



## Editorial

# Complexity in Forecasting and Predictive Models

**Jose L. Salmeron<sup>1</sup>, Marisol B. Correia<sup>2</sup>, and Pedro R. Palos-Sanchez<sup>3</sup>**

<sup>1</sup>*Universidad Pablo de Olavide de Sevilla, Spain*

<sup>2</sup>*Universidade do Algarve e Centro de Investigação, Desenvolvimento e Inovação em Turismo (CiTUR) e CEG-IST, Instituto Superior Técnico, Universidade de Lisboa, Portugal*

<sup>3</sup>*International University of La Rioja, Spain*

Correspondence should be addressed to Pedro R. Palos-Sánchez; [pedro.palos@unir.net](mailto:pedro.palos@unir.net)

Received 20 May 2019; Accepted 21 May 2019; Published 10 June 2019

Copyright © 2019 Jose L. Salmeron et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. Introduction

The challenge of this special issue has been to know the state of the problem related to forecasting modeling and the creation of a model to forecast the future behavior that supports decision making by supporting real-world applications.

This issue has been highlighted by the quality of its research work on the critical importance of advanced analytical methods, such as neural networks, soft computing, evolutionary algorithms, chaotic models, cellular automata, agent-based models, and finite mixture minimum squares (FIMIX-PLS)

Mainly, all the papers are focused on triggering a substantive discussion on how the model predictions can face the challenges around the complexity field that lie ahead. These works help to better understand the new trends in computing and statistical techniques that allow us to make better forecasts. Complexity plays a prominent role in these trends, given the increasing variety and changing data flows, forcing academics to adopt innovative and hybrid methods.

The papers address the recent advances in methodological issues and practices related to the complexity of forecasting and models. All of them are increasingly focused on heterogeneous sources.

Among these techniques, FIMIX-PLS has been applied to study unobserved heterogeneity [1]. This technique was extended by Sarstedt et al. [2]. FIMIX-PLS is the first latent class approach for PLS-SEM [3] and is an exploration tool that allows obtaining the appropriate number of segments in which the sample will be divided. In this sense, the

FIMIX-PLS technique allowed making decisions regarding the number of segments, based on pragmatic reasons and considering practical issues related to current research [4].

FIMIX-PLS calculates the probabilities of belonging to a given segment in which each observation is adjusted to the predetermined numbers of segments by means of the estimation of separate linear regression functions and the belonging of objects corresponding to several segments. Different cases are assigned to the segment with greater probability.

## 2. Static Control Model

A static control model based on Wavelet Transform Weighted Twin Support Vector Regression (WTWTSVR) is proposed in the paper “End-Point Static Control of Basic Oxygen Furnace (BOF) Steelmaking Based on Wavelet Transform Weighted Twin Support Vector Regression” by C. Gao et al. The first approach of the authors was to add a new weighted matrix and coefficient vector into the objective functions of Twin Support Vector Regression (TSVR) to improve the performance of the algorithm, and then a static control model was established based on WTWTSVR and 220 samples in real plant, which consists of prediction models, control models, regulating units, controller, and Basic Oxygen Furnace (BOF).

The results indicated that the control error bound with  $800 \text{ Nm}^3$  in the oxygen blowing volume and 5.5 tons in the weight of auxiliary materials can achieve a hit rate of 90% and 88%, respectively. In conclusion, the proposed model can provide a significant reference for real BOF applications

and can be extended to the prediction and control of other industry applications.

### 3. Hybrid Models

In the paper “Development of Multidecomposition Hybrid Model for Hydrological Time Series Analysis” by H. M. Nazir et al., two hybrid models were developed to improve the prediction precision of hydrological time series data based on the principal of three stages as denoising, decomposition, and decomposed component prediction and summation. The performance of the proposed models was compared with the traditional single-stage model (without denoised and decomposed), with the hybrid two-stage model (with denoised), and with the existing three-stage hybrid model (with denoised and decomposition). Three evaluation measures were used to assess the prediction accuracy of all models (Mean Relative Error (MRE), Mean Absolute Error (MAE), and Mean Square Error (MSE)). The proposed architecture was applied on daily rivers inflow time series data of Indus Basin System and the results showed that the three-stage hybrid models had shown improvement in prediction accuracy with minimum MRE, MAE, and MSE for all cases and that the accuracy of prediction was improved by reducing the complexity of hydrological time series data by incorporating the denoising and decomposition.

