Is it possible to describe the climatic variability of wind speed using just three parameters?

2. Methodology

The statistical uncertainty enclosed within the wind process is quantified through a Monte Carlo approach. The analysis is based on the assumptions that the ratio of the annual mean wind speed divided by the annual standard deviation is a stationary process, normally distributed and that it follows one of the most commonly used stochastic models in geophysics, i.e., Markov and HK (including the White Noise process for H=0.5). These assumptions are not only parsimonious but also considered conservative since any non-stationary approach would increase the complexity of the system, the probability function is likely it has a non-Gaussian tail and the stochastic structure cannot be any less complex that the Markov and HK one-parameter models, which entail all exponential as well as a power-type behaviours. Furthermore, the analysis is applied for all climatic zones described in the Koppen system. Moreover, each mean annual value is considered valid when it is estimated from more than 1200 h, i.e. 4 measurements per day for at least 10 months. For the synthesis of the stochastic timeseries, we use the 3×AR(1) technique described in [3]:

The stationary process is produced as a sum of 3 stationary Markov processes, xi = Ai + Bi + Ci. The processes A, B, C have the following characteristics:

Autocorrelation coefficient for lag 1:

$$\rho_a = 1.52\ (H - 0.5)^{1.32}$$
$$\rho_b = 0.953 - 7.69\ (1 - H)^{3.85}$$
$$\rho_c = 0.932 + 0.087\ H, \ \text{for}\ H < 0.76,$$
$$\rho_c = 0.993 + 0.007\ H, \ \text{for}\ H > 0.76$$

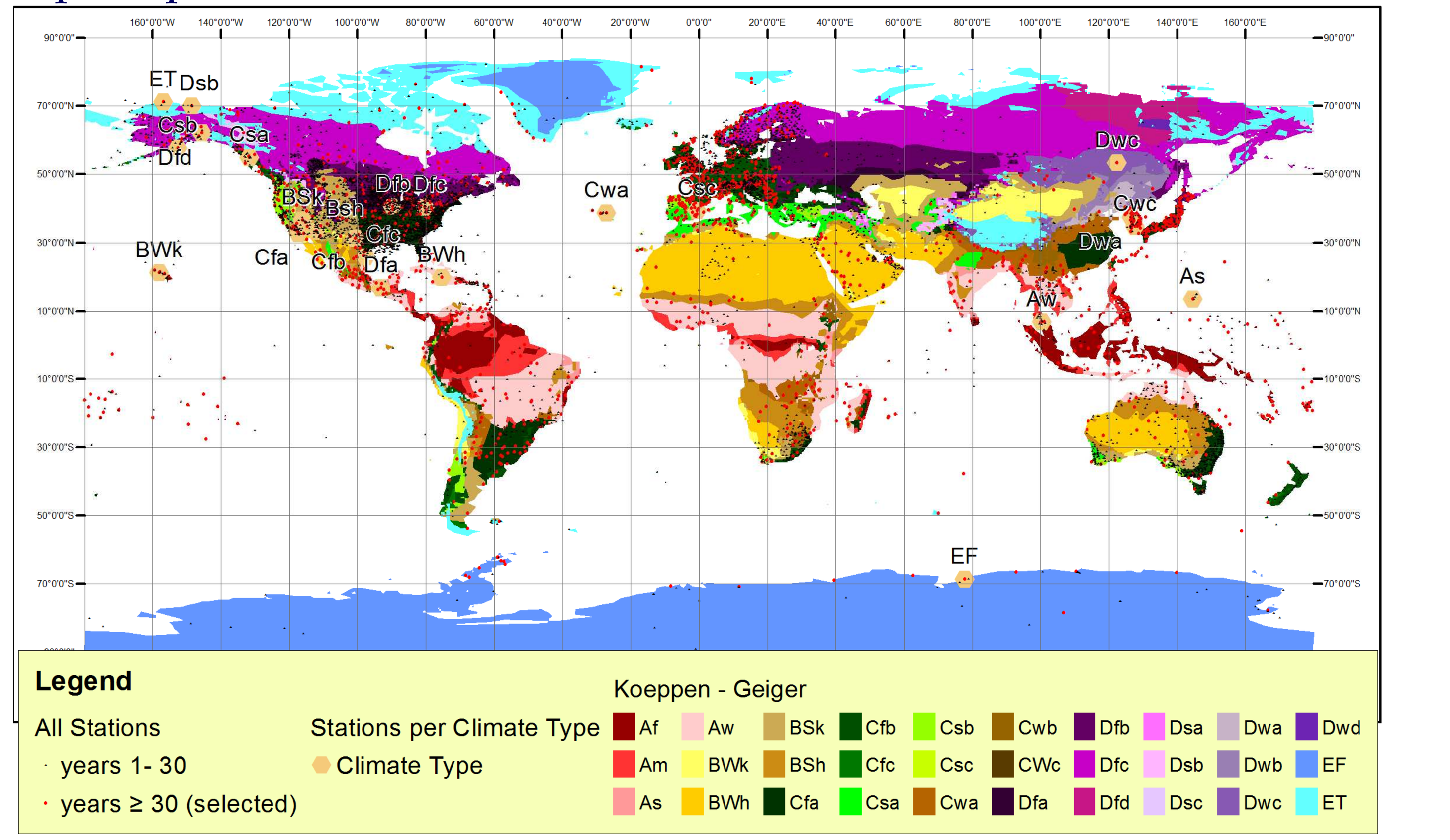
Variance:

