Bootstrap Methods for a Measurement Estimation Problem

José I. De la Rosa, *Member, IEEE*, and Gilles A. Fleury

***Abstract*—In this paper, a new approach for the statistical characterization of a measurand is presented. A description of how different bootstrap techniques can be applied in practice to estimate successfully a measurand probability density function (pdf) is given. When the direct observation of a quantity of interest is practically impossible such as in nondestructive testing, it is necessary to estimate such quantity, which is also called measurand. The statistical characterization of any estimator is important, because all the uncertainty features can be accessible to qualify such estimator. On the other hand, most of the time, the large-scale repetition of an experiment is not economically feasible, so that the Monte Carlo methods cannot be used directly for uncertainty characterization. Bootstrap methods have proved to be a potentially useful alternative. Moreover, a biased bootstrap recent technique, with which robust parameter estimates are obtained, is used. This technique is extended to be used in measurand estimation. An extended nested bootstrap improvement for the measurand pdf estimation is also presented. These techniques are applied to a realistic multidimensional measurand-estimation problem of groove dimensioning using remote field eddy current inspection. Measurand uncertainty characterization using the bootstrap techniques generally gives an accurate pdf estimation.**

***Index Terms*—Bootstrap, indirect measurement, Monte Carlo simulation, nonlinear regression, probability density function (pdf) estimation.**

NOTATIONS AND ACRONYMS

a, $\mathbf{a}$, $\mathbf{A}$	Scalar, vector, and matrix.
$\mathbf{a}^\top$, $\mathbf{A}^\top$	Transposed vector and matrix.
y_i, $\mathbf{y}$	Vector of observations (data).
x_i, $\boldsymbol{x}$	Experimental protocol or instrumental parameters.
$f(\boldsymbol{x}, \boldsymbol{\theta})$, $f(\cdot)$	Model function parameterized by $\boldsymbol{x}$ and $\boldsymbol{\theta}$.
$\boldsymbol{\theta}$	Unknown parameter vector.
$g(\boldsymbol{\theta})$, $g(\cdot)$	Measurand and parameter functional relationship.
$G_k(\boldsymbol{\theta})$, $G_k(\cdot)$	Bijective relation between measurand and parameters.
$\wp(\boldsymbol{\theta})$	Parameter distribution.
$\wp(m)$, $\wp(\boldsymbol{m})$	Measurand distribution.
$\wp(\boldsymbol{e})$	Errors or noise distribution.
$\widehat{\boldsymbol{\theta}}$, $\bar{\boldsymbol{\theta}}$	Parameter estimator and real parameters.
p	Dimensions of vector $\boldsymbol{\theta}$.
n	Dimensions of vectors $\boldsymbol{y}$, $\boldsymbol{x}$, and $\boldsymbol{e}$.
$\widehat{e}_i$, $\widehat{\boldsymbol{e}}$	Residuals vector.
$\widehat{\wp}_e$, $\widehat{\wp}(\boldsymbol{e})$	Residuals empiric distribution.
e_i^*, $\boldsymbol{e}^*$	Bootstrap residuals vector.
y_i^*, $\mathbf{y}^*$	Bootstrap fictive data.
$\widehat{\boldsymbol{\theta}}^*$	Bootstrap parameter estimators.
$\widehat{m}^*$, $\widehat{\boldsymbol{m}}^*$	Bootstrap measurand estimator.
$\wp(\widehat{m}^*)$, $\wp(\widehat{\boldsymbol{m}}^*)$	Bootstrap measurand distribution.
$\wp(\widehat{\boldsymbol{\theta}}^*)$	Bootstrap parameter distribution.
h_i, w_i	Vectors of weights used by the external and the biased bootstraps.
$\xi(\varepsilon)$	Breakdown function.
$\gamma(w_i)$	Data dispersion measure.
$\Delta_H(\cdot)$	Hellinger's distance.
B	Number of simulations (iterations) for the bootstrap.
B_1, B_2	Number of simulations for the internal and the external bootstraps.
BBQ	Biased bootstrap with quadratic norm function.
BBH	Biased bootstrap with a Huber-like norm function.
BOOT	Bootstrap procedure.
NLS	Nonlinear least squares estimation.
pdf	Probability density function.
PMC	Primitive Monte Carlo.
RFEC	Remote field eddy current.

I. INTRODUCTION

IN MANY industrial applications, direct access to a measurand ($\boldsymbol{m}$) is not possible. Yet, as an estimation of the measurand is needed, the process must be treated as an inverse problem [1]. The characterization of the whole statistical knowledge about this quantity of interest $\boldsymbol{m}$ (quantity to be measured) is naturally given by the pdf $\wp(\boldsymbol{m})$. Most of the time, the large-scale repetition of an experiment is not economically feasible. Therefore, Monte Carlo methods cannot be used. Bootstrap methods have proven to be a potentially useful alternative in accessing $\wp(\boldsymbol{m})$. Moreover, it is well known in practice that every observation in a data set does not play the same role in determining estimates, tests, or other statistics. This is due, for example, to a certain failure in the data acquisition or transmission processes. Such data are called outliers. When outliers are present in the data, a robust nonlinear regression strategy called biased bootstrap [2], [3] could be used for parameter and measurement estimation. The aim of this paper is then to appraise the measurand pdf using bootstrap techniques

Manuscript received May 3, 2005; revised October 7, 2005.
J. I. De la Rosa is with the Signal Processing Laboratory, Engineering Faculty, Universidad Autónoma de Zacatecas, Zacatecas, Mexico (e-mail: ismaelrv@ieee.org).

G. A. Fleury is with the École Supérieure d'Électricité—Department of Signal Processing and Electronic Systems, Plateau de Moulon, Gif-sur-Yvette Cedex 91192, France (e-mail: Gilles.Fleury@supelec.fr).
Digital Object Identifier 10.1109/TIM.2006.873779