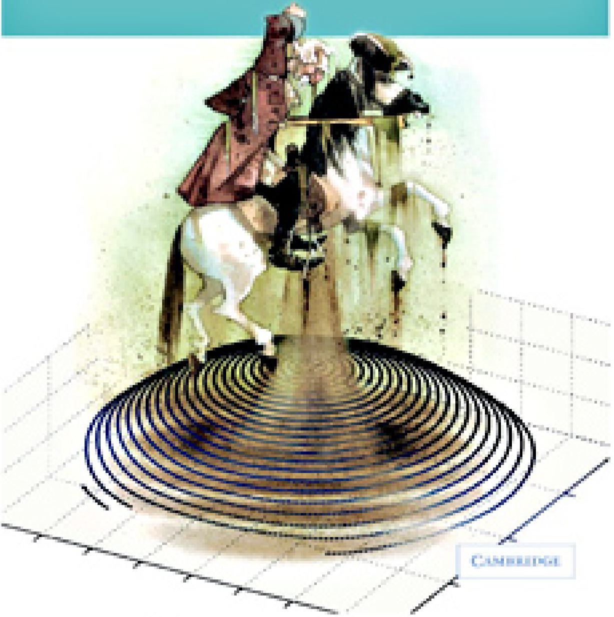
Bootstrap Techniques for Signal Processing

Abdelhak Zoubir and D. Robert Iskander



CAMBRIDGE www.cambridge.org/9780521831277

This page intentionally left blank

BOOTSTRAP TECHNIQUES FOR SIGNAL PROCESSING

A signal processing practitioner often asks themselves, "How accurate is my parameter estimator?" There may be no answer to this question if an analytic analysis is too cumbersome and the measurements sample is too small. The statistical bootstrap, an elegant solution, re-uses the original data with a computer to re-estimate the parameters and infer their accuracy.

This book covers the foundations of the bootstrap, its properties, its strengths, and its limitations. The authors focus on bootstrap signal detection in Gaussian and non-Gaussian interference as well as bootstrap model selection. The theory presented in the book is supported by a number of useful practical examples written in MATLAB.

The book is aimed at graduate students and engineers, and includes applications to real-world problems in areas such as radar, sonar, biomedical engineering and automotive engineering.

Abdelhak Zoubir is Professor of Signal Processing at Darmstadt University of Technology, Germany. He held positions in industry and in academia in Germany and Australia. He has published over 180 technical papers in the field of statistical methods for signal processing. He has maintained his research interest in the bootstrap since the late 1980s. He also regularly gives courses and tutorials on the bootstrap and its application for engineers.

D. Robert Iskander received a Ph.D. degree in signal processing from Queensland University of Technology (QUT), Australia, and holds the position of a principal research fellow in the Centre for Health Research, QUT. He has published over 60 technical papers in the field of statistical signal processing and its application to optometry. He also has several patents in the area of visual optics.

BOOTSTRAP TECHNIQUES FOR SIGNAL PROCESSING

ABDELHAK M. ZOUBIR
D. ROBERT ISKANDER

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521831277

First published in print format 2004

ISBN-13 978-0-511-19247-0 eBook (Adobe Reader)
ISBN-10 0-511-19247-9 eBook (Adobe Reader)

ISBN-13 978-0-521-83127-7 hardback
ISBN-10 0-521-83127-X hardback

Contents

Preface

The bootstrap genesis is generally attributed to Bradley Efron. In 1977 he wrote the famous Rietz Lecture on the estimation of sampling distributions based on observed data (Efron, 1979a). Since then, a number of outstanding and nowadays considered classical statistical texts have been written on the topic (Efron, 1982; Hall, 1992; Efron and Tibshirani, 1993; Shao and Tu, 1995), complemented by other interesting monographic exposés (LePage and Billard, 1992; Mammen, 1992; Davison and Hinkley, 1997; Manly, 1997; Barbe and Bertail, 1995; Chernick, 1999).