$$\sigma^2_a = (1 - c_1 - c_2)\ \gamma_0$$
$$\sigma^2_b = c_1\ \gamma_0$$
$$\sigma^2_c = c_2\ \gamma_0$$

where γ_0: the variance of real time series and c_1 and c_2: calculated in a way that the correlation coefficient of the real time series be the same as the synthetic's for hysteresis 1 and 100.

Based on the Monte Carlo results, we estimate the "**prediction error**" or "**prediction interval**" of each 30-year mean, standard deviation, minimum and maximum values. The prediction interval is actually a measurement ranging from zero to one that compares the 30-year values observed by each station with the ones predicted from the model. In this manner, we are able to capture any large, medium or low 30-year climatic variability that occurred in approximately the last 100 years.

3. Map of spatial distribution of selected stations



Map 1: Location of wind stations that still operate. Background based [5].

4. Table of selected stations for each Koppen

Station ID	Number of years	Mean (m/s)	Standard Deviation (m/s)	Skewness (m/s)	Hurst	Height (m)	Koppen	Location
2681	70	5,5	8,0	0,8	0,87	9	ET	*Winter Trail, Alaska, N. America*
5577	42	5,8	20,7	1,9	0,86	13	EF	*Antarctica*
2255	42	1,8	3,7	1,8	0,91	433	Dwc	*Daxinganling Heilongjiang, China*
1988	70	2,7	4,0	1,0	0,66	20	Dwa	*Incheon, South Korea, Asia*
2687	42	5,7	10,9	0,8	0,93	18	Dsc	*Anchorage, Alaska, N. America*
2745	70	2,5	6,9	1,4	0,95	481	Dfc	*Chitina, Alaska, N. America*
4618	70	3,9	6,1	0,8	0,91	373	Dfb	*Pittsburgh, Pensylvania, USA*
4649	42	3,9	5,5	0,6	0,98	199	Dfa	*Springfield, Illinois, USA*
5240	42	3,7	12,0	1,3	0,83	528	Cwb	*Chiapas, Mexico, North America*
2014	42	4,1	7,6	0,9	0,88	68	Cwa	*Jeonnam, South Korea, Asia*
1075	68	4,6	10,1	0,8	0,87	55	Csb	*Azores, Portugal, Atlantic Ocean*
1051	42	3,2	7,7	1,4	0,67	62	Csa	*Valencia, Spain, Europe*
2779	70	4,8	11,2	1,0	0,88	34	Cfc	*Kodiak Island ,Alaska, N. America*
2802	74	4,1	8,8	1,1	0,89	34	Cfb	*British Columbia, Canada, N. America*
4118	68	5,0	7,8	0,4	0,92	6	Cfa	*Brownsville, Texas, USA*
4160	67	4,6	6,6	0,5	0,80	582	BSh	*San Angelo, Texas, USA*
4218	70	2,7	4,2	0,8	0,92	15	BSk	*San Diego, California, USA*
4189	42	2,9	3,6	1,1	0,87	337	BWh	*Phoenix, Arizona, USA*
4377	37	4,0	7,6	0,8	0,76	1006	BWk	*Las Vegas, Nevada, USA*
5257	70	3,6	4,9	1,2	0,94	16	Aw	*Guantanamo, Gulf of Mexico, N. America*
5609	76	4,9	6,2	0,2	0,74	3	As	*Pearl Harbor, Hawaii, Pacific Ocean*
2221	42	1,6	4,0	1,3	0,92	35	Am	*Kedah, Malaysia, Asia*
5616	42	4,4	5,5	0,7	0,68	76	Af	*Guam, Mariana Islands, Western Pacific Ocean*

