

INNOVATION, ENTREPRENEURSHIP AND MANAGEMENT SERIES

BIG DATA, ARTIFICIAL INTELLIGENCE AND DATA ANALYSIS SET



Volume 14

Data Analysis in the AI Era

Models, Methods and Techniques

Edited by
James R. Bozeman
Christos H. Skiadas

ISTE

WILEY

Table of Contents

[Cover](#)

[Table of Contents](#)

[Title Page](#)

[Copyright Page](#)

[Introduction](#)

[I.1. Part 1: health and demography](#)

[I.2. Part 2: applications](#)

[I.3. Part 3: statistical techniques](#)

[PART 1: Health and Demography](#)

[1 Standard and Fertility-adjusted Equitable Normal Pension Age](#)

[1.1. Introduction](#)

[1.2. Methodology](#)

[1.3. Data source and calculations](#)

[1.4. Results](#)

[1.5. Conclusion](#)

[1.6. References](#)

[2 Number of Centenarians in Brazil and Selected Countries: Evaluation of the Quality of Census Counting](#)

[2.1. Introduction](#)

[2.2. Methodology](#)

[2.3. Results of Review 2](#)

[2.4. Conclusion](#)

[2.5. References](#)

[3 Optimal Semi-Markov Model Selection for Pediatric Obstructive Sleep Apnea Patient Data](#)

[3.1. Introduction](#)

[3.2. Sleep apnea data overview](#)

[3.3. Bayes factors](#)

[3.4. Semi-Markov models for sleep states](#)

[3.5. Sleep data model comparison](#)

[3.6. Conclusion](#)

[3.7. References](#)

[4 An Efficient Determination Criterion-based Metric for Z-DNA Pattern Identification in the Epstein-Barr Virus Associated with Nasopharyngeal Carcinoma](#)

[4.1. Introduction](#)

[4.2. Results](#)

[4.3. Discussion](#)

[4.4. References](#)

[5 Investigating the Mediating Role of Dis/Trust in Government to Deal with the Pandemic in the Relationship Between Covid-19 Cases, the Impact of the Pandemic on Employment and Social Trust: Evidence From the 2020 European Social Survey for Four European Countries](#)

[5.1. Introduction](#)

[5.2. Method](#)

[5.3. Results](#)

[5.4. Conclusion](#)

[5.5. References](#)

[6 Joint Forecasting of Screening/Randomization Processes in Multicenter Clinical Trials and Predicting Optimal Enrollment Stopping Time](#)

[6.1. Introduction](#)

[6.2. Structure of screening stages](#)

[6.3. Predicting randomized patients from screening pool](#)

[6.4. Predicting the number of randomized patients](#)

[6.5. Randomization process from future arrivals](#)

[6.6. Parameter estimation](#)

[6.7. Implementation in R tool](#)

[6.8. Future considerations](#)

[6.9. Acknowledgment](#)

[6.10. References](#)

[PART 2: Applications](#)

[7 Bidding Strategies Approaches for the Greek Electricity Market](#)

[7.1. Introduction](#)

[7.2. Methodology](#)

[7.3. Stand-alone and hybrid BESS bidding strategies](#)

[7.4. Conclusion](#)

[7.5. References](#)

[8 Long-Term Memory, Event-Driven Dynamics and Predictive Power: An Analysis of Economic Policy Uncertainty and Market Volatility in the United States and UK](#)

[8.1. Introduction](#)

[8.2. Background and theoretical foundations](#)

[8.3. Methodology](#)

[8.4. Empirical results](#)

[8.5. Discussion](#)

[8.6. Conclusion and future research](#)

[8.7. References](#)

[8.8. Appendix: EPU US Daily Policy Data](#)

[9 Spatial Dependence in Extreme Precipitation: An Extreme Value Analysis in R](#)

[9.1. Introduction](#)

[9.2. Asymptotic models for spatial extremes](#)

[9.3. An application using the R environment – rainfall data in the North of Portugal](#)

[9.4. Final remarks](#)

[9.5. Acknowledgments](#)

9.6. References

10 Time Series and Cluster Analysis of Groundwater Level in Slovakia

10.1. Introduction

10.2. Data specifications and methods

10.3. Analysis of trend and periodic components in groundwater levels

10.4. The residuals component and ARMA modeling

10.5. Probe cluster analysis and visualization

10.6. Conclusion

10.7. Acknowledgment

10.8. References

11 Coupon Collector Problem with Partial and Total Resets

11.1. Introduction and motivation

11.2. CCP with total resets

11.3. CCP with partial resets

11.4. Model comparison

11.5. Numerical results

11.6. Discussion

11.7. Acknowledgments

11.8. References

12 The Relationship Between Quality of Life and Brain Drain

12.1. Introduction

12.2. Data and methodology

12.3. Results

12.4. Conclusion

12.5. Acknowledgments

12.6. References

PART 3: Statistical Techniques

13 A Unified Model for SPC and Maintenance in High Investigation Cost Processes with Multiple Quality Disruptions

13.1. Introduction

13.2. Problem statement

13.3. Mathematical model

13.4. Cost structure analysis

13.5. Optimization framework

13.6. Numerical analysis

13.7. Conclusion

13.8. References

14 A Note on Spectral Risk Measures when Systemic Risk is Present

14.1. Introduction and motivation

14.2. Preliminaries

14.3. Theoretical framework

14.4. Contribution

14.5. A suggested application

14.6. Conclusion

14.7. References

14.8. Appendix A

14.9. Appendix B

15 Mutual Information and Pearson's Correlation in Analysis of Statistical Tests Suites: A Case Study

15.1. Introduction

15.2. ENT battery analysis

15.3. Conclusion

15.4. References

16 Binary Regression and Non-monotonic Link Functions

16.1. Introduction

16.2. Proposed method

[16.3. Methodology](#)

[16.4. Comparison](#)

[16.5. Conclusion](#)

[16.6. References](#)

[List of Authors](#)

[Index](#)

[Other titles from iSTE in Innovation, Entrepreneurship and Management](#)

[End User License Agreement](#)

List of Tables

Chapter 2

[Table 2.1. Classification of scores to assess the quality of the census...](#)

[Table 2.2. Reference intervals for selected indicators by categories of...](#)

[Table 2.3. Mean age, standard deviation, minimum and maximum values of ...](#)

[Table 2.4. Classification of the quality of census data on centenarians...](#)

[Table 2.5. Mean age, standard deviation, minimum and maximum values of ...](#)

[Table 2.6. Census data quality rating of centenarians for selected coun...](#)

Chapter 4

[Table 4.1. From left to right: \(i\) sequence; \(ii\) GenBank and version o...](#)

[Table 4.2. Estimated transition probabilities – equation \[4.2\] – from t...](#)

[Table 4.3. ds values – definition 4.1 – s = cg, with \(left\) and \(righ...](#)

[Table 4.4. Estimated transition probabilities – equation \[4.2\] – from t...](#)

[Table 4.5. ds values – definition 4.1 – s = tg, with \(left\) and \(righ...](#)

[Table 4.6. Estimated transition probabilities – equation \[4.2\] – from t...](#)

[Table 4.7. ds values – definition 4.1 – s = ca, with \(left\) and \(righ...](#)

Chapter 5

[Table 5.1. The demographic and social characteristics of participants: ...](#)

[Table 5.2. Path analyses performed with maximum likelihood: goodness-of...](#)

[Table 5.3. Path analyses performed with maximum likelihood: standardize...](#)

[Table 5.4. Path analyses: standardized indirect effects \(SIE\), their 95...](#)

Chapter 7

[Table 7.1. Key metrics for DAM and IDM](#)

[Table 7.2. DAM, IDA and RES curtailment correlations. For a color versi...](#)

Chapter 8

[Table 8.1. US EPU Hurst exponents by subperiod](#)

[Table 8.2. UK EPU Hurst exponents by subperiod](#)

Chapter 9

[Table 9.1. Examples of valid correlation functions \$\rho\(h\)\$ commonly used i...](#)

[Table 9.2. Identification of the 18 stations plotted in Figure 9.1](#)

Chapter 10

[Table 10.1. Trend and periodic components for some probes](#)

[Table 10.2. Linear models for some probes](#)

Chapter 11

[Table 11.1. Expected waiting time and variance of the waiting time in C...](#)

[Table 11.2. Expected waiting time and variance of the waiting time in C...](#)

Chapter 12

[Table 12.1. List of variables](#)

Chapter 13

[Table 13.1. Transition probability matrix](#)

[Table 13.2. Set parameters and results](#)

[Table 13.3. Numerical example: breakdown of cost and decision variables...](#)

Chapter 15

[Table 15.1. Correlation matrix of samples of p-values](#)

[Table 15.2. Significance matrix obtained from the applied K-S test \(cas...](#)

[Table 15.3. Mutual information results](#)

[Table 15.4. Significance matrix obtained from the applied K-S test \(cas...](#)

Chapter 16

[Table 16.1. Estimated parameters using different starting values](#)

[Table 16.2. Inference](#)

List of Illustrations

Chapter 1

[Figure 1.1. Completed cohort fertility for birth cohorts 1965–2025....](#)

[Figure 1.2. Variants of pension age for birth cohorts 1965–2025....](#)

[Figure 1.3. Adjusted old-age dependency ratio 2030–2100 – low fertility...](#)

[Figure 1.4. Adjusted old-age dependency ratio 2030–2100 – low fertility...](#)

[Figure 1.5. Adjusted old-age dependency ratio 2030–2100 – medium fertil...](#)

[Figure 1.6. Adjusted old-age dependency ratio 2030–2100 – high fertilit...](#)

Chapter 3

[Figure 3.1. The sleep states for patient CF001 \(top\) and CF050 \(bottom\).](#)

[Figure 3.2. The empirical distributions of the durations for the awake ...](#)

[Figure 3.3. Boxplots of the 75 Bayes factors comparing the negative bin...](#)

[Figure 3.4. A scatter plot of the posterior probabilities of the 75.pat...](#)

Chapter 4

[Figure 4.1. Dendrograms of ds values \(definition 4.1\) for s = gc \(compl...](#)

[Figure 4.2. Dendrograms of ds values \(definition 4.1\) for s = at \(compl...](#)

[Figure 4.3. Dendrograms of ds values \(definition 4.1\) for s = cg \(compl...](#)

[Figure 4.4. Dendrograms of ds values \(definition 4.1\) for s = tg \(compl...](#)

[Figure 4.5. Dendrograms of ds values \(definition 4.1\) for s = ca \(compl...](#)

Chapter 5

[Figure 5.1. The theoretical structure of the path analysis model](#)

[Figure 5.2. The frequency_\(%\) of respondents having had Covid-19: Europ...](#)

[Figure 5.3. The frequency_\(%\) of anyone living with respondents having ...](#)

[Figure 5.4. The frequency_\(%\) of the impact of the pandemic on the empl...](#)

[Figure 5.5. The mean values of the variables assessing the trust in gov...](#)

[Figure 5.6. The mean values of the variables assessing the social trust...](#)

Chapter 6

[Figure 6.1. Diagram of screening flow, showing individual stages, possi...](#)

[Figure 6.2. Example implementation of screening rule, showing metrics o...](#)

[Figure 6.3. Histogram showing results of advised stopping day over 1,00...](#)

[Figure 6.4. Example of a predicted enrollment trajectory, using same th...](#)

Chapter 7

[Figure 7.1. Intraday \(IDAs\) auctions](#)

[Figure 7.2. XBID auctions](#)

[Figure 7.3. Superpose of IDM and RES curtailment](#)

Chapter 8

[Figure 8.1. US Economic Policy Uncertainty \(EPU\) Index Over Time](#)

[Figure 8.2. UK Economic Policy Uncertainty \(EPU\) Index Over Time](#)

[Figure 8.3. Lagged correlation between US EPU and VIX.](#)

[Figure 8.4. Granger Causality Tests](#)

[Figure 8.5. DFA log-log.plot](#)

[Figure 8.6. DFA log-log.plot \(pre-2008 Crisis\)](#)

[Figure 8.7. DFA log-log.plot \(2008 Financial Crisis\)](#)

[Figure 8.8. DFA log-log.plot \(Eurozone Crisis\)](#)

[Figure 8.9. DFA log-log.plot \(Brexit Period\)](#)

[Figure 8.10. DFA log-log.plot \(post-Brexit/pre-Covid\)](#)

[Figure 8.11. DFA log-log.plot \(Covid-19 pandemic\)](#)

[Figure 8.12. DFA log-log.plot \(war shock/recent events\)](#)

[Figure 8.13. DFA log-log.plot \(UK policy data\)](#)

[Figure 8.14. DFA log-log.plot \(pre-2008 crisis – UK EPU\)](#)

[Figure 8.15. DFA log-log.plot \(2008-financial crisis – UK EPU\)](#)

[Figure 8.16. DFA log-log.plot \(eurozone crisis – UK EPU\)](#)

[Figure 8.17. DFA log-log.plot \(Brexit Period – UK EPU\)](#)

[Figure 8.18. DFA log-log.plot \(post-Brexit/pre-Covid – UK EPU\)](#)

[Figure 8.19. DFA log-log.plot \(Covid-19 pandemic – UK EPU\)](#)

[Figure 8.20. DFA log-log.plot \(war shocks/recent events – UK EPU\)...](#)

Chapter 9

[Figure 9.1. Annual maxima of daily rainfall values \(mm\) across municipa...](#)

[Figure 9.2. Diagnostic plots for Deilão \(left\) and Rego da Murta \(right...](#)

[Figure 9.3. Shape, location and scale parameters plotted for all the st...](#)

[Figure 9.4. Shape, location and scale parameters plotted against longit...](#)

[Figure 9.5. Pairwise extremal coefficient estimates and loess curves fr...](#)

[Figure 9.6. One simulation over a grid from Smith \(left\) and Schlather ...](#)

Chapter 10

[Figure 10.1. Part of hydrological network for groundwater quantity moni...](#)

[Figure 10.2. Daily data, probe 624.](#)

[Figure 10.3. Monthly data, probe 624.](#)

[Figure 10.4. Seasonal subseries plot for probe 624.](#)

[Figure 10.5. Periodogram for probe 624.](#)

[Figure 10.6. Dendrogram of probes.](#)

[Figure 10.7. Optimal number of clusters](#)

[Figure 10.8. Cluster visualization.](#)

Chapter 11

[Figure 11.1. Expected waiting time in CCPTR in terms of pR for differen...](#)

[Figure 11.2. Expected waiting time in CCPPR1 in terms of pR1 for differ...](#)

[Figure 11.3. Expected waiting time in CCPPR2 in terms of pR2 for differ...](#)

Chapter 12

[Figure 12.1. Proportion of variance explained by the first component](#)

[Figure 12.2. Dynamics of the first component weight loads.](#)

[Figure 12.3. Heatmap of quality of life.](#)

[Figure 12.4. Correlation between economic well-being and brain drain ov...](#)

[Figure 12.5. Countries that change their income group](#)

[Figure 12.6. The correlation between quality of life and brain drain. ...](#)

Chapter 13

[Figure 13.1. Proposed SPC approach control chart](#)

[Figure 13.2. Optimized action limits and cost sensitivity analysis.](#)

Chapter 15

[Figure 15.1. Dispersion matrix of the p-values](#)

Chapter 16

[Figure 16.1. Different shapes](#)

[Figure 16.2. ROC curve comparison](#)

Big Data, Artificial Intelligence and Data Analysis Set

Volume 14

Data Analysis in the AI Era

Models, Methods and Techniques

Edited by

James R. Bozeman

Christos H. Skiadas

ISTE

WILEY

First published 2026 in Great Britain and the United States by ISTE Ltd and John Wiley & Sons, Inc.

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms and licenses issued by the CLA. Enquiries concerning reproduction outside these terms should be sent to the publishers at the undermentioned address:

ISTE Ltd
27-37 St George's Road
London SW19 4EU
UK

www.iste.co.uk

John Wiley & Sons, Inc.
111 River Street
Hoboken, NJ 07030
USA

www.wiley.com

© ISTE Ltd 2026

The rights of James R. Bozeman and Christos H. Skiadas to be identified as the authors of this work have been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s), contributor(s) or editor(s) and do not necessarily reflect the views of ISTE Group.

Library of Congress Control Number: 2026943231

British Library Cataloguing-in-Publication Data
A CIP record for this book is available from the British Library
ISBN 978-1-83669-139-6

The manufacturer's authorized representative according to the EU General Product Safety Regulation is Wiley-VCH GmbH, Boschstr. 12, 69469 Weinheim, Germany, e-mail:
Product_Safety@wiley.com.

Introduction

This book describes the process of exploring and analyzing data from various sources for the purpose of extracting information, building models and making decisions and inferences. The artificial intelligence (AI) and data analysis field is a scientific area that has grown enormously, especially during the last few decades following a rapidly growing computer industry and the wide applicability of computational techniques in conjunction with new advances in analytic tools. Modeling enables data analysts to apply various statistical methods to the data they are investigating, identify relationships between variables, make predictions about future sets of data, and understand, interpret and visualize the extracted information more strategically. New research results have been developed and published and many more are currently in progress. The topic is also exhibited widely in many international scientific conferences and workshops. One such conference, Applied Stochastic Modelling and Data Analysis (ASMDA) 2025 and Demographics 2025 Workshop, held 10–13 June 2025 at the University of Piraeus, Greece, is where most of the work in this volume was presented. These facts are the evidence of increasing interest in this area. Therefore, we believe that a collection of chapters from experts in these fields will fulfill the demand from wide circles of the scientific community.

This book is a collective work by a number of leading scientists, analysts, engineers, mathematicians and statisticians who have been working on the front end of data analysis and applications of models, whether they arise from data science, operations research, engineering, machine learning or statistics. The chapters included represent a cross-section of current concerns and research interests in the above scientific areas. The material is divided in a form to provide the reader with both theoretical and applied information on data analysis methods, models and techniques along with appropriate applications in AI.

I.1. Part 1: health and demography

This portion of the book consists of six application chapters specific to the fields of health and demography. Two chapters refer to different end-of-life issues: *Standard and fertility-adjusted equitable normal pension age*, by Tomáš Fiala, Jitka Langhamrová and Jana Vrabcová, and *Number of centenarians in Brazil and selected countries: Evaluation of the quality of census counting*, by Neir Antunes Paes, Rozane Pereira de Sousa and Vivian Conceição Alves Leite Pereira do Lago. These chapters are followed by *Optimal semi-Markov model selection for pediatric obstructive sleep apnea patient data*, by Matthew Fischer, Adam B Kashlak, Mathieu Chalifour and Giseon Heo. The next chapter, *An efficient determination criterion-based metric for Z-DNA pattern identification in the Epstein-Barr virus associated with nasopharyngeal carcinoma*, by James R. Bozeman, V.A. González-López, Jesús E. García and D.S. Santos, investigates if Z-DNA, the left-handed version, is present in the Epstein–Barr virus. This chapter is followed by *Investigating the mediating role of dis/trust in government to deal with the pandemic in the relationship between Covid-19 cases, the impact of the pandemic on employment and social trust: Evidence from the 2020 European social survey for four European countries*, by Aggeliki Yfanti and Catherine Michalopoulou. This section concludes with a chapter about clinical trials, *Joint forecasting of screening/randomization in multicenter clinical trials and predicting optimal enrollment stopping time*, by Volodymyr Anisimov and Leo Gortsak.

I.2. Part 2: applications

This part of the book includes six important chapters applying data analysis in various areas. The first two refer to markets, namely *Bidding strategies approaches for Greek electricity market*, by Emmanuel Karapidakis, Marios Nikologiannis and George Matalliotakis, and *Long-term memory, event-driven dynamics and predictive power: An analysis of economic policy uncertainty and market volatility in the US and UK*, by Christos Floros, George Matalliotakis and Fotis Papatheofanous. Next are two climate-related chapters: *Spatial dependence in extreme precipitation: An extreme value analysis in R*, by Dora Prata Gomes and Maria Manuela Neves, and *Time series and cluster analysis of groundwater level in Slovakia*, by Jana Kalicka, Dominika Sonak Ballova and Michaela Cervenanska. This is followed by *Coupon collector problem with partial and total resets*, by Jelena Jocković and Bojana Todić, about a long-standing problem in mathematics. This section concludes with *The relationship between quality of life and brain drain*, by Aleksei N. Kurbatskii and Marina V. Mironenkova.

I.3. Part 3: statistical techniques

The last section of the book, consisting of four parts, introduces new methods. It starts with *A unified model for SPC and maintenance in high investigation cost processes with multiple quality disruptions*, by Konstantina Tsiota and Konstantinos A. Tasias. The next chapter is *A note on spectral risk measures when systemic risk is present*, by Zachos C. Georgios. This chapter is followed by *Mutual information and Pearson's correlation in analysis of statistical tests suites: A case study*, by Elena Almaraz Luengo. The book concludes with *Binary regression and non-monotonic link functions*, by Gloria Gheno.

We thank the authors for their contributions, and we pass on our sincere appreciation to the referees for their hard work in making this a better book. Many thanks to the secretariat and, of course, to the publisher. We are happy to have edited, organized and provided this material in its final form for the volume.

Note

Chapter written by James R. BOZEMAN and Christos H. SKIADAS.

PART 1

Health and Demography

1

Standard and Fertility-adjusted Equitable Normal Pension Age

Raising the retirement age is an often-used measure for the financial stability and sustainability of the pension system. It is an unpopular measure and therefore needs to be implemented in such a way that it can be considered intergenerationally fair. One such proposal is the idea of equitable normal pension age (ENPA), which assumes the same relation between the (expected) total years of receiving a pension and the (expected) total years of economic activity for all cohorts. Cohort life tables should be used for correct computations.

In many countries (including Czechia), the old-age pension system is based not on the cohort, but on the pay-as-you-go (PAYG) principle. In such a case, the ENPA would guarantee the stability of the pension system only in the case of a stationary-type population, that is, under fertility at the replacement level (or adequate compensation of lower fertility by foreign immigration). The fertility-adjusted ENPA could be determined by comparing the relation between the total years of receiving a pension of a (parents') cohort and the total years of economic activity of the cohort of their children. The size of this cohort depends on the completed fertility of the parents' cohort.

A very simple and rough indicator of the financial burden of this pension system is the adjusted old-age dependency ratio (AOADR), defined as the ratio of the population at retirement age to the population at productive age using the actual retirement age threshold instead of the standard 65 years.

The chapter shows the calculation of standard ENPA and fertility-adjusted ENPA and corresponding AOADR for Czechia in various scenarios of future demographic development.

1.1. Introduction

Population aging is one of the most important and very often discussed demographic phenomena in recent decades. It, and especially its economic, social and political consequences, is broadly mentioned not only by experts and politicians, but also by the public and media. All (not only) economically advanced populations face or will very soon face the aging process. Most of the analytical approaches to population aging are based on the assumption that the threshold of old age is unchanged in time, fixed and equal usually at 65 years of age (which is the threshold of the retirement age in many countries). Chronological age is a very simple and easily understood characteristic, and is thus regarded as useful in the specification of public policies. However, this “traditional” approach contradicts the continuing changes in demographic development, especially the decrease in mortality rates, the increase in life span and the increase in healthy life span.

Almost 50 years ago, Ryder ([1976](#)) established a new way of determining the old age threshold. He suggested the point of entry into old age as the age at which the (remaining) life expectancy is equal to a given, relatively low, value, say 10 years. The idea of an old age threshold flexible in time, depending not on the standardly used chronological age, but on the remaining life expectancy, was used later by several authors, for example, by Fuchs ([1984](#)). Thirty years ago, Siegel ([1993](#)) proposed determining the old-age threshold as the age when life expectancy equals 15 years.

This alternative concept of the definition of human age was very particularly focused on by Sanderson and Scherbov in several papers. They introduced a new forward-looking definition of age (so-called prospective age) and argued that, along with the usual backward-looking concept of age, this definition provides a more informative basis to discuss population aging (Sanderson and Scherbov [2005](#), [2007](#)). The use of the prospective instead of the traditionally used chronological age threshold of old means that the old age threshold is (assuming declining mortality rates) rising in time. The rise of many indicators of aging will then not be so rapid in comparison with indicators defined in the standard way (Sanderson and Scherbov [2013](#), [2014](#)). The calculations of some prospective

indicators for the population of EU countries were published by Šídlo et al. ([2020a](#)).

One of the most serious consequences of an aging population is the issue of financing old-age pensions. In the case of the fixed retirement age threshold, the decrease in mortality would result in an increase in the length of time of pension receipt, which could be a serious threat to the financial sustainability of pension systems. Pension reforms in many states include the idea of further increasing retirement age above the standard threshold of 65 years. The concept of prospective age can be used for the determination of the flexible retirement age threshold, reflecting the previous, as well as expected future permanent increase in life expectancy. However, the simplest idea to determine the retirement age threshold as the age at which the remaining life expectancy would reach a fixed value (e.g. 15 or 20 years) would mean that the period of economic activity would grow while the period of pension receipt would remain constant, so the relative period of pension receipt would drop.

Therefore, Sanderson and Scherbov introduced the concept of the so-called intergenerationally ENPA, where the retirement age threshold is determined in such a way that the ratio of the (expected) number of person-years lived by members of the cohort in the pension phase of life to the (expected) number of person-years lived in the economic activity period remains stable over time and space for all population cohorts, regardless of their mortality rates. Calculations of the intergenerationally equitable pension age for Czechia, Hungary, Poland and Slovakia (V4 countries) were published by Šídlo et al. ([2020b](#)).

Fixed normal pension ages are already being replaced by variable ones in many countries. For people to understand, accept and voluntarily adjust to changes in national pension ages, two criteria must be met. The rationale for pension age policy must be compelling, simple and transparent, and the resulting policies must be clearly intergenerationally equitable. The so-called intergenerationally ENPA is an analytically based determination of pension age policy that meets both of those criteria. The proposed pension age (varying with changing mortality conditions) expressed the idea that members of each birth cohort must receive

approximately as much money in pension benefits as they contribute to the pension system. At the same time, the pension tax rate, as well as the ratio of annual pension income to the wage in the pre-pension period, should be unchanged over time. The pension age is then determined in such a way that the proportion of time spent by the cohort in retirement age to the proportion of adult life is the same for each cohort (Sanderson and Scherbov [2014](#)).

In many countries (including Czechia), the old-age pension system is based not on the cohort, but on the pay-as-you-go (PAYG) principle. In such a case, the ENPA would guarantee the financial stability of the pension system only in the case of a stationary type population, that is, under fertility at the replacement level (or adequate compensation of lower fertility by foreign immigration). The ENPA reflecting fertility could be determined by stabilizing the ratio of the number of person-years lived by members of the cohort in the pension phase of life to the number of person-years lived in the economic activity period by the cohort of their children. The size of this child cohort depends on the completed fertility rate of the parent cohort.

A very simple and rough indicator of the financial burden of the PAYG pension system is the adjusted old-age dependency ratio (AOADR), defined as the ratio of the population at retirement age to the population at productive age using the actual retirement age threshold instead of the standard 65 years.

The chapter compares retirement age in Czechia according to current legislation with standard ENPA and ENPA reflecting fertility for Czechia for birth cohorts born in 1966 and later (retiring after 2030), and presents corresponding values of AOADR in various scenarios of future demographic development.

1.2. Methodology

We assume that the economic activity period starts at the age of 20 and lasts until retirement age is reached. If $\alpha(g)$ denotes the ENPA of birth cohort g , its value can then be determined as the (unique) solution of the equation (Sanderson and Scherbov [2014](#))

$$\frac{T_{\alpha(g)}^{(g)}}{T_{20}^{(g)} - T_{\alpha(g)}^{(g)}} = \frac{T_{\alpha(g_0)}^{(g_0)}}{T_{20}^{(g_0)} - T_{\alpha(g_0)}^{(g_0)}}, \quad [1.1]$$

where g_0 is some reference cohort. The cohort life tables should be used for calculations; $T_x^{(g)}$ denotes the T_x value of the life tables for the birth cohort g . The left-hand side term of equation [1.1] can be interpreted as the ratio of the number of person-years lived by members of the cohort in retirement age to the number of person-years lived by the same cohort in productive age. The right-hand side term of equation [1.1] can be replaced by an appropriately chosen constant value. The solution $\alpha(g)$ of equation [1.1] is found using the linear interpolation method.

It is clear that the value of ENPA depends only on mortality trends, not on fertility and migration trends. It could be financially sustainable in the long term in a funded pension system. In a pay-as-you-go system, the assumption of fertility on replacement level (or corresponding “replacement” migration level) would be necessary to ensure the long-term functioning of this system. In the case of declining fertility and low net migration, the system based on the ENPA idea may be financially unsustainable. Due to the population decline, the ratio of the number of person-years lived by members of some cohorts of retirement age to the number of person-years lived by the cohort of their children (whose contributions pay for the pensions of the parent generation) of productive age may be significantly higher than that indicated by equation [1.1].

A possible modification of the ENPA reflecting fertility level (ENPARF) may be performed as follows. The denominator of fraction (1) will be replaced by the total number of years lived in productive age, not by the cohort g , but by the cohort of their children. The ratio of the size of the cohort of children to the cohort of their parents can be estimated as the ratio of the completed fertility (CCF) of the cohort to the value of 2.1, which ensures replacement level fertility.

The value of ENPARF for birth cohort g can then be determined as the solution of the following equation:

$$\frac{T_{\alpha(g)}^{(g)}}{(T_{20}^{(g)} - T_{\alpha(g)}^{(g)}) \cdot \frac{CCF^{(g)}}{2.1}} = \frac{T_{\alpha(g_0)}^{(g_0)}}{(T_{20}^{(g_0)} - T_{\alpha(g_0)}^{(g_0)}) \cdot \frac{CCF^{(g_0)}}{2.1}}, \quad [1.2]$$

where $CCF^{(g)}$ denotes the completed cohort fertility of birth cohort g .

Cohort life tables are usually unavailable. They can be calculated using a series of period life tables covering all years of living of the cohort. The cohort probabilities of dying can be estimated as the average of appropriate period probabilities of dying by the following formula:

$$q_x^{(g)} = \frac{q_{g+x,x} + q_{g+x+1,x}}{2}, \quad [1.3]$$

where $q_x^{(g)}$ is the cohort table probability of dying at age x for birth cohort g and $q_{t,x}$ is the period table probability of dying at age x in year t . Unisex probabilities of dying (the average of males' and females' values) were used for further calculations.

The cohort life tables are then calculated by known standard formulas. Values of T_x depend on mortality rates of ages equal to x and higher. The value of ENPA does not depend on mortality rates until 20 years of age, so sufficient data for calculations of ENPA for birth cohort g are probabilities of dying in years $(g + 20) - (g + \omega)$, where ω denotes the highest assumed age, for example, 110 years. Cohort fertility rates were estimated using period values in the same way.

A very simple and rough indicator of the financial burden of the PAYG pension system is the adjusted old-age dependency ratio defined as the ratio of the population size at retirement age to the population size at productive age.

$$AOADR = \frac{S_{t,r(t)+}}{S_{t,20-r(t)}}, \quad [1.4]$$

where $S_{t,x}$ is the population size at time t at age x , and $r(t)$ denotes the retirement age threshold at time t . The linear interpolation principle is used for calculations in the case of the non-integer value of retirement age.

The ENPA reflecting fertility does not take into account the fact that in Czechia, as well as in many other countries, low fertility is usually partially compensated by positive net migration. In such a case, the increase in the retirement age determined according to the ENPARF principle could be slowed down somewhat, provided that the financial stability of the pension system is maintained.

The financial stability of the pension system would be ensured in the case of long-term stable AOADR values. The financial stability of the PAYG pension system in the period $(t_1; t_2)$ can be defined by the requirement that the ratio of the sum of the number of people of retirement age in each year of the period to the sum of the number of people of working age in the same period be equal to the required value R , that is,

$$\frac{\sum_{t=t_1}^{t_2} S_{t,r(t)}}{\sum_{t=t_1}^{t_2} S_{t,20-r(t)}} = R. \quad [1.5]$$

The value of reduced ENPARF (RENPARF) for birth cohort g can then be defined by the following formula:

$$\text{RENPARF}^{(g)} = \text{ENPARF}^{(g_0)} + (\text{ENPARF}^{(g)} - \text{ENPARF}^{(g_0)}) \cdot k \quad [1.6]$$

where the value of k is such that equation [1.5] holds.

1.3. Data source and calculations

Until 1995, the retirement age in Czechia was 60 for men and 53–57 for women (depending on the number of children raised). Since 1996, the retirement age has been gradually increasing for each subsequent birth cohort: by two months for men, and by four

months and since 2019 by six months for women, until reaching the retirement age for men. In 2030, the retirement age for men will reach 65, and men born in 1965 will retire.

The birth cohort of 1965 was therefore chosen as the reference cohort (with a retirement age of 65). For the birth cohorts from 1966–2025, the values of ENPA, ENPARF and RENPARF together with three variants of “mechanical” development of retirement age were calculated: maintaining the level of 65, increasing by one month for each subsequent birth cohort only until the age of 67 is reached (by present legislation), and increasing by one month for each subsequent birth cohort without limitation. The corresponding AOADR values for the period 2030–2100 were calculated (the values of RENPARF were calculated assuming that the ratio (5) for the period 2030–2101 would be 0.4, which is the approximate AOADR value in 2030).

The data source for the probabilities of dying in the period 1920–2024 was the (period) life tables of the Czech Statistical Office (CZSO [2025](#)). The age-specific fertility rates in 1925–2023 were calculated or directly taken from demographic yearbooks (CZSO [2024](#)), and the data on population by sex and age as of January 1 in 1920–2025 are from the same source.

Values for the period 2024–2100 come from the authors’ projection (in several variants of fertility development), which was a slight adjustment to the last Czech Statistical Office projection (CZSO [2023](#)) reflecting the latest available data from 2023 and 2024, respectively. After 2100, the probabilities of dying (necessary for calculation of cohort life tables) are supposed to be equal to the values of 2100.

The adjusted projection of the Czech population had the following scenarios:

- Base population as of January 1, 2025.
- 2024: total fertility rate (TFR) of 1.37 (preliminary real value). Three variants of future fertility development until 2050 are as follows:
 - low: drop of TFR to 1.25;

- medium: raise of TFR to 1.50;
- high: raise of TFR to 1.75;

since 2050 fertility rates are expected to be unchanged.

- Mortality: medium variant only:
 - 2024: Life expectancy of males and females is 77.15 and 83.14 (real values); until 2100, it increases to 89.4 for males and to 92.6 for females.
- Migration:
 - 2024: net migration is 36,845 (preliminary real value).
 - 2025–2027: gradual return of some Ukrainian refugees: negative net migration of about –31 000 each year; since 2028, baseline positive net migration is 25,000 each year.

Replacement migration by low fertility (preventing population decline): since 2028, there has been a positive net migration of about 64,000 each year.

1.4. Results

The significant decline in female fertility in Czechia, which occurred after the change in the political, economic and social system after 1989, was only partially compensated for in subsequent years. As a result, there has been and will continue to be a decline in the completed fertility of cohorts born in the second half of the 20th century and in this century. While the cohort born in 1965 had a completed fertility of almost 1.94, cohorts born between 1980 and 1985 are expected to only have rates of around 1.62–1.65. The completed fertility of subsequent cohorts will increasingly depend on the development of female fertility after 2024. In the low variant, there will be a gradual decline to 1.25, while in the medium variant, there will be a decline to 1.50. Only in the high variant would the completed fertility of subsequent cohorts gradually increase to 1.75 (see [Figure 1.1](#)).

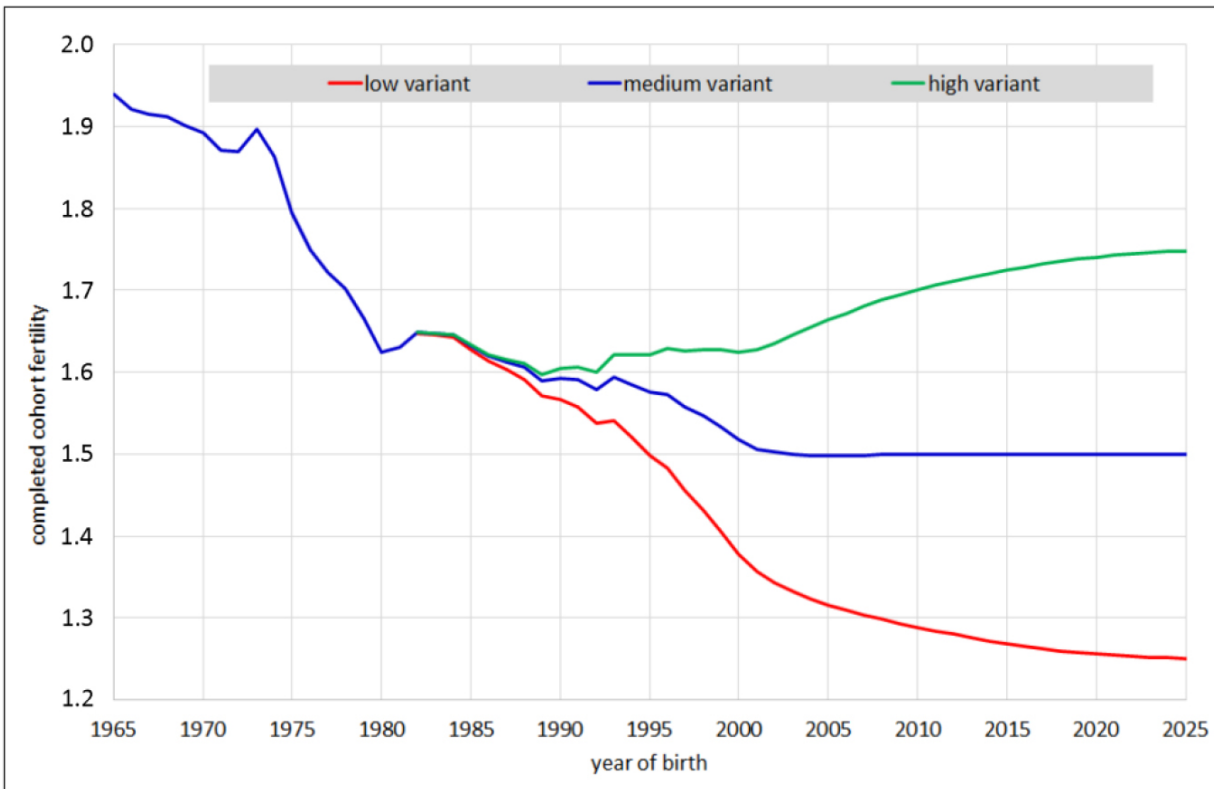


Figure 1.1. Completed cohort fertility for birth cohorts 1965–2025.

Source: authors' calculations and graph

Given this large decline in cohort fertility and the continuing decline in mortality, the ENPARF for cohorts born after 1965 would rise very quickly, with values much higher than those of standard ENPA. People born in 1980 would reach retirement age at around 69, and the ENPARF value for people born in 2000 would be around 71.5 years in the high-fertility variant and around 74 years in the low-fertility variant. For people born in this century, ENPARF values would be relatively stable in the case of a high fertility variant, while with a medium fertility variant, people born in 2025 would have an ENPARF value of 74 years, and with a low fertility variant, as high as 76.5 years. The RENPARF values would be only a few months lower than the values of ENPARF. RENPARF values exceed 72 years in all scenarios except those assuming unrealistically high net migration (more than 60,000 per year). Under such a migration level, RENPARF remains below 72 years, even with low fertility (see [Figure 1.2](#)).

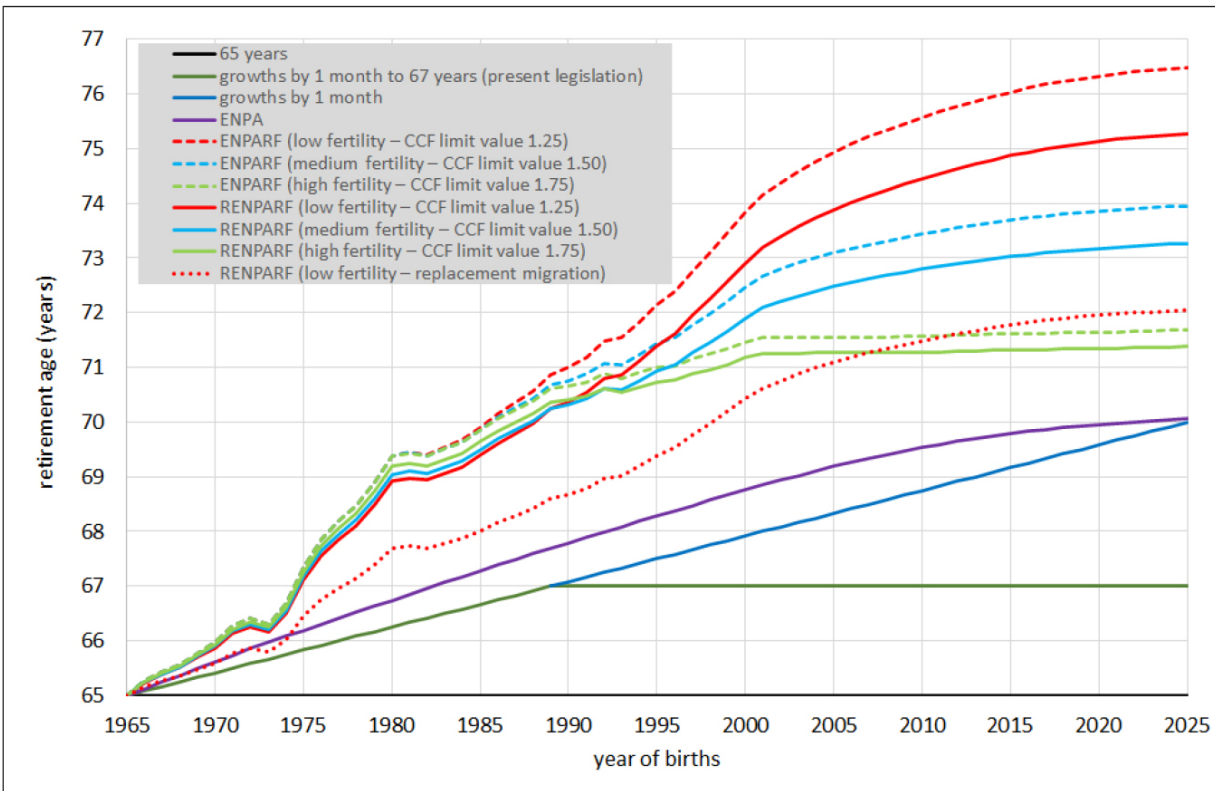


Figure 1.2. *Variants of pension age for birth cohorts 1965–2025.*

Source: authors' calculations and graph

After 2030, only the equitable normal pension age reflecting fertility (RENPARF) – which, however, seems unrealistically high – could ensure the financial stability of the pension system; its unreduced variant (ENPARF) would result in a certain reduction in the financial burden on the pension system, regardless of the fertility development scenario (see [Figures 1.3–1.6](#)). In the low fertility scenario, the standard ENPA would lead to a gradual increase in the AOADR value from less than 0.4 to more than 0.6, which would mean an increase in the number of people of retirement age per 100 people of working age by more than 50%. Similar values would be achieved with a permanent increase in the retirement age by one month for each subsequent birth cohort. If the increase in the retirement age were to stop at 67, the AOADR value would reach almost 0.75 by the end of the century; if the retirement age were to remain at 65, the number of people of retirement age per 100 people of working age would more than double by the end of the century compared to 2030 (see [Figure 1.3](#)).

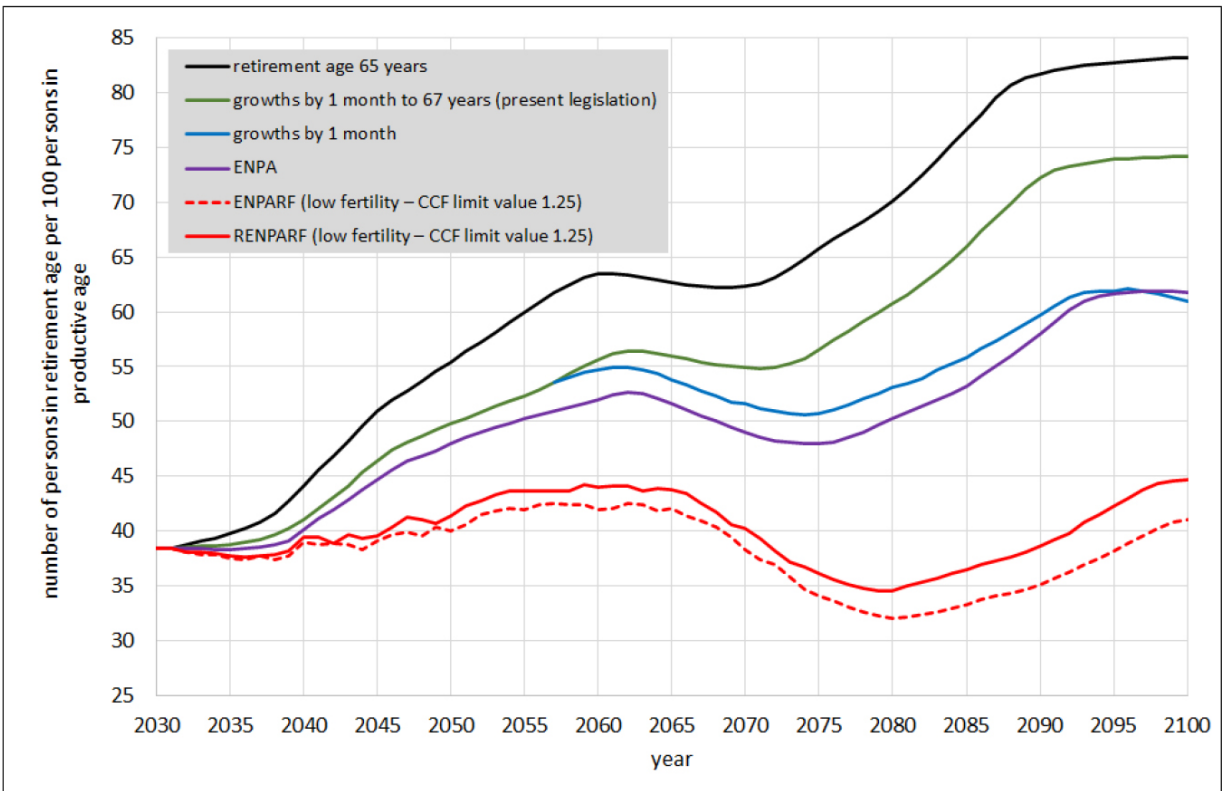


Figure 1.3. *Adjusted old-age dependency ratio 2030–2100 – low fertility variant.*

Source: authors' calculations and graph

Unrealistically high replacement migration would only partially improve the situation in the low fertility scenario, as it is assumed that migrants would remain in Czechia after the end of their economic activity and receive old-age pensions.

The AOADR values would gradually increase to more than 0.5 under ENPA and similarly under a permanent increase in the retirement age by one month for each subsequent year of birth. If the increase in the retirement age were to stop at 67, the AOADR value would approach almost 0.65 by the end of the century, while if the retirement age were to remain at 65, this value would exceed 0.7 (see [Figure 1.4](#)).

Very similar values would be achieved with a medium fertility variant and a more realistic net migration of 25,000 people per year (see [Figure 1.5](#)).

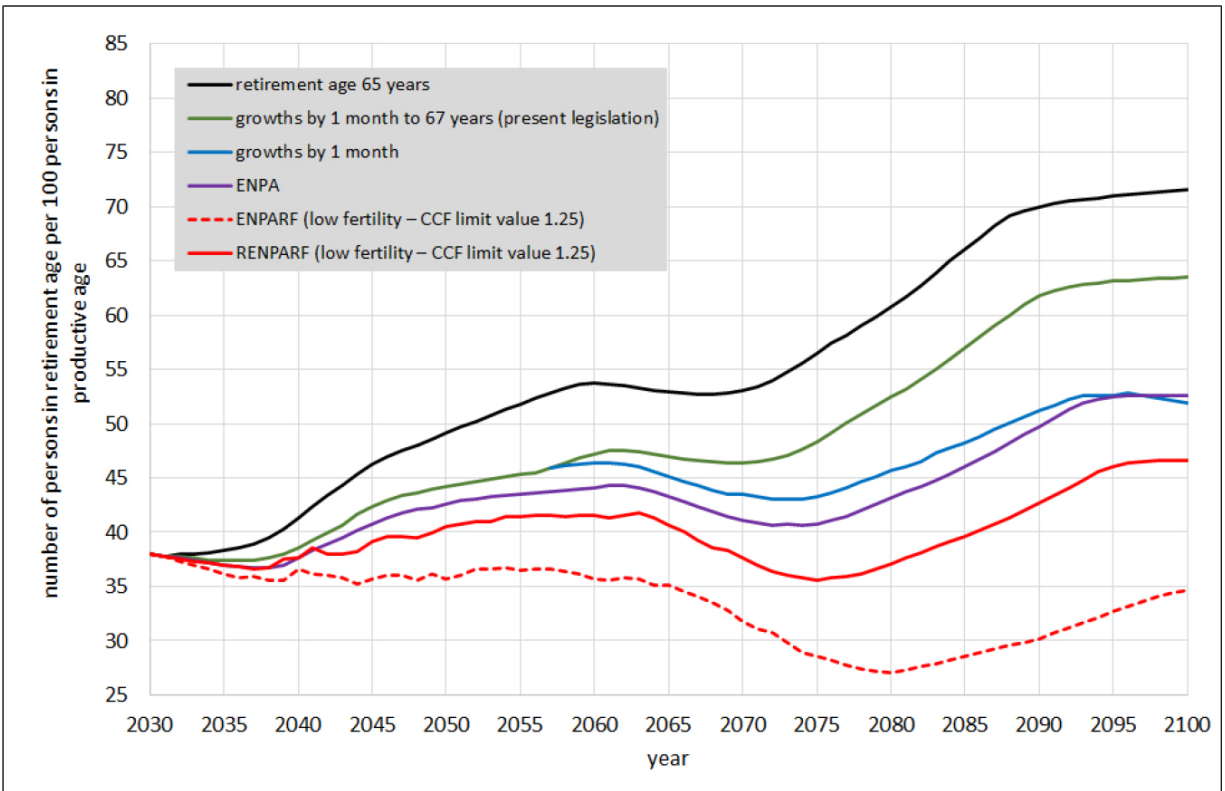


Figure 1.4. Adjusted old-age dependency ratio 2030–2100 – low fertility variant, replacement migration.

Source: authors' calculations and graph

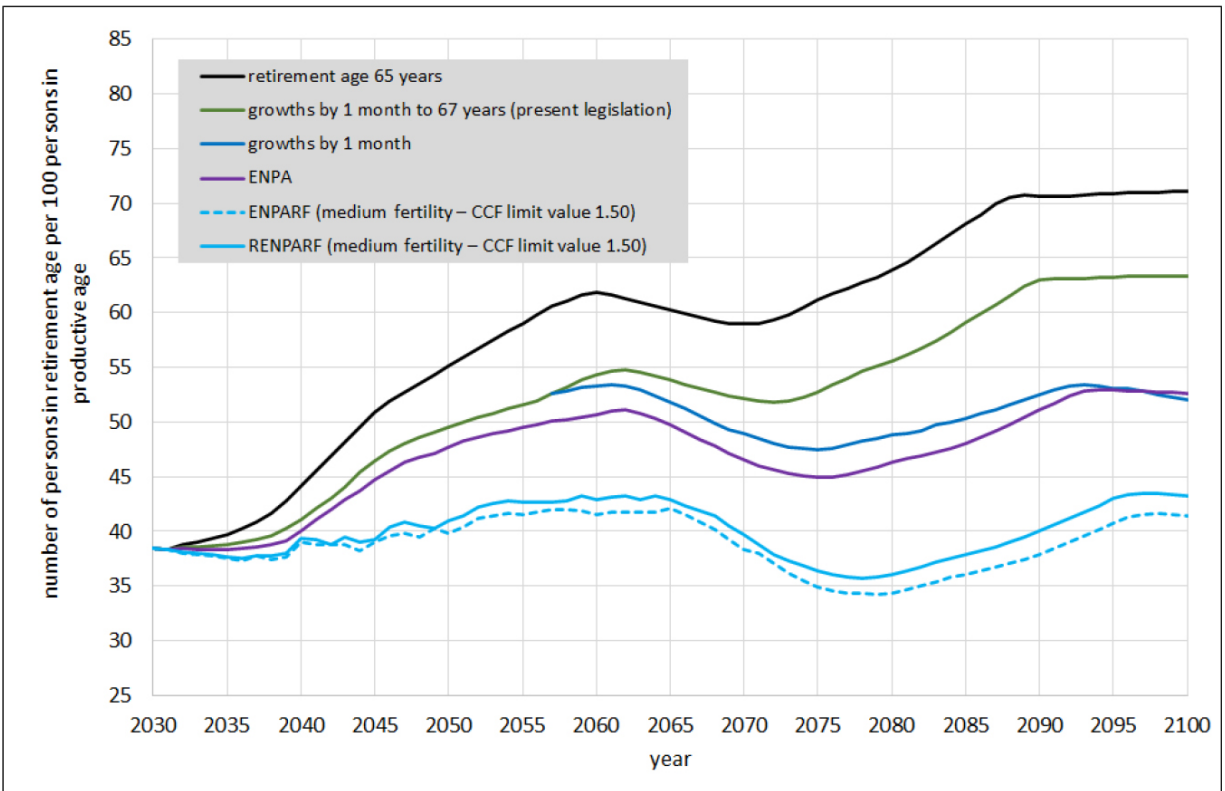


Figure 1.5. *Adjusted old-age dependency ratio 2030–2100 – medium fertility variant.*

Source: authors' calculations and graph

The most significant contribution to ensuring the financial stability of the pension system would be the high fertility variant assuming a gradual increase in the total fertility rate to 1.75 (which, however, is not currently considered very realistic). Under the standard ENPA, AOADR values would rise to 0.5 by 2060 (when the numerous population cohorts born in the mid-1970s will gradually retire) and then fall to around 0.45. A similar trend with slightly higher values should be seen in AOADR when the retirement age is increased by one month for each subsequent birth cohort. If the retirement age were to stop rising at 67, AOADR values at the end of the century would be above 0.55, while if the retirement age were to remain at 65, they would slightly exceed 0.6 (see [Figure 1.6](#)).

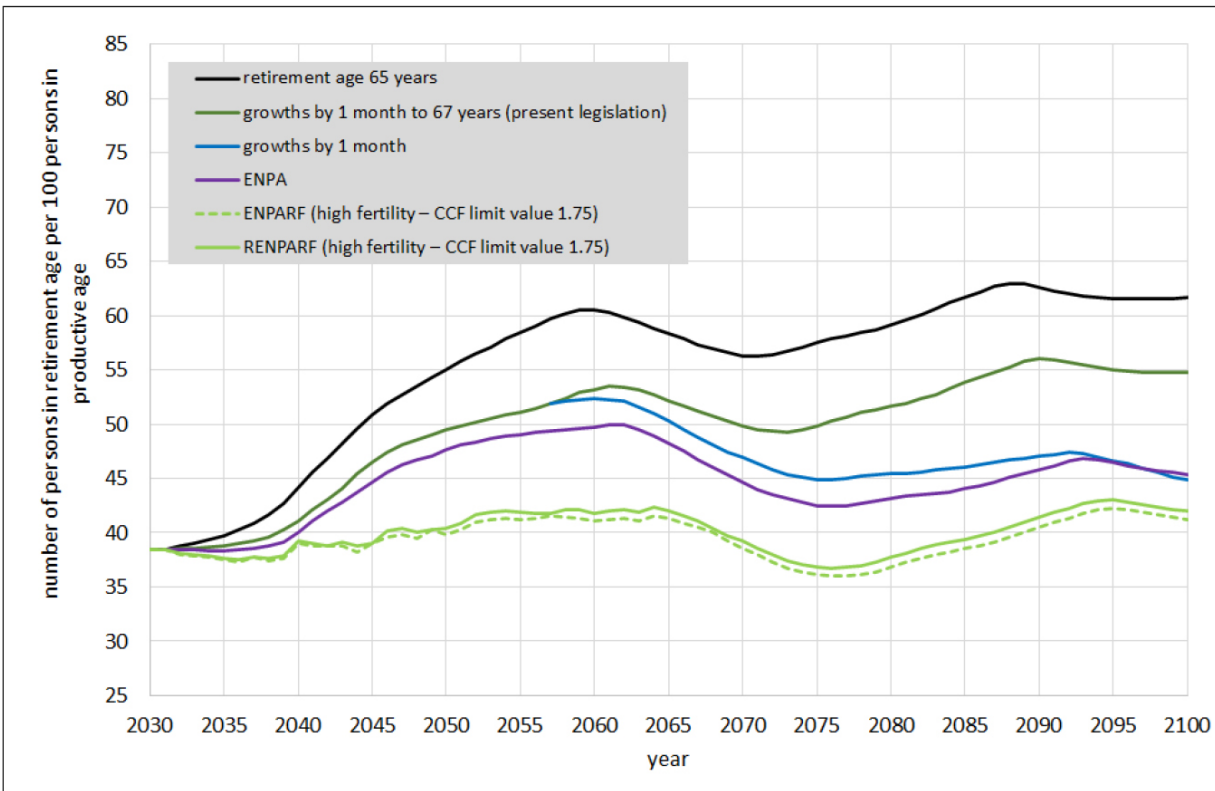


Figure 1.6. *Adjusted old-age dependency ratio 2030–2100 – high fertility variant.*

Source: authors' calculations and graph

1.5. Conclusion

It will not be possible to achieve financial stability in the Czech pension system after 2030 simply by raising the retirement age. Given the expected low fertility, the retirement age would have to rise relatively quickly, with people born after 1990 having to retire at 70 or even later. The question is whether these people would be able to perform their jobs at such an advanced age (especially in certain professions), whether suitable employment would be available for them and whether there would be demand for them in the labor market.

In addition, there is an important objection that it would be unfair to set the retirement age reflecting fertility at the same level for everyone, regardless of the number of children raised. Properly raising children involves, among other things, significant financial

costs and a temporary reduction or complete loss of income for at least one of the parents. Therefore, it would be much fairer if the retirement age and the pension system reflecting fertility were set differently according to the number of children each person has. People with more children could have slightly lower contributions to the pension system and slightly higher pensions than childless people.

It would also be appropriate to look for sources of financing for the pension system other than just income contributions. One option would be to cover part of pensions from taxes, for example, disability and survivor pensions. However, this would only be possible after a tax reform consisting of an increase in tax progression or the introduction of additional taxes, such as environmental taxes.

1.6. References

CZSO (Czech Statistical Office) (2023). Projekce obyvatelstva České republiky (Demographic projection of the Czech Republic) – 2023–2100. ČSÚ, Prague [Online]. Available at: <https://csu.gov.cz/produkty/projekce-obyvatelstva-ceske-republiky-2023-2100>.

CZSO (Czech Statistical Office) (2024). Demografické ročenky (Demographic yearbooks) (pramenná díla) 2024–1919. ČSÚ, Prague [Online]. Available at: https://csu.gov.cz/casova_rada_demografie.

CZSO (Czech Statistical Office) (2025). Public database. Complete life tables. CZSO, Prague. [Online]. Available at: <https://vdb.czso.cz/vdbvo2/faces/en/index.jsf?page=vystup-objekt-parametry&sp=A&skupId=2462&pvokc=&katalog=32592&pvo=DEMD003-CR&z=T>.

Fuchs, V.R. (1984). Though much is taken: Reflections on aging, health, and medical care. *Milbank Memorial Fund Quarterly/Health and Society*, 62(2), 142–166 [Online]. Available at: <http://www.jstor.org/stable/3349821>.

- Ryder, N.B. (1976). Notes on stationary populations. *Population Index*, 41(1), 3–28 [Online]. Available at: <http://www.jstor.org/stable/2734140>.
- Sanderson, W.C. and Scherbov, S. (2005). Average remaining lifetimes can increase as human populations age. *Nature*, 435(7), 811–813. doi: 10.1038/nature03593.
- Sanderson, W.C. and Scherbov, S. (2007). A new perspective on population aging. *Demographic Research*, 16, 27–58. doi: 10.4054/DemRes.2007.16.2.
- Sanderson, W.C. and Scherbov, S. (2013). The characteristics approach to the measurement of population aging. *Population and Development Review*, 39(4), 673–685. doi: 10.1111/j.1728-4457.2013.00633.x.
- Sanderson, W.C. and Scherbov, S. (2014). An easily understood and intergenerationally equitable normal pension age. Interim Report IR-14-020, International Institute for Applied Systems Analysis (IIASA), Laxenburg.
- Šídlo, L., Šprocha B., Ďurček, P. (2020a). A retrospective and prospective view of current and future population ageing in the European Union 28 countries. *Moravian Geographical Reports*, 28(3), 187–207. doi: 10.2478/mgr-2020-0014.
- Šídlo, L., Šprocha B., Ďurček, P. (2020b). Prospective dimension of population ageing and potential use in pension security in the V4 countries. *Ekonomický časopis (Journal of Economics)*, 68(6), 601–621.
- Siegel, J.S. (1993). *A Generation of Change: A Profile of America's Older Population*. Russell Sage Foundation, New York.

Note

Chapter written by Tomáš FIALA, Jitka LANGHAMROVÁ and Jana VRABCOVÁ.

2

Number of Centenarians in Brazil and Selected Countries: Evaluation of the Quality of Census Counting

Debates on data quality have raised questions on the accuracy of demographic data at older ages. Thus, any study on the centenarian population must consider metrics to assess the quality of the data. The main objective of this study was to evaluate the quality of centenarian censuses in Brazilian regions and selected countries between 2000 and 2023. A secondary objective was to review and update the method proposed by Medeiros and Paes. The study applied the original proposal for evaluating the quality of census data on centenarian counts, which suggests an index classified into five categories with weights ranging from high to low quality of the data on the count of centenarians. Three demographic indicators were calculated to generate the index: “centenarians” sex ratio and the ratio of the population aged 100 years and over to the population aged 85 years or over for males and females. The study collected data from the regions of Brazil and from a block of developed and developing countries, according to the United Nations classification. This updated review classification showed good internal consistency. Based on the most recent centenarian census data available for the selected countries, very good and good classifications occurred mainly in countries with high human development index (HDI) and life expectancy at birth. Moderate, low, and very low categories were observed mainly for those with intermediate or low levels in these indicators, including Brazil.

2.1. Introduction

Considering the increasing longevity observed in recent decades, the United Nations Population Division (UNDP) has systematically monitored the growth in the number of centenarians since 1990. This phenomenon is further substantiated by research findings that indicate these individuals constitute one of the most rapidly expanding demographic segments (Robine and Cuabynes [2017](#)).

Understanding these demographic trends is of significant practical importance. The resulting metrics are essential tools for global institutions to project and plan future survival in old age. This phenomenon exerts a direct influence on critical decisions concerning health infrastructure and

expenditure, safeguarding against longevity risks and numerous other pivotal domains.

Nevertheless, debates on data quality have raised questions on the accuracy of demographic data at older ages. This debate has occurred in countries such as Brazil, the most populated country in Latin America, with a population of more than 203 million people in the latest census ([2022](#)).

Hence, the main objective of this study was to evaluate the quality of centenarian censuses in Brazilian regions and selected countries between 2000 and 2024. A secondary objective was to review and update the method proposed by Medeiros and Paes ([2022](#)) to assess the quality of the census data counting of centenarians.

2.2. Methodology

2.2.1. Background

Medeiros and Paes ([2022](#)) developed a method for assessing the quality of centenarian counts in population censuses, which can be applied to any region of the world. Data were collected from the population census of 193 countries based on the latest census available for each one (mostly conducted in 2010 or 2011) through the United Nations Statistics Division – UNdata (United Nations Statistics Division (UNSD) [2025](#)). This method was reviewed and updated in 2023 (Review 1), whose basic data were the most recent estimates of centenarian populations in the same countries, available from UNSD (Medeiros and Paes [2025](#)). It was again reviewed and updated (Review 2) using data from the centenarian counts from the latest available demographic censuses (mostly conducted in 2020 to 2023), with reference to the same data source.

The inclusion of countries in the first and second studies depended on meeting the first three criteria described below. The fourth criterion was added for Review 2:

1. Having specific information on the number of centenarians registered, as well as the population aged 85 or over, for both men and women.
2. Having at least one male and female centenarian registered in the last census.
3. Having at least 100,000 inhabitants.
4. Having political-administrative-economic autonomy, being considered an independent country.

Following the model's original methodology, 85 countries among the 193 members of the United Nations at the time met the first three established criteria.

Three indicators related to the centenarian population were calculated for each selected country: (1) sex ratio (SR) of centenarians (total number of centenarian men divided by total number of centenarian women); (2) ratio of the male population aged 100 or over in relation to the men aged 85 or over, multiplied by 100 [$P_m(100+/85+)$]; and (3) ratio of the female population aged 100 or over in relation to the women aged 85 or over, multiplied by 100 [$P_f(100+/85+)$].

The study investigated the relationship between the human development index (HDI) and life expectancy at birth (e^0) with the three indicators related to the centenarian population described above, still following the original methodology.

The central hypothesis was that the quality of census data related to centenarians and the population aged 85 or over tends to be better in countries with higher levels of development, represented by HDI and e^0 . This association was tested by applying the simple linear regression model, which showed a significant association at the 1% level.

The multivariate cluster analysis was performed based on the five indicators mentioned above for the 85 countries (Ward [2021](#)), classified into four clusters:

- Cluster 1: countries with high HDI and e^0 , representing a “high data quality standard”.
- Clusters 2 and 3: countries with intermediate HDI and e^0 , characterizing an “intermediate data quality standard”.
- Cluster 4: countries with low HDI and e^0 , indicating a “low data quality standard”.

Weights were assigned to each of the cluster categories: Cluster 1 (weight 3), Clusters 2 and 3 (weight 2) and Cluster 4 (weight 1). The final quality of the centenarian count was then classified based on the total score (sum of the weights) using a scale of very good (VG), good (G), moderate (M), low (L) and very low (VL). [Table 2.1](#) summarizes the final classification of the weights.

Table 2.1. *Classification of scores to assess the quality of the census count of centenarians, according to the weights assigned to the selected indicators.*

Source: Medeiros and Paes (2022)

Classification	Sum of indicator weights SR^1, $P_m(100+/85+)^2$, $P_f(100+/85+)^3$
Very good (VG)	9
Good (G)	8 or 7
Moderate (M)	6
Low (L)	5 or 4
Very low (VL)	3

NOTE: ¹SR= sex ratio; ² $P_m(100+/85+)$: ratio of the male population aged 100 years or over to the population aged 85 years and over; ³ $P_f(100+/85+)$: ratio of the female population aged 100 years and over to the population aged 85 years and over.

2.2.2. Review 2

Regarding Review 2, priority was given to data from the most recent national census available for each country (mostly conducted between 2020 and 2023). In cases whose census data were not available for more recent years, the most up-to-date population estimate was considered.

Continuing with the methodological steps and in accordance with the established criteria (including the fourth criterion), the updated data collection resulted in a sample of 87 countries among the 193 countries made available by the UNdata system (UNSD 2025) (Table 2.6). The vast majority of the 85 countries originally collected by Medeiros and Paes (2022) were reselected.

For Brazil, the quality of centenarian data was analyzed considering its five political-administrative regions (North, Northeast, Central-West, Southeast and South). For these regions, and for Brazil as a whole, it used the municipal human development index (MHDI), which adapts the global HDI methodology to the Brazilian context and the availability of national indicators, maintaining the three HDI dimensions (health, education and income) (Institute of Applied Economic Research (IPEA) 2016).

The data were compiled into a database created in Microsoft Excel. Free R Studio version 4.3.3 software was used to implement the calculations of the centenarian indicators, execute the regression models and generate visualizations that helped interpret the results.

2.3. Results of Review 2

Cluster 1 included 29 of the 87 countries selected in Review 2, Clusters 2 and 3 included 44 and Cluster 4 included 14. The study revised the variation ranges of the three indicators among the selected countries, as detailed in [Table 2.1](#), which were very close to those found in the two previous versions.

In Cluster 4, very high values were found for Vanuatu and India, reaching a maximum for the Ivory Coast (16.37 for $P_m(100+/85+)$ and 16.13 for $P_f(100+/85+)$). These values were higher than the maximum limits of the ranges assigned for P_m and P_f in both the original proposal (7.607 and 9.514) and in Review 1 (7.595 and 9.182). Due to the possibility of countries with values close to or even higher than those appearing in other investigations, it was decided not to establish an upper limit to the ranges for these two indicators for countries classified in Cluster 4.

The SR limits ranged from 0.140 to 0.385 to define the reference intervals for the 29 countries in Cluster 1. However, the magnitudes for four of them (Israel, Iceland, Singapore and Malta) differed from those of the other countries. Therefore, they were removed from the sample. In the case of Malta, SR and P_m magnitudes were VL compared to the other countries, with a significant number of centenarian women. For the first three countries, the magnitudes of these same indicators were much higher than those in the other countries.

Some studies (Perls [2017](#); Lozada-Martínez et al. [2024](#)) suggest that the SR among centenarians usually does not exceed 0.25. However, nine of the 29 countries in Cluster 1 exceeded this magnitude (Germany, Australia, the United States, Ireland, Cyprus, Greece, Iceland, Singapore and Israel).

This behavior suggests a possible change in the specific pattern of extreme longevity, whose proportion of women is traditionally more pronounced. The observation of atypical SR values for these countries may indicate a trend of more men reaching very advanced ages in relation to women. This led to the hypothesis that the proportion of male centenarians may be increasing in several parts of the world.

Similarly, the construction of the Cluster 4 intervals required the exclusion of some of its 14 countries, as they had discrepant values in each indicator: Kenya, Zimbabwe, Malawi and Guatemala for SR; Bhutan, Bangladesh, and Guatemala for P_m ; and Bhutan for P_f . The presence of these outliers made it difficult to define representative intervals for most of the other countries, justifying their removal and the need for a specific analysis of centenarian census data.

Table 2.2. *Reference intervals for selected indicators by categories of quality standards for centenarian census data for countries selected according to the Original Proposal and Reviews 1 and 2.*

Source: UNSD (2025); Original proposal: Medeiros and Paes (2022); Review 1: Medeiros and Paes (2025)

Quality of centenarian census	Sex ratio	P_m (100+/85+) ¹	P_f (100+/85+) ²	Weight
High (Cluster 1)				
Original proposal	0.138; 0.353	0.203; 0.762	0.536; 1.405	3
Review 1	0.173; 0.322	0.214; 0.713	0.541; 1.347	
Review 2 ³	0.140; 0.385	0.278; 0.743	0.593; 1.691	
Low (Cluster 4)				
Original proposal	0.588; 1.102	2.410; 7.607	2.952; 9.514	1
Review 1	0.591; 1.103	2.518; 7.595	2.967; 9.182	
Review 2 ³	0.524; 1.122	≥3.076	≥2.936	
Intermediate (Clusters 2 and 3)				
Original proposal	Other ranges	Other ranges	Other ranges	2
Review 1	Other ranges	Other ranges	Other ranges	
Review 2 ³	Other ranges	Other ranges	Other ranges	

NOTE: ¹ $P_m(100+/85+)$: ratio of the male population aged 100 years or over to the population aged 85 years and over; ² $P_f(100+/85+)$: ratio of the female population aged 100 years and over to the population aged 85 years and over. ³Excluded from the sample: Cluster 1 = (SR: Malta, Iceland, Singapore, Cyprus and Israel), (P_m : Malta, Cyprus, Singapore and Israel) and (P_f : Singapore and Israel); Cluster 4 = (SR: Kenya, Zimbabwe, Malawi and Guatemala), (P_m : Bhutan, Bangladesh, Guatemala) and (P_f : Bhutan).

The possibilities that could justify such atypical values may include erroneous data declaration, problems in civil registration, overestimation of age, failures in the system input process, issues related to migration and variations in survival rates. There may also have been age coding errors, a phenomenon

well documented in the literature, which tends to accumulate at advanced ages and artificially inflate the frequency at certain ages (Newman [2018](#); Newman [2024](#)).

The ranges for indicators in Clusters 2 and 3 were specified so that the values were outside the other ranges of Clusters 1 and 4.

The descriptive statistics of the indicators that make up Clusters 1 (high HDI and e^0) and 4 (low HDI and e^0) are presented in [Table 2.3](#) to illustrate the development characteristics (HDI and e^0) and the quality of the selected indicators (SR, P_m and P_f) of the 87 countries of Review 2.

