Detection of Meteorological Variables in a Wind Farm Influencing the Extreme Wind Speed by Heterogeneous Granger Causality

Kateřina Hlaváčková-Schindler[1], Irene Schicker[2], Kejsi Hoxhallari[1], Claudia Plant[1]
[1] Faculty of Computer Science, University of Vienna, Vienna
[2] Geosphere Austria, Vienna, Austria

Summary

- Knowledge of meteorological parameters influencing wind speed is crucial for efficient wind power generation in a wind farm.
- We investigate temporal effects of 11 wind speed related processes within the wind farm using the Heterogeneous Graphical Granger model (HGGM).
- We use the ERA5 meteorological reanalysis to generate wind farm power production data in Eastern Austria.
- Six different scenarios for the hydrological half-year period, based on temporal intervals of moderate wind speed, low or high extreme wind speed as target variables were evaluated.
- The splitting into scenarios allows to carry out causal reasoning about possible causes of extreme wind speed in a wind farm.
- The discovered causal parameters for each scenario provide information for power generation management under adverse weather conditions.

Heterogeneous Granger Model by Minimum Message Length

The Heterogeneous Graphical Granger model (HGGM) [2], extends the multivariate Granger causality [1] for Gaussian processes to processes from exponential family. For each target variable (process) i in time t holds

$$x_i^t = \eta_i(X_{t,d}^{Lag}\beta_i')$$

where $X_{t,d}^{Lag} = (x_1^{t-d}, .., x_1^{t-1}, .., x_p^{t-d}, .., x_p^{t-1})$ p is number of processes and η_i an exponential link function. *xj Granger–causes time series xi for a given lag d, and denote xj → xi, ≡ if at least one of the d coefficients in j-th row of estimate of βi non-zero.* Method HMML [3] based on the minimum message length achieves superior precision in causal inference w.r.t. baseline methods.

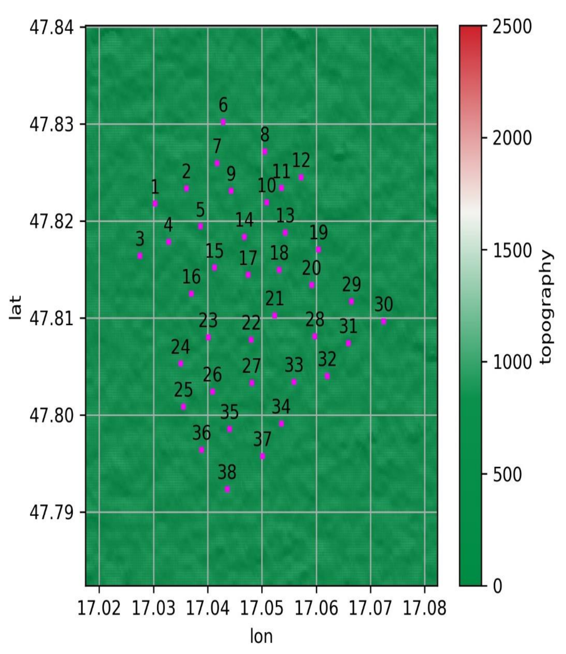


Figure 1: A map of wind park with 38 turbines near Andau, Eastern Austria

Hydrological Half-years in 2000	Winter	Summer
High extreme wind ≥15m/s	01-16 - 01-19	10-28 - 10-31
Low extreme wind ≤ 2m/s	01-01 - 01-04	07-22 - 07-25
Moderate (6m/s, 8m/s)	12-02 - 12-06 06	06-16 - 06-20

Table 1: Scenarios and time intervals considered for year 2000

Wind Farm Data and HMML

Data and scenarios

Wind power production data and 11 related meteorological parameters from the ERA5 data [4] each for 38 individual wind turbines in years 2000-2020 measured hourly are considered: geopotential in m^2/s^2 (z), boundary layer height measured in m (blh), dew point temperature at 2m in K (d2m), rel. humidity in % (rel-h), wind speed at 135 m in m/s(wspeed135m), divergence in s^{-1} (d), cloud coverage in % (cc), ozone mixing ratio in kg/kg(o3), potential vorticity in $m^2/s/kg$ (pv), temperature at 135 in K (t135m), relative vorticity in $m^2/s/kg$ (vo). **The target variable is 96-hour wind speed time series at 135m (**hub height of the turbines). HMML was used for each turbine separately as well as the following six scenarios (at 135m):