Table 1: Selected stations with high credibility for each Koppen (hourly observations).

5. Quantile-Quantile plots

The observed distribution function is compared to the Gaussian one. For each observed standardized value (-3 to +3), we estimate the value of N(0,1) corresponding to the same probability of occurrence.

The results are illustrated in the Figure 1. We notice that the average as well as the Q25, Q50, Q75 quartiles almost collide with each other. Particularly, they are close to N(0,1) for values ranging from -2 to 2. This means that, for values ranging from μ-2σ to μ+2σ, the probability density is very close to normality.

- ✓ For X<μ-2σ, the results are quite different, probably for technical reasons; the anemometers do not work effectively for low values of wind speed.
- ✓ For X>μ+2σ, the fitting is also poor indicating that the tail of the distribution (corresponding to extremely high values of wind) is much heavier than the normal one.

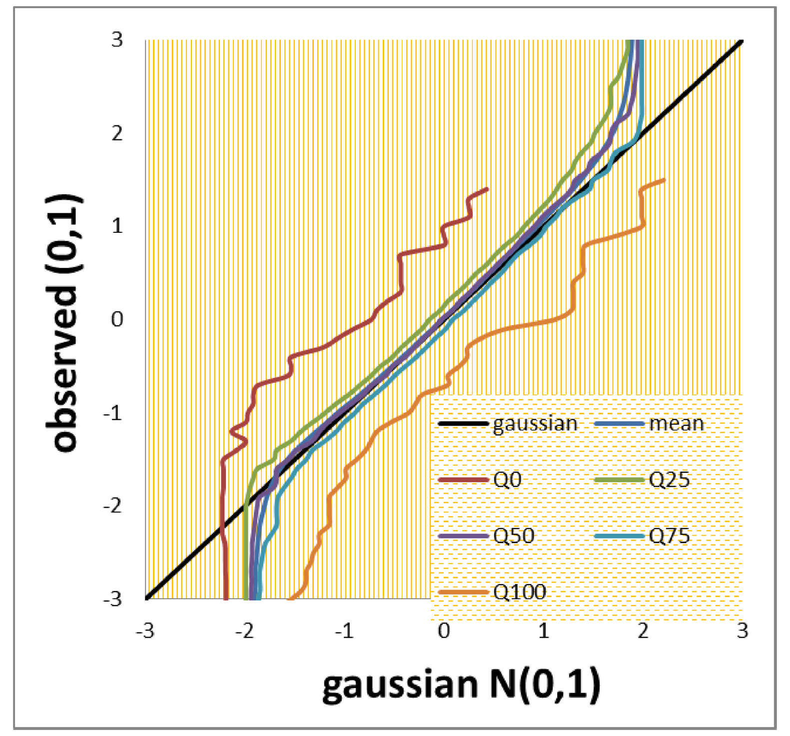


Figure 1: quantile-quantile plot of annual mean wind speed

6. Climacograms

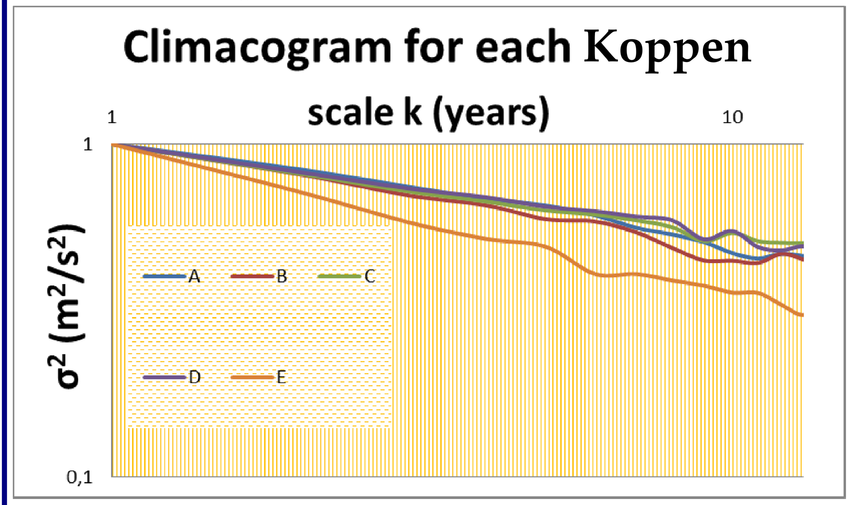


Figure 2: climacogram for each climate type.

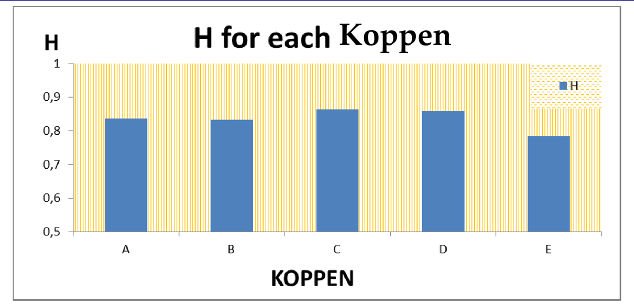


Figure 3: Hurst coefficient for each climate type.

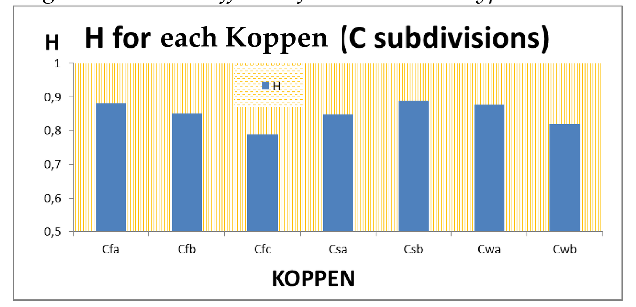


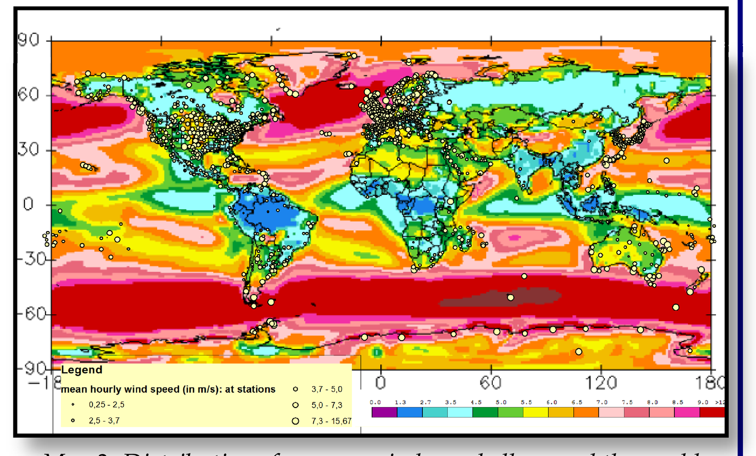
Figure 4: Hurst coefficient for C climate subdivisions

For each Koppen class, a climacogram is estimated, as described in Figure 2. By using the equation H=1+d/2, where H is the Hurst coefficient and d is the loglog slope of the climacogram, we estimate the Hurst coefficient for each climatic type. Actually, the five climatic types can be described in groups, with types A and B exhibiting H=0.83, while C and D a slightly higher value of H=0.86 and, finally, a lower value H=0.78 for E (Figure 3).