Efron and Tibshirani (1993) state in the Preface of their book *Our goal in this book is to arm scientists and* **engineers***, as well as statisticians, with computational techniques that they can use to analyze and understand complicated data sets.* We share the view that Efron and Tibshirani (1993) have written an outstanding book which, unlike other texts on the bootstrap, is more accessible to an engineer. Many colleagues and graduate students of ours prefer to use this text as the major source of knowledge on the bootstrap. We believe, however, that the readership of (Efron and Tibshirani, 1993) is more likely to be researchers and (post-)graduate students in mathematical statistics than engineers.

To the best of our knowledge there are currently no books or monographs on the bootstrap written for electrical engineers, particularly for signal processing practitioners. Therefore the decision for us to fill such a gap. The bootstrap world is a great one and we feel strongly for its discovery by engineers. Our aim is to stimulate interest by engineers to discover the power of bootstrap methods. We chose the title *Bootstrap Techniques for Signal Processing* not only because we work in this discipline and because many of the applications in this book stem from signal processing problems, but also owing to the fact that signal processing researchers and (post-)graduate students are the more likely engineers to use the book. In particular, we would like to reach researchers and students in statistical signal processing such as those working on problems in areas that include radar, sonar, telecommunications and biomedical engineering.

We have made every attempt to convey the "how" and "when" to use the bootstrap rather than mathematical details and proofs. The theory of the bootstrap is well established and the texts mentioned above can give the necessary details if the reader so wishes. We have included at least one example for every introduced topic. Some of the examples are simple, such as finding a confidence interval for the mean. Others are more complicated like testing for zero the frequency response of a multiple-input single-output linear time-invariant system.

It was difficult to decide whether we should include MATLAB† codes for the examples provided. After some deliberation, and given the fact that many graduate students and researchers ask for MATLAB codes to reproduce published results, we decided to include them. We have also provided a MATLAB toolbox which comprises frequently used routines. These routines have been purposely written for the book to facilitate the implementation of the examples and applications. All the MATLAB routines can be found in the Appendices.

A few tutorial papers on the bootstrap for signal processing exist. The interested readers can refer to the work of Zoubir (1993); Zoubir and Boashash (1998), and Zoubir (1999).

We are grateful to our colleagues Hwa-Tung Ong, Ramon Brcich, and Christopher Brown for making very helpful comments and suggestions on the manuscript. Additional words of thanks are for Hwa-Tung Ong for his help in the development of the Bootstrap MATLAB Toolbox. We would like to thank our current and past colleagues and graduate students who contributed directly or indirectly to the completion of the book. In particular, we would like to thank Johann Böhme, Don Tufts, David Reid, Per Pelin, Branko Ristic, Jonathon Ralston, Mark Morelande, Said Aouada, Amar Abd El-Sallam and Luke Cirillo. The authors are grateful to all honours and PhD students and colleagues of the Communications and Signal Processing Group at Curtin University of Technology in Perth, Australia and special thanks are due to Tanya Vernon for her continued support to the group.

Many government agencies and industries supported our research on the bootstrap over the years. Thanks are due to Boualem Boashash at Queensland University of Technology and John Hullett and Zigmantas Budrikis at

† MATLAB is a registered trademark of The MathWorks, Inc.

Curtin University of Technology. The data from the JINDALEE over-the-horizon radar system was provided by the Defence Science and Technology Organisation (DSTO) in Edinburgh, South Australia. Special thanks are due to Stuart Anderson and Gordon Frazer for their generous support. The data from the GPR system was also collected at the DSTO in Edinburgh, South Australia. Words of gratitude are for Ian Chant and Canicious Abeynayake for their support with the landmine project. Zoubir acknowledges the financial support of the Australian Research Council. We also thank Michael Collins from the Centre for Eye Research at Queensland University of Technology for his support and encouragement during the preparation of this book. The encouragement of Phil Meyler from the Cambridge University Press is also gratefully acknowledged.

We thank our families, wives and children for their support, understanding and love. Without their patience this work could not be completed.

Last, we wish to refer the reader to a recent exposition by Bradley Efron (2002) on the role of bootstrap methods in modern statistics and wish the reader a “happy bootstrapping”.

Notations

This list gives in alphabetical order the symbols that are frequently used throughout this book. Special notation that is used less frequently will be defined as needed.

