

# Applied Functional Data Analysis: Methods and Case Studies

James O. Ramsay and Bernard W. Silverman

J.O. Ramsay  
Department of Psychology  
McGill University  
Montreal, Quebec H3A 1B1  
Canada  
ramsay@psych.mcgill.ca

B.W. Silverman  
Department of Mathematics  
University of Bristol  
Bristol BS8 1TW  
United Kingdom  
b.w.silverman@bristol.ac.uk

Library of Congress Cataloging-in-Publication Data

Ramsay, J.O. (James O).

Applied functional data analysis : methods and case studies / J.O. Ramsay, B.W. Silverman.  
p. cm. — (Springer series in statistics)

Includes bibliographical references and index.

ISBN 0-387-95414-7 (pbk. : alk. paper)

I. Multivariate analysis. I. Silverman, B.W., 1952– II. Title. III. Series.

QA278 .R35 2002

519.5 35—dc21

2002022924

ISBN 0-387-95414-7

Printed on acid-free paper.

© 2002 Springer-Verlag New York, Inc.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed in the United States of America.

9 8 7 6 5 4 3 2

SPIN 10944752

Typesetting: Pages created by the authors using a Springer T<sub>E</sub>X macro package.

[www.springer-ny.com](http://www.springer-ny.com)

Springer-Verlag New York Berlin Heidelberg  
A member of BertelsmannSpringer Science+Business Media GmbH

# Preface

Almost as soon as we had completed our previous book *Functional Data Analysis* in 1997, it became clear that potential interest in the field was far wider than the audience for the thematic presentation we had given there. At the same time, both of us rapidly became involved in relevant new research involving many colleagues in fields outside statistics.

This book treats the field in a different way, by considering case studies arising from our own collaborative research to illustrate how functional data analysis ideas work out in practice in a diverse range of subject areas. These include criminology, economics, archaeology, rheumatology, psychology, neurophysiology, auxology (the study of human growth), meteorology, biomechanics, and education—and also a study of a juggling statistician.

Obviously such an approach will not cover the field exhaustively, and in any case functional data analysis is not a hard-edged closed system of thought. Nevertheless we have tried to give a flavor of the range of methodology we ourselves have considered. We hope that our personal experience, including the fun we had working on these projects, will inspire others to extend “functional” thinking to many other statistical contexts. Of course, many of our case studies required development of existing methodology, and readers should gain the ability to adapt methods to their own problems too.

No previous knowledge of functional data analysis is needed to read this book, and although it complements our previous book in some ways, neither is a prerequisite for the other. We hope it will be of interest, and accessible, both to statisticians and to those working in other fields. Similarly, it should appeal both to established researchers and to students coming to the subject for the first time.

Functional data analysis is very much involved with computational statistics, but we have deliberately not written a computer manual or cookbook. Instead, there is an associated Web site accessible from [www.springer-ny.com](http://www.springer-ny.com) giving annotated analyses of many of the data sets, as well as some of the data themselves. The languages of these analyses are MATLAB, R, or S-PLUS, but the aim of the analyses is to explain the computational thinking rather than to provide a package, so they should be useful for those who use other languages too. We have, however, freely used a library of functions that we developed in these languages, and these may be downloaded from the Web site.

In both our books, we have deliberately set out to present a personal account of this rapidly developing field. Some specialists will, no doubt, notice omissions of the kind that are inevitable in this kind of presentation, or may disagree with us about the aspects to which we have given most emphasis. Nevertheless, we hope that they will find our treatment interesting and stimulating. One of our reasons for making the data, and the analyses, available on the Web site is our wish that others may do better. Indeed, may others write better books!

