

EGU2020-7125

<https://doi.org/10.5194/egusphere-egu2020-7125>

EGU General Assembly 2020

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



Connecting big data from climate and biology using statistics and machine learning techniques

Susanne Pfeifer, Katharina Bülow, and Lennart Marien

Climate Service Center (GERICS), Helmholtz-Zentrum Geesthacht, Hamburg, Germany (susanne.pfeifer@hzg.de)

Within the Hamburg Cooperation project „HYBRIDS – Chances and Challenges of New Genomic Combinations“ (<https://www.biologie.uni-hamburg.de/en/forschung/verbundvorhaben/hybride-mehr-infos.html>), one subproject deals with the problem of finding relations between the existence of hybrid plant species and the climate and its variability at the same location. For this, biological and climatic data is brought together and statistical and machine learning techniques are applied to derive climatic differences between those regions where both parent species, but no hybrid species are found, and those regions where both parent species and the hybrid species were found.

Both the climate data (here daily gridded E-OBS temperature (mean, min, max) and precipitation on ~10 km grid resolution for the period of 1970 to 2006 (Haylock et al., 2008, Cornes et al., 2018)) and the plants data (Hybrid Flora of the British Isles, 700 taxa, 6 112 847 lines of data, (Stace et al., 2015)) can be considered as „big data“. However, the peculiarities of both data are very different and so are the issues to be considered when tackling the data.

We will present the first results of this interdisciplinary effort, discuss the methodological issues and elaborate on the chances and challenges of interpreting the findings.

Cornes, R., G. van der Schrier, E.J.M. van den Besselaar, and P.D. Jones. 2018: An Ensemble Version of the E-OBS Temperature and Precipitation Datasets, *J. Geophys. Res. Atmos.*, 123.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D. & M. New (2008): A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research Atmospheres*, 113(20). <https://doi.org/10.1029/2008JD010201>

Stace, C.A., Preston, C.D. & D.A. Pearman (2015): Hybrid flora of the British Isles. *Botanical Society of Britain & Ireland*. 501pp.