General notation

A	scalar
$\boldsymbol{A}$	column vector or matrix
$\boldsymbol{A}'$	transpose of a vector or a matrix $\boldsymbol{A}$
$\overline{\boldsymbol{A}}$	complex conjugate of a vector or a matrix
$\boldsymbol{A}^H$	Hermitian operation (transpose and complex conjugate) on a vector or a matrix
$\boldsymbol{A}^{-1}$	inverse of a matrix
$\\|\boldsymbol{A}\\|$	Euclidean vector norm
$\lvert A \rvert$	magnitude of A
$\lfloor A \rfloor$	largest integer $\leq A$
$\lceil A \rceil$	largest integer $\geq A$
$\widehat{\boldsymbol{A}}$	estimator or estimate of $\boldsymbol{A}$
j	imaginary unit, $j^2 = -1$
E	expectation operator
mod	modulo operator
Prob	probability
Prob_*	probability conditioned on observed data
tanh	hyperbolic tangent
var	variance operation

Latin symbols

$c_{XX}(t)$	covariance function of a stationary signal X_t
$C_{XX}(\omega)$	spectral density of a stationary signal X_t
F	distribution function
h	kernel width or bandwidth
$\boldsymbol{I}$	identity matrix
$I_{XX}(\omega)$	periodogram of an observed stationary signal X_t
k	discrete frequency parameter
$K(\cdot)$	kernel function
n	size of a random sample
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean μ and variance σ^2
$o(\cdot)$	order notation: "of smaller order than"
$O(\cdot)$	order notation: "of the same order as"
P_D	probability of detection
P_F	probability of false alarm
$\mathbb{R}$	the set of real numbers
t	discrete or continuous time index
t_n	t-distribution with n degrees of freedom
T_n	test statistic
$\mathcal{U}(a, b)$	uniform distribution over $[a, b]$.
X_t	random signal
$\mathcal{X}$	random sample
$\mathcal{X}^*$	bootstrap resample of $\mathcal{X}$
$\mathbb{Z}$	the set of integers

Greek symbols

α	level of significance
$\delta(\cdot)$	Kronecker's delta function
χ^2_n	chi-square distribution with n degrees of freedom
Γ	mean-square prediction error
θ	parameter
μ	mean
τ	time delay or lag
σ^2	variance
$\Phi(x)$	standard Gaussian distribution function
ω	radial frequency

Acronyms

AIC	Akaike information criterion
AR	autoregression
CDF	cumulative distribution function
CFAR	constant false alarm rate
CFD	central finite difference
FIR	finite impulse response
GPR	ground penetrating radar
GPS	global positioning system
HF	high frequency
IF	instantaneous frequency
iid	independent and identically distributed
MDL	minimum distance length (criterion)
MISO	multiple input single output
MLE	maximum likelihood estimator/estimation
ROC	receiver operating characteristic
SNR	signal-to-noise ratio
SRB	sequentially rejective Bonferroni (procedure)
UMP	uniformly most powerful (test)

1
Introduction

Signal processing has become a core discipline in engineering research and education. Many modern engineering problems rely on signal processing tools. This could be either for filtering the acquired measurements in order to extract and interpret information or for making a decision as to the presence or absence of a signal of interest. Generally speaking, statistical signal processing is the area of signal processing where mathematical statistics is used to solve signal processing problems. Nowadays, however, it is difficult to find an application of signal processing where tools from statistics are not used. A statistician would call the area of statistical signal processing time series analysis.

In most statistical signal processing applications where a certain parameter is of interest there is a need to provide a rigorous statistical performance analysis for parameter estimators. An example of this could be finding the accuracy of an estimator of the range of a flying aircraft in radar. These estimators are usually computed based on a finite number of measurements, also called a sample. Consider, for example, a typical radar scenario, in which we aim to ascertain whether the received signal contains information about a possible target or is merely interference. The decision in this case, based on calculating the so-called test statistic, has to be supported with statistical measures, namely the probability of detection and the probability of false alarm. Such a decision can be made if the distribution of the test statistic is known in both cases: when the received signal contains target information and when the target information is absent.

Another important problem in signal processing is to make certain probability statements with respect to a true but unknown parameter. For example, given some estimator of a parameter, we would like to determine upper and lower limits such that the true parameter lies within these limits with a preassigned probability. These limits constitute the so-called confidence interval (Cramér, 1999).