There are many people to whom we are deeply indebted. Particular acknowledgment is due to the distinguished paleopathologist Juliet Rogers, who died just before the completion of this book. Among much other research, Juliet's long-term collaboration with BWS gave rise to the studies in Chapters 4 and 8 on the shapes of the bones of arthritis sufferers of many centuries ago. Michael Newton not only helped intellectually, but also gave us some real data by allowing his juggling to be recorded for analysis in Chapter 12. Others whom we particularly wish to thank include Darrell Bock, Virginia Douglas, Zmira Elbaz-King, Theo Gasser, Vince Gracco, Paul Gribble, Michael Hermanussen, John Kimmel, Craig Leth-Steenson, Xiaochun Li, Nicole Malfait, David Ostry, Tim Ramsay, James Ramsey, Natasha Rossi, Lee Shepstone, Matthew Silverman, and Xiaohui Wang. Each of them made a contribution essential to some aspect of the work we report, and we apologize to others we have neglected to mention by name. We are very grateful to the Stanford Center for Advanced Study in the Behavioral Sciences, the American College Testing Program, and to the McGill students in the Psychology 747A seminar on functional data analysis. We also thank all those who provided comments on our software and pointed out problems.

Montreal, Quebec, Canada  
Bristol, United Kingdom  
January 2002

Jim Ramsay  
Bernard Silverman

# Contents

<b>Preface</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Why consider functional data at all? . . . . .	1
1.2 The Web site . . . . .	2
1.3 The case studies . . . . .	2
1.4 How is functional data analysis distinctive? . . . . .	14
1.5 Conclusion and bibliography . . . . .	15
<b>2 Life Course Data in Criminology</b>	<b>17</b>
2.1 Criminology life course studies . . . . .	17
2.1.1 Background . . . . .	17
2.1.2 The life course data . . . . .	18
2.2 First steps in a functional approach . . . . .	19
2.2.1 Turning discrete values into a functional datum . . . . .	19
2.2.2 Estimating the mean . . . . .	21
2.3 Functional principal component analyses . . . . .	23
2.3.1 The basic methodology . . . . .	23
2.3.2 Smoothing the PCA . . . . .	26
2.3.3 Smoothed PCA of the criminology data . . . . .	26
2.3.4 Detailed examination of the scores . . . . .	28
2.4 What have we seen? . . . . .	31

2.5	How are functions stored and processed? . . . . .	33
2.5.1	Basis expansions . . . . .	33
2.5.2	Fitting basis coefficients to the observed data . . . . .	35
2.5.3	Smoothing the sample mean function . . . . .	36
2.5.4	Calculations for smoothed functional PCA . . . . .	37
2.6	Cross-validation for estimating the mean . . . . .	38
2.7	Notes and bibliography . . . . .	40
<b>3</b>	<b>The Nondurable Goods Index</b>	<b>41</b>
3.1	Introduction . . . . .	41
3.2	Transformation and smoothing . . . . .	43
3.3	Phase-plane plots . . . . .	44
3.4	The nondurable goods cycles . . . . .	47
3.5	What have we seen? . . . . .	54
3.6	Smoothing data for phase-plane plots . . . . .	55
3.6.1	Fourth derivative roughness penalties . . . . .	55
3.6.2	Choosing the smoothing parameter . . . . .	55
<b>4</b>	<b>Bone Shapes from a Paleopathology Study</b>	<b>57</b>
4.1	Archaeology and arthritis . . . . .	57
4.2	Data capture . . . . .	58
4.3	How are the shapes parameterized? . . . . .	59
4.4	A functional principal components analysis . . . . .	61
4.4.1	Procrustes rotation and PCA calculation . . . . .	61
4.4.2	Visualizing the components of shape variability . . . . .	61
4.5	Varimax rotation of the principal components . . . . .	63
4.6	Bone shapes and arthritis: Clinical relationship? . . . . .	65
4.7	What have we seen? . . . . .	66
4.8	Notes and bibliography . . . . .	66
<b>5</b>	<b>Modeling Reaction-Time Distributions</b>	<b>69</b>
5.1	Introduction . . . . .	69
5.2	Nonparametric modeling of density functions . . . . .	71
5.3	Estimating density and individual differences . . . . .	73
5.4	Exploring variation across subjects with PCA . . . . .	76
5.5	What have we seen? . . . . .	79
5.6	Technical details . . . . .	80
<b>6</b>	<b>Zooming in on Human Growth</b>	<b>83</b>
6.1	Introduction . . . . .	83
6.2	Height measurements at three scales . . . . .	84
6.3	Velocity and acceleration . . . . .	86
6.4	An equation for growth . . . . .	89
6.5	Timing or phase variation in growth . . . . .	91
6.6	Amplitude and phase variation in growth . . . . .	93

