UNIVERSITI PUTRA MALAYSIA

GENERALIZED SPLINES SMOOTHING IN GENERALIZED ADDITIVE MODELS VIA SIMULATION STUDIES

MOSTAFA BEHZADI

GENERALIZED SPLINES SMOOTHING IN GENERALIZED ADDITIVE MODELS VIA SIMULATION STUDIES

By

MOSTAFA BEHZADI

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

April 2015

DEDICATIONS

To spirits of my dear Mum, Gohar Azad and my dear Dad, Ali Behzady whom I can feel every where, every time.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

GENERALIZED SPLINES SMOOTHING IN GENERALIZED ADDITIVE MODELS VIA SIMULATION STUDIES

By

MOSTAFA BEHZADI

April 2015

Chair: Associate Professor, Mohd Bakri Bin Adam, Ph.D.

Faculty: Science

In general, real life's effects are not linear. To identify and interpret better the phenomena of real life, a flexible statistical approach is needed. Hence, in order to interpret the real phenomena, among many approaches, generalized additive model, GAM, seems to be a good tool to describe the non-linear effects. GAMs are similar to generalized linear models, GLM, in which the linear combination of explanatory variable is replaced by linear combination of scatter plot smoothers.

This research aims to study a restriction of GAM which concentrates to investigate the parameter of location. Therefore, the method in this research is based on GAM approach. Univariate generalized additive model is applied over special data which are generated from extreme value families. The simulated data are in stationary and non-stationary cases. Therefore, in stationary case, the study has focused over measuring the accuracy of estimation of parameter of location, μ. Also, in non-stationary cases the research has focused on measuring the accuracy of estimation of parameter of location, μ_t. Recall that the stationary case has no trend, while the structure of non-stationary cases are based on trends. The simulated data are belong to generalized extreme value distribution, GEV, distribution of Gumbel and special case of generalized pareto distribution, GPD. The GEV and Gumbel distributions are simulated in four types: stationary case and non-stationary cases which have the property of non-stationary in location, non-stationary in scale and non-stationary in location and scale simultaneously. The special case of GPD distribution is simulated in two types: stationary and non-stationary cases. Thus, there are ten types of special data which are investigated during this research.

Finally, to evaluate and measure of accuracy of estimation of parameter of location, a measure of spread is needed. Root mean square of errors as a measure of spread is applied for these measurements and evaluations. The result of this research strongly illustrate that the measure of accuracy of estimation of parameter of location which is obtained based on estimation of univariate GAM, is better than the alternative calculation which obtains based on maximum likelihood estimation.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk Ijazah Sarjana Sains

MODEL TAMBAHAN TERITLAK UNIVARIAT DALAM SIMULASI DATA MELAMPAU

Oleh

MOSTAFA BEHZADI

April 2015

Pengerusi: Profesor Madya, Mohd Bakri Bin Adam, Ph.D.

Fakulti: Sains

Secara umum, kesan dalam kehidupan sebenar tidak linear. Untuk mengenal pasti dan mentafsir fenomena kehidupan sebenar dengan lebih baik, ber pendekatan statistik yang fleksibel diperlukan. Oleh itu, untuk mentafsir fenomena yang sebenar, di kalangan banyak pendekatan, model tambahan umum, GAM, merupakan menjadi alat yang baik untuk menggambarkan kesan tidak linear. GAM mempungai persamaan dengan model linear teritlak, GLM, di mana kombinasi linear pembolehubah terokaan digantikan dengan kombinasi linear plot berselerak licin.

