

Posudek na diplomovou práci Lukáše Drápala

Multivariate extreme value models and their application in hydrology

vypracovaný Prof. RNDr. Danielou Jaruškovou, CSc.

Hledání vhodného rozdělení pro modelování vícerozměrných dat, pro která je třeba brát v úvahu závislosti mezi jednotlivými složkami, je jedním z nejvíce potřebných (ale zároveň nejobtížnějších) problémů stochastického modelování. Pokud nepředpokládáme normalitu studovaných dat, jsme ve velmi složité situaci. V poslední době je oblíbený postup, kde se nejprve modelují marginální rozdělení jednotlivých složek a pak se modeluje závislost mezi složkami pomocí kopulí. V hydrologii a klimatologii se často objevuje problém, který spočívá v hledání vícerozměrného rozdělení, které dobře modeluje rozdělení jen pro velké hodnoty argumentů, či - li jak se říká, modelování chvostu rozdělení. Autoři John, Smith a Weissman (1992) a Davison a Smith (1990) navrhli postup, který opět spočívá v separátním modelování marginálních rozdělení a závislostí, přičemž pro modelování závislostí se používá vícerozměrné extrémální rozdělení. Jejich postup se nazývá v literatuře „point process approach“. Bohužel vícerozměrné extrémální rozdělení netvoří parametrickou třídu. Situace je obzvláště obtížná, přejeme-li si modelovat stochastické chování chvostu rozdělení vektoru, který má více než dvě složky, neboť bychom byli rádi, kdyby byla dobře zachycena, jak závislost mezi všemi složkami, tak i závislosti „marginální“. Nejčastěji používaná rozdělení jsou rozdělení logistické a rozdělení extrémální Dirichletovo. V roce 2010 byl autory Cooley, Davis a Naveau (2010) publikován článek zavádějící nový typ vícerozměrného extrémálního rozdělení, které autoři nazvali pairwise beta. V roce 2011 byl autory Ballani a Schlather (2011) publikován článek, který zobecňoval postup Cooley et al (2010) k vytváření dalších vícerozměrných extrémálních rozdělení.

Úkolem diplomanta bylo porozumět tomuto postupu, především vlastnostem nově definovaných modelů a zpracovat srážková data, která byla v minulosti zpracovaná jinými modely, viz Jarušková (2009), pomocí nových modelů, které vzniknou postupem navrženým Ballanim a Schlatherem (2011).

Diplomant se musel seznámit s částí statistiky, která není součástí studijního plánu, především s teorií vícerozměrných extrémálních rozdělení, přičemž se měl speciálně zaměřit na konstrukční postup navržený Ballanim a Schlatherem (2011). To znamená, nejen porozumět formální stránce důkazu, ale i pochopit jeho myšlenku. To se diplomantovi podařilo natolik, že by schopen pomocí něho navrhnout i další extrémální rozdělení, které je schopno v některých případech popsat závislost mezi souřadnicemi lépe než pairwise beta rozdělení. Navržení nového rozdělení považuji za největší přínos diplomové práce. Diplomant při hledání různých přístupů k modelování, našel rovněž článek Sabourin et al (2013), který doporučuje modelovat závislost pomocí směsi dvou či více rozdělení, přičemž používá Bayesův princip, to je předpokládá se znalost apriorních rozdělení pravděpodobností zastoupení, tak i apriorní rozdělení jednotlivých parametrů. Pro odhad aposteriorního rozdělení se používá MCMC. Tento přístup byl také použit v diplomové práci. Pro konkrétní výpočty používal diplomant programové prostředí R, přičemž musel některá makra upravit a vytvořit nová.

Diplomant svou práci prokázal, že je schopen samostatně porozumět odborné literatuře, navrhnout nový statistický postup, vytvořit nová makra v rámci R a rovněž zpracovat reálná data.

Literatura

Ballani F. and Schlather M.: A construction principle for multivariate extreme value distributions, *Biometrika* (2011), 1-12.

Cooley D., Davis R. and Naveau P.: The pairwise beta distribution: A flexible parametric models for extremes. *J. Multivariate Anal.* 101, (2010), 2103-2117.

Davison A.C. and Smith R.L.: Models for exceedances over high threshold (with discussion), *J.R.Statist. Soc. B* 52, 393-442.

Jarušková D.: Modeling multivariate extremes of precipitation series in northern Moravia, *Environmetrics* 20, (2009), 751-775.

Joe H., Smith R.L. and Weissman: Bivariate threshold methods for extremes, *J.R. Statist. Soc. B* 54, (1992), 171-183.

Sabourin A., Naveau P. and Fougères A.: Bayesian model averaging for multivariate extremes. *Extremes*, 16(3):325-350, 2013

V Praze 27.7.2014

Daniela Jarušková