IIT, Madras

PREFACE

Preface for special issue on "Data Analysis: Techniques and Applications"

**Raghunathan Rengaswamy ·
Lakshminarayanan Samavedham**

Modern process plants are equipped with data acquisition systems that are capable of collecting operating data very frequently. This is a result of the availability of inexpensive storage and computational resources, further translating to unprecedented data archival and warehousing. Process understanding and insights have to be culled from this data deluge. Several techniques are being developed to address this problem of data interpretation and analysis. While ideas from multivariate statistical methods underlie many of these techniques, other concepts such as extraction of qualitative features are also being pursued. The spectrum of applications for these techniques are varied and include speech and image processing, biomedical signal processing, process fault diagnosis, bioinformatics, envirometrics and chemometrics. In this special issue we have assembled a collection of papers that deal with data analysis techniques that are generic in nature and relevant in a variety of applications. Some of these papers address novel applications of data analysis techniques in experimental systems.

The first paper in this issue by Narasimhan and Narasimhan discusses data reconciliation of linear systems when the underlying model is uncertain. While data reconciliation is a widely studied area, this aspect of data reconciliation seems to have received much less attention. The paper by Kuppuraj and Rengaswamy addresses the multiple model identification problem. In this work, the

data that is being processed is assumed to have been generated by multiple models operating in non-intersecting partitions of the input space. Both static and dynamic cases are addressed. The paper by Chee and Srinivasan advocates the development of an artificial immune system (AIS) inspired framework for adaptive fault diagnosis and recovery once a fault has occurred. A simulated distillation column example is used to highlight the advantages of the AIS solution approach. The paper by Babu and Narasimhan proposes the use of iterated principal component analysis (IPCA) as a consistent pre-processing technique for independent component analysis. This pre-processing is invariant to data scaling and also handles heteroscedastic errors. A clustering method that identifies the maximal number of distinct clusters from candidate patterns is proposed by Jonnalagada and Srinivasan. The efficacy of this approach is tested using gene expression data sets.

The paper by Sumana et al. compares two approaches for fault diagnosis in nonlinear systems. These are the Kernel PCA and nonlinear transformation of data followed by a correspondence analysis. The Tennessee-Eastman benchmark problem is used to compare the two approaches. The paper by Villez et al. discusses the application of a technique that is based on extraction of qualitative trends from process data for fault detection and identification of a phosphate analyzer that is used for water quality monitoring. In the paper by Karthik and Lakshminarayanan, a multi-objective optimization based approach is used to estimate the length and width of crystal structures from images. This is an important problem in monitoring and control of crystallization processes. The paper by Razak et al. is on the development of representative models for self powered neutron detectors using PCA based techniques. In particular, the standard PCA technique and the IPCA technique are compared using data from an experimental system.

R. Rengaswamy (✉)
Department of Chemical Engineering, Texas Tech University,
Lubbock, TX, USA
e-mail: raghu.rengasamy@ttu.edu

L. Samavedham
Department of Chemical Engineering, National University
of Singapore, Singapore, Singapore
e-mail: laksh@nus.edu.sg

As can be seen, the applications of data analysis tools could be in several different fields ranging from crystallization to nuclear reactors. We hope that through this collection of papers, we have been able to provide the readers a glimpse of the several exciting application areas and research issues that will continue to drive this field forward.