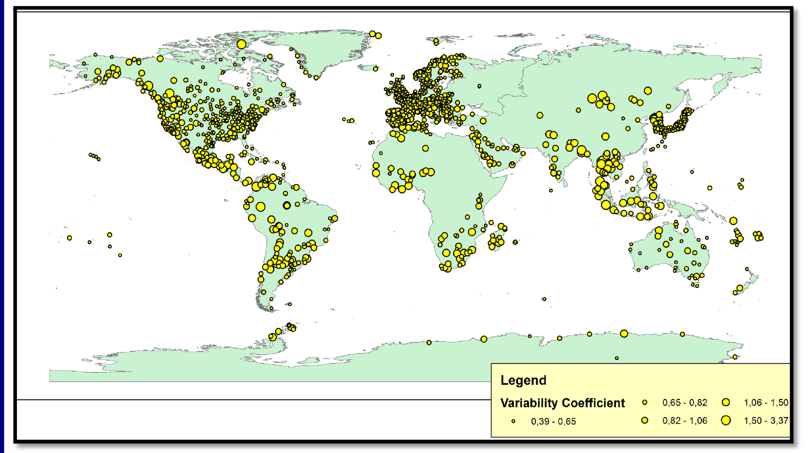
In addition, the sub-climatic types of C are further analyzed, since half of the world's selected wind stations (661) are characterized by C. The results show that the Hurst coefficient is generally constant around 0.85, except for the sub-categories Cfc, Cwb, which however represent only a few stations (22 in total) and have H=0.8 (as in [6]).

7. Average wind speed and variability coefficient

Both maps and data depicted indicate that near the Poles the wind speed is quite high and lower at the middle zones. These observations are expected by comparing them to the background map.



Map 2: Distribution of average wind speed all around the world.
Source of background: http://www.climate-charts.com/World-Climate-Maps.html#wind-speed



Map 3: Distribution of variability coefficient all around the world

The variability coefficient is quite high (>1) in Central and South America, Africa, Asia and Indonesia. In contrast, it is much less than 1 in North America and Europe.

8. Estimation of the prediction interval for a station of high credibility (1)

The following analysis is based on observations measured at a station located in Winter Trail, Alaska. This station is characterized by Koppen E climate type, and particularly sub-category ET. It is close to the sea level (height at 9 m) and near the Arctic ocean. It has measurements for 70 continuous years and it constitutes one of the most reliable stations of this climatic type and one of the most reliable of all the stations analyzed.

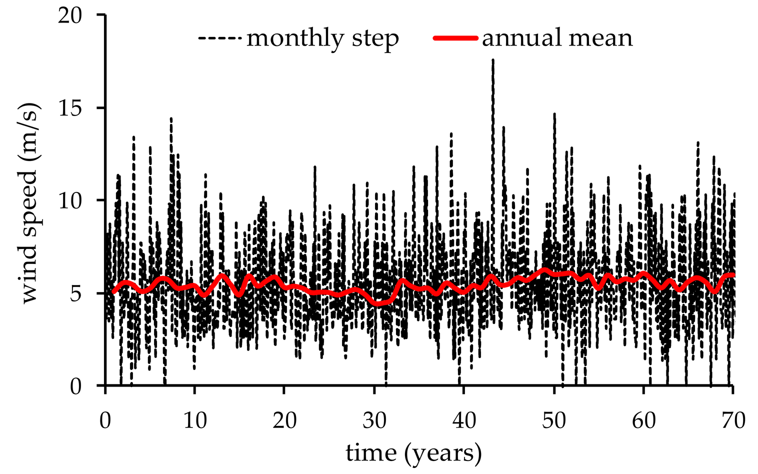


Figure 5: Time series of wind speed for the examined station.

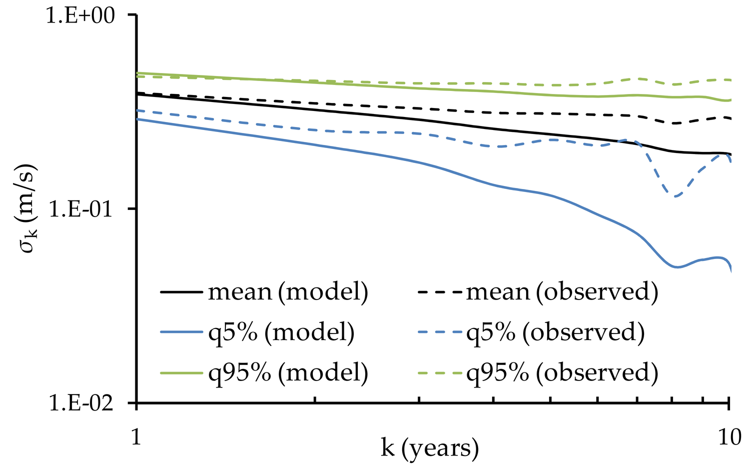


Figure 6: Climacograms for the model and observed time series of the examined station.