Two main questions arise in a parameter estimation problem. Given a number of measurements and a parameter of interest:

(i) What estimator should we use?

(ii) Having decided to use a particular estimator, how accurate is it?

A signal processing practitioner would first attempt to use the method of maximum likelihood or the method of least squares to answer the first question (Scharf, 1991; Kay, 1993). Having computed the parameter estimate, its accuracy could be measured by the variance of the estimator or a confidence interval for the parameter of interest. In most cases, however, techniques available for computing statistical characteristics of parameter estimators assume that the size of the available set of samples is sufficiently large, so that asymptotic results can be applied. Techniques that invoke the Central Limit Theorem and the assumption of Gaussianity of the noise process are examples of such an approach (Bhattacharya and Rao, 1976; Serfling, 1980).

Let us consider an example where we are interested in finding the 100α, $0 < \alpha < 1$, percent confidence interval for the power spectral density of a stationary real-valued signal, given a finite number of observations. If we were to assume that the number of observations is large, we would use an asymptotic approach so that the spectral density estimates at distinct frequencies could be considered independent with a limiting scaled χ^2 distribution.

Let us mathematically formalise this problem. Assume $X_1, \ldots, X_n$ to be a finite set of observations from a real-valued, strictly stationary signal X_t, $t \in \mathbb{Z}$, with mean zero and a finite variance. Define the spectral density of X_t by

$$C_{XX}(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \mathsf{E}\left[X_t X_{t-|\tau|}\right] e^{-j\omega\tau}, \qquad (1.1)$$

where $\mathsf{E}[\cdot]$ denotes expectation, and let

$$I_{XX}(\omega) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} X_t e^{-j\omega t} \right|^2, \qquad -\pi < \omega \leq \pi, \qquad (1.2)$$

denote the periodogram of the sample (Brillinger, 1981; Marple Jr, 1987).

Consider estimating the spectral density $C_{XX}(\omega)$ by a kernel spectral density estimate (the smoothed periodogram), given by

$$\hat{C}_{XX}(\omega;\, h) = \frac{1}{n\,h} \sum_{k=-M}^{M} K\left(\frac{\omega - \omega_k}{h}\right) I_{XX}(\omega_k), \qquad -\pi < \omega \leq \pi, \quad (1.3)$$

where the kernel $K(\cdot)$ is a symmetric, nonnegative function on the real line, h is its bandwidth, and M is the largest integer less than or equal to $n/2$. Let the discrete frequencies ω_k be given by

$$\omega_k = 2\,\pi\,k/n, \qquad -M \leq k \leq M\,.$$

A variety of kernels can be used in Equation (1.3) but let us choose the Bartlett-Priestley window (Priestley, 1981, p. 444) for $K(\cdot)$. Given the estimate of the power spectral density (1.3), one can approximate its distribution asymptotically, as $n \to \infty$ by

$$C_{XX}(\omega)\,\chi^2_{4\,m+2}/(4\,m+2),$$

where $m = \lfloor(h\,n - 1)/2\rfloor$, $\lfloor\cdot\rfloor$ denotes the floor operator and $\chi^2_{4\,m+2}$ is the χ^2 distribution with $4\,m+2$ degrees of freedom. This leads to the following 100α percent confidence interval (Brillinger, 1981)

$$\frac{(4\,m+2)\,\hat{C}_{XX}(\omega;\,h)}{\chi^2_{4\,m+2}\left(\dfrac{1+\alpha}{2}\right)} < C_{XX}(\omega) < \frac{(4\,m+2)\,\hat{C}_{XX}(\omega;\,h)}{\chi^2_{4\,m+2}\left(\dfrac{1-\alpha}{2}\right)}, \quad (1.4)$$

where $\chi^2_\nu(\alpha)$ denotes a number such that the probability

$$\mathsf{Prob}\left[\chi^2_\nu < \chi^2_\nu(\alpha)\right] = \alpha\,.$$

The analytical result in (1.4) is pleasing, but it assumes that n is sufficiently large so that $\hat{C}_{XX}(\omega_1), \ldots, \hat{C}_{XX}(\omega_M)$ are independent χ^2 random variables. In many signal processing problems, as will be seen throughout the book, large sample methods are inapplicable. This is either because of time constraints or because the signal of interest is non-stationary and stationarity can be assumed over a small portion of data only.

