Paul Wiemann and Thomas Kneib.

Using a fully Bayesian approach to analyze multiple responses and their interrelation.

In the past, most regression models have focused on one-dimensional response distributions. Thereby, the interest in statistical applications was often limited to explaining the expected value of the response as a linear combination of the available covariate information. Recently, the focus has shifted towards (i) analyzing all distributional aspects of the response distribution and thereby (ii) studying multiple response variables simultaneously and their interrelation.

Generalized additive models for location, shape and scale, or structured additive distributional regression models facilitate the first aspect by relating structured additive predictors to all distributional parameters. Besides the linear effects of the covariates, the predictors support other effect types, such as smooth functional effects, random effects, or spatial effects. This model class is not limited to response distributions from the exponential family and it supports any parametric distribution.

To match the second aspect, i.e. simultaneously analyzing multiple response variables, copula-based regression approaches are indeed helpful since they allow us to separate the marginal distributions and the dependence structure. We elaborate on a fully Bayesian regression model for multidimensional responses in which we separate the marginal distributions and the dependency structure via a copula while keeping the scope and the flexibility of distributional regression. In the model, all parameters can be modeled as functions of the covariate variables via structured additive predictors. Furthermore, we present an efficient MCMC algorithm for parameter estimation that avoids manual tuning.

Additionally, we demonstrate how the presented copula-approach can be employed to analyze data with a distributional regression model when observations are non-randomly selected. The model proposed accounts for the selection process to avoid biased estimates effectively.

We illustrate the model class presented, as well as the extension towards correcting for sample selection, with two applications on childhood undernutrition and advice weighting.