Lecture Notes in Statistics 217

Edited by P. Bickel, P. Diggle, S.E. Fienberg, U. Gather,
I. Olkin, S. Zeger

More information about this series at
http://www.springer.com/series/694

Anestis Antoniadis • Jean-Michel Poggi •
Xavier Brossat
Editors

Modeling and Stochastic Learning for Forecasting in High Dimensions

Springer

Editors
Anestis Antoniadis
Department of Statistics
University Joseph Fourier
Grenoble, France

Jean-Michel Poggi
Laboratoire de Mathématiques
Université Paris-Sud
Orsay Cedex, France

Xavier Brossat
Electricité de France R & D, OSIRIS
Clamart Cedex, France

ISSN 0930-0325 ISSN 2197-7186 (electronic)
Lecture Notes in Statistics
ISBN 978-3-319-18731-0 ISBN 978-3-319-18732-7 (eBook)
DOI 10.1007/978-3-319-18732-7

Library of Congress Control Number: 2015941890

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media
(www.springer.com)

Preface

Forecasting and time series prediction have seen a great deal of development and attention over the last few decades, fostered by an impressive improvement in observational capabilities and measurement procedures. Time series prediction is a challenge in many fields. In finance, one forecasts stock exchange or stock market indices; data processing specialists forecast the flow of information on their networks; producers of electricity forecast the electric load, and hydrologists forecast river floods. Many methods designed for time series prediction and forecasting perform well (depending on the complexity of the problem) on a rather short-term horizon but are rather poor on a longer-term one. This is due to the fact that these methods are usually designed to optimize the performance at short-term prediction using relatively low dimensional models. For long-term prediction, linear and nonlinear methods and tools for the analysis and predictive modeling of high-dimensional phenomena are necessary and more useful.

From an industrial point of view, forecasting the electricity/natural gas consumption of consumers is yet another challenge. In order to be able to provide accurate forecasts on different horizons (from short term, the next hour, to middle term, several years) and at different levels of aggregation (from complete portfolio to local demand), these forecasts have to be done in an evolving environment. Indeed, national demand is increasing, uses are changing, the number of available data is significantly growing (thanks to smart meters), the individual renewable electricity generation is becoming important, and smart grids are developing. On the other side of electricity markets, renewables provide a bigger and bigger part of electricity generation. Forecasting wind or solar power is also difficult because of their inner variability, the need for accurate and local meteorological forecasts at any horizon – from a few hours up to several days ahead – as well as the rapid evolution of operating installations.

On June 5–7, 2013, an international workshop on Industry Practices for Forecasting was held in Paris, France, organized and supported by the OSIRIS Department of EDF R&D, located in Clamart, France. OSIRIS stands for Optimization Simulation Risks and Statistics for energy markets. The meeting was the second in the series of the WIPFOR conferences and was attended by several researchers (academics, industrial and professionals, and other interested parties) from several countries. Following tradition, both theoretical statistical results and practical contributions of this active field of statistical research and on forecasting issues in a fast evolving industrial environment were presented. The program and abstracts are available on the conference website (http://conferences-osiris.org/wipfor/13-Main-page).

The editors of this volume hope that these lecture notes reflect the broad spectrum of the conference, as it includes 16 articles contributed by specialists in various areas in this field. The material compiled is fairly wide in scope and ranges from the development of results on forecasting in industry and in time series, on nonparametric and functional methods, on online machine learning for forecasting, and on tools for high-dimensional and complex data analysis.

The articles are arranged and numbered in alphabetical order by author rather than subject matter. Papers 1, 3, 5, and 9 are dedicated to nonparametric techniques for short-term load forecasting in the industry and include classical curve linear regression, sparse functional regression based on dictionaries, as well as a new estimation procedure based on iterative bias reduction. Papers 11 and 14 focus on electrical system changes: the first is dedicated to large-scale electrical load simulation for smart grids using GAM modeling; the second focuses on space-time trajectories of wind power generation including parameterized precision matrices based on a Gaussian copula approach. Paper 7 provides flexible and dynamic modeling of dependencies via copulas. Papers 6, 8, and 13 explore different aspects of online learning ranging from the most operational for online residential baseline estimation to the more theoretical, which focuses on oracle bounds for prediction errors related to aggregation strategies. The third one is dedicated to the aggregation of experts proposing some resampling ideas to enlarge the basic family of the so-called experts. Papers 2, 4, 7, 12, and 16 study some general approaches to high-dimensional and complex data (inference in high-dimensional models, graphical models and model selection, adaptive spot volatility estimation for high-frequency data, functional classification and prediction). Finally, papers 10 and 15 deal with some special topics in time series, namely, modeling and prediction of time series arising on a graph and optimal reconciliation of contemporaneous hierarchical time series forecasts.

We would like to address our gratitude to the keynote speakers and all the contributors for accepting to participate in these Lecture Notes, and we greatly appreciate the time that they have taken to prepare their papers.

To conclude, we would like to acknowledge the following distinguished list of reviewers who helped improve the papers by providing general and focused feedback to the authors: U. Amato, R. Becker, R. Cao, J. Cugliari, G. Fort, P. Fryzlewicz, F. Gamboa, B. Ghattas, I. Gijbels, S. Girard, E. Gobet, Y. Goude, R. Hyndman, I. Kojadinovic, S. Lambert-Lacroix, J.-M. Loubes, E. Matzner-Løber, G. Oppenheim, P. Pinson, G. Stoltz, A. Verhasselt, A. Wigington, and Q. Yao. We thank them again for their work, dedication, and professional expertise.

Finally, to the editorial office of Springer-Verlag and particularly to Eva Hiripi, we are grateful for their efficient assistance.

Grenoble, France — Anestis Antoniadis
Clamart, France — Xavier Brossat
Paris, France — Jean-Michel Poggi
Paris, June 2014.

Contents