There are cases where small sample results, obtained analytically, do exist (Shenton and Bowman, 1977; Field and Ronchetti, 1990). However, more often it is the case that these results cannot be attained and one may have to resort to Monte Carlo simulations (Robert and Casella, 1999). We will recall the problem of estimating the 100α percent confidence interval for power spectral densities in Chapter 2, where a bootstrap-based solution is described.

The bootstrap was introduced by Bradley Efron (1979a,b, 1981, 1982) more than two decades ago, mainly to calculate confidence intervals for parameters in situations where standard methods were not applicable (Efron and Gong, 1983). An example of this would be a situation where the number of observations is so small that asymptotic results are unacceptably inaccurate. Since its invention, the bootstrap has seen many more applications and has been used to solve problems which would be too complicated to be solved analytically (Hall, 1992; Efron and Tibshirani, 1993). Before we continue, let us clarify several questions that we have been frequently asked in the past.

What is the bootstrap? Simply put, the bootstrap is a method which does with a computer what the experimenter would do in practice if it were possible: they would repeat the experiment. With the bootstrap, a new set of experiments is not needed, instead, the original data is reused. Specifically, the original observations are randomly reassigned and the estimate is recomputed. These assignments and recomputations are done a large number of times and considered as repeated experiments. One may think that the bootstrap is similar to Monte Carlo simulations. However, this is not the case. The main advantage of the bootstrap over Monte Carlo simulations is that the bootstrap does not require the experiment to be repeated.

From a data manipulation point of view, the main idea encapsulated by the bootstrap is to simulate as much of the "real world" probability mechanism as possible, substituting any unknowns with estimates from the observed data. Through the simulation in the "bootstrap world", unknown entities of interest in the "real world" can be estimated as indicated in Figure 1.1. The technical aspects of such simulation are covered in the next chapter.

Why is the bootstrap so attractive? The bootstrap does not make the assumption of a large number of observations of the signal. It can answer many questions with very little in the way of modelling, assumptions or analysis and can be applied easily. In an era of exponentially declining computational costs, computer-intensive methods such as the bootstrap are becoming a bargain. The conceptual simplicity of bootstrap methods can sometimes undermine the rich and difficult theory upon which they are based (Hall, 1992; Shao and Tu, 1995). In the next chapter, we will provide a review of the bootstrap theory in a manner more accessible to signal processing practitioners.

What can I use the bootstrap for? In general, the bootstrap is a methodology for answering the question we posed earlier, that is, *how accurate is a parameter estimator*? It is a fundamental question in many signal processing problems and we will see later how one, with the bootstrap, can

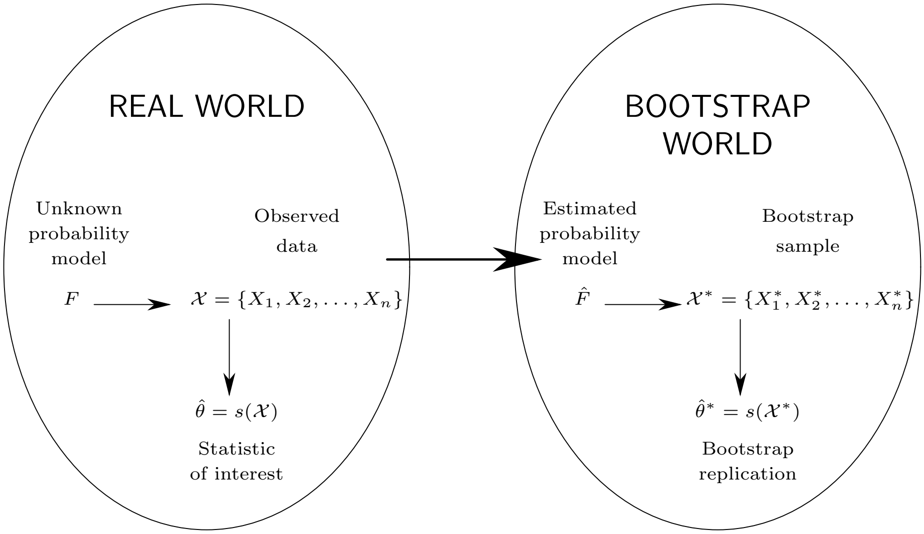


Fig. 1.1. The bootstrap approach, adapted from Efron and Tibshirani (1993, Fig. 8.3). See Chapter 2 for a technical interpretation of this figure.