Table 2.3. *Mean age, standard deviation, minimum and maximum values of selected indicators, human development index (HDI) and life expectancy at birth (e^0), of countries grouped in Cluster 1 and Cluster 4, according to Review 2.*

Source: UNSD ([2025](#))

Measure	SR ¹	P _m ²	P _f ³	HDI	e ⁰
Cluster 1					
Mean age	0.206	0.459	1.105	0.939	82.6
SD	0.052	0.128	0.263	0.021	1.1
Minimum	0.140	0.278	0.593	0.890	79.3
Maximum	0.385	0.743	1.691	0.972	84.7
Cluster 4					
Mean age	0.698	7.590	7.892	0.605	67.2
SD	0.199	4.146	3.901	0.084	5.4
Minimum	0.524	3.076	2.936	0.416	55.1
Maximum	1.122	16.370	16.135	0.741	74.7

NOTE: ¹SR = sex ratio; ²P_m(100+/85+): ratio of the male population aged 100 years or over to the population aged 85 years and over; ³P_f(100+/85+): ratio of the female population aged 100 years and over to the population aged 85 years and over. Excluded: Cluster 1 (SR: Malta, Iceland, Singapore, Cyprus and Israel), (P_m: Malta, Cyprus, Singapore and Israel) and (P_f: Singapore and Israel); Cluster 4 (SR: Kenya, Zimbabwe, Malawi and Guatemala), (P_m: Bhutan, Bangladesh and Guatemala) and (P_f: Bhutan).

[Table 2.3](#) shows that the mean age of the SR in the countries that form Cluster 4 was much lower (3.3 times lower) than that of Cluster 1, with a much higher dispersion of values in the latter. Similarly, the mean P_m and P_f were lower in Cluster 1 (16.5 times lower for P_m and 7.1 times lower for P_f) than in Cluster 4.

Likewise, the highest HDI and e^0 values recorded in Cluster 4 (0.741 and 74.7 years, respectively) were significantly lower than the lowest values in Cluster 1 (0.890 and 79.3 years, respectively).

Similar analogies were also observed when analyzing these same indicators for the countries that form Clusters 2 and 3. In other words, their mean values were lower than those of Cluster 1 and higher than those of Cluster 4.

The difference in these indicators highlights a striking distinction between the groups, suggesting that the clusters represent quite distinct realities in terms of survival that influence longevity. This may reflect structural inequalities, such as access to health services, education, basic sanitation and general socioeconomic conditions, which directly impact the social determinants of health and, consequently, the life expectancy of the population.

Census data on centenarians in Brazil and its regions were used as an illustration of the revised methodology (Review 2) based on the last demographic census of 2022 (Brazilian Institute of Demography and Statistics (IBGE) [2022](#)). [Table 2.4](#) details the three indicators, the weights attributed to each one and the final classification of the quality of the census data for Brazil and its regions.

Table 2.4. *Classification of the quality of census data on centenarians in Brazil and its regions, Census [2022](#).*

Source: IPEA ([2016](#)); IBGE ([2022](#))

Region	SR [*]	W ^{**}	P _m ^{***}	W ^{**}	P _f ^{****}	W ^{**}	Sum W	C ^{*****}
Brazil	0.45	2	1.41	2	1.71	2	6	M
N ¹	0.48	2	2.45	2	3.63	1	5	L
NW ²	0.44	2	2.29	2	3.10	1	5	L
SW ³	0.30	3	0.88	2	1.46	3	8	G
S ⁴	0.34	3	0.84	2	1.27	3	8	G
CW ⁵	0.50	2	1.47	2	2.00	2	6	M

NOTE: *SR = sex ratio; **W = weight: assigned according to the reference range to which the calculation value corresponds in [Table 2.1](#); ***P_m(100+/85+): ratio of the male population aged 100 years or over to the population aged 85 years and over; ****P_f(100+/85+): ratio of the female population aged 100 years and over to the population aged 85 years and over; *****C = classification: presented according to the scale proposed by Medeiros and Paes ([2022](#)).

¹N = North; ²NW = Northeast; ³SW = Southeast; ⁴S = South; ⁵CW = Central-West.

The high (Low) values for P_m and P_f in the North and Northeast may indicate an atypical concentration of people aged 100+ within the population aged 85+ in these regions, raising questions about the reliability of the census data

quality for this specific age group. A better (Good) situation was found for the South and Southeast of the country, and a moderate situation was found for the Central-West.

The final classification for Brazil as moderate also suggests that census data on centenarians in these areas may have inconsistencies or problems that require cautious interpretation and use. In this context of potential weaknesses in census data, Reis and Turra ([2026](#)) point to challenges in the quality of these data, noting that the census showed a reduction in the number of centenarians, from 24,576 in 2000 to 24,236 in 2010, attributed to coverage problems and incorrect age declaration. In 2022, this number increased to 37,814 centenarians (IBGE [2022](#)).

[Table 2.5](#) shows the descriptive statistics for the regions of Brazil for SR, P_m , P_f , MHDI, and e^0 in the 2022 census year.

The North and Northeast had mean MHDI and e^0 values below those found for the South and Southeast, while the Central-West was in between them. These positions support the hypothesis that better data quality ([Table 2.4](#)) is related to better levels of global development.

Table 2.5. Mean age, standard deviation, minimum and maximum values of the selected indicators. MHDI and of the regions of Brazil, Census 2022.

Source: IPEA (2016); IBGE (2022)

Measure	SR ¹	P _m (100+/85+) ²	P _f (100+/85+) ³	MHDI	e ⁰
North					
Mean age	99.857	2.616	3.487	0.703	74.16
SD	0.838	0.775	1.207	0.014	1.238
Minimum	98.720	1.149	1.882	0.688	72.62
Maximum	101.29	3.257	5.679	0.731	75.78
Northeast					
Mean age	93.538	2.315	3.030	0.702	74.61
SD	1.799	0.636	0.843	0.020	1.588
Minimum	91.210	1.544	2.173	0.676	72.27
Maximum	96.580	3.548	4.742	0.734	77.10
Southeast					
Mean age	93.190	0.996	1.597	0.778	79.14
SD	2.743	0.299	0.286	0.019	0.907
Minimum	89.390	0.701	1.209	0.762	78.02
Maximum	95.230	1.318	1.835	0.806	80.00
South					
Mean age	95.220	0.861	1.250	0.777	79.79
SD	1.905	0.139	0.108	0.013	1.027
Minimum	93.400	0.700	1.128	0.769	78.96
Maximum	97.200	0.947	1.332	0.792	80.94
Central-West					
Mean age	96.453	1.448	2.075	0.757	77.00
SD	4.198	0.472	0.466	0.037	1.968
Minimum	91.060	0.966	1.530	0.736	75.30
Maximum	101.31	2.086	2.467	0.814	79.70

NOTE: ¹SR = sex ratio; ²P_m(100+/85+): ratio of the male population aged 100 years or over to the population aged 85 years and over; ³P_f(100+/85+): ratio of the female population aged 100 years and over to the population aged 85 years and over.

The study applied the Review 2 reference intervals as a quality assessment standard and defined the weights and their sum to obtain the revised final classification for the 87 selected countries ([Table 2.5](#)). The countries were ranked by continent, from best (VG) to worst quality (VL).

It was identified that some countries had values very close to the limits of the predefined ranges for each category (VL, L, M, G and VG). In these borderline cases, it was decided to assign the highest rating between the two adjacent categories. This methodological decision aimed to ensure that countries with indicators on the threshold of a higher category were not unduly classified at a lower level, reflecting a conservative approach to assessing data quality.

The assessment of the quality of centenarian census data also considered speculations about the increasing proportion of males among centenarians. It considered indications of a greater representation of men in some populations of centenarians with high-quality data (Israel, Iceland, Singapore and Malta). Thus, these countries, whose SR showed increasingly more male centenarians and whose other indicators (P_m and P_f) pointed to good quality, were classified as VG.

Europe and Oceania generally had a higher concentration of countries with VG and G ratings. Of the 40 countries on these continents, only six were categorized as M or L, except Vanuatu, which received a VL rating. This finding may reflect more robust investments in civil registration systems, statistical infrastructure and advanced census methodologies in these countries. In contrast, all African countries were rated as VL or L.

Countries in the Americas and Asia had a mixed panorama, covering all classification categories from VG to VL. This heterogeneity demands specific analyses of each national context, considering historical, cultural, economic and political factors.

Countries with lower data quality ratings (M, L or VL) used estimates as their primary source. This raises questions about the reliability of the data and the need to improve census system projections in these countries.

This concern about the accuracy of demographic data at older ages is corroborated by Newman ([2024](#)), who analyzed survival patterns in old age on a global scale and revealed the presence of significant anomalies in late survival data across all scales and periods. In this context, Kestenbaum ([2021](#)) warns of the risks of using projected data about people extremely advanced in age without questioning and carefully evaluating them, which can lead to erroneous conclusions and ineffective policies.

Table 2.6. *Census data quality rating of centenarians for selected countries, 2010–2023.*

Source: UNSD (2025)

Classification	Africa	Americas	Asia	Europe	Oceania
<i>Very good</i>		Argentina**	Azerbaijan	Albania	Australia
		Canada	Brunei Darussalam	Austria	New Zealand
		United States	Israel***	Belgium	
		Uruguay*	Japan	Cyprus	
			Mongolia	Denmark	
			Singapore***	Estonia	
			South Korea	Finland	
			Tajikistan*	France	
			Thailand	Germany	
			Turkey	Greece	
			Uzbekistan	Iceland***	
				Ireland	
				Italy	
				Luxembourg	
				Malta***	
				Netherlands	
				Norway	
				Poland	
				Portugal	
				Slovenia	
				Spain	
				Sweden	
				Switzerland	
				United Kingdom	
<i>Good</i>		Peru	China**	Belarus	
			Kyrgyzstan	Bulgaria	
				Croatia	

Classification	Africa	Americas	Asia	Europe	Oceania
				Czech Republic	
				Hungary	
				Latvia	
				Romania	
				Russian Federation	
				Slovakia	
<i>Moderate</i>		Brazil	Bhutan	Georgia	
		Chile	Iran	Lithuania	
		Guatemala	Kazakhstan	Serbia	
		Mexico	Philippines		
		Panama*			
<i>Low</i>	Chad	Colombia	Bangladesh	Macedonia	
	Kenya	Ecuador		Montenegro	
	Malawi			Ukraine	
	Zimbabwe				
<i>Very low</i>	Ivory Coast	Cuba	India		Vanuatu
	Rwanda	Dominican Republic	Nepal		
	Senegal	Venezuela	Oman		
	South Africa		Saudi Arabia		

*Data obtained from the National Institute of Statistics, 2025.

**Classified according to the best adjacent level.

****Very good* quality was attributed based on the theory of an increasing proportion of male centenarians.

Furthermore, the application of the methodology to assess the quality of centenarian census data revealed significant challenges in countries with outdated data.

This dependence on estimates negatively impacted the quality of the data assessed by the methodology used in this study. Countries with estimated data tended to overestimate values for the centenarian population, which differed from the projections for the total population.

The loss of data quality in countries with outdated census information or with evident methodological inconsistencies highlights the importance of having recent censuses and rigorous data collection and processing for the demographic analysis of vulnerable populations such as centenarians.

The methodology applied in this study proved to be sensitive to these limitations, highlighting the need for caution in interpreting results based on estimated data or data with methodological flaws.

2.4. Conclusion

Based on the most recent centenarian census data available for the selected countries, it was found that the classification categorized as VG and G occurred mainly in countries with high HDI and e^0 . M, L and VL categories were observed mainly for those with intermediate or low levels in these indicators, including Brazil.

Although the ranking puts Brazil ahead of most developing countries, its position as a moderate still falls short of the quality found in more developed countries, since there are marked regional disparities in centenarian counts.

The results of this study raise the hypothesis of an increasing proportion of male centenarians. Countries with presumably more reliable registries tend to have a higher proportion of men among centenarians. This pattern suggests that the SR at extremely advanced ages may function as a parameter for the accuracy of age registries, which justifies its exploration in future investigations aimed at validating the quality of demographic data.

Despite the limitations (e.g. outdated data and estimates for some countries and the established criteria), the Review 2 classification showed good internal consistency, satisfactory, easy-to-understand and simple application, use of few indicators, and no need for sophisticated computer programs.

Therefore, managers and researchers are advised to exercise caution when using census counts of the centenarian population to plan and evaluate public policies. Thus, this proposal can be a good contribution to analyzing centenarian census counts in any region of the world once the necessary data are available.

2.5. References

Brazilian Institute of Demography and Statistics (IBGE) (2022). *Census 2022*. IBGE, Rio de Janeiro.

Institute of Applied Economic Research (IPEA) (2016). *Human Development in Brazilian Macroeconomic Regions*. PNUD, Brazil.

- Kestenbaum, B. (2021). Mortality of centenarians in the United States. In *Exceptional Lifespans*, Maier, H., Jeune, B., Vaupel, J.W. (eds). Springer, Cham. doi: 10.1007/978-3-030-49970-9_6.
- Lozada-Martínez, I., Marín, J., Castelblanco-Toro, S., Mazenett-Granados, E., Suárez, J., Sarmiento, M., Anaya, J. (2024). Demographics and clinical characteristics of a new population of Centenarians in Colombia. The COOLCEN cohort. *Archives of Gerontology and Geriatrics Plus*, 1(1), 100006. doi: 10.1016/j.aggp.2024.100006.
- Medeiros, J.B. and Paes, N.A. (2022). Classification proposal to evaluate the quality of census data of centenaries. *Latin American Population Magazine*, 16, e202207. doi: 10.31406/relap2022.v16.e202207.
- Medeiros, J.B. and Paes, N.A. (2025). Updating the proposal for assessing the quality of census count of centenaries. In *Quantitative Methods and Data Analysis in Applied Demography – Volume 1: Health, Mortality and Covid-19*, Skiadas, C.H. and Skiadas, C. (eds). Springer, Cham. doi: 10.1007/978-3-031-82275-9.
- Newman, S.J. (2018). Errors as a primary cause of late-life mortality deceleration and plateaus. *PLoS Biology*, 16. doi: 10.1371/journal.pbio.2006776.
- Newman, S.J. (2024). The global pattern of centenarians highlights deep problems in demography. *medRxiv*. doi: 10.1101/2024.09.06.24313170.
- Perls, T.T. (2017). Male centenarians: How and why are they different from their female counterparts? *Journal of the American Geriatrics Society*, 65(9), 1904–1906. doi: 10.1111/jgs.14978.
- Reis, C.S. and Turra, C.M. (2026). Spatial distribution of centenarians in Brazil: An exploratory analysis of data quality from the 2000 and 2010 censuses. *Espinhaço Magazine (UFVJM)*, 5(1), 48–56 [Online]. Available at: <https://revistas.ufvjm.edu.br/revista-espin-haco/article/view/76>.
- Robine, J.M. and Cubaynes, S. (2017). Worldwide demography of centenarians. *Mechanisms of Ageing and Development*, 165, 59–67. doi: 10.1016/j.mad.2017.03.004.
- United Nations Statistics Division (UNSD) (2025). UNdata: Demographic Statistics Database. United Nations [Online]. Available at: <https://data.un.org/Data.aspx?d=POP&f=tableCode%3a22>.
- Ward, J.H. (2021). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236–244. doi: 10.1080/01621459.1963.10500845.

Note

Chapter written by Neir Antunes PAES, Rozane PEREIRA DE SOUSA and Vivian Conceição Alves Leite PEREIRA DO LAGO.

3

Optimal Semi-Markov Model Selection for Pediatric Obstructive Sleep Apnea Patient Data

Sleep is critically important to the health and growth of all people especially in children who are undergoing rapid growth and development. Quick and accurate diagnosis of sleep disorders is of critical importance in pediatric patients so that appropriate treatment can be undertaken. Most approaches to diagnosis model the progression of a child's sleep states as a Markov process. However, we demonstrate using Bayesian model selection that carefully chosen semi-Markov models are superior for modeling the trajectory of a child's sleep. Hence, we argue that such more sophisticated modeling is crucial to improving the health outcomes of pediatric and adult patients.

3.1. Introduction

The sleep cycles that the human body progresses throughout a night of sleep are critical for healthy bodily function, repair and growth. Being able to accurately model and track a patient's sleep patterns allows for the diagnosis of sleep disorders that can negatively affect a patient's physical and mental health. This is especially important for pediatric patients whose bodies are constantly growing and developing during the first 18 years of life.

Medical researchers have identified three main types of sleep being wakefulness, rapid eye-movement (REM) and non-rapid eye-movement (NREM). NREM sleep accounts for about 75–85% of total sleep and REM sleep accounts for 15–25%. However, these proportions change a lot over the first decade of life making modeling of sleep states for children particularly challenging based on how quickly their bodies change. Within NREM sleep, there are three separate states referred to as NREM1, NREM2 and NREM3.

NREM3 is deep sleep, whereas NREM1 and NREM2 are lighter transitional sleep states (Iber et al. [2007](#)).

From a stochastic modeling perspective, we look to model a discrete-time discrete-state-space process where, in particular, we have five sleep states to model: Awake, REM, NREM1, NREM2 and NREM3. Markov models have had a long track record of success in the modeling how the human brain transitions between sleep states during a night of sleep (Geessinck et al. [2018](#); Meng et al. [2025](#); Pietzsch et al. [2015](#); Nowlen and Scheffler [2025](#); Sankar et al. [2020](#); Okubo et al. [2015](#)). Furthermore, the use of hidden Markov models (HMMs) is natural for identifying the (hidden) sleep states based on a patient's biological signals data (Fonseca et al. [2018](#); Sadoughi et al. [2022](#); Rossetto et al. [2023](#)). This includes some of our own recent research into HMMs in functional spaces (Kashlak et al. [2023](#)), which demonstrated strong clustering accuracy in automated labeling of pediatric sleep states.

However, very few works have considered extending from Markov models to semi-Markov models (Yang and Hirsch [1973](#); Tan et al. [2015](#)) in the context of sleep research. Hence, in this work, we undertake a comprehensive Bayesian model comparison study comparing various semi-Markov state duration distributions to see if any can outperform the classic Markov model. The classic Markov model corresponds to the geometric state duration distribution, which necessitates faster state switching that often occurs in practice. Alternatively, the Poisson distribution allows for longer sojourns in a given state, but suffers from the variance being equal to the mean. The negative binomial distribution generalizes both of these distributions for more flexible modeling albeit at the cost of using two parameters instead of one. In this work, we demonstrate that using the negative binomial distribution for modeling state durations significantly outperforms all other duration distributions considered even after accounting for the inclusion of more parameters in this model. Our model selection criteria is the Bayes factor, which has a long history of use in model selection (Kass and Raftery [1995](#)).

In this chapter, section [3.2](#) provides a brief description of the data we considered, which was collected during a medical study conducted at

the University of Alberta hospital. Section 3.3 derives formulae for computing Bayes factors under a variety of duration distributions. Section 3.4 outlines the mathematical form of the semi-Markov model. Lastly, section 3.5 compares the various semi-Markov models tested against our 75 pediatric patients.

Both python and R code can for this research can be found at <https://github.com/matthew-fischer/Research-Summer24> and https://github.com/cachelack/SemiMarkovFitting_BayesFactors, respectively. The data are not publicly available at this time.

3.2. Sleep apnea data overview

The data we considered in this work were collected via a sleep study performed at the University of Alberta hospital in Edmonton, Alberta with the specific intent to diagnose obstructive sleep apnea (OSA). Hence, most of the patients do not exhibit regular healthy sleep patterns. The clinical study was approved by the Health Research Ethics Board of the University of Alberta (Pro00057638). There were 75 patients (58 male and 17 female) ranging from age 2–18 years. A large amount of signals data was collected including electroencephalogram (EEG), electromyogram (EMG), electrocardiogram (ECG), breath rate, heart rate and others. The electroencephalogram and related signals data measure the strength of electrical signals at different parts of the body such as the brain (EEG), muscles (EMG), and heart (ECG). The signals were cleaned and the sleep states were manually labeled by a sleep specialist. The manually assigned labels are what we focus on in this work and not the raw signals.

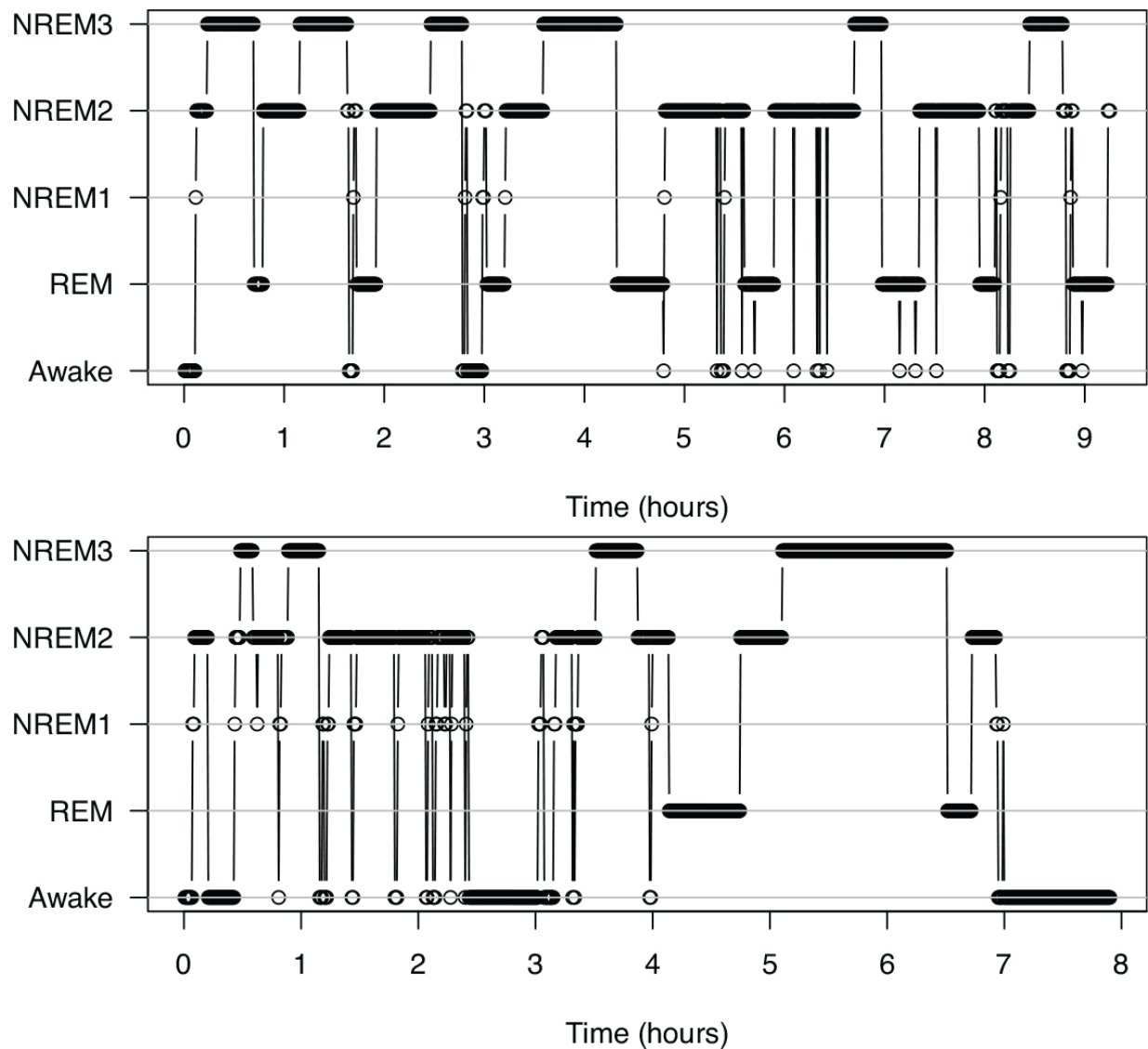


Figure 3.1. *The sleep states for patient CF001 (top) and CF050 (bottom)*

Each patient slept in the sleep lab for a range of 8–10 h. This time frame was partitioned into 30-s *epochs*, and each epoch was assigned one of the five sleep states by a sleep specialist looking at the signals data. Hence, for this chapter, we consider discrete-time discrete-state-space time series of sleep states, the lengths of which differ for each of the 75 patients but range between 960 and 1,200 epochs that are split among the five sleep states. Two examples are displayed in [Figure 3.1](#). Patient CF001 (top plot) begins with a more typical progression between REM and NREM sleep, whereas CF050 has a

very chaotic sleep progression, which begins with a lot of wake ups and then a long period of deep sleep later in the night.

3.3. Bayes factors

In order to determine which duration distribution is best for modeling sleep states via a semi-Markov model, we compute the posterior probabilities of the data given each duration distribution. This is done by imposing the Jeffreys conjugate prior on the unknown parameters for the geometric and Poisson distributions. For the negative binomial distribution, if the number of successes r is assumed known, then we have a standard conjugate beta prior for the parameter p . However, we do not wish to assume that r is known and hence, impose an improper gamma prior on it. Similarly, the logarithmic distribution does not have an easy to use conjugate prior, so we nevertheless impose a beta prior on the unknown parameter p and proceed numerically to compute the desired probabilities. A compendium of non-informative priors can be found in Yang and Berger ([1996](#)). Much detailed information on the following discrete distributions can be found in Johnson et al. ([2005](#)).

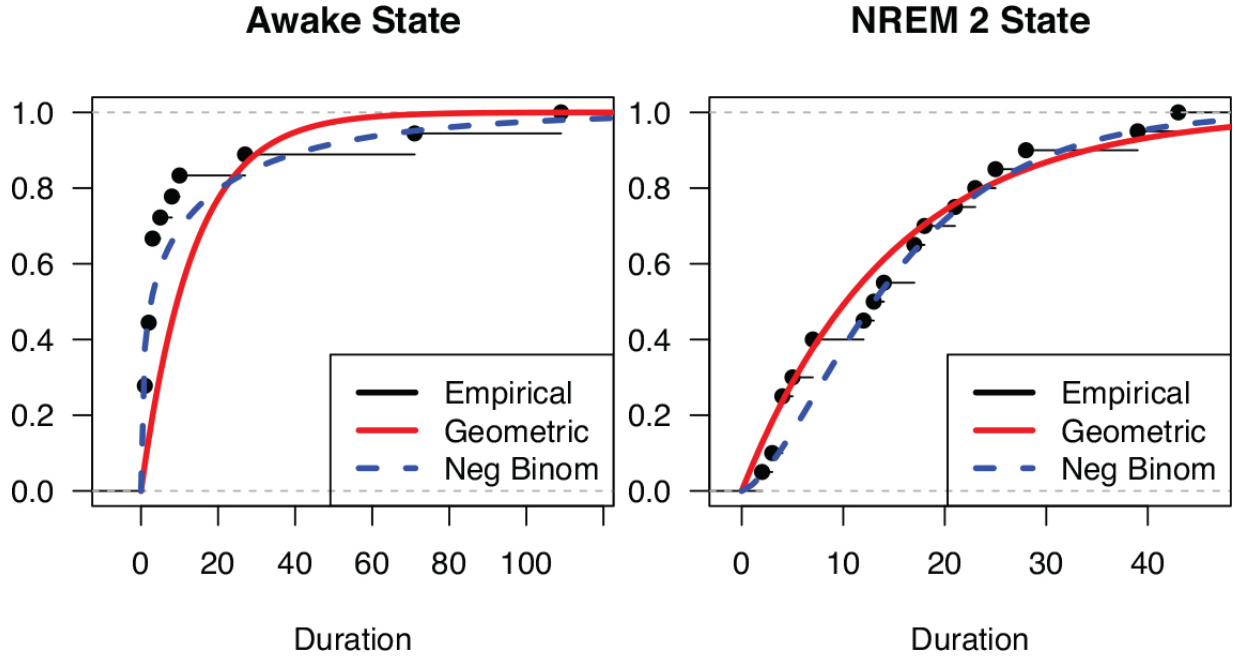


Figure 3.2. *The empirical distributions of the durations for the awake state (left) and the NREM2 state (right) with fitted geometric and negative binomial distributions plotted to compare the fit. These data came from patient CF050.*

To motivate moving beyond the classic Markov model, which corresponds to the geometric duration distribution, we see the empirical distributions of the durations for the awake and NREM2 states for patient CF050 in [Figure 3.2](#). The dashed blue line corresponding to the negative binomial distribution better captures the nuances in the empirical distribution than the geometric distribution does. While the negative binomial is, in fact, a generalization of the geometric distribution, we will see below that nevertheless the modeling performance drastically increases when switching from the geometric to the negative binomial distributions.

3.3.1. Geometric distribution

The probability mass function for the geometric distribution, given a parameter $p \in [0, 1]$ is:

$$f(d; p) = (1 - p)^{d-1} p.$$

We impose a beta distribution on p using the Jeffreys prior $p \sim \text{Beta}(1/2, 1/2)$. Given observed durations $d_1, \dots, d_n$, we update the prior to get $p|d_1, \dots, d_n \sim \text{Beta}(1/2 + n, 1/2 + \sum_{i=1}^n d_i - n)$. Therefore, the posterior probability of the durations $d_1, \dots, d_n$ under the geometric distribution with a beta prior is:

$$\begin{aligned} P(d_1, \dots, d_n | M_{\text{geo}}) &= E_p \left[\prod_{i=1}^n (1-p)^{d_i-1} p \right] \\ &= \int_0^1 (1-p)^{\sum d_i - n} p^n \frac{(1-p)^{1/2 + \sum d_i - n - 1} p^{1/2 + n - 1}}{B(1/2 + n, 1/2 + \sum d_i - n)} dp \\ &= \frac{B(1/2 + 2n, 1/2 + 2 \sum d_i - 2n)}{B(1/2 + n, 1/2 + \sum d_i - n)} \end{aligned}$$

3.3.2. Poisson distribution

The probability mass function for the Poisson distribution, given a parameter $\lambda > 0$ is:

$$f(d; \lambda) = \frac{\lambda^d e^{-\lambda}}{(d-1)!}.$$

We impose a (improper) gamma distribution on λ using the Jeffreys prior $\lambda \sim \text{Gamma}(1/2, 0)$, which is proportional to $\lambda^{-1/2}$. Given observed durations $d_1, \dots, d_n$, we update the prior to get the posterior $p|d_1, \dots, d_n \sim \text{Gamma}(1/2 + \sum_{i=1}^n d_i - n, n)$. Therefore, the posterior probability of the durations $d_1, \dots, d_n$ under the Poisson distribution with a gamma prior is:

$$\begin{aligned}
P(d_1, \dots, d_n | M_{\text{pois}}) &= E_\lambda \left[\prod_{i=1}^n \frac{\lambda^{d_i-1} e^{-\lambda}}{(d_i-1)!} \right] \\
&= \int_0^\infty \frac{\lambda^{\sum d_i - n} e^{-n\lambda}}{\prod (d_i-1)!} \frac{n^{1/2 + \sum d_i - n}}{\Gamma(1/2 + \sum d_i - n)} \lambda^{1/2 + \sum d_i - n - 1} e^{-n\lambda} d\lambda \\
&= \frac{n^{1/2 + \sum d_i - n} \Gamma(1/2 + 2 \sum d_n - 2n)}{\prod (d_i-1)! \Gamma(1/2 + \sum d_i - n) (2n)^{1/2 + 2 \sum d_n - 2n}}
\end{aligned}$$

3.3.3. Negative binomial distribution

The probability mass function for the negative binomial distribution, given parameters $p \in [0, 1]$ and $r > 0$ is:

$$f(d; p, r) = \binom{d+r-2}{d-1} (1-p)^{d-1} p^r.$$

If r is known, then the conjugate prior for p is the beta distribution with Jeffreys prior Beta $(1, 1/2)$. However, there is no tractable conjugate prior for the parameter tuple (r, p) . So we proceed numerically by imposing the prior $p \sim \text{Beta}(1, 1/2)$ and a (improper) Gamma prior $r \sim \text{Gamma}(1/2, 0)$, which was similarly used for the Poisson distribution above and is proportional to $r^{-1/2}$. The posterior density for (r, p) given the durations $d_1, \dots, d_n$ is proportional to:

$$r^{-1/2} p^{nr} (1-p)^{\sum d_i - n - 1/2} \prod_{i=1}^n \binom{d_i + r - 2}{d_i - 1}.$$

Therefore, the posterior probability of the durations $d_1, \dots, d_n$ under the negative binomial distribution with the above priors is:

$$\begin{aligned}
P(d_1, \dots, d_n | M_{\text{nb}}) &= E_{r,p} \left[\prod_{i=1}^n \binom{d_i + r - 2}{d_i - 1} (1-p)^{d_i-1} p^r \right] \\
&= \frac{1}{C} \int_0^1 \int_0^\infty \left[\prod_{i=1}^n \binom{d_i + r - 2}{d_i - 1} (1-p)^{d_i-1} p^r \right] \times \\
&\quad r^{-1/2} p^{nr} (1-p)^{\sum d_i - n - 1/2} \prod_{i=1}^n \binom{d_i + r - 2}{d_i - 1} dr dp \\
&= \frac{\iint_{[0,1]^2} \left[\prod_{i=1}^n \binom{d_i - \log u - 2}{d_i - 1} \right]^2 (-\log u)^{-1/2} p^{2n \log u} (1-p)^{2 \sum d_i - 2n - \frac{1}{2}} \frac{du}{u} dp}{\iint_{[0,1]^2} \left[\prod_{i=1}^n \binom{d_i - \log u - 2}{d_i - 1} \right] (-\log u)^{-1/2} p^{n \log u} (1-p)^{\sum d_i - n - \frac{1}{2}} \frac{du}{u} dp}
\end{aligned}$$

We substitute $r = -\log u$, so that these integrals can be estimated via a Riemann sum on $[0, 1] \times [0, 1]$.

The negative binomial distribution can be considered as an overdispersed generalization of the Poisson distribution, that is, the variance is larger than the mean. Hence, in our model fitting, we compute $\bar{d}$, the sample mean, and s^2 , the sample variance, of the durations. If $\bar{d} < s^2$, then we numerically compute the above posterior probability for the negative binomial distribution. If instead $\bar{d} \geq s^2$, then we revert back to the Poisson distribution from the previous section.

3.3.4. Logarithmic distribution

The probability mass function for the logarithmic distribution, given a parameter $p \in [0, 1]$, is:

$$f(d; p) = \frac{1}{-\log(1-p)} \frac{p^d}{d}.$$

There is no obvious conjugate prior for this distribution, so we choose the standard arcsine distribution as a prior on p being $p \sim \text{Beta}(1/2, 1/2)$. Given observed durations $d_1, \dots, d_n$, we update the prior to get a posterior distribution proportional to $p^{\sum d_i - 1/2} (1 - p)^{-1/2} [-\log(1 - p)]^{-n}$. Therefore, the posterior probability of the durations $d_1, \dots, d_n$ under the logarithmic distribution with a beta prior is:

$$\begin{aligned} P(d_1, \dots, d_n | M_{\log}) &= E_p \left[\prod_{i=1}^n \frac{1}{-\log(1 - p)} \frac{p^{d_i}}{d_i} \right] \\ &= \frac{1}{C} \int_0^1 \frac{1}{[-\log(1 - p)]^n} \frac{p^{\sum d_i}}{\prod d_i} \frac{p^{\sum d_i - 1/2} (1 - p)^{-1/2}}{[-\log(1 - p)]^n} dp \\ &= \frac{\int_0^1 \frac{p^{2 \sum d_i - 1/2} (1 - p)^{-1/2}}{[\log(1 - p)]^{2n}} dp}{[\prod d_i] \int_0^1 \frac{p^{\sum d_i - 1/2} (1 - p)^{-1/2}}{[-\log(1 - p)]^n} dp}. \end{aligned}$$

In practice, these integrals are estimated via a Riemann sum on $[0, 1]$. The arcsine prior ensures that the integrand does not diverge as p approaches either 0 or 1.

3.4. Semi-Markov models for sleep states

The sleep sequence data are best represented using a *run length encoding*. That is, for each of the 75 patients, we have a sequence of distinct sleep states $s_1, \dots, s_n$ taking values in $\{1, \dots, 5\}$ for awake, REM, NREM1, NREM2, NREM3, respectively, and a sequence of integer durations $d_1, \dots, d_n$ indicating how many epochs the patient remained in a given state before switching. That is, d_i epochs were spent in state s_i before transitioning to a new state.

The semi-Markov transition matrix is, in our case, a 5×5 stochastic matrix with zeros on the diagonal, which determines how a given

state transitions into a new distinct state. Once the process enters a new state, the duration distribution, such as those discussed in section 3.3, determines how long the process persists in its current state. We note that each state of the semi-Markov model can have its own parameters for the state durations. Given transition probabilities p_{ij} for transitioning from state i to state j and a duration distribution $f_i(d) = P(d \mid \text{state } i)$, the probability of a given state sequence is:

$$P((d_1, \dots, d_n), (s_1, \dots, s_n)) = \prod_{t=1}^{n-1} [f_{s_t}(d_t) p_{s_t, s_{t+1}}]$$

In general, we would require an initial state probability in the above equation. However, in our study, the initial state is always fixed to be state 1 = awake. Similarly, a general semi-Markov model has a censoring problem with the final state as it is unknown how long that state would persist. However, in our data, the final state is always state 1 = awake when the study ends. Thus, we ignore the final state in the above probability calculation.

For the sake of this research, the transition probabilities can be ignored as they are independent of the state duration probabilities, which is the main focus of this work. When we compute Bayes factors to compare different semi-Markov models, the transition probabilities will cancel out. Ultimately, the log Bayes factor will be used as a model selection criterion for determining which semi-Markov model best fits our sleep states data. It takes on the following compact form:

$$\log \text{BF}(M_1, M_2) = \sum_{s=1}^5 \log P(\{d_i\}_{s_i=s} \mid M_1) - \sum_{s=1}^5 \log P(\{d_i\}_{s_i=s} \mid M_2)$$

where the probabilities are computed as discussed above in section 3.3.

3.5. Sleep data model comparison

To compare the collection of semi-Markov models, we compute the log Bayes factors as per the previous section over all 75 patients' data. The boxplots in [Figure 3.3](#) display the results for the negative binomial distribution and the logarithmic distribution as compared to the classic Markov model, that is, the geometric distribution.

The logarithmic distribution is seen to have comparable performance to the geometric distribution. In contrast, the negative binomial distribution has significantly stronger Bayes factors overall. Using the classic criterion of Jeffreys, 96% of the patients scored a $\log_{10}$ Bayes factor greater than 1 indicating strong evidence in favor of the negative binomial distribution over the geometric distribution. Going further, 89.3% of the patients scored above a 2 indicating very strong evidence in favor of the negative binomial duration distribution.

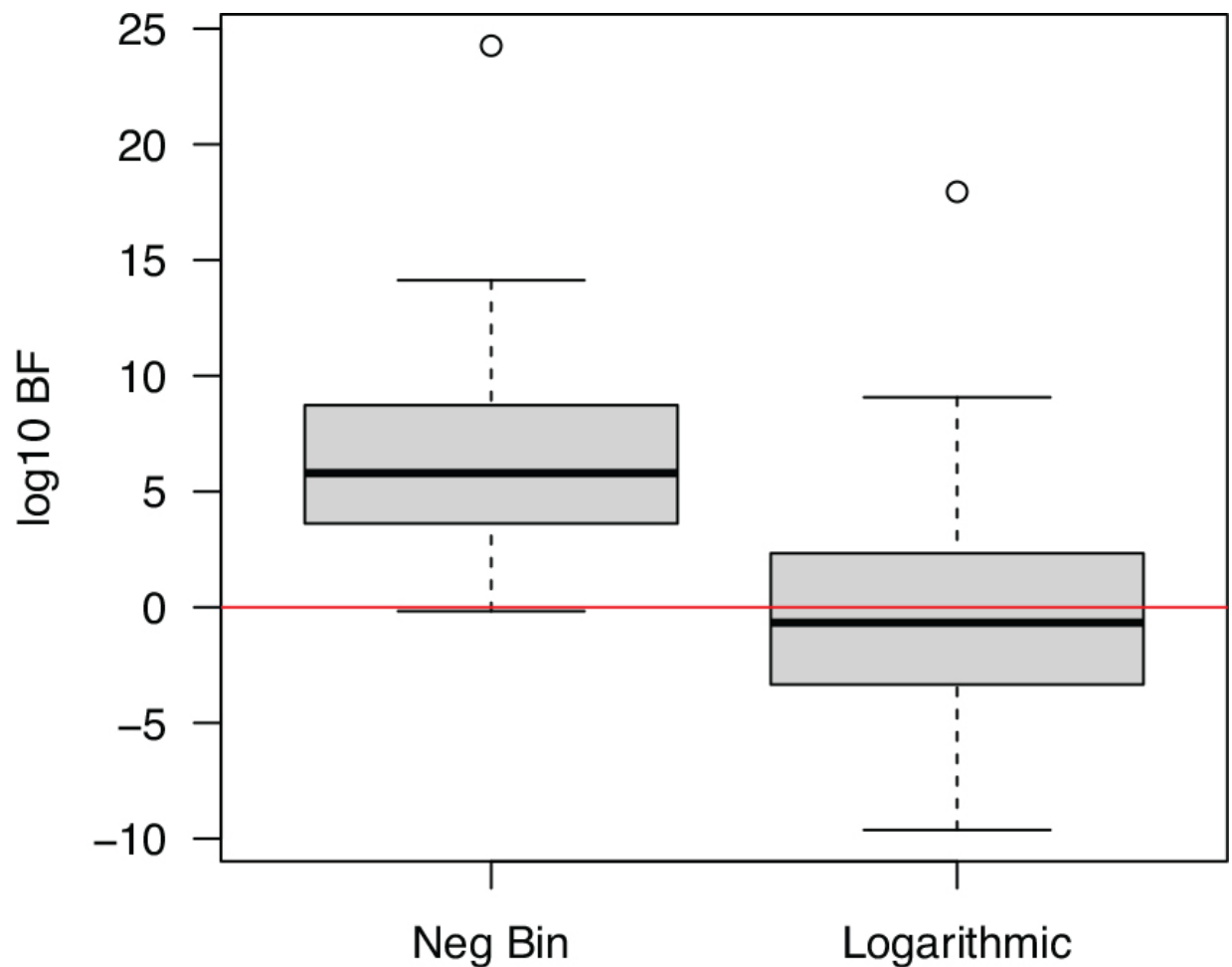


Figure 3.3. *Boxplots of the 75 Bayes factors comparing the negative binomial (left) and the logarithmic (right) to the geometric distribution*

[Figure 3.4](#) displays the posterior probabilities under the geometric distribution plotted against the posterior probabilities under the negative binomial distribution. Barring a few outliers, the majority of the patients are scoring similarly under both models with a noticeable outlier. The patient on the far right had the largest Bayes factor in favor of the negative binomial, but overall did not fit either model overly well.

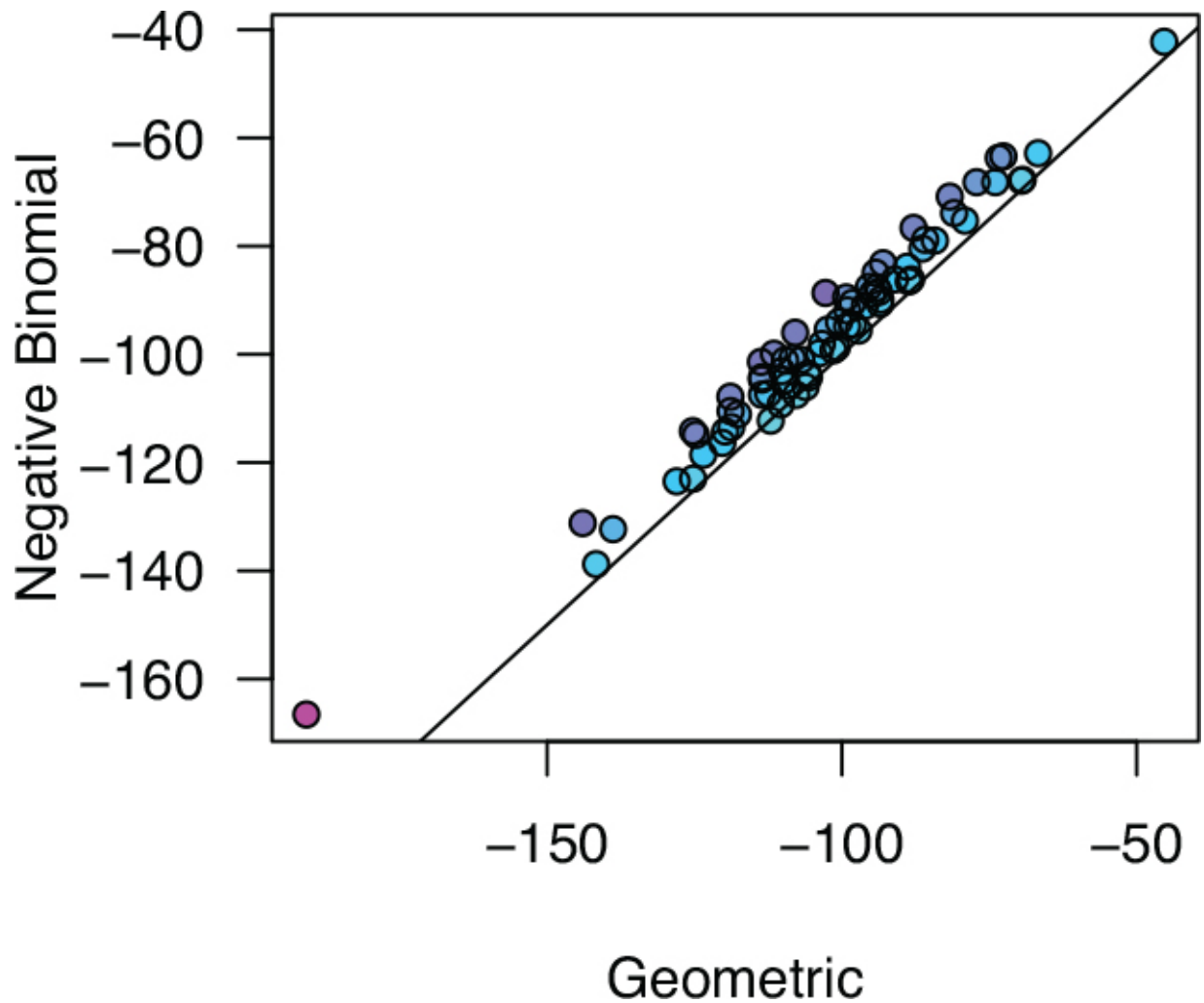


Figure 3.4. *A scatter plot of the posterior probabilities of the 75 patients with geometric on the horizontal axis and negative binomial on the vertical axis.*

The Poisson distribution was also considered in this analysis. However, it scored extremely poorly indicating that a solely Poisson semi-Markov model is not a good fit to our pediatric sleep state data. Nevertheless, the Poisson distribution still is utilized in our negative binomial semi-Markov model when there is no evidence of overdispersion. Hence, the Poisson distribution is still useful in certain sub-cases of our negative binomial semi-Markov model.

3.6. Conclusion

Markov models have had wide application and many successes in the modeling of sleep states in hospital patients. While the semi-Markov model has been in existence at least since the 1970s and has been applied to sleep state data, it is still mostly ignored in the medical community. In this work, we demonstrated that the negative binomial semi-Markov model has overall superior performance in modeling the transitions between sleep states for our cohort of 75 pediatric patients.

Going further, most sleep data are not initially labeled. Instead, long time series of signals are collected that are then manually or algorithmically labeled as discrete sleep states. One of the main tools for achieving this is the HMM, which has been used in countless works including our own (Kashlak et al. [2023](#)). This chapter motivates the reinvestigation of such work on HMMs with the hidden semi-Markov model (HSMM) in its place. This will inevitably lead to more complex fitting algorithms and modeling choices. We predict that such HSMM models will have superior performance in many future investigations into sleep states and other stochastic processes.

3.7. References

- Fonseca, P., den Teuling, N., Long, X., Aarts, R.M. (2018). A comparison of probabilistic classifiers for sleep stage classification. *Physiol Meas*, 39(5), 055001. doi: 10.1088/1361-6579/aabbc2.
- Geessinck, F.A.J., Pleijhuis, R.G., Mentink, R.J., van der Palen, J., Koffijberg, H. (2018). Cost-effectiveness analysis of the DiagnOSAS screening tool compared with polysomnography diagnosis in Dutch Primary Care. *J Clin Sleep Med.*, 14(6), 1005–1015. doi: 10.5664/jcsm.7170.
- Iber, C., Ancoli-Israel, S., Chesson, A., Quan, S.F. (2007). *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. American Academy of Sleep Medicine, Darien, IL.

- Johnson, N.L., Kemp, A.W., Kotz, S. (2005). *Univariate Discrete Distributions*. John Wiley & Sons, Hoboken, NJ.
- Kashlak, A.B., Loliencar, P., Heo, G. (2023). Topological hidden Markov models. *Journal of Machine Learning Research*, 24(340), 1–49.
- Kass, R.E. and Raftery, A.E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773–795.
- Meng, F., Ang, G.Y., Chang, R.R.Y., Lee, C.P., Tan, K.B., Abisheganaden, J.A. (2025). Cost-effectiveness analysis of continuous positive airway pressure treatment for obstructive sleep apnea in Singapore from a health system perspective. *J Sleep Res.*, 34(2), e14326. doi: 10.1111/jsr.14326.
- Nowlen, F. and Scheffler, P. (2025). Cost-effectiveness of hypoglossal nerve stimulation for pediatric severe obstructive sleep apnea in Down syndrome patients. *Otolaryngol Head Neck Surg*, 172(2), 661–667. doi: 10.1002/ohn.1030.
- Okubo, R., Kondo, M., Hoshi, S.L., Yamagata, K. (2015). Cost-effectiveness of obstructive sleep apnea screening for patients with diabetes or chronic kidney disease. *Sleep Breath*, 19(3), 1081–1092. doi: 10.1007/s11325-015-1134-x.
- Pietzsch, J.B., Liu, S., Garner, A.M., Kezirian, E.J., Strollo, P.J. (2015). Long-term cost-effectiveness of upper airway stimulation for the treatment of obstructive sleep apnea: A model-based projection based on the STAR trial. *Sleep*, 38(5), 735–744. doi: 10.5665/sleep.4666.
- Rossetto, A., Midelet, A., Baillieul, S., Tamisier, R., Borel, J.C., Prigent, A., Bailly, S., Pépin, J.L. (2023). Factors associated with residual apnea-hypopnea index variability during CPAP treatment. *Chest*, 163(5), 1258–1265. doi: 10.1016/j.chest.2022.12.048.
- Sadoughi, A., Shamsollahi, M.B., Fatemizadeh, E. (2022). Automatic detection of respiratory events during sleep from

polysomnography data using layered hidden Markov model. *Physiol Meas*, 43(1). doi: 10.1088/1361-6579/ac45e1.

Sankar, A., Dixon, P.R., Sivanathan, L., Memtsoudis, S.G., de Almeida, J.R., Singh, M. (2020). Cost-effectiveness analysis of preoperative screening strategies for obstructive sleep apnea among patients undergoing elective inpatient surgery. *Anesthesiology*, 133(4), 787–800. doi: 10.1097/ALN.0000000000003429.

Tan, K.B., Toh, S.T., Guilleminault, C., Holty, J.E. (2015). A cost-effectiveness analysis of surgery for middle-aged men with severe obstructive sleep apnea intolerant of CPAP. *J Clin Sleep Med*, 11(5), 525–535. doi: 10.5664/jcsm.4696.

Yang, R. and Berger, J.O. (1996). *A Catalog of Noninformative Priors (Volume 1018)*. Institute of Statistics and Decision Sciences, Duke University, Durham, NC.

Yang, M.C.K. and Hursch, C.J. (1973). The use of a semi-Markov model for describing sleep patterns. *Biometrics*, 29(4), 667–676.

Note

Chapter written by Matthew FISCHER, Adam B. KASHLAK, Mathieu CHALIFOUR and Giseon HEO.

4

An Efficient Determination Criterion-based Metric for Z-DNA Pattern Identification in the Epstein-Barr Virus Associated with Nasopharyngeal Carcinoma

In this chapter, we apply a metric based on the efficient determination criterion (EDC) (García et al. [2024](#)) to investigate phenomena associated with Z-DNA occurrences, such as those observed in cancer (Bozeman et al. [2025](#)). The proposed metric allows for the comparison of DNA sequences by analyzing their transition probabilities related to specific states, enabling precise identification of similarities and differences in their underlying stochastic structures (see also García et al. ([2018](#))). We use this approach to analyze three Epstein–Barr virus (EBV) DNA sequences in FASTA format: B95-8, GD2 and HKNPC1. Building on previous studies, specific nucleotide pairs, including *gc*, *ca*, *at*, *cg* and *tg*, have been identified as exhibiting patterns suggestive of B-Z DNA conformational transitions (Shin et al. [2016](#)).

Our analysis confirms that, for the nucleotide pairs *cg*, *tg* and *ca*, the carcinoma-derived EBV strains GD2 and HKNPC1 differ significantly from the reference strain B95-8. Given that *cg* and *tg* are known (Kim et al. [2021](#)) to participate in B-DNA to Z-DNA transitions, these findings provide evidence that such conformational changes may contribute to the observed differences.

4.1. Introduction

We analyze the similarities and differences between an EBV sequence (see Baer et al. ([1984](#))) and two additional EBV-derived sequences associated with nasopharyngeal carcinoma (see Liu et al. ([2011](#)) and Kwok et al. ([2012](#))). All three sequences are well known within the EBV research community and are provided in FASTA format,

allowing for convenient analysis within a stochastic process framework.

To achieve this, we employ metrics based on the Efficient Determination Criterion (Zhao et al. [2001](#)), treating the FASTA-formatted sequences as samples derived from stochastic processes. These metrics are designed to identify structural similarities and divergences in the conditional behavior of the sequences, specifically given a state of interest.

Formally, let $(Z_t)_{t \in \mathbb{N}}$ be a discrete time order o Markov chain on a finite and discrete alphabet $\mathcal{E}$, with memory o such that $o < \infty$. Let us call $\mathcal{E}^o = \mathcal{E} \times \dots \times \mathcal{E}$ the state space on which the process operates. The following notation allows the representation of realizations or samples of the process $(Z_t)_{t \in \mathbb{N}}$. Denote the concatenation $z_k z_{k+1} \dots z_m$ by z_k^m , where $z_i \in \mathcal{E}$, $k \leq i \leq m$. For example, in the case of sequences in format FASTA, the alphabet is $\mathcal{E} = \{a, c, g, t\}$ composed of the nucleotide bases of the genetic alphabet. If the order of the process is $o = 2$, the state space is the cartesian product $\mathcal{E}^o = \{a, c, g, t\} \times \{a, c, g, t\}$, and each state $s \in \mathcal{E}^o$ can be expressed by the concatenation of two elements coming from $\{a, c, g, t\}$, for instance, $s = at$ is one case.

Once the value of o is known (or estimated), a stochastic process is fully defined by its transition probabilities. For each $a \in \mathcal{E}$ and $s \in \mathcal{E}^o$, the transition probability from the state s to a is:

$$P(a|s) = \text{Prob}(Z_t = a | Z_{t-o}^{t-1} = s). \quad [4.1]$$

Next, the collection $\{P(a|s) : a \in \mathcal{E}, s \in \mathcal{E}^o\}$ defines the process $(Z_t)_{t \in \mathbb{N}}$. To estimate these probabilities, it is necessary to consider a sample of the process, a procedure that we describe as follows.

In a given sample $z_1^n = z_1 z_2 \dots z_n$, of size n , coming from the stochastic process $(Z_t)_{t \in \mathbb{N}}$, given the state $s \in \mathcal{E}^o$ and the element of the alphabet $a \in \mathcal{E}$, we denote the number of occurrences of s followed by a in the sample z_1^n by:

$$N_n(s, a) = |\{t : o < t \leq n, z_{t-o}^{t-1} = s, z_t = a\}|$$

and

$$N_n(s) = \sum_{a \in \mathcal{E}} N_n(s, a)$$

which is the number of occurrences of s in the sample z_1^n . Then, for each $a \in \mathcal{E}$ and $s \in \mathcal{E}^o$, the estimation of the transition probability (equation [4.1]) is given by the likelihood estimator:

$$\hat{P}(a|s) = \frac{N_n(s, a)}{N_n(s)}, \text{ when } N_n(s) > 0. \quad [4.2]$$

In the following subsection, we describe the phenomenon we aim to investigate, employing tools derived from stochastic processes.

4.1.1. Z-DNA phenomenon

In this chapter, we analyze three viral sequences in FASTA format, originally included in the work of Bozeman et al. (2025) and obtained from <https://www.ncbi.nlm.nih.gov/>. One of the primary objectives of the study in Bozeman et al. (2025) was to investigate the prevalence of specific DNA configurations associated with Z-DNA formation. According to prior literature, the propensity for Z-DNA formation follows a descending order of prevalence: *gc*, *ca* and *at* (Shin et al. 2016). This pattern was also observed in the EBV sequences analyzed in Bozeman et al. (2025), where *gc* pairs appeared more frequently than *ca* pairs, which in turn were more frequent than *at* pairs (see Table 1 in Bozeman et al. (2025)). Moreover, transition sites between the B-DNA and Z-DNA conformations (and vice versa) were investigated, such as *cg*, *tg* and *at*, (see Ha et al. (2005), Kim et al. (2021) and Ravichandran et al. (2019)).

In this chapter, we want to establish if there is or is not a difference among those three sequences (coming from EBV) in the cases: *gc*, *ca*, *at* and *cg*, *tg* (transition states that includes *at*). We postulate that the sequence-generating processes follow Markov chain dynamics

over the specified alphabet $\mathcal{E} = \{a, c, g, t\}$, with memory $o = 2$ and state space $\mathcal{E} \times \mathcal{E}$. The three sequences that we compare are as follows: B95-8 (EBV), GD2 (EBV plus carcinoma) HKNPC1 (EBV plus carcinoma).

The objective is to detect differences between B95-8 and GD2, and between B95-8 and HKNPC1. We also compare GD2 and HKNPC1.

Here, we provide a description of each sequence: B95-8 (GenBank NC_007605, https://www.ncbi.nlm.nih.gov/nuccore/NC_007605.1/), as reported in Baer et al. (1984), is the first complete genome sequence of EBV derived from a type 1 strain and serves as the reference EBV sequence. This sequence was extracted from a North American case of infectious mononucleosis. GD2 (GenBank HQ020558, <https://www.ncbi.nlm.nih.gov/nuccore/HQ020558.1>), as described in Liu et al. (2011), is a type 1, EBV strain derived from a tumor in a nasopharyngeal carcinoma patient from Guangdong province, southern China. HKNPC1 (GenBank JQ009376, <https://www.ncbi.nlm.nih.gov/nuccore/JQ009376.2>), as documented in Kwok et al. (2012), is a strain type 1 of EBV, isolated from a primary nasopharyngeal carcinoma in a Chinese patient from Hong Kong. Table 4.1 summarizes details of the three sequences.

Our objective is to examine the set of states $\{gc, ca, at, cg, tg\} \subset \mathcal{E} \times \mathcal{E}$. We aim to demonstrate that when comparing scenarios with and without cancer, the discrepancy between sequences becomes detectable in varying degrees, depending on the metric used. In doing so, we seek to highlight that the presence of a cancer diagnosis could influence the frequencies of these states. In practical terms, we will treat the sequences listed in Table 4.1, expressed in FASTA format, as three samples: $z_{1,1}^{n_1}$ (GenBank: NC_007605.1), $z_{2,1}^{n_2}$ (GenBank: HQ020558.1) and $z_{3,1}^{n_3}$ (GenBank: JQ009376.2), each originating from a stochastic process. The comparisons we propose are $z_{1,1}^{n_1}$ versus $z_{2,1}^{n_2}$ and $z_{1,1}^{n_1}$ versus $z_{3,1}^{n_3}$, and also, as a reference $z_{2,1}^{n_2}$ versus $z_{3,1}^{n_3}$, with $n_1 = 171, 823$, $n_2 = 164, 669$ and $n_3 = 171, 553$, according to Table 4.1.

Table 4.1. From left to right: (i) sequence; (ii) GenBank and version of the sequence; (iii) size of the sequence, alphabet $\mathcal{E} = \{a, c, g, t\}$; (iv) main reference

Sequence	GenBank	Size n	Reference
B95-8	NC_007605.1	171823	(Baer et al. 1984)
GD2	HQ020558.1	164669	(Liu et al. 2011)
HKNPC1	JQ009376.2	171553	(Kwok et al. 2012)

The next subsection introduces the stochastic tool we use to compare the sequences.

4.1.2. The family of metrics

Consider two independent Markov chains $(Z_{i,t})_{t \in \mathbb{N}}$ and $(Z_{j,t})_{t \in \mathbb{N}}$, of finite order o , arranged on the finite alphabet $\mathcal{E}$ with state space $\mathcal{E}^o$, $i, j = 1, 2, 3, i \neq j$. Given $s \in \mathcal{E}^o$ denote by $\{P_i(a|s)\}_{a \in \mathcal{E}}$ and $\{P_j(a|s)\}_{a \in \mathcal{E}}$ the sets of transition probabilities of $(Z_{i,t})_{t \in \mathbb{N}}$ and $(Z_{j,t})_{t \in \mathbb{N}}$, respectively. Consider also two independent samples $z_{i,1}^{n_i}, z_{j,1}^{n_j}$ of $(Z_{i,t})_{t \in \mathbb{N}}$ and $(Z_{j,t})_{t \in \mathbb{N}}$, respectively.

Define

$N_{n_i+n_j}(s, a) = N_{n_i}(s, a) + N_{n_j}(s, a)$, $N_{n_i+n_j}(s) = N_{n_i}(s) + N_{n_j}(s)$, where N_{n_i} and N_{n_j} are given as before, computed from the samples $z_{i,1}^{n_i}$ and $z_{j,1}^{n_j}$, respectively.

Next, we introduce the notion that allows us to investigate the proximity/discrepancy between processes $(Z_{i,t})_{t \in \mathbb{N}}$ and $(Z_{j,t})_{t \in \mathbb{N}}$, through their samples, given by $z_{i,1}^{n_i}$ and $z_{j,1}^{n_j}$, respectively.

DEFINITION 4.1.– Consider two Markov chains $(Z_{i,t})_{t \in \mathbb{N}}$ and $(Z_{j,t})_{t \in \mathbb{N}}$, of finite order o , with finite alphabet $\mathcal{E}$, state space $\mathcal{E}^o$ and independent samples $z_{i,1}^{n_i}, z_{j,1}^{n_j}$, respectively. Define, for the state $s \in \mathcal{E}^o$:

$$d_s(z_{i,1}^{n_i}, z_{j,1}^{n_j}) = \psi_{i,j}^{-1} \sum_{a \in \mathcal{E}} \left\{ \sum_{k=i,j} N_{n_k}(s, a) \ln \left(\frac{N_{n_k}(s, a)}{N_{n_k}(s)} \right) - N_{n_i+n_j}(s, a) \ln \left(\frac{N_{n_i+n_j}(s, a)}{N_{n_i+n_j}(s)} \right) \right\}$$

with $\psi_{i,j} = \frac{(|\mathcal{E}|-1)}{\alpha} c_{n_i+n_j}$, with $\{c_{n_i+n_j}\}$ as a sequence of positive numbers depending on $n_i + n_j$ and α as a positive value.

Note that the notion given by definition 4.1 is a metric in the sample space (see for example García et al. (2018)), meaning that it also satisfies the triangle inequality. It is obtained from the Bayesian information criterion (BIC), when $c_{n_i+n_j} = \ln(n_i + n_j)$, and it is obtained from the EDC (García et al. 2024), otherwise.

Each member of this family, indexed by a choice of $c_{n_i+n_j}$, operates as follows: according to the model selection criterion, it indicates that the samples originate from the same stochastic law if and only if the value of $d_s(z_{i,1}^{n_i}, z_{j,1}^{n_j})$ is less than or equal to 1; otherwise, it concludes that the samples follow different stochastic laws. These results hold under the assumption of sufficiently large sample sizes, ensuring statistical consistency. To achieve such consistency, it is known that the sequence $c_{n_i+n_j}$ must satisfy certain conditions, such as:

$$\frac{c_{n_i+n_j}}{\ln(\ln(n_i + n_j))} \xrightarrow{\min(n_i, n_j) \rightarrow \infty} \infty, \text{ and } \frac{c_{n_i+n_j}}{n_i + n_j} \xrightarrow{\min(n_i, n_j) \rightarrow \infty} 0, \quad [4.3]$$

or be of the form $\ln(\ln(n_i + n_j))$. For this reason, we focus on two specific cases: $\ln(n_i + n_j)$, which satisfies the conditions [4.3] and corresponds to the BIC, and $\ln(\ln(n_i + n_j))$, which represents our second case of interest.

The following section shows the results of the comparisons.

4.2. Results

As explained above, we present our results using two versions of the metric. The first version, given in definition 4.1, has a penalty term equal to $\ln(n_i + n_j)$. The second version also follows definition 4.1, but with a penalty term equal to $\ln(\ln(n_i + n_j))$. In both cases, we use $\alpha = 2$, and the underlying theory is discussed in Schwarz (1978).

We construct dendrograms based on the d_s values (see Figures 4.1–4.5). For states where d_s does not capture discrimination between B95-8 and GD2 (B95-8 and HKNPC1), we report only the dendrograms. This situation is observed for state *gc* (Figure 4.1) and for state *at* (Figure 4.2). In contrast, for states where d_s does reveal discrimination between B95-8 and GD2 and between B95-8 and HKNPC1, we present the dendrograms, tables with the corresponding d_s values, and the estimated transition probabilities – equation [4.2] – for each one of the sequences B95-8, GD2 and HKNPC1. This situation is observed for state *cg* (see Figure 4.3, Tables 4.2 and 4.3), and for state *tg* (see Figure 4.4, and Tables 4.4 and 4.5); it is also observed for state *ca* (see Figure 4.5, Tables 4.6 and 4.7).

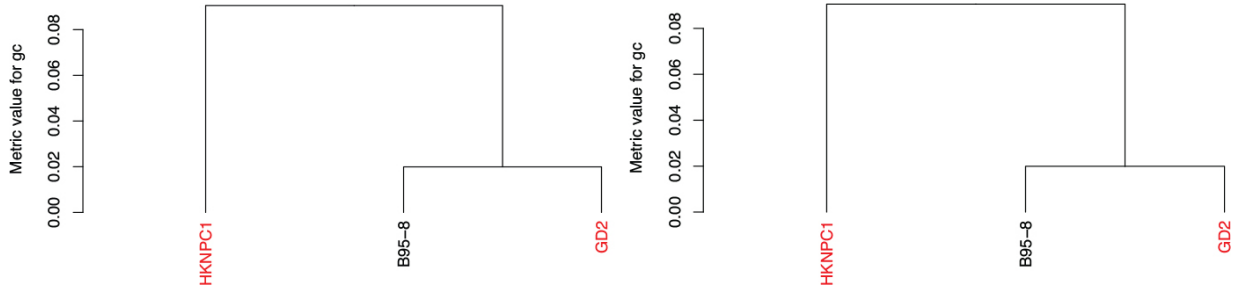


Figure 4.1. Dendrograms of d_s values (definition 4.1) for $s = gc$ (complete linkage) with $c_{n_i+n_j} = \ln(n_i + n_j)$ (left) and $c_{n_i+n_j} = \ln(\ln(n_i + n_j))$ (right), n_i and n_j denoting the samples sizes, $i, j = 1, 2, 3, i \neq j$. Sequences associated with carcinomas are highlighted in red (aource: authors' own work)

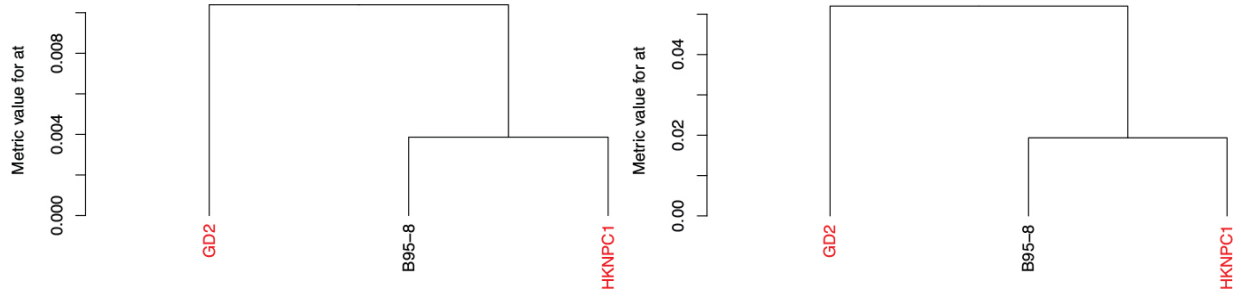


Figure 4.2. Dendrograms of d_s values (definition 4.1) for $s = at$ (complete linkage) with $c_{n_i+n_j} = \ln(n_i + n_j)$ (left) and $c_{n_i+n_j} = \ln(\ln(n_i + n_j))$ (right), n_i and n_j denoting the samples sizes, $i, j = 1, 2, 3, i \neq j$. Sequences associated with carcinomas are highlighted in red (

Source:authors' own work)

We show now the dendrograms of the cases in which the sequences are not distinguished based on cancer diagnosis: gc and at .