### 4. Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)

Y. Zhou et al. proposed an approach that integrates complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and extreme gradient boosting (XGBOOST), the so-called CEEMDAN-XGBOOST, for forecasting crude oil prices in the paper “A CEEMDAN and XGBOOST-Based Approach to Forecast Crude Oil Prices”. To demonstrate the performance of the proposed approach, the authors conducted extensive experiments on the West Texas Intermediate (WTI) crude oil prices. Finally, the experimental results showed that the proposed CEEMDAN-XGBOOST outperforms some state-of-the-art models in terms of several evaluation metrics.

### 5. Machine Learning (ML) Algorithms

The ability to handle large amounts of data incrementally and efficiently is indispensable for Machine Learning (ML) algorithms. The paper “An Incremental Learning Ensemble Strategy for Industrial Process Soft Sensors” by H. Tian et al. addressed an Incremental Learning Ensemble Strategy (ILES) that incorporates incremental learning to extract information efficiently from constantly incoming data. The ILES aggregates multiple sublearning machines by different weights for better accuracy.

The weight updating rules were designed by considering the prediction accuracy of submachines with new arrived data, so that the update can fit the data change and obtain

new information efficiently. The sizing percentage soft sensor model was established to learn the information from the production data in the sizing of industrial processes and to test the performance of ILES, where the Extreme Learning Machine (ELM) was selected as the sublearning machine. The results of the experiments demonstrated that the soft sensor model based on the ILES has the best accuracy and ability of online updating.

### 6. Bass Innovation Diffusion Model

M. L. Bertotti and G. Modanese in “The Bass Diffusion Model on Finite Barabasi-Albert Networks” used a heterogeneous mean-field network formulation of the Bass innovation diffusion model and exact results by Fotouhi and Rabbat [5] on the degree correlations of Barabasi-Albert (BA) networks to compute the times of the diffusion peak and to compare them with those on scale-free networks, which have the same scale-free exponent but different assortativity properties.

The authors compared their results with those obtained by D’Agostino et al. [6] for the SIS epidemic model with the spectral method applied to adjacency matrices and turned out that diffusion times on finite BA networks were at a minimum. They believe this may be due to a specific property of these networks, e.g. whereas the value of the assortativity coefficient is close to zero, the networks look disassortative if a bounded range of degrees is solely considered, including the smallest ones, and slightly assortative on the range of the higher degrees.

Finally, the authors found that if the trickle-down character of the diffusion process is enhanced by a larger initial stimulus on the hubs (via a inhomogeneous linear term in the Bass model), the relative difference between the diffusion times for BA networks and uncorrelated networks is even larger.

### 7. Forecast Copper Treatment Charges (TC)/Refining Charges (RC)

In order to overcome the lack of research about the price at which mines sell copper concentrate to smelters, the paper “Looking for Accurate Forecasting of Copper TC/RC Benchmark Levels” by F. J. Díaz-Borrego et al. provides a tool to forecast copper Treatment Charges (TC)/Refining Charges (RC) annual benchmark levels, in which a three-model comparison was made by contrasting different measures of error.

The results indicated that the Linear Exponential Smoothing (LES) model is the one that has the best predictive capacity to explain the evolution of TC/RC in both the long and the short term. Finally, the authors believe this suggests a certain dependency on the previous levels of TC/RC, as well as the potential existence of cyclical patterns in them and that this model enables a more precise estimation of copper TC/RC levels, which is useful for smelters and mining companies.