Kajian ini bertujuan untuk mengkaji sekatan terhadap GAM yang memberi tumpuan kepada kajian terhadap parameter lokasi. Oleh itu, kaedah dalam kajian ini adalah berdasarkan kepada pendekatan GAM. Univariat model teritlak digunakan ke atas data khas yang terhasil daripada famili nilai yang ekstrim. Data simulasi adalah dalam kes pegun dan tidak pegun. Dalam kes pegun, kajian ini memberi lebih tumpuan dalam mengukur ketepatan anggaran parameter lokasi, μ. Manakala, dalam kes yang tidak pegun, penyelidikan ini memfokuskan kepada mengukur ketepatan anggaran parameter lokasi, μ_t. Imbas kembali, kes yang pegun tidak mempunyai trend, manakala struktur bagi kes tidak pegun adalah berdasarkan kepada trend. Data simulasi ini tergolong kepada taburan nilai ekstrem umum, GEV, Gumbel dan kes khas taburan pareto umum, GPD. Taburan GEV dan Gumbel telah disimulasikan dalam empat jenis: kes pegun dan kes yang tidak pegun yang mana mempunyai ciri seperti tidak pegun di lokasi, dalam skala, di lokasi dan skala secara serentak. Kes khas taburan GPD telah disimulasikan dalam dua jenis: kes pegun dan tidak pegun. Oleh itu, terdapat sepuluh jenis data khas yang digunakan dalam penyelidikan ini.

Akhir sekali, untuk menilai dan mengukur ketepatan anggaran parameter lokasi, ukuran bagi serakan diperlukan. Punca bagi ralat min kuasa dua sebagai alat mengukur serakan telah digunakan kepada pengiraan dan penilaian ini. Hasil daripada kajian ini jelas menggambarkan bahawa ukuran ketepatan anggaran parameter lokasi yang diperolehi berdasarkan anggaran GAM univariat, adalah tepat daripada pengiraan alternatif yang diperoleh berdasarkan anggaran kemungkinan maksimum.

ACKNOWLEDGEMENTS

First and foremost I want to thanks God to give me another opportunity to continue my education again. Then, I would like to extend my appreciation to my dear supervisor Assoc. Prof. Dr. Mohd Bakri Bin Adam for his continuous and endless supervisions and encouragements. My deep gratitude to my co-supervisor Dr. Anwar Fitrianto of his pure assistance which never be forgotten. My special thanks goes to my dear uncle and my dear sister. I would also to express my acknowledgement to Nur Afzan, Mohammad Firdaus, Mohammad Rostami, Tofiq, Mohammad Hesam Hesamian, Hoseinali Livani, Moin Asgari, Payam Farzan and who somehow helped me during school days and at the end I would like to say thank you to my friends who have given me moral support.

I certify that a Thesis Examination Committee has met on 13 April of 2015 to conduct the final examination of Mostafa Behzadi on his thesis entitled "Generalized Splines Smoothing in Generalized Additive Models via Simulation Studies" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Master of Science.

Members of the Thesis Examination Committee were as follows:

Jayanthi A/P Arasan, Ph.D.
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Chairperson)

Mohd Rizam Bin Abu Bakar, Ph.D.
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Internal Examiner)

Ahmad Mahir Bin Razali, Ph.D.
Associate Professor
Faculty of Science and Technology
Universiti Kebangsaan Malaysia
(External Examiner)

Yong Zulina Zubairi, Ph.D.
Associate Professor
Centre For Foundation Studies in Science
University of Malaya
(External Examiner)

ZULKARNAIN ZAINAL, Ph.D.
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 13 May 2015

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science.

The members of the Supervisory Committee were as follows:

Mohd Bakri Bin Adam, Ph.D.
Assoxiate Professor
Faculty of Science
Universiti Putra Malaysia
(Chairperson)

Anwar Fitrianto, Ph.D.
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Member)

BUJANG KIM HUAT, Ph.D.
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Unievrsiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: ____________________________ Date: ____________________________

Name and Matric No: **MOSTAFA BEHZADI** **GS31372**

Declaration by Members of Supervisory Committee

This is to confirm that:
- the research conducted and the writing of this thesis was under our supervision;
- supervision responsbilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature:______________________________

Name of Chairman of Supervisory Committee
Mohd Bakri Adam, Ph.D.
Associate Professor