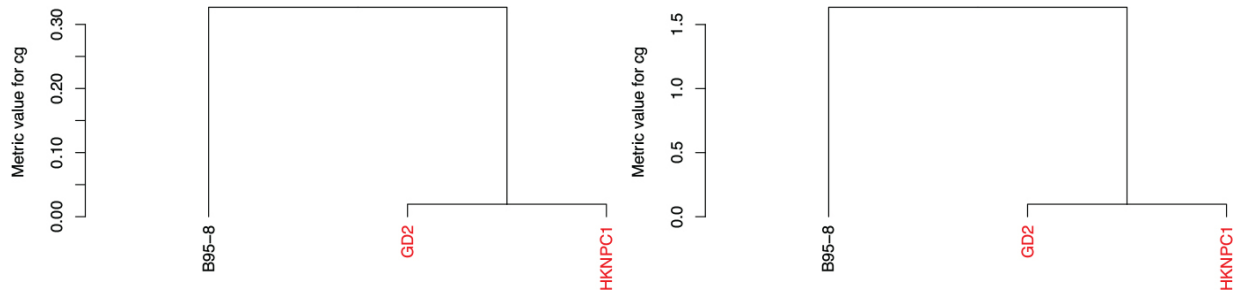


Figure 4.3. Dendrograms of d_s values (definition 4.1) for $s = cg$ (complete linkage) with $c_{n_i+n_j} = \ln(n_i + n_j)$ (left) and $c_{n_i+n_j} = \ln(\ln(n_i + n_j))$ (right), n_i and n_j denoting the samples sizes, $i, j = 1, 2, 3, i \neq j$. Sequences associated with carcinomas are highlighted in red (

Source:authors' own work)

Table 4.2. Estimated transition probabilities – equation [4.2] – from the state cg to elements of $\mathcal{E}$

	$\hat{P}_i(a cg)$	$\hat{P}_i(c cg)$	$\hat{P}_i(g cg)$	$\hat{P}_i(t cg)$
B95-8 ($i = 1$)	0.1427	0.2823	0.3814	0.1936
GD2 ($i = 2$)	0.1531	0.2766	0.3626	0.2076
HKNPC1 ($i = 3$)	0.1514	0.2726	0.3612	0.2075

Table 4.3. d_s values – definition 4.1 – $s = cg$, with $c_{n_i+n_j} = \ln(n_i + n_j)$ (left) and $c_{n_i+n_j} = \ln(\ln(n_i + n_j))$ (right), n_i and n_j denoting the samples sizes, $i, j = 1, 2, 3, i \neq j$

$\ln(n_i + n_j)$	B95-8	GD2	HKNPC1	$\ln(\ln(n_i + n_j))$	B95-8	GD2	HKNPC1
B95-8	-	0.3267	0.2656	B95-8	-	1.6346	1.3302
GD2	-	-	0.0194	GD2	-	-	0.0973
HKNPC1	-	-	-	HKNPC1	-	-	-

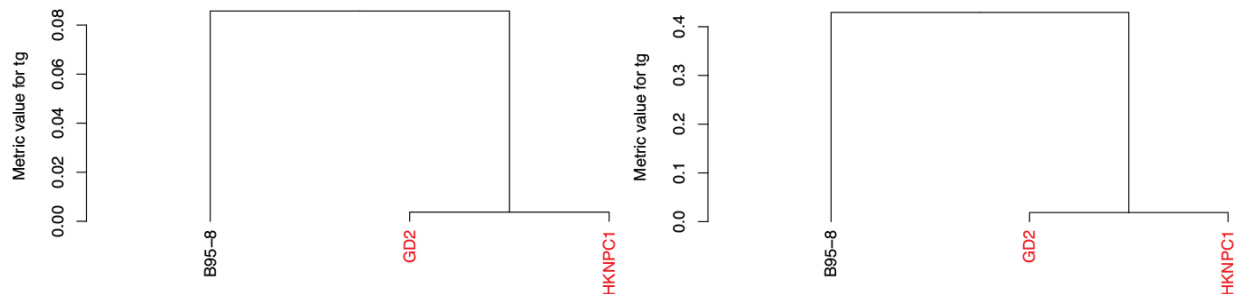


Figure 4.4. Dendrograms of d_s values (definition 4.1) for $s = tg$ (complete linkage) with $c_{n_i+n_j} = \ln(n_i + n_j)$ (left) and $c_{n_i+n_j} = \ln(\ln(n_i + n_j))$ (right), n_i and n_j denoting the samples sizes, $i, j = 1, 2, 3, i \neq j$. Sequences associated with carcinomas are highlighted in red (

Source: authors' own work)

Table 4.4. Estimated transition probabilities – equation [4.2] – from the state tg to elements of $\neq$

	$\hat{P}_i(a tg)$	$\hat{P}_i(c tg)$	$\hat{P}_i(g tg)$	$\hat{P}_i(t tg)$
B95-8 ($i = 1$)	0.1892	0.2440	0.3470	0.2200
GD2 ($i = 2$)	0.1930	0.2434	0.3376	0.2260
HKNPC1 ($i = 3$)	0.1947	0.2417	0.3377	0.2259

Table 4.5. d_s values – definition 4.1 – $s = tg$, with $c_{n_i+n_j} = \ln(n_i + n_j)$ (left) and $c_{n_i+n_j} = \ln(\ln(n_i + n_j))$ (right), n_i and n_j denoting the samples sizes, $i, j = 1, 2, 3, i \neq j$

$\ln(n_i + n_j)$	B95-8	GD2	HKNPC1	$\ln(\ln(n_i + n_j))$	B95-8	GD2	HKNPC1
B95-8	-	0.0737	0.0857	B95-8	-	0.3690	0.4293
GD2	-	-	0.0037	GD2	-	-	0.0187
HKNPC1	-	-	-	HKNPC1	-	-	-

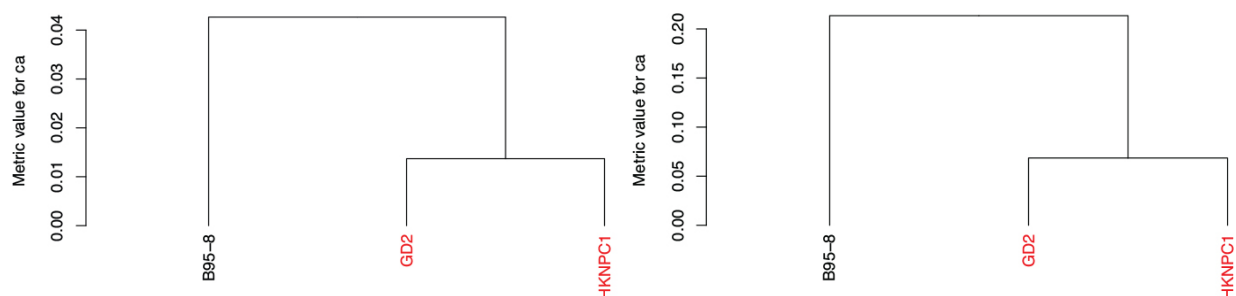


Figure 4.5. Dendrograms of d_s values (definition 4.1) for $s = ca$ (complete linkage) with $c_{n_i+n_j} = \ln(n_i + n_j)$ (left) and $c_{n_i+n_j} = \ln(\ln(n_i + n_j))$ (right), n_i and n_j denoting the samples sizes, $i, j = 1, 2, 3, i \neq j$. Sequences associated with carcinomas are highlighted in red (

Source: authors' own work)

Table 4.6. *Estimated transition probabilities – equation [4.2] – from the state ca to elements of $\mathcal{E}$*

	$\hat{P}_i(a ca)$	$\hat{P}_i(c ca)$	$\hat{P}_i(g ca)$	$\hat{P}_i(t ca)$
B95-8 ($i = 1$)	0.1688	0.2568	0.3795	0.1949
GD2 ($i = 2$)	0.1722	0.2574	0.3719	0.1984
HKNPC1 ($i = 3$)	0.1709	0.2549	0.3764	0.2175

The dendrograms corresponding to both cases, gc and at (left and right), reveal a similar hierarchical organization.

For cases in which the sequences are distinguished based on cancer diagnosis (cg , tg and ca), we present the results in that order, beginning with those associated with state cg , as one of the metrics yields d_s values greater than 1, between sequence B95-8 and sequence GD2, as well as between sequence B95-8 and sequence HKNPC1 (see [Table 4.3](#) (bottom)). Next, we focused on states tg and ca , as both retain the ability to distinguish carcinoma sequences from the viral reference sequence ([Figures 4.4](#) and [4.5](#)).

Table 4.7. d_s values – definition 4.1 – $s = ca$, with $c_{n_i+n_j} = \ln(n_i + n_j)$ (left) and $c_{n_i+n_j} = \ln(\ln(n_i + n_j))$ (right), n_i and n_j denoting the samples sizes, $i, j = 1, 2, 3, i \neq j$

$\ln(n_i + n_j)$	B95-8	GD2	HKNPC1	$\ln(\ln(n_i + n_j))$	B95-8	GD2	HKNPC1
B95-8	-	0.0427	0.0156	B95-8	-	0.2135	0.0780
GD2	-	-	0.0137	GD2	-	-	0.0685
HKNPC1	-	-	-	HKNPC1	-	-	-

Across all three states, cg , tg and ca , we observe that the magnitude of the d_s values increases proportionally when the penalty term $\ln(\ln(n_i + n_j))$ is applied, compared to the results obtained using penalty term $\ln(n_i + n_j)$. This pattern is consistently reflected when comparing the top and bottom sections of each d_s value table for states cg , tg and ca ([Tables 4.3](#), [4.5](#) and [4.7](#)).

The transition probability tables for these cases ([Tables 4.2, 4.4](#) and [4.6](#)) do not provide clear evidence of a discrepancy between B95-8 and GD2, nor between B95-8 and HKNPC1. This highlights the need for auxiliary tools to aid in such detections, such as the family of metrics introduced in definition 4.1.

4.3. Discussion

Given the established association between EBV and cancer (Stebbing and Bower [2009](#)), and the identification of Z-DNA patterns in Bozeman et al. ([2025](#)), it is pertinent to investigate the extent to which viral strains differ between individuals with and without a cancer diagnosis. In our analysis, three out of five configurations examined revealed discrepancies. Notably, two of these configurations, *cg* and *tg*, are known to participate in transitions between B-DNA and Z-DNA conformations. This observation may indicate a potential link between such structural transitions and the detected sequence differences.

This work raises several practical questions. First, we assessed the consistency between the decisions suggested by two metrics: one based on the BIC, using a penalty term of $\ln(n_i + n_j)$, and the other based on the EDC, with a penalty of $\ln(\ln(n_i + n_j))$. The resulting dendrograms exhibit similar structural patterns, with the primary difference being the magnitude of the scores (d_s), which tend to be proportionally larger under the EDC than under the BIC. It is important to note that any member of the family of notions introduced in definition 4.1 satisfies the properties of a metric. Therefore, differences in magnitude primarily reflect the level of stringency imposed when deciding whether two sequences follow the same underlying stochastic process (i.e. transition probabilities). As observed in [Tables 4.2, 4.4](#) and [4.6](#), even small discrepancies can be indicative of meaningful differences, an interpretation supported by the structures revealed in dendrograms [Figures 4.3–4.5](#). As a second point, note that by relaxing the penalty imposed by the BIC criterion, from $\ln(n_i + n_j)$ to the lower bound $\ln(\ln(n_i + n_j))$, we adopt the EDC-based framework. This relaxation enables a more sensitive detection of differences. In the case of the *cg* configuration, this approach

allows us to use a cutoff value of 1 for the d_s metric. Over this threshold, the discrepancies between B95-8 and GD2, as well as between B95-8 and HKNPC1, are particularly pronounced, indicating a strong divergence in the underlying stochastic behavior.

4.4. References

- Baer, R., Bankier, A.T., Biggin, M.D., Deininger, P.L., Farrell, P.J., Gibson, T.J., Hatfull, G., Hudson, G.S., Satchwell, S.C., Séguin, C., Tuffnell, P.S., Barrell, B.G. (1984). DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature*, 310, 207–211. doi: 10.1038/310207a0.
- Bozeman, J.R., García, J.E., González-López, V.A. (2025). Probability of Z-DNA forming subsequences and B-Z transition sites in the Epstein-Barr virus and others. *WSEAS Transactions on Biology and Biomedicine*, 22, 152–168 [Online]. Available at: [https://wseas.com/journals/bab/2025/a345108-015\(2025\).pdf](https://wseas.com/journals/bab/2025/a345108-015(2025).pdf).
- García, J.E., Gholizadeh, R., González-López, V.A. (2018). A BIC-based consistent metric between Markovian processes. *Appl Stochastic Models Bus Ind*, 34(6), 868–878. doi: 10.1002/asmb.2346.
- García, J.E., González-López, V.A., Gomez Sanchez, J.I. (2024). A metric based on the efficient determination criterion. *Entropy*, 26(6), 526. doi: 10.3390/e26060526.
- Ha, S.C., Lowenhaupt, K., Rich, A., Kim, Y.G., Kim, K.K. (2005). Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases. *Nature*, 437(7062), 1183–1186. doi: 10.1038/nature04088.
- Kim, S.H., Jung, H.J., Lee, I.B., Lee, N.K., Hong, S.C. (2021). Sequence-dependent cost for Z-form shapes the torsion-driven B-Z transition via close interplay of Z-DNA and DNA bubble. *Nucleic Acids Research*, 49(7), 3651–3660. doi: 10.1093/nar/gkab153.
- Kwok, H., Tong, A.H., Lin, C.H., Lok, S., Farrel, P.J., Kwong, D.L., Chiang, A.K. (2012). Genomic sequencing and comparative

analysis of Epstein-Barr virus genome isolated from primary nasopharyngeal carcinoma biopsy. *PLoS ONE*, 7(5), e36939. doi: 10.1371/journal.pone.0036939.

Liu, P., Fang, X., Feng, Z., Guo, Y.M., Peng, R.J., Liu, T., Huang, Z., Feng, Y., Sun, X., Xiong, Z. et al. (2011). Direct sequencing and characterization of a clinical isolate of Epstein-Barr virus from nasopharyngeal carcinoma tissue by using next-generation sequencing technology. *Journal of Virology*, 85(21), 11291–11299. doi: 10.1128/jvi. 0082311.

Ravichandran, S., Subramani, V.K., Kim, K.K. (2019). Z-DNA in the genome: From structure to disease. *Biophys Rev*, 11(3), 383–387. doi: 10.1007/s12551-019-00534-1.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464. doi: 10.1214/aos/1176344136.

Shin, S.I., Ham, S., Park, J., Seo, S.H., Lim, C.H., Jeon, H., Huh, J., Roh, T.Y. (2016). Z-DNA-forming sites identified by ChIP-Seq are associated with actively transcribed regions in the human genome. *DNA Research*, 23(5), 477–486. doi: 10.1093/dnares/dsw031.

Stebbing, J. and Bower, M. (2009). Epstein-Barr virus in Burkitt's lymphoma: The missing link. *The Lancet Oncology*, 10(4), 430. doi: 10.1016/S1470-2045(09)70045-2.

Zhao, L.C., Dorea, C.C.Y., Gonçalves, C.R. (2001). On determination of the order of a Markov chain. *Statistical Inference for Stochastic Processes*, 4(3), 273–282. doi: 10.1023/A:1012245821183.

Note

Chapter written by James R. BOZEMAN, Jesús E. GARCÍA, V.A. GONZÁLEZ-LÓPEZ and D.S. SANTOS.

5

Investigating the Mediating Role of Dis/Trust in Government to Deal with the Pandemic in the Relationship Between Covid-19 Cases, the Impact of the Pandemic on Employment and Social Trust: Evidence From the 2020 European Social Survey for Four European Countries

In the literature, trust in government is considered a crucial component for societies to function well. However, during a pandemic, trust is thought to have an ambiguous effect on outcomes. In this respect, the purpose of this study is to explore the mediating role of dis/trust in government to address the impact of the pandemic on the relationship between Covid-19 cases and employment and social trust. The analysis was based on the 2020 (Round 10) European Social Survey (ESS) datasets for four European countries: Finland, Greece, Italy and the Netherlands. In all four countries, the separate path analysis models provided adequate model fit. In all countries, social trust was positively predicted by trust in government to address the impact of the coronavirus pandemic. In all countries, trust in government was positively affected by the respondents having Covid-19. In all countries, the findings showed weak mediation effects of the respondents having Covid-19 by trust in government and work hours reduced by trust in government. Moreover, the findings showed weak mediation effects of unpaid vacation by trust in government in Finland, Greece and the Netherlands. In the cases of Greece and Italy, weak mediation effects of income reduced by trust in government were observed. These results indicate that further research is necessary.

5.1. Introduction

In the literature, trust in government is considered a crucial component for societies to function well. However, as Demertzis et al. ([2023](#), p. 6) pointed out, during a national emergency such as a pandemic, the experience of the pandemic and the government's handling of it can "contribute to both enhancement and decline of public trust". In such cases, greater support and trust are generated in key political institutions, and the relevant literature

has shown that this “rally effect associated with the pandemic lasted on average 9 months and in many cases was over by the end of 2020” (Demertzis et al. [2023](#), p. 6). In line with this, Reiersen et al. ([2022](#)) argued that during a pandemic, trust is thought to have an ambiguous effect on its outcomes:

On one hand, a high level of trust in society may lead to greater acceptance among citizens for public measures that aim to combat a virus. If people believe that their respective governments implement unbiased and well-informed measures and people also believe that their fellow citizens will follow these measures, this may lead to a high general compliance in society and less people will be infected. On the other hand, trust may affect people’s perception of risk and hence their behaviour. If people believe that most people are trustworthy, they may be less willing to think of everyone else as a potential health threat. If people also trust the government to manage the pandemic in a competent way, their perception of the risks related to the pandemic weaken (Reiersen et al. [2022](#), p. 1).

Accordingly, political trust in governments handling the pandemic and social trust are crucial for any public policy measures taken to succeed (Toshkov et al. [2022](#)).

Devine et al. ([2021](#), p. 275) pointed out that, in the literature, although there is widespread disagreement on the measurement of political or social trust,

how citizens come to trust judgements and what the consequences are of trust being present or absent for governing. Yet, engaging with these questions is imperative in the face of a life-threatening – and certainly life-changing – pandemic. Understandably, there has been interest both in the consequences of the pandemic and government measures for levels of trust as well as the role trust plays in societal responses to it.

Referring to Van Bavel et al. (2020), they suggested that “this is important since high levels of trust are seen to be a necessary condition for the implementation of restrictive policies and for public compliance with them” (Devine et al. [2021](#), p. 275).

Demertzis et al. ([2023](#), pp. 5–6) proposed that focusing on trust in government and people’s experiences of the coronavirus pandemic is important, as

societal and governmental responses to Covid-19 have been highly contingent upon trust so much so that countries exhibiting high levels of trust more often than not have tended to implement less stringent lockdowns. These governments also relied largely on the voluntary compliance of citizens to governmental guidelines. The flip side of this is that high levels of social trust may prompt one to believe that others are closely observing stringency measures so that s/he is likely to follow a sort of free-rider effect concerning vaccination, and other protective measures. Moreover, high trust in the government and other authorities may facilitate underestimation of the risk of the pandemic and a perception that individual action is redundant. Thus, people may think that the government will do all necessary jobs with the pandemic.

Bieszk-Stolorz and Dmytrów ([2022](#), p. 3) studied the impact of Covid-19 on EU labor markets and noted that

in times of crisis, the labour market becomes one of the first to experience severe turbulence. Employability is one of the parameters that has changed significantly due to the existence of the global health crisis. People employed in flexible forms of employment, which are hardly subject to any legal protection, are the first to lose their jobs [...]. Informal workers, youth, and women [...], as well as small traders, the self-employed, migrant workers, and daily wage earners [...] were the first to experience employment problems.

All relevant research concurs with these findings (see, for example Fana et al. [2020](#); Cook and Grimshaw [2021](#); Abrahám and Vošta [2022](#)).

In this chapter, with the above theoretical and empirical insights in mind, we investigate the mediating role of dis/trust in government to address the impact of the pandemic on the relationship between Covid-19 cases and employment and social trust. The analysis was based on the 2020 (Round 10) ESS datasets for four European countries: Finland, Greece, Italy and the Netherlands.

5.2. Method

5.2.1. Procedure and participants

The ESS questionnaire comprises a core questionnaire that is repeated in all rounds of the survey and one or two rotated modules. In the 2020 ESS questionnaire, a module on the Covid-19 pandemic was included. The analysis was based on the 2020 ESS datasets for the following four countries that had completed the Covid-19 module: Finland, Greece, Italy and the Netherlands. The ESS implements all the strict methodological prerequisites

for overtime and cross-national comparability (Kish [1994](#); Carey [2000](#)) by applying probability sampling, minimum effective achieved sample sizes in all participating countries, and a maximum target non-response rate of 30% (The ESS Sampling Expert Panel [2016](#)). Applying the same definition to all rounds, the ESS defines the survey population as all individuals aged 15+ residing within private households in each country, regardless of their nationality, citizenship or language. The surveys of all rounds exclusively used face-to-face interviewing as the data collection method, but in the 2020 survey, due to the pandemic, national teams implemented self-completion methods for the first time. Therefore, two different datasets were provided for the countries implementing face-to-face interviewing (European Social Survey European Research Infrastructure (ESS ERIC) [2022a](#)) and self-completion (European Social Survey European Research Infrastructure (ESS ERIC) [2022b](#)) methods. The ESS standardized questionnaire is designed centrally in British English, including a core questionnaire that is repeated in all rounds of the survey and a rotating module employing rigorous translation strategies (Harkness et al. [2010](#)). In the 2020 survey, the questionnaire was used in the same manner as in all previous surveys for the face-to-face interviewing (European Social Survey [2020a](#)) and a shorter questionnaire was used for the surveys applying self-completion (European Social Survey [2020b](#)) methods. Therefore, the two ESS surveys are not comparable. However, as the four countries under investigation used face-to-face interviewing, their findings are comparable.

In [Table 5.1](#), a summary of the participants' demographic and social characteristics is presented. As shown, gender was almost equally distributed in most countries apart from Greece and Italy, where the samples included more women than men. The mean age ranged from 48 (Netherlands) to 52 (Finland). Married participants and those who had completed secondary education ranged from 44.3% (Italy) to 47.2% (Greece) and 50.0% (Finland) to 77.8% (Italy), respectively. Participants in paid work ranged from 48.8% (Italy) to 65.8% (Netherlands).

Furthermore, note that the achieved sample sizes ranged from 1,470 (Netherlands) to 2,799 (Greece). The response rates were as follows: Finland (41.1%), Greece (48.0%), Italy (49.8%) and the Netherlands (35.7%). Therefore, in all countries, the response rates did not meet the ESS requirement of 70%.

Table 5.1. *The demographic and social characteristics of participants: European Social Survey, 2020*

		Men	Women	Age	Married	Educ*	In paid work**
Country	<i>N</i>	(%)	(%)	$\bar{X}$ (SD)	(%)	(%)	(%)
Finland	1,577	49.5	50.5	52.61 (19.323)	45.1	50.0	52.8
Greece	2,799	47.7	52.3	50.38 (16.974)	47.2	65.2	54.4
Italy	2,640	47.5	52.5	51.59 (18.690)	44.3	77.8	48.8
Netherlands	1,470	51.0	49.0	48.62 (18.502)	45.5	52.2	65.8

*Educ = secondary education or lower (European Social Survey version of ISCED, ES-ISCED, ES-ISCED I – ES-ISCED IIIa).

**The reference period for the respondent’s main activity was defined as during the last 7 days.

5.2.2. Measures

In the analyses, one item from the core questionnaire worded as “Most people would try to take advantage of you if they got the chance (0), or would they try to be fair (10)” (A5) was used as an indicator of social trust labeled “Peoplefair”. The level of measurement of this variable is interval.

The Covid-19 module comprised Section K of the 2020 ESS questionnaire (Gemenis and Littvay [2020](#); Helbling et al. [2020](#); Hanson et al. [2021](#)). It included two questions (K17, K18) on whether the respondents or anyone living with them had Covid-19, labeled “RespCovid19” and “Living RespCovid19”, respectively. Both of these variables are dichotomies.

A submodule on “government authority and legitimacy in the age of a pandemic” (Helbling et al. [2020](#); see also Hanson et al. [2021](#)) included a question assessing the dis/trust in government to deal with the impact of the coronavirus pandemic (K15) worded as follows: “Using this card, please tell me to what extent you trust the national government in [country] to deal with the impact of the coronavirus pandemic. Please answer on a score of 0-10 where 0 means you do not trust the government at all and 10 means you have complete trust”. The level of measurement of this variable labeled “Govhandling” is interval.

Finally, a multiple response question was used to assess the impact of the pandemic on employment (K19), worded as follows: “Using this card, please tell me if any of the following happened to you as a result of the coronavirus pandemic? Please include anything that has happened at any time since the start of the pandemic, even if it is no longer impacting you. Select all that apply: I was made redundant/lost my job (1); The income from my job was reduced (2); My working hours were reduced (3); I was furloughed (4); I was forced to take unpaid leave/holiday (5); None of these (6); I have not been in work at any time since the start of the pandemic (7)”. For analysis purposes, a variable was created allocating a single answer per respondent. The level of measurement of this variable is nominal. Moreover, from this variable, seven dummy (dichotomy) variables were created labeled “Redundant”, “Incomereduced”, “Workhreduced”, “Furlough”, “Unpaidvacation”, “Noneofthese” and “Notinwork”.

5.2.3. Statistical analyses

5.2.3.1. Preliminary decisions, tests and procedures

Preliminary decisions, tests and procedures refer to the adequacy of the sample size for performing path analyses and how outliers and missing values will be addressed.

Kline ([1998](#)) proposed that a sample size of 10 times the number of parameters in a model is adequate for performing path analysis and that a sample size of 20 times the number of parameters in path analysis is best. However, it should be pointed out that path analysis is a large sample method (Tabachnick and Fidell [2007](#)).

Data screening for outliers was based on the variables’ boxplots. Cases were eliminated if they were shown in the boxplots as outliers (Tabachnick and Fidell [2007](#)).

Missing data analysis was performed, and the size and pattern of missing values for each variable were investigated. Missing data had to be replaced because complete datasets are required for SPSS Amos.

5.2.3.2. The theoretical structure of the path analysis model

The theoretical structure of the path analysis model under consideration was defined as follows ([Figure 5.1](#)): the two variables measuring experiences of the pandemic RespCovid-19 and Living RespCovid-19, and the seven dummy variables assessing the impact of the pandemic on employment – Redundant, Incomereduced, Workhreduced, Furlough, Unpaidvacation, Noneofthese and Notinwork – defined the independent variables (IV); the question on dis/trusting the national government to deal with the impact of

the coronavirus pandemic Govhandling was the mediator; and the indicator of social trust Peoplefair was the dependent variable (DV).

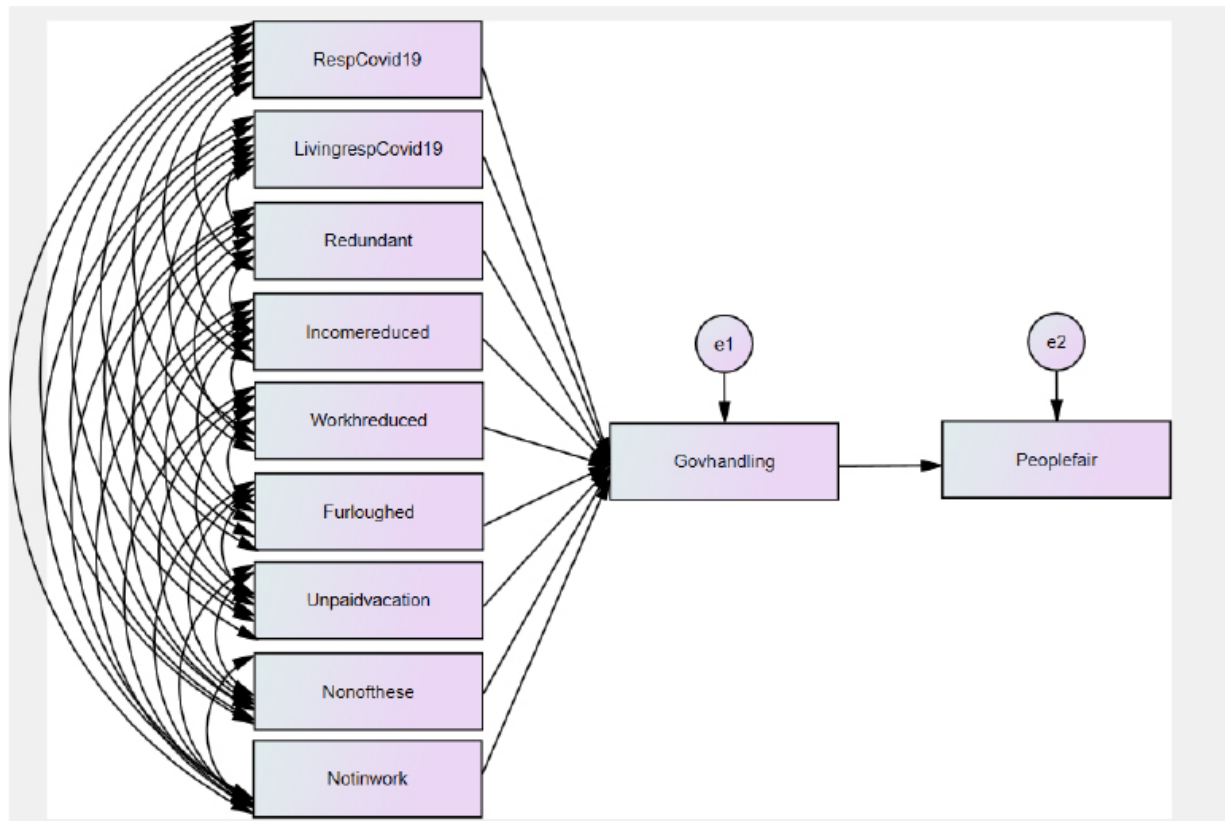


Figure 5.1. *The theoretical structure of the path analysis model*

The level of measurement of the DV and mediator variables is interval, and all IVs are dichotomies. The background variable measuring educational attainment (F15, pseudointerval) was used as a control.

5.2.3.3. Path analysis

Initially, univariate statistics were computed for each measure, and their distributional properties were inspected (testing for normality) to decide on the appropriateness of the methods to be used.

Separate path analyses were performed with maximum likelihood for each country. Standardized indirect effects were considered the “index of mediation” (Preacher and Hayes [2008](#), p. 38; see also Cheung [2009](#)) and were tested with bootstrapping methods.

To generate bootstrap confidence intervals for the standardized indirect effects, 5,000 bootstrap samples were used (Preacher and Hayes [2008](#); Cheung [2009](#)). A standardized indirect effect was considered statistically

significant only if its 95% bootstrap confidence interval excluded zero (Hayes [2018](#)).

Model fit was considered adequate if chi-square tests were not significant, the comparative fit index (CFI) and Tucker–Lewis index (TLI) values were greater than or close to 0.95 and the root-mean-square error approximation (RMSEA) ≤ 0.06 with the 90% confidence interval (CI) upper limit ≤ 0.06 (Bollen [1989](#); Hu and Bentler [1999](#); Tabachnick and Fidell [2007](#)). Model fit was considered acceptable if chi-square tests were not significant, CFI and TLI values were > 0.90 and RMSEA < 0.08 with the 90% CI upper limit < 0.08 (Hu and Bentler [1999](#); Marsh et al. [2004](#)).

Statistical analyses and path analyses were performed using IBM SPSS Statistics Version 29 and IBM SPSS Amos Version 29, respectively.

5.3. Results

5.3.1. Covid-19 cases and the impact of the pandemic on employment

In [Figures 5.2](#) and [5.3](#), the frequency distributions of the variables measuring whether the respondents or anyone living with them had Covid-19 are presented. As shown in [Figure 5.2](#), the respondents declaring that they had tested positive for Covid-19 ranged from 1.3% (Finland) to 25.7% (Netherlands). Those who thought they had Covid-19 but were not tested or did not test positive ranged from 2.5% (Finland) to 6.1% (Netherlands). More than 50.0% of the respondents did not have Covid-19, ranging from 68.2% (Netherlands) to 96.2% (Finland).

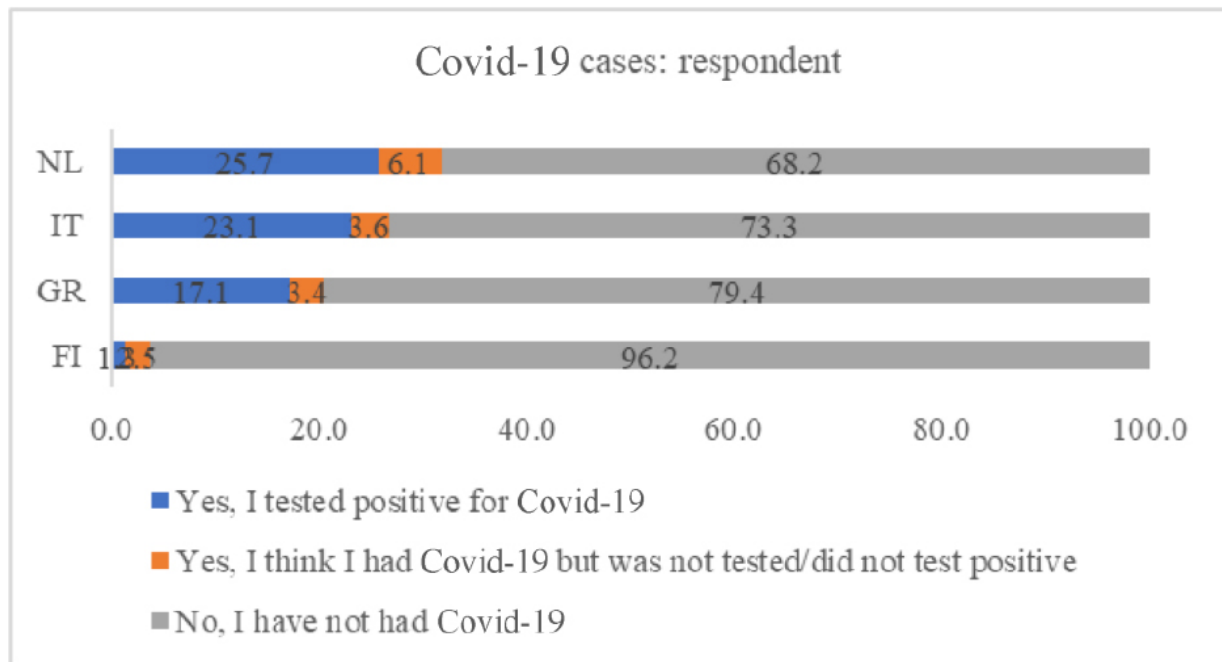


Figure 5.2. *The frequency (%) of respondents having had Covid-19: European Social Survey, 2020.*

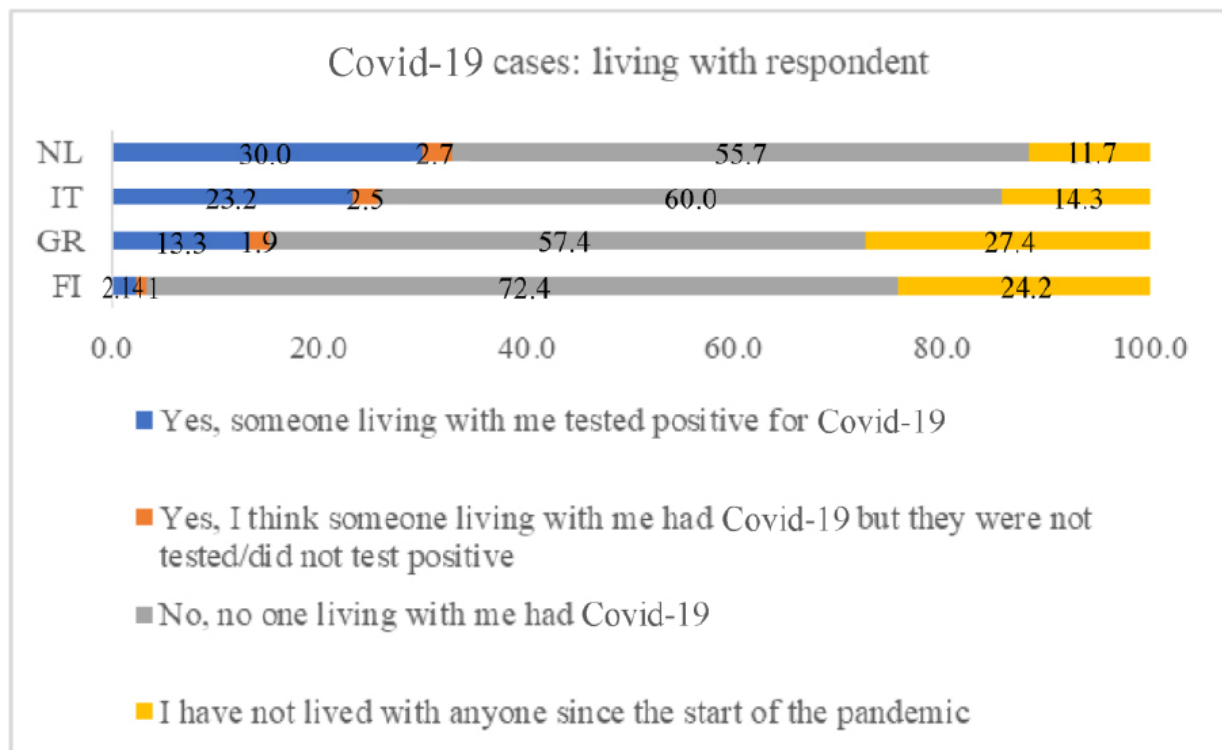


Figure 5.3. *The frequency (%) of anyone living with respondents having had Covid-19: European Social Survey, 2020.*

The respondents declaring that someone living with them had tested positive for Covid-19 ([Figure 5.3](#)) ranged from 2.4% (Finland) to 30.0% (Netherlands). Those who thought that someone living with them had Covid-19 but they were not tested/did not test positive ranged from 1.1% (Finland) to 2.7% (Netherlands). The respondents who reported that no one living with them had Covid-19 ranged from 55.7% (Netherlands) to 72.4% (Finland).

In [Figure 5.4](#), the frequency distribution of the impact of the pandemic on employment is presented. As shown, most of the employed respondents in Finland (47%) and Greece (45%) declared that they were furloughed and most of those in the Netherlands (47%) declared that their working hours were mainly reduced. Most Italians declared that their working hours were reduced (29%) and that they were furloughed (28%).

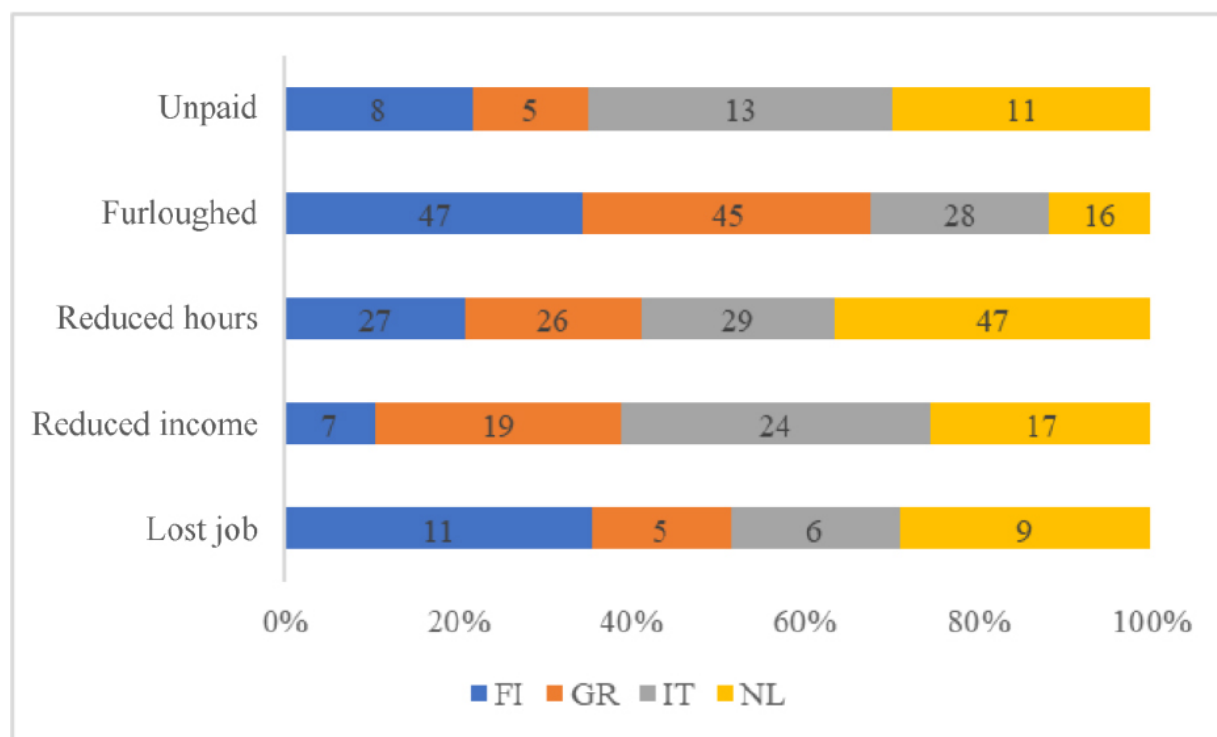


Figure 5.4. *The frequency (%) of the impact of the pandemic on the employed: European Social Survey, 2020.*

5.3.2. Political and social trust in the age of a pandemic

In [Figures 5.5](#) and [5.6](#), the mean values of the variables assessing dis/trust in government to address the impact of the coronavirus pandemic and social trust are presented, respectively.

As shown in [Figure 5.5](#), the lowest and highest levels of political dis/trust in government to deal with the pandemic were observed in Greece (4.49) and

Finland (6.805), respectively. The same pattern was repeated in the levels of social trust ([Figure 5.6](#)), albeit with rather higher values, as the mean scores of Greece and Finland were 4.75 and 7.073, respectively.

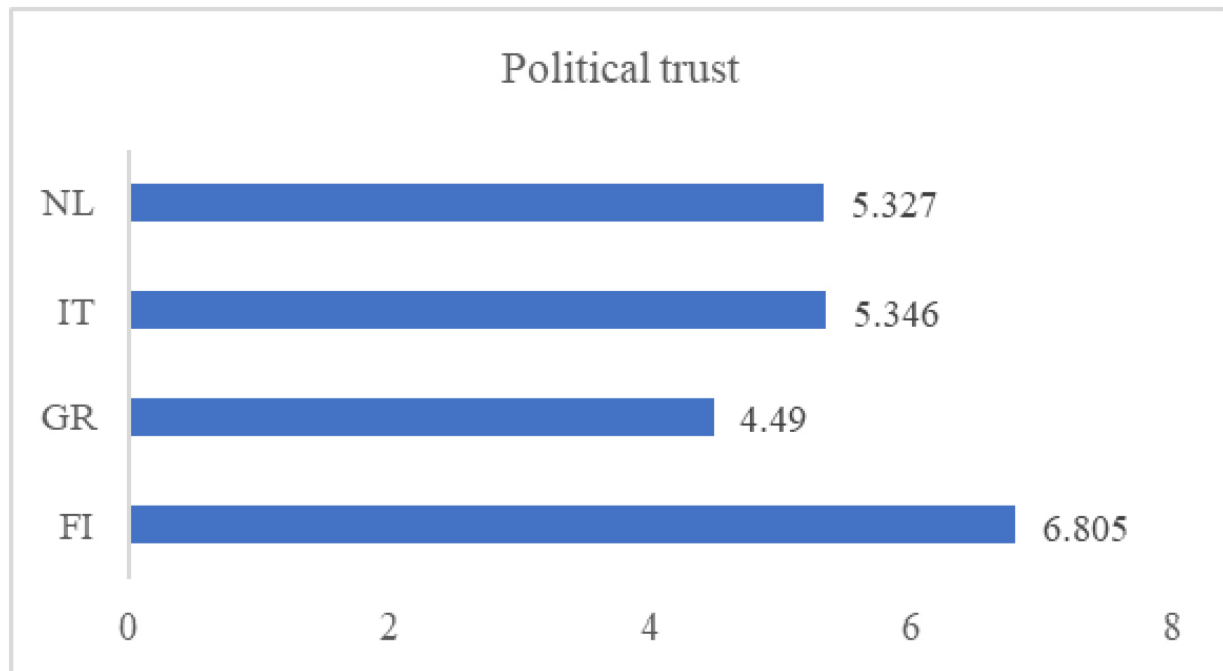


Figure 5.5. *The mean values of the variables assessing the trust in government to deal with the impact of the coronavirus: European Social Survey, 2020*

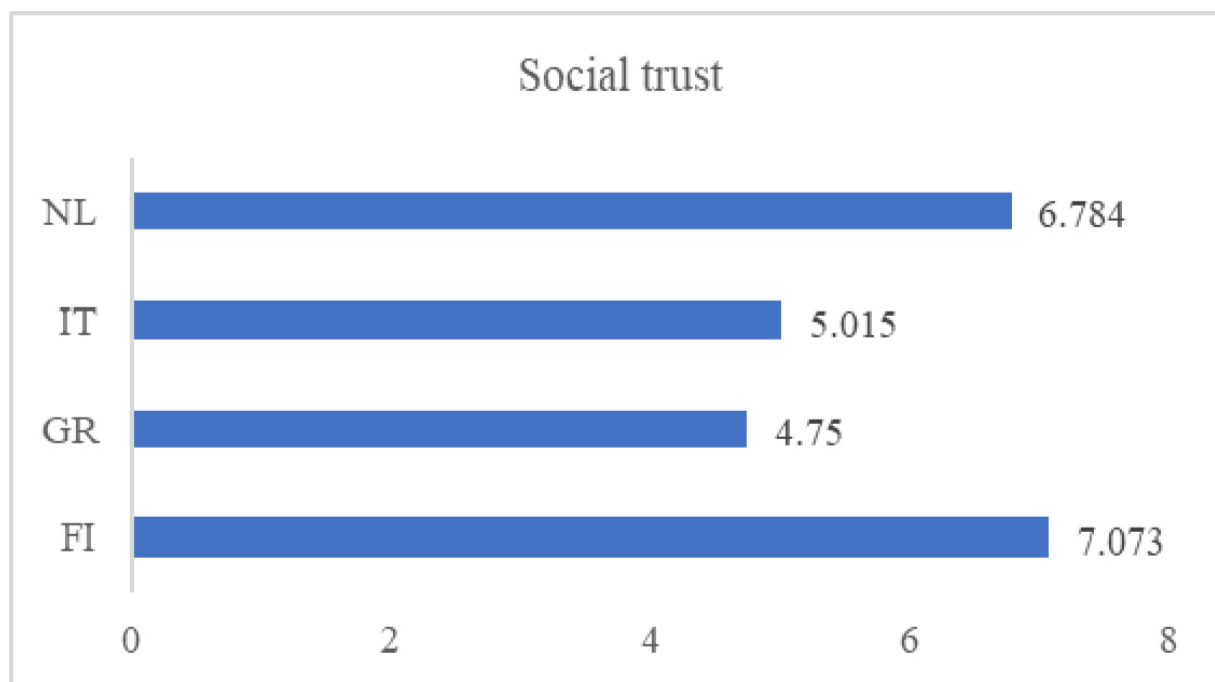


Figure 5.6. *The mean values of the variables assessing the social trust in the age of the coronavirus pandemic: European Social Survey, 2020*

5.3.3. Mediation analysis

The sample sizes ([Table 5.1](#)) ranged from 1,470 (Netherlands) to 2,799 (Greece) and were considered large enough for carrying out path analyses separately in each country. The screening of samples detected only a negligible number of outlying cases, and it was decided not to reject them from the analysis. In every country, respondents had used the full range of possible responses for all variables. All missing values were replaced by the mean values of each variable because complete datasets were required for SPSS Amos.

In [Table 5.2](#), the goodness of fit indices for the path analyses models performed with maximum likelihood are presented for each country. As shown, the models of all countries provided adequate model fit – except for the TLI value in the case of Greece, which indicates acceptable model fit.

Table 5.2. *Path analyses performed with maximum likelihood: goodness-of-fit indices*

Country	χ^2 (df)	<i>p</i>	CFI	TLI	RMSEA (90% CI)
Finland	20.890 (9)	0.013	0.990	0.926	0.029 (0.013–0.045)
Greece	60.198 (9)	0.000	0.983	0.873	0.045 (0.035–0.056)
Italy	17.345 (9)	0.044	0.997	0.979	0.019 (0.007–0.044)
Netherlands	18.089 (9)	0.034	0.996	0.968	0.026 (0.007–0.044)

df = degrees of freedom; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root-mean-square error of approximation; CI = confidence interval. Model fit is considered adequate if χ^2 is not significant, CFI and TLI values are greater than or close to 0.95 and RMSEA $\leq$ 0.06 with the 90% CI upper limit $\leq$ 0.06. Model fit is considered acceptable if χ^2 is not significant, CFI > 0.90, TLI > 0.90 and RMSEA < 0.08 with the 90% CI upper limit < 0.08.

In [Table 5.3](#), the path analysis results are presented. As shown, in all countries, social trust was positively predicted by dis/trust in government to address the impact of the coronavirus pandemic. In all countries, political dis/trust was positively affected by having Covid-19. In the cases of Finland and Greece, political trust was positively affected by anyone living with the respondent having Covid-19 and negatively affected in Italy and the Netherlands. In all countries, political dis/trust was negatively affected by all seven variables assessing the impact of the pandemic on employment, except in Italy for “working hours reduced” and “none of these”. The squared multiple correlations for social trust and dis/trust in government to address the impact of the coronavirus pandemic ranged from 0.028 (Italy) to 0.049 (Netherlands) and 0.024 (Greece, Italy) to 0.038 (Finland), respectively.

Table 5.3. Path analyses performed with maximum likelihood: standardized coefficients and squared multiple correlations

			β			
Model			Finland	Greece	Italy	Netherlands
Govhandling	←	RespCovid19	0.065	0.087	0.093	0.111
Govhandling	←	LivingrespCovid19	0.016	0.002	– 0.077	–0.038
Govhandling	←	Redundant	–0.007	–0.017	– 0.030	–0.030
Govhandling	←	Incomereduced	–0.023	–0.097	– 0.078	–0.034
Govhandling	←	Workhreduced	–0.121	0.082	– 0.033	–0.108
Govhandling	←	Furlough	–0.070	–0.001	0.019	–0.036
Govhandling	←	Unpaidvacation	–0.100	–0.048	– 0.082	–0.052
Govhandling	←	Noneofthese	–0.145	0.018	– 0.028	–0.100
Govhandling	←	Notinwork	–0.053	–0.033	– 0.033	–0.024
DV: Peoplefair	←	Govhandling	0.150	0.173	0.153	0.161
			R^2			
		Variable	Finland	Greece	Italy	Netherlands
		Govhandling	0.038	0.024	0.024	0.031
		Peoplefair	0.030	0.035	0.028	0.049

In [Table 5.4](#), the mediating effects are presented for each country. As shown, weak simple mediation effects of respondents having Covid-19 by dis/trust in government to deal with the impact of the coronavirus pandemic were detected in all countries with significant standardized indirect effects: 0.010 (Finland), 0.015 (Greece), 0.014 (Italy) and 0.018 (Netherlands). Moreover, weak simple mediation effects of working hours reduced by political dis/trust were observed in all countries except Italy, with significant standardized indirect effects: –0.018 (Finland), 0.014 (Greece) and –0.017 (Netherlands).

Furthermore, weak simple mediation effects of unpaid vacation by political dis/trust were detected in all countries except the Netherlands with significant standardized indirect effects: -0.015 (Finland), -0.008 (Greece) and -0.013 (Italy). Simple weak mediation effects of income reduced by political dis/trust were observed in Greece and Italy, with significant standardized indirect effects of -0.017 and -0.012 , respectively.

Additionally, weak simple mediation effects of “none of these” by political dis/trust were observed in Finland and the Netherlands, with significant standardized indirect effects of -0.022 and -0.016 , respectively. Only in Italy were weak simple mediation effects of anyone living with the respondent having Covid-19 by political dis/trust detected, with significant standardized indirect effects of -0.012 . Additionally, only in Finland were weak simple mediation effects of furlough by political dis/trust detected, with significant standardized indirect effects of -0.011 .

There were no mediating effects in any country for those made redundant by political dis/trust.

Table 5.4. *Path analyses: standardized indirect effects (SIE), their 95% bootstrap confidence intervals (CI) and significance*

	Finland			Greece		
Variables	SIE	95% CI	p value	SIE	95% CI	p value
RespCovid19	0.010	0.009, 0.144	0.022	0.015	0.008, 0.023	0.000
LivingrespCovid19	0.002	-0.020, 0.038	0.560	0.000	-0.007, 0.008	0.896
Redundant	-0.001	-0.118, 0.085	0.799	-0.003	-0.009, 0.003	0.365
Incomereduced	-0.003	-0.206, 0.074	0.436	-0.017	-0.025, -0.009	0.000
Workhreduced	-0.018	-0.301, -0.079	0.000	0.014	0.007, 0.022	0.000
Furlough	-0.011	-0.151, -0.026	0.005	0.000	-0.007, 0.007	0.972
Unpaidvacation	-0.015	-0.512, -0.091	0.002	-0.008	-0.017, -0.001	0.030
Noneofthese	-0.022	-0.131, -0.041	0.000	0.003	-0.005, 0.013	0.437
Notinwork	-0.008	-0.098, 0.000	0.050	-0.006	-0.013, 0.001	0.102
	Italy			Netherlands		
RespCovid19	0.014	0.007, 0.024	0.000	0.018	0.015, 0.056	0.000
LivingrespCovid19	-0.012	-0.021, -0.005	0.001	-0.006	-0.026, 0.005	0.174
Redundant	-0.005	-0.011, 0.001	0.098	-0.005	-0.134, 0.035	0.247
Incomereduced	-0.012	-0.020, -0.005	0.000	-0.005	-0.102, 0.026	0.244
Workhreduced	-0.005	-0.012, 0.001	0.100	-0.017	-0.179, -0.041	0.001
Furlough	0.003	-0.003, 0.009	0.343	-0.006	-0.149, 0.017	0.135

	Finland			Greece		
Variables	SIE	95% CI	p value	SIE	95% CI	p value
Unpaidvacation	– 0.013	–0.022, – 0.006	0.001	– 0.008	–0.257, – 0.013	0.084
Noneofthese	– 0.004	–0.012, – 0.003	0.251	– 0.016	–0.097, – 0.017	0.002
Notinwork	– 0.005	–0.013, – 0.002	0.160	– 0.004	–0.092, – 0.032	0.389

5.4. Conclusion

In this chapter, the mediating role of dis/trust in government to address the impact of the pandemic on the relationship between Covid-19 cases and employment and social trust was investigated. The analysis was based on the 2020 (Round 10) ESS datasets for four European countries: Finland, Greece, Italy and the Netherlands.

The models of the separate path analyses provided adequate model fit for all countries. In all countries, social trust was positively predicted by dis/trust in government to address the impact of the coronavirus pandemic. In all countries, political dis/trust was positively affected by having Covid-19. In the cases of Finland and Greece, political trust was positively affected by anyone living with the respondent having Covid-19 and negatively affected in Italy and the Netherlands. In all countries, political dis/trust was negatively affected by all seven variables assessing the impact of the pandemic on employment – “made redundant”, “income reduced”, “working hours reduced”, “furloughed”, “unpaid vacation”, “none of these”, “not in work” – except in Italy for “working hours reduced” and “none of these”.

In all countries, the results showed weak simple mediation effects of respondents having Covid-19 by dis/trust in government to address the impact of the coronavirus pandemic. Moreover, weak simple mediation effects of working hours reduced by political dis/trust were observed in all countries with the exception of Italy. Furthermore, weak simple mediation effects of unpaid vacation by political dis/trust were detected in all countries except the Netherlands. Additionally, simple weak mediation effects of income reduced by political dis/trust were observed in Greece and Italy. Additionally, weak simple mediation effects of “none of these” by political dis/trust were observed in Finland and the Netherlands. Only in Italy were weak simple mediation effects of anyone living with the respondent having Covid-19 by political dis/trust detected. Additionally, only in Finland were

weak simple mediation effects of furloughed by political dis/trust detected. There were no mediating effects in any country for those made redundant by political dis/trust.

Certainly, the resulting percentages of the total variation explained by the models with social trust and dis/trust in government to address the impact of the coronavirus pandemic as the dependent and mediating variables, respectively, indicate that in future research, additional social indicators should be considered to provide a more coherent interpretation.

5.5. References

- Abrhám, J. and Vošta, M. (2022). Impact of the Covid-19 pandemic on EU. *Convergence Journal of Risk and Financial Management*, 15, 384. doi: 10.3390/jrfm15090384.
- Bieszk-Stolorz, B. and Dmytrów, K. (2022). Assessment of the similarity of the situation in the EU labour markets and their changes in the face of the Covid-19 pandemic. *Sustainability*, 14, 3646. doi: 10.3390/su14063646.
- Bollen, K.A. (1989). *Structural Equations with Latent Variables*. Wiley, New York.
- Carey, S. (ed.) (2000). Measuring adult literacy: The International Adult Literacy Survey (IALS) in the European context. Office for National Statistics, London.
- Cheung, M.W.-L. (2009). Comparison of methods for constructing confidence intervals of standardized indirect effects. *Behavior Research Methods*, 41(2), 425–438. doi: 10.3758/BRM.41.2.425.
- Cook, R. and Grimshaw, D. (2021). A gendered lens on Covid-19 employment and social policies in Europe. *European Societies*, 23(S1), S215–S227. doi: 10.1080/14616696.2020.1822538.
- Demertzis, N., Yfanti, A., Koniordos, S. (2023). Literature review on the impact of the Covid-19 pandemic on democratic attitudes and trust. Working paper No 3.1. TRUEDEM: Trust in European Democracies Project [Online]. Available at: www.TRUEDEM.eu.
- Devine, D., Gaskell, J., Jennings, W., Stoker, G. (2021). Trust and the coronavirus pandemic: What are the consequences of and for trust? An early review of the literature. *Political Studies Review*, 19(2), 274–285.
- European Social Survey (2020a). ESS Round 10 Source Questionnaire. ESS ERIC, University of London, London.

European Social Survey (2020b). ESS Round 10 Source Questionnaire (Self completion, Paper). ESS ERIC, University of London, London.

European Social Survey European Research Infrastructure (ESS ERIC) (2022a). ESS10 – integrated file, edition 2.1. Data set, Sikt - Norwegian Agency for Shared Services in Education and Research. doi: 10.18712/ess10.

European Social Survey European Research Infrastructure (ESS ERIC) (2022b). ESS10 Self-completion – integrated file, edition 1.0. Data set, Sikt - Norwegian Agency for Shared Services in Education and Research. doi: 10.18712/ess10sc.

Fana, M., Torrejón Pérez, S., Fernández-Macías, E. (2020). Employment impact of Covid-19 crisis: From short term effects to long terms prospects. *Journal of Industrial and Business Economics*, 47, 391–410. doi: 10.1007/s40812-020-00168-5.

Gemenis, K. and Littvay, L. (2020). Covid-19 conspiracy beliefs and government rule compliance. Sub-module proposal for ESS Round 10 Covid-19 module [Online]. Available at: <https://www.europeansocialsurvey.org/docs/about/Gemenis-Covid-19-Conspiracy-Beliefs-and-Government-Rule-Compliance.pdf>.

Hanson, T., Helbling, M., Maxwell, R., Trautmüller, R., Gemenis, K., Littvay, L. (2021). Developing a Covid-19 module for the European social survey. *Measurement Instruments for the Social Sciences*, 3(9), 1–6. doi: 10.1186/s42409-021-00029-4.

Harkness, J.A., Villar, A., Edwards, B. (2010). Translation, adaptation and design. In *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, Harkness, J.A., Braun, M., Edwards, B. et al. (eds). Wiley, Hoboken.

Hayes, A.F. (2018). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach*, 2nd edition. Guilford Press, New York.

Helbling, M., Maxwell, R., Trautmüller, R. (2020). Government authority and legitimacy in the age of a pandemic. Sub-module proposal for ESS Round 10 Covid-19 module [Online]. Available at: <http://www.europeansocialsurvey.org/docs/about/ESSProposalHelblingMaxwellTrautmueller.pdf>.

Hu, L. and Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives.

Structural Equation Modeling, 6(1), 1–55. doi: 10.1080/10705519909540118.

Kish, L. (1994). Multi-population survey designs: Five types with seven shared aspects. *International Statistical Review*, 62(2), 167–186.

Kline, R.B. (1998). *Principles and Practice of Structural Equation Modeling*. Guilford, New York.

Marsh, H.W., Hau, K.T., Wen, Z. (2004). In search of golden rules: Comment on hypotheses-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11(3), 320–341. doi: 10.1207/s15328007sem1103_2.

Preacher, K.J. and Hayes, A.F. (2008). Contemporary approaches to assessing mediation in communication research. In *The Sage Sourcebook of Advanced Data Analysis Methods for Communication Research*, Hayes, A.F., Slater, M.D., Snyder, L.B. (eds). Sage, Thousand Oaks, CA.

Reiersen, J., Roll, K., Williams, J.D., Carlsson, M. (2022). Trust: A double-edged sword in combating the Covid-19 pandemic? *Frontiers in Communication*, 7, 822302.

Tabachnick, B.G. and Fidell, L.S. (2007). *Using Multivariate Statistics*. Pearson Allyn & Bacon, Upper Saddle River, NJ.

The ESS Sampling Expert Panel (2016). Sampling guidelines: Principles and implementation for the European Social Survey. ESS ERIC Headquarters, London [Online]. Available at: https://www.europeansocialsurvey.org/sites/default/files/2023-06/ESS8_sampling_guidelines.pdf

Toshkov, D., Carroll, B., Yesilkagit, K. (2022). Government capacity, societal trust or party preferences: What accounts for the variety of national policy responses to the Covid-19 pandemic in Europe? *Journal of European Public Policy*, 29(7), 1009–1028. doi: 10.1080/13501763.2021.1928270.

Note

Chapter written by Aggeliki YFANTI and Catherine MICHALOPOULOU.

6

Joint Forecasting of Screening/Randomization Processes in Multicenter Clinical Trials and Predicting Optimal Enrollment Stopping Time

Accurately forecasting patient enrollment is a critical component of planning and executing late-phase multi-center clinical trials. Traditional models, such as the Poisson and Poisson–gamma (PG) models, assume a direct pathway from patient arrival to randomization, which does not reflect the complex screening procedures in many real-world trials. In particular, multi-stage screening processes, involving pre-screening as well as screening, can produce substantial uncertainty and delay between a patient’s arrival and successful enrollment (randomization), making direct enrollment predictions prone to oversampling or inefficiency.

In this chapter, we present a novel analytic methodology for joint forecasting of screening/randomization processes that extends the basic PG model for a patient’s arrival to explicitly account for multi-stage screening, incorporating probabilistic drop-off at each stage and associated delays.

We propose a decision rule for stopping new arrivals before the number of randomized patients reaches sample size while minimizing both oversampling and wasting resources on non-randomized patients. This technique is demonstrated through detailed simulations, and predictive intervals are provided to support interim decision-making. This work offers a practical tool for enrollment management in trials with complex eligibility pipelines.

6.1. Introduction

Patient enrollment remains one of the most critical challenges in modern large multicenter clinical studies that aim to enroll hundreds/thousands of participants across diverse sites with varying performance, as typically it is required to achieve a rather large sample size to prove the efficiency of a particular drug. As noted by Ken Getz, Director of the Tufts Center for Study of Drug Development, USA, patient enrollment and retention are among the greatest challenges that the clinical research enterprise faces today, and they are a major cause of drug development delays.

To address this, the field has increasingly adopted predictive and adaptive approaches to monitor and model enrollment progress. A commonly used framework is to model patient arrivals as Poisson processes with fixed rates per site. This is seen in early works such as Carter et al. ([2005](#)) and Senn ([1997](#), [1998](#)).

However, in practice, enrollment rates vary substantially between centers due to differences in capacity, geography and patient access. To capture this variability, Anisimov and Fedorov ([2007](#)) (see also Anisimov ([2011](#)); Anisimov et al. ([2007](#))) introduced the PG model, which incorporates a hierarchical Bayesian structure. Here, site-level enrollment follows a Poisson process with a rate drawn from a gamma distribution, allowing the modeling of heterogeneity across centers.

Currently, the originally developed PG model (Anisimov [2011](#); Anisimov and Fedorov [2007](#)) has emerged as one of the most widely used and robust frameworks for forecasting enrollment and in some recent papers is called “one of the most popular techniques” and an “industry standard” model.

This model was also extended to modeling trial operational characteristics (Anisimov [2020](#)), predicting event counts in oncology trials (Anisimov [2020](#)) and centralized statistical monitoring of the enrollment performance (Anisimov and Austin [2020](#)).

There are also other approaches to enrollment modeling described in the literature; however, they mainly deal with the analysis of global enrollment and therefore have some limitations as they typically require a rather large number of centers and patients and cannot be applied to predicting enrollment on center/country level (see survey

papers by Barnard et al. ([2010](#)), Heitjan et al. ([2015](#)), Anisimov ([2016](#)) and Gkioni et al. ([2019](#))).

Yet, many real-world trials involve *screening* and *pre-screening* stages prior to randomization to evaluate a patient's eligibility for a given trial, which are not explicitly captured by the traditional PG model. There can be a large number of patients in these stages at any given time, and failure to account for these in advance could lead to a large oversampling and delays. For example, stopping enrollment of new patients at the instant when the randomization target is reached, then the patients in screening at that instant would still have to complete screening, leading to oversampling and delays, or be withdrawn before randomization, leading to wasted resources on incomplete or unnecessary costly screening procedures. The existence of this issue can necessitate stopping new arrivals in advance, before the randomization target has been reached.

In this chapter, we extend the PG model to incorporate multi-stage screening pipelines. By explicitly modeling the screening process prior to randomization, we provide a realistic and flexible framework for trials where patient attrition occurs during eligibility assessment. We also introduce a stopping rule based on predictive inference, allowing investigators to stop the arrival of new patients earlier when a sufficient number of randomized patients is projected, accounting for those still in screening.

The chapter is organized as follows. The first part, sections [6.2–6.5](#), describes the methodology for modeling and forecasting jointly the arrival and screening processes for the case of screening and pre-screening stages and how to apply an earlier stopping rule, which was developed by the first author and presented at the Conference on Advances in Statistics (Anisimov [2025](#)) and the 21st ASMDA Conference, June 10–13, 2025, Piraeus, Greece. Section [6.6](#) deals with parameter estimation using interim data. The part devoted to simulation (section [6.7](#), Implementation in R), as well as the application of these results to some real-world studies, was completed by the second author. Section [6.8](#) describes some considerations for future work.

6.2. Structure of screening stages

Consider a randomized multicenter clinical trial in which each site may put a patient through a pre-screening and/or screening process. Patients can then be randomized through one of two pathways: they can either initially go directly to screening (primary screening) and then after time interval Δ_{scr} be randomized to trial with some probability; or they can first be filtered through an additional “pre-screening” stage and after interval Δ_{pre} pass into the screening process. Then, after interval Δ_{scr} this secondary screened patient will be randomized to trial with some probability. A flow diagram of this process can be seen in [Figure 6.1](#).

The patients within any of these screening stages at the time when new screening is stopped will contribute to the final number of randomized patients, and therefore should be accounted for to avoid unnecessary oversampling.

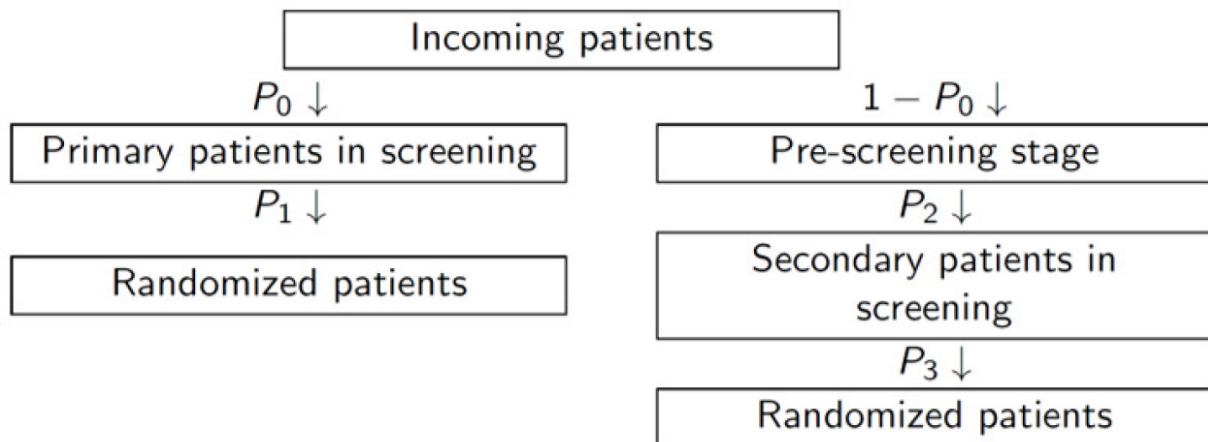


Figure 6.1. *Diagram of screening flow, showing individual stages, possible paths to randomization and corresponding probabilities passing from each stage to the next*

6.2.1. Stages of prediction

The prediction for an enrollment process with the structure described above includes two main stages:

1. Considering only the number of patients currently in the screening process, predict the final number of patients that will

be randomized from this pool. If there is a sufficiently high probability of this number surpassing the remaining number of patients yet to be randomized, then it can be recommended to stop the arrival of new patients for screening.

2. If the decision is not to stop, predict the process of patients being randomized over time, accounting for future arrivals of new patients into the screening process as well as patients currently in screening.

6.3. Predicting randomized patients from screening pool

Consider some interim time t_1 and assume we have the following data:

- n_1 : the number of patients currently in screening process (primary patients in screening);
- n_2 : the number of patients currently in pre-screening process;
- n_3 : the number of patients currently in screening process after successful pre-screening process (secondary patients in screening).

Assume that the probabilities of success for each screening stage are known and denote as follows:

- p_0 : probability of sending a patient arrived at clinical center directly to screening;
- p_1 : probability of successful randomization of patients currently in screening (primary screening patients);
- p_2 : probability of successful pre-screening of patients currently in pre-screening;
- p_3 : probability of successful randomization of patients in screening after pre-screening process (secondary screening patients).

Assume that the duration of screening for primary and secondary patients is Δ_{scr} and the duration of pre-screening is Δ_{pre} .

Denote by $\text{Bin}_j(n_j, p_j), j = 1, 2, \dots$ the independent binomial random variables with size n_j and probability p_j .

Following the structure of screening stages, the total number of successfully randomized patients out of patients in screening, pre-screening and secondary patients in screening after time $\Delta = \Delta_{pre} + \Delta_{scr}$ can be expressed as the following combination of random variables:

$$Y = \text{Bin}_1(n_1, p_1) + \text{Bin}_2(\text{Bin}_3(n_2, p_2), p_3) + \text{Bin}_4(n_3, p_3) \quad [6.1]$$

Using Y , we can now implement a *decision rule*, on whether or not the process of new arrivals should be stopped.

Typically, the decision rule may have the form: we need to check if

$$\mathbf{P}(Y > L) \geq \delta_1 \quad [6.2]$$

where L is the remaining number of patients left to be randomized and δ_1 is the desired confidence level where typically we can use $\delta_1 = 0.99$.

There may be different approaches to verify relation [6.2].

6.3.1. *Distributional approach*

To calculate the distribution of Y computationally, as it is the sum of independent random variables, we can use a convolution of the individual distributions to directly calculate the distribution of this sum.

It can be easily seen that a hierarchic variable $\text{Bin}_2(\text{Bin}_3(n_2, p_2), p_3)$ by distribution is equivalent to $\text{Bin}(n_2, p_2 \cdot p_3)$. Then Y is a sum of three independent binomial random variables, and to calculate the global distribution of Y , we can use the Base-R function `convolve()`.

6.3.2. *Normal approximation*

The mean and variance of Y can be calculated as the sum of the means and variances of the individual components:

$$E[Y] = n_1 \cdot p_1 + n_2 \cdot p_2 \cdot p_3 + n_3 \cdot p_3$$

$$\text{Var}[Y] = n_1 \cdot p_1 \cdot (1 - p_1) + n_2 \cdot p_2 \cdot p_3 \cdot (1 - p_2 \cdot p_3) + n_3 \cdot p_3 \cdot (1 - p_3)$$

Using these values, we can evaluate relation [6.2] using a normal approximation.

6.3.3. Monte Carlo

The distribution of Y can also be approximated nearly precisely using Monte Carlo simulation by simulating the binomial variables $\text{Bin}_i(\cdot, \cdot)$ in each run, calculating Y and repeating this procedure over 10^6 runs. This process will yield a sample of different realizations of Y , from which the mean, variance, various summary statistics and probabilities can be calculated.

All three approaches were compared using a few test examples and showed practically the same results. The distributional approach and Monte Carlo simulation provide the same results, as expected. This confirms the correctness of derivations and algorithms.

6.4. Predicting the number of randomized patients

If at some interim time t_1 a decision is not to stop screening, the next stage is predicting the process of patients being randomized over time, accounting for future arrivals of new patients into the screening process as well as patients currently in screening.

There are two independent directions:

- Predicting for $t > t_1$ the process $Y(t)$ of the number of randomized patients out of patients (n_1, n_2, n_3) who are currently at different screening stages.

- Predicting for $t > t_1$ the process $X(t)$ of new patients incoming for screening and how these patients will be distributed between screening stages and randomized later.

The process $Y(t)$ incorporates contributions from three independent patient flows over time:

$$Y(t) = Y_1(t) + Y_2(t) + Y_3(t),$$

where:

- $Y_1(t)$: randomized patients from n_1 patients in primary screening;
- $Y_2(t)$: randomized patients from n_2 patients in pre-screening;
- $Y_3(t)$: randomized patients from n_3 patients in secondary screening.

These components can be addressed individually, allowing the mean and variance of $Y(t)$ to be calculated in a closed form.

6.4.1. Randomization process $Y_1(t)$ from primary screening

The process $Y_1(t), t > t_1$, can be represented as:

$$Y_1(t) = \sum_{i=1}^{n_1} I(u_i \leq t) \cdot Z_i,$$

where u_i are the screening completion times in interval $[t_1, t_1 + \Delta_{scr}]$, and $Z_i \sim \text{Bernoulli}(p_1)$ indicates whether the i th patient is successfully randomized.

The future screening completion times u_i in general are not known exactly as even when we know the times of arrivals of patients at screening stage in the interval $[t_1 - \Delta_{scr}, t_1]$, in practice there is

variation in the duration of screening. Thus, we can assume that these times are uniformly distributed in $[t_1, t_1 + \Delta_{\text{scr}}]$.

6.4.2. Mean and variance of $Y_1(t)$

– For $t \in [t_1, t_1 + \Delta_{\text{scr}}]$:

Each u_i is uniformly distributed in $[t_1, t_1 + \Delta_{\text{scr}}]$, so:

$$P(u_i \leq t) = \frac{t - t_1}{\Delta_{\text{scr}}}, \quad \text{for } t \in [t_1, t_1 + \Delta_{\text{scr}}].$$

Taking expectation:

$$\mathbf{E}[Y_1(t)] = \sum_{i=1}^{n_1} \mathbf{E}[I(u_i \leq t)Z_i] = \sum_{i=1}^{n_1} P(u_i \leq t) \cdot p_1 = n_1 p_1 \frac{t - t_1}{\Delta_{\text{scr}}}. \quad [6.3]$$

For variance, using independence of u_i and Z_i and noting that $I(u_i \leq t)Z_i$ is a Bernoulli random variable with parameter $p_1(t - t_1)/\Delta_{\text{scr}}$:

$$\mathbf{Var}[Y_1(t)] = n_1 p_1 \frac{t - t_1}{\Delta_{\text{scr}}} \left(1 - p_1 \frac{t - t_1}{\Delta_{\text{scr}}} \right). \quad [6.4]$$

– For $t > t_1 + \Delta_{\text{scr}}$:

Now all n_1 patients have completed screening, each with probability p_1 , thus $Y_1(t)$ is a binomial variable with:

$$\mathbf{E}[Y_1(t)] = n_1 p_1 \quad [6.5]$$

$$\mathbf{Var}[Y_1(t)] = n_1 p_1 (1 - p_1). \quad [6.6]$$

Similar calculations can be done for $Y_2(t)$ and $Y_3(t)$, continuing with the assumption of uniformly distributed screening/pre-

screening completion times. Namely, completion times for patients in pre-screening n_2 are uniformly distributed in the interval $[t_1, t_1 + \Delta_{\text{pre}}]$; completion times for patients in secondary screening n_3 are uniformly distributed in the interval $[t_1, t_1 + \Delta_{\text{scr}}]$.