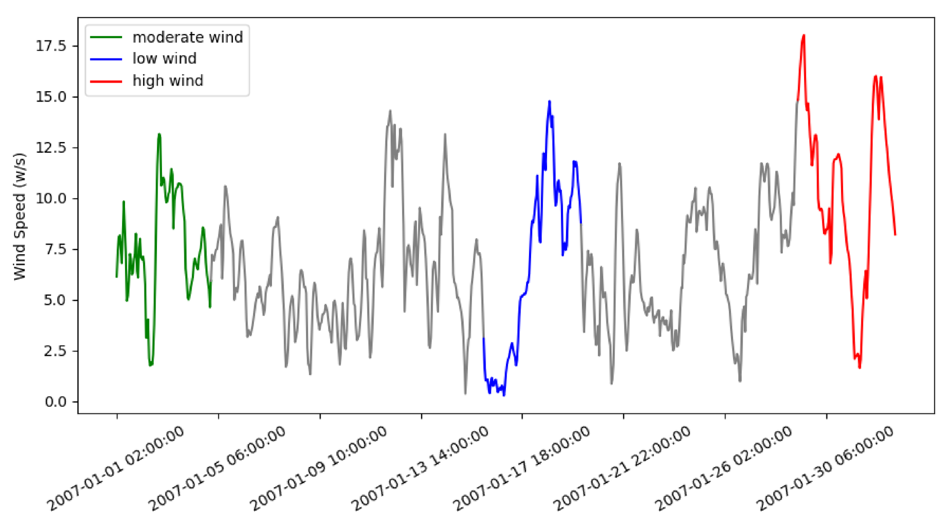


Figure 2: Wind speed in 2007. Best fitting distributions of wind speed for different scenarios

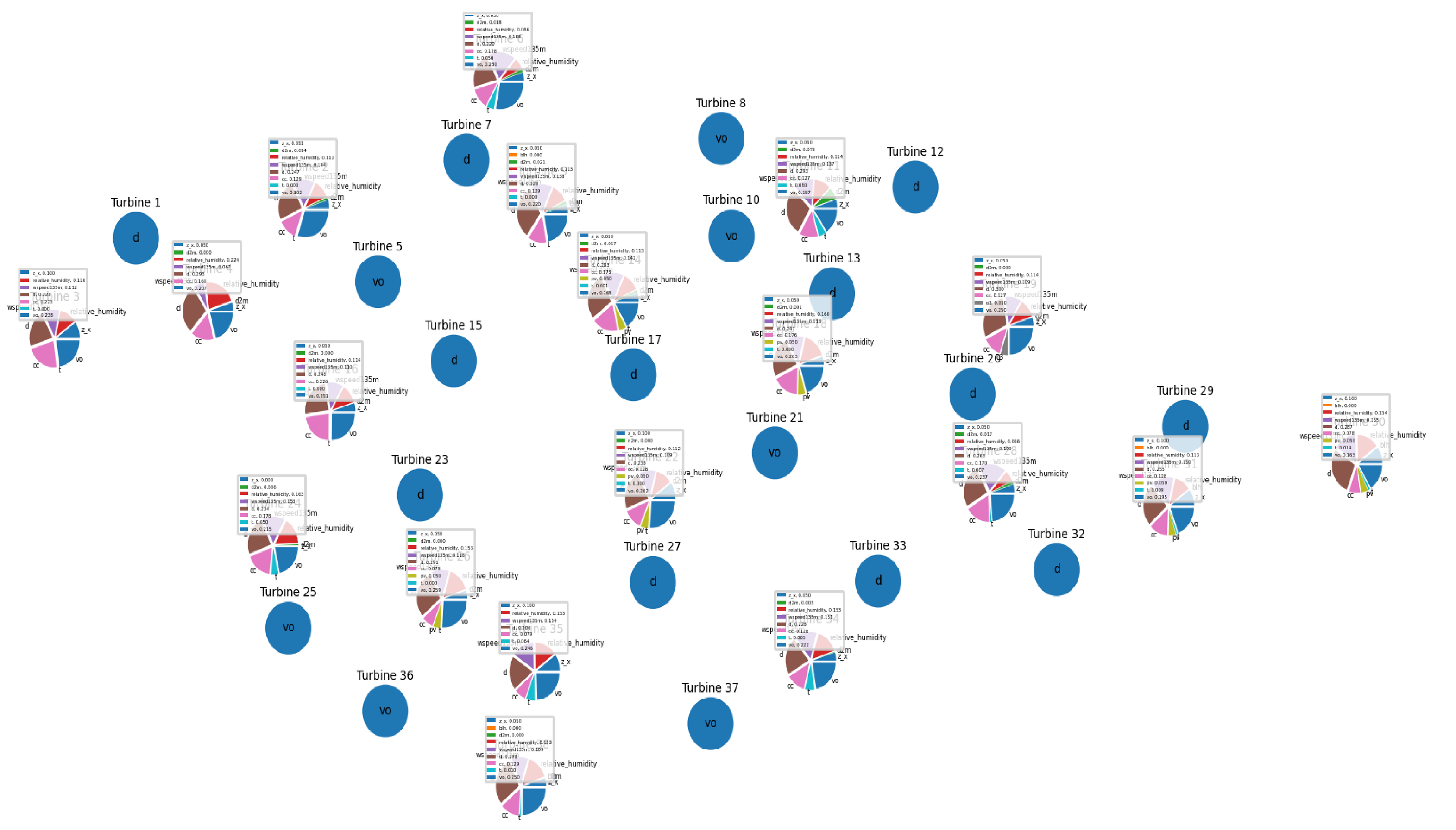
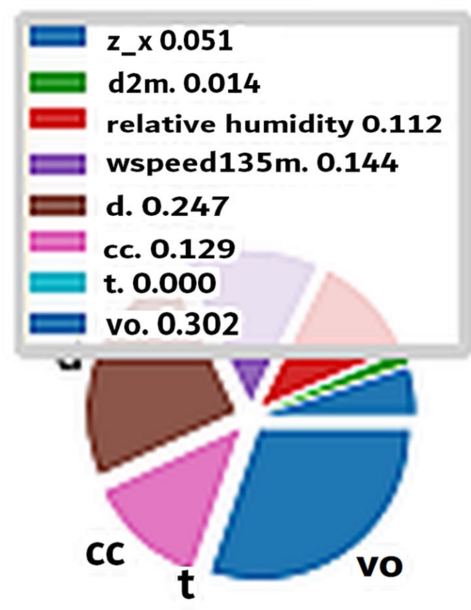