Signature:______________________________

Name of Member of Supervisory Committee
Anwar Fitrianto, Ph.D.
Senior Lecturer

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure		Page

LIST OF ABBREVIATIONS

GEV	Generalized Extreme Value
GPD	Generalized Pareto Distribution
OCV	Ordinary Cross Validation
GCV	Generalized Cross Validation
MSE	Mean Square of Errors
λ	Smoothing parameter
RMSD	Root Mean Square of Deviations
RMSE	Root Mean Square of Errors
CS	Cubic Spline
CRS	Cubic Regression Spline
PRS	Penalized Regression Spline
GAM	Generalized Additive Model
MLE	Maximum Likelihood Estimation
GSS	Generalized Smoothing Spline
X	Model Matrix
S	Matrix of Coefficients
B	Root of S
SC	Stationary Case
NSC	Non-Stationary Case
LM	Linear Model
GLM	Generalized Linear Model
GSS	Generalized Spline Smoothing
CRAN	Comprehensive-R-Archive-Network
CRAN	http://CRAN.R-project.org

CHAPTER 1

INTRODUCTION

1.1 Introduction

The conception of generalized additive models (GAM) as a type of regression modelling is so close to generalized linear model, hence, the preface about structures and aims of linear model and GLM, help to understand GAM better.

In this chapter there is a thorough introduction to explain the linear models, generalized linear models and additive models with particular concentrate on generalized additive models. This research focuses to show the univariate generalized additive model. The data via this research, have been simulated from extreme value distributions family. These distributions are generalized extreme value, Gumbel and generalized pareto distribution. One of the applications of extreme value distributions is in rare events (Chavez-Demoulin & Davison, 2005). As an instance, the phenomena such as floods, climates, stock marketings and engineering are included rare events. Therefore, the aim of this thesis is to model the simulation data of these phenomena by GAM.

1.2 Background

1.2.1 Linear Model

Linear model as a regression model can illustrate the expectation of a random variable, Y, as a linear summation or combination of functions of explanatory variables such as: $X_1, X_2, \ldots, X_n$ (Breslow & Clayton, 1993).

The structure of definition can be shown step by step as an example in the following model:

Example 1.1

$$\mu_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3.$$

In this relation i : number of rows $= 1, \ldots, n$ in matrix of $\mu_{n\times 1}$, and $\beta_0, \beta_1, \beta_2$ and β_3 are unknown coefficients or unknown parameters. The values of these unknown parameters should be estimated. By substitute μ_i in the $Y_i = \mu_i + \varepsilon_i$, a model will be obtained. With regards to the model below, Y_i, there is a cubic model of relationship between x and y:

$$Y_i = \mu_i + \varepsilon_i.$$

Linear models are known as statistical models where a univariate response is formulated as an amount of a linear predictor, plus a zero mean random error term. The linear predictor is depended upon some other predictor variables that should be estimated with the response variable and some unknown parameters. One of the most important properties of the linear predictor is that it is linearly depend on the parameters (Breslow & Clayton, 1993). Linear model indicates that the random variable Y, and the variable X are depend to each other by:

$$Y = \alpha + \beta x + \varepsilon,$$

in which α, is the intercept, β, is the slope of the predicted line, and ε, displays a random error. The error term has $\varepsilon \sim N(0, \sigma^2)$ (MacCullagh & Nelder, 1989). The usage of linear models is so broad. It means that, linear models are applied in many branches of sciences such as modelling tasks, analysis of designed experiments and polynomial regressions.

1.2.2 Generalized Linear Models

Generalized linear models (GLM) allows the expected value of the response to reduce the rigid linearity assumption of linear models. In other words, there is an assumption that the distribution of expected response, is smoothed by allowing it to follow up any distribution of the exponential class such as binomial, normal, gamma, and poisson etc (McCullagh & Nelder, 1989). The basis of the inference in GLMs, is centred on likelihood theory. Nelder & Wedderburn (1972) have specified any model that relates μ, expectation of response variable Y, to a linear summation of the explanatory variables: $x_1, x_2, \ldots, x_n$. Thus, in the structure of defined model

$$g(\mu) = \beta_0 + \beta_1 x + \cdots + \beta_n x_n,$$

where $\beta_1, \beta_2, ..., \beta_n$ are unknown parameters and g is a link function. Some instances of the link function can be also indicated as follow:

$$\begin{cases} g(\mu) = \mu, & \text{identity link,} \\ g(\mu) = \log(\mu), & \text{logarithmic link,} \\ g(p) = \log\left(\frac{p}{1-p}\right), \quad (0 \leq p \leq 1), & \text{logistic link or logit link,} \\ g(p) = \log\left\{-\log(1-p)\right\}, \quad (0 \leq p \leq 1), & \text{complementary log-log link,} \\ g(p) = -\log\left\{-\log(p)\right\}, \quad (0 \leq p \leq 1), & \text{the negative log-log link,} \\ g(p) = \tan\left\{\pi\left(p - \frac{1}{2}\right)\right\}, \quad (0 \leq p \leq 1), & \text{the inverse Cauchy-link,} \\ g(p) = \Phi^{-1}(p), \quad (0 \leq p \leq 1), & \text{probit.} \end{cases}$$