9. Estimation of the prediction interval for a station of high credibility (2)

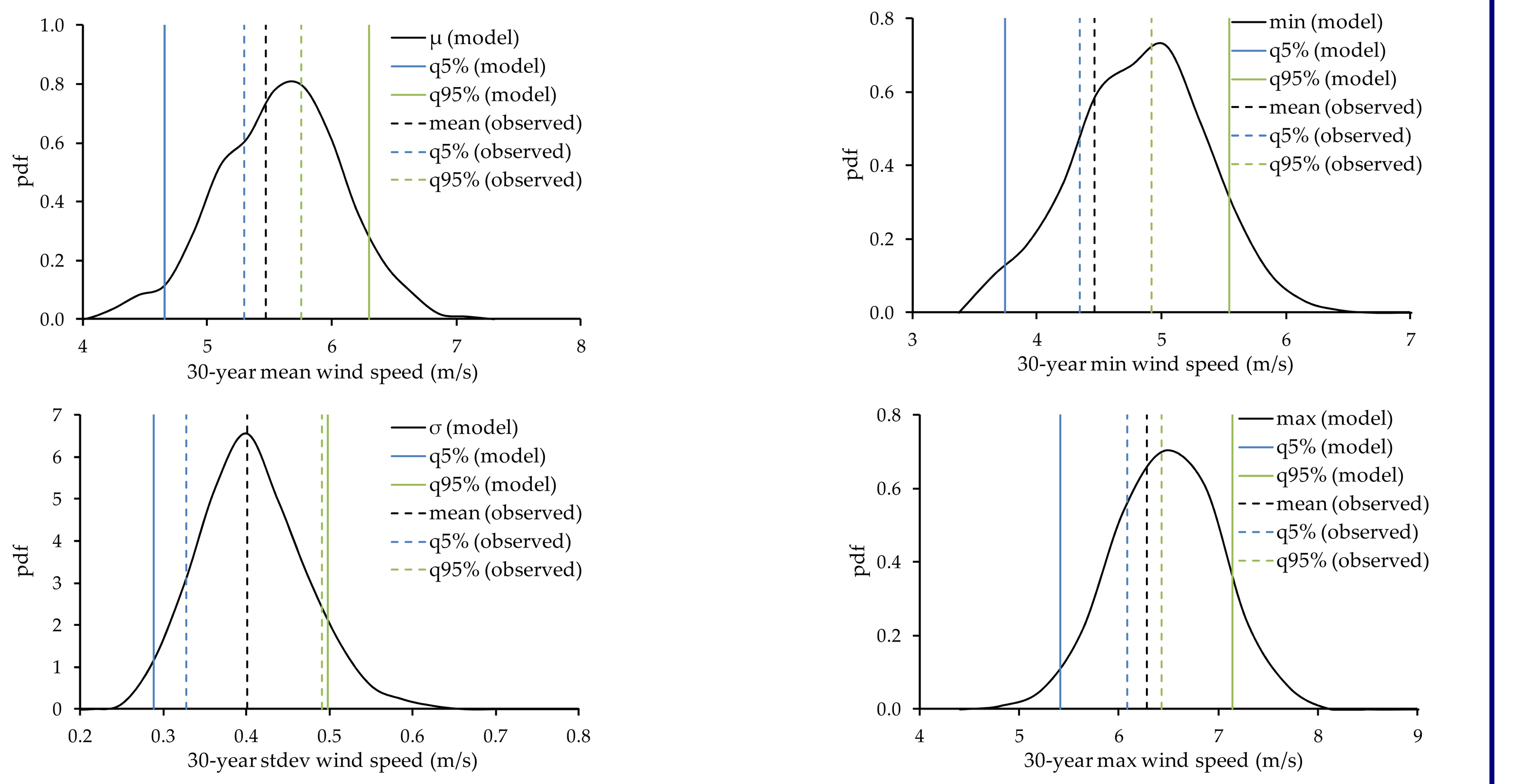
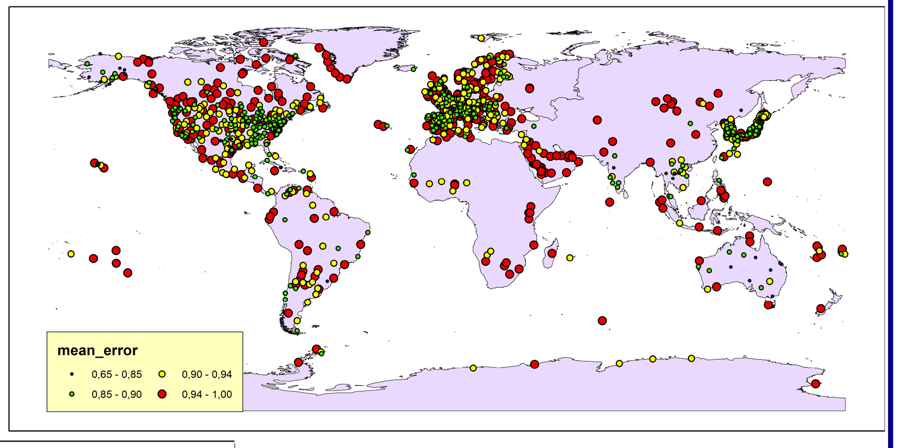