## 8. Ant Colony Optimization Algorithm

F. Liu et al. in “Prediction of Ammunition Storage Reliability Based on Improved Ant Colony Algorithm and BP Neural Network” proposed an Improved Ant Colony Optimization (IACO) algorithm based on a three-stage and a Back Propagation (BP) Neural Network algorithm to predict ammunition failure numbers, where the main affecting factors of ammunition include temperature, humidity, and storage period and where the reliability of ammunition storage is obtained indirectly by failure numbers. Experimental results show that IACO-BP algorithm has better accuracy and stability in ammunition storage reliability prediction than BP network, Particle Swarm Optimization-Back Propagation (PSO-BP), and ACO-BP algorithms.

## 9. Finite Mixture Partial Least Squares (FIMIX-PLS)

Finally, the paper “Green Start-Ups’ Attitudes towards Nature When Complying with the Corporate Law” by R. Robina-Ramírez et al. examined how Spanish green start-ups develop improved attitudes towards nature having in mind the framework of the new Spanish Criminal Code in relation to corporate compliance. Smart Partial Least Squares (PLS) Path Modelling was used to build an interaction model among variables and unobserved heterogeneity was analysed using Finite Mixture Partial Least Squares (FIMIX-PLS).

The model reveals a strong predictive power ( $R^2 = 77.9\%$ ) with a sampling of one hundred and fifty-two Spanish start-ups. This test is performed in a set of four stages. In the first one, it executes FIMIX, and in a second it calculates the number of optimal segments. In the third one, we discussed the latent variables that justify these segments, to finally estimate the model and segments [7].

The results allow concluding that Spanish start-ups should implement effective monitoring systems and new organisational standards, in order to develop surveillance measures to avoid unexpected sanctions.

## 10. Conclusion

An overview of the 8 selected papers for this special issue has been presented, reflecting the most recent progress in the research field. We hope this special issue can further stimulate interest in the critical importance of advanced analytical methods, such as neural networks, soft computing, evolutionary algorithms, chaotic models, cellular automata, agent-based models, and finite mixture minimum squares (FIMIX-PLS). More research results and practical applications on advanced analytical methods are expected in the future.

## Conflicts of Interest

The guest editor team does not have any conflicts of interest or private agreements with companies.

*Jose L. Salmeron  
Marisol B. Correia  
Pedro R. Palos-Sanchez*

## References

- [1] C. Hahn, M. D. Johnson, A. Herrmann, and F. Huber, “Capturing customer heterogeneity using a finite mixture PLS approach,” *Schmalenbach Business Review*, vol. 54, no. 3, pp. 243–269, 2002 (Russian).
- [2] M. Sarstedt, J.-M. Becker, C. M. Ringle, and M. Schwaiger, “Uncovering and treating unobserved heterogeneity with FIMIX-PLS: which model selection criterion provides an appropriate number of segments?” *Schmalenbach Business Review*, vol. 63, no. 1, pp. 34–62, 2011.
- [3] M. Sarstedt, “A review of recent approaches for capturing heterogeneity in partial least squares path modelling,” *Journal of Modelling in Management*, vol. 3, no. 2, pp. 140–161, 2008.
- [4] M. Sarstedt, M. Schwaiger, and C. M. Ringle, “Do we fully understand the critical success factors of customer satisfaction with industrial goods? - extending festge and schwaigers model to account for unobserved heterogeneity,” *Journal of Business Market Management*, vol. 3, no. 3, pp. 185–206, 2009.
- [5] B. Fotouhi and M. G. Rabbat, “Degree correlation in scale-free graphs,” *The European Physical Journal B*, vol. 86, no. 12, article no 510, 2013.
- [6] G. D’Agostino, A. Scala, V. Zlatić, and G. Caldarelli, “Robustness and assortativity for diffusion-like processes in scale-free networks,” *EPL (Europhysics Letters)*, vol. 97, no. 6, article no 68006, 2012.
- [7] P. Palos-Sánchez, F. Martín-Velicia, and J. R. Saura, “Complexity in the acceptance of sustainable search engines on the internet: an analysis of unobserved heterogeneity with FIMIX-PLS,” *Complexity*, Article ID 6561417, 19 pages, 2018.