The first two link functions, $g(\mu) = \mu$ and $g(\mu) = \log(\mu)$, are related to random variables which have normal and poisson distributions, respectively. The last five link functions provide different families of models for dealing with, as an instance, variation in the parameter of a binomial distribution (MacCullagh & Nelder, 1989).

1.3 Additive Model

The method of additive models expresses a generalization of multiple regression model. Multiple regression model itself is a particular general linear model (Hastie & Tibshirani, 1990). A linear least square fit, in linear regression is calculated for a collection of predictors variables, X, to see a dependent variable, Y. This linear regression with k predictors can be shown in the example below:

Example 1.2

Recall that the linear regression model is:

$$Y = A + \sum_{j=1}^{k} B_j X_j + \varepsilon,$$

where A is the intercept of model and B_j represent linear effects, $j = 1, 2, \ldots, k$. Hence, for additive model, it models Y, as an additive combination of non-parametric functions of the Xs :

$$Y = A + \sum_{j=1}^{k} f_j(x_j) + \varepsilon.$$

One approach of generalizing of the multiple regression model, is to maintain the additive's content of the model. This maintenance of content, is substitution the non-parametric function with coefficient, B_j, in the linear equation. Non-parametric function means that there is no accurate and definite parametric form of function. In other words, in additive models, instead of using a single coefficient for each variable, a non-parametric function is approximated for any predictor (Hastie & Tibshirani, 1993).

1.3.1 Generalized Additive Model

The concept of generalized additive models has been structured based on ideas of additive models plus generalized linear models. This combined idea is illustrated by the following formula:

$$g\big(\mu(i)\big) = \sum_i \Big(f_i(x_i) \Big),$$

where i : number of rows $= 1, \ldots, n$ in matrix of $\mu_{n\times1}$, f_is are non-parametric functions and x_is are explanatory variables. Maximizing the quality of prediction of a dependent variable, Y, from different distributions, is the most significant aim of generalized additive models. This is done by estimating non-parametric functions of the predictor variables which are connected to the dependent variable via a link function (Hastie & Tibshirani, 1986).

GAMs are a nonparametric extension of GLMs. GAMs are used often for the cases when there are no a priori reasons for choosing a particular response function (such as linear, quadratic, etc). GAMs fulfil this duty via a smoothing function, similar to locally weighted regressions. GAMs take each predictor variable in the model, then apply knots to separate model into sections. Next step is fitting polynomial functions to each section separately. In this step, there is no particular complexity on the knots. It means that second derivatives of the separated functions are equal at both sides of the knots. The number of parameters used for such fitting is obviously more than what would be necessary for a simpler parametric fit to the same data. The degrees of freedom of the model, are usually lower than the amount of expectation from a line with so much 'wiggliness' (Wood, 2006). Indeed, this is the fundamental statistical issue associated with GAM modeling: minimizing residual deviance, while maximizing lowest possible degrees of freedom. The fitted models are directly comparable with GLMs using likelihood techniques like AIC, since the model fit is based on deviance/likelihood. Even more, all the link and error structures of GLMs are accessible and useful in GAMs. A major cause why GAMs are often less preferred than GLMs, is that the results are often difficult to interpret because no parameter values are returned (Hastie & Tibshirani, 1990).

Therefore, generalized additive model is a model that is similar to a generalized linear model in which a linear combination of explanatory variables is substituted by the linear combination of scatter plot smoother (Everitt, 2005). To be able to use GAMs practically, it needs to extend the GLM structure (Green et al., 1994). There are three main methods that can be used for GAM:

1. Representation of the smooth functions (Silverman, 1985).

2. Controlling the degree of smoothness of the functions, in order to evaluate the models with different degrees of smoothness (Wang, 1998).

