

## Practical Smoothing

This is a practical guide to P-splines, a simple, flexible, and powerful tool for smoothing. P-splines combine regression on B-splines with simple, discrete, roughness penalties. They were introduced by the authors in 1996 and have been used in many diverse applications. The regression basis makes it straightforward to handle non-normal data, like in generalized linear models. The authors demonstrate optimal smoothing, using mixed model technology and Bayesian estimation, in addition to classical tools like cross-validation and AIC, covering theory and applications with code in R. Going far beyond simple smoothing, they also show how to use P-splines for regression on signals, varying-coefficient models, quantile and expectile smoothing, and composite links for grouped data. Penalties are the crucial elements of P-splines; with proper modifications they can handle periodic and circular data as well as shape constraints. Combining penalties with tensor products of B-splines extends these attractive properties to multiple dimensions. The appendices offer a systematic comparison to other smoothers.

**PAUL H. C. EILERS** is Professor Emeritus of Genetical Statistics at the Erasmus University Medical Center, Rotterdam, The Netherlands. He received his PhD in biostatistics. His research interests include high-throughput genomic data analysis, chemometrics, smoothing, longitudinal data analysis, survival analysis, and statistical computing. He has published extensively on these subjects.

**BRIAN D. MARX** is Professor in the Department of Experimental Statistics at Louisiana State University. He received his PhD in statistics. His main research interests include smoothing, ill-conditioned regression problems, and high-dimensional chemometric applications, and he has numerous publications on these topics. He is currently serving as coordinating editor for the journal *Statistical Modelling*. He is coauthor of two books and is a Fellow of the American Statistical Association.

Cambridge University Press  
978-1-108-48295-0 — Practical Smoothing  
Paul H.C. Eilers , Brian D. Marx  
Frontmatter  
[More Information](#)

---

# Practical Smoothing

## The Joys of P-splines

Paul H. C. Eilers  
*Erasmus University Medical Center*

Brian D. Marx  
*Louisiana State University*



CAMBRIDGE  
UNIVERSITY PRESS

Cambridge University Press  
978-1-108-48295-0 — Practical Smoothing  
Paul H.C. Eilers, Brian D. Marx  
Frontmatter  
[More Information](#)

## CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom  
One Liberty Plaza, 20th Floor, New York, NY 10006, USA  
477 Williamstown Road, Port Melbourne, VIC 3207, Australia  
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,  
New Delhi – 110025, India  
79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9781108482950](http://www.cambridge.org/9781108482950)

DOI: 10.1017/9781108610247

© Paul H. C. Eilers and Brian D. Marx 2021

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2021

Printed in the United Kingdom by TJ Books Ltd., Padstow Cornwall

*A catalogue record for this publication is available from the British Library.*

*Library of Congress Cataloging-in-Publication Data*

Names: Eilers, Paul H. C., 1948– author. | Marx, Brian D., 1960– author.

Title: Practical smoothing : the joys of P-splines / Paul H. C. Eilers,  
Erasmus University Medical Center, Brian D. Marx, Louisiana State University.

Description: Cambridge, UK ; New York, NY : Cambridge University Press,  
2021. | Includes bibliographical references and index.

Identifiers: LCCN 2020016638 (print) | LCCN 2020016639 (ebook) |  
ISBN 9781108482950 (hardback) | ISBN 9781108610247 (epub)

Subjects: LCSH: Smoothing (Statistics) | Spline theory.

Classification: LCC QA278 .E397 2021 (print) | LCC QA278 (ebook) |  
DDC 511/.4223–dc23

LC record available at <https://lcn.loc.gov/2020016638>

LC ebook record available at <https://lcn.loc.gov/2020016639>

ISBN 978-1-108-48295-0 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Cambridge University Press  
978-1-108-48295-0 — Practical Smoothing  
Paul H.C. Eilers , Brian D. Marx  
Frontmatter  
[More Information](#)

---

To my family (PE)  
For Arnold; to Leopold (BDM)

Cambridge University Press  
978-1-108-48295-0 — Practical Smoothing  
Paul H.C. Eilers , Brian D. Marx  
Frontmatter  
[More Information](#)

---

## Contents

---

<i>Preface</i>	<i>page xi</i>
<b>1 Introduction</b>	<b>1</b>
<b>2 Bases, Penalties, and Likelihoods</b>	<b>6</b>
2.1 Linear and Polynomial Regression	6
2.2 B-splines	9
2.3 Penalized Least Squares	14
2.4 Interpolation and Extrapolation	18
2.5 Derivatives	20
2.6 The Effective Dimension	21
2.7 Standard Errors	23
2.8 Heavy Smoothing and Polynomial Limits	24
2.9 P-splines as a Parametric Model	24
2.10 Whittaker: P-splines without B-splines	26
2.11 Equivalent Kernels	26
2.12 Smoothing of a Non-normal Response	28
2.12.1 Poisson Smoothing	28
2.12.2 Binomial Smoothing	31
2.12.3 GLM Effective Dimension and Standard Errors	32
2.13 Notes and Details	34
<b>3 Optimal Smoothing in Action</b>	<b>36</b>
3.1 Cross-Validation	37
3.2 Akaike's Information Criterion	38
3.3 Density Estimation	40
3.4 Mixed Models	41
3.5 Bayesian P-splines	45
3.6 Dangers of Automatic Smoothing	51