Using [6.3]–[6.6], we can compute other components of $Y(t)$ as follow.

6.4.3. $Y_2(t)$ and $Y_3(t)$

– For $t \in [t_1, t_1 + \Delta_{\text{scr}}]$:

$$Y_2(t) = 0$$

– For $t \in [t_1 + \Delta_{\text{scr}}, t_1 + \Delta_{\text{pre}} + \Delta_{\text{scr}}]$:

$$\mathbf{E}[Y_2(t)] = n_2 p_2 p_3 \frac{t - (t_1 + \Delta_{\text{scr}})}{\Delta_{\text{pre}}}$$

$$\mathbf{Var}[Y_2(t)] = n_2 p_2 p_3 \frac{t - (t_1 + \Delta_{\text{scr}})}{\Delta_{\text{pre}}} \left(1 - p_2 p_3 \frac{t - (t_1 + \Delta_{\text{scr}})}{\Delta_{\text{pre}}} \right)$$

– For $t > t_1 + \Delta_{\text{pre}} + \Delta_{\text{scr}}$:

$$\mathbf{E}[Y_2(t)] = n_2 p_2 p_3$$

$$\mathbf{Var}[Y_2(t)] = n_2 p_2 p_3 (1 - p_2 p_3)$$

The expressions for $Y_3(t)$ have the same form as for $Y_1(t)$, but n_1 and p_1 should be replaced by n_3 and p_3 .

The mean and the variance of the overall process $Y(t)$ are then calculated analytically as:

$$\begin{aligned}\mathbf{E}[Y(t)] &= \mathbf{E}[Y_1(t)] + \mathbf{E}[Y_2(t)] + \mathbf{E}[Y_3(t)], \\ \mathbf{Var}[Y(t)] &= \mathbf{Var}[Y_1(t)] + \mathbf{Var}[Y_2(t)] + \mathbf{Var}[Y_3(t)].\end{aligned}$$

6.5. Randomization process from future arrivals

Consider now predicting the process of randomized patients coming from the new patients arriving for screening.

6.5.1. Arrival process

To model the arrival process, we use a widely adopted PG model developed in (Anisimov [2011](#); Anisimov and Fedorov [2007](#)). Assume that at interim time t_1 , we have a set of interim data: for every center, k_i is the cumulative number of arrivals at screening in that center, and v_i is the enrollment duration in the center.

Consider a prediction in some country or region I . Assume also for simplicity that all centers from I belong to the same pool of centers (that means, the enrollment rates are gamma distributed with the same parameters). Denote by (α, β) the parameters of a PG model estimated using technique in Anisimov ([2011](#)); Anisimov and Fedorov ([2007](#)). Note that for a reliable estimation, the country I should contain a rather large number of centers (at least 15–20 is recommended).

Then for any center $i \in I$ in this country, the posterior rate $\hat{\lambda}_i$ with parameters (α_i, β_i) , where $\alpha_i = \alpha + k_i$, $\beta_i = \beta + v_i$.

The global posterior rate for the arrival process $X(t)$, $t > t_1$, in country I is:

$$\Lambda(I) = \sum_{i \in I} \hat{\lambda}_i$$

Denote the mean and the variance of this rate as

$$E(I) = \sum_{i \in I} (\alpha + k_i) / (\beta + v_i), \quad V^2(I) = \sum_{i \in I} (\alpha + k_i) / (\beta + v_i)^2 \quad [6.7]$$

Using techniques developed in Anisimov (2020); Anisimov and Austin (2020), for $t > t_1$, the predictive arrival process $X(t)$ can be approximated by a PG process $\hat{X}(t)$, which is a mixed Poisson process with gamma distributed rate $\hat{\Lambda}(I)$ with mean $E(I)$ and variance $V^2(I)$.

6.5.2. Randomization process

Consider now a construction of a randomization process, which is defined by the process of screening of incoming patients. This process consists of two stages:

1. $U_1(t)$: the patients randomized from primary screening patients sent directly to screening with probability p_0 . These patients will be successfully randomized with probability p_1 and appear in the trial with delay Δ_{scr} .
2. $U_2(t)$: the patients randomized from secondary patients sent to screening after a successful pre-screening stage. These patients are required to successfully pass a few steps: first, upon arrival, they are sent to pre-screening with probability $1 - p_0$; then with probability p_2 , each patient can successfully complete pre-screening and will be sent to the secondary screening stage; after this, with probability p_3 , these patients can successfully complete secondary screening and be randomized. These randomized patients will appear in the trial only with delay $\Delta = \Delta_{pre} + \Delta_{scr}$.

That means for each arriving patient the probabilities to successfully pass either these two stages are:

1. Stage 1, $P_1 = p_0 p_1$;

2. Stage 2, $P_2 = (1 - p_o)p_2p_3$.

Therefore, the process $U_1(t)$ is a PG process with rate $P_1\Lambda(I)$ and delayed with lag Δ_{scr} ; the process $U_2(t)$ is a PG process with rate $P_2\Lambda(I)$ and delayed with lag Δ .

As there are two lags, there are two parallel flows of new randomized patients after time t_1 , and for any $t > t_1$:

$$U_1(t) = \Pi_{P_1\Lambda(I)}[t - \Delta_{scr}]_+, \quad U_2(t) = \Pi_{P_2\Lambda(I)}[t - \Delta]_+$$

where $[a]_+ = \max(0, a)$, $\Pi_A(t)$ defines a Poisson process with rate A , and the arrival rate $\Lambda(I)$ is the same for both flows.

Thus, the global process $R(t)$ of new randomized patients from the patients arriving for screening is represented as $R(t) = U_1(t) + U_2(t)$, $t > t_1$.

Correspondingly, a cumulative rate $\Lambda_R(t)$ of $R(t)$ is:

$$\Lambda_R(t) = \Lambda(I) \left(P_1[t - \Delta_{scr}]_+ + P_2[t - \Delta]_+ \right)$$

with mean and variance

$$E_R(t) = E(I) \left(P_1[t - \Delta_{scr}]_+ + P_2[t - \Delta]_+ \right) \quad [6.8]$$

$$V_R^2(t) = V^2(I) \left(P_1[t - \Delta_{scr}]_+ + P_2[t - \Delta]_+ \right)^2.$$

Therefore, for $t > t_1$, the process $R(t)$ can be approximated by a PG process with the mean and the variance of cumulative rate, $E_R(t)$ and $V_R^2(t)$.

6.5.3. Global randomization process

Recall that we also have the process $Y(t)$ of patients randomized from screening, addressed in section [6.4](#). Thus, the overall global

randomization process is:

$$R_{Tot}(t) = Y(t) + R(t). \quad [6.9]$$

Now the mean and the variance of $R_{Tot}(t)$ can be calculated in the closed form by combining the earlier results for $\mathbf{E}[Y(t)]$ and $\mathbf{Var}[Y(t)]$ with [6.8].

We can then either use a normal approximation and predict this process with low and upper bounds and correspondingly evaluate the time $\tau(n)$ to reach a particular target n with low and upper bounds, or use a PG approximation as described in Anisimov and Austin (2020), when the number of centers is not so large.

Note that a prediction over time alone is not sufficient to decide to stop screening new patients – only the stopping rule [6.2] can determine this more precisely.

6.6. Parameter estimation

It is likely that the PG parameters (α, β) , the probabilities of passing the screening stages (p_1, p_2, p_3) as well as the probability of being initially assigned to screening/pre-screening (p_0) , are not known in advance, therefore will have to be estimated from interim data, collected at t_1 . Assume these probabilities are the same for all centers. As there are multiple binomial processes occurring, introduce a secondary subscript, $j = 1, 2, 3$, to notate j th screening stage. Assume that for each center, the following set of values $(n_{i,j}, k_{i,j}, s_{i,j}, v_i)$ is available, where at t_1 :

- $n_{i,j}$ is the number of patients in the j th screening stage at center i ;
- $k_{i,j}$ is the total number of patients that have at some point entered stage j at center i ;
- $s_{i,j}$ is the number of patients that have successfully passed stage j at center i ;

- v_i is the duration of active enrollment in center i from the initiation time until the interim time t_1 , which does not depend on index j .

The cumulative number of arrivals for a center up to time t_1 we denote as $k_{i,0}$, where $k_{i,0} = k_{i,1} + k_{i,2}$. Also note that at any given time, $k_{i,3}$ (total patients entered into secondary screening) is equal to $s_{i,2}$ (total patients succeeded pre-screening) due to the screening structure.

Then the likelihood function has the following form:

$$\prod_{i=1}^N \left[\mathbf{P}(\text{Pg}(\alpha, \beta, v_i) = k_{i,0}) \mathbf{P}(\text{Bin}(k_{i,0}, p_0) = k_{i,1}) \prod_{j=1}^3 \mathbf{P}(\text{Bin}(k_{i,j} - n_{i,j}, p_j) = s_{i,j}) \right]$$

Here, the variables (α, β) can be separated and ML estimators for (α, β) can be found in the same way as for a PG model in (Anisimov [2011](#); Anisimov and Fedorov [2007](#)).

For the multiplier $\prod_{i=1}^N \mathbf{P}(\text{Bin}(k_{i,0}, p_0) = k_{i,1})$, taking the log-likelihood and equating the derivative in p_0 to zero, we find the estimator:

$$\hat{p}_0 = \frac{\sum_{i=1}^N k_{i,1}}{\sum_{i=1}^N k_{i,0}}.$$

Similarly for the multiplier

$\prod_{i=1}^N \prod_{j=1}^3 \mathbf{P}(\text{Bin}(k_{i,j} - n_{i,j}, p_j) = s_{i,j})$, for each j , taking the log-likelihood and equating the derivative in p_j to zero, we get:

$$\hat{p}_j = \frac{\sum_{i=1}^N s_{i,j}}{\sum_{i=1}^N (k_{i,j} - n_{i,j})} \quad \text{for } j = 1, 2, 3$$

6.7. Implementation in R tool

The methodology described in this chapter aims to reduce over-sampling during monitoring the screening process and provide reliable guidance on closing new arrivals toward the end of the study. Here, we demonstrate how this may be implemented in practice. We first must generate dummy data for testing.

6.7.1. Algorithm of simulation in R

Consider a single region with N centers, where the rate λ_i in each center i is drawn from a Gamma distribution with parameters (α, β) , and u_i is the initiation time of center i . Denote vector $\text{vecu}=(u_i)$. Define values for p_o and p_j . The enrolment process follows the process described in section [6.2](#) for a given duration of T days.

1. Generate values of λ_i : `vecla=rgamma(N, shape= $\alpha$ , rate= $\beta$ ).`
2. For each center i , given `vecla[i]`, generate a sample of independent Poisson variables, `rpois(T,vecla[i])`, storing the results in a $N \times T$ matrix. Account for initiation times by setting `Xtot[i, 1:vecu[i]] <- 0`. This is the arrival matrix `Xtot`. Apply a binomial filter to each element of `Xtot` using `rbinom` with probability p_o to create primary screening and pre-screening arrival matrices `Xscr` and `Xpre`.
3. Looping through days $1:T$, create the $N \times T$ matrices of numbers in screening for each stage `Qscr` and `Qpre` and populate them by summing the corresponding arrival matrix in the relevant preceding window, e.g. `Qscr[, day] <- rowSums(X[, (screen_window+1):day])`, where `screen_window <- max(0, day- $\Delta_j$ )`, Δ_j is the screening delay.
4. Looping through days, create the $N \times T$ success matrices for each process `Yscr` and `Ypre` and populate them by applying a binomial filter to each arrival day, and adding the result to the success matrix with the relevant lag of Δ_j , e.g. `Y[, day+ $\Delta_{scr}$ ] <- sapply(Xscr[, day], function(x) rbinom(1, x,  $p_1$ ))`.

5. To get similar matrices for secondary sceening, repeat steps 3-4, using Y_{pre} as the corresponding arrival matrix X_{scr2} .
6. Convert all the arrival (X) and success (Y) matrices to a cumulative count using the `cumsum()` function.

These matrices give us the values $(n_{i,j}, k_{i,j}, s_{i,j}, v_i)$ at any interim time $1 \leq t_1 \leq T$, allowing us to estimate probabilities as shown in section [6.6](#) and evaluate the stopping rule [[6.2](#)].

6.7.2. Stopping new arrival

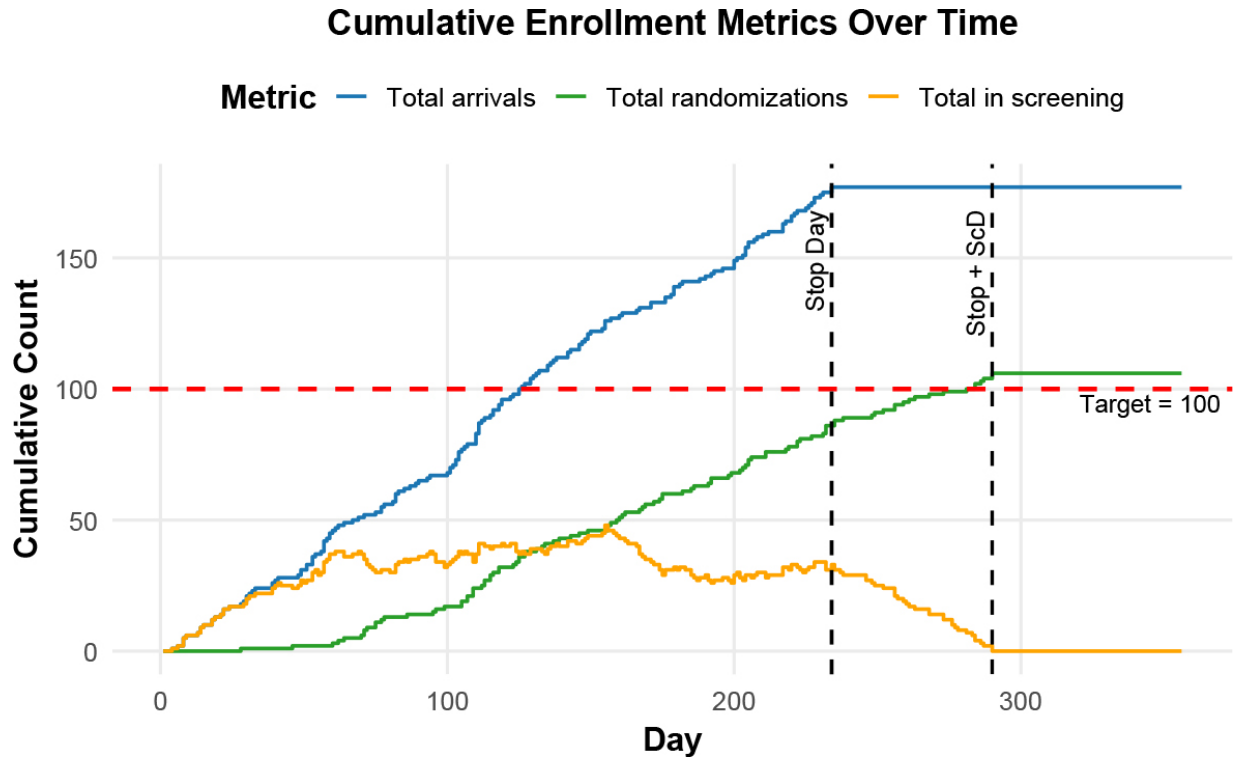


Figure 6.2. Example implementation of screening rule, showing metrics of total arrivals, randomization and total of all screening/pre-screening stages. The parameters of the simulation were: $N = 50$, $(\alpha, \beta) = (1.2, 70)$, $(p_0, p_1, p_2, p_3) = (0.2, 0.6, 0.8, 0.7)$, $(\Delta_{scr}, \Delta_{pre}) = (2 \text{ weeks}, 6 \text{ weeks})$, $vecu = rep(0, N)$ and a target of 100 patients to randomize. The stopping rule was satisfied on $t = 234$ days, with confidence $\delta_1 = 0.99$, $n_1 + n_2 + n_3 = 29$ patients across all screening stages and remaining patients $L = 14$, leading to a final randomization of 106 patients. The target of 100 was reached on $t = 281$ days.

To identify the earliest possible stopping time, the process of estimating probabilities is applied to every interim time from $[1, T]$. The first day, when [6.2] is evaluated as “true”, we consider it as `stop_day` and cease all future arrivals with `Xtot[, stop_day] <- 0` and let the simulation play out to observe the outcome. An example of this can be seen in [Figure 6.2](#).

As expected, the process of new arrivals was stopped before the target was reached (labeled “Stop day”), and the number in screening made up the remaining amount of randomized patients over

screening duration $\Delta_{scr} + \Delta_{pre}$ (labeled ScD). There is still some oversampling, as is almost inevitable due to designing the stopping rule with a high confidence δ_1 . The stopping rule was satisfied on $t = 234$ days, with confidence $\delta_1 = 0.99$, leading to a final randomization of 106 patients. The target of 100 was reached on $t = 281$ days.

There are many fluctuating statistics involved in evaluating the stopping decision, so it is interesting to see how the stopping day can vary. To demonstrate, we repeat 1,000 simulation runs with identical parameters used in [Figure 6.2](#) and record the stopping day for each run. A histogram of this is shown in [Figure 6.3](#). The variation of stopping day can be large, so it is suggested to evaluate the decision rule [\[6.2\]](#) often, as soon as at least 10–15 patients have passed each screening stage, for estimating transition probabilities more reliably.

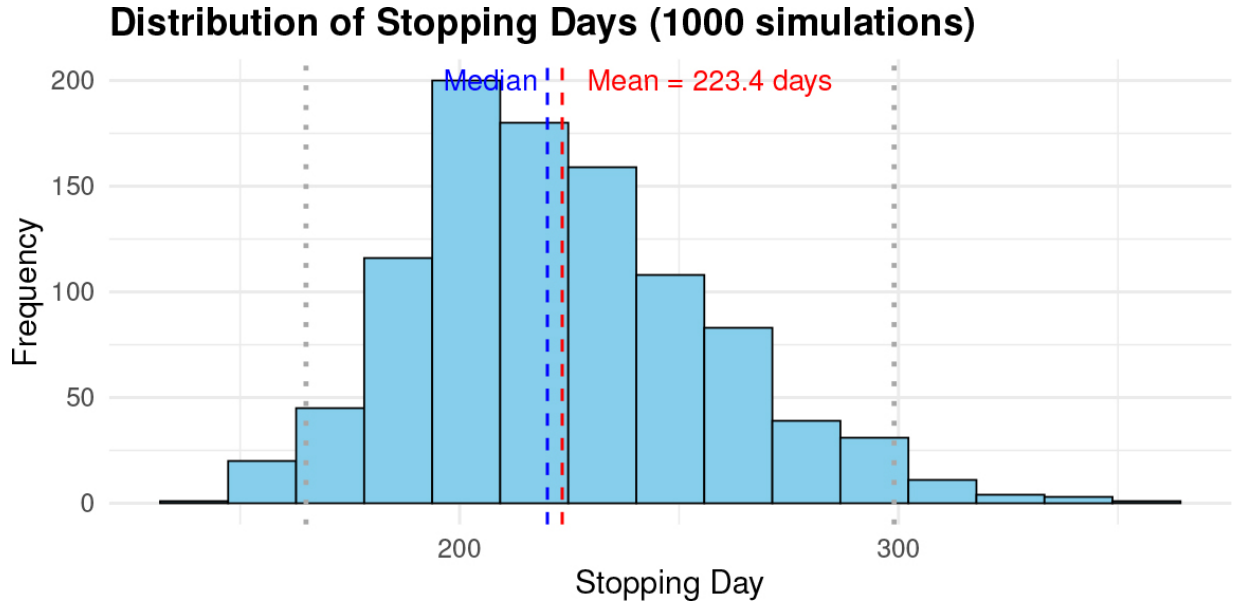


Figure 6.3. Histogram showing results of advised stopping day over 1,000 simulation runs with parameters the same as given in the caption of [Figure 6.2](#). The mean and median are 223.4 and 220 days. The central 95% interval is shown by the dotted lines as between 165 and 299 days.

6.7.3. Predicting enrollment

While the stopping rule is the only deciding factor for stopping new enrollment, it is helpful to have a prediction ahead of time of when

the target will be reached for purposes of planning. Anytime the tool is run, it will also output a prediction over time, using the methodology described in section 6.5 and estimations outlined in section 6.6. To demonstrate this output, we use the same simulation run as in Figure 6.2, but instead run the tool at interim time $t_1 = 200$ days. At this point, there are remaining $L = 32$ patients. A normal approximation is used in equation [6.9], predicting the mean enrollment with a 90% central confidence interval. The tool's predictive output is shown in Figure 6.4.

Predicting randomized patients over time

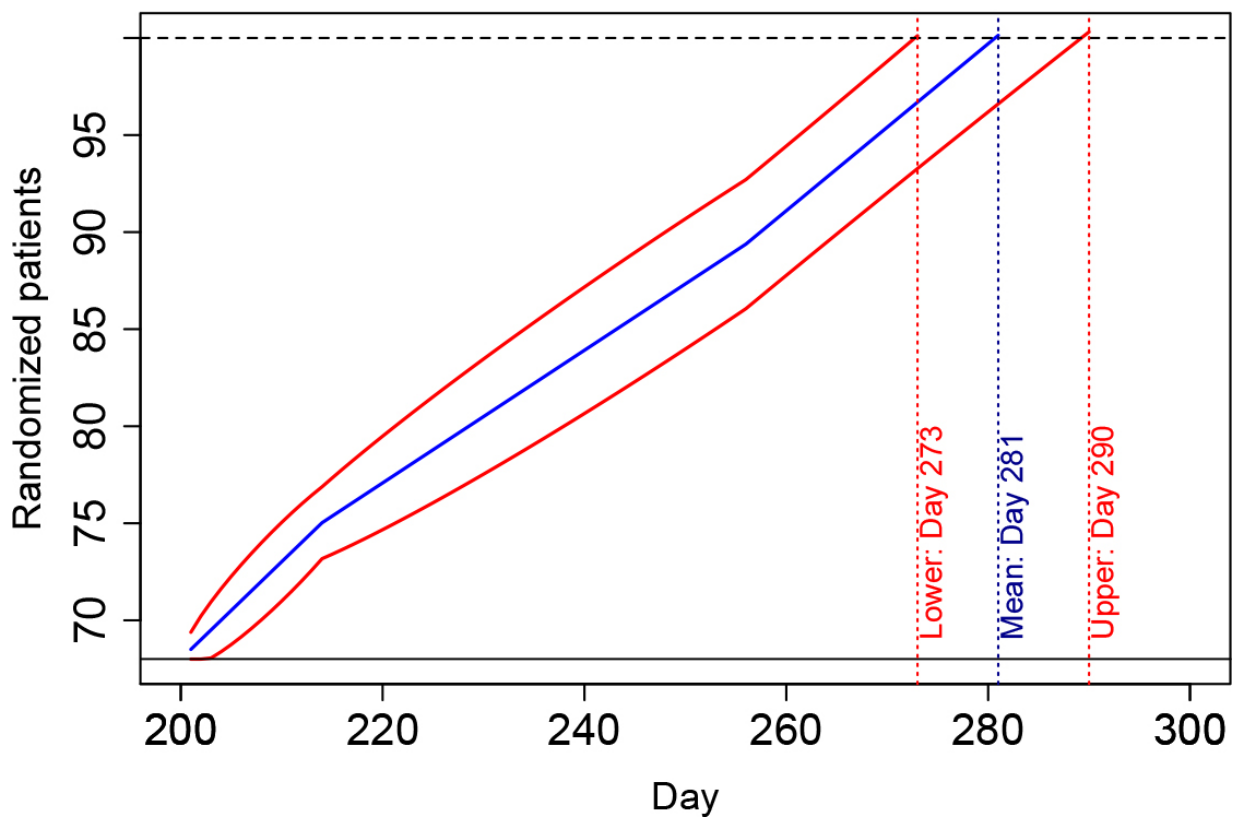


Figure 6.4. Example of a predicted enrollment trajectory, using same the simulation data as in Figure 7.1, but taken at interim time $t_1 = 200$ days. The mean of the normal approximation is shown with a 90% central confidence interval, and these trajectories are extended to meet the target of 100, giving bounds for enrollment times.

The predicted trajectory with bounds is used to evaluate the mean, lower and upper bounds of the completion time, which are 281, 273 and 290 days, respectively. This agrees with the true completion time of $t = 281$ days, shown in [Figure 6.2](#). There is a change of gradient observable at $t = t_1 + \Delta_{scr}$ days: this is explained by there being a larger number of patients in the primary and secondary screening stages than would be expected given the estimated parameters.

Thus, these patients deliver more randomized patients than expected. This is a common feature of these predictions, due to the statistical fluctuation of patients in each screening stage, and is also visible as a minor gradient change at time $t = t_1 + \Delta_{scr} + \Delta_{pre}$ due to the same reason.

6.7.4. Applications

This tool can also be adapted to evaluate the decision rule for a study with multiple distinct regions/countries, where the probabilities of passing each stage are individually estimated for each region, resulting in a unique Y_J (predicted number of randomized patients from those in screening) for each region J . These can then be summed using the techniques described in [section 6.3](#) to evaluate the global distribution. This tool was applied to a real study with the described screening/pre-screening structure and multiple regions, and successfully helped to monitor the screening process and advise on closing new arrivals and minimizing patient oversampling.

6.8. Future considerations

In this chapter, we discussed a specific pre-screening/screening procedure, however there can also be other designs. For example, a simpler design like a single screening stage, or a more complex one where patients can have some probability of re-screening if they fail a screening process initially.

We also consider only point estimates of the parameters, but it could be more accurate to model the variability in these estimators using full probability distributions or Bayesian approaches, for example, by treating the probability estimates as random variables drawn from

beta distributions. Incorporating this would provide more robust predictions and better reflect the uncertainty in enrollment projections, which is especially important for decision-making in trial monitoring and adaptive designs.

Another point is that using fixed delays Δ_{scr} and Δ_{pre} for both failures and successes may not always be accurate and it can be advisable to account in future considerations for potential variability of these variables.

6.9. Acknowledgment

The authors would like to express their sincere gratitude to Matt Austin and the Data Science Team at the Center for Design & Analysis and Operational Design Team, Amgen, for constant and useful discussions on this subject.

6.10. References

- Anisimov, V. (2011). Statistical modeling of clinical trials (recruitment and randomization). *Communications in Statistics – Theory and Methods*, 40(19–20), 3684–3699.
- Anisimov, V. (2016). Discussion on the paper “Real-time prediction of clinical trial enrollment and event counts: A review” by D.F. Heitjan et al. *Contemporary Clinical Trials*, 40, 7–10.
- Anisimov, V. (2020). Modern analytic techniques for predictive modeling of clinical trial operations. In *Quantitative Methods in Pharmaceutical Research and Development*, Marchenko, O.V. and Katenka, N.V. (eds). Springer, Cham, 361–408. doi: 10.1007/978-3-030-48555-9_8.
- Anisimov, V. (2025). Forecasting the process of screened/randomized patients and optimal early stopping time in multicentre clinical trials. In *Proc. of 11th International Conference on Advances in Statistics*, Bologna, 25–27 April.
- Anisimov, V. and Austin, M. (2020). Centralized statistical monitoring of clinical trial enrollment performance.

Communications in Statistics – Case Studies and Data Analysis, 6(4), 392–410.

Anisimov, V. and Fedorov, V. (2007). Modelling, prediction and adaptive adjustment of recruitment in multicentre trials. *Statistics in Medicine*, 26(27), 4958–4975.

Anisimov, V., Downing, D., Fedorov, V. (2007). Recruitment in multicentre trials: Prediction and adjustment. In *mODa 8 - Advances in Model-Oriented Design and Analysis*, López-Fidalgo, J., Rodríguez-Díaz, J.M., Torsney, B. (eds). Physica, Heidelberg, 1–8. doi: 10.1007/978-3-7908-1952-6_1.

Barnard, K.D., Dent, L., Cook, A. (2010). A systematic review of models to predict recruitment to multicentre clinical trials. *BMC Medical Research Methodology*, 10, 63.

Carter, R.E., Sonne, S.C., Brady, K.T. (2005). Practical considerations for estimating clinical trial accrual periods: Application to a multi-center effectiveness study. *BMC Medical Research Methodology*, 5, 11–15.

Gkioni, E., Riusd, R., Dodda S., Gamblea, C. (2019). A systematic review describes models for recruitment prediction at the design stage of a clinical trial. *Journal of Clinical Epidemiology*, 115, 141–149.

Heitjan, D.F., Ge, Z., Ying, G.S. (2015). Real-time prediction of clinical trial enrollment and event counts: A review. *Contemporary Clinical Trials*, 45(Part A), 26–33.

Senn, S. (1997). *Statistical Issues in Drug Development*. Wiley, Chichester.

Senn, S. (1998). Some controversies in planning and analysis multi-center trials. *Statistics in Medicine*, 17, 1753–1756.

Note

Chapter written by Volodymyr ANISIMOV and Leo GORTZAK.

PART 2

Applications

7

Bidding Strategies Approaches for the Greek Electricity Market

The growing penetration of renewable energy sources (RESs) in the Greek electricity market leads short-term flexibility mechanisms to be increasingly vital to maintain supply–demand balance. The pronounced RES curtailments and heightened price volatility observed in 2024 underscored the system’s limited capacity to adjust to real-time fluctuations. The forthcoming transition to a 15-min market time unit (MTU) in June 2025 seeks to enhance market responsiveness and further align Greece with the broader European market framework. This chapter evaluates the operational performance of flexibility mechanisms, and the relevant bidding strategies of stand-alone Battery Energy Storage Systems (BESS) and RES collocated with BESS units within the day-ahead market (DAM) and intraday market (IDM), emphasizing the comparative roles of structured intraday auctions (IDAs) and continuous trading via the XBID platform. Using real market data obtained from the Hellenic Energy Exchange (HEnEx) and the Independent Power Transmission Operator (IPTO), the analysis explores daily traded volumes, market responses to market clearing price (MCP) volatility and RES curtailment trends.

7.1. Introduction

The rapid integration of RESs into electricity systems has fundamentally reshaped both their operational dynamics and economic structures (Nikologiannis et al. [2025](#)). As the share of variable generation from wind and solar technologies expands, the demand for short-term flexibility becomes increasingly critical to sustaining system stability. Extensive research efforts have addressed this challenge (Karapidakis et al. [2024a](#)), aiming to enhance price stability, reliability and overall market efficiency.

In this evolving context, electricity markets must incorporate mechanisms that enable the prompt correction of forecast deviations and the optimal allocation of available capacity while simultaneously upholding competitiveness, social welfare and temporal efficiency (Acaroğlu and García Márquez [2021](#)). The Greek electricity market, administered by the HEnEx, has made substantial progress toward harmonization with the European Target Model through the establishment of DAM and IDM structures (Makrygiorgou et al. [2023](#); Karapidakis et al. [2024b](#)). Within these frameworks, participants are able to place bids in the DAM and subsequently adjust their portfolios via IDA1, IDA2 and IDA3, and continuous cross-border trading facilitated by the XBID platform. These mechanisms provide distinct levels of flexibility regarding timing, liquidity and responsiveness to fluctuations in demand and RES generation – spurring ongoing research across European electricity markets (Shinde et al. [2021](#); Alberizzi et al. [2023](#); Andreotti et al. [2024](#)).

Greece is set to introduce a 15-min MTU within both the single day-ahead coupling (SDAC) and single intraday coupling (SIDC) mechanisms by June 2025, a milestone expected to enhance temporal granularity and market adaptability. The shift from the current 60-min to a 15-min MTU will significantly improve the temporal resolution of the Greek electricity market, allowing participants to align generation and consumption more precisely with real-time system conditions. This refinement reduces exposure to forecast inaccuracies and lowers balancing costs, while also facilitating more effective integration of RES. By capturing variability within shorter intervals, the new structure encourages more responsive bidding behaviors in both the DAM and IDM. Moreover, this transition fully aligns Greece with the European Union's Target Model for market coupling, thereby improving market liquidity and cross-border interoperability.

Electricity market flexibility refers to the capability of the system to respond effectively and reliably to short-term deviations between forecasted and actual supply or demand. This feature is particularly vital in systems characterized by high RES penetration, where intermittency poses operational challenges. Flexibility can arise from supply-side adjustments (e.g. ramping of thermal units), demand-

side responses, energy storage deployment or cross-border trading. The layered architecture of the Greek market – comprising the DAM, IDAs and continuous trading through the XBID platform – provides a practical framework for assessing the contribution of each mechanism to grid stability and market efficiency.

This study examines specific, for both stand-alone and hybrid BESS with RES, power plants' bidding strategies of IDA transactions and XBID-based continuous trading as instruments of short-term flexibility within the Greek electricity market. Utilizing publicly available data from January to March 2025, the analysis explores the interactions among MCP volatility, renewable curtailment and system imbalances. By evaluating pretransition market dynamics, the research establishes a baseline for assessing how effectively the current market design supports short-term flexibility ahead of 15-min MTU implementation.

7.2. Methodology

In the present study, DAM results (EnExGroup n.d.a, n.d.b) were collected and systematically organized by category and date. The dataset includes hourly scheduled power quantities (in MW) for next-day transactions, encompassing locally generated electricity from various technologies sold to different classes of consumers, as well as cross-border exchanges with coupled markets – namely, North Macedonia, Italy, Bulgaria, Turkey and Albania. The purchasing entities comprise low-voltage (LV) and medium-voltage (MV) distribution networks, high-voltage (HV) transmission systems, pumping power requirements for hydroelectric plants (PUMPS) and system losses (LOSSES).

Each record is accompanied by the hourly MCP established for every trading hour of the subsequent day (in €/MW). Instances of MCP values recorded as “-0.01” were identified as data anomalies and subsequently corrected to “0”.

For analytical consistency, each category was further segmented into three subsamples corresponding to the respective years under investigation. The Pearson correlation coefficient was computed for

each power trading category to assess the degree of linear association with the MCP using the following expression:

$$\rho(X, Y) = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 * \sum (y - \bar{y})^2}} \quad [7.1]$$

where X and Y are the random variables, x and y are their elements and $\bar{x}$ and $\bar{y}$ are their mean values.

When the correlation coefficient deviated unexpectedly from the anticipated values, equation [7.2] was used for the moving average of the selected time series:

$$\bar{T}_t = \frac{1}{m} * \sum_{j=-k}^k x_{t+j} \quad [7.2]$$

where $m = 2*k+1$ is the order of the trend cycle.

Equation [7.3] is used for the sample mean of each random variable:

$$\sigma = \sqrt{\frac{1}{n-1} * \sum (x - \bar{x})^2} \quad [7.3]$$

Equation [7.4] is used for the standard error calculation:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad [7.4]$$

In addition, the skewness by [7.5]:

$$\gamma_1 = \frac{n}{(n-1) * (n-2)} * \sum \left(\frac{x - \bar{x}}{\sigma} \right)^3 \quad [7.5]$$

$$\gamma_2 = \frac{n * (n + 1)}{(n - 1) * (n - 2) * (n - 3)} * \sum \left(\frac{x - \bar{x}}{\sigma} \right)^4 - \frac{3 * (n - 1)^2}{(n - 2) * (n - 3)} \quad [7.6]$$

In parallel, this study utilizes detailed operational and pricing datasets encompassing daily MCPs, trading volumes from IDA1–3 and the XBID platform, as well as records of RES curtailments. The primary objective is to assess the effectiveness of IDM instruments in enhancing system flexibility and maintaining supply–demand balance amid increasing RES penetration and market volatility. To evaluate the rigidity and responsiveness of the DAM, the analysis begins by computing daily DAM price volatility (equation [7.7]), defined as the normalized difference between the daily maximum and minimum MCP values, whereas [Table 7.1](#) provides the key market and system variables.

$$Volatility_{Day} = \frac{MCP_{max,day} - MCP_{min,day}}{MCP_{Average,day}} \quad [7.7]$$

Table 7.1. Key metrics for DAM and IDM

Type	Unit	Mean	Min	Max	STd Dev
Avg. MCP	EUR/MWh	130.92	72.26	207.28	29.48
XBID Volume	MWh	15,565.87	7,528.50	33,031.20	5,025.36
IDA Volume	MWh	3,793.41	1,754.16	8,953.65	1,379.96
RES Curtailment	MWh	2,612.04	0.00	16,322.73	4,645.39

The analysis reveals a distinct hierarchy in the utilization of IDM mechanisms. Continuous trading through the XBID platform recorded substantially higher daily volumes (mean: 15,566 MWh; maximum: 33,031 MWh) compared to the IDAs, which averaged 3,793 MWh per day. This disparity indicates a strong preference among market participants for the immediacy and temporal

flexibility afforded by continuous trading, enabling more effective real-time position adjustments.

RES curtailment emerged as a notable operational challenge, averaging 2,612 MWh per day and peaking at over 16 GWh. The high standard deviation (4,645 MWh) suggests that curtailment events were not uniformly distributed, but were concentrated during periods of system stress. These results underscore the critical role of IDM – particularly the XBID mechanism – as a key flexibility instrument that mitigates the limitations of the DAM and absorbs variability associated with elevated RES penetration. It is worth mentioning that the persistence of curtailment highlights enduring structural constraints, mainly due to restricted storage capacity.

[Figures 7.1](#) and [7.2](#) present the daily trading volumes of IDAs and XBID transactions, respectively, illustrating their roles in facilitating real-time system balancing.

These trends highlight the pivotal role of the IDM – and particularly continuous trading via the XBID platform – as a key flexibility channel for mitigating the real-time effects of RES forecast deviations and market imbalances. The concurrent use of both XBID and IDAs reflects a multilayered flexibility strategy within the Greek electricity market, where structured and continuous mechanisms operate in parallel to manage uncertainty. Nevertheless, the continued occurrence of substantial RES curtailment, even during periods of elevated IDM activity, underscores persistent challenges related to grid congestion, forecasting accuracy and the sufficiency of flexibility resources within the transmission network.

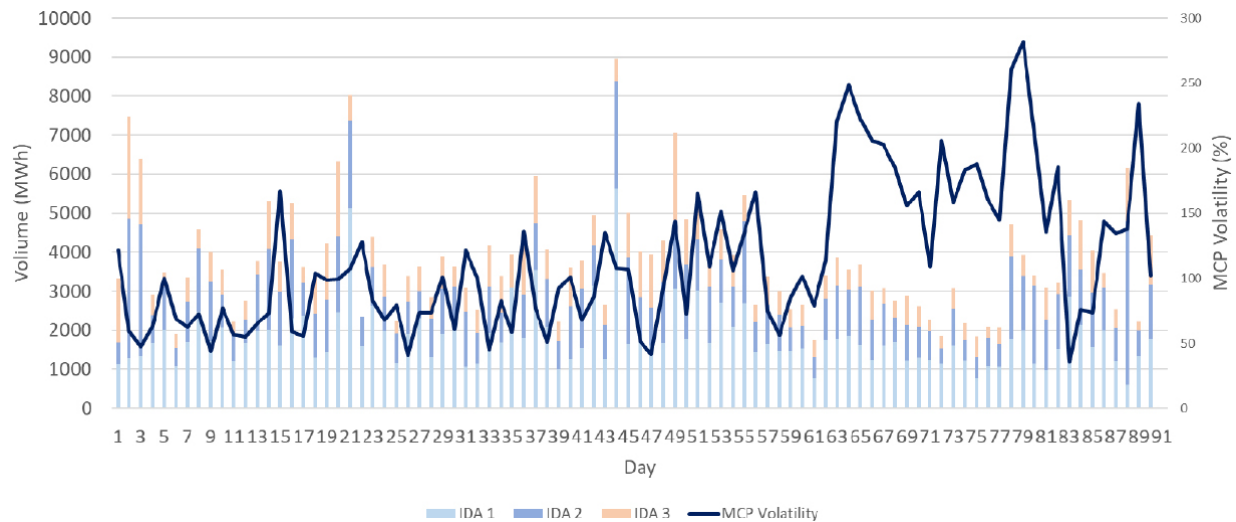


Figure 7.1. *Intraday (IDAs) auctions*

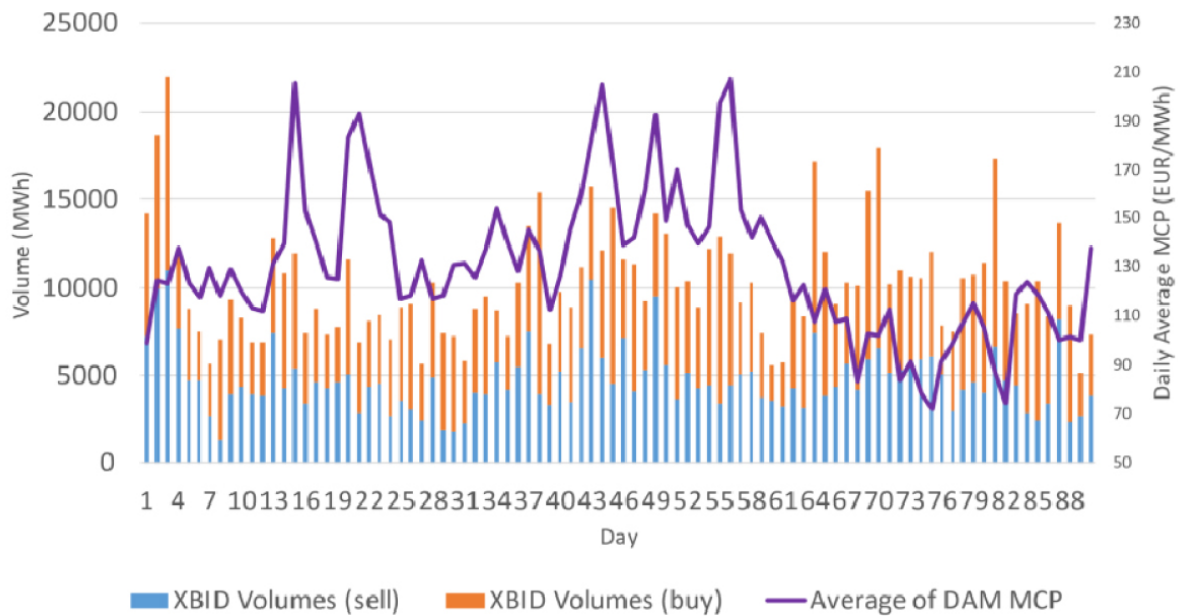


Figure 7.2. *XBID auctions*

[Figure 7.3](#) superimposes curtailment onto total IDM energy (including IDAs and XBID transactions), highlighting periods when intraday adjustments may have mitigated the need for forced RES disconnections.

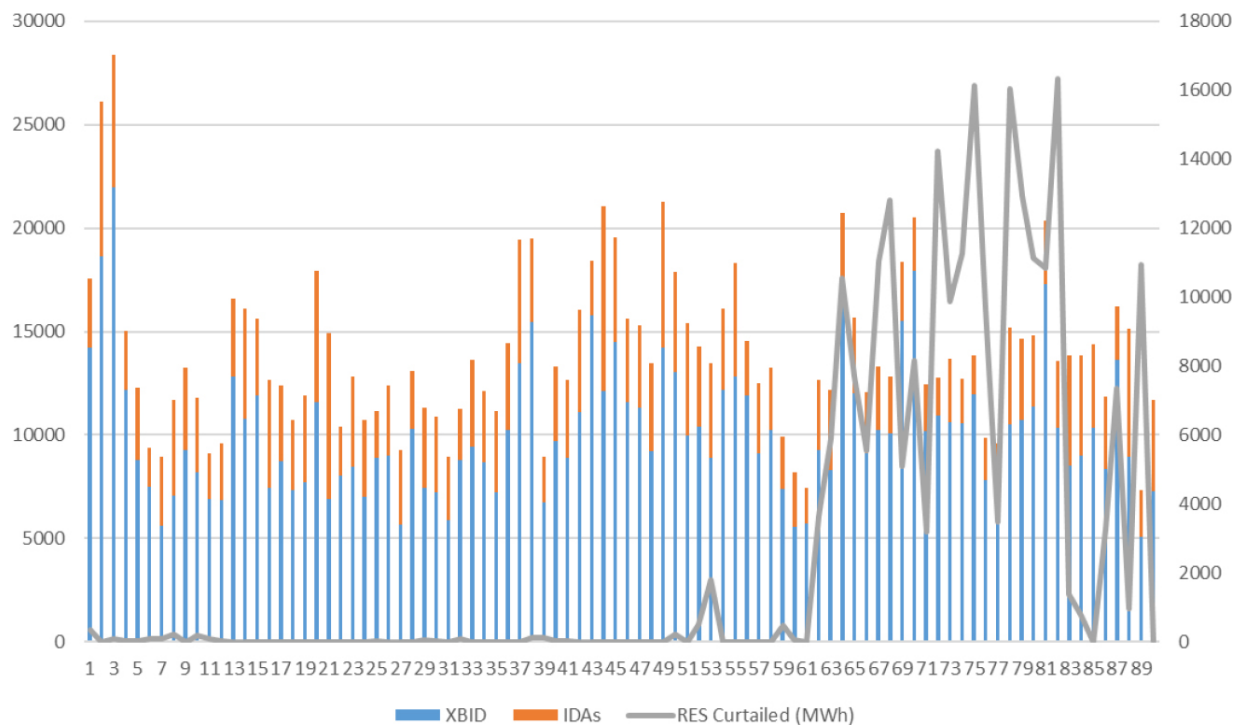


Figure 7.3. *Superpose of IDM and RES curtailment*

Table 7.2. *DAM, IDA and RES curtailment correlations.*

	XBID sell	XBID buy	IDAs	RES Curtailed	Average MCP
XBID sell	1.00				
XBID buy	0.82	1.00			
IDAs	0.28	0.27	1.00		
RES Curtailed (MWh)	0.19	0.18	-0.32	1.00	
Average MCP	0.06	0.00	0.51	-0.64	1.00

In conclusion, [Table 7.2](#) illustrates the correlations between MCP volatility and average daily MCP with IDA volumes, XBID trading volumes and RES curtailment. The analysis of correlations with average MCP indicates that IDA volumes display a moderate positive relationship with higher daily prices, reflecting a degree of sensitivity to price fluctuations.

Conversely, XBID volumes exhibit a weaker association with average MCP, suggesting that continuous trading activity is primarily influenced by underlying system conditions rather than price signals.

7.3. Stand-alone and hybrid BESS bidding strategies

The integration of BESS and hybrid RES–BESS units within the HEnEx offers new strategic pathways for optimizing both market performance and system flexibility. Building on the analysis of DAM and IDM dynamics presented in the previous sections, this part develops optimal bidding strategies for (a) stand-alone BESS and (b) collocated RES–BESS assets, emphasizing their operational and economic contributions under the forthcoming 15-min MTU regime.

7.3.1. Stand-alone BESS

For stand-alone BESS assets, strategic participation in the DAM and IDM focuses on price arbitrage, multimarket optimization and state-of-charge (SoC) management. The objective is to maximize daily profit by exploiting price spreads between charging and discharging periods while adhering to physical and market constraints.

The daily profit function Π_t for a storage unit participating in DAM and IDM can be expressed as:

$$\max_{\{P_t^{\text{ch}}, P_t^{\text{dis}}\}} \Pi_t = \sum_{t=1}^T \left(\lambda_t^{\text{sell}} P_t^{\text{dis}} \eta^{\text{dis}} - \lambda_t^{\text{buy}} P_t^{\text{ch}} / \eta^{\text{ch}} \right) \quad [7.8]$$

Subject to:

$$\begin{aligned} E_{t+1} &= E_t + \eta^{\text{ch}} P_t^{\text{ch}} \Delta t - \frac{1}{\eta^{\text{dis}}} P_t^{\text{dis}} \Delta t \\ 0 &\leq E_t \leq E^{\text{max}}, \quad 0 \leq P_t^{\text{ch}}, P_t^{\text{dis}} \leq P^{\text{max}} \\ E_T &= E_0 \end{aligned} \quad [7.9]$$

where:

- $\lambda_t^{\text{buy}}, \lambda_t^{\text{sell}}$ = MCP at hour (or 15-min interval) t in DAM/IDM;

- $P^{\text{ch}}_t, P^{\text{dis}}_t$ = charging and discharging power (MW);
- $n^{\text{ch}}, n^{\text{dis}}$ = charge/discharge efficiency;
- E_t = stored energy at time t ;
- $E^{\text{max}}, P^{\text{max}}$ = rated capacity and power limits;
- Δt = time resolution (0.25 h for 15-min MTU).

The BESS operator submits buy bids in low-price periods and sell offers during high MCP intervals, balancing market exposure between the DAM (predictable spreads) and IDM (short-term volatility).

Given the observed weak correlation between XBID volume and MCP ([Table 7.2](#)), BESS units should maintain partial SoC flexibility after DAM commitments to exploit unforeseen price spikes in IDM. The expected value of flexibility (V_f) can be formulated as:

$$V_f = \mathbb{E} \left[\max(\lambda_t^{\text{IDM}} - \lambda_t^{\text{DAM}}, 0) \cdot P_t^{\text{dis}} \right] \quad [7.10]$$

A higher V_f implies stronger incentives for deferred IDM participation, particularly during high volatility or imbalance hours. To ensure operational readiness, the BESS should maintain SoC within a responsive mid-range, from 0.4 to 0.6 of E_{max} .

7.3.2. Collocated RES with BESS

Hybrid configurations coupling RES with BESS introduce an additional decision layer: determining whether to export, store or curtail generation. These systems can optimize their participation by absorbing forecast errors, reducing curtailment and maximizing arbitrage value across markets.

For a hybrid RES–BESS asset, the bidding problem can be formulated as:

$$\max_{\{P_t^{\text{res}}, P_t^{\text{ch}}, P_t^{\text{dis}}\}} \Pi_t = \sum_{t=1}^T \lambda_t (P_t^{\text{res}} + P_t^{\text{dis}} - P_t^{\text{ch}}) \Delta t \quad [7.11]$$

subject to:

$$P_t^{\text{res}} + P_t^{\text{dis}} - P_t^{\text{ch}} \leq P^{\text{max}} \quad [7.12]$$

$$E_{t+1} = E_t + \eta^{\text{ch}} P_t^{\text{ch}} \Delta t - \frac{1}{\eta^{\text{dis}}} P_t^{\text{dis}} \Delta t$$

$$0 \leq P_t^{\text{res}} \leq P_t^{\text{res,avail}} = \hat{P}_t^{\text{res}} (1 - \delta_t^{\text{curt}})$$

where δ_t represents the curtailment ratio determined by system constraints.

The hybrid system strategically chooses whether to inject energy, store excess generation or trade intraday based on price signals and forecast accuracy.

The optimal strategy for RES–BESS hybrids in the Greek market should follow a sequential DAM–IDA–XBID bidding framework:

- DAM – submit a conservative generation forecast P_t^{Res} with flexible sell bids to mitigate imbalance exposure.
- IDAs – adjust positions as improved weather and system data become available.
- XBID – optimize real-time participation by discharging or purchasing based on MCP deviations, maximizing revenue and minimizing curtailment.

The incremental bidding adjustment across markets can be represented as:

$$\Delta P_t^{\text{XBID}} = P_t^{\text{opt}} - (P_t^{\text{DAM}} + \sum_{i=1}^3 \Delta P_t^{\text{IDA}_i}) \quad [7.13]$$

Given curtailment peaks exceeding 16 GWh, as shown in section [7.2](#), storage systems must be sized to absorb surplus generation during

midday RES surges. The curtailment reduction ratio (CRR) may be defined as:

$$\text{CRR} = 1 - \frac{\sum_t P_t^{\text{curt, with BESS}}}{\sum_t P_t^{\text{curt, without BESS}}} \quad [7.14]$$

Higher CRR values indicate greater effectiveness of storage in mitigating forced disconnections and stabilizing system operation.

In conclusion, the modeling results emphasize the value of multimarket, flexibility-oriented bidding strategies. The optimal design integrates:

- predictive DAM participation for baseline commitments;
- dynamic IDM responsiveness through XBID participation;
- real-time optimization using SoC-aware algorithms.

For the upcoming 15-min MTU implementation, algorithmic bidding frameworks capable of real-time market interaction will be essential. These models should integrate MCP forecasts, volatility indicators and congestion signals to dynamically optimize bids while ensuring SoC stability.

7.4. Conclusion

This chapter investigated the operational dynamics of the Greek electricity market under high RES penetration, emphasizing the role of flexibility mechanisms and optimal bidding strategies within the DAM and IDM. Using real datasets from the HEnEx and the Independent Power Transmission Operator (IPTO), the analysis identified a clear hierarchy between structured IDA1–3 and continuous XBID trading, with XBID demonstrating substantially higher trading volumes and faster responsiveness to market conditions. However, persistent RES curtailment, reaching daily peaks up to 16 GWh, exposed the system’s limited adaptability and underscored the need for enhanced flexibility solutions.

The research investigates bidding strategies for both stand-alone BESS and collocated RES with BESS hybrid configurations through mathematical formulations that maximize profit and mitigate curtailment while respecting operational constraints. Stand-alone BESS strategies centered on price arbitrage, maintaining a responsive state of charge to capitalize on MCP differentials between DAM and IDM sessions. Hybrid RES with BESS strategies introduced joint optimization of generation, storage and intraday bidding to absorb forecast deviations and reduce forced disconnections.

These findings confirm that the Greek electricity market operates under a layered flexibility structure, wherein structured (IDA-based) and continuous (XBID) mechanisms complement one another to manage real-time uncertainty. The upcoming transition to a 15-min MTU is expected to further enhance temporal granularity and system adaptability. Nevertheless, enduring challenges – including limited storage capacity, grid congestion and forecasting inaccuracies – continue to constrain full market efficiency.

Both stand-alone and hybrid BESS configurations are thus positioned as key enablers of flexibility in Greece's power market. Their optimal integration supports system reliability, reduces RES curtailment, and enhances alignment with the European Union's decarbonization and market integration objectives.

Future research should extend this analysis through higher resolution modeling under the 15-min MTU, along with expanding the analysis to include BESS participation across balancing (Karapidakis [2007](#)) and ancillary service (Bouzounierakis et al. [2019](#)) markets based on previous studies. Joint optimization models can quantify the combined technical and economic benefits of storage across energy and system services, improving overall grid resilience.

7.5. References

Acaroğlu, H. and García Márquez, F.P. (2021). Comprehensive review on electricity market price and load forecasting based on wind energy. *Energies*, 14, 7473.

- Alberizzi, A., Zani, A., Barba, P.D. (2023). Impact of renewables resources into the Italian intraday electricity market. *International Journal of Smart Grid and Clean Energy*, 12, 4.
- Andreotti, D., Spiller, M., Scrocca, A., Bovera, F., Rancilio, G. (2024). Modeling and analysis of BESS operations in electricity markets: Prediction and strategies for day-ahead and continuous intra-day markets. *Sustainability*, 16, 7940.
- Bouzounierakis, N., Katsigiannis, Y., Fiorentzis, K., Karapidakis, E. (2019). Effect of hybrid power station installation in the operation of insular power systems. *Inventions*, 4(3), 38.
- EnExGroup (n.d.a). DAM and IDM archive [Online]. Available at: <https://www.enexgroup.gr/web/guest/dam-idm-archive>.
- EnExGroup (n.d.b). Day ahead market publications [Online]. Available at: <https://www.enexgroup.gr/web/guest/markets-publications-el-day-ahead-market>.
- Karapidakis, E. (2007). Machine learning for frequency estimation of power systems. *Applied Soft Computing Journal*, 7(1), 105–114.
- Karapidakis, E., Katsigiannis, Y., Blazakis, K., Nikologiannis, M., Matalliotakis, G., Stavrakakis, G., Venianakis, N., Bonfini, P. (2024a). Assessing the impact of renewable energy sources on energy economics: A non-linear regression analysis of hellenic energy exchange market clearing prices. In *SMTDA Conference*, Chania, Crete, 4–7 June.
- Karapidakis, E., Katsigiannis, Y., Blazakis, K., Nikologiannis, M., Matalliotakis, G., Stavrakakis, G., Venianakis, N. (2024b). Evaluation of renewable energy sources on energy economics: A correlation analysis of clearing prices in the Hellenic energy exchange market. In *CHAOS Conference*, Chania, Crete, 11–14 June.
- Makrygiorgou, J.J., Karavas, C.-S., Dikaiakos, C., Moraitis, I.P. (2023). The electricity market in Greece: Current status, identified challenges, and arranged reforms. *Sustainability*, 15, 3767.

Nikologiannis, M., Karapidakis, E., Iliadis, I., Yfanti, S. (2025). Assessing flexibility in the greek day-ahead and intraday electricity markets. *E3S Web of Conf.*, 638, 02010.

Shinde, P., Boukas, I., Radu, D., Manuel De Villena, M., Amelin, M. (2021). Analyzing trade in continuous intra-day electricity market: An agent-based modeling approach. *Energies*, 14, 3860.

Note

Chapter written by Emmanuel KARAPIDAKIS, Marios NIKOLOGIANIS and George MATALLIOTAKIS.

8

Long-Term Memory, Event-Driven Dynamics and Predictive Power: An Analysis of Economic Policy Uncertainty and Market Volatility in the United States and UK

In this chapter, we investigate the persistence of economic policy uncertainty (EPU) indices in the United States and the United Kingdom for the period of 2001 to 2025. Furthermore, we examine the relationship between the US EPU and the volatility index (VIX). In our analysis, we use time series techniques such as the Hurst exponent, detrended fluctuation analysis, augmented Dickey–Fuller tests, correlation analysis and finally Granger causality tests. We apply these techniques in our analysis, as we obtain clear evidence of long memory behavior in both EPU time series. We find Hurst estimates consistently above 0.5, while further subperiod analysis clearly shows that the persistence varies across different economic regimes. While it is usual for financial crises to weaken persistence, we note that prolonged structural events such as Brexit and the Covid-19 pandemic generated stronger and more sustained uncertainty dynamics. We notice in our analysis that although the correlation between US EPU and the VIX is relatively modest, there is significant bidirectional Granger causality, with US EPU leading the VIX by approximately four days. Our findings suggest that policy-related uncertainty may contain useful information about potential changes in market volatility during periods of economic and market stress.

8.1. Introduction

EPU indices introduced by Baker et al. ([2016](#)) are now standard tools to quantify uncertainty around policy decisions and their likely implications. In empirical macroeconomic research, EPU indices are

frequently used as risk proxies to measure policy-related uncertainty that would otherwise be difficult to directly measure. The persistence properties of EPU indices, that is, shocks fading quickly or rather lingering over time, carry important implications. If uncertainty shocks exhibit long memory, this could potentially influence investment behavior, financial market volatility, investment decisions and the broader transmission of macroeconomic policy (Bloom [2009](#); Arouri et al. [2016](#)). This is an important distinction for modeling. In standard macroeconomic frameworks, transitory shocks can often be treated as short-lived fluctuations; however, persistent shocks are very different. It is well documented that persistent shocks can influence expectations, affect risk perceptions and spread through the economy for extended periods of time.

An extensive body of work looking at the persistence of EPU time series since 2010 is now available, in which researchers have explored various methods. There are studies using techniques usually associated with the field of econophysics, such as estimation of the Hurst exponent, detrended fluctuation analysis, and multifractal scaling approaches. Others rely on standard econometric frameworks, including fractional integration models, ARFIMA processes and unit root testing procedures. Despite differences in methodologies, the conclusion that EPU tends to be persistent is broadly reached by most studies in this literature.

However, a closer study reveals something slightly more complex and even slightly contradictory. A significant number of studies report Hurst exponents above 0.5 or fractional integration parameters greater than zero, and these results are typically interpreted as evidence of long-range dependence. At the same time, other studies report the presence of mean reversion, which is often reflected in fractional integration parameters below 1, or identify episodes in which the series appears closer to anti-persistent behavior (Gil-Alana and Payne [2016](#); Solarin and Gil-Alana [2021](#)). Thus, it appears that empirical research is not entirely uniform. Persistence is often present, but its strength and direction may vary depending on the sample period, the country examined, or the methodology applied.

Other researchers focus on the interaction between policy uncertainty and financial market volatility, but the causal links between the two are still not clearly established. There are difficulties in determining the direction of influence and potential lead–lag structures between them.

In this chapter, we address this issue directly by examining the relationship between US EPU and the VIX. The analysis proceeds with several complementary statistical approaches, including Pearson correlation, lagged correlation analysis and Granger causality testing. Our main goal in this research has less to do with the two variables moving together, but mainly we wish to identify if one variable may systematically precede the other in time. If this is true, this could contain predictive information regarding future market volatility.

There are three main objectives in our study:

- First, our goal is to evaluate whether the EPU time series for both the United States and the United Kingdom display evidence of long-term memory. It is important for researchers to establish if this property is valid since persistent uncertainty shocks could cause prolonged macroeconomic and financial effects in the market.
- The second goal would be to examine whether different major economic and geopolitical events could influence the persistence characteristics of EPU indices. There are events, such as financial crises, political realignments and other global disruptions, that could alter the way uncertainty behaves and evolves through time.
- The third research goal would be to investigate the empirical relationship between the US EPU and the VIX. In our study, we place particular emphasis on correlation patterns, potential lead–lag dynamics and causal interactions as captured by Granger causality tests.

The remainder of our research proceeds as follows. In section [8.2](#), we provide the relevant background and the theoretical foundation that we use in our analysis. In section [8.3](#), we describe the data as well as

the methodology employed in our empirical investigation. In section [8.4](#), we present and discuss the main empirical research results, while in section [8.5](#), we consider the broader implications of these findings. We are looking to understand the relationship between policy uncertainty and financial market dynamics. Finally, in section [8.6](#), we offer the final conclusions, along with a summary of the key insights, and we highlight possible directions for future research.

8.2. Background and theoretical foundations

8.2.1. Long memory and persistence

The study of uncertainty in financial markets and the economy is a primary research area that integrates macroeconomic and quantitative analysis. The empirical study as well as the analysis of business cycles and economic policy effectiveness requires certain EPU measures. When researchers study economic systems, they tend to analyze EPU patterns through the Hurst exponent, as this could help them better understand their behavior.

Time series data show long memory when their autocorrelation values decrease at a slow rate, which commonly occurs in economic systems. The Hurst exponent (H) serves as an effective method to determine the level of persistence that exists in time series data. Long memory stands when the autocorrelations of a time series slowly decay, typically at a hyperbolic rate (Beran [1994](#)).

Autocorrelation decay under long memory:

$$\rho(k) \sim Ck^{2H-2}, \text{ as } k \rightarrow \infty$$

where:

- $\rho(k)$ = autocorrelation at lag k ;
- H = Hurst exponent;
- C = constant.

This hyperbolic decay contrasts with the exponential decay observed in short-memory processes.

As a result, observations far apart in time remain statistically dependent. This contrasts with short-memory processes, where autocorrelations decline exponentially and the influence of shocks dissipates relatively quickly.

One formal approach to modeling long memory is fractional integration. In fractionally integrated processes, the integration parameter d lies between 0 and 0.5 for stationary long-memory behavior (Granger and Joyeux [1980](#); Hosking [1981](#)).

$$(1 - L)dX_t = \varepsilon_t$$

where:

- L = lag operator;
- d = fractional differencing parameter;
- $\varepsilon_t \sim iid(0, \sigma^2)$.

This framework, although powerful, still requires the correct specification of the underlying data-generating process, and for uncertainty indices, this is not always clear in practice.

There is a well-known and flexible approach that relies on the Hurst exponent, denoted by H . The Hurst exponent takes values between 0 and 1 and provides a direct measure of persistence (Hurst [1951](#); Zournatzidou and Floros [2023](#)).

Long memory is typically characterized by:

- Hurst exponent (H): for $H > 0.5$, the series is persistent (positive autocorrelations decay slowly); $H = 0.5$ indicates no long-range dependence (random walk); $H < 0.5$ suggests anti-persistence (mean reversion stronger than random walk) (Hurst [1951](#); Beran [1994](#)).
- Fractional integration parameter (d): a series is fractionally integrated of order d , denoted $I(d)$, if it becomes stationary after differencing d times. For $d \in (0, 0.5)$, the series is stationary with long memory and mean reversion; for $d \in (0.5, 1)$, it is non-

stationary but mean reverting; $d = 1$ corresponds to a unit root (Baillie [1996](#); Gil-Alana and Payne [2016](#)).

The relationship between H and d is $H = d + 0.5$ for fractionally integrated processes (Granger and Joyeux [1980](#)).

For uncertainty indices, high Hurst values suggest that uncertainty is clustered and that shocks fade slowly. The degree of persistence in EPU has direct economic consequences:

- Persistent EPU ($H > 0.5$, $d > 0$): uncertainty shocks have long-lasting effects on investment, employment and output. Policy interventions to reduce uncertainty may take extended periods to materialize (Bloom [2009](#); Solarin and Gil-Alana [2021](#)).
- Mean-reverting EPU ($d < 1$): shocks eventually dissipate, although the speed depends on the magnitude of d . This implies that extreme uncertainty episodes are temporary (Gil-Alana and Payne [2016](#); Solarin and Gil-Alana [2021](#)).
- Modeling implications: ignoring long memory in EPU can bias impulse response functions in VAR models, overstating the magnitude and duration of EPU effects on macroeconomic variables (Bloom [2009](#); Arouri et al. [2016](#)).

There is a question that naturally arises about the reason uncertainty exhibits long memory. There are several economic factors that could explain long memory in uncertainty indices. Many sources of uncertainty evolve slowly over time. There are reasons ranging from political and regulatory frameworks to policy regimes, as they often change gradually, which implies that uncertainty could persist for longer periods. Second, information frictions and learning dynamics may contribute to persistence. Economic agents update their beliefs gradually, as a new flow of information is distributed in the markets. Therefore, news about uncertain events often accumulates rather than being resolved instantaneously. This can create a prolonged adjustment process in measured uncertainty. Third, uncertainty is closely linked to the business cycle. There are periods of high uncertainty that tend to coincide with recessions or financial stress, during which feedback mechanisms can reinforce persistence. When uncertainty is high, firms pause hiring and investments, which

causes a weak economy, leading to even more uncertainty. Together, these mechanisms suggest that long-range dependence on uncertainty indices implies an underlying, substantive economic cause.

8.2.2. Academic literature

Uncertainty indices play a major role in empirical macroeconomics and finance (Bloom [2009](#); Baker et al. [2016](#)). Arouri et al. ([2016](#)) found that an increase in policy uncertainty significantly reduces stock returns; this effect is stronger and persistent during extreme volatility periods. Furthermore, a common empirical finding is that uncertainty indices display a high degree of persistence, often exceeding what can be adequately captured by standard short-memory time-series models; when uncertainty rises, it persistently exhibits high levels.

The analysis of macroeconomic and financial systems depends on uncertainty indices that include EPU measures. The indices show persistent trends that standard short-memory time series models cannot forecast. Hence, it is important to examine the long memory phenomenon within uncertainty indices and test whether economic shocks to uncertainty will fade out at a slow rate and need to be considered by both modelers and policymakers.

Previous research has extensively studied the impact of EPU on various macroeconomic indicators and financial markets. However, a deeper understanding of the long-term memory properties of EPU series, particularly how these properties evolve during significant economic and geopolitical events, remains. We aim to fill this gap in the study by conducting a comprehensive analysis of EPU persistence using advanced time series techniques, including the Hurst exponent and detrended fluctuation analysis.

Gil-Alana and Payne ([2016](#)) showed that the EPU index is persistent but mean reverting in the United States. Caporale et al. ([2018](#)) studied the persistence in uncertainty series and found time-varying persistence, indicating inefficiency during crises. In addition, Plakandaras et al. ([2019](#)) examined the persistence and chaotic behavior of 72 uncertainty indices over the period of the mid-1980s to 2017. They reported that almost all uncertainty indices are

persistent, while chaotic dynamics are detected only for certain indices during recessions of economic turbulence. Solarin and Gil-Alana (2021) explained that the examination of the persistence of EPU is important. They reported that the size of the uncertainty persistence will determine how large the negative effect of an uncertainty shock is on the financial market as well as on the economy. Their results show a long memory and mean reverting behavior. Furthermore, Sheng et al. (2022) analyzed the persistence of economic and policy-related uncertainty of the 50 US states in a panel data setup over the monthly period of 1984:03 to 2019:12. They reported that the degree of persistence of uncertainty is higher under the upper regime of climate risks. Recently, Marques and Pinto (2024) examined the persistence dynamics in Brazilian economic uncertainty; they reported that economic uncertainty in Brazil is non-stationary but mean-reverting, with its persistence only transitorily affected by economic and political crises.

8.3. Methodology

To achieve the research objectives, a comprehensive suite of time series analytical techniques was employed.

8.3.1. Data

To study long memory with the Hurst exponent, we consider daily EPU measures for both the United States and United Kingdom. The source is www.policyuncertainty.com/, which presents a description of both indices as follows: the daily news-based EPU Index for the United States is based on a broader set of newspapers derived from the Access World News NewsBank service. The number of articles contains at least one term from each of three sets of terms. The first set is economic or economy. The second is uncertain or uncertainty. The third set is legislation, deficit, regulation, congress, federal reserve or the White House.

Similarly, the daily newspaper-based EPU Index for the United Kingdom draws on the digital archives of the Access World News NewsBank service. This service includes coverage of approximately 650 UK newspapers, ranging from large national papers such as the

Guardian to small local newspapers across the United Kingdom. According to www.policyuncertainty.com/, to construct the index, the daily count of articles that contain at least one term from each of three term sets (economic or 1) economy, 2) uncertainty, 3) spending, deficit, regulation, budget, tax, policy, or Bank of England) is considered.

The sample considers a long period of uncertainty, that is, from 2001 to 2025. More specifically, the period 2001–2002 includes the so-called Dot-com Bubble period, coupled with the September 11 attacks, which caused a significant nine-month recession. The period of 2007–2009 is called The Great Recession triggered by the subprime mortgage crisis and the collapse of the US housing bubble; this led to a global financial crisis (GFC) and included bank failures and economic downturns. The period of 2010–2019 is about the Eurozone Debt Crisis and includes a sovereign debt crisis of PIIGS, resulting in severe austerity measures. The period of 2020–2022 includes the Covid-19 pandemic and inflation shock, while the recent period of 2023–2025 includes the energy crisis following the Russian invasion of Ukraine and high energy costs. Finally, 2025 faces a complex mix of risks, such as economic, geopolitical and environmental risks due to trade wars, high levels of debt, high inflation and slow growth as well as climate-related financial risks.

8.3.2. *Hurst exponent and detrended fluctuation analysis*

The Hurst exponent (H) serves as a crucial measure of long-term memory or persistence in a time series. A value of $H < 0.5$ suggests anti-persistence, implying that increases tend to be followed by decreases and vice versa. A value of $H = 0.5$ indicates a random walk, where past movements have no predictive power for future directions. Conversely, $H > 0.5$ signifies persistence or long-term memory, meaning that past trends are likely to continue.

Rescaled Range statistic:

$$\frac{R(n)}{S(n)} \propto n^H$$

where:

- $R(n)$ = range of cumulative deviations;
- $S(n)$ = standard deviation;
- c = constant.

Detrended fluctuation analysis was used to estimate the Hurst exponent. We find that DFA is particularly robust for analyzing non-stationary time series since it can effectively quantify long-range correlations. This is done by detrending the series and observing fluctuation behavior across different time scales. DFA removes polynomial trends before computing fluctuation scaling, making it robust to non-stationarities (Peng et al. [1994](#)). Multifractal DFA (MF-DFA) extends this to capture scale-dependent persistence.

Yao et al. ([2020](#)) used MF-DFA to analyze US EPU, finding generalized Hurst exponents significantly greater than 0.5, indicating persistence. Zhao and Cui ([2019](#)) employed multifractal detrended cross-correlation analysis (MF-DCCA) to study US EPU and exchange rates, reporting Hurst exponents ranging from 0.64 to 0.89 across different scales, with stronger persistence in long-term correlations.