3. Some methods are required to choose the most suitable degree of smoothness, if the models be applicable for merely exploratory and analytical studies (Bowman & Azzalini, 1997).

Generalized additive model investigates three general areas of research. The first area is to address the usage of basis developments of smooth functions (Craven & Wahba, 1978; Hutchinson & De Hoog, 1985). The second area is stated by estimating models with penalized likelihood maximization, in which wiggly models are more penalized in comparison to smooth models (Gu & Kim, 2002; Fahrmeir et al., 2004). The third area is implemented by applying methods that are based on cross validation by Hastie & Tibshirani (1990).

In this research, R is used as a statistical software which is accessible in CRAN (Team et al., 2005).

1.4 Problem Statement

There are some points of view about modelling by univariate generalized additive. One of this point is estimation of smoothing parameters and the other point for discussion is coefficients via a penalized regression spline. Practically, the solution of the problem of penalized regression splines is removed by penalized regression methods. The cross validation is one of the solution to estimate the smoothing parameters. Determining a suitable degree of smoothness for smooth functions, f_j, has a crucial role. This role is similar to role of coefficients in a linear regression (Wood, 2006).

The data which are applied into models in this research belong to GEV, Gumbel and GPD in stationary and non-stationary cases. It is clear that the duty of a statistical model is description of the population of a collection of data. Hence, it is really important to evaluate this ability of the model. A suitable statistical measurement to check the model accuracy is essential, otherwise, the probability of model's wrong fitting, will be raised.

Accordingly, the measurement of accuracy of model to estimate the parameters, is arguable. In other words, after implementation univariate generalized additive models in simulated extreme data, it should be evaluate the accuracy of estimated parameters. Whereas in this research it is focused on parameter of location, having another approach to estimate this parameter for comparison is necessary. Hence, MLE is applied to estimate the parameter of location via the extreme data. Therefore, this model is a linear model which its parameters obtained by maximum likelihood estimator.

As the topic of this research presents, it is worked on univariate generalized additive model. Since, it is appeared that the measure of accuracy of estimation of parameter of location among extreme value data by MLE is not enough, hence, to increase the accuracy of estimation of parameter of location, it is important to find a new method with better estimation. In this work, it is intended to solve this problem by suggesting a method based on GAM. The novelty of this thesis is to display the ability of univariate generalized additive model to calculate the accuracy of estimation of parameter of location among stationary and non-stationary GEV, Gumbel and GPD data. Then, a comparison between univariate GAM and linear model based on MLE is applied to show the accuracy of estimation of parameter of location. The benchmark for this comparison is root mean square of errors, RMSE. This research is able to show that the univariate GAM can give, an alternative promising of modelling through GEV, Gumbel and GPD models. In addition, in this thesis, a comparison between empirical λ, and the λ which is based on minimum of GCV function is illustrated for the first time.

1.5 Objective of the Thesis

This research concentrates on accuracy of estimation of parameter of locations, μ and μ_t, for stationary and non-stationary cases respectively. The estimation is accomplished by univariate GAM. Therefore, the objectives of this thesis are:

- To identify appropriate number of knots and suitable λ for cubic spline, CB, and for penalized regression spline, PRS, respectively, as a suitable empirical methods which are used in GAM.
- To introduce an alternative promising of modelling to estimate the measure of accuracy of parameter of location for GEV, Gumbel and GPD by univariate GAM.
- To calculate the MLE functions of simulated stationary and non-stationary of GEV, Gumbel and GPD data to obtain the optimized value of parameter of location to use in linear models: μ and μ_t, in order to comparison with GAM.
- To identify better estimator to the parameter of location, the estimator which is based on MLE or the estimator which is based on GAM.

1.6 Organization of the Thesis

This thesis has five chapters:

Chapter One is an introduction which explains step by step the background of generalized additive models.

Chapter Two is allocated to literature review which introduces the context of generalized additive models and some of its applications. In addition, there is a preface about stationary and non-stationary processes as a basic statistical modelling. Likewise, the extreme value theory is mentioned and reminded some discussion about simulation and its implementation in R. This chapter ends by description of method of maximum likelihood.