solve many more problems encountered by a signal processing engineer today, for example, signal detection. This text will also provide an answer to the question regarding the choice of an estimator from among a family of estimators using the bootstrap. We briefly discuss this topic in Chapter 5, where we consider the optimisation of trimming for the trimmed mean in a radar application (see also the work of Léger *et al.* (1992), for example).

Is the bootstrap always applicable? Theoretical work on the bootstrap and applications have shown that bootstrap methods are potentially superior to large sample techniques. A danger, however, does exist. The signal processing practitioner may well be attracted to apply the bootstrap in an application to avoid the use of methods that invoke strong assumptions, such as asymptotic theory, because these are judged inappropriate. But in this case the bootstrap may also fail (Mammen, 1992; Young, 1994). Special care is therefore required when applying the bootstrap in real-life situations (Politis, 1998). The next chapter provides the fundamental concepts and methods needed by the signal processing practitioner to decide when and how to apply the bootstrap successfully.

Applications of bootstrap methods to real-life engineering problems have been reported in many areas, including radar and sonar signal processing, geophysics, biomedical engineering and imaging, pattern recognition and computer vision, image processing, control, atmospheric and environmental research, vibration analysis and artificial neural networks. In almost all these fields, bootstrap methods have been used to approximate the distribution of an estimator or some of its characteristics. Let us list in no particular order some of the bootstrap engineering applications that we found interesting.

Radar and sonar: The bootstrap has been applied to radar and sonar problems for more than a decade. Nagaoka and Amai (1990, 1991) discuss an application in which the distribution of the estimated "close approach probability" is derived to be used as an index of collision risk in air traffic control. Hewer *et al.* (1996) consider a wavelet-based constant false alarm rate (CFAR) detector in which the bootstrap is used to derive the statistics of the detector from lexicographically ordered image vectors. Anderson and Krolik (1998a,b, 1999) use the bootstrap in a hidden Markov model approximation to the ground range likelihood function in an over-the-horizon radar application.

Ong and Zoubir (1999a,b, 2000b, 2003) consider bootstrap applications in CFAR detection for signals in non-Gaussian and correlated interference, while Zoubir *et al.* (1999) apply bootstrap methods to the detection of landmines. Böhme and Maiwald (1994) apply bootstrap procedures to signal detection and location using sensor arrays in passive sonar and to the analysis of seismic data.

Krolik *et al.* (1991) use bootstrap methods for evaluating the performance of source localisation techniques on real sensor array data without precise *a priori* knowledge of true source positions and the underlying data distribution (see also (Krolik, 1994)). Reid *et al.* (1996) employ bootstrap based techniques to determine confidence bounds for aircraft parameters given only a single acoustic realisation, while Bello (1998) uses the bootstrap to calculate cumulative receiver operating characteristic (ROC) curve confidence bounds for sets of side-scan sonar data.

Geophysics: A similar interest in bootstrap methods has taken place in geophysics. Fisher and Hall (1989, 1990, 1991) apply the bootstrap to the problem of deciding whether or not palaeomagnetic specimens sampled from a folded rock surface were magnetised before or after folding occurred. They conclude that the bootstrap method provides the only feasible approach in this common palaeomagnetic problem. Another application of bootstrap methods in palaeomagnetism has been reported by Tauxe *et al.* (1991). Kawano and Higuchi (1995) estimate with the bootstrap the average component in the minimum variance direction in space physics.

Ulrych and Sacchi (1995) propose an extended information criterion based on the bootstrap for the estimation of the number of harmonics actually present in geophysical data. Later, Sacchi (1998) uses the bootstrap for high-resolution velocity analysis.

