# A Statistical Inference Comparison for Measurement Estimation Using Stochastic Simulation Techniques

José Ismael De la Rosa, *Member, IEEE*, Gerardo Miramontes, *Member, IEEE*, Lyle E. McBride, José de Jesús Villa, Gilles A. Fleury, and Marie-Eve Davoust

***Abstract*—The purpose of this paper is to present a comparison of different techniques for making statistical inference about a measurement system model. This comparison involves results when two main assumptions are made: 1) the unknowable behavior of the probability density function (pdf) $\wp(e)$ of errors since the real measurement systems are always exposed to continuous perturbations of an unknown nature and 2) the assumption that, after some experimentation, one can obtain sufficient information that can be incorporated into the modeling as prior information. The first assumption leads us to propose the use of two approaches, which permit building hybrid algorithms; such approaches are the nonparametric bootstrap and the kernel methods. The second assumption makes possible the exploration of a Bayesian framework solution and Monte Carlo Markov chain auxiliary that is used to access the *a posteriori* pdf of the measurement. For both assumptions over $\wp(e)$ and the model, different classical criteria can be used; one also uses an extension of a recent criterion of entropy minimization. The entropy criterion is constructed on the basis of a symmetrized kernel estimate $\widehat{\wp}_{n,h}(e)$ of $\wp(e)$. Finally, a comparison between results obtained with the different schemes proposed by De la Rosa is presented.**

***Index Terms*—Bootstrap, indirect measurement, Monte Carlo Markov chain (MCMC), nonlinear regression, nonparametric probability density function (pdf) estimation, robust estimation.**

## I. Introduction

The uncertainty characterization of a model is a whole complex problem that depends, among other characteristics, on the degrees of freedom of such a model, its behavior, and its structure. The case of nonlinear modeling is a special case treated by the Department of Measurement (recently renamed Department of Signal Processing and Electronics), École Supérieure d'Électricité (SUPÉLEC) [18], to deal with some problems in an indirect measurement framework (e.g., instrumentation). Tackling inverse problems according to a statistical point of view permits the proposed methods to take advantage of the diversity and different characteristics of the collected data and all information about the collection procedure. According to previous conditions, one will have more or less information about the whole measurement system, and then, different considerations can be taken according to the model structure.

*Minimum of information:* In cases where limited information is available, the most natural proposition is to take a small number of hypotheses and take advantage of all the information contained in the data itself using methods such as the bootstrap (parametric or nonparametric) [15]–[17], [19]–[22], [38], [39], the kernel methods or nonparametric estimation [6]–[9], and all the related methods [1], [11]–[14], [27], [28], [36] (see also some of the work of Parzen–Rosenblatt dating back to 1968).

*Well-known information:* In other cases, one may dispose of sufficient but well-known information, and if the complexity of the modeling problem can be bounded, the modeling problem can be summarized using classical parametric methods, which could lead to analytical models [2]–[4], even if, for practical problems, the models are more complex than they seem.

*Maximum of information:* This last case describes ideal conditions that any researcher would like to reach, even if these conditions can be attained only after some experience with the modeling considerations and with the treated data. In recent years, this way of tackling problems points toward the use of Bayesian methods or even ideas about data fusion. When modeling deals with complex problems but there is a maximum of information, and the problem is studied in a Bayesian framework, then Monte Carlo procedures known as Monte Carlo Markov chains (MCMCs) [32] provide a set of tools to obtain practical solutions of the proposed models [9], [10].

The purpose of using Monte Carlo simulations such as bootstrap or MCMC is to obtain the empirical probability density function (pdf) [31], which, in general, will lead to establishing different statistics about the treated data and, more specifically, about the measurements of interest. The measurement modeling has been considered in various works [2]–[10]. The principal contribution of this paper is presented in the following, where a comparison of the different schemes for estimation and uncertainty characterization proposed here is presented.

Measurement systems are formalized by two equations [4], [5], [10]:

1) *Observation equation*, which is described by the classical nonlinear regression model

$$y = f(\boldsymbol{x}, \boldsymbol{\theta}) + \boldsymbol{e} \tag{1}$$

Manuscript received September 19, 2007; revised March 7, 2008. This work was supported in part by the Mexican Program for Professor's Technical Improvement (PROMEP) under Grant UAZAC PTC 24-103.5/03/1127.

J. I. De la Rosa, G. Miramontes, and J. de Jesús Villa are with the Signal Processing Laboratory, Faculty of Electrical Engineering, Universidad Autónoma de Zacatecas, Zacatecas 98600, México (e-mail: ismaelrv@ieee.org).

L. E. McBride is with California State University, Chico, CA 95929 USA.

G. A. Fleury and M.-E. Davoust are with the Department of Signal Processing and Electronic Systems, École Supérieure d'Électricité, 91192 Gif sur Yvette, France.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2008.922098