Chapter Three deals with the applied methodology. It discusses about idea, relevant theory, development of method and improving steps of proposed method.

Chapter Four is related to results and discussion of univariate GAM over special data. These data are belong to GEV, Gumbel and GPD functions. The data are divided into two parts, stationary and non-stationary cases.

The final chapter summarizes the obtained results and makes an overall conclusion with the glance to future work and activities in order to investigate the other parameters of extreme value functions.

BIBLIOGRAPHY

ARMSTRONG, J. S. & COLLOPY, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting* 8 69–80.

ARNOLD, B. C. (2008). Pareto and generalized pareto distributions. In *Modeling Income Distributions and Lorenz Curves*. Springer, 119–145.

BARNDORFF-NIELSEN, O. (2014). *Information and exponential families in statistical theory*. John Wiley & Sons.

BOWMAN, A. W. & AZZALINI, A. (1997). Applied smoothing techniques for data analysis .

BRESLOW, N. E. & CLAYTON, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* 88 9–25.

CAI, Y. & HAMES, D. (2010). Minimum sample size determination for generalized extreme value distribution. *Journal of Communications in Statistics, Simulation and Computation®* 40 87–98.

CASTILLO, E. & HADI, A. S. (1997). Fitting the generalized pareto distribution to data. *Journal of the American Statistical Association* 92 1609–1620.

CHAVEZ-DEMOULIN, V. & DAVISON, A. C. (2005). Generalized additive modelling of sample extremes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 54 207–222.

COLES, S., BAWA, J., TRENNER, L. & DORAZIO, P. (2001). *An introduction to statistical modeling of extreme values*, vol. 208. Springer.

COLES, S., ROBERTS, G. & JARNER, S. (2002). Computer intensive methods. *Lecture Notes, University of Lancaster* .

COLES, S. G. & DIXON, M. J. (1999). Likelihood-based inference for extreme value models. *Extremes* 2 5–23.

COOK, E. R. & PETERS, K. (1981). The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. *Tree-ring bulletin* 41 45–53.

COX, D. D. R. & ISHAM, V. (1980). *Point processes*, vol. 12. CRC Press.

COX, D. R. & HINKLEY, D. V. (1979). *Theoretical statistics*. CRC Press.

CRAVEN, P. & WAHBA, G. (1978). Smoothing noisy data with spline functions. *Numerische Mathematik* 31 377–403.

DARGAHI-NOUBARY, G. (1989). On tail estimation: An improved method. *Mathematical geology* 21 829–842.

DAVISON, A. C. & SMITH, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)* 393–442.

DE BOOR, C. ET AL. (1978). A practical guide to splines .

DE HAAN, L. & PICKANDS III, J. (1986). Stationary min-stable stochastic processes. *Probability Theory and Related Fields* 72 477–492.

DIACONIS, P. & EFRON, B. (1983). Computer-intensive methods in statistics. *Scientific American* 248 116–130.

DING, Y., CHENG, B. & JIANG, Z. (2008). A newly-discovered gpd-gev relationship together with comparing their models of extreme precipitation in summer. *Advances in Atmospheric Sciences* 25 507–516.

DOMINICI, F., MCDERMOTT, A., ZEGER, S. L. & SAMET, J. M. (2002). On the use of generalized additive models in time-series studies of air pollution and health. *American journal of epidemiology* 156 193–203.

DURRLEMAN, S. & SIMON, R. (1989). Flexible regression models with cubic splines. *Statistics in medicine* 8 551–561.

EFRON, B. & HINKLEY, D. V. (1978). Assessing the accuracy of the maximum likelihood estimator: Observed versus expected fisher information. *Biometrika* 65 457–483.

EILERS, P. H. & MARX, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical science* 89–102.

EILERS, P. H. & MARX, B. D. (2010). Splines, knots, and penalties. *Wiley Interdisciplinary Reviews: Computational Statistics* 2 637–653.

EMBRECHTS, P. (1997). *Modelling extremal events: for insurance and finance*, vol. 33. Springer.