Step 1: profile construction:

$$Y(i) = t = 1 \sum i (X_t - \bar{X})$$

Step 2: fluctuation function:

$$F(n) = \sqrt{\frac{1}{N} \sum [Y(i) - Y_n(i)]^2}$$

where:

- $Y_n(i)$ = local polynomial trend in window n .

8.3.3. Augmented Dickey–Fuller test

The augmented Dickey–Fuller test is a statistical test used to determine the presence of a unit root in a time series, which would be an indication of non-stationarity. The null hypothesis of the ADF test is that a unit root exists (i.e. the series is non-stationary). A p value less than 0.05 typically leads to the rejection of the null hypothesis, indicating that the series is stationary. Stationarity is a desirable property for many econometric models and forecasting techniques.

$$\Delta X_t = \alpha + \beta t + \gamma X_{t-1} + \sum_{i=1}^p \phi_i \Delta X_{t-i} + \varepsilon_t$$

Null hypothesis:

$$H_0: \gamma = 0$$

8.3.4. Subperiod analysis

To investigate the impact of significant economic and geopolitical events on the persistence characteristics of EPU, both US and UK EPU data were segmented into several distinct subperiods. These periods were selected to capture major events that could plausibly alter the dynamics of policy uncertainty. The subperiods analyzed included the following:

- Pre-2008 crisis.
- 2008 financial crisis.
- Eurozone crisis.
- Brexit period.
- Post-Brexit/pre-COVID.
- Covid-19 pandemic.
- War shocks/recent events.

For each subperiod, the Hurst and DFA Hurst exponents were calculated to observe changes in persistence.

8.3.5. Pearson correlation

The Pearson correlation coefficient was used to quantify the strength and direction of the linear relationship between US EPU and VIX. The coefficient ranges from +1 (perfect positive correlation) to -1 (perfect negative correlation), with 0 indicating no linear correlation.

8.3.6. Lagged correlation analysis

Lagged correlation analysis was applied to explore potential lead-lag relationships between US EPU and VIX. This technique involves calculating the correlation between the two time series at various time shifts (lags) to identify whether one series consistently precedes movements in the other.

8.3.7. Granger causality test

The Granger causality test is a statistical hypothesis test used to determine if one time series is useful in forecasting another. The null hypothesis is that the past values of one series do not help predict the future values of another. Rejection of the null hypothesis suggests a causal relationship in the Granger sense. This test was performed to understand the directional predictive power between US EPU and VIX.

Testing if X Granger-causes Y:

$$Y_t = \alpha_0 + \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{i=1}^p \beta_i X_{t-i} + \varepsilon_t$$

Null hypothesis:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

8.4. Empirical results

8.4.1. Overall EPU persistence and stationarity

There are studies published after 2010 that indicate that both the US and UK EPU indices exhibit persistence. This is usually shown in Hurst exponents being above 0.5 or in fractional integration parameters between 0 and 1. In our data presented here, there is

strong evidence of this general pattern, which indicates the long memory properties of the EPU time series.

We performed a full-sample analysis that indicates pronounced persistence in both countries. In the United States, the EPU has a DFA-based Hurst exponent of the row series of 1.0181, while the standard Hurst estimate is 0.853. We have similar findings for the UK index, where the DFA-Hurst exponent is approximately 0.9831. These values exceed the 0.5 level, pointing to trend-reinforcing behavior. We notice there are periods of increased uncertainty that are statistically more likely to continue and declining phases that are also likely to persist when they happen. On the other hand, in the United States, the time series exhibits slightly stronger persistence. This is an indication that past trends are somewhat more influential in shaping movements in the future.

We assessed stationarity using the augmented Dickey–Fuller test. Our results confirmed that both series are stationary. The US EPU yields a p value of 7.301×10^{-5} with an ADF statistic of -4.733 , while the UK EPU has a p value of 4.037×10^{-9} and an ADF statistic of -6.694 . The p values are well below the 0.05 threshold, providing a rejection of the null hypothesis of non-stationarity. In practice, this means that although the series do exhibit strong persistence, they follow a stationary random walk, while their statistical properties remain stable over time.

8.4.2. Presentation of data

The persistence we discussed in the previous sections is clearly shown in the visualization of the US EPU Index over time.

We can observe by looking at the plot:

- *Long-lasting trends*: there are periods where the EPU index stays relatively high or low for extended durations, rather than fluctuating randomly around a mean. A persistent series does have this characteristic, a “memory” of past states.
- *Clustering of values*: when the index goes up, it tends to stay up for a while. When it goes down, it tends to stay down. This

clustering of similar values reinforces the idea that the series is not undergoing frequent, sudden reversals.

This visual evidence is in line with the calculated Hurst exponents (DFA Hurst: ~ 1.018 , Hurst Exp: ~ 0.867), both of which indicated strong persistence. A higher Hurst exponent suggests a stronger tendency for the series to maintain its direction, which is what we can see in these sustained upward or downward movements.

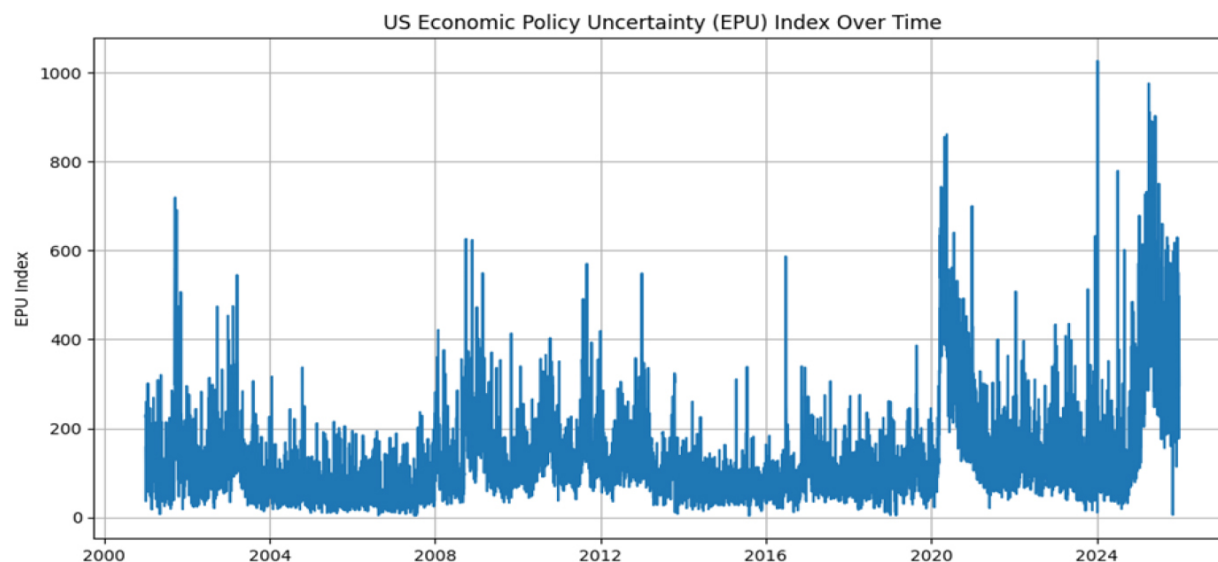


Figure 8.1. *US Economic Policy Uncertainty (EPU) Index Over Time*

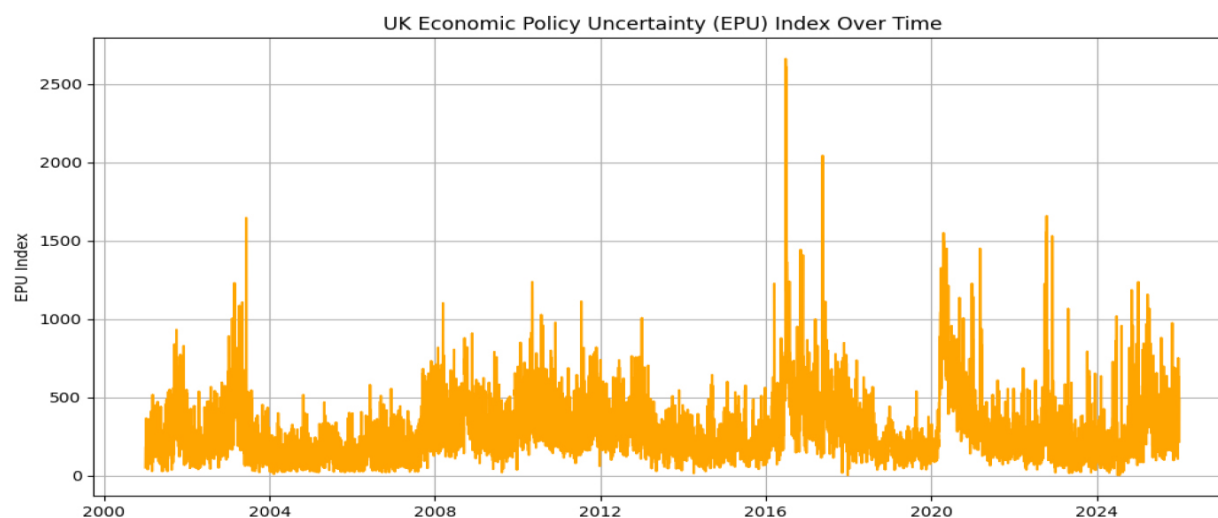


Figure 8.2. *UK Economic Policy Uncertainty (EPU) Index Over Time*

The UK EPU series (shown in orange) exhibits clear visual persistence in a similar manner. Once trends emerge, they tend to continue for extended periods, which agrees with the calculated Hurst exponents (DFA ≈ 0.983 , standard Hurst ≈ 0.853).

- *Overall pattern*: the UK EPU plot mirrors the behavior observed in the US series. Both display periods of relatively stable uncertainty, interrupted by sharp spikes, followed by gradual declines or sustained phases of elevated uncertainty.
- *Magnitude and timing*: although the general shapes look similar, the exact magnitude of peaks and troughs differs, reflecting country-specific economic and political developments. However, the underlying tendency for trends to persist appears very similar across both series.

Together, the visual evidence reinforces our statistical findings. It is clear for both EPU indices in the United States and the United Kingdom that they do have a long-term memory, an indication that past levels of uncertainty significantly influence future values, instead of following a random walk. This visual inspection adds further strength to the conclusions from our Hurst exponent analysis in this study.

Here is a summary of what we have found:

1. *DFA Hurst exponent*:
 - US EPU: ~ 1.018 .
 - UK EPU: ~ 0.983 .
 - Interpretation: both series show strong persistence (long-term memory), meaning past trends are likely to continue. The values being close to 1 suggest a robust trend-reinforcing behavior.
2. Hurst exponent (R/S method):
 - US EPU: ~ 0.867 .
 - UK EPU: ~ 0.853 .

- Interpretation: these values also confirm strong persistence for both series, consistent with the DFA results.

3. ADF test for stationarity:

- US EPU: stationary (p value = $7.301e-05$, ADF Stat = -4.733).
- UK EPU: stationary (p value = $4.037e-09$, ADF Stat = -6.694).
- Interpretation: both series are statistically stationary, meaning their statistical properties (such as the mean and variance) are constant over time. This is a desirable property for forecasting and statistical modeling.

4. Hurst exponent on log-transformed data:

- UK EPU (log-transformed): ~ 0.856 .
- Interpretation: log transformation did not significantly alter the estimated Hurst exponent, reinforcing the robustness of the persistence finding.

5. Visualizations:

- Line plots of both US and UK EPU data visually confirmed the persistence, showing clear periods of sustained high or low uncertainty, aligning with the statistical Hurst exponent findings.

Both the US and UK EPU time series are stationary and exhibit strong persistence. This suggests that while the series fluctuates around a stable mean, its movements have a long memory, making past trends a good indicator of future direction. This makes them interesting candidates for further time series modeling and forecasting.

8.4.3. Subperiod analysis of EPU persistence

The subperiod analysis revealed how the persistence characteristics of EPU shifted in response to major economic and geopolitical events.

8.4.3.1. US EPU subperiod analysis

The DFA and Hurst exponents for the US EPU across different subperiods are presented in [Table 8.1](#).

Table 8.1. *US EPU Hurst exponents by subperiod*

Subperiod	DFA Hurst	Hurst Exp
Pre-2008 crisis	0.8804	0.6985
2008 Financial crisis	0.7800	0.6394
Eurozone crisis	0.8744	0.7138
Brexit period	0.8406	0.5410
Post-Brexit/pre-Covid	0.8011	0.6961
Covid-19 pandemic	0.9445	0.6853
War shocks/recent events	0.7897	0.6703

Observations from the US EPU subperiod analysis indicate that persistence generally decreased during the 2008 Financial crisis and the War shocks/recent events periods. Conversely, persistence remained notably high, particularly the DFA Hurst exponent, during the Covid-19 pandemic, reaching 0.9445.

8.4.3.2. UK EPU subperiod analysis

The DFA and Hurst exponents for the UK EPU across different subperiods are presented in [Table 8.2](#).

Table 8.2. *UK EPU Hurst exponents by subperiod*

Subperiod	DFA Hurst	Hurst Exp
Pre-2008 crisis	0.8565	0.7286
2008 Financial crisis	0.7209	0.5212
Eurozone crisis	0.7281	0.5372
Brexit period	1.0130	0.7502
Post-Brexit/pre-Covid	0.7875	0.7374
Covid-19 pandemic	0.9151	0.7051
War shocks/recent events	0.8181	0.6371

For the UK EPU, similar to the US EPU, there was a decrease in persistence in the 2008 Financial crisis and the Eurozone crisis. However, the Brexit period showed exceptionally high persistence, with a DFA Hurst exponent of 1.0130, indicating a very strong trending behavior. The Covid-19 pandemic also showed high persistence for UK EPU, with a DFA Hurst of 0.9151.

8.4.4. Relationship between US EPU and VIX

8.4.4.1. Pearson correlation

We calculated the Pearson correlation coefficient between the US EPU and the VIX, which was found to be 0.2552. This is an indication of a weak positive linear relationship. While both indices reflect uncertainty, they do not move linearly together in a strong relationship.

8.4.4.2. Lagged correlation

The lagged correlation analysis we performed revealed a lead–lag relationship in which the US EPU tends to lead the VIX. The maximum correlation value of 0.3230 occurred when the US EPU led the VIX by 4 days. This could be an indication that changes in EPU might precede changes in market volatility.

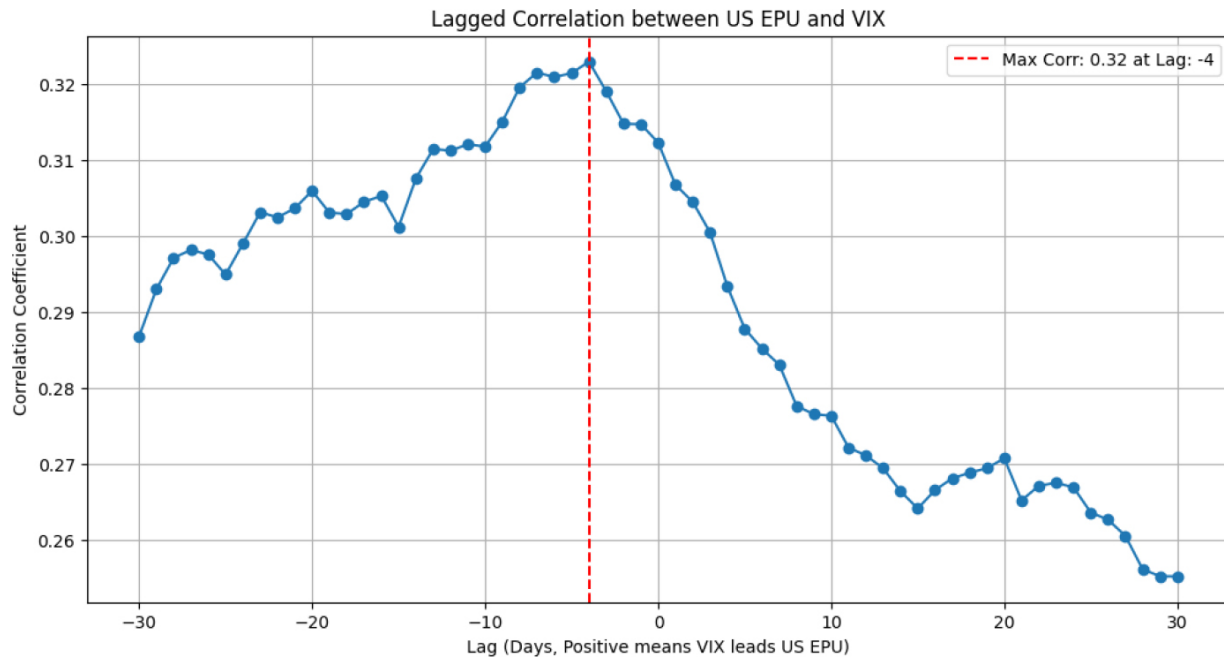


Figure 8.3. *Lagged correlation between US EPU and VIX.*

- Pearson correlation coefficient between US EPU and VIX: 0.3123.
- Maximum correlation: the plot highlights the maximum correlation of 0.32 occurring at a lag of -4 days.

The correlation remains positive across most lags, although it is generally weak to moderate. This indicates that the US EPU and VIX have the tendency to move in the same direction. This practically means that higher EPU is generally associated with higher market volatility, and vice versa.

However, the most significant finding from this plot is the peak correlation at -4 days. In this context:

- A negative lag means that the *first series (US EPU)* leads the *second series (VIX)*.
- A maximum correlation at -4 days suggests that changes in US EPU tend to precede changes in VIX by approximately 4 days.
- This is a suggestion that an increase in US EPU might be followed by an increase in market volatility (VIX) approximately

4 days later. Conversely, a decrease in US EPU might precede a decrease in VIX.

While the maximum correlation of 0.32 is not quite strong, it is an indication of a statistically significant relationship. The results of our analysis suggest that EPU is one of the factors influencing movements in market volatility. Further evidence points to a modest leading effect. In other words, changes in policy-related uncertainty appear to precede shifts in market volatility by a short interval.

Practically, this pattern may offer us a rather small but potentially useful predictive signal. If we monitor developments in the US EPU index, this provides us with a brief lead, roughly 4 days, before similar movements become visible in the VIX. For market participants, policymakers, or risk managers, even a short period of time, such as this one, may prove valuable when attempting to anticipate changes in market sentiment linked to policy developments.

At the same time, this kind of relationship should be approached with some caution. A correlation, even with a clear lead–lag structure, does not necessarily indicate a direct causal mechanism. We should examine this kind of relationship more carefully since the results point to a sequential dynamic between the two series. This could be done through formal causality testing, such as the Granger causality framework, to assess whether past values of policy uncertainty contain statistically meaningful information about future movements in market volatility.

8.4.4.3. Granger causality test results

The Granger causality tests indicated a significant bidirectional relationship between US EPU and VIX, as shown by the highly significant p values across various lags. This will help us determine if past values of one series are statistically significant in predicting future values of the other.

Summary of Granger causality test results (p value < 0.05 indicates Granger causality):

1. Testing whether US EPU Granger-causes VIX:

- Lag 1: p value = 0.0000 (US EPU Granger-causes VIX? Yes).
- Lag 2: p value = 0.0000 (US EPU Granger-causes VIX? Yes).
- Lag 3: p value = 0.0000 (US EPU Granger-causes VIX? Yes).
- Lag 4: p value = 0.0001 (US EPU Granger-causes VIX? Yes)
- Lag 5: p value = 0.0007 (US EPU Granger-causes VIX? Yes)

We reject the null hypothesis that US EPU does not Granger-cause VIX at all tested lags. This means that past values of US EPU are statistically significant in predicting future values of VIX.

2. *Testing whether VIX Granger-causes US EPU:*

- Lag 1: p value = 0.0000 (VIX Granger-causes US EPU? Yes).
- Lag 2: p value = 0.0000 (VIX Granger-causes US EPU? Yes).
- Lag 3: p value = 0.0000 (VIX Granger-causes US EPU? Yes).
- Lag 4: p value = 0.0000 (VIX Granger-causes US EPU? Yes).
- Lag 5: p value = 0.0000 (VIX Granger-causes US EPU? Yes).

We reject the null hypothesis that VIX does not Granger-cause US EPU at all tested lags. This means that past values of VIX are statistically significant in predicting future values of US EPU.

The results suggest a bidirectional Granger causality between US EPU and VIX. This means that:

- EPU has predictive power over market volatility (VIX): this finding agrees with the idea that uncertainty about government policies can lead to increased fear and risk aversion in financial markets.

- Market volatility (VIX) also has predictive power over EPU: this suggests that periods of high market fear and volatility might influence policy discussions, decisions and the overall perception of policy stability.

This finding builds upon our lagged correlation analysis. While lagged correlation showed a stronger correlation when EPU led VIX by 4 days, the Granger test indicates that both series have predictive information for each other. This suggests a complex interplay rather than a simple one-way causal relationship. It is an interesting dynamic where each can signal future movements in the other.

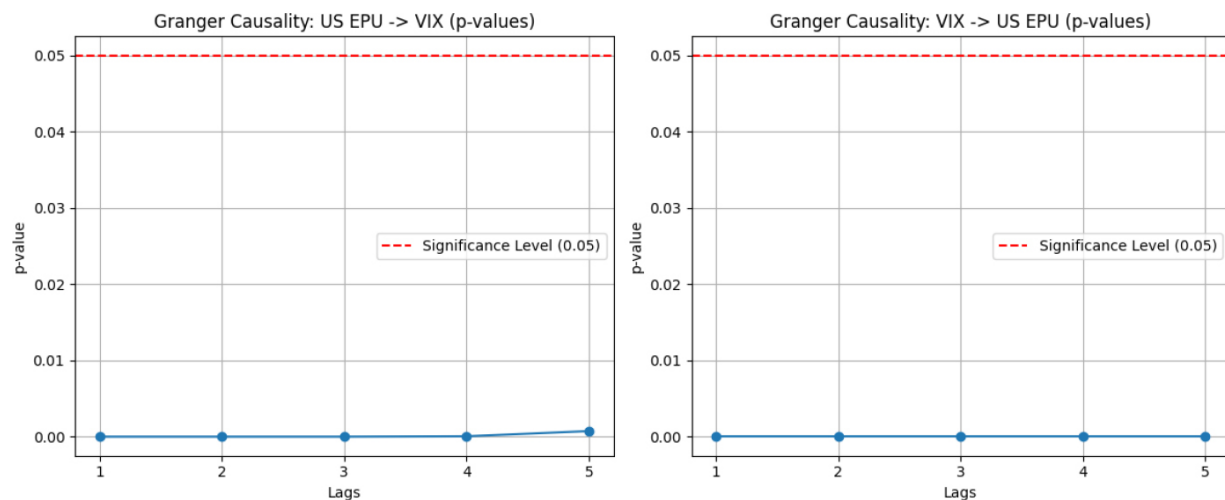


Figure 8.4. Granger Causality Tests

We create two plots to visualize the Granger causality test results. Each plot shows the p values for different lags, with a horizontal line indicating the 0.05 significance level. This will allow us to identify the lags for which the null hypothesis of no Granger causality is rejected.

The plots visually confirm the results of our Granger causality tests.

On the left, “Granger causality: US EPU -> VIX (p values)”:

- You can see that the p values for all lags (1 through 5) are well below the 0.05 significance level (the red dashed line).
- This visually reinforces our conclusion that past values of US EPU significantly predict future VIX values.

On the right, “Granger causality: VIX -> US EPU (p values)”:

- Similarly, the p values for all lags (1 through 5) are also well below the 0.05 significance level.
- This visually confirms that past VIX values significantly predict future US EPU values.

8.4.4.4. Overall interpretation of the visualization

Our findings reported from the hypothesis tests earlier are clearly shown in the visualizations of the Granger causality tests. For example, when the p values across the tested lags are plotted, we notice a consistent pattern. We find that for both directions, US EPU predicting VIX and VIX predicting US EPU, the p values remain extremely small and stay below conventional significance thresholds across all tested lags.

In simple terms, the graphical evidence of our analysis offers further support to the formal test results, since there appears to be strong statistical support for the presence of predictive relationships in both directions. For example, there is information within changes of the US EPU Index that impact subsequent movements in the VIX. At the same time, when there are movements in the VIX, they appear to have information about future movements in the US EPU.

This kind of pattern suggests that the relationship between policy uncertainty and market volatility is not straightforward or one-dimensional. More accurately, the two variables seem to interact in a dynamic and mutually reinforcing way. Practically, this is an indication of a more complex structure of uncertainty and the way it transmits within financial markets. It is fair to say that policy developments and market sentiment evolve together rather than independently.

8.5. Discussion

We gain several insights from the empirical findings of this study about how EPU behaves over time. We also learn how policy uncertainty interacts with financial market volatility. There are some results that are broadly consistent with the literature. However, there

are others indicating characteristics that require a much closer analysis.

8.5.1. *Persistent nature of EPU*

In our analysis, one of the most consistent outcomes is the strong persistence that we observe in both the US and the UK EPU time series. Across the full sample, both the estimated Hurst exponents and the DFA Hurst exponents remain well above the benchmark value of 0.5. This is typically an indication that the series exhibits persistent dynamics in the context of long memory processes, and it is not moving randomly.

Practically, our findings suggest that EPU tends to follow its own historical trajectory. There are periods of elevated uncertainty, which are more likely to be followed by further uncertainty, and there are calm periods that tend to persist as well when they do happen.

This kind of behavior is not a simple statistical phenomenon that matters only for academic researchers. It carries important implications for decision-making. There are firms, investors and policymakers that operate in regimes where expectations matter. In these kinds of environments, uncertainty could shape expectations for extended periods. For example, investment decisions may be postponed, risk premiums may be adjusted and policy responses may become more cautious when uncertainty appears likely to persist rather than disappear quickly.

There is also another persistent issue from a modeling standpoint. These results also reinforce the importance of taking into consideration long-term memory behavior when constructing forecasting frameworks for uncertainty-related variables.

8.5.2. *Impact of major events on EPU persistence*

A full sample analysis reveals clear evidence of persistence, but the subperiod results suggest that this type of behavior is not entirely stable in time. Instead, the persistence properties of EPU appear to change depending on the broader economic and geopolitical narratives.

8.5.2.1. *Financial crises (2008, Eurozone)*

There were major financial crises during our sample period, which included both the GFC and the Eurozone crisis. During those times, the estimated persistence of the EPU index declined slightly. Our studies revealed that during those episodes, the Hurst exponent moved closer to the value of 0.5, which indicates behavior that is closer to a random process.

Sudden systemic shocks could provide one possible interpretation. As they disrupt the typical dynamics of uncertainty, as events are unfolding and markets, as well as policymakers, are reacting to rapidly changing events, the usual path of uncertainty becomes more erratic and less trend-like. In those situations, uncertainty does not move forward gradually. It usually spikes suddenly and may fluctuate sharply as new information arrives in the market.

8.5.2.2. *Brexit period*

During the Brexit period, we noticed a very different pattern that emerged in the data. In the United Kingdom, the DFA Hurst index rose above one, indicating an extremely strong persistence.

This outcome is probably not surprising since the Brexit crisis represented a prolonged process that was characterized by repeated negotiations, political debate and uncertainty regarding the future. Under these conditions, policy uncertainty did not appear as a short-lived shock, but it was rather a sustained state of uncertainty. This is visible in the data, as strong persistence is clearly observed.

8.5.2.3. *Covid-19 pandemic*

The Covid-19 pandemic offers another case in which persistence increased in the US as well as the UK EPU time series. There was a difference from the Financial Crisis of 2008, which unfolded mostly within the financial system. The 2019 pandemic, on the contrary, generated a broad and prolonged global disruption that impacted public health, economic policy and international trade at the same time. During this period, uncertainty evolved gradually as governments introduced different policy responses, lockdown measures and fiscal interventions. Therefore, the resulting uncertainty appears to have remained at high levels for an extended

duration, and this is consistent with the high persistence estimates observed in our analysis.

When we consider all of these inputs together, the results highlight an important point. The persistence of EPU is not fixed. It appears instead to vary depending on the nature of the underlying shock, as well as the time horizon over which that uncertainty unfolds.

8.5.3. Relationship between US EPU and VIX

The analysis of the interaction between US EPU and the VIX reveals a relationship that is more complex than a simple correlation might suggest.

8.5.3.1. Weak positive correlation

The Pearson correlation coefficient between the two series is approximately 0.2552. This indicates only a modest positive relationship. However, this relatively weak correlation suggests that the two indices do not move together in a strict linear fashion.

There is one reasonable explanation for this, as each index captures different kinds of dimensions of uncertainty. For example, the IPU index reflects uncertainty, which is related specifically to government policy and institutional development.

On the other hand, this is calculated from stock prices in equity markets, and it is often seen as a forward-looking measure of market volatility and investor sentiment. Obviously, the two concepts are clearly related, but they are not identical.

8.5.3.2. Lagged relationship

The lagged correlation analysis we performed also offers additional insights. Our findings suggest that the US EPU tends to lead movements in the VIX by approximately 4 days, and the maximum correlation reaches approximately 0.3230 at that lag.

The magnitude of this correlation remains moderate, but the timing is noteworthy. This suggests that shifts in policy-related uncertainty may precede changes in market volatility. This could potentially provide an early signal of rising risk conditions in financial markets.

Especially for investors and policymakers, such a lead–lag structure could be very useful and important. Keeping track of developments in policy uncertainty could provide useful information about emerging volatility pressures.

8.5.3.3. *Bidirectional Granger causality*

The Granger causality tests we performed reveal an even more interesting pattern. These results indicate statistically significant bidirectional causality between the two series.

The US EPU appears to Granger cause VIX, which suggests that past movements in policy uncertainty hold information that helps predict future market volatility. On the other hand, the VIX also Granger causes US EPU. This is another indication that periods of increased market volatility may, in turn, feedback into perceptions of policy uncertainty.

When we take these two together, these findings indicate a feedback mechanism rather than a one-way causal relationship. Policy uncertainty can influence financial markets. Equally, financial market turbulence could also shape expectations regarding future policy action.

It is important to understand this kind of interaction because it suggests that uncertainty in modern financial markets may transmit through multiple channels simultaneously. This links economic policy developments with market dynamics in a continuous cycle.

8.6. Conclusion and future research

In this chapter, we examined the persistence properties of EPU indices in the United States and the United Kingdom. We also explored the interaction between the US EPU index and the market VIX. The empirical results, when taken together, point to a clear and consistent pattern. We note that both indices exhibit an important degree of persistence.

Our analysis indicates that EPU indices for both the US and the UK are stationary. This is a suggestion that uncertainty tends to evolve gradually rather than disappearing quickly after a shock. There are

periods characterized by elevated policy uncertainty that tend to persist. This is shaping expectations and potentially influences economic behavior for extended periods of time.

There is one very informative aspect of this study that has to do with the subperiod analysis. This highlights how the persistence of uncertainty can vary depending on the nature of the underlying shock. There are episodes of significant financial stress, such as the GFC or the Eurozone crisis, that appear to reduce the predictability of EPU dynamics.

During these times, persistence measures tend to move closer to values associated with random behavior. In contrast, prolonged and structurally focused sources of uncertainty produce a very different pattern. For example, during Brexit or the Covid-19 pandemic, uncertainty evolved more slowly and remained elevated over longer periods of time, resulting in a particular strong persistence.

These different outcomes make an important point that EPU is not governed by a single stable dynamic process. Its behavior instead appears to depend heavily on the context in which uncertainty happens. Sudden systemic shocks and prolonged structural disruptions may influence uncertainty in fundamentally different ways.

Our study analysis also provides new insights into the relationship between policy uncertainty and the financial market VIX. While the correlation between US EPU and the VIX is relatively modest, the dynamic relationship between the two variables is considerably more informative for us. The lag analysis suggests that movements in US policy uncertainty tend to occur before changes in the VIX by roughly 4 days. This may act as an early signal of rising volatility in financial markets.

It is even more important that Granger causality results reveal that bidirectional predictive relationships exist. US EPU appears to contain information that could help us forecast future movements in the VIX. However, the opposite is also true. Past volatility in financial markets could provide predictive information for future policy uncertainty. Our analysis findings point toward a feedback mechanism rather than a simple one-directional causality. Policy developments influence financial markets and in turn may shape

perceptions, and in turn, market conditions may shape perceptions of policy uncertainty.

The implications of our findings could practically impact several different market participants. Policymakers would benefit from recognizing that policy-deviated uncertainty can persist for significant periods of time. Economists and researchers may wish to analyze long memory dynamics more carefully when modeling uncertainty or forecasting macroeconomic values. Furthermore, financial participants may find it useful to monitor EPU changes as a possible leading indicator of future financial market volatility.

There are, of course, several different opportunities for further research. Future work could extend the analysis by examining a broader set of countries. For example, we could compare the persistence properties of different national EPU indices. There is also another direction that could involve exploring sector-specific or industry-specific reactions to policy uncertainty.

Furthermore, the relationship between EPU indices and other financial indicators, such as credit spreads, currency volatility, or sovereign risk measures, could provide further information about how uncertainty is transmitted in financial markets.

It is an important challenge for both researchers and policymakers to understand the dynamics of policy uncertainty. As the global economic and geopolitical environment is constantly evolving, it would become extremely important in the future to better understand the role of uncertainty in shaping financial markets.

8.7. References

- Arouri, M., Estay, C., Rault, C., Roubaud, D. (2016). Economic policy uncertainty and stock markets: Long-run evidence from the US. *Finance Research Letters*, 18, 136–141.
- Baillie, R.T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73(1), 5–59.

- Baker, S.R., Bloom, N., Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, 131(4), 1593–1636.
- Beran, J. (1994). *Statistics for Long-Memory Processes*. Chapman & Hall, London.
- Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77(3), 623–685.
- Caporale, G.M., Gil-Alana, L.A., Trani, T. (2018). Persistence in US economic policy uncertainty. *Applied Economics*, 50(51), 5484–5497.
- Dickey, D.A. and Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366), 427–431.
- Gil-Alana, L.A. and Payne, J.E. (2016). Measuring the persistence of economic policy uncertainty. *Applied Economics Letters*, 23(15), 1107–1112.
- Granger, C.W.J. and Joyeux, R. (1980). An introduction to long memory time series models. *Journal of Time Series Analysis*, 1(1), 15–29.
- Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, 68(1), 165–176.
- Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116, 770–808.
- Marques, A. and Pinto, P. (2024). Persistence dynamics in Brazilian economic uncertainty. *Economic Modelling*, 128, 106456.
- Peng, C.K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E., Goldberger, A.L. (1994). Mosaic organization of DNA nucleotides. *Physical Review E*, 49(2), 1685–1689.
- Plakandaras, V., Gupta, R., Wohar, M.E. (2019). The long memory property of economic uncertainty indices. *Physica A*, 526,

120900.

Sheng, X., Gupta, R., Cepni, O. (2022). Climate risks and persistence of policy uncertainty across US states. *Journal of Financial Stability*, 60, 101009.

Solarin, S.A. and Gil-Alana, L.A. (2021). Persistence of economic policy uncertainty. *Journal of Policy Modeling*, 43(2), 356–371.

Yao, C.-Z., Liu, C., Ju, W.-J. (2020). Multifractal analysis of the WTI crude oil market, US stock market and EPU. *Physica A: Statistical Mechanics and its Applications*, 550, 124096.

Zhao, L. and Cui, H. (2019). Multifractal cross-correlation between policy uncertainty and exchange rates. *Physica A*, 523, 464–476.

Zournatzidou, G. and Floros, C. (2023). Hurst exponent analysis: Evidence from volatility indices and the volatility of volatility indices. *J. Risk Financ. Manag.*, 16, 272.

8.8. Appendix: *EPU US_Daily_Policy_Data*

– DFA Hurst exponent: 1.0181409379455173.

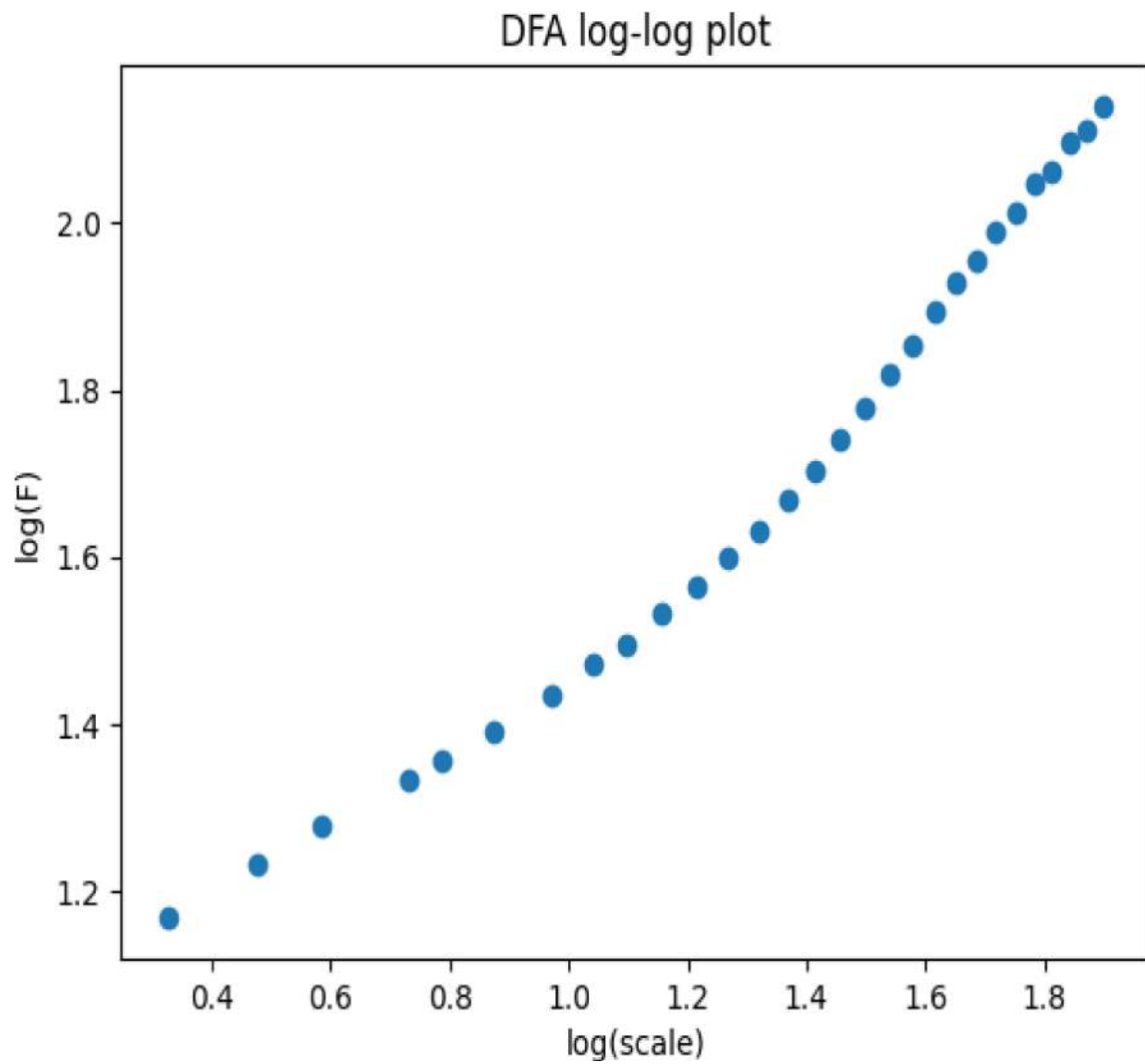


Figure 8.5. *DFA log-log plot*

- Pre-2008 crisis.
- DFA Hurst exponent: 0.8804350117002042.
- Simple Hurst exponent: 0.6985236701876388.

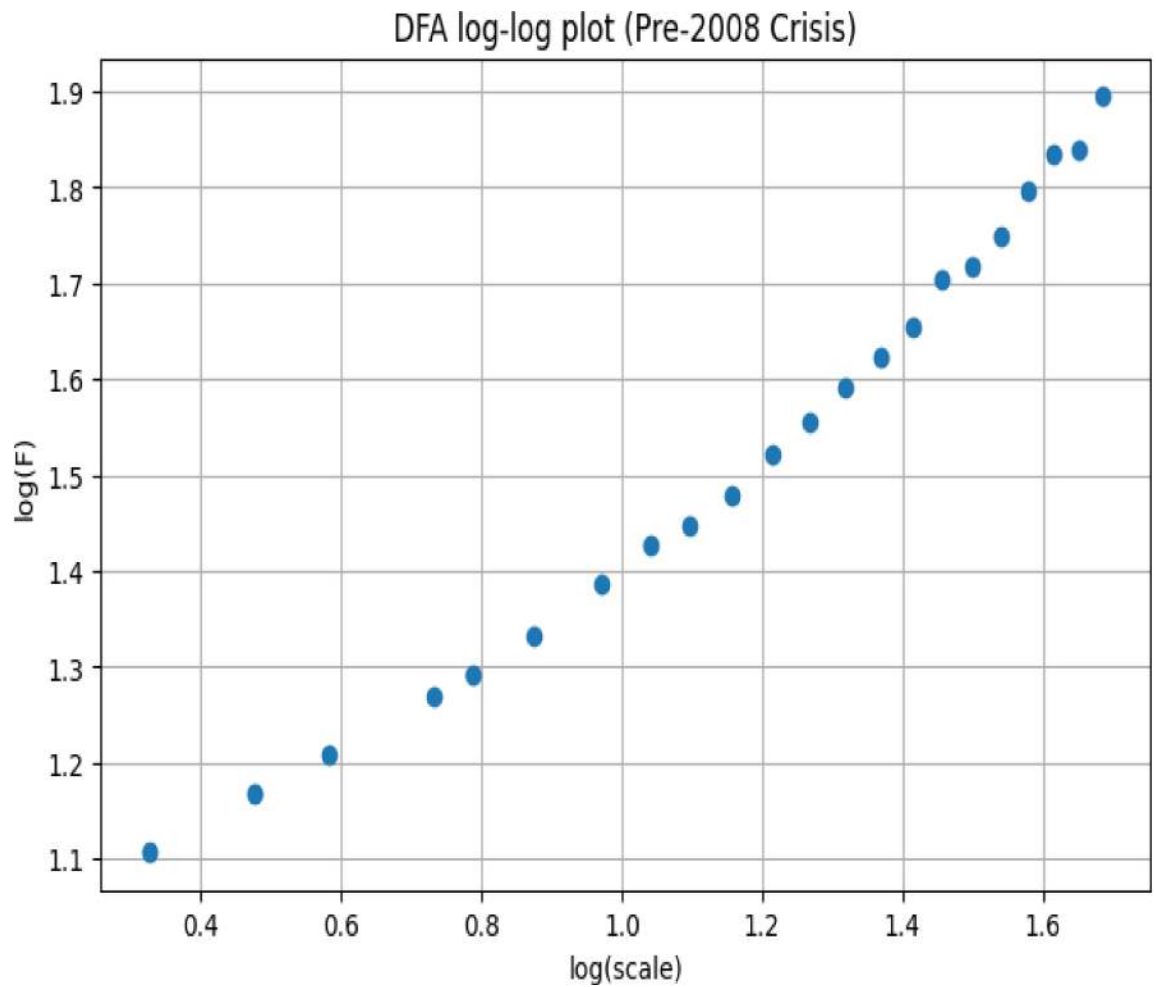


Figure 8.6. *DFA log-log plot (pre-2008 Crisis)*

- 2008 Financial crisis.
- DFA Hurst exponent: 0.7799816575586822.
- Simple Hurst exponent: 0.6394174799851401.

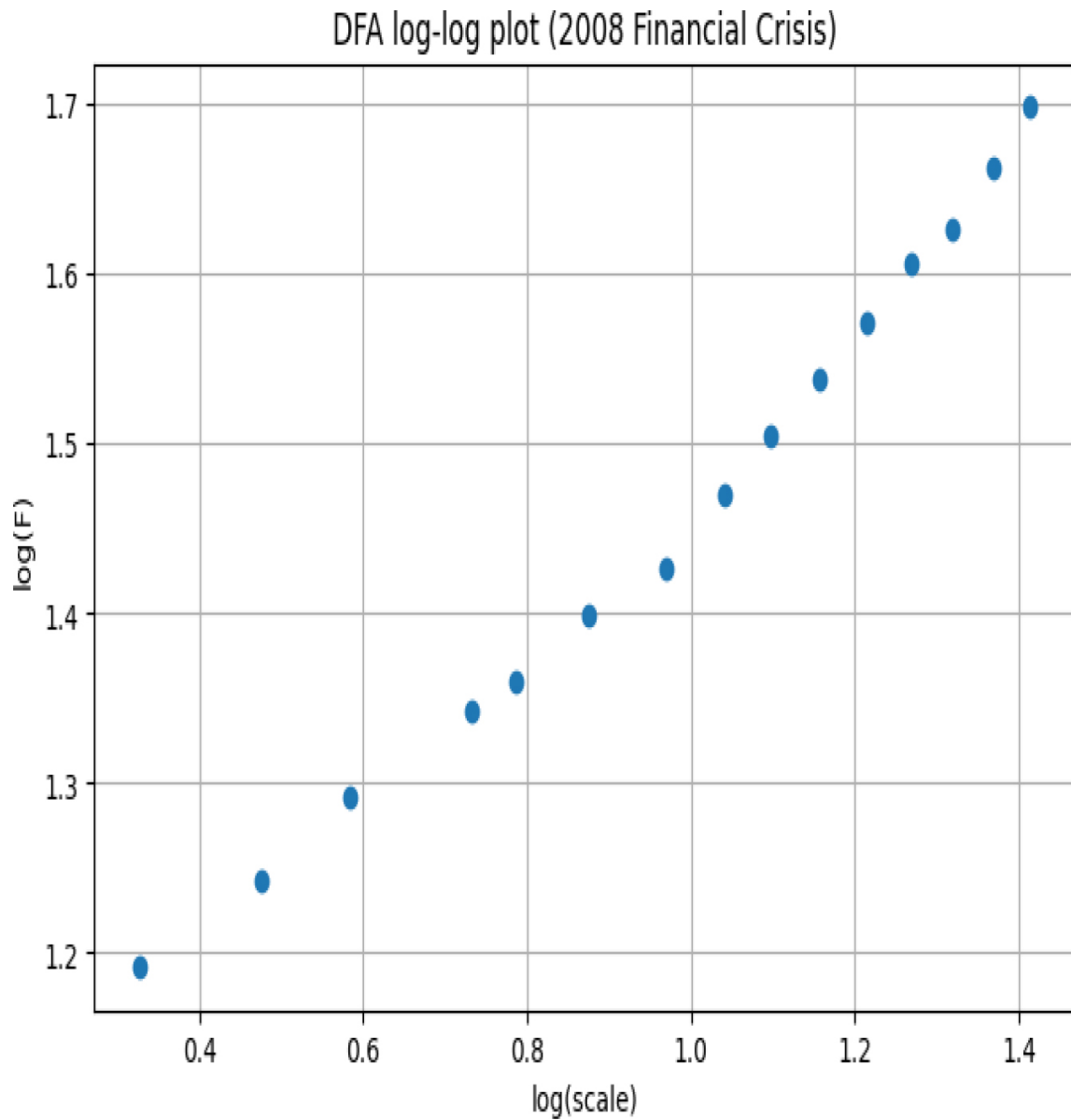


Figure 8.7. *DFA log-log plot (2008 Financial Crisis)*

- Eurozone crisis.
- DFA Hurst exponent: 0.8744420197191005.
- Simple Hurst exponent: 0.7137665272132953.

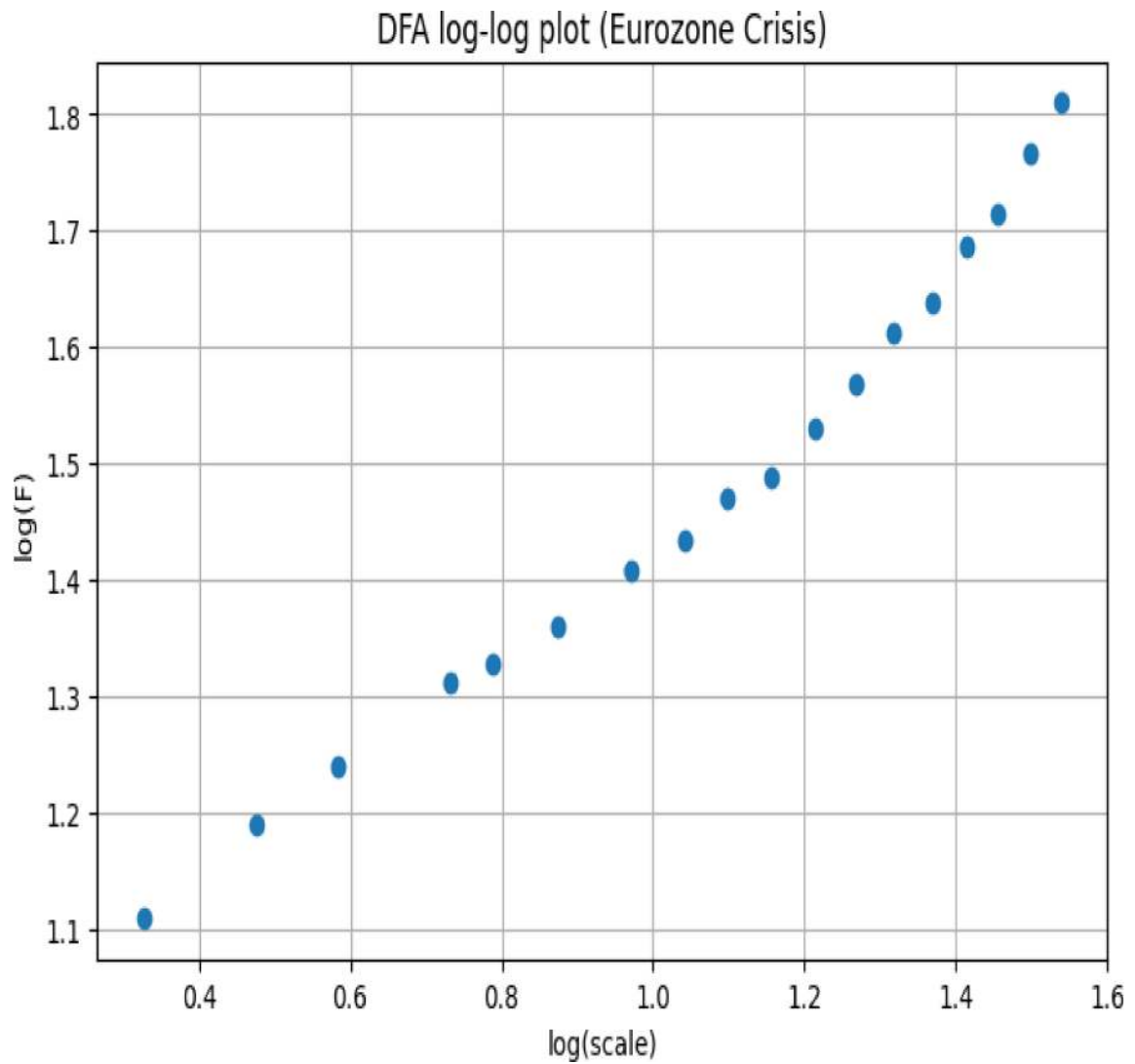


Figure 8.8. *DFA log-log plot (Eurozone Crisis)*

- Brexit period.
- DFA Hurst exponent: 0.8406134227395853.
- Simple Hurst exponent: 0.5410200388312578.

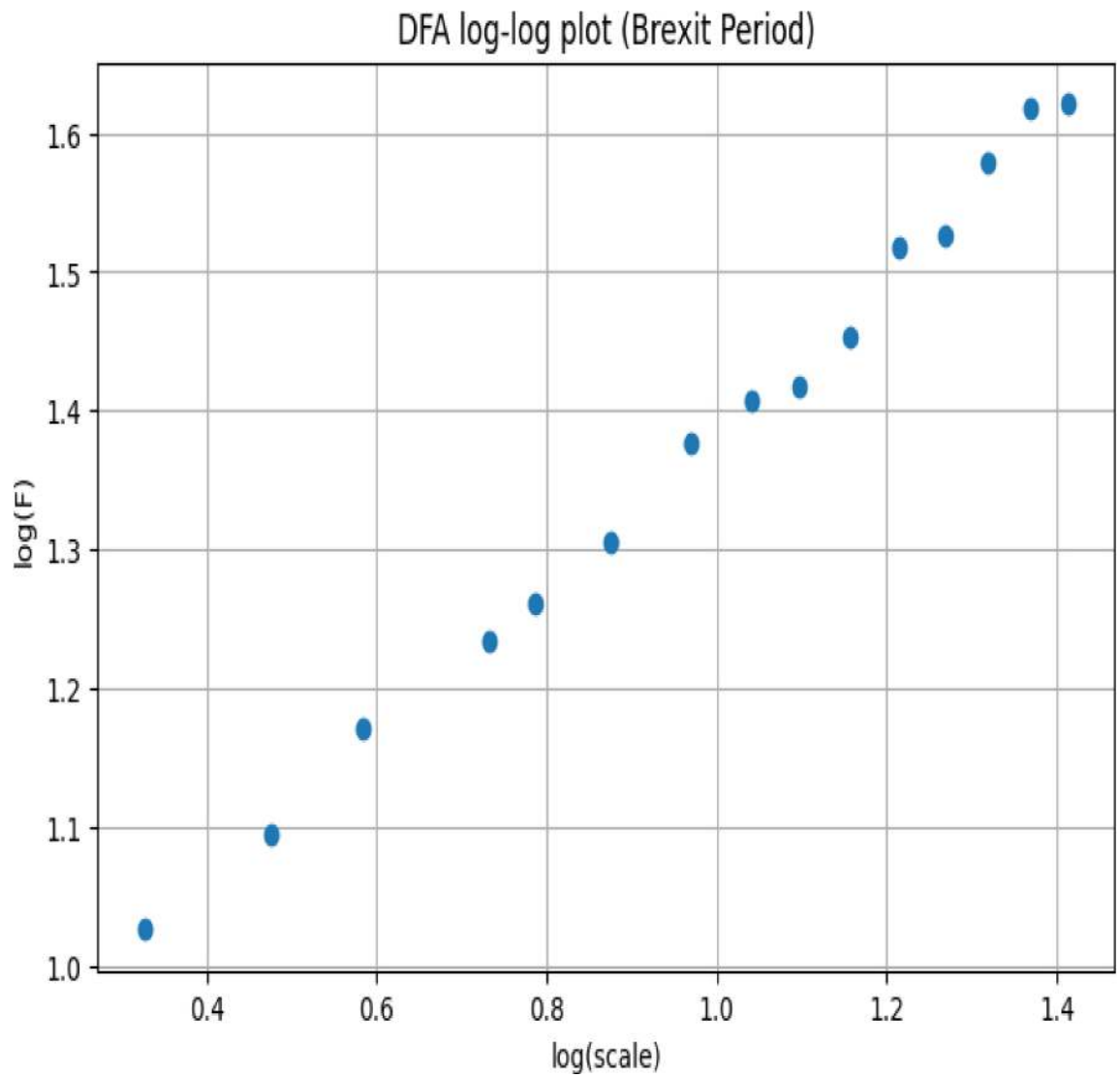


Figure 8.9. *DFA log-log plot (Brexit Period)*

- Post-Brexit/pre-Covid.
- DFA Hurst exponent: 0.8011252365392719.
- Simple Hurst exponent: 0.6960904255421574.

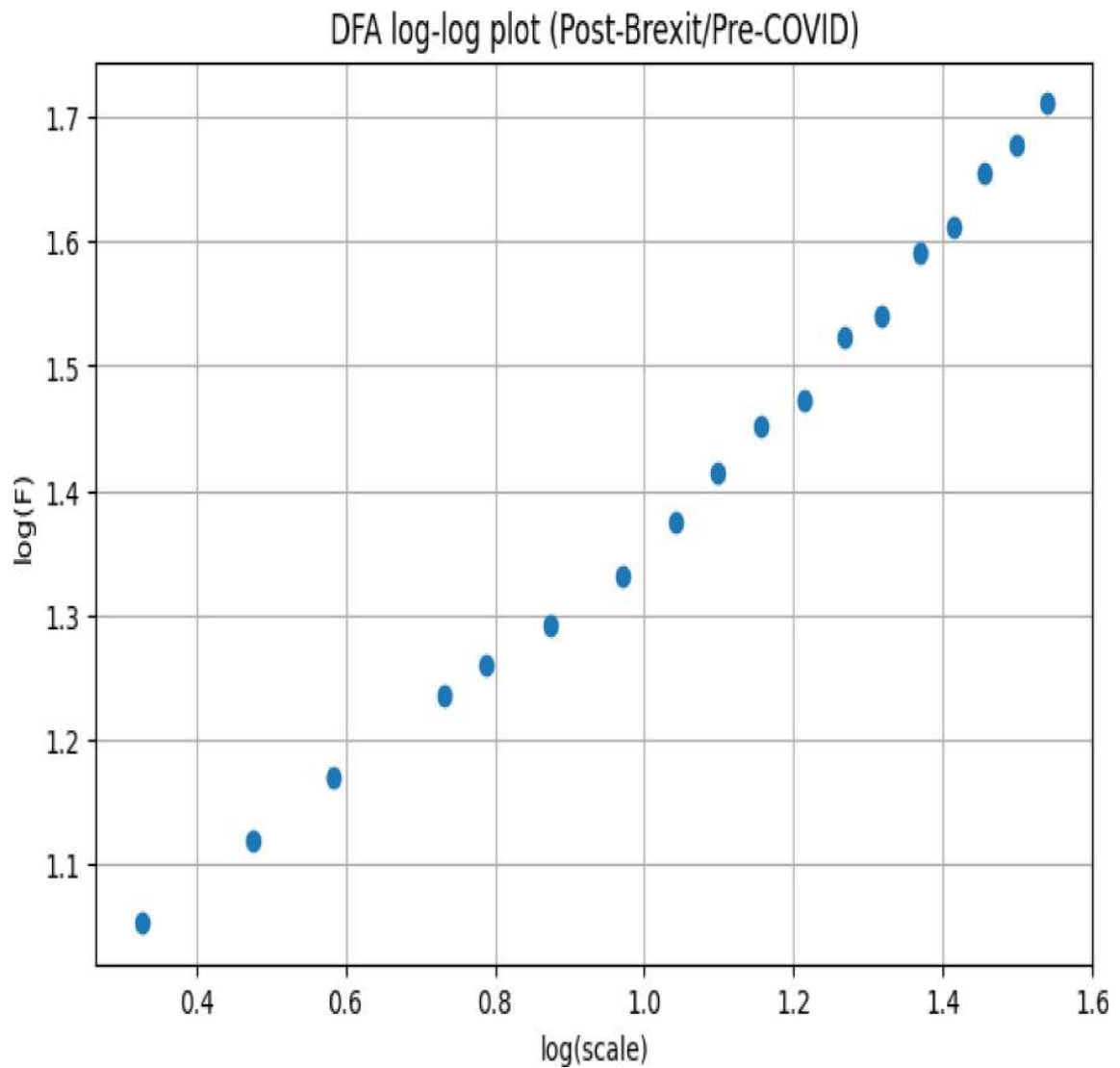


Figure 8.10. *DFA log-log plot (post-Brexit/pre-Covid)*

- Covid-19 pandemic.
- DFA Hurst exponent: 0.9445266651371033.
- Simple Hurst exponent: 0.6852596419500245.

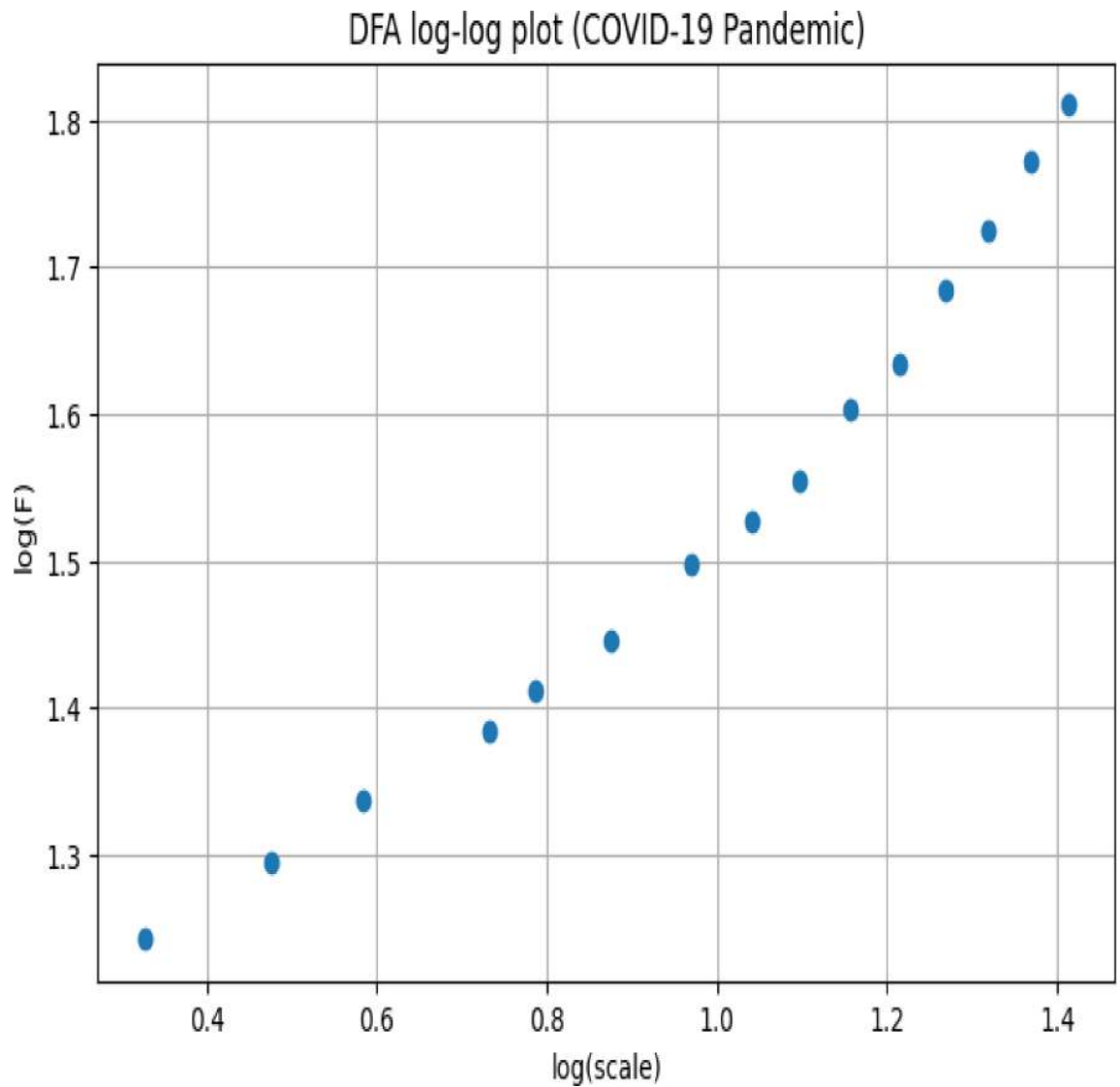


Figure 8.11. *DFA log-log plot (Covid-19 pandemic)*

- War shocks/recent events.
- DFA Hurst exponent: 0.7897184158529152.
- Simple Hurst exponent: 0.6702598992309345.

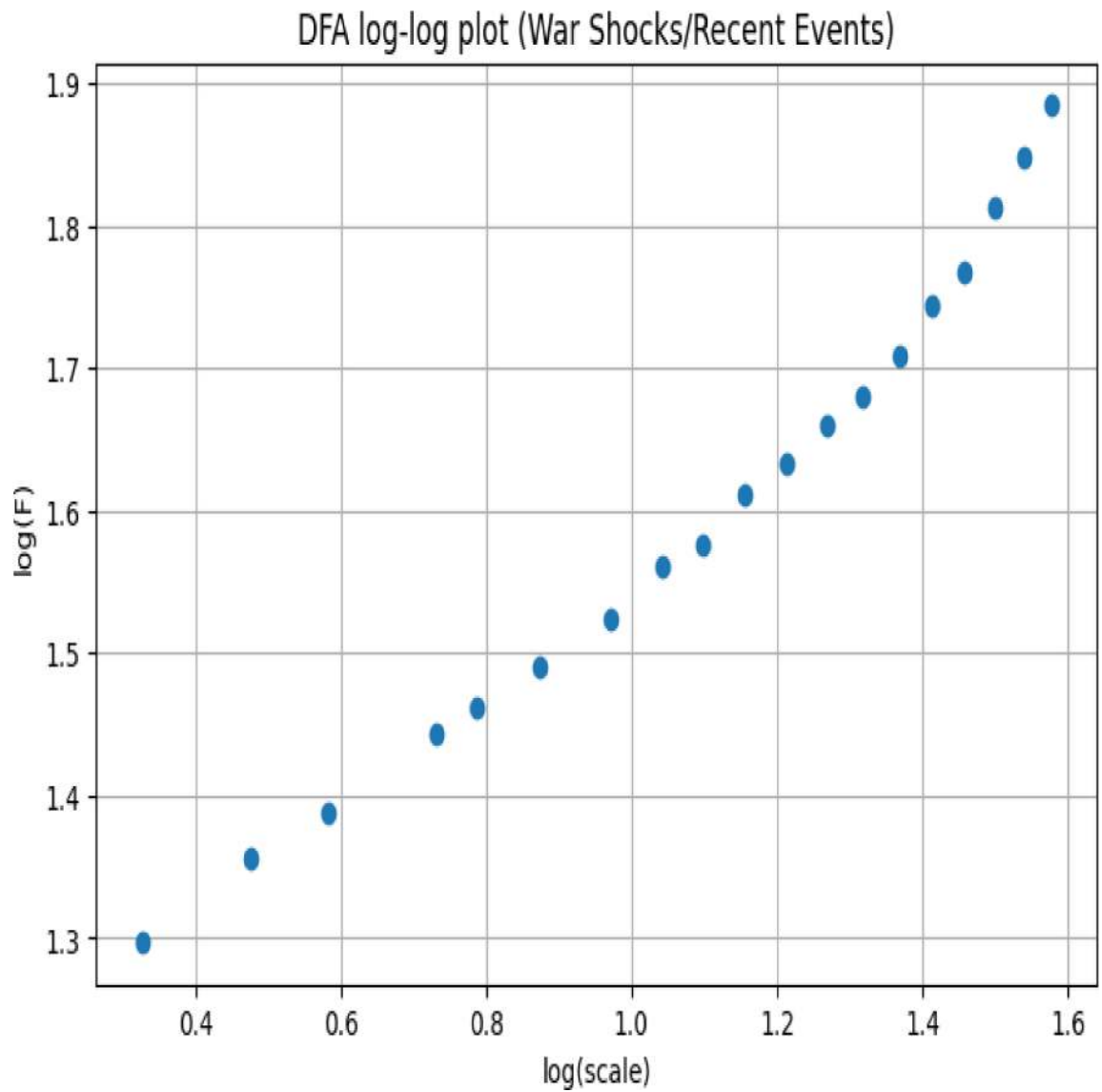


Figure 8.12. *DFA log-log plot (war shock/recent events)*

- EPU UK_Daily_Policy_Data.
- DFA Hurst exponent: 0.9831394910410282.

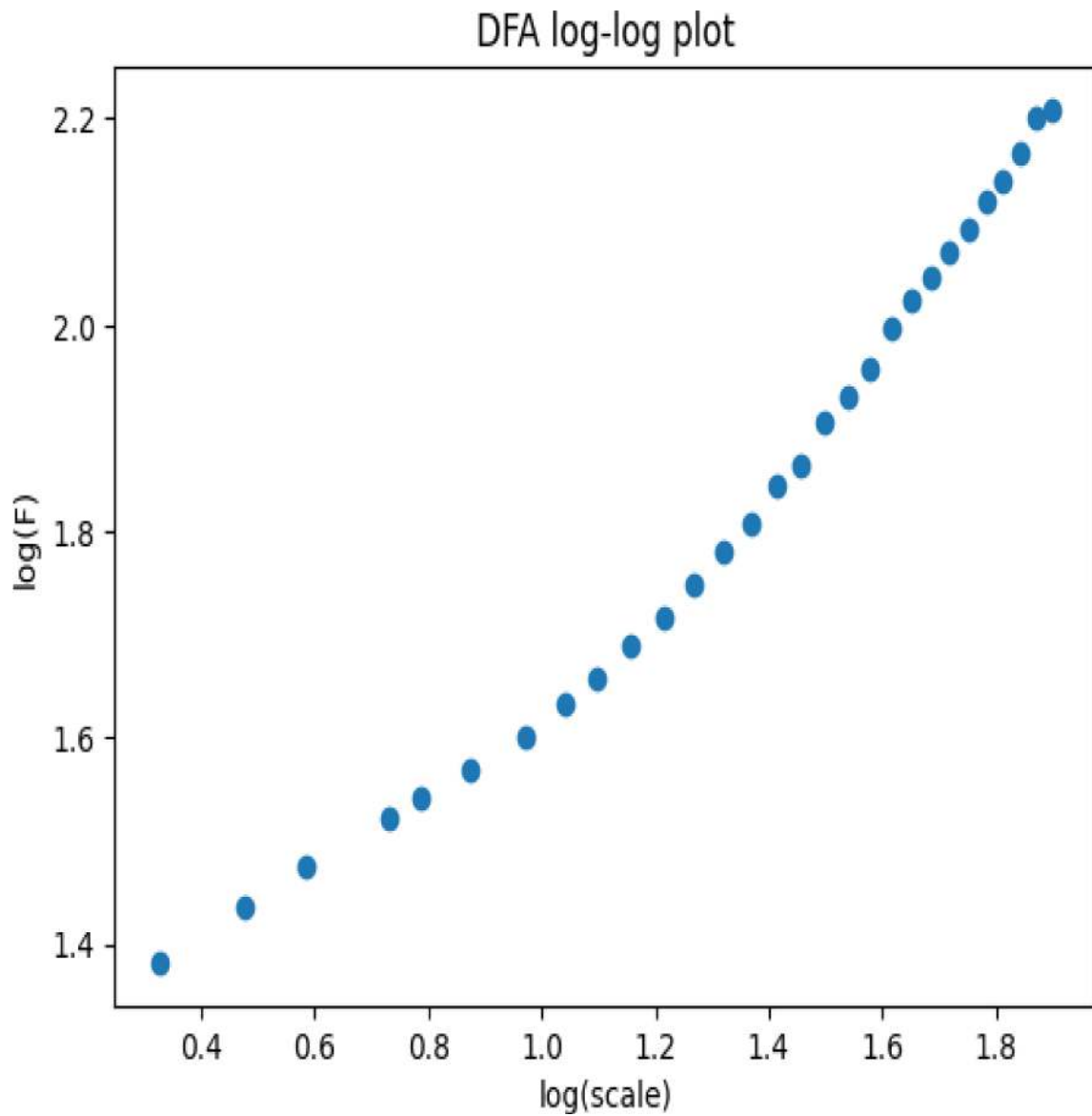


Figure 8.13. *DFA log-log plot (UK policy data)*

- Pre-2008 crisis (UK EPU).
- DFA Hurst exponent: 0.8565260458182776.
- Simple Hurst exponent: 0.7286230972490214.