Lanz *et al.* (1998) perform quantitative error analyses using a bootstrap technique while determining the depth and geometry of a landfill's lower boundary. Mudelsee (2000) uses bootstrap resampling in ramp function regression for quantifying climate transitions, while Rao (2000) uses the bootstrap to assess and improve atmospheric prediction models.

Biomedical engineering: Biomedical signal and image processing has been another area of extensive bootstrap applications. Haynor and Woods (1989) use the bootstrap for estimating the regional variance in emission tomography images. Banga and Ghorbel (1993) introduce a bootstrap sampling scheme to remove the dependence effect of pixels in images of the human retina. Coakley (1996) computes bootstrap expectation in the reconstruction of positron emission tomography images. Locascio *et al.* (1997) use the bootstrap to adjust p-values in multiple significance tests across pixels in magnetic resonance imaging. Maitra (1998) applies the bootstrap in estimating the variance in parametric biomedical images. Verotta (1998) investigates the use of the bootstrap to obtain the desired estimates of variability of system kernel and input estimates, while Bullmore *et al.* (2001) use bootstrap based techniques in the time and wavelet domains in neurophysiological time series analysis. Another interesting application of the bootstrap is reported by Iskander *et al.* (2001) where it is used to find the optimal parametric model for the human cornea. Recently, Chen *et al.* (2002) have employed the bootstrap for aiding the diagnosis of breast cancer in ultrasound images.

Image processing: Bootstrap methods have been widely applied in image processing, pattern recognition and computer vision. Jain *et al.* (1987) apply several bootstrap techniques to estimate the error rate of nearest-neighbour and quadratic classifiers. Hall (1989b) calculates confidence regions for hands in degraded images. Archer and Chan (1996) apply the bootstrap to problems in blind image restoration where they calculate confidence intervals for the true image. Cabrera and Meer (1996) use the bootstrap to eliminate the bias of estimators of ellipses, while Saradhi and Murty (2001) employ the bootstrap technique to achieve higher classification accuracy in handwritten digit recognition.

Control: Bootstrap methods have found applications in statistical control. Dejian and Guanrong (1995) apply bootstrap techniques for estimating the distribution of the Lyapunov exponent of an unknown dynamic system using its time series data. Seppala *et al.* (1995) extend the bootstrap percentile method to include a series of subgroups, which are typically used in assessing process control limits. They show that the method achieves comparatively better control limit estimates than standard parametric methods. Ming and Dong (1997) utilise the bootstrap to construct a prediction interval for future observations from a Birnbaum–Saunders distribution that is used as a failure time model. Jones and Woodall (1998) use the bootstrap in control chart procedures, while Aronsson *et al.* (1999), apply bootstrap techniques to control linear stochastic systems. They derive the optimal future control signal so that the unknown noise distribution and uncertainties in parameter estimates are taken into account. Recently, Tjärnström and Ljung (2002) have used the bootstrap to estimate the variance of an undermodelled structure that is not flexible enough to describe the underlying control system without the need for Monte Carlo simulations.

Environmental engineering: Bootstrap techniques have found several applications in atmospheric and environmental research. Hanna (1989) uses the related jackknife procedure and the bootstrap for estimating confidence limits for air quality models. The resampling procedures have been applied to predictions by a number of air quality models for the Carpentaria coastal dispersion experiment. Downton and Katz (1993) use the bootstrap to compute confidence intervals for the discontinuity in variance of temperature time series, while Krzyscin (1997) infers about the confidence limits of the trend slope and serial correlation coefficient estimates for temperature using the bootstrap.

Artificial neural networks: Bootstrap techniques have also been applied in the area of artificial neural networks. Bhide *et al.* (1995) demonstrate the use of bootstrap methods to estimate a distillation process bottoms' composition. Tibshirani (1996) discusses a number of methods for estimating the standard error of predicted values from a multilayered perceptron. He finds that bootstrap methods perform best, partly because they capture variability due to the choice of starting weights. LeBaron and Weigend (1998) use the bootstrap to compare the uncertainty in the solution stemming from the data splitting with neural-network specific uncertainties in application to financial time series. Franke and Neumann (2000) investigate the bootstrap methods in the context of artificial neural networks used for estimating a