EVERITT, B. S. (2005). *Generalized Additive Model*. Wiley Online Library.

FAHRMEIR, L., KNEIB, T. & LANG, S. (2004). Penalized structured additive regression for space-time data: a bayesian perspective. *Statistica Sinica* 14 731–762.

FISHER, R. A. & TIPPETT, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 24. Cambridge Univ Press, 180–190.

FORTIN, M.-J., JACQUEZ, G. M. & SHIPLEY, B. (2006). Computer-intensive methods. *Encyclopedia of Environmetrics* .

GOLUB, G. H. & VAN VAN LOAN, C. F. (1996). Matrix computations (johns hopkins studies in mathematical sciences) .

GREEN, P. J., SILVERMAN, B. W., SILVERMAN, B. W. & SILVERMAN, B. W. (1994). *Nonparametric regression and generalized linear models: a roughness penalty approach*. Chapman & Hall London.

GU, C. (1992). Cross-validating non-gaussian data. *Journal of Computational and Graphical Statistics* 1 169–179.

GU, C. (2002). *Smoothing spline ANOVA models*. Springer.

GU, C. & KIM, Y.-J. (2002). Penalized likelihood regression: general formulation and efficient approximation. *Canadian Journal of Statistics* 30 619–628.

GU, C. & WAHBA, G. (1991a). Discussion: multivariate adaptive regression splines. *The Annals of Statistics* 19 115–123.

GU, C. & WAHBA, G. (1991b). Minimizing gcv/gml scores with multiple smoothing parameters via the newton method. *SIAM Journal on Scientific and Statistical Computing* 12 383–398.

GUMBEL, E. J. (1954). *Statistical theory of extreme values and some practical applications: a series of lectures*, vol. 33. US Government Printing Office Washington.

GUMBEL, E. J. (1958). *Statistics of extremes*. Dover Publications. com.

GUMBEL, E. J. (1960). Multivariate extremal distributions. *Bull. Inst. Internat. de Statistique* 37 471–475.

GUMBEL, E. J. & VON SCHELLING, H. (1950). The distribution of the number of exceedances. *The Annals of Mathematical Statistics* 21 247–262.

HAMILTON, J. D. (1994). *Time series analysis*, vol. 2. Cambridge Univ Press.

HASTIE, T. & TIBSHIRANI, R. (1986). Generalized additive models. *Statistical science* 297–310.

HASTIE, T. & TIBSHIRANI, R. (1987). Generalized additive models: some applications. *Journal of the American Statistical Association* 82 371–386.

HASTIE, T. & TIBSHIRANI, R. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society. Series B (Methodological)* 757–796.

HASTIE, T. J. & TIBSHIRANI, R. J. (1990). *Generalized additive models*, vol. 43. CRC Press.

HOSKING, J., WALLIS, J. R. & WOOD, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics* 27 251–261.

HOSKING, J. R. & WALLIS, J. R. (1987). Parameter and quantile estimation for the generalized pareto distribution. *Technometrics* 29 339–349.

HUTCHINSON, M. (1995). Interpolating mean rainfall using thin plate smoothing splines. *International journal of geographical information systems* 9 385–403.

HUTCHINSON, M. F. & DE HOOG, F. (1985). Smoothing noisy data with spline functions. *Numerische Mathematik* 47 99–106.

HYNDMAN, R. J. & KOEHLER, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting* 22 679–688.

KARL, T. R. & KNIGHT, R. W. (1998). Secular trends of precipitation amount, frequency, and intensity in the united states. *Bulletin of the American Meteorological society* 79 231–241.

KIM, Y.-J. & GU, C. (2004). Smoothing spline gaussian regression: more scalable computation via efficient approximation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 66 337–356.

KOTZ, S. & NADARAJAH, S. (2000). *Extreme value distributions*. World Scientific.

LOCKHART, R. A. & STEPHENS, M. A. (1994). Estimation and tests of fit for the three-parameter weibull distribution. *Journal of the Royal Statistical Society. Series B (Methodological)* 491–500.

LORANCE, P., PAWLOWSKI, L. & TRENKEL, V. M. (2010). Standardizing blue ling landings per unit effort from industry haul-by-haul data using generalized additive models. *ICES Journal of Marine Science: Journal du Conseil* 67 1650–1658.