Figure 7: Estimation of prediction intervals (mean, standard deviation, minimum and maximum wind speed) for the examined station.

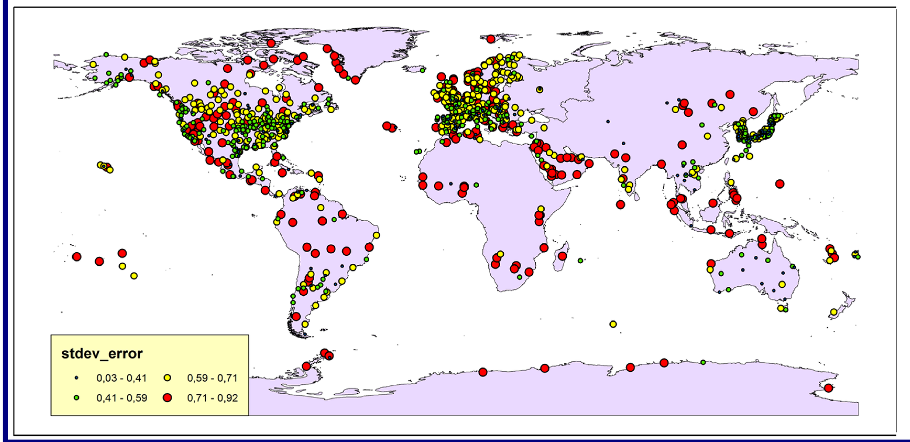
10. Prediction measures, mean & standard deviation

The prediction intervals for the annual mean values are exceptional, varying from 0.65 to 1.

- ✓ 99% of the prediction intervals of the mean values is higher than 0.75.



Map 5: prediction interval of wind velocity annual standard deviation.



Map 4: prediction interval of wind velocity annual mean.

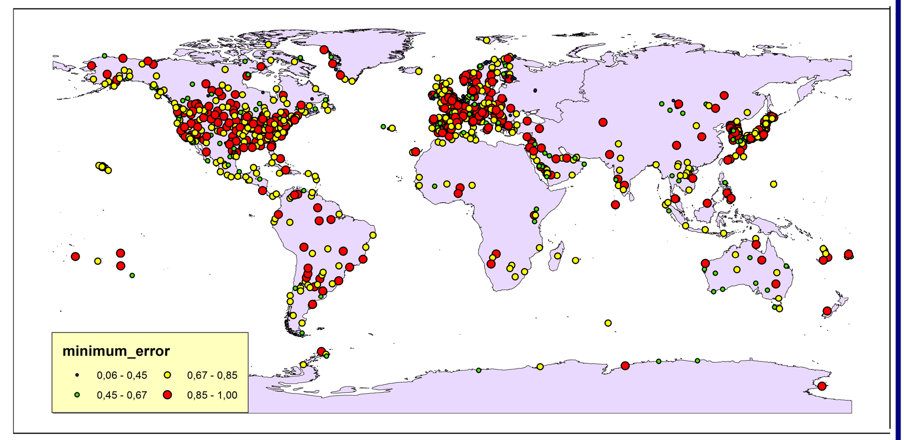
The prediction intervals for the annual standard deviation are generally lower than the ones for the mean values.