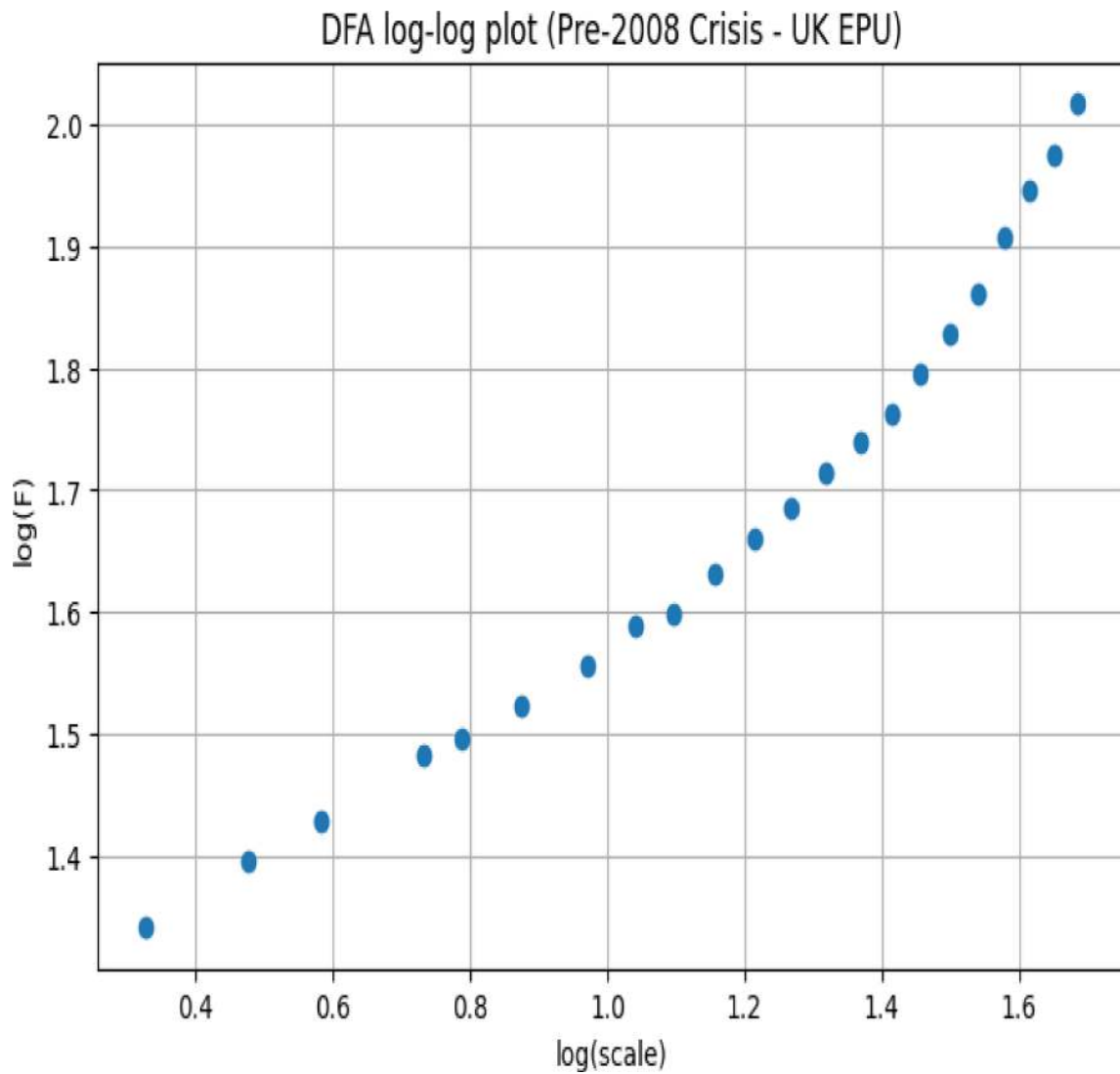


Figure 8.14. *DFA log-log plot (pre-2008 crisis – UK EPU)*

- 2008 Financial crisis (UK EPU).
- DFA Hurst exponent: 0.7208563006542942.
- Simple Hurst exponent: 0.5212473370036819.

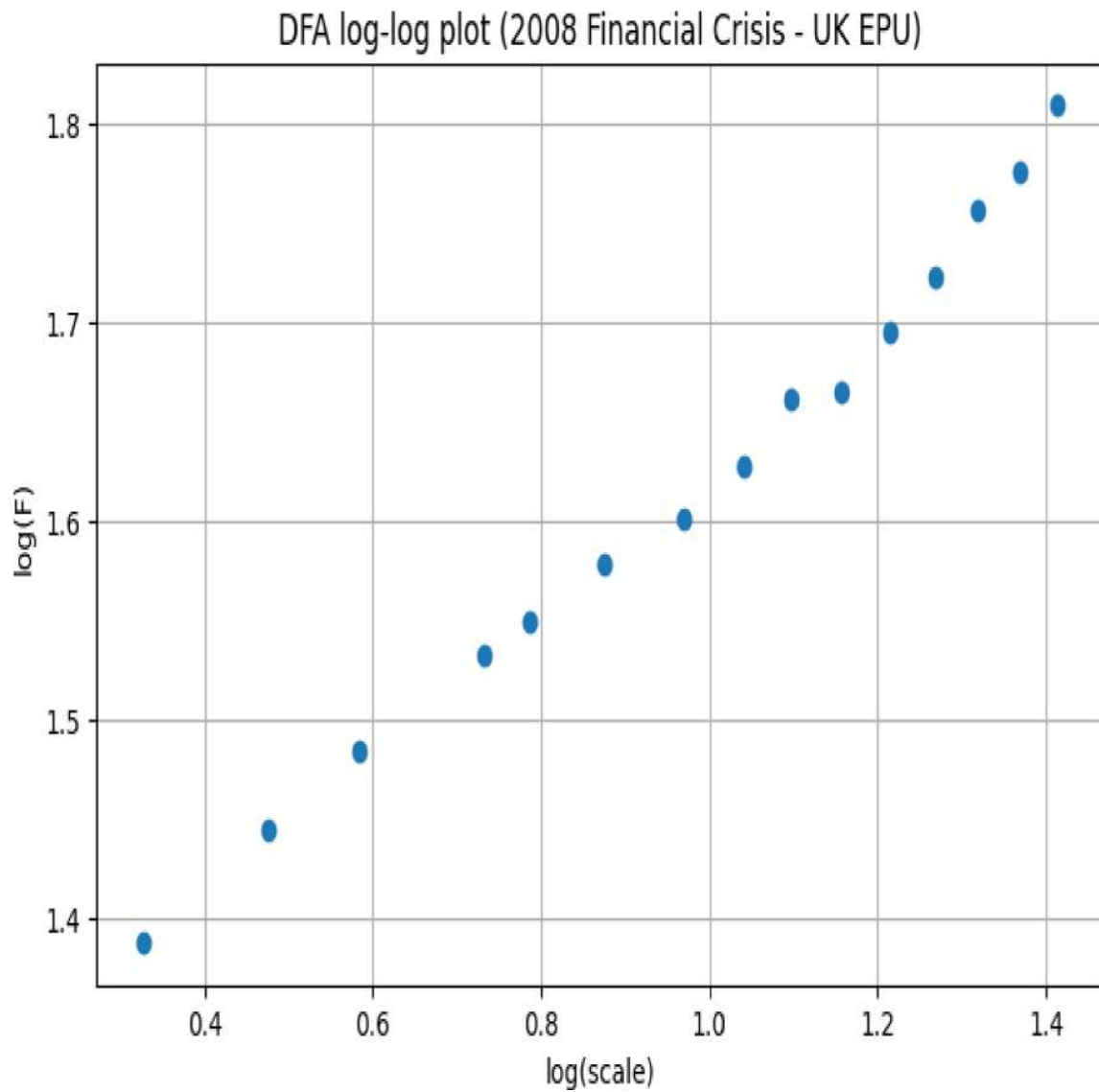


Figure 8.15. *DFA log-log plot (2008-financial crisis – UK EPU)*

- Eurozone crisis (UK EPU).
- DFA Hurst exponent: 0.728106981284717.
- Simple Hurst exponent: 0.5371580232840014.

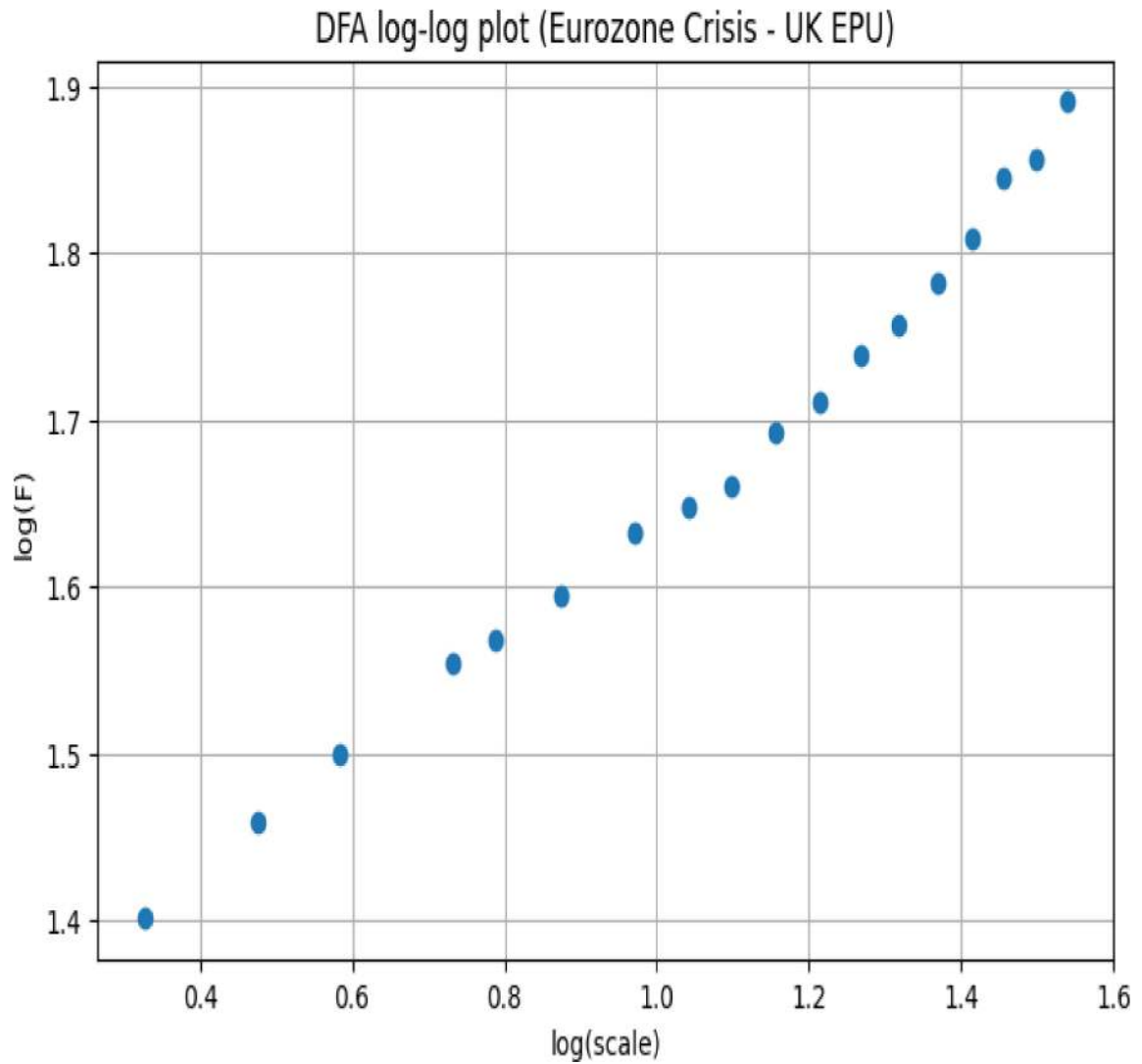


Figure 8.16. *DFA log-log plot (eurozone crisis – UK EPU)*

- Brexit period (UK EPU).
- DFA Hurst exponent: 1.0130346651571496.
- Simple Hurst exponent: 0.750208434986978.

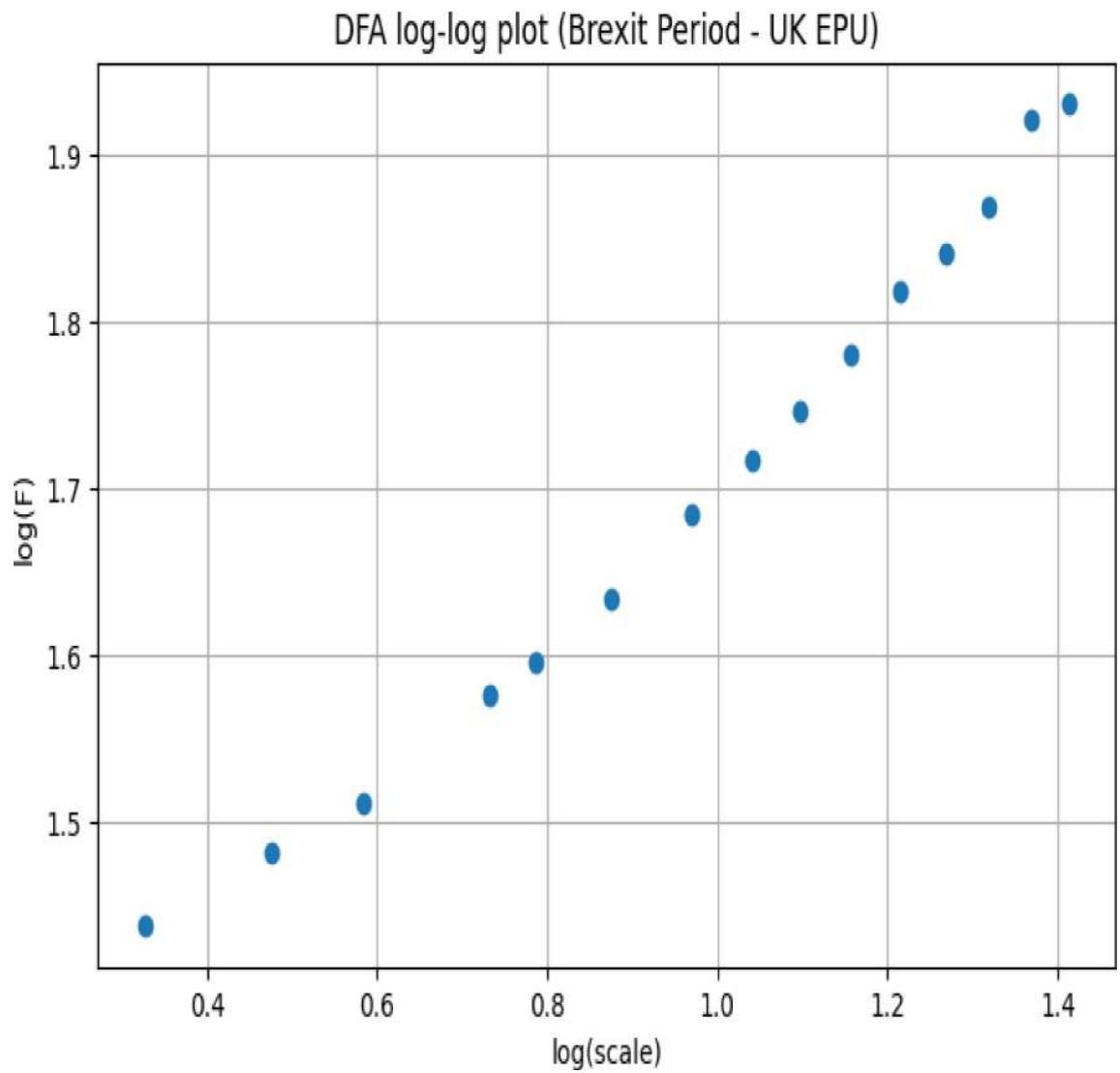


Figure 8.17. *DFA log-log plot (Brexit Period – UK EPU)*

- Post-Brexit/Pre-Covid (UK EPU).
- DFA Hurst exponent: 0.7874567201509983.
- Simple Hurst exponent: 0.7373626464884875.

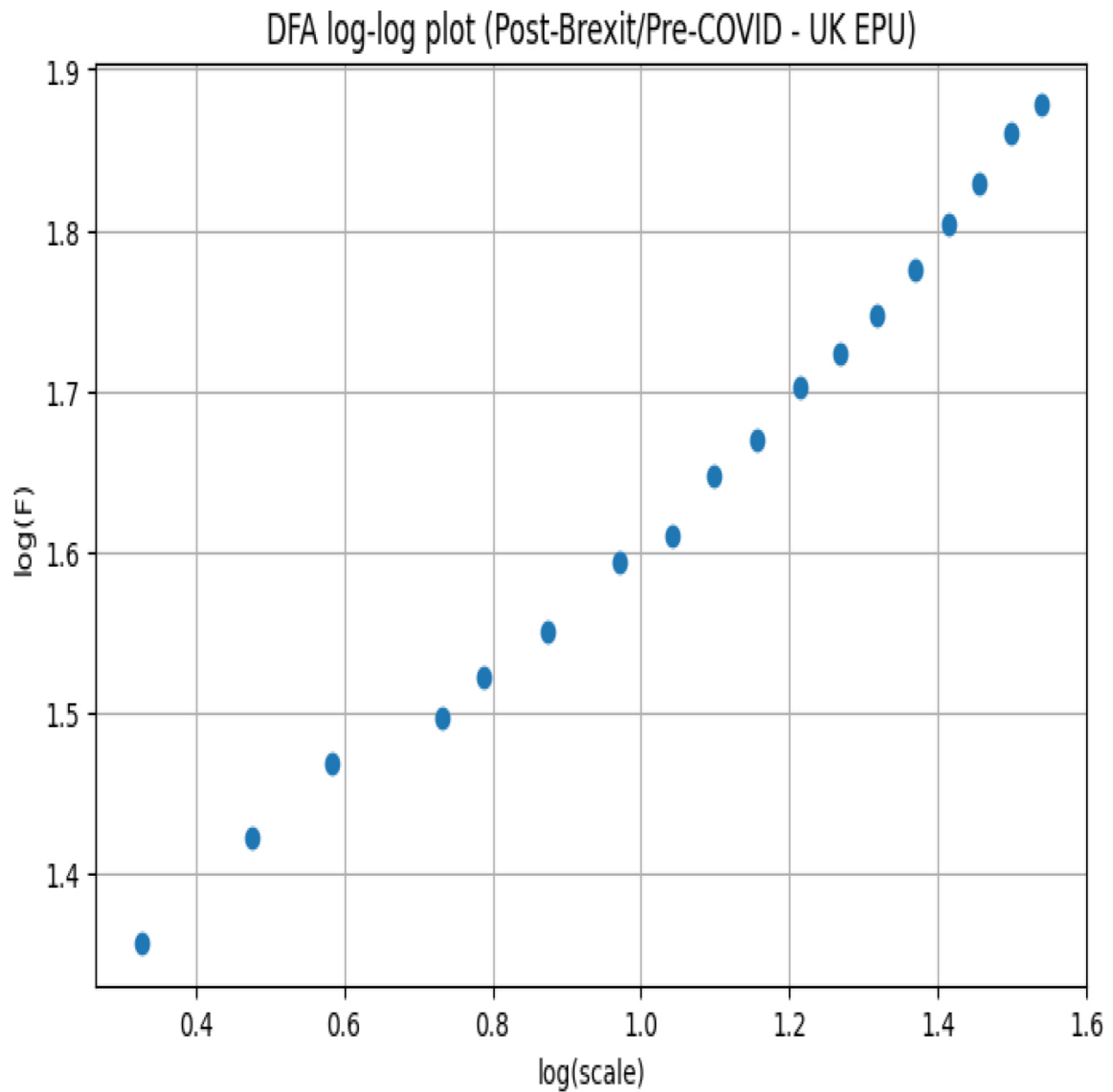


Figure 8.18. *DFA log-log plot (post-Brexit/pre-Covid – UK EPU)*

- Covid-19 pandemic (UK EPU).
- DFA Hurst exponent: 0.9150679818469877.
- Simple Hurst exponent: 0.7050696110766509.

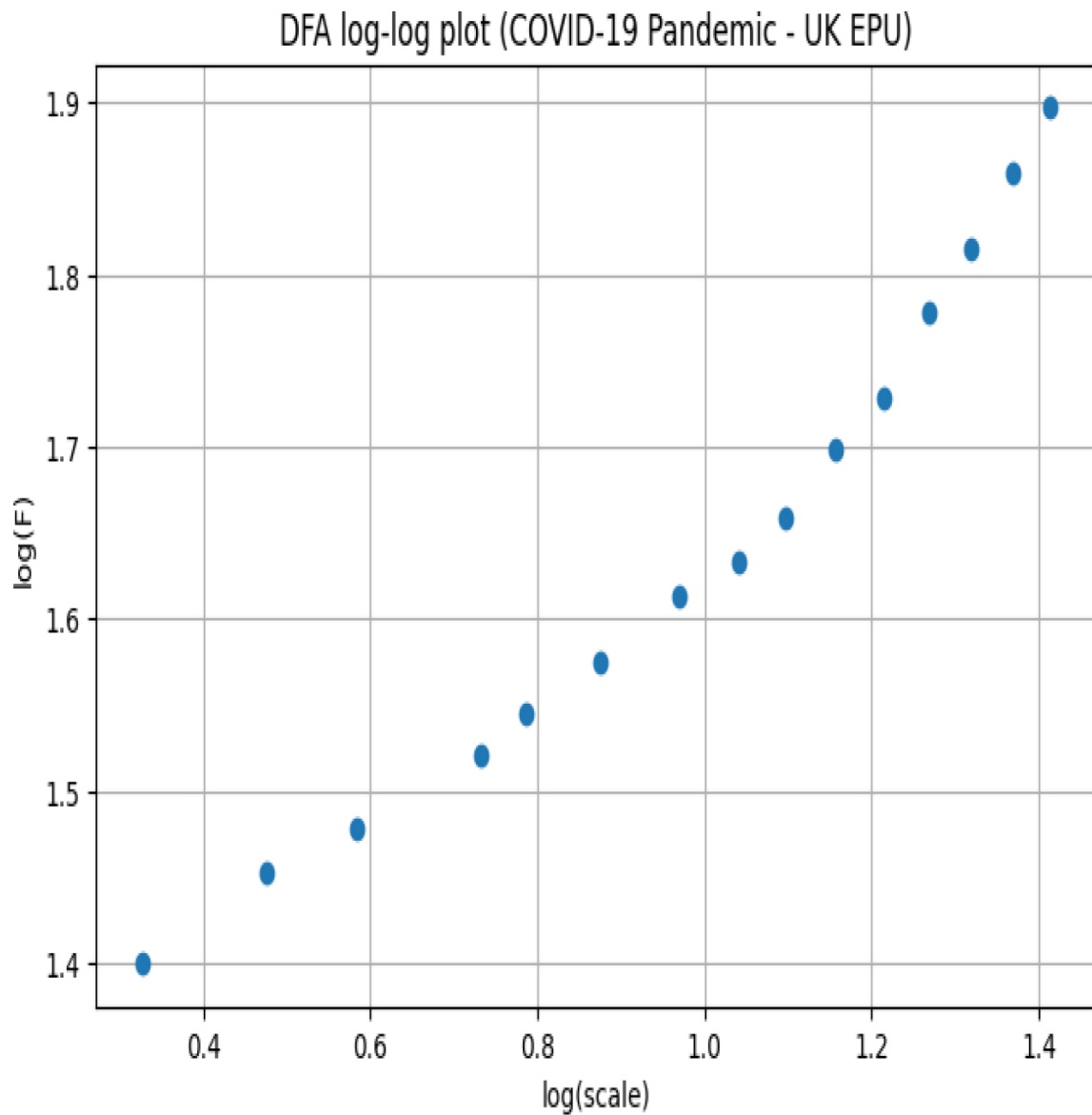


Figure 8.19. *DFA log-log plot (Covid-19 pandemic – UK EPU)*

- War shocks/recent events (UK EPU).
- DFA Hurst exponent: 0.8180734449161045.
- Simple Hurst exponent: 0.6371455258651094.

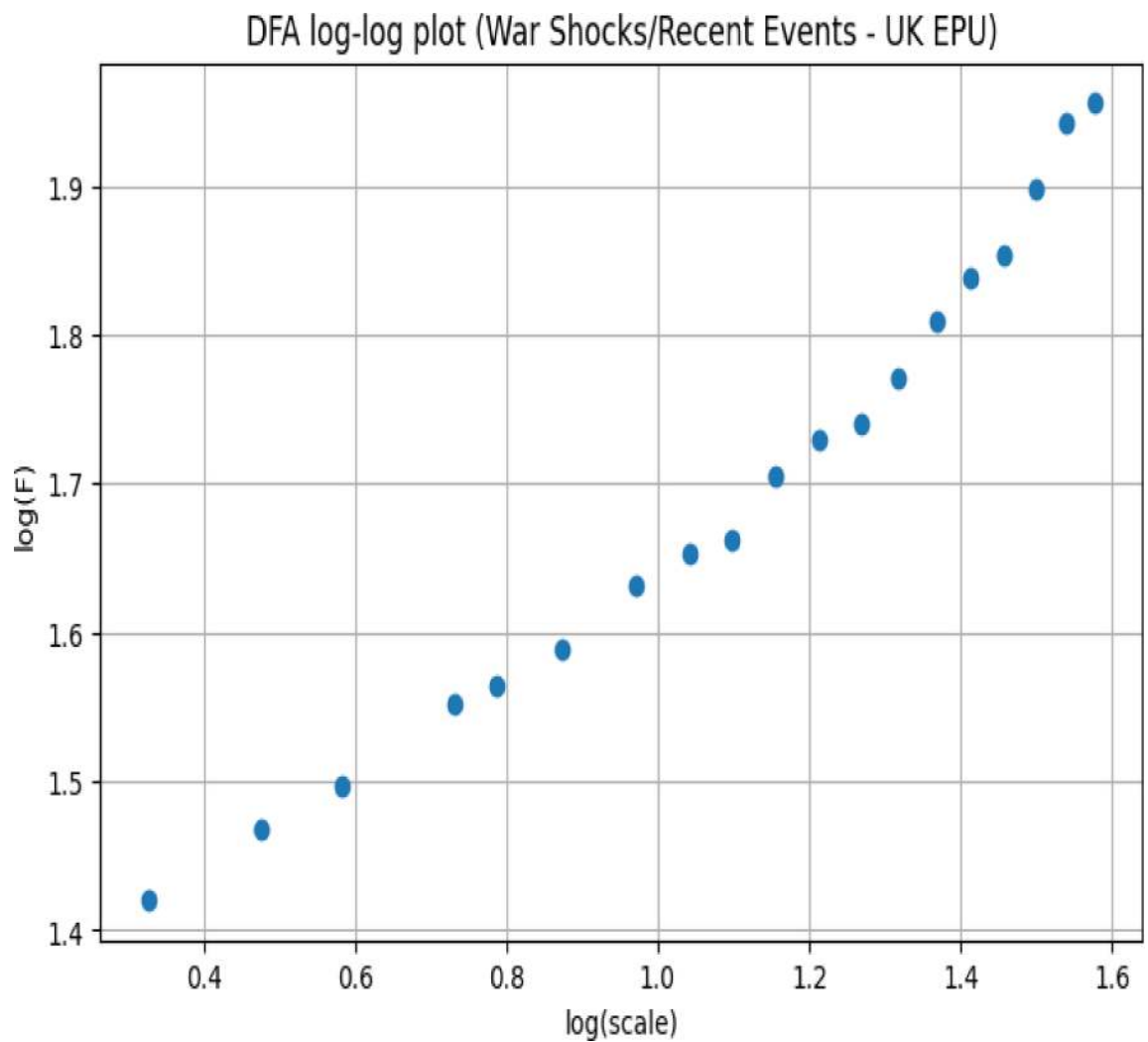


Figure 8.20. *DFA log-log plot (war shocks/recent events – UK EPU)*

Note

Chapter written by Christos FLOROS, George MATALLIOTAKIS and Fotis PAPATHEOFANOUS.

9

Spatial Dependence in Extreme Precipitation: An Extreme Value Analysis in R

Understanding and quantifying extreme meteorological events is crucial due to their significant implications for public safety, agriculture and the economy. Extreme value theory (EVT) provides a framework for analyzing rare events with potentially severe consequences. In meteorology, it is common for a single extreme event to impact multiple locations, making the assumption of independence often unrealistic. Although stationary and weakly dependent time series share the same limiting distribution as independent series, their parameters are affected by dependence structures. When data are collected across various spatial locations, accounting for spatial dependence becomes essential. This study first reviews max-stable processes – spatial extensions of multivariate EVT – and then illustrates their application through an exploratory analysis using R software. The analysis focuses on annual maxima of daily precipitation recorded from 1943 to 2023 at 18 meteorological and rainfall stations across northern Portugal, incorporating estimation of dependence measures and inference on model parameters.

9.1. Introduction

A key aim of spatial analysis in environmental sciences is to quantify and describe the behavior of phenomena such as rainfall, wind speed or daily temperature records. Although various general frameworks for spatial modeling exist (e.g. Cressie ([1993](#)); Barndorff-Nielsen et al. ([1998](#)); Ripley ([2004](#))), these often emphasize average process behavior and may not be well suited for capturing extremal characteristics. Understanding and predicting extreme environmental events – including hurricanes, storms and floods – requires dedicated methods for spatial extremes. With growing

concerns about climate change, the need for robust mathematical and statistical tools in this area has become increasingly evident. While the theory and statistical practice for univariate extremes is well established, practical guidance for modeling spatial extremes remains limited, which is a challenge given that many environmental variables naturally exhibit spatial structure.

In this study, we focus on time series of componentwise maxima observed at $k = 1, \dots, K$, sites within a defined region. Here, each observation $\{y_{n,k}\}$ represents the maximum of m measurements over $n = 1, \dots, N$ temporal blocks – for instance, when daily data are used and $m = 365$, $\{y_{n,k}\}$ represents annual maxima and N is the number of years.

The spatial counterpart to multivariate extreme-value models is the class of max-stable processes (de Haan [1984](#); de Haan and Pickands [1986](#); Resnick 1987), which share the same asymptotic basis as the univariate generalized extreme value (GEV) distribution (von Mises [1954](#); Jenkinson [1955](#)) and allow for modeling extremes while accounting for temporal or spatial dependence. Foundational work on statistical methods for max-stable processes and their practical application can be found in Smith ([1990](#)); Coles ([1993](#)); Coles and Walshaw ([1994](#)) and Coles and Tawn ([1996](#)). However, standard likelihood approaches are challenging to apply because the joint density function for these models is intractable except in very simple cases, which complicates their practical use for spatial extremes.

9.2. Asymptotic models for spatial extremes

9.2.1. *Marginal modeling of extremes*

The theory of univariate extremes is now well established and widely applied across many fields; see Coles ([2001](#)) for a clear introduction. Suppose $Y_1, Y_2, \dots$ is a sequence of independent random variables sharing a common distribution F , with finite upper endpoint $y_F = \sup\{y \in \mathbb{R}: F(y) < 1\}$. Let $M_n = \max(Y_1, Y_2, \dots, Y_n) \sim F^n$, as $n \rightarrow \infty$. A classical result in EVT states that if there exist normalizing sequences $a_n > 0$ and $b_n \in \mathbb{R}$, then the linearly

rescaled maximum M_n converges in distribution, as $n \rightarrow \infty$, to a non-degenerate limit. This fundamental result forms the basis for modeling the behavior of extreme values, that is:

$$\frac{M_n - b_n}{a_n} \xrightarrow{D} Z \sim G, \quad [9.1]$$

then G must be a non-degenerate limit distribution, with $\xrightarrow{D}$ denoting convergence in distribution. The limit law can be written as:

$$G(z) = \exp \left\{ - \left[1 + \xi \frac{(z - \mu)}{\sigma} \right]^{-1/\xi} \right\}, \quad \text{for } 1 + \xi \frac{(z - \mu)}{\sigma} > 0, \quad [9.2]$$

where $\mu \in \mathbb{R}$, $\sigma > 0$, and $\xi \in \mathbb{R}$ are the location, scale and shape parameters, respectively. This family, known as the generalized extreme value (GEV) distribution, describes tail behavior: $\xi < 0$ implies a finite upper tail, $\xi = 0$ indicates a light tail and $\xi > 0$ represents a heavy tail.

A key feature of G is its property of max-stability: for any $t > 0$, there exist constants $\alpha_t > 0$ and $\beta_t \in \mathbb{R}$ such that for all z :

$$G^t(\alpha_t z + \beta_t) = G(z). \quad [9.3]$$

This means that the distributions G and G^t share the same type; in other words, they belong to the same location-scale family. Put differently, both G and G^t are GEV distributions with an identical shape parameter ξ , differing only in their location and scale parameters.

9.2.2. Max-stable processes

By analogy with equation [9.3], a stochastic process $Z(\mathbf{s})$, defined over a region $\mathbf{s} \in \mathcal{S} \subset \mathbb{R}^d$, is said to be *max-stable* if, for any finite collection of locations $\mathbf{s}_1, \dots, \mathbf{s}_D \in \mathcal{S}$ and any positive real

number $t > 0$, there exist functions $\alpha_t(\mathbf{s}) > 0$ and $\beta_t(\mathbf{s})$ such that, for all values $z_1, \dots, z_D$, the following equality holds:

$$\begin{aligned} \Pr \{Z(\mathbf{s}_1) \leq \alpha_t(\mathbf{s}_1)z_1 + \beta_t(\mathbf{s}_1), \dots, Z(\mathbf{s}_D) \leq \alpha_t(\mathbf{s}_D)z_D + \beta_t(\mathbf{s}_D)\}^t &= \\ \Pr \{Z(\mathbf{s}_1) \leq z_1, \dots, Z(\mathbf{s}_D) \leq z_D\}. \end{aligned} \quad [9.4]$$

Comparing equation [9.4] with equation [9.3], one sees that *max-stable* processes necessarily have marginal distributions of the GEV type. To better analyze the spatial dependence structure, it is convenient to standardize the process so that all margins are on a common scale. Assuming that the process $Z(\mathbf{s})$ has unit Fréchet marginal distributions, that is:

$$\Pr\{Z(\mathbf{s}) \leq z\} = \exp\{-1/z\}, \quad z > 0, \quad [9.5]$$

which corresponds to a GEV distribution with parameters $\mu = \sigma = \xi = 1$, one can take $\alpha_t(\mathbf{s}_j) = t$ and $\beta_t(\mathbf{s}_j) = 0$, reducing equation [9.4] to the simpler form:

$$\Pr \{Z(\mathbf{s}_1) \leq tz_1, \dots, Z(\mathbf{s}_D) \leq tz_D\}^t = \Pr \{Z(\mathbf{s}_1) \leq z_1, \dots, Z(\mathbf{s}_D) \leq z_D\}. \quad [9.6]$$

Just as the GEV distribution is the only non-degenerate limiting distribution for appropriately rescaled maxima of independent and identically distributed random variables, *max-stable* processes arise as the only non-degenerate limits for normalized maxima of random fields with non-degenerate marginal distributions.

Let $\mathbf{s} \in \mathcal{S} \subset \mathbb{R}^d$, and consider a sequence of independent and identically distributed stochastic processes $Y_1(\mathbf{s}), Y_2(\mathbf{s}), \dots$. Define the pointwise maximum process by $M_n(\mathbf{s}) = \max_{i=1, \dots, n} Y_i(\mathbf{s})$. If there exist sequences of normalizing functions $a_n(\mathbf{s}) > 0$ and $b_n(\mathbf{s})$ such that the rescaled process $a_n(\mathbf{s})^{-1} \{M_n(\mathbf{s}) - b_n(\mathbf{s})\}$ converges, as $n \rightarrow \infty$, in distribution to a process $Z(\mathbf{s})$ with non-degenerate margins, then

$Z(\mathbf{s})$ is a *max-stable* process that satisfies equation [9.4]. This asymptotic characterization justifies the use of max-stable processes as a fundamental tool for modeling spatial extremes.

An effective approach for constructing and characterizing *max-stable* processes is based on the spectral representation proposed by de Haan (1984). For more details, see also Schlather (2002) and Chapter 9 of de Haan and Ferreira (2006).

Let $R_1, R_2, \dots$ be points of a Poisson process on the interval $[0, +\infty)$ with intensity $r^{-2} dr$, and let $W_1(s), W_2(s), \dots$ be independent replicates of a non-negative stochastic process $W(s)$, defined such that $W(s) \geq 0$ and $\mathbb{E}[W(s)] = 1$. Then, a *max-stable* process $Z(s)$ with unit Fréchet margins can be constructed through the representation:

$$Z(s) = \sup_{i=1,2,\dots} R_i W_i(s). \quad [9.7]$$

Intuitively, this formulation can be interpreted as a pointwise supremum of infinitely many random fields of the form $R_i W_i(\mathbf{s})$, where each term can be viewed as representing a “storm” characterized by amplitude R_i and spatial profile $W_i(\mathbf{s})$. Due to the heavy-tailed behavior inherent in the power-law intensity of $\{R_i\}$, the process exhibits asymptotic dependence and, by construction, satisfies the property of max-stability.

From representation [9.7], the finite-dimensional distributions of the random vector $\{Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_D)\}^T$ are given by:

$$G(z_1, \dots, z_D) = \Pr\{Z(\mathbf{s}_1) \leq z_1, \dots, Z(\mathbf{s}_D) \leq z_D\} = \exp\{-V(z_1, \dots, z_D)\}, \quad [9.8]$$

where $V(z_1, \dots, z_D)$ is referred to as the exponent function and can be expressed as:

$$V(z_1, \dots, z_D) = \mathbb{E} \left[\max \left\{ \frac{W(s_1)}{z_1}, \dots, \frac{W(s_D)}{z_D} \right\} \right]^{[9.9]}.$$

By choosing different specifications for the process W , various families of max-stable models can be derived. Notable examples include the Schlather model and its extensions (see Schlather (2002); Davison and Gholamrezaee (2012)), the Brown–Resnick model (Brown and Resnick 1977; Kabluchko 2009) and the extremal- t model (Opitz 2013).

9.2.3. Spatial models

Max-stable processes provide a flexible framework for modeling spatial extremes. Two important spatial models within this class are the *Smith model* and the *Schlather model*, both of which rely on different stochastic process constructions to induce spatial dependence with unit Fréchet margins.

9.2.3.1. Smith model

A classical example of a spatial *max-stable* process that fits within the spectral representation framework is the *Smith model*. In this case, each stochastic replicate $W_i(s)$ is defined as a shifted probability density function, that is:

$$W_i(s) = g(s - X_i),$$

where g is typically chosen to be a multivariate normal density on $\mathbb{R}^d$, and $\{X_i\}$ are points of a homogeneous Poisson process on $\mathbb{R}^d$. Each term $R_i W_i(s)$ in the representation:

$$Z(s) = \sup_{i=1,2,\dots} R_i W_i(s),$$

can be interpreted as a “storm” centered at location X_i , with intensity R_i and spatial structure governed by the kernel g . The resulting process $Z(s)$ captures the maximum impact over an infinite number of such random storms. The *Smith model* is particularly appealing

due to its interpretability and its elliptical level sets, which naturally arise from the use of Gaussian densities. When g is the multivariate normal density, the model is often referred to as the *Smith model*, originally introduced by Smith (1990).

If g is chosen as a multivariate Gaussian density with covariance matrix Σ , the resulting process is termed a *Gaussian extreme value process*. The joint distribution at two locations $\mathbf{s}_1$ and $\mathbf{s}_2$ is then given by:

$$\begin{aligned} \Pr\{Z(\mathbf{s}_1) \leq z_1, Z(\mathbf{s}_2) \leq z_2\} = \exp \left\{ -\frac{1}{z_1} \Phi \left(\frac{a}{2} + \frac{1}{a} \log \frac{z_2}{z_1} \right) \right. \\ \left. - \frac{1}{z_2} \Phi \left(\frac{a}{2} + \frac{1}{a} \log \frac{z_1}{z_2} \right) \right\}, \end{aligned} \quad [9.10]$$

where $a = \sqrt{(\mathbf{s}_1 - \mathbf{s}_2)^\top \Sigma^{-1} (\mathbf{s}_1 - \mathbf{s}_2)}$ is the Mahalanobis distance between the locations, and Φ is the standard normal cumulative distribution function. The parameter a governs the strength of spatial dependence: as $a \rightarrow 0$, the process becomes perfectly dependent, while $a \rightarrow \infty$ leads to independence.

9.2.3.2. Schlather model

An alternative specification within the spectral representation framework is the *Schlather model*, where each process $W_i(\mathbf{s})$ corresponds to a transformation of a stationary standard Gaussian process with correlation function $\rho(h)$.

Specifically, $W_i(\mathbf{s}) = \max\{0, \varepsilon_i(\mathbf{s})\}$, where $\varepsilon_i(\mathbf{s})$ are independent realizations of a mean-zero, unit-variance Gaussian process satisfying $\mathbb{E}[\max\{0, \varepsilon_i(\mathbf{s})\}] = 1$, to ensure unit Fréchet margins in the resulting process.

The *max-stable* process is then constructed as:

$$Z(\mathbf{s}) = \sup_{i=1,2,\dots} R_i \max\{0, \varepsilon_i(\mathbf{s})\}, \quad [9.11]$$

where the R_i are points from a Poisson process on $[0, +\infty)$ with intensity $r^{-2}dr$, and the $\varepsilon_i(\mathbf{s})$ have spatial correlation $\rho(h)$, with h denoting the Euclidean distance between spatial locations.

Under this construction, the bivariate exponent function for $Z(\mathbf{s}_1)$ and $Z(\mathbf{s}_2)$, denoted $V(z_1, z_2)$, takes the form:

$$V(z_1, z_2) = \frac{1}{2} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \left(1 + \left[1 - 2 \cdot \frac{(\rho(h) + 1)z_1 z_2}{(z_1 + z_2)^2} \right]^{1/2} \right),$$

where $h = \|\mathbf{s}_1 - \mathbf{s}_2\|$ and $\rho(h)$ is a valid correlation function. The *Schlather model* is particularly attractive in spatial statistics, as it permits the use of many standard correlation functions, as sketched in [Table 9.1](#).

Table 9.1. *Examples of valid correlation functions $\rho(h)$ commonly used in the construction of the Schlather model*

Family	Correlation function $\rho(h)$	Range of validity
Whittle–Matérn	$\rho(h) = \{2^{\kappa-1}\Gamma(\kappa)\}^{-1}(h/\lambda)^\kappa K_\kappa(h/\lambda)$	$\kappa > 0$
Cauchy	$\rho(h) = \{1 + (h/\lambda)^2\}^{-\kappa}$	$\kappa > 0$
Stable	$\rho(h) = \exp\{-(h/\lambda)^\kappa\}$	$0 < \kappa \leq 2$
Exponential	$\rho(h) = \exp(-h/\lambda)$	–

9.2.4. Extremal coefficient

To summarize the strength of extremal dependence between different spatial locations, one often relies on scalar measures such as the *extremal coefficient*. This coefficient captures the tendency of extreme events to occur simultaneously across multiple sites, and is particularly useful when dealing with *max-stable* processes of the

form $Z(s) = \sup_{i \geq 1} R_i W_i(s)$, with unit Fréchet margins. Given a set of sites $\{s_1, \dots, s_D\}$, the joint distribution of the associated random vector $\{Z(s_1), \dots, Z(s_D)\}^T$ can be written as:

$$\Pr\{Z(s_1) \leq z, \dots, Z(s_D) \leq z\} = \exp \left\{ -\frac{V(1, \dots, 1)}{z} \right\} = \exp \left\{ -\frac{\theta_D}{z} \right\}, \quad [9.12]$$

for any $z > 0$, where $\theta_D := V(1, \dots, 1) \in [1, D]$ is the *extremal coefficient*. This value quantifies the effective number of independent components in the maxima: $\theta_D = 1$ corresponds to complete dependence, while $\theta_D = D$ indicates independence.

In the bivariate case $D = 2$, the extremal coefficient $\theta(h)$ depends only on the spatial lag $h = \|s_1 - s_2\|$, assuming stationarity and isotropy. An equivalent interpretation is given by the limit:

$$\lim_{z \rightarrow \infty} \Pr\{Z(s_2) > z \mid Z(s_1) > z\} = 2 - \theta(h),$$

which represents the conditional probability of an extreme event at s_2 given an extreme event at s_1 . Hence, lower values of $\theta(h)$ imply stronger extremal dependence.

To estimate $\theta(h)$ from data, one common approach involves the so-called F -madogram (Cooley et al. [2006](#)), defined as:

$$\nu_F(h) = \frac{1}{2} \mathbb{E} [|F(Z(s_1)) - F(Z(s_2))|],$$

where $F(z) = \exp\{-1/z\}$ denotes the unit Fréchet distribution function. The madogram $\nu_F(h) \in [0, 1/2]$ summarizes the variation between sites and is related to the *extremal coefficient* by the identity:

$$\theta(h) = \frac{1 + 2\nu_F(h)}{1 - 2\nu_F(h)}.$$

This relationship is bijective, allowing empirical estimates of $\nu_F(h)$ to be translated directly into estimates of $\theta(h)$. The F -madogram remains finite even when working with heavy-tailed marginals, and has been extended to spatial frameworks by Naveau (2009), enabling estimation of extremal dependence over a continuous spatial domain.

9.2.4.1. Extremal coefficient for the Smith model

For the *Smith model*, where the process $W_i(\mathbf{s})$ is derived from a multivariate Gaussian density centered at random storm locations, the *extremal coefficient* for a pair of locations $\mathbf{s}_1, \mathbf{s}_2$ depends on the Mahalanobis distance between them:

$$a(h) = (h^\top \Omega^{-1} h)^{1/2}, \quad \text{where } h = \|\mathbf{s}_1 - \mathbf{s}_2\|,$$

and Ω is the covariance matrix of the Gaussian kernel $g(\cdot)$. The bivariate *extremal coefficient* is then given by:

$$\theta_{\text{Smith}}(h) = 2 \Phi \left(\frac{a(h)}{2} \right),$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function. This expression satisfies $\theta(h) \rightarrow 1$ as $h \rightarrow 0$, indicating perfect dependence at zero distance, and $\theta(h) \rightarrow 2$ as $h \rightarrow \infty$, which corresponds to asymptotic independence.

9.2.4.2. Extremal coefficient for the Schlather model

In the *Schlather model*, where $W_i(s) = \max\{0, \varepsilon_i(s)\}$ and $\varepsilon_i(s)$ is a stationary standard Gaussian process with correlation function $\rho(h)$, the *extremal coefficient* is:

$$\theta_{\text{Schlather}}(h) = 1 + \left(\frac{1 - \rho(h)}{2} \right)^{1/2},$$

where $\rho(h) \in [0, 1]$. This form also satisfies $\theta(h) = 1$ when $\rho(h) = 1$ (perfect dependence), but due to the constraint $\rho(h) \geq 0$ from the requirement of positive definiteness, the maximum attainable value of $\theta(h)$ is $1 + 1/\sqrt{2} \approx 1.707$ and never reaches full independence ($\theta = 2$). Hence, the *Schlather model* cannot fully capture asymptotic independence at large distances.

To contextualize the methodological framework discussed in this study, we present an illustrative example drawn from environmental data, highlighting the type of application our work aims to support. We outline the key steps involved in implementing these methods in R software, focusing on the estimation of dependence measures and parameter inference within a spatial context.

9.3. An application using the R environment – rainfall data in the North of Portugal

In this section, we apply the methodology outlined previously to model extreme geostatistical rainfall data for a region located in northern Portugal. The analysis is conducted using the R software and several packages, such as `evd`, `ismev` and mainly `SpatialExtremes` (Ribatet [2022](#)).

This region is among the wettest in the country, a result of its proximity to the Atlantic Ocean combined with the influence of mountainous terrain. Accurately modeling the spatial distribution of extreme rainfall is critical for informing disaster preparedness and mitigation strategies, particularly in light of the heightened risk of flooding and other severe weather events. Reliable estimates of extreme precipitation are also essential for engineers involved in designing resilient infrastructure capable of withstanding such conditions.

A significant challenge in this context is the limited spatial coverage of rainfall data, as measurements are typically available only from a

finite number of fixed monitoring stations. To effectively model extreme precipitation and estimate upper quantiles, it is important to incorporate spatial dependence between observation sites.

In this study, we use rainfall data collected from 18 meteorological and hydrological stations. The data sources include the Portuguese Institute for Sea and Atmosphere (Instituto Português do Mar e da Atmosfera (IPMA)) and the National Water Resources Information System (Sistema Nacional de Informação de Recursos Hídricos (SNIRH)), administered by the Portuguese Water Resources Institute.

While additional monitoring stations are present within the study area, many exhibit several missing values in their historical records, limiting their suitability for rigorous statistical analysis. [Figure 9.1](#) illustrates the geographical distribution of the 18 stations selected for this study and shows spatial variability in precipitation patterns across the region.

To specifically target extreme precipitation events, we extracted the annual maxima from the daily rainfall series at each station, spanning the period from 1943 to 2023. This results in a time series of up to 81 annual maximum values per location.

Nonetheless, some stations have missing data for certain years, leading to incomplete records. Although the precise effect of these missing observations on the accuracy of the statistical modeling is not fully understood, their presence underscores a key limitation and points to a valuable direction for future investigation.

[Table 9.2](#) gives the names of the 18 meteorological stations.

The data must be introduced in a numeric matrix, with annual extreme values in each site in columns and another matrix with the coordinates of each site per row.

It is necessary to check if the GEV fits our data. The `evd` and `ismev` packages with functions `gev.fit()` for each locale followed by `gev.diag()` allow us to perform that study. For our data, it is shown to be a reasonable fit; as an example, see [Figure 9.2](#) for two sites.

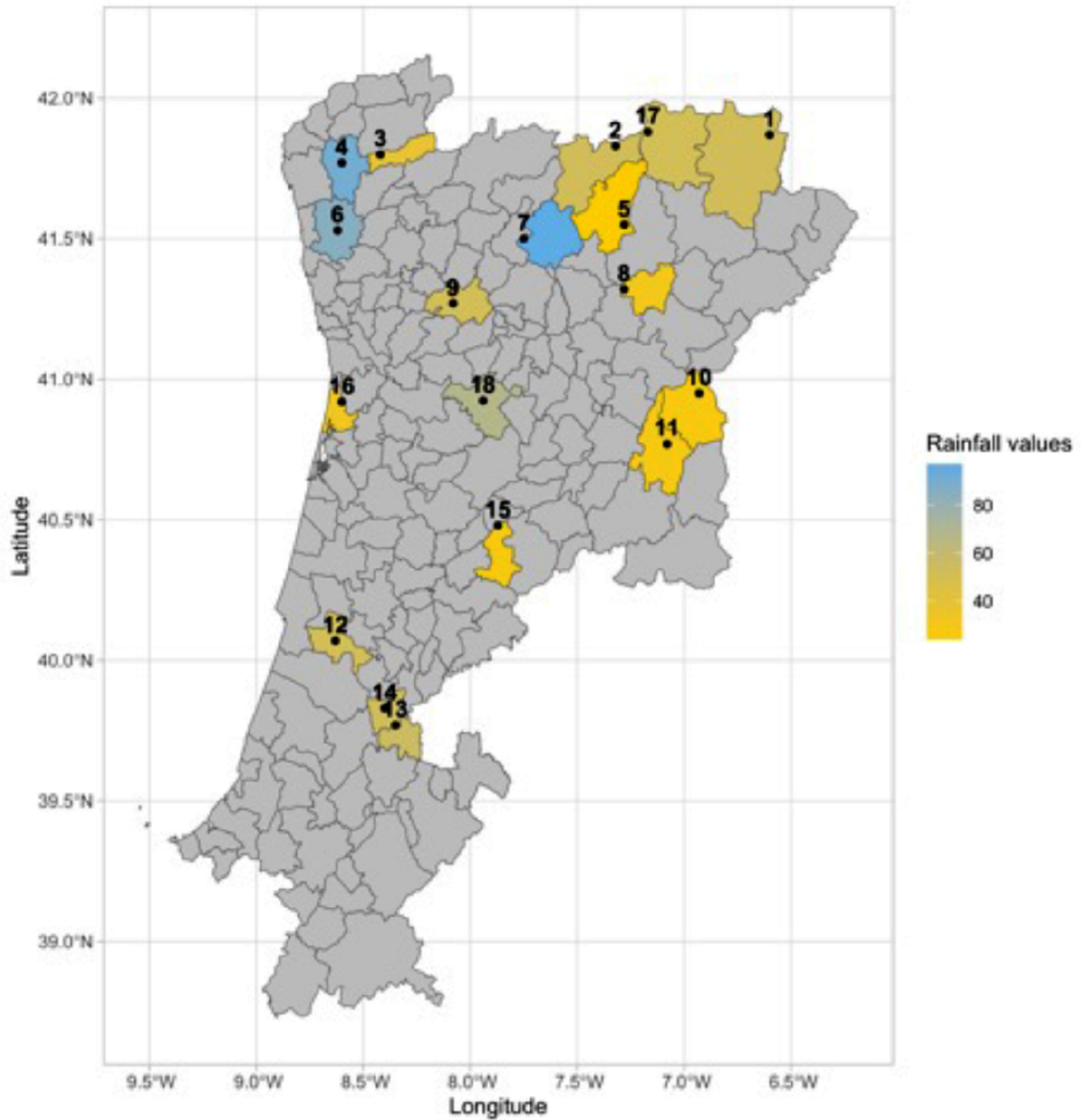


Figure 9.1. Annual maxima of daily rainfall values (mm) across municipalities in northern Portugal for the year 1999 (shown for illustrative purposes), with dots representing the rainfall stations analyzed in this study, as listed in [Table 9.2](#).

Table 9.2. *Identification of the 18 stations plotted in [Figure 9.1](#)*

Number	Name	Number	Name
1	Deilão	10	Escalhão
2	Travancas	11	Pinhel
3	Ponte da Barca	12	Soure
4	Ponte de Lima	13	Rego da Murta
5	Rio Torto	14	Alvaiázere
6	Barcelos	15	Caldas de Felgueiras
7	Sta Marta Montanha	16	Espargo
8	Folgares	17	Gestosa
9	Amarante	18	Castro Daire

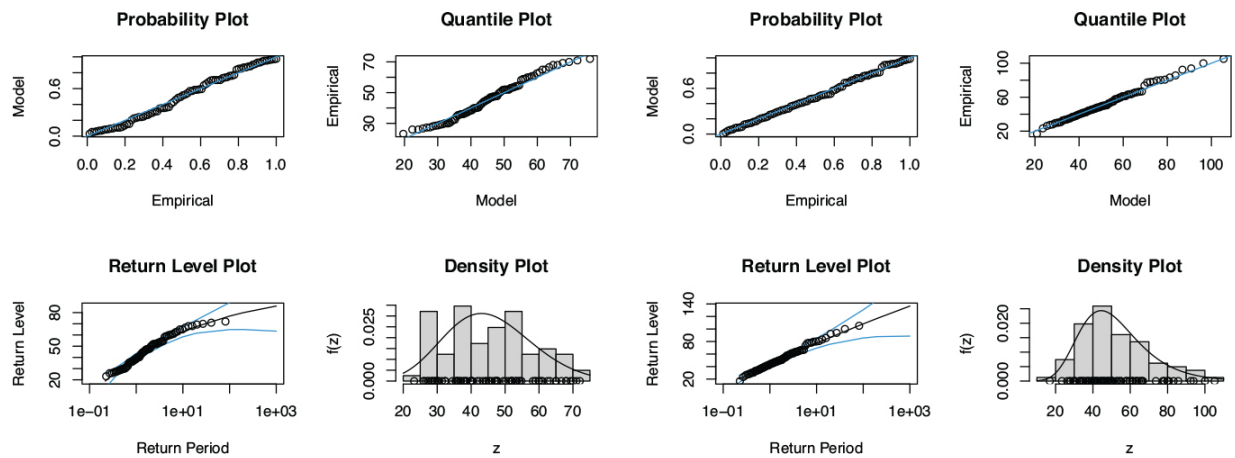


Figure 9.2. *Diagnostic plots for Deilão (left) and Rego da Murta (right) sites.*

The GEV distribution was fitted to the annual maximum values at each station, yielding station-specific estimates of the location (μ), shape (ξ) and scale (σ) parameters, which are plotted in [Figure 9.3](#).

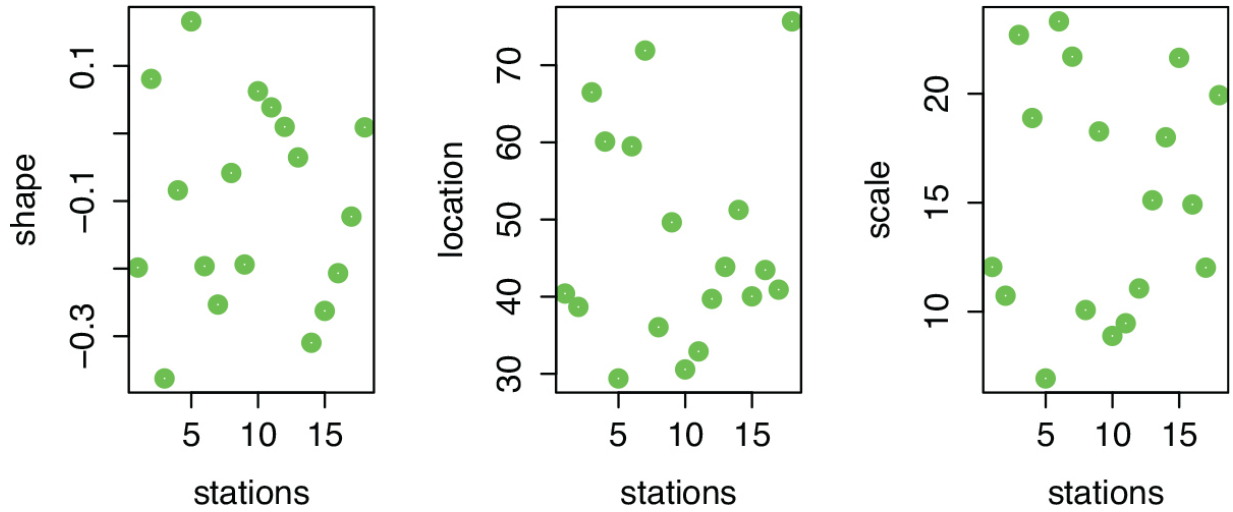


Figure 9.3. *Shape, location and scale parameters plotted for all the stations observed*

Shape, location and scale parameters were also plotted in [Figure 9.4](#) against longitude and latitude for each station. [Figure 9.4](#) reveals a non-homogeneous pattern relative to each station, possibly indicating the presence of a relationship between longitude and the parameters of location (μ), shape (ξ) and scale (σ).

These exploratory plots suggested investigating the optimal trend surface for the GEV location (μ), shape (ξ) and scale (σ), parameters as functions of latitude and longitude. For each station x , several trend surfaces were explored. One model that fits our data well is:

$$\hat{\mu}(x) = -44.58 - 11.69 \ln(x) \quad [9.13]$$

$$\hat{\sigma}(x) = -25.472 - 5.197 \ln(x) \quad [9.14]$$

$$\hat{\xi}(x) = 1.186 + 0.162 \ln(x) \quad [9.15]$$

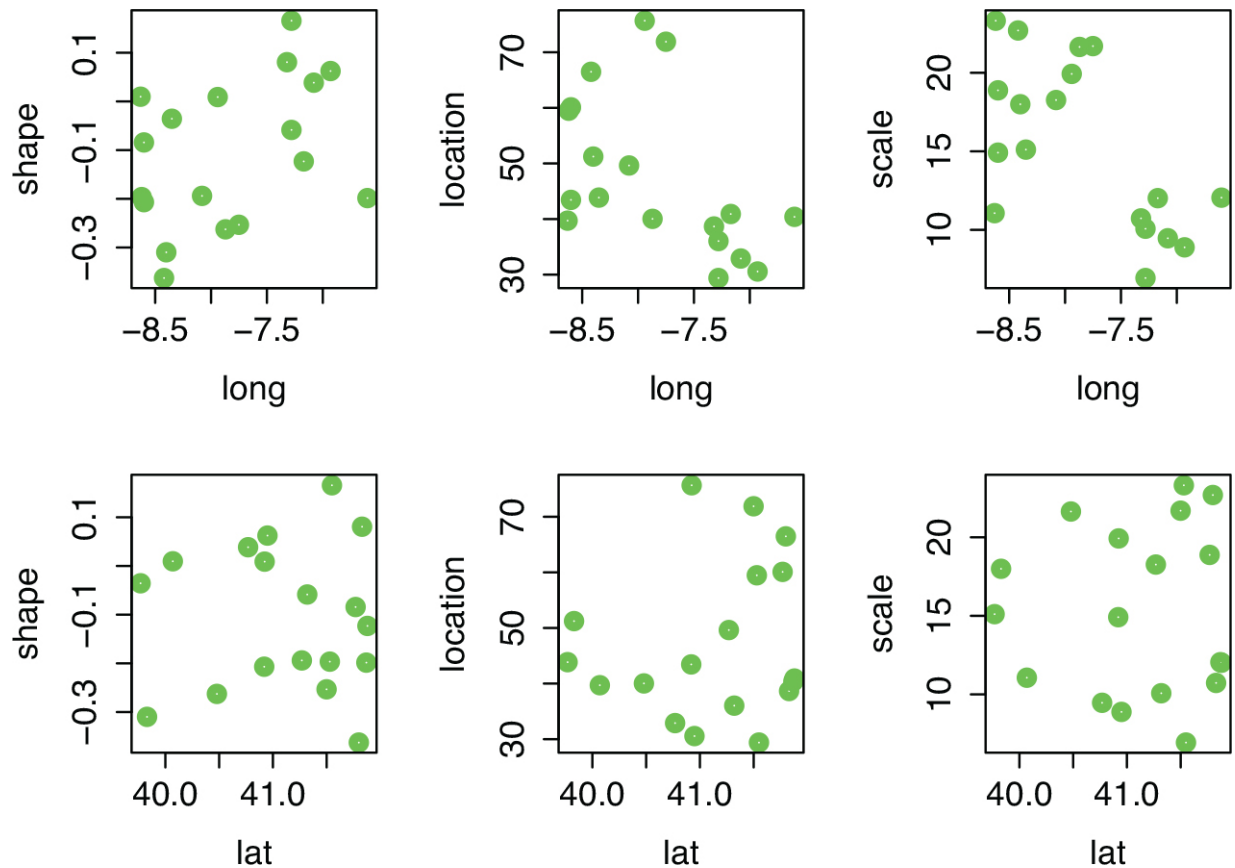


Figure 9.4. *Shape, location and scale parameters plotted against longitude and latitude for each station*

To analyze the dependence structure and fit classical max-stable models, the marginal distributions at each station were first transformed to the unit Fréchet scale using the function `gev2frech()` from the `SpatialExtremes` package.

The `SpatialExtremes` package provides the most comprehensive collection of functions for exploration and statistical inference with max-stable processes for spatial maxima data in geographical space ($d = 2$). While standard full likelihoods are not tractable even for a moderate number of locations with the common models, pairwise likelihood has become the standard approach for fitting max-stable processes, with implementations in `SpatialExtremes`.

The simulation and estimation of the pairwise extremal coefficient were performed using the following R code:

```
par(mfrow = c(1, 2))
fitextcoeff(data, coord, loess = FALSE)
fitextcoeff(data, coord, estim = "Smith", loess = FALSE)
```

[Figure 9.5](#) presents the pairwise extremal coefficient estimates obtained for the Smith and Schlather models. In line with previous studies on precipitation extremes, these estimates indicate a clear spatial dependence in the maximum rainfall values across stations. Accounting for both this dependence and the spatial trends captured by the estimated surfaces is essential for producing reliable predictions of extreme rainfall.

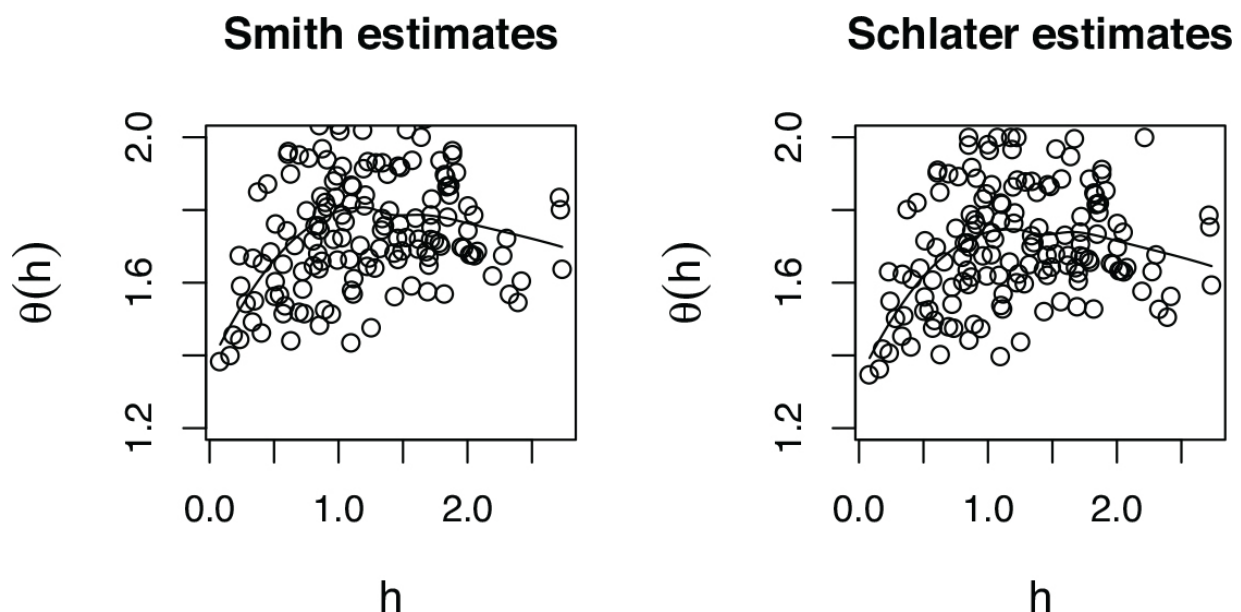


Figure 9.5. *Pairwise extremal coefficient estimates and loess curves from Smith (left) and Schlather (right)*

The `rmaxstab` function allows for the simulation of random fields based on some defined models. [Figure 9.6](#) displays two simulated random fields over a grid. One of the models is the Smith model, for which it is necessary to indicate the estimated covariance matrix Σ . The other model is the Schlather model, for which it is necessary to indicate the nugget, range and smooth estimated values. The code used to produce these figures, for our data, is shown below.

```

x3 <- seq(-10,5, 1)
y3 <- seq(38,42,0.26)
coord3 <- cbind(long =x3, lat = y3)

##1. Simulate a random field from the Smith model

data3 <- rmaxstab(1,coord3,"gauss", cov11 =..., cov12 =...,
  cov22 =...,grid = TRUE)

##1. Simulate a random field from the Schlather model
data4 <- rmaxstab(1,coord3, "powexp", nugget=...,range=...,
smooth=..., grid = TRUE)

par(mfrow=c(1,2))
long=x3;lat=y3
image(long, lat, log(data3[,]), col = terrain.colors(32))
image(long, lat, log(data4[,]), col = terrain.colors(32))

```

In both examples, the argument specifying the covariance model (e.g. "gauss" or "powexp") plays a key role in defining the spatial dependence structure of the simulated field.

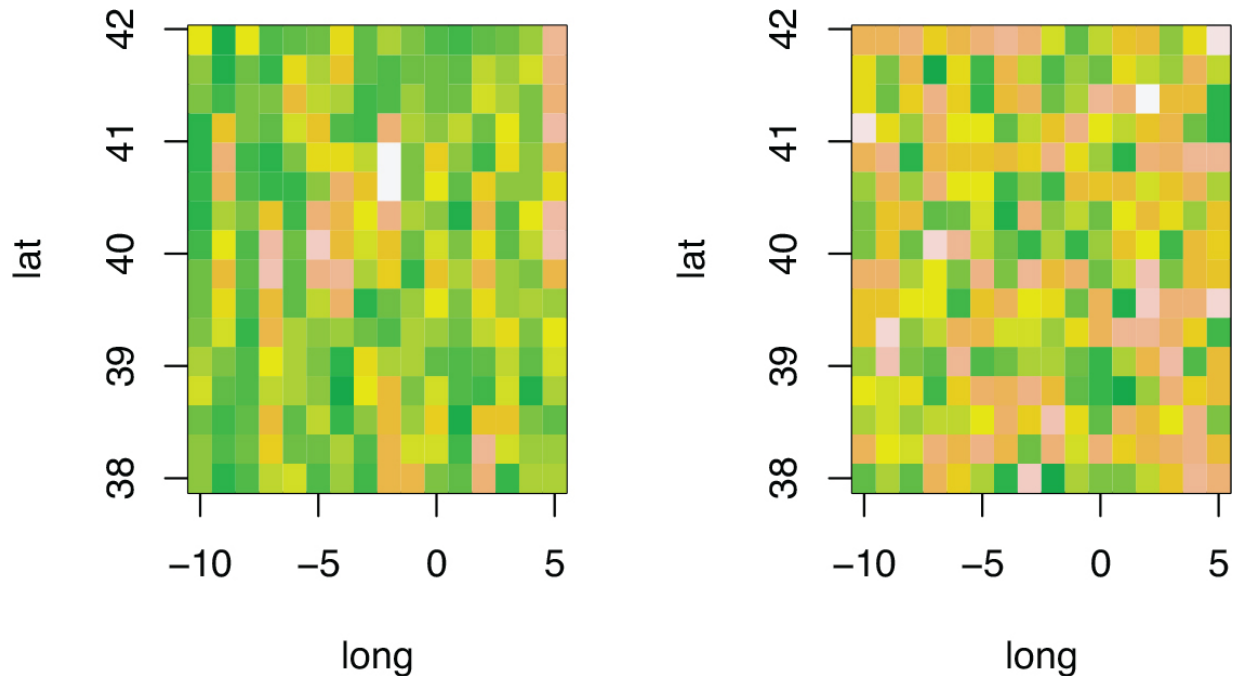


Figure 9.6. *One simulation over a grid from Smith (left) and Schlather (right).*

9.4. Final remarks

In previous studies, see Neves and Prata Gomes ([2011](#)) and Prata Gomes et al. ([2016](#)), preliminary investigations for other locations were presented. However, the current study utilizes a longer series of data collected from 18 stations, thus providing a more robust framework for evaluating spatial dependence in extreme precipitation events. Max-stable processes provide a natural extension of extremal dependence structures and serve as a valuable framework for modeling multivariate extremes. However, the complexity of the multivariate density function often limits practical inference to simple scenarios, such as the bivariate case. In more general settings, substantial approximations and high computational demands are typically required. Despite these challenges, the results obtained highlight the potential of max-stable processes to capture the spatial characteristics of extreme rainfall, offering valuable insights for hydrological planning and climate adaptation strategies. Overall, this study emphasizes the relevance of integrating rigorous statistical theory with applied case studies, showing how max-stable models can be effectively implemented with real hydrological data.

The findings reinforce the importance of accounting for spatial dependence when estimating risk measures associated with extreme rainfall, which is essential for designing resilient infrastructure and supporting decision-making under climate uncertainty. Future research may extend this work by exploring alternative dependence structures, incorporating covariate effects or adopting Bayesian and hierarchical approaches that can better accommodate missing data and improve computational efficiency.

9.5. Acknowledgments

The authors thank the editors for their constructive comments and valuable suggestions, which led to substantial improvements to this work.

This work is funded by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the scope of the projects UID/00297/2025 (<https://doi.org/10.54499/UID/00297/2025>) and UID/PRR/00297/2025 (<https://doi.org/10.54499/UID/PRR/00297/2025>), within the Center for Mathematics and Applications.

9.6. References

- Barndorff-Nielsen, O.E., Gupta, V.K., Pérez-Abreu, V., Waymire, E. (1998). *Stochastic Methods in Hydrology*. World Scientific, Singapore.
- Brown, B.M. and Resnick, S.I. (1977). Extreme values of independent stochastic processes. *Journal of Applied Probability*, 14, 732–739.
- Coles, S.G. (1993). Regional modelling of extreme storms via max-stable processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, 55, 797–816.
- Coles, S.G. (2001). *An Introduction to Statistical Modelling of Extreme Values*. Springer, London.
- Coles, S.G. and Tawn, J.A. (1996). Modelling extremes of the areal rainfall process. *Journal of the Royal Statistical Society: Series B*

(*Methodological*), 58(2), 329–347.

Coles, S.G. and Walshaw, D. (1994). Directional modelling of extreme wind speeds. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 43(1), 139–158.

Cooley, D., Naveau, P., Poncet, P. (2006). Variograms for spatial max-stable random fields. In *Dependence in Probability and Statistics*, Bertail, P., Soulier, P., Doukhan, P. (eds). Springer, New York.

Cressie, N.A.C. (1993). *Statistics for Spatial Data*. Wiley, New York.

Davison, A.C. and Gholamrezaee, M.M. (2012). Geostatistics of extremes. *Proceedings of the Royal Society of London: Series A*, 468, 581–608.

de Haan, L. (1984). A spectral representation for max-stable processes. *The Annals of Probability*, 12, 1194–1204.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer, New York.

de Haan, L. and Pickands, J. (1986). Stationary min-stable stochastic processes. *Probability Theory and Related Fields*, 74, 477–492.

Jenkinson, A.F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 87, 158–171.

Kabluchko, Z., Schlather, M., de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*, 37, 2042–2065.

von Mises, R. (1954). La distribution de la plus grande de n valeurs. In *Selected Papers, Volume II*. American Mathematical Society, Providence, RI, 271–294.