3.7	L- and V-curves	54
3.8	Transformation of the Independent Variable	56
3.9	Notes and Details	58
<b>4</b>	<b>Multidimensional Smoothing</b>	<b>59</b>
4.1	Generalized Additive Models	60
4.2	Varying Coefficient Models	63
4.3	Tensor Product Models	67
4.4	Tensor Product Bases	69
4.5	Two-Dimensional Penalties	71
4.6	Interpolation and Extrapolation	73
4.7	Smoothing on Large Grids	74
4.8	Generalized Two-Dimensional Smoothing	75
4.9	Optimal Two-Dimensional Smoothing	77
4.10	Issues with Isotropic Smoothing	79
4.11	Higher Dimensions	79
4.12	Nested Bases and PS-ANOVA	79
4.13	Notes and Details	83
<b>5</b>	<b>Smoothing of Scale and Shape</b>	<b>84</b>
5.1	Quantile Smoothing	85
5.2	Expectile Smoothing	91
5.3	Models for Shape and Scale Parameters	97
5.4	Baseline Estimation	101
5.5	Notes and Details	102
<b>6</b>	<b>Complex Counts and Composite Links</b>	<b>103</b>
6.1	Histograms with Wide Bins	104
6.2	Histograms and Scale Transformation	107
6.3	Individual Censoring	109
6.4	Latent Mixtures	110
6.5	Notes and Details	112
<b>7</b>	<b>Signal Regression</b>	<b>114</b>
7.1	A Chemical Calibration Problem	115
7.2	Extensions to the Generalized Linear Model	120
7.3	Multidimensional Signal Regression	122
7.4	Further Extensions	126
7.5	Notes and Details	129
<b>8</b>	<b>Special Subjects</b>	<b>131</b>
8.1	The Proper B-spline Basis	132
8.2	Harmonic Smoothing	132



	Contents	ix
8.3	Circular Smoothing	135
8.4	Signal Separation with Penalties	138
8.5	Double Penalties	141
8.6	Piecewise Constant Smoothing	143
8.7	Shape Constraints	146
8.8	Variable and Adaptive Penalties	152
8.9	Survival Analysis and Mortality Modeling	155
8.10	Notes and Details	158
	<b>Appendix A P-splines for the Impatient</b>	159
	<b>Appendix B P-splines and Competitors</b>	161
	<b>Appendix C Computational Details</b>	168
	<b>Appendix D Array Algorithms</b>	174
	<b>Appendix E Mixed Model Equations</b>	176
	<b>Appendix F Standard Errors in Detail</b>	182
	<b>Appendix G The Website</b>	184
	<i>References</i>	188
	<i>Index</i>	196

Cambridge University Press  
978-1-108-48295-0 — Practical Smoothing  
Paul H.C. Eilers , Brian D. Marx  
Frontmatter  
[More Information](#)

---

## Preface

---

It is rare to see the word *joy* in the title of a statistics book. The only example we know is *The Joy of Statistics* by Selvin (2019). We will do our best to be a worthy runner-up.

Our first goal is to present a practical, consistent, and versatile tool. We are not modest: our claim is that P-splines are the best smoother on the planet. Our motto is “show, don’t tell,” so we present many applications, illustrating the power of this tool. For the skeptics, we provide an appendix with a tabular comparison to the competition. However, our confidence stems from external validation: the many citations that our original paper received in the past quarter century or so. Many of them appeared among the work of other statisticians, yet the majority came from applications in a wide variety of fields of science. Our tool is being used and appreciated.

The plan to write this book is over a decade old; the delay is due to a variety of reasons. Fortunately, over the last few years progress toward this book project changed, and the writing wheel starting spinning. Such a long incubation period can have its drawbacks, but we now have the benefit of writing a much riper book, one that reflects our own distilled understanding of the subject, in combination with a much more mature literature base.

In the chapters to come, we present theory and applications in an easygoing way. We show many graphs to illustrate concepts and applications. The R code for every graph is available as a set of independent short programs on the website <https://psplines.bitbucket.io>. Not only does this make our results reproducible, but it also gives the reader an accessible and smooth start to analyze their own data. A small R package (JOPS) provides supporting functions.

We have aimed to keep the programs small and self-contained. We make little use of large packages like `mgcv` and `gamlss`. Of course, these packages

are excellent, and we do not want to compete with them. We believe that our approach provides better insight in the simplicity and power of P-splines. We have done our best to make our programs easy to read.

We have been working together with many colleagues in a number of countries, many of whom we have met at many editions of the International Workshop on Statistical Modelling. Our first contact was at the Workshop in Trento in 1989. Over the years, an active band has formed, of sisters and brothers who enjoy getting together each year, especially in Spain, and discussing developments in theory and practice of P-splines, which indeed has brought joy to our work together.

We thank (in alphabetical order): Martin Boer, Giancarlo Camarda, Iain Currie, Maria Durbán, Gianluca Frasso, Jutta Gampe, Oswaldo Gressani, Philippe Lambert, Dae-Jin Lee, and Maria-Xosé Rodríguez Alvarez.

Iain Currie and Jutta Gampe reviewed a late version of the manuscript. This was a mixed blessing. They are excellent editors, and it was sobering to read their many comments. But they led to an improved book.

We also thank Lauren Cowles, our main contact with Cambridge University Press. She was patient, consistent, and supportive for well over a decade, and sometimes it looked as if she was more confident about the final result than we were ourselves.

Over the years, when we presented several courses on P-splines and B-splines, people asked where their names came from. The prosaic explanation is that the *P* stands for penalties and B-splines is the name de Boor gave to his invention. Jokingly we said that *P* stands for Paul and *B* for Brian. Now that the book is ready, reflecting our long experience, we are confident to say that *B* stands for *best* and *P* for *practical*.

We wish you an enjoyable and very useful journey in the land of P-splines. Don't forget to take a look at the website <https://psplines.bitbucket.io>, where you will find a package with data and useful functions, as well as the scripts for all graphs in this book. Play with them to improve your understanding, and adapt them to analyze your own data. Please acknowledge us in a fair way.