EGU22-11075
https://doi.org/10.5194/egusphere-egu22-11075
EGU General Assembly 2022

Communicating Uncertainties in Bi- and Multivariate Distributions

Benedikt Gräler
52°North Spatial Information Research GmbH, Münster, Germany (b.graeler@52north.org)

Any model we use is an approximation of the real world and associated with some uncertainty about the transition of the model result into the real world. This issue is even stressed in situations where models are used to predict a set of variables over time. A prominent and widely used example are runs of climate models that generate ensembles of possible future paths of our climate including several correlated variables. Typically, these ensembles are presented and assessed in an univariate approach where e.g. spaghetti plots of single variables depict the variability within the future paths. What remains hidden are the dependencies among the variables. As the correlation measures might be ambiguous and summarize the possibly complex dependence structures in a single value, we use the concept of copulas to illustrate the variability of multivariate distributions. Going back to Sklar's theorem any continuous multivariate distribution can be decomposed into its univariate margins and their copula describing the dependence between the univariate marginals. Copulas can serve two purposes in this context, to model and quantify i) the multivariate variability within the ensemble and ii) the variability in the dependence between the variables among the ensembles.

We illustrate the effect of different copula families on the inherent uncertainties in the ensemble based on synthetic data. Furthermore, we use climate predictions of the current century to identify and study the dependence and its variability in the data set. In order to increase the data and to avoid the expectation that we can do an assessment precisely for each year, the data is grouped by decades. Our approach is illustrated in R and an exploratory analysis is supported through an interactive Shiny application. The Shiny application also serves to communicate the multivariate uncertainty in the data.