- ✓ 12% of the standard deviation prediction intervals is higher than 0.75.

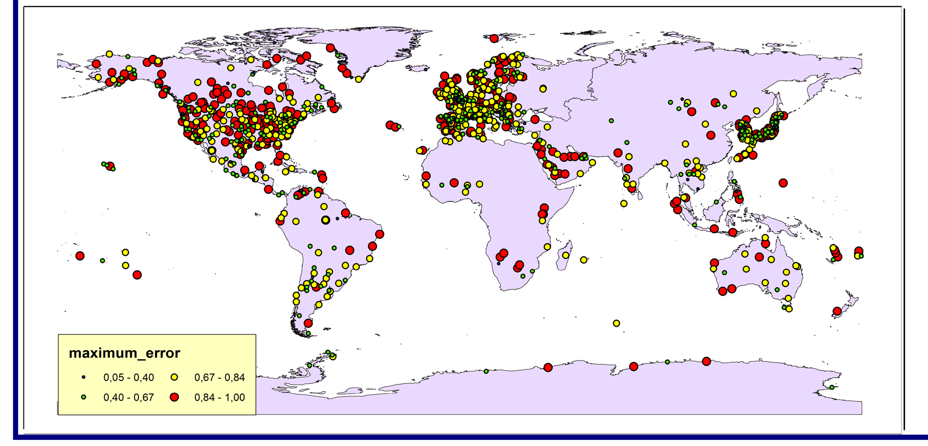
11. Prediction measures, minimum and maximum values

The prediction intervals for the annual minimum values are really good.

- ✓ 65% of the prediction intervals of the minimum values is higher than 0.75.



Map 7: prediction interval of wind velocity annual maximum values.



Map 6: prediction interval of wind velocity annual minimum values.

Stations' maximum values' prediction intervals are also very good.

- ✓ 60% of the stations' prediction interval is more than 0.75.

12. Conclusions

The three parameters mean, standard deviation and Hurst coefficient permit us to describe adequately the climatic variability of wind speed. The Monte Carlo simulation is used to quantify the stochastic uncertainties of the model. A different Hurst coefficient is used for each Koppen climatic type, since a slightly different behaviour is observed from data. Generally, the annual mean wind speed distribution for all stations is very close to normality, especially between μ-2σ and μ+2σ values. As a result, the selection of annual mean and standard deviation, in combination with the Hurst coefficient (which indicates a strong long-term persistence around the globe), constitutes a challenging way to identify the wind variability along with the over safety assumption of gaussianity.

Indeed, the mean, standard deviation, minimum and maximum values of 30-year periods appear to have quite high prediction intervals for the large majority of wind stations. Particularly:

Mean prediction interval: 90% for 71% of stations and 75% for 99% of stations.

Stdev prediction interval: 70% for 30% of stations and 50% for 80% of stations.

Min prediction interval: 80% for 53% of stations and 60% for 85% of stations.

Max prediction interval: 80% for 50% of stations and 60% for 80% of stations.

References

[1]: Philip Ball, Global Greenhouse affects air pressure - Climate-change predictions may be an underestimate, *Nature*, 59 (5), 955–958, doi:10.1038/news030317-6 , 2003.
[2]: O'Connell, P.E., D. Koutsoyiannis, H. F. Lins, Y. Markonis, A. Montanari, and T.A. Cohn, The scientific legacy of Harold Edwin Hurst (1880 – 1978), *Hydrological Sciences Journal, doi:10.1080/02626667.2015.1125998, 2015.*
[3]: D. Koutsoyiannis, The Hurst phenomenon and fractional Gaussian noise made easy, *Hydrological Sciences Journal, 47 (4), 573–595, 2002.*
[4]: Dimitriadis, P., and D. Koutsoyiannis, Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes, Stochastic Environmental Research & Risk Assessment, 29 (6), 1649–1669, doi:10.1007/s00477-015-1023-7, 2015.
[5]: Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World map of the Koppen-Geiger climate classification updated, Meteorol. Zeitschr., 15(3), 259–263, 2006.
[6]: Tsekouras, G., and D. Koutsoyiannis, Stochastic analysis and simulation of hydrometeorological processes associated with wind and solar energy, *Renewable Energy, 63, 624–633, 2014.*