MACCULLAGH, P. & NELDER, J. A. (1989). *Generalized linear models*, vol. 37. CRC press.

MARRA, G. & WOOD, S. N. (2011). Practical variable selection for generalized additive models. *Computational Statistics & Data Analysis* 55 2372–2387.

MARRA, G. & WOOD, S. N. (2012). Coverage properties of confidence intervals for generalized additive model components. *Scandinavian Journal of Statistics* 39 53–74.

MARX, B. D. & EILERS, P. H. (1998). Direct generalized additive modeling with penalized likelihood. *Computational Statistics & Data Analysis* 28 193–209.

MCCULLAGH, P. & NELDER, J. A. (1989). *Generalized linear models*, vol. 2. Chapman and Hall London.

MCLEAN, M. W., HOOKER, G., STAICU, A.-M., SCHEIPL, F. & RUPPERT, D. (2014). Functional generalized additive models. *Journal of Computational and Graphical Statistics* 23 249–269.

NELDER, J. A. & WEDDERBURN, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A (General)* 370–384.

PENG, L. (1998). Asymptotically unbiased estimators for the extreme-value index. *Statistics & Probability Letters* 38 107–115.

PRESCOTT, P. & WALDEN, A. (1980). Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika* 67 723–724.

PRESCOTT, P. & WALDEN, A. (1983). Maximum likeiihood estimation of the parameters of the three-parameter generalized extreme-value distribution from censored samples. *Journal of Statistical Computation and Simulation* 16 241–250.

RAMESH, N. & DAVISON, A. (2002). Local models for exploratory analysis of hydrological extremes. *Journal of Hydrology* 256 106–119.

RIPLEY, B. D. (2009). *Stochastic simulation*, vol. 316. Wiley. com.

SHI, D. (1995). Fisher information for a multivariate extreme value distribution. *Biometrika* 644–649.

SIJBERS, J., DEN DEKKER, A. J., SCHEUNDERS, P. & VAN DYCK, D. (1998). Maximum-likelihood estimation of rician distribution parameters. *Medical Imaging, IEEE Transactions on* 17 357–361.

SILVERMAN, B. W. (1985). Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *Journal of the Royal Statistical Society. Series B (Methodological)* 1–52.

SMITH, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika* 72 67–90.

SMITH, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science* 4 367–377.

SMITH, R. L. (1990). Extreme value theory. *Handbook of applicable mathematics* 7 437–471.

TANNER, M. A. (1991). *Tools for statistical inference: observed data and data augmentation methods*. Springer-Verlag New York.

TEAM, R. D. C. ET AL. (2005). R: A language and environment for statistical computing.

VOGEL, R. M. (1986). The probability plot correlation coefficient test for the normal, lognormal, and gumbel distributional hypotheses. *Water Resources Research* 22 587–590.

WAHBA, G. (1975). Smoothing noisy data with spline functions. *Numerische Mathematik* 24 383–393.

WAHBA, G. (1980). Spline bases, regularization, and generalized cross-validation for solving approximation problems with large quantities of noisy data. *Approximation theory III* 2.

WAHBA, G. (1985). A comparison of gcv and gml for choosing the smoothing parameter in the generalized spline smoothing problem. *The Annals of Statistics* 1378–1402.

WAHBA, G. (1990). *Spline models for observational data*. 59. Siam.

WANG, Y. (1998). Smoothing spline models with correlated random errors. *Journal of the American Statistical Association* 93 341–348.

WATKINS, D. S. (2004). *Fundamentals of matrix computations*, vol. 64. John Wiley & Sons.

WOOD, S. (2006). *Generalized additive models: an introduction with R*. CRC press.

WOOD, S. N. (2000). Modelling and smoothing parameter estimation with multiple quadratic penalties. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 62 413–428.

WOOD, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65 95–114.

WOOD, S. N. (2008). Fast stable direct fitting and smoothness selection for generalized additive models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 70 495–518.

WOOD, S. N. & AUGUSTIN, N. H. (2002). Gams with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecological modelling* 157 157–177.