Figure 3: The farm in high wind extreme scenario in summer

The most influential variable is denoted at each turbine. The values of the three most significant variables for each scenario are in Table 2. Figure 3 illustrates the farm in high wind extreme scenario in summer. All figures for all scenarios and turbines and tables of βi-proportionalities are in supplementary material [5].



Experiments

To achieve a statistical validity of the causal values, for all scenarios, each turbine and each variable i, we calculated the corresponding arithmetical mean of βi-proportionality over all 21 years. βi-proportionality, i.e. relative strength of causality with respect to all 11 variables (all add up to 1). Our Python code together with implementation of HMML can be found in [5].

Wind speed	Winter half year	Summer half year
High extreme wind	rh (0.184), vo (0.153), cc (0.127)	d (0.261), vo (0.231), rel-h (0.134)
Low extreme wind	cc (0.212), rel-h (0.206), d (0.154)	d (0.241), cc (0.212), rel-h (0.178)
Moderate wind	d (0.363), vo (0.216), rel-h (0.193)	cc (0.374), rel-h (0.183), d (0.141)

Table 2: Most significant causal variables and their average values βi-proportionality for ws135 for each scenario

	Storm			Control		
name	length n	lag d	link function	length n	lag d	link function
Jeanett	240	*15*	*Gaussian*	240	*15*	*inverse Gaussian*
Kyrill	312	*16*	*inverse Gaussian*	312	*16*	*inverse Gaussian*
Paula	168	*10*	*gamma*	168	*10*	*gamma*
Emma	288	*16*	*inverse Gaussian*	288	*16*	*gamma*
Andrea	240	*15*	*Gaussian*	240	*15*	*gamma*
Niklas	288	*16*	*Gaussian*	288	*16*	*gamma*
Egon	120	*7*	*inverse Gaussian*	120	*7*	*Gaussian*
Herwart	96	*5*	*Gaussian*	96	*5*	*gamma*
Ana	240	*15*	*gamma*	240	*15*	*Gaussian*
Friederike	144	*8*	*inverse Gaussian*	144	*8*	*inverse Gaussian*
Hervé	240	*15*	*Gaussian*	240	*15*	*gamma*
Ciar	288	*16*	*Gaussian*	288	*16*	*Gaussian*

Figure 5: Well-known European wind storms as examples of high extreme wind speed scenarios and the best fitting exponential distributions to them

References

[1] Granger, C. W. (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica: journal of the Econometric Society, 424–438.
[2] Behzadi, S., Hlaváčková-Schindler, K., Plant C. (2019) Granger causality for heterogeneous processes. Advances in Knowledge Discovery and Data Mining: PAKDD 2019,Springer, pp. 463–475.
[3] Hlaváčková-Schindler, K., Plant C. (2020) Heterogeneous graphical Granger causality by minimum message length. Entropy 22, 12 , 1400.
[4] Herschbach, H. et al. (2020) The era5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 146, 730 (2020), 1999–2049.
[5] Supplementary material: https://anonymous.4open.science/r/Wind_farm_HMML--20BB/README.md

Acknowledgement: This work was funded in part within the Austrian Climate and Research Programme (ACRP) project KR19AC0K17614.