Naveau, P., Guillou, A., Cooley, D., Diebolt, J. (2009). Modeling pairwise dependence of maxima in space. *Biometrika*, 96, 1–17.

- Neves, M.M. and Prata Gomes, D. (2011). Geostatistics for spatial extremes. A case study of maximum annual rainfall in Portugal. *Proc. Environ. Sci.*, 7, 246–251.
- Opitz, T. (2013). Extremal t processes: Elliptical domain of attraction and a spectral representation. *Journal of Multivariate Analysis*, 122, 409–413.
- Prata Gomes, D., Neves, M.M., Moreira, E. (2016). An exploratory study of spatial annual maximum of monthly precipitation in the northern region of Portugal. *Phys. Chem. Earth Parts A/B/C*, 94, 77–84.
- Resnick, S.I. (2022). *Extreme Values, Point Processes and Regular Variation*. Springer, New York.
- Ribatet, M. (2022). SpatialExtremes: Modelling spatial extremes. R package version 2.1-0.
- Ripley, B.D. (2004). *Spatial Statistics*. John Wiley & Sons, Hoboken, NJ.
- Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes*, 5, 33–44.
- Smith, R.L. (1990). Max-stable processes and spatial extremes. Unpublished manuscript.

Note

Chapter written by Dora PRATA GOMES and Maria Manuela NEVES.

10

Time Series and Cluster Analysis of Groundwater Level in Slovakia

In this chapter, we deal with time series and cluster analysis of the groundwater level on Great Rye Island. Great Rye Island is a large river island located in southwestern Slovakia and the largest in Europe. Groundwater level on the island is monitored daily by probes and recorded in the form of a time series. We examined the form of the trend seasonally also residual components as part of our analysis. Within the deterministic component of the time series, our analysis focused on identifying the formula of the trend, the number of significant periods and the duration of each period in the seasonal component. Only linear $ARMA(p, q)$ models were used to model the residual component. These time series' monitored statistical properties served as input for the cluster analysis. The goal of the cluster analysis was to create groups of probes with similar monitored properties and to compare whether the groups are related to the location of the probes in the given area. All results were obtained by analysis in statistical software R and Wolfram Mathematica (R Core Team [2025](#); Wolfram Research, Inc. [2024](#)).

10.1. Introduction

Groundwater represents a crucial and frequently the most economical source of drinking water globally. Its relative ease of capture, utilization, quality assurance and protection often makes it preferable to surface water. In Slovakia, a country with a well-established history of groundwater-centric water management, it reigns as the primary source for drinking water. Fluctuations in groundwater levels directly determine the available quantities of this water, and evaluating these quantities is impossible without long-term groundwater monitoring data.

Groundwater has been analyzed extensively on a global scale, and Slovakia is no exception. Trend analysis of groundwater levels in India has been conducted by Halder et al. ([2020](#)) and Thakur and Thomas ([2011](#)). A slight increasing trend in groundwater levels in western Jilin Province, China, was identified by Lu et al. ([2014](#)). Zuansi and Ofterdinger ([2016](#)) examined how groundwater levels respond to rainfall in Ireland.

Jasechko et al. ([2024](#)) analyzed the global trend of declining groundwater levels, which is most pronounced in dry agricultural areas and is accelerating in many regions. At the same time, it highlights successful examples where changes in policies and water management measures have led to the recovery of underground resources.

Ballová et al. ([2024](#)) studied global groundwater trends using change-point detection first, a key method for identifying sudden data changes and ensuring accurate trend analysis. Furthermore, this methodology is applied to analyze the groundwater resources in Slovakia.

10.2. Data specifications and methods

Rye Island represents the largest river island in Europe, situated in the south-western part of the Slovak Republic, and its geographical boundaries are defined by the dominant flows of the Danube River to the south and the Little Danube to the north. From a geomorphological perspective, it is an extensive alluvial plain with highly fertile soils, determined by the long-term sedimentation processes of the river systems. Specifically, the study area is located in the lower Rye Island, bordered by the Váh and Small Danube rivers and several drainage channels. The complex hydrological regime of the area is attributed to the presence of two major rivers and a comprehensive network of smaller watercourses and canals. The extensive quaternary hydrogeological structure contains a significant reserve of high-quality drinking water of Central European importance, crucial for regional water supplies. Despite intensive agricultural land use, Rye Island supports locally

significant biotopes, particularly in riparian areas and less cultivated lands, which provide habitat for diverse plant and animal species.

Groundwater dynamics within this environment are monitored by the state hydrological network for groundwater quantity monitoring, also known as the groundwater observation network, operated by the Slovak Hydrometeorological Institute (see [Figure 10.1](#)). This network, one of the most extensive environmental monitoring systems within the institution by the number of observation points, primarily aims to continuously or with a daily to weekly periodicity track changes in the groundwater level regime and temperature, facilitating the provision of assessments, expert evaluations and scientific studies. Groundwater levels are measured at designated observation points, commonly termed as probes.

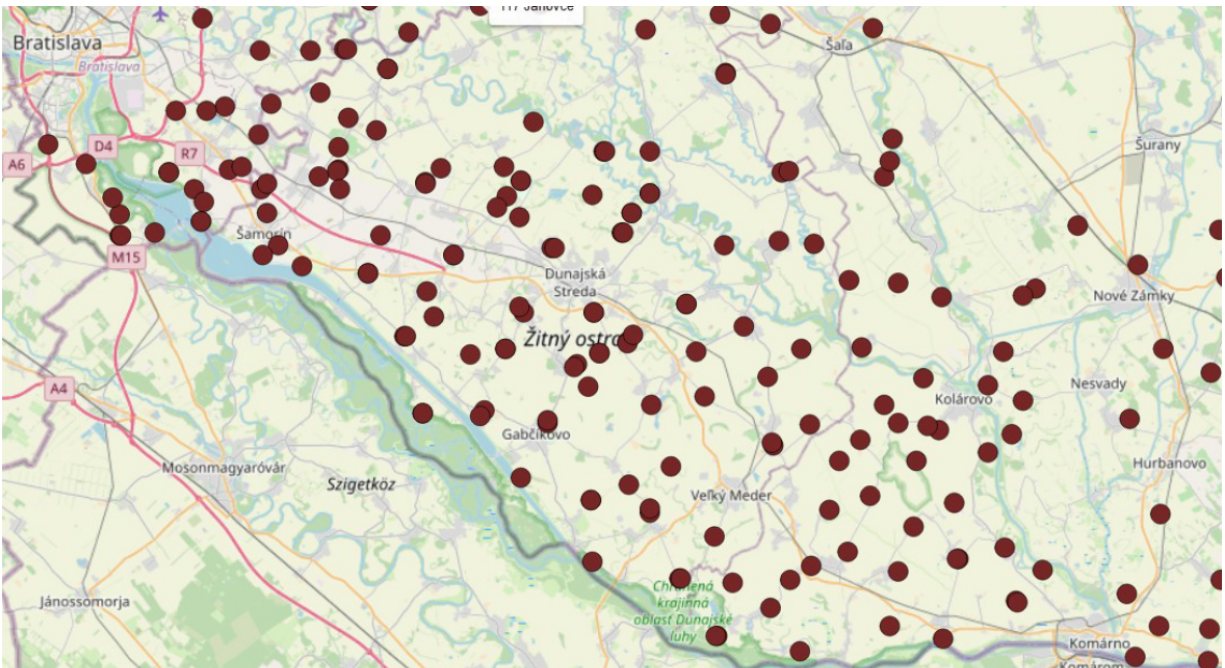


Figure 10.1. Part of hydrological network for groundwater quantity monitoring (Slovak Hydrometeorological Institute: Catalogue of Groundwater Resources of Slovakia).

Groundwater level dynamics were analyzed using the time series dataset $\{x_t\}_{t=1}^n$ where each x_t represents the monthly mean groundwater level derived from the aggregation of daily measurements. This dataset, comprising monthly averages from 21 monitoring wells, covers the temporal interval from January 2007 to

October 2014 (see [Figures 10.2](#) and [10.3](#)). To decompose the temporal structure of the measurements, the initial phase involved the application of time series analysis techniques to identify and quantify the trend and seasonal components. Subsequently, the residual series, obtained after the removal of trend and seasonality, was subjected to statistical modeling. Finally, cluster analysis was performed on the feature space defined by the estimated trend, seasonal parameters and residual characteristics to group the observation points.

10.3. Analysis of trend and periodic components in groundwater levels

Time series analysis is a statistical technique used to analyze data points collected over a specific period and ordered chronologically (Box and Jenkins [1970](#); Shumway and Stoffer [2024](#)). The primary goal is to identify recurring patterns, trends, seasonal variations and irregularities within the data. By understanding these components, we can gain insights into the underlying processes that generate the time series and make informed forecasts about future values. This chapter focuses on the application of time series additive decomposition methods to groundwater level data with the aim of accurately identifying and characterizing their trend and periodical (seasonal and cyclical) components.

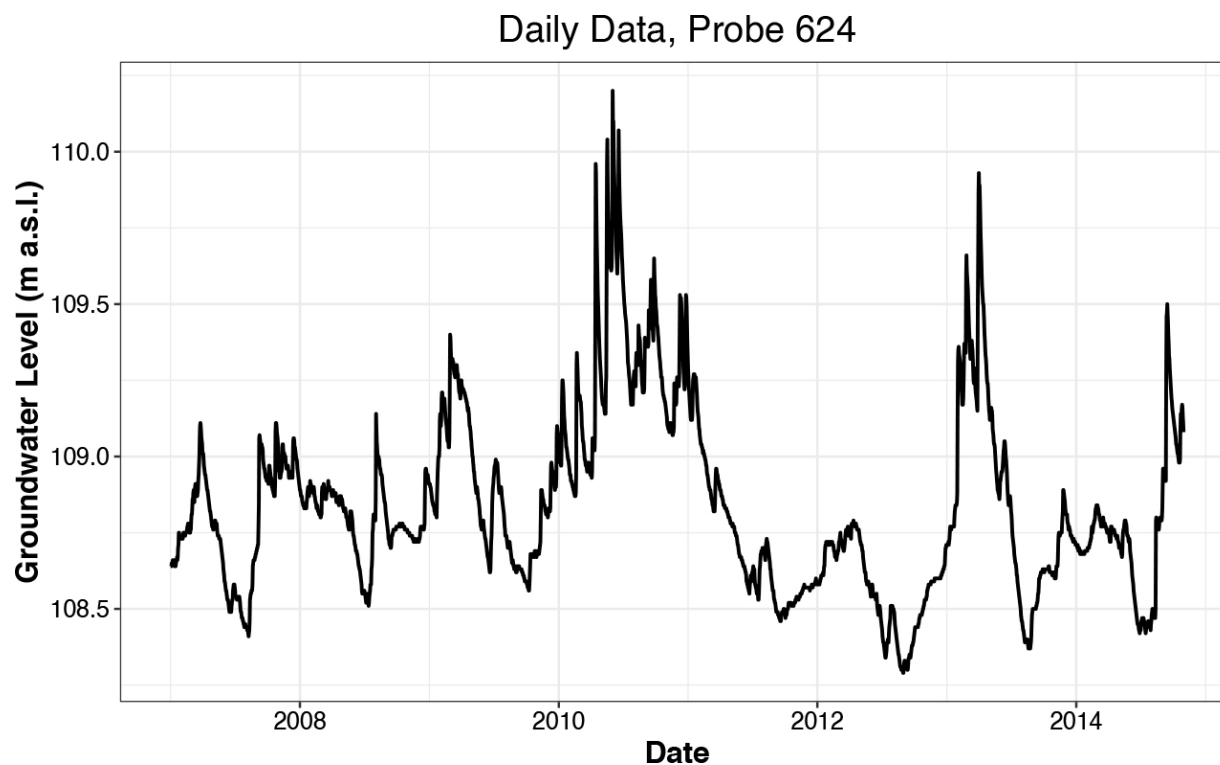


Figure 10.2. *Daily data, probe 624*

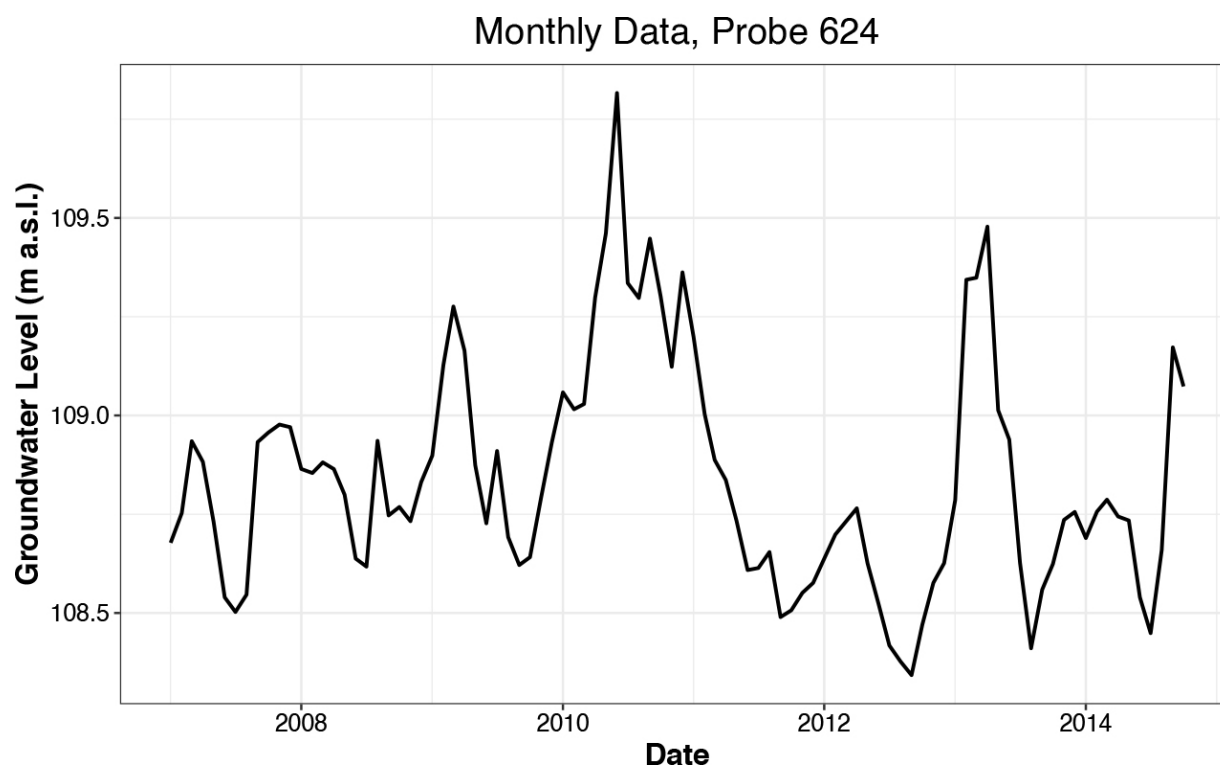


Figure 10.3. *Monthly data, probe 624*

Trend T_t describes the long-term development in the average behavior of the time series. It arises as a consequence of forces that systematically act in the same direction. It thus expresses a long-term increase (or decrease). The trend in time series can be described using trend functions. This expression is used if the development of the time series corresponds to a certain function of time, such as linear, quadratic, etc.

The long-term trend in the time series of monthly groundwater level measurements was analyzed using polynomial regression, with the maximum polynomial degree limited to three.

$$T_t(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$$

Seasonal Groundwater Level Variation (Probe 624)

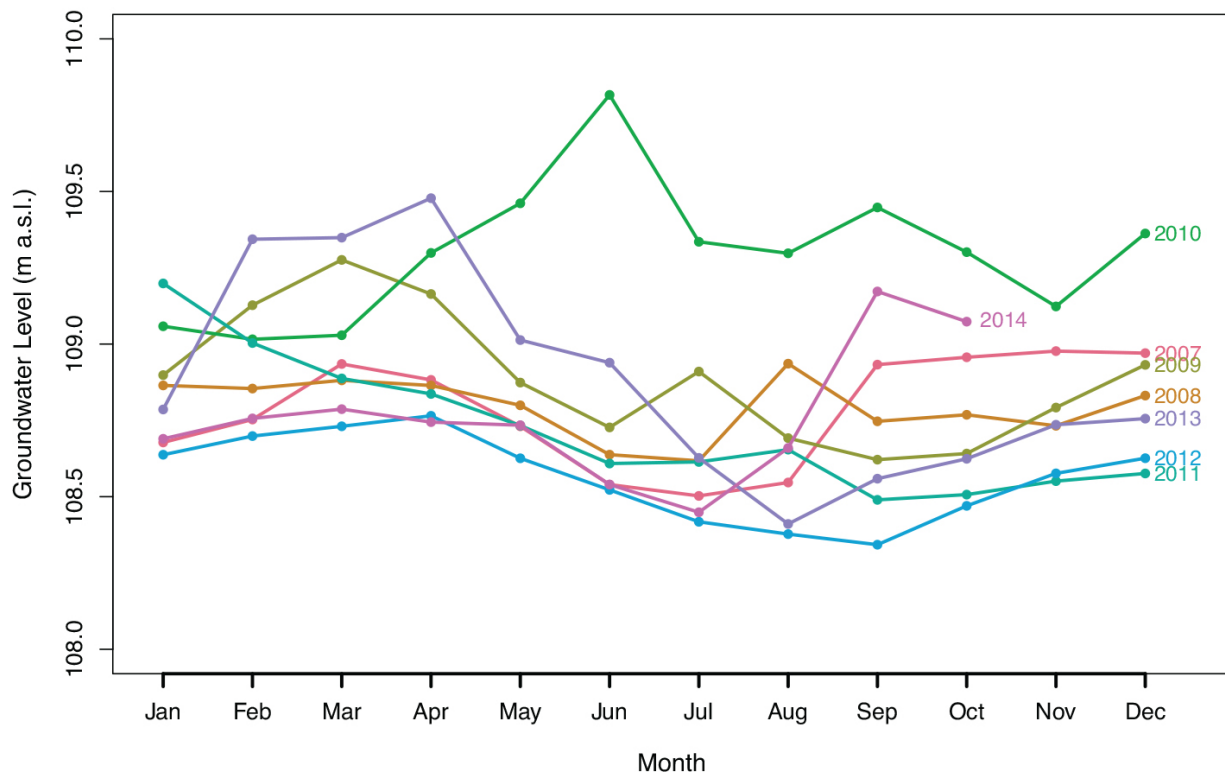


Figure 10.4. Seasonal subseries plot for probe 624.

The coefficients β_i , $i = 1, 2, 3$ for each model were estimated using the *OLS* method, which minimizes the sum of squared differences between observed x_t and modeled values $T_t(t)$, $t = 1, 2, \dots, n$. The goodness of fit of the individual polynomial models can be evaluated

using metrics such as the residual sum of squares RSS , information criterion, coefficient of determination R^2 or adjusted coefficient of determination R^2_{adj} . For selecting the most suitable model, we opted for an information criterion BIC due to its stronger theoretical foundation and its ability to better balance model fit and complexity compared to R^2_{adj} . The BIC information criterion is calculated as:

$$BIC = k \cdot \ln(n) - 2\ln(\mathcal{L})$$

where k represents the number of parameters in the model, n is the number of measurements and $\mathcal{L}$ is the maximum value of the log-likelihood of the given model.

Seasonal component S_t describes periodic changes in the time series, whose period L is equal to a certain standard unit of time or its constant multiple. The seasonal component is also characterized as a correlation dependence between the i th and $(i + k)$ th element of the time series, where k is the shift. The period L is determined from the correlogram, specifically at values (and their multiples) where the correlogram reaches local maxima. The seasonal component can be described by dummy variables or a harmonic function.

The cyclical component C_t is a periodic component whose period does not correspond to calendar units. It represents irregular fluctuations around the trend, alternating between growth and decline phases. This component is expressed as the sum of sines and cosines with different frequencies and amplitudes. The frequency ω is given in radians per unit of time and the length of the period L , or the time for one cycle, is given by $L = \frac{2\pi}{\omega}$. Significant frequencies are determined using spectral analysis, typically from the sample spectral density graph (periodogram), and confirmed using Fisher's test for periodicity (Fisher [1929](#); Anděl [1976](#); Siegel [1980](#)). The null and alternative hypotheses are as follows:

- H_0 : The time series is only white noise
 $X_t = \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2)$;

$$- H_1 : \\ X_t = \mu + \sum_{j=1}^p (\alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)) + \varepsilon_t, t = 1, 2, \dots, n.$$

where μ , α_j and β_j are unknown parameters and ω_j are significant frequencies, $\varepsilon_t \sim N(0, \sigma^2)$. Test statistics are expressed in the form of:

$$W = \max Y_j, j = 1, 2, \dots, m < \frac{n-1}{2}$$

where

$$Y_j = \frac{I(\omega_j^*)}{\sum_{i=1}^m I(\omega_i^*)}, j = 1, 2, \dots, m$$

and $I(\omega_j^*)$ are periodogram values calculated from the dataset $\{x_t\}_{t=1}^n$ for the frequencies $\omega_j^* = \frac{2\pi j}{n}, j = 1, 2, \dots, m$. The null hypothesis is rejected, if $W \geq g_F$, where g_F is critical value tabulated in Fisher ([1929](#), [1940](#)).

Some results of trend and periodic component analysis for probes are in [Table 10.1](#).

10.4. The residuals component and ARMA modeling

The residual component R_t represents the unexplained variability in a time series after the systematic elements have been extracted. It captures the irregular, unsystematic fluctuations that do not exhibit a discernible pattern or periodicity. often being regarded as the noise or random error within the observed data. While the primary objective of time series decomposition is to isolate the residual component as unexplained variability or noise, this component

frequently exhibits a discernible structure amenable to further modeling. For this purpose, linear stationary processes, specifically autoregressive moving average models of order p and q ($ARMA(p, q)$) are particularly effective. Their systematic application is a cornerstone of the widely recognized Box–Jenkins methodology (Box and Jenkins [1970](#)). This methodology emphasizes the critical analytical role of the residual component, which, even after the removal of systematic time series elements, often comprises correlated or dependent random variables.

Table 10.1. *Trend and periodic components for some probes*

Probe	Trend	Number of significant periods	Most significant period (months)
624	Constant	4	42
626	Quadratic	6	12
2607	Quadratic	5	42
2611	Cubic	6	42
2628	Constant	4	42

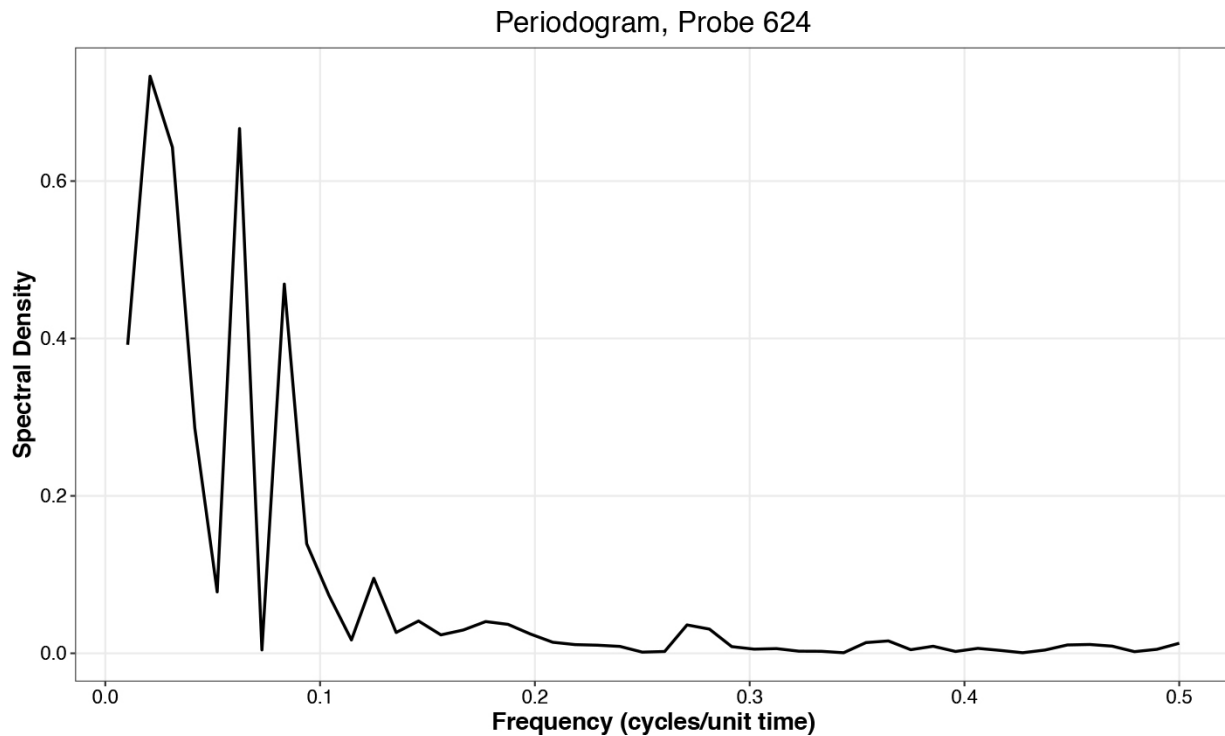


Figure 10.5. *Periodogram for probe 624*

The $ARMA(p, q)$ model structure is expressed in the form of:

$$R_t = \phi_1 R_{t-1} + \phi_2 R_{t-2} + \dots + \phi_p R_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_p \epsilon_{t-p}$$

where $\phi_i, i = 1, 2, \dots, p, \theta_j, j = 1, 2, \dots, q$ are parameters of $ARMA(p, q)$ and ϵ_t is white noise.

Model identification, the foundational step of the Box–Jenkins methodology, involved initially determining the optimal orders p and q for the $ARMA(p, q)$ model that best captured the time series' residual characteristics. The BIC criterion then guided the selection of the most parsimonious model. Following this, the $ARMA(p, q)$ model parameters were estimated using the MLE method in the second procedural step. Finally, diagnostic checking was performed by subjecting the model's residuals z_t to the Ljung–Box test to confirm its statistical adequacy (Ljung and Box [1978](#)).

The null and alternative hypotheses are as follows:

- H_0 : residuals z_t are independently distributed.
- H_1 : residuals z_t exhibit serial correlation.

The test statistics are in the form of:

$$Q = n(n + 2) \sum_{k=1}^h \frac{r_k^2}{n - k}$$

where n is the sample size, r_k is the sample autocorrelation at lag k and h is the number of lags being tested. The null hypothesis is rejected (for significance level α), if $Q \geq \chi_h^2(1 - \alpha)$, where

$Q \geq \chi_h^2(1 - \alpha)$ is $(1 - \alpha)$ quantile of the χ^2 distribution with h degrees of freedom. Some results of Box–Jenkins analysis for probes are presented in [Table 10.2](#).

Table 10.2. *Linear models for some probes*

Probe	Linear model
624	$ARMA(1, 1)$
626	$ARMA(1, 3)$
2607	$ARMA(1, 1)$
2611	$ARMA(1, 3)$
2628	$ARMA(1, 2)$

10.5. Probe cluster analysis and visualization

This chapter details the application of cluster analysis to probes, utilizing a feature space comprising statistical properties extracted from their time series, probe elevation, and borehole depth. The core aim was twofold: to identify distinct groupings of probes based on their shared monitored characteristics and to examine the relationship between these clusters and the probes' geographical distribution within the study area. The complete linkage method was employed for clustering, which defines cluster dissimilarity by the maximum Euclidean distance between any two points from respective clusters. This method favors the formation of compact, spherical clusters (Ward [1963](#); Everitt et al. [2011](#)). The resulting hierarchical structure was then visualized using a dendrogram (see [Figure 10.6](#)).

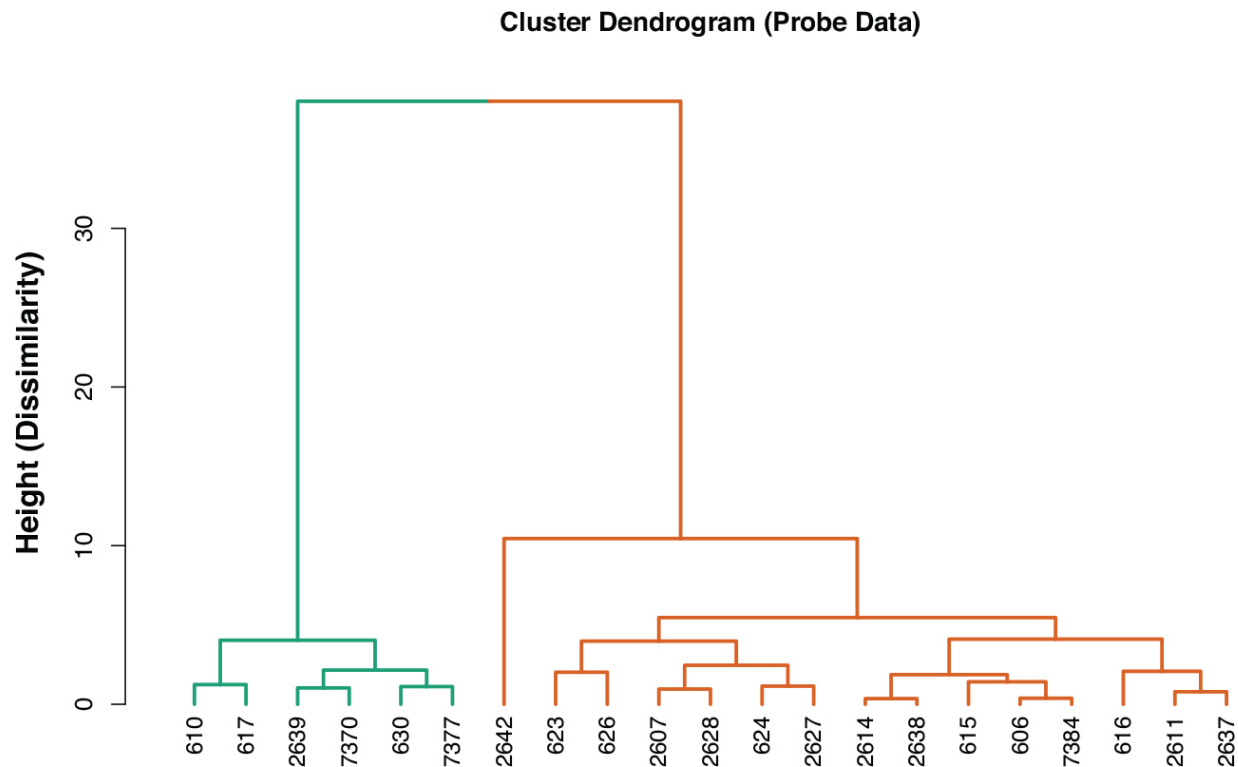


Figure 10.6. *Dendrogram of probes.*

The optimal number of clusters (2) was determined using the silhouette method, which assesses clustering quality by evaluating how well each object is matched to its own cluster versus a neighboring cluster (Rousseeuw [1987](#); Maechler et al. [2025](#)) (see [Figure 10.7](#)).

Cluster 1 probes display distinct temporal and spatial characteristics. Temporally, they are defined by a constant trend and exhibit five to six significant periods, with the most prominent period consistently having a length of 12 months. Spatially, these probes are located on the periphery of the band formed by cluster 2 probes.

In contrast, cluster 2 probes exhibit distinct characteristics. Temporally, cluster 2 probes are characterized by any of the considered trends, indicating a broader range of underlying behaviors compared to the consistent trend observed in cluster 1. They show four, five or six significant periods, with the crucial distinction that none of their significant periods have a length of 12 months, directly differentiating them from cluster 1's prominent 12-month periodicity. Cluster 2 probes form a well-defined linear band

within the study area, as visually confirmed by the accompanying geographical distribution plot ([Figure 10.8](#)). This band can be approximated by a straight line, representing a primary spatial feature. In contrast, cluster 1 probes are consistently positioned on the periphery of this cluster 2 band. This spatial pattern suggests a central organizational role for cluster 2, around which the probes of cluster 1 are distributed.

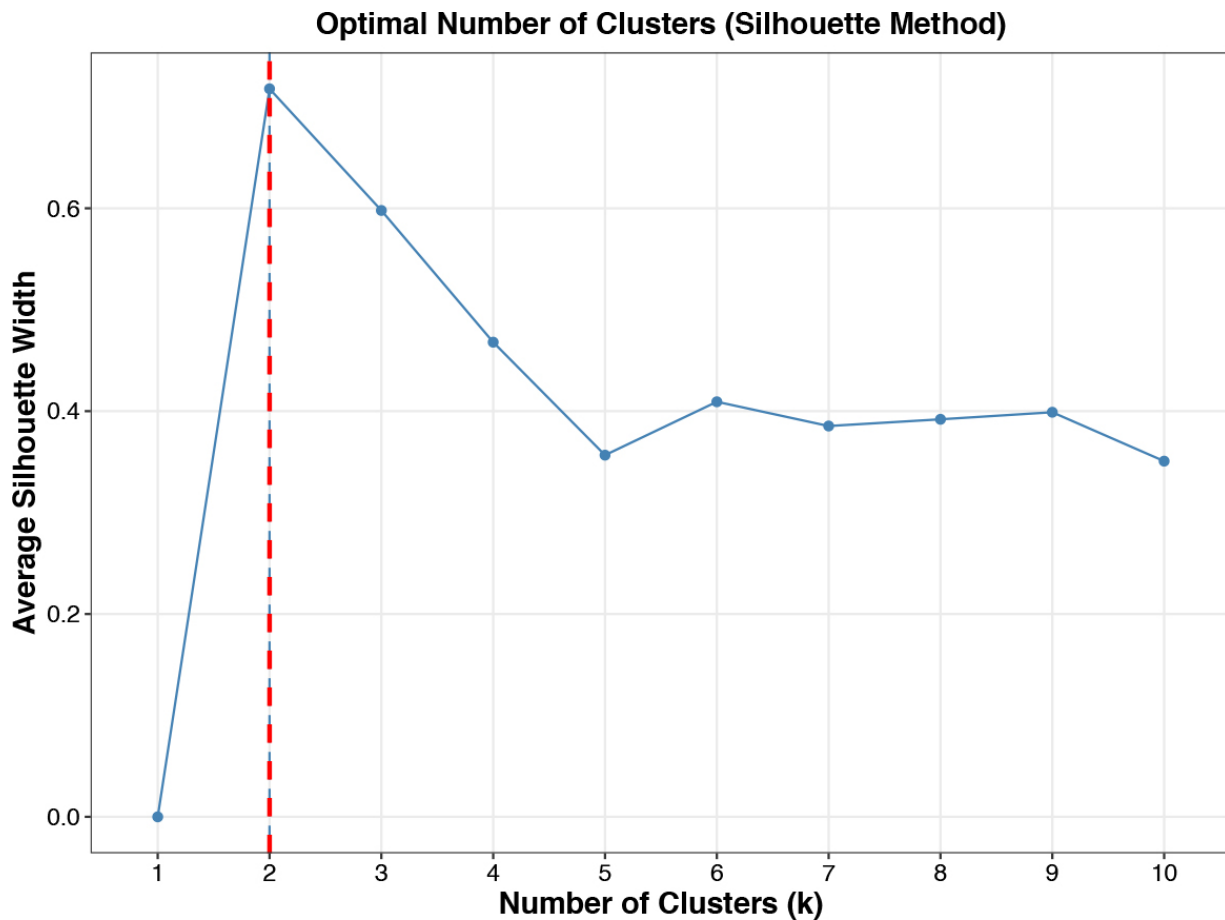


Figure 10.7. *Optimal number of clusters*

10.6. Conclusion

This study characterized groundwater level probes by analyzing their statistical time series properties and spatial distribution. The primary objective was not to predict future groundwater levels, but rather to gain a deeper, more nuanced understanding of the intrinsic hydrogeological behavior within the specific area of Great Rye

Island. Our analysis distinctly identified that probes exhibiting similar temporal characteristics, such as specific trends and seasonal patterns, also tend to cluster spatially. This correlation unequivocally indicates a direct connection between underlying hydrogeological processes and the physical location of the monitoring points. These findings were instrumental in enhancing our comprehension of the probes' intricate spatial arrangement and functional differentiation across the island. The identified spatiotemporal patterns provide valuable, foundational insights into the complex dynamics of groundwater on this significant European river island, thereby laying a robust groundwork for future, more targeted hydrological investigations and effective water resource management strategies.

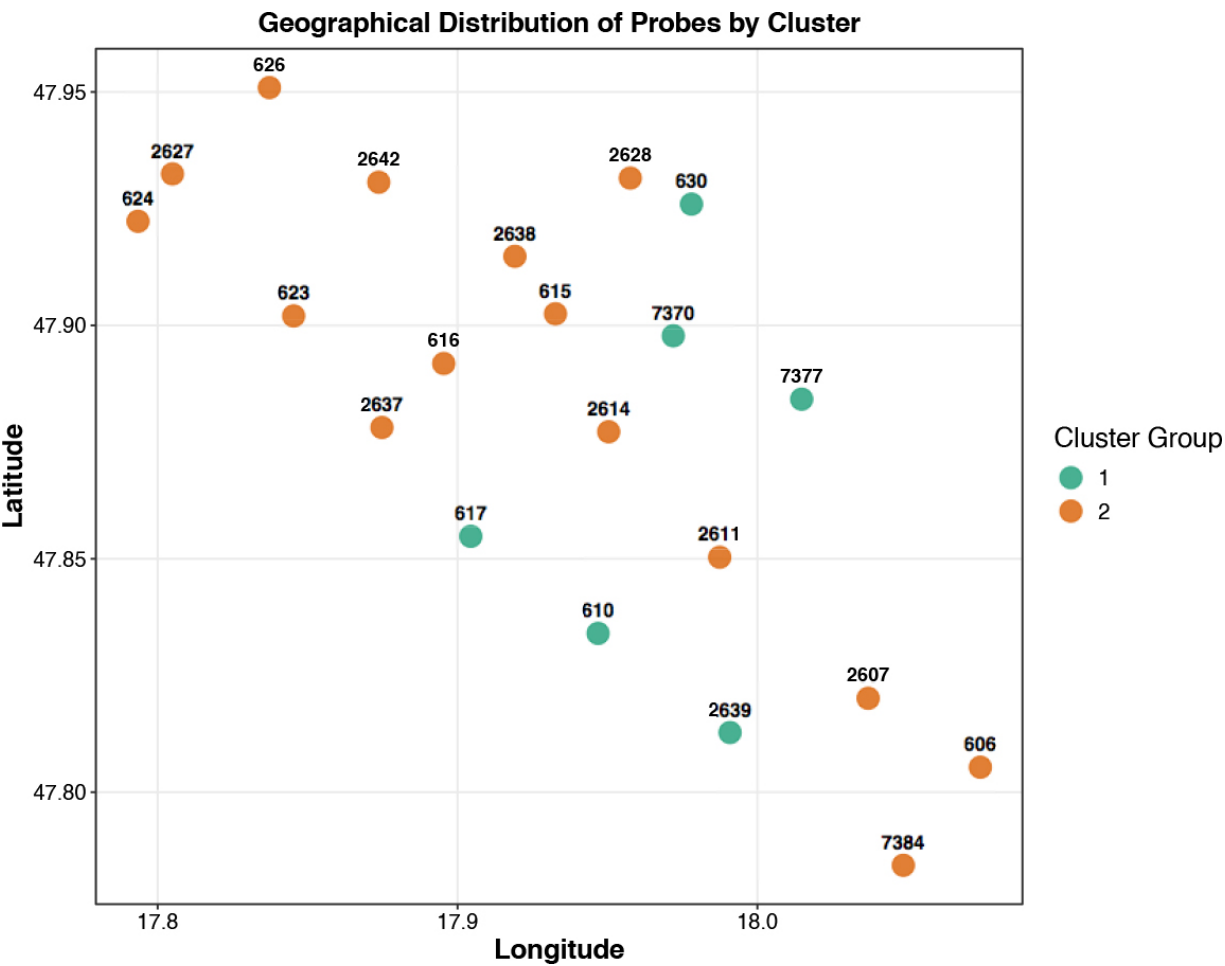


Figure 10.8. *Cluster visualization.*

10.7. Acknowledgment

This work was supported by the grant VEGA-1/0036/23.

10.8. References

- Anděl, J. (1976). *Statistická analýza časových řad*. SNTL, Prague.
- Ballová, D. Sónak, Kalická, J., Červeňanská, M. (2024). Statistical analysis of groundwater level in Slovakia. In *Data Analysis and Related Applications 4*, Dimotikalis, Y. and Skiadas, C.H. (eds). ISTE Ltd, London and Wiley, Hoboken, NJ, 19–28.
- Box, G. and Jenkins, G. (1970). *Time Series Analysis, Forecasting, and Control*. Holden–Day, San Francisco, CA.
- Everitt, B.S. Landau, S., Leese, M., Stahl, D. (2011). *Cluster Analysis*, 5th edition. Wiley-Blackwell, Chichester.
- Fisher, R.A. (1929). Test of significance in harmonic analysis. *Proc. Roy. Soc. London A*, 125, 54–59.
- Fisher, R.A. (1940). On the similarity of the distributions found for the test of significance in harmonic analysis and in Steven’s problem in geometrical probability. *Annals of Eugenics*, 10, 14–17.
- Grolemund, G. and Wickham, H. (2017). *R for Data Science*. O’Reilly Media, Sebastopol, CA.
- Halder, S., Roy, M.B., Roy, P.K. (2020). Analysis of groundwater level trend and groundwater drought using Standard Groundwater Level Index: A case study of an eastern river basin of West Bengal, India. *SM Applied Science*, 2, 507.
- Jasechko, S., Seybold, H., Perrone, D., Fan, Y., Shamsudduha, M., Taylor, R.G., Fallatah, O., Kirchner, J.W. (2024). Rapid groundwater decline and some cases of recovery in aquifers globally. *Nature*, 625, 715–721.
- Kassambara, A. and Mundt, F. (2020). factoextra: Extract and visualize the results of multivariate data analyses. R package

version 1.0.7 [Online]. Available at: <https://CRAN.R-project.org/package=factextra>.

Ljung, G.M. and Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303.

Lu, W.X., Zhao, Y., Chu, H.B., Yang, L.L. (2014). The analysis of groundwater levels influenced by dual factors in western Jilin Province by using time series analysis method. *Applied Water Science*, 4, 251–260.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K. (2025). cluster: Cluster analysis basics and extensions. R package version 2.1.8.1 [Online]. Available at: <https://CRAN.R-project.org/package=cluster>.

R Core Team (2025). *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna [Online]. Available at: <https://www.R-project.org/>.

Rousseeuw, P.J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–56.

Shumway, R.H. and Stoffer, D.S. (2024). *Time Series Analysis and Its Applications. With R Examples*. Springer, Cham.

Siegel, A.F. (1980). Testing for periodicity in a time series. *J. Amer. Statist. Assoc.*, 75, 345–348.

Slovak Hydrometeorological Institute (n.d.). *Catalogue of Groundwater Resources of Slovakia* [Online]. Available at: <https://www.shmu.sk/en/?page=1598>.

Thakur, G. and Thomas, T. (2011). Analysis of groundwater levels for detection of trend in Sagar district, Madhya Pradesh. *Journal of the Geological Society of India*, 77, 303–308.

Ward, Jr., J.H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236–244.

Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.

Wolfram Research, Inc. (2024). Mathematica, Version 14.0. Wolfram Research, Inc., Champaign, IL.

Zuansi, C. and Ofterdinger, U. (2016). Analysis of groundwater-level response to rainfall and estimation of annual recharge in fractured hard rock aquifers, NW Ireland. *Journal of Hydrology*, 2, 535.

Note

Chapter written by Jana KALICKÁ, Dominika Sónak BALLOVÁ and Michalea ČERVENANSKÁ.

11

Coupon Collector Problem with Partial and Total Resets

Coupon collector problem (CCP) has been considered extensively in mathematical literature for several decades and has several recent generalizations. One line of generalizations, with possible application in biology, is obtained by introducing the possibility of a partial or a total reset, which means that some or all coupons collected up to time t are lost (removed from the collection) with a given probability. We analyze both versions of the problem through the Markov chain approach and provide several exact and asymptotic results related to the expected waiting time until the full collection is sampled despite the resets.

11.1. Introduction and motivation

The CCP is an old and well-known combinatorial problem that dates back to the 18th century and the work of Abraham de Moivre, “De mensura sortis” (“On the measurement of chance”). Nevertheless, the CCP is still extensively treated in mathematical literature and appears in various fields of application.

The simplest (classical) version of the CCP can be explained as follows: a collector buys a coupon every day (or at discrete moments) in order to complete a full collection of n different coupons from the set $N_n = \{1, 2, \dots, n\}$, assuming that all coupons are drawn with the same probability and that sampling is done with replacement.

The classical CCP can be generalized or modified in various ways (e.g. by redefining the goal of the collection process, considering multiple collectors or group draws, including coupons with special purposes in the coupon set, etc.), which may be motivated by possible applications.

A specific application in the field of biology, the modeling of parasitism, serves as motivation for this work (Zoroa et al. [2017](#); Singh [2021](#)). Put simply, there is a population of parasitoids and a (finite) population of hosts, and the process of parasitism can be formalized as a sequence of parasitoid attacks on hosts. When a host is attacked by parasitoids, it cannot be attacked again, which is equivalent to drawing a new coupon (see

Zorua et al. (2017)). In this context, attacks are therefore draws, and attacked (parasitized) hosts are drawn coupons.

From this, various quantities relevant to this research area can be calculated, for example, the exact or mean number of hosts parasitized in a given time interval, or the number of hosts that have escaped being parasitization (because they can produce offspring and keep the population alive).

11.2. CCP with total resets

Here, we briefly explain the generalization of CCP introduced in Jocković and Todić (2024a), which is based on the idea of having additional coupons with special purposes in the coupon set. Other generalizations of CCP that follow the same idea are, for example, Anceaume et al. (2015); Anceaume et al. (2016), Jocković and Todić (2024b,c,d), Nakata and Kubo (2006) and Todić (2024).

More precisely, we assume that there is a special coupon (so-called reset coupon) that does not belong to the set $\mathbb{N}_n$ and acts as a reset button, which means that the collection of coupons drawn up to time (day) t becomes empty if the reset coupon is drawn on day $t + 1$. There is therefore an extended set of available coupons:

$$\mathbb{N}_n^{(R)} = \{1, 2, \dots, n, R\} = \mathbb{N}_n \cup \{R\},$$

where R denotes the reset coupon. We assume that R is drawn with probability p_R , $0 \leq p_R < 1$, so each standard coupon (from $\mathbb{N}_n$) is drawn with probability $\frac{1-p_R}{n}$. This version of CCP will be called CCP with total reset and will be denoted as CCPTR in the rest of the text.

Let X_t be the number of distinct types of standard coupons sampled after t days. We can easily conclude that $\{X_t, t = 1, 2, \dots\}$ is a Markov chain on the state space:

$$S = \{0, 1, \dots, n\}, \tag{11.1}$$

and the transition probability matrix:

$$\mathbf{P}^{(TR)} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & \dots & \dots & n-1 & n \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ \vdots \\ n-2 \\ n-1 \\ n \end{matrix} & \begin{pmatrix} p_R & 1-p_R & 0 & 0 & \dots & \dots & 0 & 0 \\ p_R & p & (n-1)p & 0 & \dots & \dots & 0 & 0 \\ \vdots & \vdots & & & & & & \vdots \\ p_R & 0 & 0 & 0 & \dots & (n-2)p & 2p & 0 \\ p_R & 0 & 0 & 0 & \dots & 0 & (n-1)p & p \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix} \end{matrix},$$

When dealing with any version of the CCP, one of the main mathematical objects of interest is the waiting time to complete a collection of n distinct coupons from the set $\mathbb{N}_n$. In the case of CCPTR, this is the waiting time until a complete collection of standard coupons is collected despite the resets, denoted as $W_n^{(TR)}$.

In Jocković and Todić (2024a), the authors study the distribution of the waiting time $W_n^{(TR)}$ using Markov chain techniques and properties of special functions and derive a simple expression for $E(W_n^{(TR)})$ in terms of the beta function. This result is stated in the following theorem.

THEOREM 11.1 (Jocković and Todić (Jocković and Todić 2024a)).– If $0 < p_R < 1$, for the waiting time $W_n^{(TR)}$, the following equality holds:

$$E(W_n^{(TR)}) = \frac{1}{p_R} \left(\frac{1}{np_R B\left(n, \frac{np_R}{1-p_R}\right)} - 1 \right),$$

where $B(\cdot)$ denotes the beta function.

Theorem 11.1 gives a simple way to obtain asymptotic properties of the expected waiting time $E(W_n^{(TR)})$ by using asymptotic properties of the beta function, which is explained in detail in Jocković and Todić (2024a).

The CCPTR problem can be formulated in another way, without using the notion of reset coupon. More precisely, one can say that at a point in time (day) with probability p_R , the collection of coupons drawn up to this

moment (day) becomes empty and no additional coupon is drawn at that moment (day).

In the context of modeling parasitism, this version of the problem corresponds to the situation that at any point in time the entire population of hosts recovers from the parasitoid attacks with probability p_R .

11.3. CCP with partial resets

When considering the CCPTR for modeling parasitism, one can argue that from a biological perspective it is not realistic to assume that the entire host population recovers from parasitism when the reset occurs, but that this happens partially. This motivates us to modify the CCPTR problem to take this possibility into account.

Next, we propose two models that account for the possibility of a partial reset.

11.3.1. Partial reset (1)

In this version of the problem at each point of time j out of i coupons disappear from the collection with probability $d_{i,j}$ and then a new coupon is drawn.

Again, we can conclude that $\{X_t, t \geq 0\}$, with X_t being the number of distinct coupons sampled after t units of time, forms a Markov chain on the state space [11.1], and the transition probability matrix:

$$\mathbf{P}^{(PR1)} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & \dots & n-1 & n \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ n-2 \\ n-1 \\ n \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & p_{11} & p_{12} & 0 & \dots & 0 & 0 \\ 0 & p_{21} & p_{22} & p_{23} & \dots & 0 & 0 \\ \vdots & \vdots & & & & \vdots & \\ 0 & p_{n-2,1} & p_{n-2,2} & p_{n-2,3} & \dots & p_{n-2,n-1} & 0 \\ 0 & p_{n-1,1} & p_{n-1,2} & p_{n-1,3} & \dots & p_{n-1,n-1} & p_{n-1,n} \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \end{matrix},$$

where the transition probabilities are:

$$p_{i,i+1} = \frac{n-i}{n}d_{i,0}, 0 \leq i \leq n-1, \quad [11.2]$$

and

$$p_{i,i-k} = \frac{n-i+k+1}{n}d_{i,k+1} + \frac{i-k}{n}d_{i,k}, \quad [11.3]$$

for $0 \leq i \leq n-1$, $0 \leq k \leq i-1$. The term d_{ij} is the probability that j out of i available coupons disappear from the collection, at any time point. In general, $D_i = \{d_{ij}, j = 0, 1, \dots, i\}$ is an arbitrary probability distribution over the set $\{0, 1, \dots, i\}$. REMARK 11.1.– Transition probabilities [11.2] and [11.3] are obtained as follows:

– Transition from i to $i+1$:

No coupons out of i available disappear from the collection with probability $d_{i,0}$, and a new coupon is drawn with probability $\frac{n-i}{n}$.

– Transition from i to $i-k$, $0 \leq k \leq i-1$:

Exactly $k+1$ coupons out of i available disappear from the collection with probability $d_{i,k+1}$, and a new coupon is drawn with probability $\frac{n-i+k+1}{n}$, because any of $k+1$ lost coupons can be drawn again.

Exactly k coupons out of i available disappear from the collection with probability $d_{i,k}$ and an existing coupon is drawn with probability $\frac{i-k}{n}$.

REMARK 11.2.– We have:

$$\begin{aligned}
\sum_{k=1}^{i+1} p_{i,k} &= \sum_{k=0}^{i-1} p_{i,i-k} + p_{i,i+1} \\
&= \sum_{k=0}^{i-1} \frac{n-i+k+1}{n} d_{i,k+1} + \sum_{k=0}^{i-1} \frac{i-k}{n} d_{i,k} + \frac{n-i}{n} d_{i,0} \\
&= \frac{n-i}{n} \sum_{k=0}^{i-1} d_{i,k+1} + \sum_{k=0}^{i-1} \frac{k+1}{n} d_{i,k+1} + \sum_{k=0}^{i-1} \frac{i-k}{n} d_{i,k} + \frac{n-i}{n} d_{i,0} \\
&= \frac{n-i}{n} \sum_{k=0}^i d_{i,k} + \frac{i}{n} \sum_{k=0}^{i-1} d_{i,k} + \sum_{k=0}^{i-1} \frac{k+1}{n} d_{i,k+1} - \sum_{k=0}^{i-1} \frac{k}{n} d_{i,k} \\
&= \frac{n-i}{n} \sum_{k=0}^i d_{i,k} + \frac{i}{n} \sum_{k=0}^{i-1} d_{i,k} + \frac{i}{n} d_{i,i} \\
&= \frac{n-i}{n} \sum_{k=0}^i d_{i,k} + \frac{i}{n} \sum_{k=0}^i d_{i,k} \\
&= 1.
\end{aligned}$$

This confirms that $\mathbf{P}^{(PR1)}$ is indeed a transition probability matrix, regardless of the choice of the distribution D_i .

REMARK 11.3.– By setting:

$$d_{i,j} = \begin{cases} 1, & j = 0; \\ 0, & j \neq 0, \end{cases}$$

this version of the problem is reduced to CCP.

Next, we will propose two useful and intuitively clear choices for the distribution D_i .

EXAMPLE 11.1.– The simplest case is obtained when D_i has uniform distribution over the set $\{0, 1, \dots, i\}$, that is:

$$d_{i,j} = \frac{1}{i+1},$$

which means that any number of coupons will disappear from the collection with the same probability.

EXAMPLE 11.2.– Let D_i be the binomial distribution with the parameter $0 \leq p_{R1} \leq 1$:

$$d_{i,j} = \binom{i}{j} p_{R1}^j (1 - p_{R1})^{i-j}.$$

This means that each coupon already collected disappears from the collection with probability p_{R1} .

This example is relevant because it can easily be generalized to the case of the CCP with unequal probabilities by assuming that each coupon $i, i \in \mathbb{N}_n$ disappears from the collection with probability p_i . This is of great importance in the context of modeling parasitism, as different hosts have different predispositions to being parasitized.

11.3.2. *Partial reset (2)*

In this version of the problem at each point of time j out of i coupons disappear from the collection with probability $d_{i,j}$, after a new coupon is drawn. This is more realistic from a biological point of view than the previous model, as some host species can be very resilient (unable to be attacked by parasites).

The transition probability matrix is:

$$\mathbf{P}^{(PR2)} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & \dots & n-1 & n \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ n-2 \\ n-1 \\ n \end{matrix} & \begin{pmatrix} p_{00} & p_{01} & 0 & 0 & \dots & 0 & 0 \\ p_{10} & p_{11} & p_{12} & 0 & \dots & 0 & 0 \\ p_{20} & p_{21} & p_{22} & p_{23} & \dots & 0 & 0 \\ \vdots & & & & & & \vdots \\ p_{n-2,0} & p_{n-2,1} & p_{n-2,2} & p_{n-2,3} & \dots & p_{n-2,n-1} & 0 \\ p_{n-1,0} & p_{n-1,1} & p_{n-1,2} & p_{n-1,3} & \dots & p_{n-1,n-1} & p_{n-1,n} \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \end{matrix},$$

where the transition probabilities are:

$$p_{i,i+1} = \frac{n-i}{n} d_{i+1,0}, \quad 0 \leq i \leq n-1, \quad [11.4]$$

and

$$p_{i,i-k} = \frac{n-i}{n} d_{i+1,k+1} + \frac{i}{n} d_{i,k}, \quad [11.5]$$

with $0 \leq i \leq n-1$, $0 \leq k \leq i$, and d_{ij} being the probability that j out of i available coupons disappear from the collection (as in the previous model).

REMARK 11.4.– Transition probabilities [11.4] and [11.5] are obtained as follows:

- Transition from i to $i+1$:

With i coupons in the collection, a new coupon is drawn with probability $\frac{n-i}{n}$, and no coupons disappear with probability $d_{i+1,0}$.

- Transition from i to $i-k$, $0 \leq k \leq i-1$:

With i coupons in the collection, a new coupon is drawn with probability $\frac{n-i}{n}$, and after that $k+1$ coupons are lost with probability $d_{i+1,k+1}$. An existing coupon is drawn with probability $\frac{i}{n}$, and after that k coupons are lost with probability $d_{i,k}$.

REMARK 11.5.– Noting that:

$$\begin{aligned}
\sum_{k=0}^{i+1} p_{i,k} &= \sum_{k=0}^i p_{i,i-k} + p_{i,i+1} \\
&= \sum_{k=0}^i \frac{n-i}{n} d_{i+1,k+1} + \sum_{k=0}^i \frac{i}{n} d_{i,k} + \frac{n-i}{n} d_{i+1,0} \\
&= \frac{n-i}{n} \sum_{k=0}^{i+1} d_{i+1,k} + \frac{i}{n} \sum_{k=0}^i d_{i,k} \\
&= 1.
\end{aligned}$$

confirms that $\mathbf{P}^{(PR2)}$ is indeed a transition probability matrix, regardless of the choice of the distribution D_i .

By setting:

$$d_{i,j} = \frac{1}{i+1}, \quad [11.6]$$

and

$$d_{i,j} = \binom{i}{j} p_{R2}^j (1 - p_{R2})^{i-j}, \quad 0 \leq p_{R2} \leq 1, \quad [11.7]$$

respectively, we obtain the examples analogous to those for the previous model. Remark 11.3 is also valid.

11.4. Model comparison

We illustrate the behavior of the three models analyzed in sections 11.2 and 11.3 by comparing them according to three criteria:

- probability of completing a full collection of coupons in the minimum number of steps;
- expected waiting time until completing a full collection;
- variance of the waiting time until completing a full collection.

11.4.1. Probability of completing a full collection of coupons in the minimum number of steps

The probability of completing a full collection of coupons in the minimum number of steps (denoted as $P^{(min)}$) is calculated for each variant of the CCP as the product of the elements above the main diagonal in the corresponding transition probability matrix, that is,

$$P^{(min)} = \prod_{j=0}^{n-1} p_{j,j+1}.$$

In the context of modeling parasitism, $P^{(s)}$ gives the probability that the entire host population will be parasitized in the smallest time horizon.

Next, we give the expression $P^{(s)}$ for the three generalizations of the CCP analyzed in the previous section.

- CCP with total reset:

$$P_{TR}^{(min)} = (1 - p_R)^n \cdot \frac{(n-1)!}{n^{n-1}}; \quad [11.8]$$

- CCP with partial reset (1):

$$P_{PR1}^{(min)} = \frac{(n-1)!}{n^{n-1}} \prod_{j=0}^{n-1} d_{j,0};$$

- CCP with partial reset (2):

$$P_{PR2}^{(min)} = \frac{(n-1)!}{n^{n-1}} \prod_{j=0}^{n-1} d_{j+1,0}.$$

In particular, when D_i is the binomial distribution, the last two expressions reduce to:

$$P_{PR1}^{(min)} = \frac{(n-1)!}{n^{n-1}} (1 - p_{R1})^{\frac{n(n-1)}{2}} \quad [11.9]$$

and

$$P_{PR2}^{(min)} = \frac{(n-1)!}{n^{n-1}} (1 - p_{R2})^{\frac{n(n+1)}{2}}, \quad [11.10]$$

respectively, and the three models become directly comparable.

11.4.2. Expected value and variance of the waiting time until completing a full collection

The expected waiting time and the variance of the waiting time to complete a collection are calculated from the corresponding fundamental matrices, which is a standard technique for analyzing absorbing Markov chains.

Let us introduce some notations. In the rest of the text, $\mathbf{0}$ denotes the zero matrix, $\mathbf{I}$ is the identity matrix and $\mathbf{1}$ denotes the matrix with all entries equal to 1. Next, $\mathbf{A}^{(sq)}$ denotes the matrix whose elements are squared elements of the matrix $\mathbf{A}$ and $S_1(\mathbf{A})$ denotes the sum of the entries of the first row of the matrix $\mathbf{A}$.

Consider a finite Markov chain with r absorbing and t transient states. The transition matrix $\mathbf{P}$ of this Markov chain can be written in the following canonical block matrix form:

$$\mathbf{P} = \left(\begin{array}{c|c} \mathbf{Q}_{t \times t} & \mathbf{R}_{t \times r} \\ \hline \mathbf{0}_{r \times t} & \mathbf{I}_{r \times r} \end{array} \right),$$

where the matrix $\mathbf{Q}_{t \times t}$ describes the probabilities of motion among the transient states, and the matrix $\mathbf{R}_{t \times r}$ gives the probabilities of absorption in one step (going from one of the transient states to one of the absorbing states in a single step).

The fundamental matrix $\mathbf{F}$ is obtained as:

$$\mathbf{F} = (\mathbf{I} - \mathbf{Q})^{-1} = \sum_{k=0}^{+\infty} \mathbf{Q}^k,$$

where the i th row of the matrix $\mathbf{F}$ consists of the expected number of times the chain is in each transient state before being absorbed, starting from transient state i (see, for example, Kemeny and Snell (1976)). Then, the matrices consisting of the corresponding expected waiting times and

variances of the waiting times, respectively, starting from transient states, are:

$$\mathbf{E} = \mathbf{F} \cdot \mathbf{1}_{t \times 1}, \quad \mathbf{V} = (2\mathbf{F} - \mathbf{I})\mathbf{E} - \mathbf{E}^{(sq)}.$$

Finally, a way to compute the expectation and the variance of the waiting time $W_n^{(\cdot)}$ for any variant of the CCP, when we assume starting from an empty collection, that is, state 0, is:

$$E\left(W_n^{(\cdot)}\right) = S_1(\mathbf{E}), \quad Var\left(W_n^{(\cdot)}\right) = S_1(\mathbf{V}). \quad [11.11]$$

For more details, see Kemeny and Snell (1976).

11.5. Numerical results

In this section, we present numerical results related to the expected value and variance of the waiting time $W_n^{(\cdot)}$ in generalizations of the CCP analyzed in section 11.2 and section 11.3. More precisely, we compare the CCPTR with versions of the CCP with partial reset (1) obtained when choosing uniform distribution (Example 11.1) and binomial distribution (Example 11.2, referred to as CCPPR1 in the rest of the text) and partial reset (2), which is obtained by choosing uniform distribution (equation [11.6]) and binomial distribution (equation [11.7], referred to as CCPPR2 in the rest of the text).

In all scenarios considered, it is assumed that the set of available coupons consists of 10 different standard coupons ($\mathbb{N}_{10} = \{1, 2, \dots, 10\}$).

All numerical calculations were carried out using the statistical software R.

11.5.1. $E\left(W_n^{(\cdot)}\right)$ and $Var\left(W_n^{(\cdot)}\right)$ for different values of reset probabilities

In Table 11.1 we show the expected waiting time and variance of the waiting time $W_n^{(\cdot)}$ in CCPTR, CCPPR1 and CCPPR2, respectively, for:

$$p_R = p_{R1} = p_{R2} \in \left\{ \frac{1}{3}, \frac{1}{4}, \frac{1}{8}, \frac{1}{10}, \frac{1}{15}, \frac{1}{50}, \frac{1}{100}, 0 \right\},$$

obtained using [11.11].

Table 11.1. Expected waiting time and variance of the waiting time in CCPTR, CCPPR1 and CCPPR2 in terms of $p = p_R = p_{R1} = p_{R2}$

p	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{20}$	$\frac{1}{50}$	$\frac{1}{100}$	0
$E(W_n^{(1)})$	9006	1768	192	127	75	59	38	33	29
$V_{ar}(W_n^{(1)})$	17392517	505677	3658	1497	513	325	169	144	126
$E(W_n^{(2)})$	3.751e+10	97586665	22799	4524	566	217	54	38	29
$V_{ar}(W_n^{(2)})$	4.152e+20	2.481e+15	1.121e+08	4279627	65083	9411	493	229	126
$E(W_n^{(3)})$	2.144e+12	1.697e+09	81559	12031	1017	323	59	40	29
$V_{ar}(W_n^{(3)})$	1.321e+24	7.299e+17	1.387e+09	29249227	203889	20315	613	251	126

The corresponding results for the CCP with partial reset (1) and CCP with partial reset (2) in the case of uniform distribution, are:

- for partial reset (1),

$$E(W_n^{(1)}) = 2623315604, V_{ar}(W_n^{(1)}) = 2.4097e + 18;$$

- for partial reset (2),

$$E(W_n^{(2)}) = 28560721303, V_{ar}(W_n^{(2)}) = 2.6886e + 20.$$

11.5.2. $E(W_n^{(\cdot)})$ and $V_{ar}(W_n^{(\cdot)})$ for the same $P^{(min)}$

In Table 11.2 we show the expected waiting time and variance of the waiting time $W_n^{(\cdot)}$ in CCPTR, CCPPR1 and CCPPR2, respectively, for

$$p_R \in \left\{ \frac{1}{3}, \frac{1}{4}, \frac{1}{8}, \frac{1}{10}, \frac{1}{15}, \frac{1}{50}, \frac{1}{100}, 0 \right\},$$

and p_{R1}, p_{R2} computed from the condition

$P_{TR}^{(min)} = P_{PR1}^{(min)} = P_{PR2}^{(min)}$ using equations [11.8], [11.9] and [11.10]. $E \left(W_n^{(\cdot)} \right)$ and $Var \left(W_n^{(\cdot)} \right)$ are obtained using equation [11.11].

Table 11.2. Expected waiting time and variance of the waiting time in CCPTR, CCPPR1 and CCPPR2, for different values of p_R and p_{R1}, p_{R2} obtained from $P_{TR}^{(min)} = P_{PR1}^{(min)} = P_{PR2}^{(min)}$

p	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{20}$	$\frac{1}{50}$	$\frac{1}{100}$	0
$E \left(W_{10}^{(TR)} \right)$	9006	1768	192	127	75	59	38	33	29
$Var \left(W_{10}^{(TR)} \right)$	17392517	505677	3658	1497	513	325	169	144	126
p_{R1}	0.0862	0.0619	0.0292	0.0231	0.0152	0.0113	0.0045	0.0022	0
$E \left(W_{10}^{(PR1)} \right)$	1878	428	78	61	45	40	33	31	29
$Var \left(W_{10}^{(PR1)} \right)$	727887	36991	1115	644	336	252	161	141	126
p_{R2}	0.0711	0.05096	0.0239	0.0189	0.0125	0.0093	0.0037	0.0018	0
$E \left(W_{10}^{(PR2)} \right)$	1396	344	72	57	44	39	33	31	29
$Var \left(W_{10}^{(PR2)} \right)$	384898	23084	915	555	308	237	157	140	126

11.5.3. $E \left(W_n^{(\cdot)} \right)$ and $Var \left(W_n^{(\cdot)} \right)$ for different values of n

In Figures 11.1, 11.2 and 11.3 we illustrate the dependence of the expected waiting time $E \left(W_n^{(\cdot)} \right)$ in CCPTR, CCPPR1 and CCPPR2, respectively, on the probabilities p_R, p_{R1}, p_{R2} , for different values of n . We have:

$$p_R, p_{R1}, p_{R2} \in \left\{ \frac{1}{3}, \frac{1}{4}, \frac{1}{8}, \frac{1}{10}, \frac{1}{15}, \frac{1}{50}, \frac{1}{100}, 0 \right\}.$$

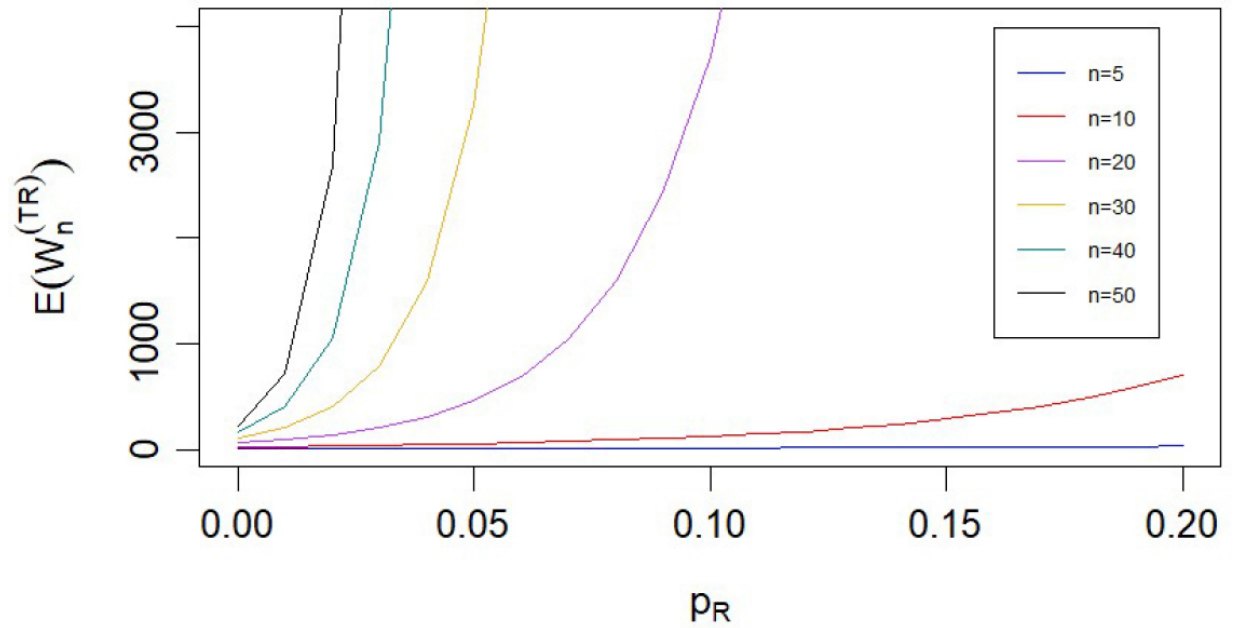


Figure 11.1. Expected waiting time in CCPTR in terms of p_R for different values of n

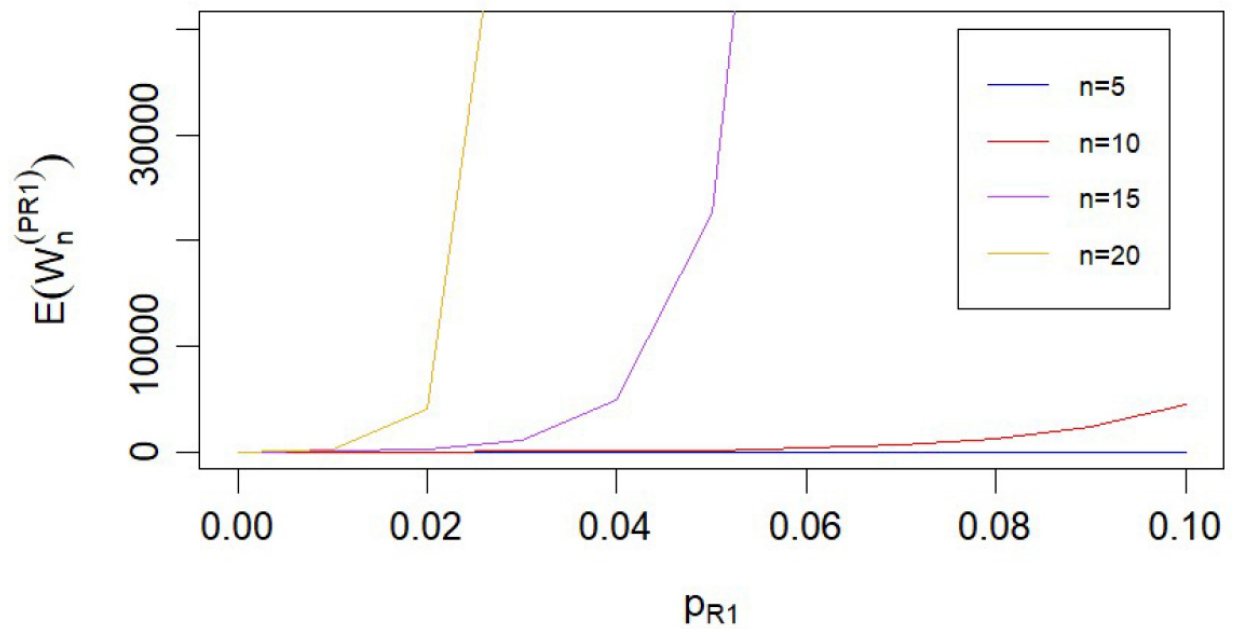


Figure 11.2. Expected waiting time in CCPPR1 in terms of p_{R1} for different values of n

11.5.4. Conclusion

We can draw the following conclusions from the numerical experiments:

- While obtaining the results, we have encountered several numerical difficulties (or even failure to calculate any results). They do not seem to be related to the version of the problem, but rather to the range of the reset probability and the relationship between the reset probability and n . More specifically, they occur when the reset probability (p_R , p_{R1} , or p_{R2}) is greater than 0.5, or, for $n \geq 20$ and the reset probability close to $\frac{1}{n}$. Greater difficulties were encountered with partial resets with uniform distribution, when we did not obtain a result for $n \geq 15$. We suspect that these problems are due to the form of the transition probability matrices (many very small entries when calculating higher powers and inverses), which requires the use of more subtle calculation methods or some kind of approximation.
- In general, the behavior of the expected waiting time matches the intuition we have about the problem: the expected waiting time increases as n increases (as a larger number of standard coupons extends the waiting time) and it also increases with the probability of a reset (partial or total), since a reset removes the collected coupons, which increases the waiting time. In several cases, the expected waiting time appears to grow exponentially with n .
- If we compare the CCPTR, CCPPR1 and CCPPR2 models, we can see that CCPPR2 generally has the largest expected waiting time and the largest variance of the waiting time, which is also intuitively clear, since the corresponding Markov chain has the largest number of (reachable) states.