6.7	What we have seen? . . . . .	96
6.8	Notes and further issues . . . . .	97
6.8.1	Bibliography . . . . .	97
6.8.2	The growth data . . . . .	98
6.8.3	Estimating a smooth monotone curve to fit data . . . . .	98
<b>7</b>	<b>Time Warping Handwriting and Weather Records</b>	<b>101</b>
7.1	Introduction . . . . .	101
7.2	Formulating the registration problem . . . . .	102
7.3	Registering the printing data . . . . .	104
7.4	Registering the weather data . . . . .	105
7.5	What have we seen? . . . . .	110
7.6	Notes and references . . . . .	110
7.6.1	Continuous registration . . . . .	110
7.6.2	Estimation of the warping function . . . . .	113
<b>8</b>	<b>How Do Bone Shapes Indicate Arthritis?</b>	<b>115</b>
8.1	Introduction . . . . .	115
8.2	Analyzing shapes without landmarks . . . . .	116
8.3	Investigating shape variation . . . . .	120
8.3.1	Looking at means alone . . . . .	120
8.3.2	Principal components analysis . . . . .	120
8.4	The shape of arthritic bones . . . . .	123
8.4.1	Linear discriminant analysis . . . . .	123
8.4.2	Regularizing the discriminant analysis . . . . .	125
8.4.3	Why not just look at the group means? . . . . .	127
8.5	What have we seen? . . . . .	128
8.6	Notes and further issues . . . . .	128
8.6.1	Bibliography . . . . .	128
8.6.2	Why is regularization necessary? . . . . .	129
8.6.3	Cross-validation in classification problems . . . . .	130
<b>9</b>	<b>Functional Models for Test Items</b>	<b>131</b>
9.1	Introduction . . . . .	131
9.2	The ability space curve . . . . .	132
9.3	Estimating item response functions . . . . .	135
9.4	PCA of log odds-ratio functions . . . . .	136
9.5	Do women and men perform differently on this test? . . . . .	138
9.6	A nonlatent trait: Arc length . . . . .	140
9.7	What have we seen? . . . . .	143
9.8	Notes and bibliography . . . . .	143
<b>10</b>	<b>Predicting Lip Acceleration from Electromyography</b>	<b>145</b>
10.1	The neural control of speech . . . . .	145
10.2	The lip and EMG curves . . . . .	147

10.3	The linear model for the data . . . . .	148
10.4	The estimated regression function . . . . .	150
10.5	How far back should the historical model go? . . . . .	152
10.6	What have we seen? . . . . .	155
10.7	Notes and bibliography . . . . .	155
<b>11</b>	<b>The Dynamics of Handwriting Printed Characters</b>	<b>157</b>
11.1	Recording handwriting in real time . . . . .	157
11.2	An introduction to dynamic models . . . . .	158
11.3	One subject's printing data . . . . .	160
11.4	A differential equation for handwriting . . . . .	162
11.5	Assessing the fit of the equation . . . . .	165
11.6	Classifying writers by using their dynamic equations . . . . .	166
11.7	What have we seen? . . . . .	170
<b>12</b>	<b>A Differential Equation for Juggling</b>	<b>171</b>
12.1	Introduction . . . . .	171
12.2	The data and preliminary analyses . . . . .	172
12.3	Features in the average cycle . . . . .	173
12.4	The linear differential equation . . . . .	176
12.5	What have we seen? . . . . .	180
12.6	Notes and references . . . . .	181
	<b>References</b>	<b>183</b>
	<b>Index</b>	<b>187</b>