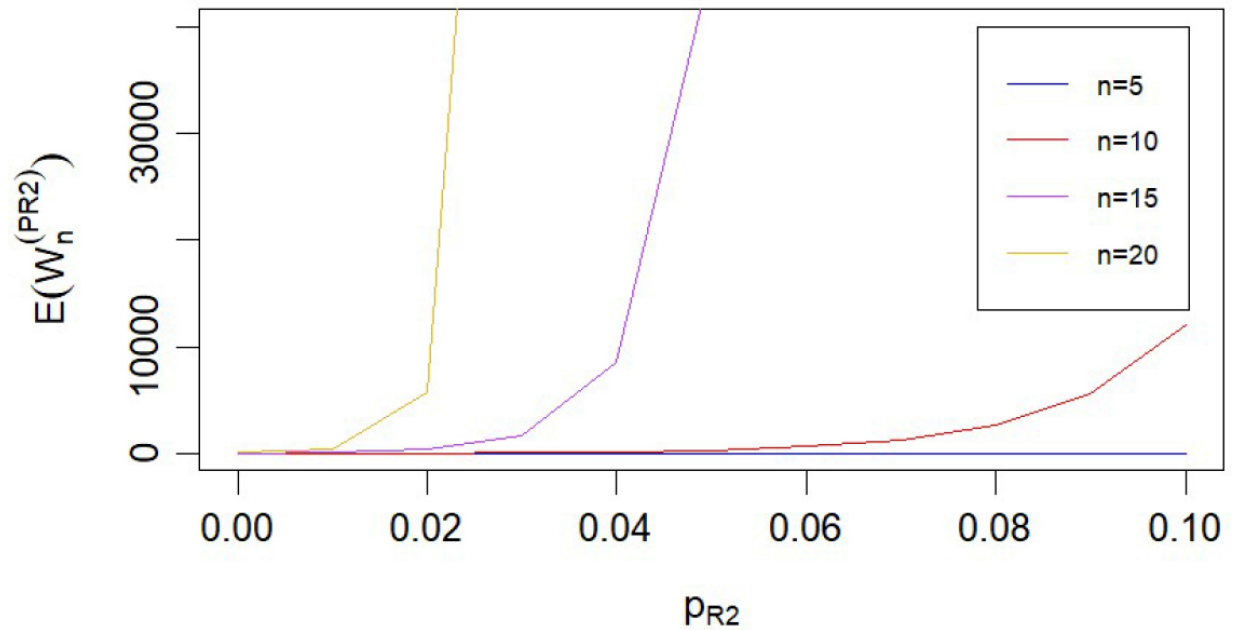


Figure 11.3. Expected waiting time in CCPPR2 in terms of p_{R2} for different values of n

11.6. Discussion

In this chapter, we have considered three different generalizations of the classical CCP that include the possibility that some or all of the already collected coupons disappear from the collection. To our knowledge, the versions of the problem with partial resets, motivated by the particular application in modeling parasitism, have not been considered so far. We obtained some theoretical properties and numerical comparisons for these models.

There are several possibilities to extend this work.

In the context of modeling parasitism, questions of interest would be how to estimate the reset probability from real data related to parasitism, or how to generalize these models for the case of the CCP with unequal probabilities (which corresponds to the realistic situation where different hosts have different probabilities of being parasitized).

In the probabilistic context, the question of obtaining a closed form and asymptotic properties of the expected value and variance of the waiting times in the versions of the problem with partial resets may be of interest.

11.7. Acknowledgments

This work was supported by Ministry of Science, Technological Development and Innovation of the Republic of Serbia: Grant No. 451-03-136/2025-03/200104.

11.8. References

- Anceaume, E., Busnel, Y., Sericola, B. (2015). New results on a generalized coupon collector problem using Markov chains. *Journal of Applied Probability*, 52, 2, 405–418.
- Anceaume, E., Busnel, Y., Schulte-Geers, E., Sericola, B. (2016). Optimization results for a generalized coupon collector problem. *Journal of Applied Probability*, 53, 622–629.
- Jocković, J. and Todić, B. (2024a). Coupon collector problem with reset button. *Mathematics*, 12(2), 239. doi: 10.3390/math12020239.
- Jocković, J. and Todić, B. (2024b). Markov chain approach to the coupon collector problem with universal coupon. *Filomat*, 38(24), 8565–8581.
- Jocković, J. and Todić, B. (2024c). Some bounds on the coupon collector problem with universal coupon. *Analele Stiintifice ale Universitatii Ovidius Constanta: Seria Matematica*, 32(2), 99–113.
- Jocković, J. and Todić, B. (2024d). Waiting time for a small subcollection in the coupon collector problem with universal coupon. *Journal of Theoretical Probability*, 37(3), 1945–1957.
- Kemeny, J. and Snell, L. (1976). *Finite Markov Chains*. Springer, New York.
- Nakata, T. and Kubo, I. (2006). A coupon collector's problem with bonuses. In *Proceedings of the 4th Colloquium on Mathematics and Computer Science, DMTCS Proceedings*. Springer, New York..
- Singh, A. (2021). Stochasticity in host-parasitoid models informs mechanisms regulating population dynamics. *Scientific Reports*, 11, 16749.
- Todić, B. (2024). Coupon collector problem with penalty coupon. *Matematički Vesnik*, 76(1–2), 15–28.
- Zoróa, N., Lesigne, E., Fernández-Sáez, M.J., Zoróa, P., Casas, J. (2017). The coupon collector urn model with unequal probabilities in ecology and evolution. *Journal of The Royal Society Interface*, 14, 20160643.

Note

Chapter written by Jelena JOCKOVIĆ and Bojana TODIĆ.

12

The Relationship Between Quality of Life and Brain Drain

This chapter explores the relationship between quality of life and brain drain across countries with varying income levels over the period 2000–2022. Due to the limited availability of direct data on skilled emigration, we use a perception-based indicator from the World Yearbook, which reflects the extent to which brain drain is seen as a barrier to national competitiveness. To measure quality of life, we construct a composite index based on a set of economic, social and demographic indicators sourced from the World Bank. Our analysis reveals a consistent positive correlation between quality of life and the perception of reduced negative impact from brain drain at the global level. However, when countries are grouped by income, the relationship appears to diverge significantly. In low-income countries, improvements in quality of life are associated with a growing negative impact of brain drain, suggesting that rising living standards may increase the mobility of skilled individuals without improving institutional capacity to retain them. In contrast, middle- and high-income countries show a positive correlation, indicating that better living conditions may help mitigate the adverse effects of brain drain or reflect higher resilience to talent outflows. These findings highlight the importance of considering country-specific economic and institutional contexts when assessing the dynamics of skilled migration. They also suggest that enhancing quality of life alone may not be sufficient to reduce the economic costs associated with brain drain in lower income settings.

12.1. Introduction

Why do people migrate? There are numerous reasons why individuals choose to relocate. These reasons can generally be grouped into several categories: economic, educational,

environmental, political, familial and others. When employment opportunities are lacking in a given location, people are likely to move elsewhere in search of a livelihood (Fantazzini et al. [2021](#)). However, a depressed economy can also act as a restraining factor, reducing migration flows (Bergh et al. [2015](#)). Relocation may be motivated by a desire for better political institutions and improved access to human rights. It is common for young people to migrate in pursuit of higher quality education (Bagdonaitė et al. [2022](#)). In some cases, individuals move to countries with more suitable climates due to medical conditions. Moreover, the motivations behind migration are not fixed and may evolve over time (Carling [2024](#)).

All of the above can likely be integrated into a broader category – migration aimed at improving the individual's quality of life (Geist and McManus [2012](#)). However, studies exploring the relationship between quality of life and migration indicate that there is no consistent or universal link between the two across all countries. No general pattern appears to govern migration flows based on quality-of-life indicators. Brain drain, as a specific type of migration, has attracted particular scholarly interest. Since the term typically refers to the emigration of highly skilled professionals – such as doctors, engineers and researchers – the donor country must be concerned not only with the loss of high-potential labor but also with broader systemic issues, such as the structure and efficacy of its educational system (Gibson and McKenzie [2011](#)).

The relationship between quality of life and brain drain is actively studied within the frameworks of migration and institutional economic theories. A number of studies suggest that a high quality of life in receiving countries and a low quality of life in sending countries are key determinants of high-skilled emigration (Beine et al. [2008](#); Docquier and Rapoport [2012](#)). In this context, quality of life is interpreted broadly – it encompasses economic well-being, access to education and healthcare, political stability and civil liberties. The higher the perceived overall quality of life in a given country, the more attractive it becomes for mobile professionals (Czaika and Parsons [2017](#)).

However, there is no consensus in the literature regarding the consequences of brain drain. Some studies argue that the emigration

of skilled workers can weaken the development potential of low-income countries and slow their economic modernization (Lowell and Findlay [2001](#)). At the same time, an alternative perspective has emerged, suggesting that emigration does not necessarily lead to permanent loss. In contrast, it may initiate a process of brain circulation (Saxenian [2005](#)). Under this paradigm, skilled migrants return to their home countries or maintain transnational ties, contributing to the transfer of knowledge, technology and investment.

Lee and Kurbatskii ([2024](#)) provides critical empirical evidence on the economic and social consequences of migration in Russia – a context shaped by ongoing brain drain and shifting migration flows. By utilizing panel data from 80 Russian regions, it quantifies the effects of skilled and general migration on both productivity and income levels, highlighting how outflows suppress growth while inflows boost regional economies. Importantly, it reveals the heterogeneity of these effects: less populated regions gain disproportionately from new arrivals, while densely populated areas see muted benefits. These insights are vital for designing effective migration and innovation policies aimed at improving quality of life and mitigating the loss of human capital.

The concept of brain circulation is especially relevant in the context of high-income countries, where both state institutions and the private sector have developed mechanisms not only to retain domestic talent but also to attract skilled professionals back. These economies often have institutional frameworks that enable them to partially offset brain drain through educational exchanges, research funding programs and the repatriation of diaspora professionals (Kapur and McHale [2005](#)). Moreover, in countries with a high quality of life, emigration may not be perceived as a threat but rather as part of the global redistribution of human capital.

The aim of this chapter is to examine the relationship between quality of life – measured using principal component analysis (PCA) – and a proxy variable for brain drain.

12.2. Data and methodology

Reliable data on the actual scale of brain drain are largely unavailable. With rare exceptions, we do not know why individuals cross national borders – whether for leisure or employment purposes. Official migration statistics provide information on the volume of migration flows but not on the quality of the workforce.

Therefore, in this study, we use a proxy variable – indicator 3.2.21 “Brain drain” – sourced from the World Competitiveness Yearbook (WCY) published by the International Institute for Management Development (IMD) in Switzerland. To assess a country’s competitiveness, the IMD surveys business executives, asking them questions related to various aspects of their country’s economic environment. One such question is: “Brain drain (well-educated and skilled people) does not hinder competitiveness in your economy?” Responses are given on a 10-point scale, where higher values indicate that brain drain is perceived to have a less negative impact on the national economy.

Table 12.1. *List of variables*

	Model name	Source: World Bank
1	GDP_per_capita	GDP per capita, PPP (current international \$)
2	Lab_effect	Labor productivity - GDP per person employed (constant 2021 PPP \$)
3	Consumption	Households and NPISHs Final consumption expenditure per capita (constant 2015 US\$)
4	Literacy_rate	Literacy rate, adult total (% of people ages 15 and above)
5	Fundsup	Income share held by highest 20%
6	Fundsdown	Income share held by lowest 20%
7	Funds	Income share held by highest 20%/Income share held by lowest 20%, percent
8	Inflation_GDP_deflator	Inflation, GDP deflator (annual %)
9	Expectancy	Life expectancy at birth, total (years)
10	Mortality	Mortality rate, infant (per 1,000 live births)
11	Carbon	Carbon dioxide (CO ₂) emissions (total) excluding LULUCF (% change from 1990)
12	Resdev	Research and development expenditure (% of GDP)
13	Fertility	Birth rate, total (births per woman)
14	Gini	Gini index
15	GNI	Gross national income per capita, PPP (current international \$)
	Model name	Source: World Competitiveness Yearbook

	Model name	Source: World Bank
16	Brain drain	3.2.21. <i>Brain drain (well-educated and skilled people) does not hinder competitiveness in your economy</i>

To assess quality of life, we used data provided by the World Bank, covering the period from 2000 to 2022 inclusive. The methodology for calculating the quality-of-life index follows the approach proposed in Aivazian ([2016](#)) and Kurbatskii and Mironenkov ([2023](#)). The central idea of the method is to reduce the dimensionality of the indicator space while preserving the informativeness of the resulting composite index. The full list of determinant indicators is presented in [Table 12.1](#). For the period 2000–2022, we selected those countries from the World Bank database for which, across all 15 tracked indicators, the number of missing values in any given year did not exceed five. Only countries for which both macroeconomic data and brain drain data were simultaneously available were included in the analysis.

As previously noted, the dataset available to the authors contained a relatively small number of missing values. These were imputed using singular value decomposition (SVD), a method applied in several studies (e.g. Troyanskaya et al. [2001](#); Emmanuel et al. [2021](#)). Missing values are initially imputed with zeros. A rank-1 SVD approximation of the matrix is then computed. Only the missing entries are updated based on the approximation. This process is repeated iteratively until convergence is achieved, as measured by the relative error. This approach represents a classical iterative method for matrix completion based on low-rank approximation.

After imputing the missing values, we applied PCA. [Figure 12.1](#) presents the explained variance of the first principal component, which accounts for 48.9% in 2022 and exceeds 50% in all other years.

Proportion of Variance Explained by PC 1

Annual Analysis (2000-2022)

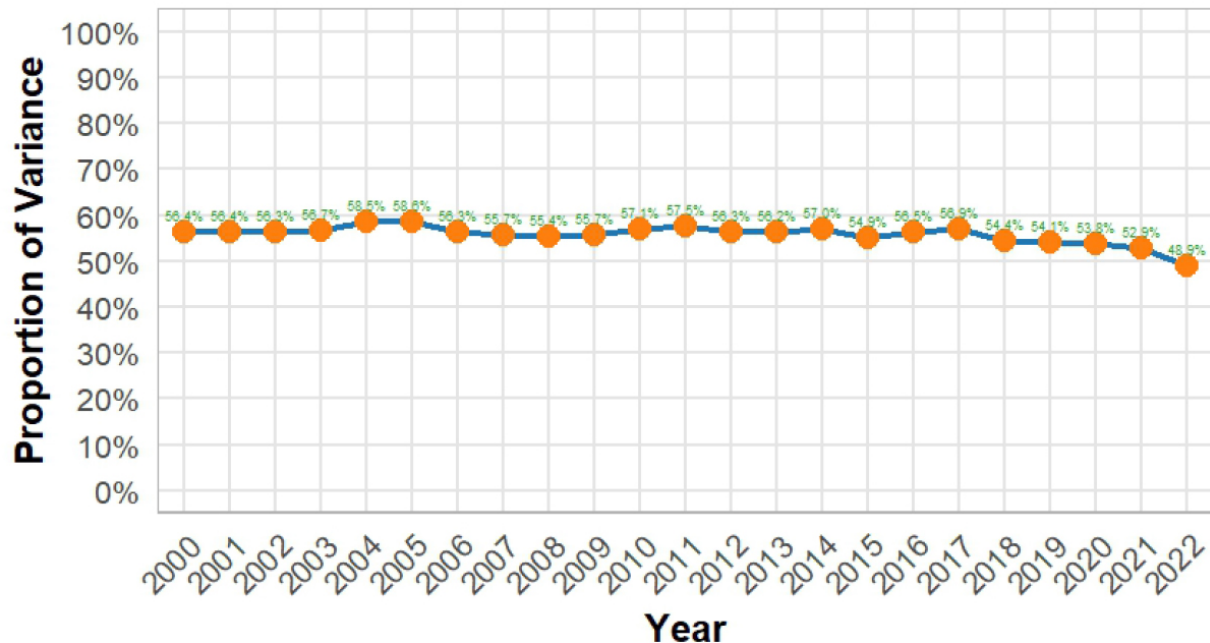


Figure 12.1. *Proportion of variance explained by the first component*

The component loadings remain relatively stable over time, with a noticeable change only in 2022, potentially reflecting the impact of global crises related to intensified military conflicts ([Figure 12.2](#)).

A set of indicators demonstrates consistently positive loadings, including GDP per capita, consumption per capita, life expectancy at birth, labor effectiveness, literacy rate and research and development expenditure. Conversely, consistently negative loadings are observed for the inflation rate (GDP deflator), Gini index, mortality rate and income ratio between the top 20% and the bottom 20% of the population. It is worth noting that birth rate exhibits a negative loading in our analysis. This implies that an increase in birth rate is associated with a lower quality-of-life index. This is likely because the analysis does not incorporate a longitudinal (lifecycle) perspective; rather, it assesses welfare at a single point in time. From this static viewpoint, having children may reduce current economic well-being. Moreover, we do not have access to subjective data that would allow us to assess how the presence of children affects

individuals' perceived well-being or happiness. Additionally, over the study period, the absolute value of the negative loading for carbon dioxide emissions (% change from 1990) has increased by a factor of 2.7, from -0.1 in 2002 to -0.271 in 2022. Visual inspection of the data indicates that the negative impact of emissions on quality of life has been steadily intensifying over time. Another important observation is the sharp decline in the negative effect of the Gini coefficient in 2022. While in previous years its loading was consistently around -0.3 , in 2022 it abruptly shifted to a positive value of $+0.006$.

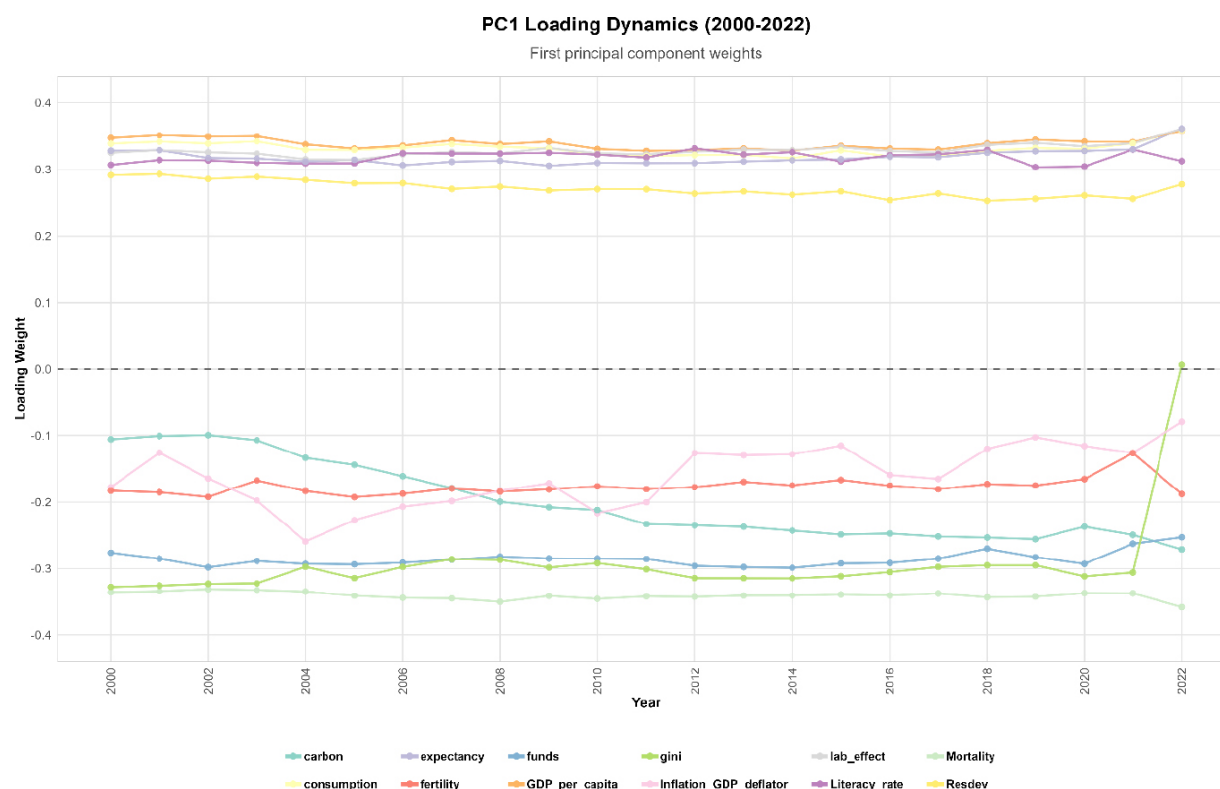


Figure 12.2. *Dynamics of the first component weight loads.*

12.3. Results

Following the application of PCA, we calculated the quality of life index for the 40 countries under study. The index was normalized so that the country with the lowest quality of life was assigned a value of 0, and the country with the highest quality of life received a value of 10. The heatmap displaying the calculated index values is presented in [Figure 12.3](#).

An analysis of the heatmap for the period 2000–2022 reveals the existence of stable clusters of countries according to their overall quality of life levels. Countries with a high quality of life – such as Luxembourg, Switzerland, Norway, Sweden, Finland, Denmark, and Austria – consistently demonstrate the highest index values, close to the upper bound of the scale. Their position at the top of the chart, along with the dark blue color of the heatmap cells, reflects a consistently high perceived standard of living and the stability of their socioeconomic systems.

The second group includes countries with a medium level of quality of life, such as Slovenia, Spain, Israel, Greece, Poland and Hungary. This group is characterized by moderate index values (approximately 4.5–7), along with positive dynamics observed in several cases – particularly in Eastern Europe – potentially reflecting the impact of institutional and economic reforms over the past two decades.

The third group comprises countries with low index values (below 3), including the Philippines, Brazil, Peru and Indonesia. These countries show light tones on the heatmap throughout the observed period, corresponding to lower quality of life scores. While there are instances of local improvement, the overall situation remains relatively stable and less favorable compared to other clusters.

In terms of temporal variation, most countries demonstrate a high degree of stability in their quality-of-life scores. These values change gradually over time without significant fluctuations, which is consistent with expectations for aggregate socioeconomic indicators. Notably, countries in the medium cluster exhibit more pronounced positive trends, whereas highly developed and least developed countries tend to show relative stagnation.

The ranking we obtained aligns reasonably well with traditional classifications, such as the Human Development Index (HDI), as well as with those produced using alternative approaches to constructing composite quality of life indicators (Mironenkov et al. [2024](#)).

The correlation analysis conducted in this study revealed a consistent presence of a moderately strong relationship between quality of life and brain drain. To assess the relationship between the level of economic well-being and the perception of the consequences of brain drain, a correlation analysis was performed with time lags ranging

from 0 to 10 years. As an indicator of brain drain, we used a variable where higher values reflect a less negative perceived impact of brain drain, while lower values indicate a more detrimental perception.

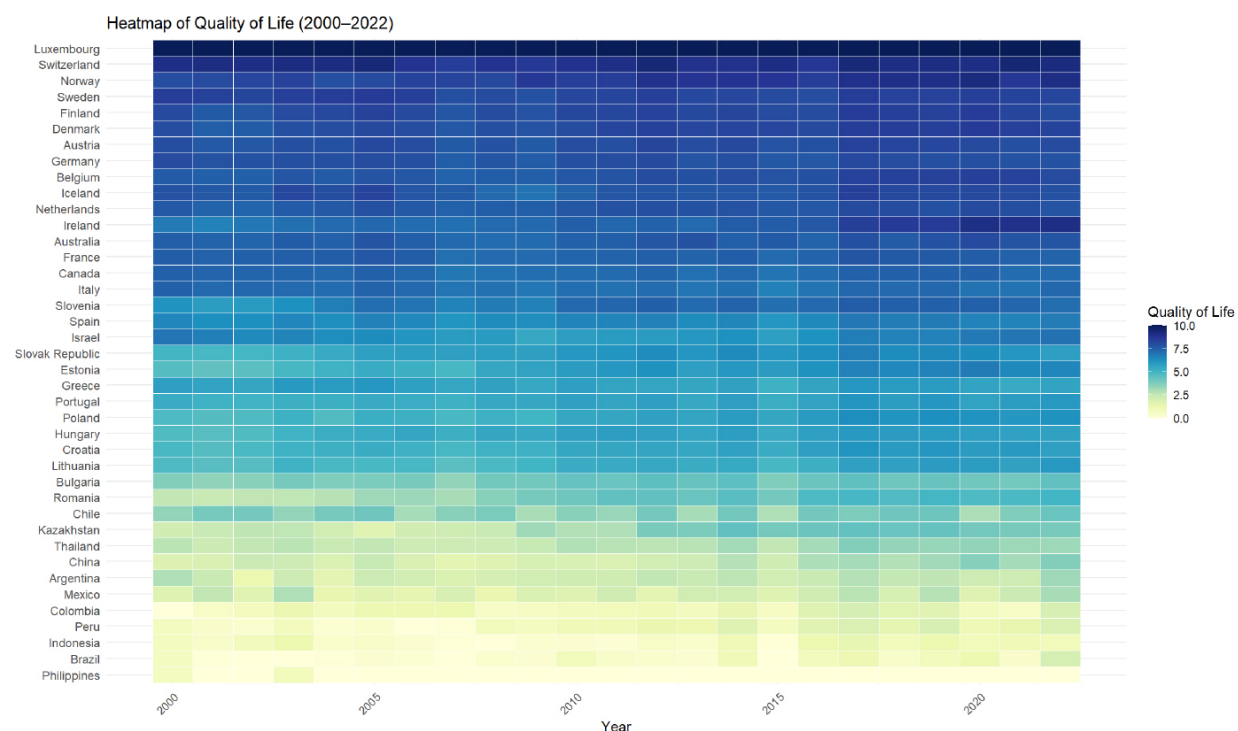


Figure 12.3. *Heatmap of quality of life.*

The results, presented in the form of a heatmap ([Figure 12.4](#)), demonstrate a stable positive correlation between quality of life and the subsequent decrease in the perceived threat of brain drain. The correlation coefficients range from 0.35 to 0.65, with the strongest relationship observed at time lags of 1–3 years. This suggests the presence of a delayed effect, in which an improvement in a country’s quality of life influences the perception of highly skilled migration with a time lag of 1–3 years.

It is worth noting that in countries with higher levels of quality of life, brain drain is not perceived as a significant barrier to economic competitiveness. This may be due to stronger institutional resilience, the ability to compensate for the outflow of human capital and integration into global flows of knowledge and talent.

The dynamics of the correlation values over time are also of interest: during the mid-2010s (2011–2015), the correlation coefficients declined, possibly reflecting the impact of global economic and

political shocks. However, after 2016, a steady increase in correlation values is observed, which may indicate a shift in the perception of highly skilled migration and a more positive attitude toward it against the backdrop of recovered economic stability in a number of countries.

In addition to the temporal analysis, we aimed to identify country-specific patterns. To this end, we classified the countries under study into three groups based on their per capita gross national income (GNI) adjusted for purchasing power parity (PPP) using data from the World Bank. The use of GNI per capita (PPP) in this context appears more appropriate than the more commonly used GDP, as it provides a measure that is closer to the potential income available to a country’s residents. Unlike GDP, GNI accounts for income earned by residents abroad (such as remittances and investment returns) and excludes income repatriated by foreign-owned companies. This makes GNI a more accurate indicator of the actual well-being of the population, particularly in countries with substantial labor migration (e.g. where migrant workers send remittances home) or where a significant share of profits is transferred to foreign investors.

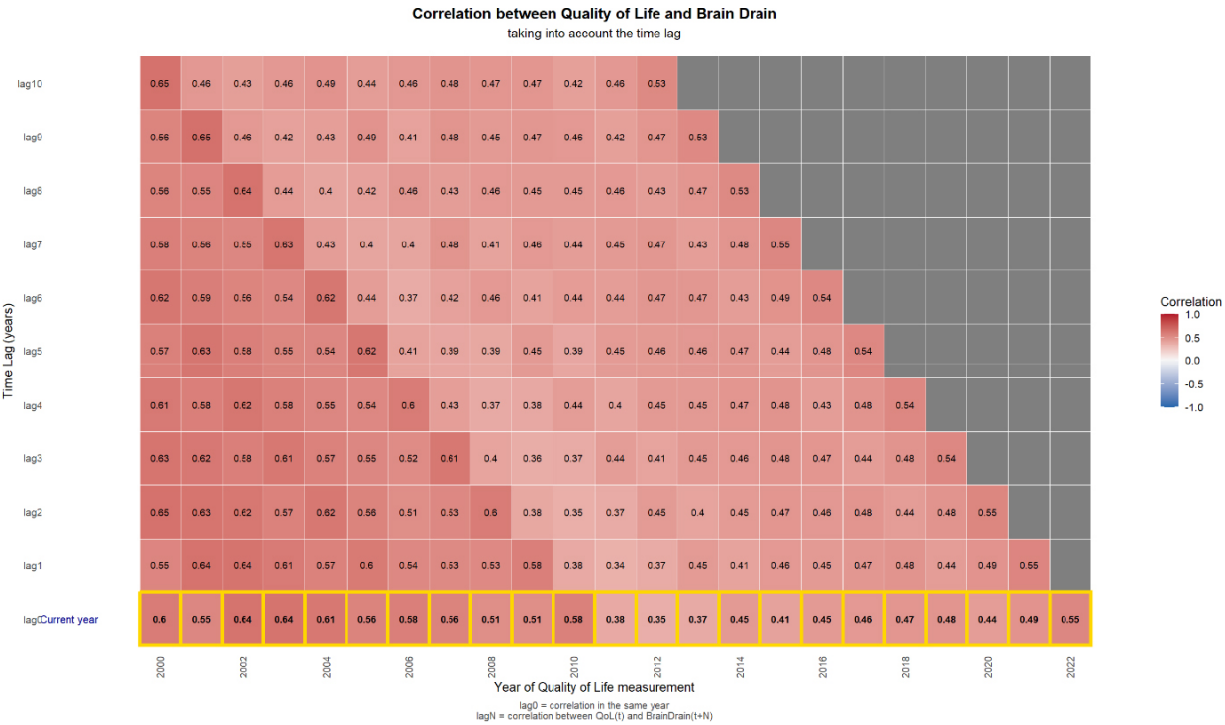


Figure 12.4. *Correlation between economic well-being and brain drain over the years.*

The grouping was determined annually according to the lower, middle and upper terciles of the income distribution. Consequently, group composition varied from year to year, reflecting changes in the underlying statistics. A number of countries consistently remained within the same income group throughout the entire period. The low-income group consistently included Brazil, Bulgaria, Chile, China, Colombia, Indonesia, Kazakhstan, Peru, Philippines and Thailand. The middle-income group was stably comprised of Hungary, Israel, Portugal, Slovenia, Spain. The high-income group consistently included Austria, Belgium, Denmark, Germany, Luxembourg, Netherlands, Norway, Sweden, Switzerland. Countries that shifted between groups over time are shown in [Figure 12.5](#).

Following the income-based grouping, a repeated correlation analysis was conducted. [Figure 12.6](#) presents a heatmap of the correlation between quality of life and the brain drain indicator, which reflects the perceived impact of brain drain on a country's economic competitiveness.

In low-income countries, a stable negative correlation between quality of life and brain drain has been observed since approximately 2008. This suggests that as the quality of life improves, the negative impact of brain drain on the economy increases (i.e. the indicator value decreases). A possible explanation is that even modest improvements in living conditions in poorer countries make citizens – particularly well-educated and skilled professionals – more mobile and capable of emigrating. As a result, countries with limited institutional and economic capacity may face greater losses due to the outflow of talent, despite overall improvements in living standards.

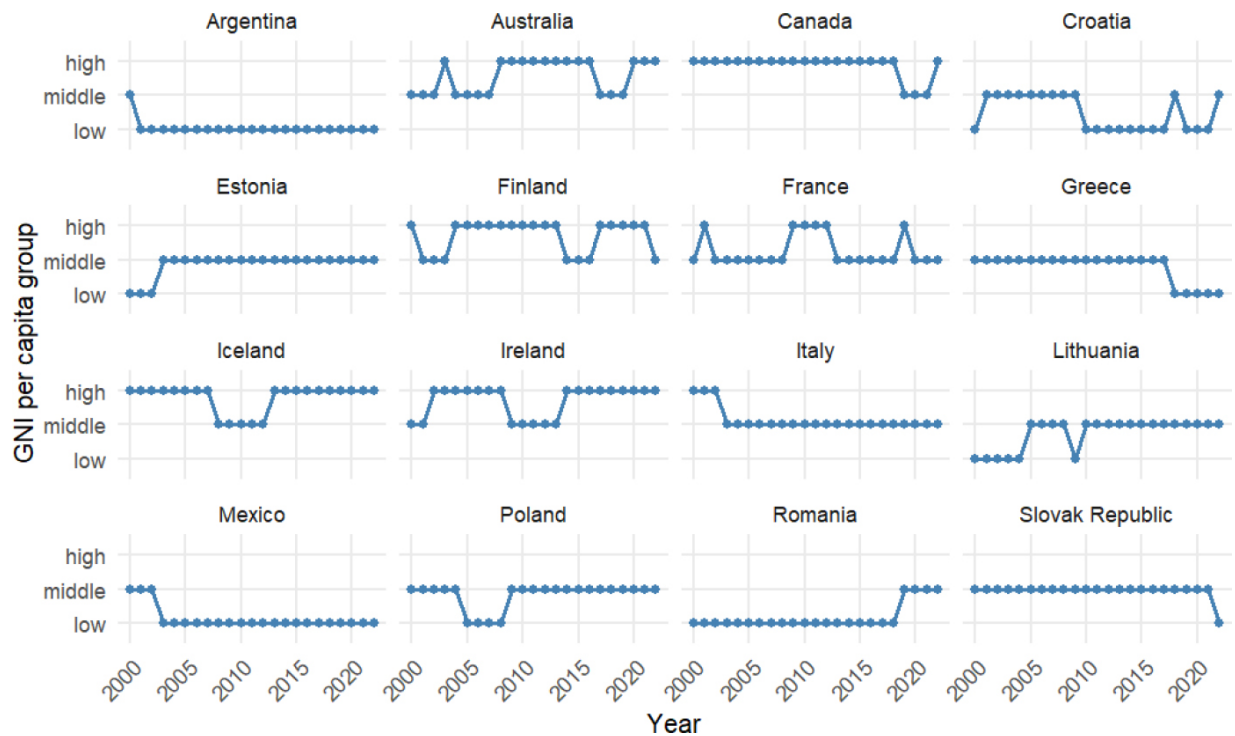


Figure 12.5. *Countries that change their income group*

In middle-income countries, the correlation remains consistently positive throughout the observed period. This indicates that improvements in quality of life in these countries are associated with a reduction in the negative consequences of brain drain. One possible reason is the gradual strengthening of institutional resilience and the development of a more attractive environment for skilled professionals, which helps to retain talent or offset its departure through bidirectional migration flows.

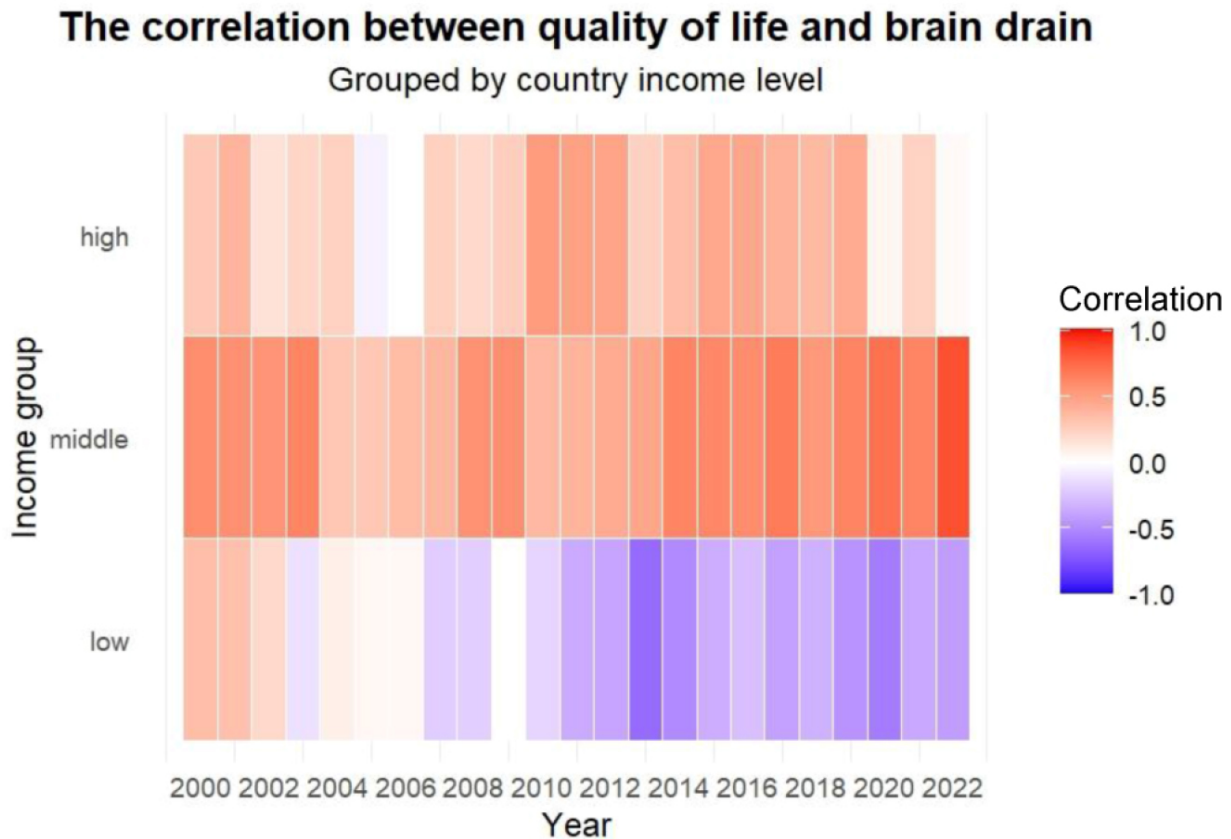


Figure 12.6. *The correlation between quality of life and brain drain.*

In high-income countries, the correlation is also positive, although weaker than in the middle-income group. This may suggest that in developed economies, quality of life is not the primary determinant of the consequences of brain drain. These countries typically possess strong institutional frameworks and economic capacity, allowing them to absorb or compensate for the loss of human capital, often through inflows of foreign talent or internal resources.

12.4. Conclusion

This study examined the relationship between quality of life and the perceived impact of brain drain across countries with varying income levels over the period 2000–2022. Using a composite quality-of-life index constructed through PCA and a perception-based brain drain indicator from the World Competitiveness Yearbook, we identified a complex and nonlinear association between these two phenomena.

At the global level, we observed a moderately strong positive correlation between improvements in quality of life and reduced negative perceptions of brain drain, with a lagged effect of approximately 1–3 years. However, when stratified by income, the patterns diverge considerably. In low-income countries, rising living standards are paradoxically associated with increased perceived harm from brain drain, potentially reflecting the enhanced capacity of skilled individuals to emigrate, without a parallel strengthening of institutions to retain them. In contrast, middle- and high-income countries display a consistently positive relationship, suggesting that better socioeconomic conditions may either mitigate the adverse effects of brain drain or indicate higher systemic resilience.

These findings emphasize the need for a nuanced, context-dependent approach to policymaking. For countries with lower income levels, improving quality of life alone is unlikely to curb skilled emigration unless accompanied by reforms that strengthen institutional capacity, professional opportunities and national innovation systems. For wealthier nations, brain drain may increasingly be embedded within global talent circulation, and policies should focus on maximizing the benefits of transnational knowledge exchange rather than solely on retention.

Future research could benefit from more granular data on actual skilled migration flows, as well as qualitative insights into migrants' motivations. Incorporating subjective well-being indicators and longitudinal tracking of professional trajectories could further enrich our understanding of how quality of life influences, and is influenced by, talent mobility in the global economy.

12.5. Acknowledgments

This research is funded by the Russian Science Foundation, project No. 25-18-00319 “Economic and demographic analysis of the population quality of life in Russia”.

12.6. References

- Aivazian, S. (2016). *Quality of Life and Living Standards Analysis: An Econometric Approach*. De Gruyter, Berlin, Boston. doi: 10.1515/9783110316254.
- Aytaç, S. and Aydın, G.Ç. (2019). Analysis of the effects of pull and push factors on brain drain: An example of Turkey. *Journal of Social Economics Research*, 6(2), 74–84.
- Bagdonaitė, R., Gudavičiūtė, K., Laertas, M., Stanevičius, R. (2022). Assessment of the reasons for migration of young people of working age. In *Individual. Society. State: Proceedings of the International Student and Teacher Scientific-Practical Conference*, 7–16. Rezekne Academy of Technologies, Rēzekne.
- Beine, M., Docquier, F., Rapoport, H. (2008). Brain drain and human capital formation in developing countries: Winners and losers. *The Economic Journal*, 118(528), 631–652.
- Bergh, A., Mirkina, I., Nilsson, T. (2015). Pushed by poverty or by institutions? Determinants of global migration flows. IFN Working Paper, No. 1077, Research Institute of Industrial Economics (IFN), Stockholm.
- Carling, J. (2024). Why do people migrate? Fresh takes on the foundational question of migration studies. *International Migration Review*, 58(4), 1757–1791.
- Czaika, M. and Parsons, C.R. (2017). The gravity of high-skilled migration policies. *Demography*, 54(2), 603–630.
- Docquier, F. and Rapoport, H. (2012). Globalization, brain drain, and development. *Journal of Economic Literature*, 50(3), 681–730.
- Emmanuel, T., Maupong, T., Mpoeleng, D., Semong, T., Mphago, B., Tabona, O. (2021). A survey on missing data in machine learning. *Journal of Big Data*, 8(1), 140.
- Fantazzini, D., Pushchelenko, J., Mironenkov, A., Kurbatskii, A. (2021). Forecasting internal migration in Russia using Google trends: Evidence from Moscow and Saint Petersburg. *Forecasting*, 3, 774–803.

- Geist, C. and McManus, P.A. (2012). Different reasons, different results: Implications of migration by gender and family status. *Demography*, 49(1), 197–217.
- Gibson, J. and McKenzie, D. (2011). Eight questions about brain drain. *Journal of Economic Perspectives*, 25(3), 107–128.
- Kapur, D. and McHale, J. (2005). The global migration of talent: What does it mean for developing countries. CGD Brief, Center for Global Development, Washington, D.C.
- Kurbatskii, A. and Mironenkov, A. (2023). Estimating the quality of life using weighted principal components method. *Montenegrin Journal of Economics*, 19(1), 7–17.
- Lee, E.S. (1966). A theory of migration. *Demography*, 3(1), 47–57.
- Lee, H.-S. and Kurbatskii, A.N. (2024). The effect of migration on economic and productivity growth in Russia. *Journal of Eurasian Studies*, 16(1), 123–138.
- Lowell, B.L. and Findlay, A. (2001). Migration of highly skilled persons from developing countries: Impact and policy responses. *International Migration Papers*, 44(25), 1–34.
- Mironenkov, A.A., Kurbatsky, A.N., Mironenkova, M.V. (2024). Measuring the quality of life of the population with a stochastic choice of weights for the weighted principal component. *Journal of Applied Economic Research*, 23(1), 82–109 (In Russian).
- Saxenian, A. (2005). From brain drain to brain circulation: Transnational communities and regional upgrading in India and China. *St Comp Int Dev*, 40, 35–61.
- Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D. and Altman, R.B. (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics*, 17(6), 520–525.
- Wong, Q.Y.A. and Chew, F.T. (2024). The association between migration and prevalence of allergic diseases: A systematic review

and meta-analysis. *International Archives of Allergy and Immunology*, 185(11), 1099–1122.

Note

Chapter written by Aleksei N. KURBATSKII and Marina V. MIRONENKOVA.

PART 3

Statistical Techniques

13

A Unified Model for SPC and Maintenance in High Investigation Cost Processes with Multiple Quality Disruptions

In modern manufacturing systems, integrating process quality control and equipment maintenance is crucial for maintaining operational efficiency and reliability. Traditional statistical process control (SPC) methods often initiate investigations in response to alarms, which can be impractical in environments with high investigation costs or frequent disruptions. This study introduces a unified framework that integrates SPC with preventive maintenance to address these challenges. Designed for manufacturing environments characterized by high investigation costs and multiple quality shifts, this framework leverages quality monitoring data to guide both quality policy and maintenance decisions. A probabilistic decision-making model dynamically balances quality control and maintenance efforts, optimizing cost efficiency and reducing downtime. A detailed numerical analysis demonstrates the model's effectiveness in improving cost efficiency, reducing downtimes and optimizing decision-making in complex industrial settings.

13.1. Introduction

In modern manufacturing, achieving consistent, high-quality production while balancing economic and sustainability goals demands advanced strategies for fault detection and maintenance. Traditional approaches struggle with this balance, as evidenced by persistent issues of unplanned downtime, safety risks and quality losses stemming from reactive maintenance policies and outdated monitoring methods (Cassady et al. [2000](#); Woodall and Faltin [2019](#)). These challenges are compounded by two key limitations in current practice.

First, while adaptive maintenance strategies and SPC have advanced significantly as separate domains, their integration remains underdeveloped. Although researchers recognize the fundamental relationship between equipment conditions and product quality, most studies still examine these aspects separately. Tracing the evolution of

research in this area, Tagaras (1988) and Ben-Daya (1999) established early SPC-maintenance integration frameworks, while Cassady et al. (2000) demonstrated the potential of joint optimization over standalone approaches. Subsequent developments by Linderman et al. (2005) and Yeung et al. (2007) introduced adaptive quality control policies and economic optimization, respectively, yet critical gaps in handling real-world complexities still exist in the literature.

Second, conventional SPC methods assume that all alarms require investigation and that diagnoses are always accurate. However, in practice, investigations are expensive, time consuming and prone to errors, especially when multiple potential causes must be explored. This underscores the need for more pragmatic, cost-aware approaches to quality control and maintenance that:

1. prioritize alarms based on severity and process deterioration levels;
2. isolate likely failure zones among multiple causes;
3. recommend targeted replacements to enhance restoration efficiency;
4. support optimized maintenance planning;
5. enable economically sound inspection and maintenance decisions based on the actual process conditions.

To address these limitations, this study proposes a unified framework integrating

1. action-driven SPC with prioritized alarms and multizone control limits,
2. condition-based corrective maintenance (CM) for targeted restoration and
3. age-based PM optimized via a Markov decision process. By dynamically balancing monitoring and maintenance, the proposed model minimizes operational costs while reducing false alarms and unplanned downtime. The policy optimizes key parameters, including the sampling size, action limits and PM interval, through a probabilistic cost model validated via numerical analysis.

13.2. Problem statement

Let us consider a continuously operating manufacturing process where product quality is characterized by a critical-to-quality variable X , modeled as a normally distributed random variable $X \sim N(\mu_0, \sigma_0^2)$.

Initially, the process operates in an optimal in-control (IC) state, with the mean and standard deviation conforming to target specifications.

However, the process is subject to two assignable causes that induce shifts in the process mean, leading the process to an out-of-control (OOC) operating state (OOC₁, OOC₂). The assignable cause 1 shifts the process mean by a multiple d_1 of the standard deviation, and the assignable cause 2 by d_2 ($d_2 > d_1 > 0$) :

$$\mu_1 = \mu_0 + d_1 \cdot \sigma_0 \quad [13.1]$$

$$\mu_2 = \mu_0 + d_2 \cdot \sigma_0 \quad [13.2]$$

The occurrence of assignable cause 1 does not prevent the occurrence of assignable cause 2, which further deteriorates the process performance.

To monitor and manage process quality shifts, an $\bar{X}$ control chart is selectively implemented when justified by process evaluation. When SPC is activated, sampling and charting occur at discrete time points. Let $s_i = \{0, 1\}$ be a binary decision variable for each time period i following maintenance, where $s_i = 1$ indicates that sampling is performed at period i , while $s_i = 0$ is not. The complete SPC strategy is defined by vector $S = (s_1, s_2, \dots, s_{\tau-1})$.

The control chart operates as follows: when sampling is activated ($s_i = 1$), a sample of n products is measured, and the sample mean is calculated. The control chart employs two action limits ($k_1 < k_2$) that define three decision zones: (1) if $\bar{x} < k_1$, the process remains IC with no action needed; (2)

$k_1 \leq \bar{x} < k_2$, the process is OOC₁, and restoration action A is triggered to address assignable cause 1; (3) if $\bar{x} \geq k_2$, the more severe OOC₂ condition is identified, requiring restoration action B. This two-threshold design directly links specific OOC conditions to their corresponding CM actions, enabling targeted responses to different process shift magnitudes.

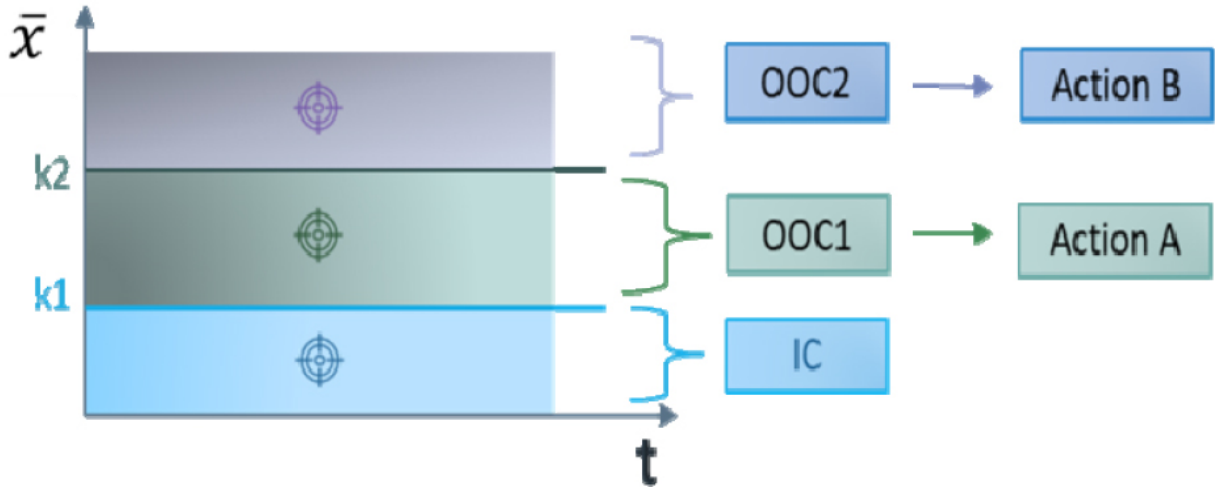


Figure 13.1. *Proposed SPC approach control chart*

The PM policy follows an age-replacement strategy with threshold τ : if no failure is detected within τ consecutive periods, the process is halted for PM.

Unlike traditional age-replacement models, where failures are immediately observable, our framework accounts for latent failures that require shutdown and investigation for verification. Here, τ represents the maximum allowable operational age before mandatory renewal, ensuring proactive intervention even in the absence of SPC alarms.

13.3. Mathematical model

13.3.1. Failure probabilities

Let i represent the number of time intervals elapsed since the last maintenance action (CM or PM). If the process operates IC, three possible failure scenarios may occur: direct transition to OOC₁, direct transition to OOC₂ and progressive deterioration, that is, a shift to OOC₁ and OOC₂:

$$\text{IC} \rightarrow \text{OOC}_1 \quad P_{1_01}(i) = p_{01} + (1 - p_{01}) \cdot (1 - \delta_{01}^{i-1}), \quad i=1,2,\dots \quad [13.3]$$

$$\text{IC} \rightarrow \text{OOC}_2 \quad P_{2_02}(i) = p_{02} + (1 - p_{02}) \cdot (1 - \delta_{02}^{i-1}), \quad i=1,2,\dots \quad [13.4]$$

$$\text{OOC}_1 \rightarrow \text{OOC}_2 \quad P_{2_12}(i) = p_{12} + (1 - p_{12}) \cdot (1 - \delta_{12}^{i-1}), i=1,2,\dots \quad [13.5]$$

where p_{01} , p_{02} and p_{12} are the initial probabilities of failure and δ_{01} , δ_{02} and δ_{12} are constants that reflect the gradual failure rates ($0 < \delta_{01}, \delta_{02}, \delta_{12} < 1$).

Likewise, the probability that no shift occurs equals:

$$q(i) = 1 - (P_{1_01}(i) + P_{2_02}(i)) \quad [13.6]$$

13.3.2. Error probabilities

The sampling-based control chart is subject to three types of statistical decision errors: (i) false alarm: incorrect OOC signal when the process is IC; (ii) missed detection: failure to signal when the process is OOC; and (iii) misclassification error: correct OOC signal but wrong corrective action. Let $\Phi(\cdot)$ denote the standard normal random CDF. The probabilities of these scenarios are:

$$\text{Process: IC, Action: N/A} \quad b_1 = \Phi(k_1) \quad [13.7]$$

$$\text{Process: IC, Action: A} \quad b_2 = \Phi(k_2) - \Phi(k_1) \quad [13.8]$$

$$\text{Process: IC, Action: B} \quad b_3 = 1 - \Phi(k_2) \quad [13.9]$$

$$\text{Process: OOC}_1, \text{ Action: N/A} \quad b_4 = \Phi(k_1 - d_1 \cdot \sqrt{n}) \quad [13.10]$$

$$\text{Process: OOC}_1, \text{ Action: A} \quad b_5 = \Phi(k_2 - d_1 \cdot \sqrt{n}) - \Phi(k_1 - d_1 \cdot \sqrt{n}) \quad [13.11]$$

$$\text{Process: OOC}_1, \text{ Action: B} \quad b_6 = 1 - \Phi(k_2 - d_1 \cdot \sqrt{n}) \quad [13.12]$$

$$\text{Process: OOC}_2, \text{ Action: N/A} \quad b_7 = \Phi(k_1 - d_2 \cdot \sqrt{n}) \quad [13.13]$$

$$\text{Process: OOC}_2, \text{ Action: A} \quad b_8 = \Phi(k_2 - d_2 \cdot \sqrt{n}) - \Phi(k_1 - d_2 \cdot \sqrt{n}) \quad [13.14]$$

$$\text{Process: OOC}_2, \text{ Action: B} \quad b_9 = 1 - \Phi(k_2 - d_2 \cdot \sqrt{n}) \quad [13.15]$$

13.3.3. Transition and steady-state limiting probabilities

The system's stochastic behavior is characterized by two probability classes: transition probabilities, associated with transitions between states and limiting probabilities, which represent the long-run probability of the process operating at each state.

$$- (i, \text{IC}) \text{ to } (i+1, \text{IC}): \quad \mathbf{Pr}_{[16]} = s_i \cdot b_1 \cdot q(i) - (1 - s_i) \cdot q(i) \quad [13.16]$$

$$- (i, \text{IC}) \text{ to } (1, \text{IC}): \quad \mathbf{Pr}_{[17]} = s_i \cdot (1 - b_1) \cdot q(1) \quad [13.17]$$

$$- (i, \text{IC}) \text{ to } (i+1, \text{OOC}_1): \quad \mathbf{Pr}_{[18]} = s_i \cdot (b_4 + b_6) \cdot P_{1_01}(i) + (1 - s_i) \cdot P_{1_01}(i) \quad [13.18]$$

$$- (i, \text{IC}) \text{ to } (i+1, \text{OOC}_2): \quad \mathbf{Pr}_{[19]} = s_i \cdot (b_7 + b_8) \cdot (P_{2_02}(i) + P_{2_12}(i)) + (1 - s_i) \cdot (P_{2_02}(i) + P_{2_12}(i)) \quad [13.19]$$

$$- (i, \text{IC}) \text{ to } (1, \text{OOC}_1): \quad \mathbf{Pr}_{[20]} = s_i \cdot b_6 \cdot P_{1_01}(i) \quad [13.20]$$

$$- (i, \text{IC}) \text{ to } (1, \text{OOC}_2): \quad \mathbf{Pr}_{[21]} = s_i \cdot b_8 \cdot (P_{1_01}(i) + P_{2_02}(i) + P_{2_12}(i)) \quad [13.21]$$

$$- (i, \text{OOC}_1) \text{ to } (i+1, \text{IC}): \quad \mathbf{Pr}_{[22]} = s_i \cdot b_5 \cdot q(1) \quad [13.22]$$

$$- (i, \text{OOC}_2) \text{ to } (i+1, \text{IC}): \quad \mathbf{Pr}_{[23]} = s_i \cdot b_9 \cdot q(1) \quad [13.23]$$

$$- (i, \text{OOC}_1) \text{ to } (i+1, \text{OOC}_1): \quad \mathbf{Pr}_{[24]} = s_i \cdot b_4 + (1 - s_i) \quad [13.24]$$

$$- (i, \text{OOC}_2) \text{ to } (i+1, \text{OOC}_2): \quad \mathbf{Pr}_{[25]} = s_i \cdot b_7 + (1 - s_i) \quad [13.25]$$

$$- (i, \text{OOC}_1) \text{ to } (i+1, \text{OOC}_2): \quad \mathbf{Pr}_{[26]} = s_i \cdot (b_4 + b_7) + (1 - s_i) \quad [13.26]$$

$$- (i, \text{OOC}_1) \text{ to } (1, \text{OOC}_1): \quad \mathbf{Pr}_{[27]} = s_i \cdot b_6 \cdot P_{1_01}(i) \quad [13.27]$$

$$- (i, \text{OOC}_2) \text{ to } (1, \text{OOC}_2): \quad \mathbf{Pr}_{[28]} = s_i \cdot b_8 \cdot P_{2_02}(i) \quad [13.28]$$

$$- (i, \text{OOC}_1) \text{ to } (1, \text{OOC}_2): \quad \mathbf{Pr}_{[29]} = s_i \cdot b_6 \cdot P_{2_12}(i) \quad [13.29]$$

$$- (\tau-1, \text{IC}) \text{ to } \tau: \quad \mathbf{Pr}_{[30]} = s_{\tau-1} \cdot b_1 + (1 - s_{\tau-1}) \quad [13.30]$$

$$- (\tau-1, \text{OOC}_1) \text{ to } \tau: \quad \mathbf{Pr}_{[31]} = s_{\tau-1} \cdot b_4 + (1 - s_{\tau-1}) \quad [13.31]$$

$$- (\tau-1, \text{OOC}_2) \text{ to } \tau: \quad \mathbf{Pr}_{[32]} = s_{\tau-1} \cdot b_7 + (1 - s_{\tau-1}) \quad [13.32]$$

The steady-state probabilities characterize long-term system behavior: $\pi_{(i, \text{IC})}$ represents the probability of operating in IC at period i , while $\pi_{(i, \text{OOC}_1)}$ and $\pi_{(i, \text{OOC}_2)}$ denote OOC probabilities induced by assignable causes 1 and 2, respectively. The PM triggering state occurs with probability $\pi_{(\tau)}$, where the age threshold is reached. For operational states, the action space reduces to binary decisions, either “do nothing” ($s_i = 0$) or “initiate sampling” ($s_i = 1$). These probabilities are derived by solving the linear system defined by the transition matrix (Nenes et al. [2015](#)).

We formulate the proposed model as a discrete-time Partially Observable Markov Decision Process (POMDP) with $3\tau-2$ distinct states. The state space includes operational quality states (i, IC) , (i, OOC_1) and (i, OOC_2) for each time period $i = 1, 2, \dots, \tau - 1$ since last maintenance, plus the PM triggering state τ . State transitions follow the structured transition probability matrix shown in [Table 13.1](#), where non-zero entries advance

time indices for valid state progressions, while zeros represent invalid transitions. Time-dependent failure probabilities (increasing with equipment age i) create non-stationary dynamics, with sampling decisions $s_i \in \{0, 1\}$ influencing each transition. All movements obey equations [13.16]–[13.32] probabilities, terminating at maintenance state τ when $i = \tau - 1$. This POMDP structure optimally balances sampling costs against maintenance actions while managing quality state uncertainty

13.4. Cost structure analysis

To develop an economically optimal unified policy, we define a cost structure comprising four components: (1) per-unit inspection cost c_s , (2) PM cost $c_{m,IC}$, (3) CM costs c_{m,OOC_1} and c_{m,OOC_2} for OOC_1 and OOC_2 states, respectively, and (4) poor quality cost c_q . These costs follow a logical hierarchy reflecting real-world applications, where prevention is cheaper than correction, major failures are more expensive to remove than minor ones and quality losses outweigh intervention costs:

$$0 < c_s < c_{m,IC} < c_{m,OOC_1} < c_{m,OOC_2} < c_q \quad [13.33]$$

This structure enables a precise trade-off analysis between monitoring effort, maintenance timing and quality performance. When the process remains IC, sampling activities ($s_i = 1$) incur two key expenses: (1) the sampling cost for quality testing and (2) potential false alarm costs in the case of an OOC_1 or OOC_2 signal when the process is actually in control ($c_{m,IC}$). Consequently, the expected cost of the process operating IC at different periods, that is, (1, IC), (2, IC), ..., ($\tau - 1$, IC), equals:

$$C_{IC} = s_i \cdot [c_s \cdot n + c_{m,IC} \cdot (1 - b_1)] \quad [13.34]$$

Table 13.1. *Transition probability matrix*

	1, IC	1, OOC ₁	1, OOC ₂	2, IC	2, OOC ₁	2, OOC ₂	...	$\tau - 1, IC$	$\tau - 1, OOC_1$	$\tau - 1, OOC_2$	τ
1, IC	Pr _[17]	Pr _[20]	Pr _[21]	Pr _[16]	Pr _[18]	Pr _[19]	...	0	0	0	0
1, OOC ₁	Pr _[22]	Pr _[27]	Pr _[28]	0	Pr _[24]	Pr _[25]	...	0	0	0	0
1, OOC ₂	Pr _[23]	0	Pr _[29]	0	0	Pr _[26]	...	0	0	0	0
2, IC	Pr _[17]	Pr _[20]	Pr _[21]	0	0	0	...	0	0	0	0
2, OOC ₁	Pr _[22]	Pr _[27]	Pr _[28]	0	0	0	...	0	0	0	0
2, OOC ₂	Pr _[23]	0	Pr _[29]	0	0	0	...	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...
$\tau - 1, IC$	Pr _[17]	Pr _[20]	Pr _[21]	0	0	0		0	0	0	Pr _[30]
$\tau - 1, OOC_1$	Pr _[22]	Pr _[27]	Pr _[28]	0	0	0		0	0	0	Pr _[31]
$\tau - 1, OOC_2$	Pr _[23]	0	Pr _[29]	0	0	0		0	0	0	Pr _[32]
τ	$q(1)$	$P_{1_01}(1)$	$P_{2_02}(1)$	0	0	0		0	0	0	0

Conversely, OOC operation introduces CM costs and poor-quality costs. In particular, the expected cost of OOC operation at different time periods is given by:

$$C_{OOC} = s_i \cdot \left[c_s \cdot n + c_{m, OOC_1} \cdot b_5 + c_{m, OOC_2} \cdot b_9 + c_q \cdot (b_4 + b_6) \right] + (1 - s_i) \cdot c_q \quad [13.35]$$

The unified cost model combines all operational components into a single long-run expected cost per period, accounting for state-dependent costs, failure probabilities and steady-state probabilities, and is computed by:

$$\begin{aligned} \gamma(S, n, k_1, k_2, \tau) = & \sum_{i=1}^{t-1} C_{IC} \cdot \pi_{(i, IC)} + \sum_{i=1}^{t-1} \sum_{y=1}^2 C_{OOC} \cdot \pi_{(i, OOC_y)} + \\ & + \pi_{(\tau)} \cdot \left[P_{1_01}(\tau) \cdot c_{m, OOC_1} + P_{2_02}(\tau) \cdot c_{m, OOC_2} \right] \end{aligned} \quad [13.36]$$

13.5. Optimization framework

The proposed optimization procedure aims to minimize the long-run expected cost by jointly determining the optimal parameters for SPC and PM. The decision variables include the sampling strategy

$S = (s_1, s_2, \dots, s_{\tau-1})$, sample size n , action limits (k_1, k_2) , and PM age threshold τ , while treating cost coefficients and process shift parameters as fixed inputs.

The solution algorithm employs an iterative search process: beginning with $\tau = 2$, it evaluates the long-run cost function $C(S, n, k_1, k_2, \tau)$ (equation [13.36]) across feasible ranges for each τ . At each iteration, the transition probability matrix is constructed, steady-state probabilities are calculated and the corresponding long-term cost is evaluated. A full enumeration approach is used to identify the cost-minimizing configuration for the given τ .

Due to the cost function's convexity with respect to τ , the procedure terminates when an increase in cost is detected, ensuring that the identified solution is globally optimal:

$$C(S, n, k_1, k_2, \tau + 1) > C(S, n, k_1, k_2, \tau) \quad [13.37]$$

The optimization framework was implemented via a custom MATLAB algorithm that systematically evaluates the long-run cost function across the parameter space.

13.6. Numerical analysis

The model's performance is evaluated through an extensive set of benchmark scenarios. The benchmarking systematically varied all cost parameters to assess their impact on optimal policy selection. The results demonstrated consistent policy patterns: lower sampling costs favored frequent monitoring (higher n , active S) and extended PM intervals (τ),

while high-quality risks resulted in tighter action limits (k_1, k_2) and shorter PM intervals. [Table 13.2](#) presents four representative cases from our analysis demonstrating the model's behavior across different operational conditions. As the cost function is characterized by convexity with respect to τ , the procedure terminates when an increase in cost is detected, ensuring that the identified solution is globally optimal. To illustrate this behavior, an additional line item has been included in [Table 13.2](#): following the optimal cost (C_τ) achieved at the optimal τ policy, the cost for $\tau + 1$ is also calculated and reported ($C_{\tau+1}$), highlighting the increasing trend.

Table 13.2. *Set parameters and results*

		Case 1	Case 2	Case 3	Case 4
Input parameters	c_s	1	25	40	5
	$c_{m, IC}$	50	50	50	400
	c_{m, OOC_1}	100	100	150	500
	c_{m, OOC_2}	150	150	600	600
	c_q	200	200	700	700
Decision variables	τ	7	4	3	7
	n	10	5	5	15
	k_1	1	0.5	0.5	1.5
	k_2	2	1	4	2.5
	S	[1,1,1,1,1,1]	[0,0,0]	[0,1]	[1,1,1,1,1,1]
Objective function	C_τ	25.305	85.577	222.210	132.126
	$C_{\tau+1}$	25.403	100.941	225.209	133.100
Optimal	$C_{optimal}$	25.305	85.577	222.210	132.126

In Case 1, characterized by relatively low-cost parameters, the optimal policy favors an extended preventive maintenance interval ($\tau = 7$) combined with a frequent monitoring strategy ($S = [1,1,1,1,1,1]$). The low cost per sample also enables the use of a higher number of samples ($n = 10$) at each inspection point, enhancing the monitoring precision without a substantial increase in overall cost. This configuration leads to the lowest total cost among the evaluated scenarios. In Case 2, where sampling and quality-related costs are higher, the model responds by

shortening the optimal preventive maintenance interval ($\tau = 4$) and tightening the action limits k_1 and k_2 . This adjustment indicates a more precautionary approach, where PM is implemented more frequently to avoid the potentially high costs associated with OOC states. Case 3 represents a case where quality-related costs, particularly those associated with major OOC conditions and poor quality, are relatively high. In response, the model adopts a more aggressive strategy by selecting a shorter optimal PM interval ($\tau = 3$) and implementing stricter action limits to minimize the likelihood of costly deviations. Although the sampling cost is relatively high, it remains lower than the associated maintenance costs. As a result, the model balances the overall cost by recommending a conservative sampling strategy, initially skipping sampling and testing, followed by selective implementation at the next stage, to effectively mitigate the risk of entering and remaining in an expensive OOC state. Finally, Case 4 represents a case with high maintenance costs. In response, the model selects a longer optimal preventive maintenance interval ($\tau = 7$) to reduce the frequency of costly interventions. Given that the sampling cost is set on the lower side, the algorithm recommends an active sampling policy involving testing at each control point ($S = [1,1,1,1,1,1]$) and a larger sample size ($n = 15$). The use of larger sample sizes serves to enhance detection accuracy, thereby supporting cost optimization by reducing the likelihood of false alarms and minimizing the risk of entering and remaining in an expensive OOC state.

[Table 13.3](#) provides a breakdown of the cost evaluation and decision variable integration using the parameters from Case 1. For each iteration from $\tau = 1$ up to the optimal PM interval, the total cost and the corresponding decision variables are presented.

The optimization process reveals several key insights. Beginning with $\tau = 1$, where PM occurs every period at a fixed cost of \$120 without any SPC implementation, we observe the upper bound of operational expenses. As the algorithm explores longer intervals, costs decrease significantly, dropping to \$68.302 at $\tau = 2$ and further to \$41.101 at $\tau = 3$, where SPC first becomes economically viable. This trend continues through $\tau = 4$ (\$36.664) and subsequent intervals, demonstrating how extended maintenance windows coupled with strategic SPC activation can reduce costs. The optimal configuration emerges at $\tau = 7$, achieving the minimum expected cost through full SPC implementation at all periods. Beyond this point, at $\tau = 8$, costs begin to rise, confirming $\tau = 7$ as the global optimum due to the function's convexity.

Table 13.3. Numerical example: breakdown of cost and decision variables (Case 1)

Cost (\$)	τ	k_1	k_2	n	s_1	s_2	s_3	s_4	s_5	s_6	s_7
120	1	—	—	—	—	—	—	—	—	—	—
68.302	2	—	—	—	0	—	—	—	—	—	—
41.101	3	0.5	1	5	1	1	—	—	—	—	—
36.664	4	0.5	1	5	1	1	1	—	—	—	—
27.484	5	1	2	10	1	1	1	1	—	—	—
25.503	6	1	2	10	1	1	1	1	1	—	—
25.305	7	1	2	10	1	1	1	1	1	1	—
25.403	8	1	2.5	10	1	1	1	1	1	1	1

Our analysis has demonstrated that the action limits significantly influence the cost function. To illustrate this relationship, we examine Cases 1 and 4 through [Figure 13.2](#). The upper contour plot illustrates how combinations of k_1 and k_2 affect total costs, with a pronounced global minimum (marked in red) validating our optimization methodology. This visual confirmation shows how specific action limit pairs minimize long-run cost. This finding is further supported by the sensitivity analysis shown in the right graphs, which reveals a well-defined cost minimum, demonstrating both significant cost savings and the robustness of the optimal solution.

13.7. Conclusion

This study extends traditional SPC methodologies by developing a unified quality-maintenance framework that addresses critical limitations in high-cost manufacturing environments. By introducing (1) multizone control charts for differentiated OOC responses, (2) condition-based maintenance triggers and (3) cost optimization, we demonstrate consistent cost reductions across various scenarios. The proposed Markov decision process model balances the trade-off between monitoring intensity and maintenance frequency. For industrial practitioners, this work provides three key insights. First, the action-limit control chart design significantly reduces unnecessary investigations in high-cost environments. Second, the dynamic maintenance policy prevents both excessive PM frequency and delayed interventions. Third, the sampling strategy optimization proves that variable-frequency monitoring often outperforms fixed-

interval approaches when investigation costs are considered. Future research should explore two promising extensions of this work: (a) expanding the cost function to include energy consumption and material waste metrics would align the system with circular economic goals; (b) machine learning could strengthen root cause analysis by processing equipment operational data (vibration, temperature, power consumption patterns) to predict quality shifts.

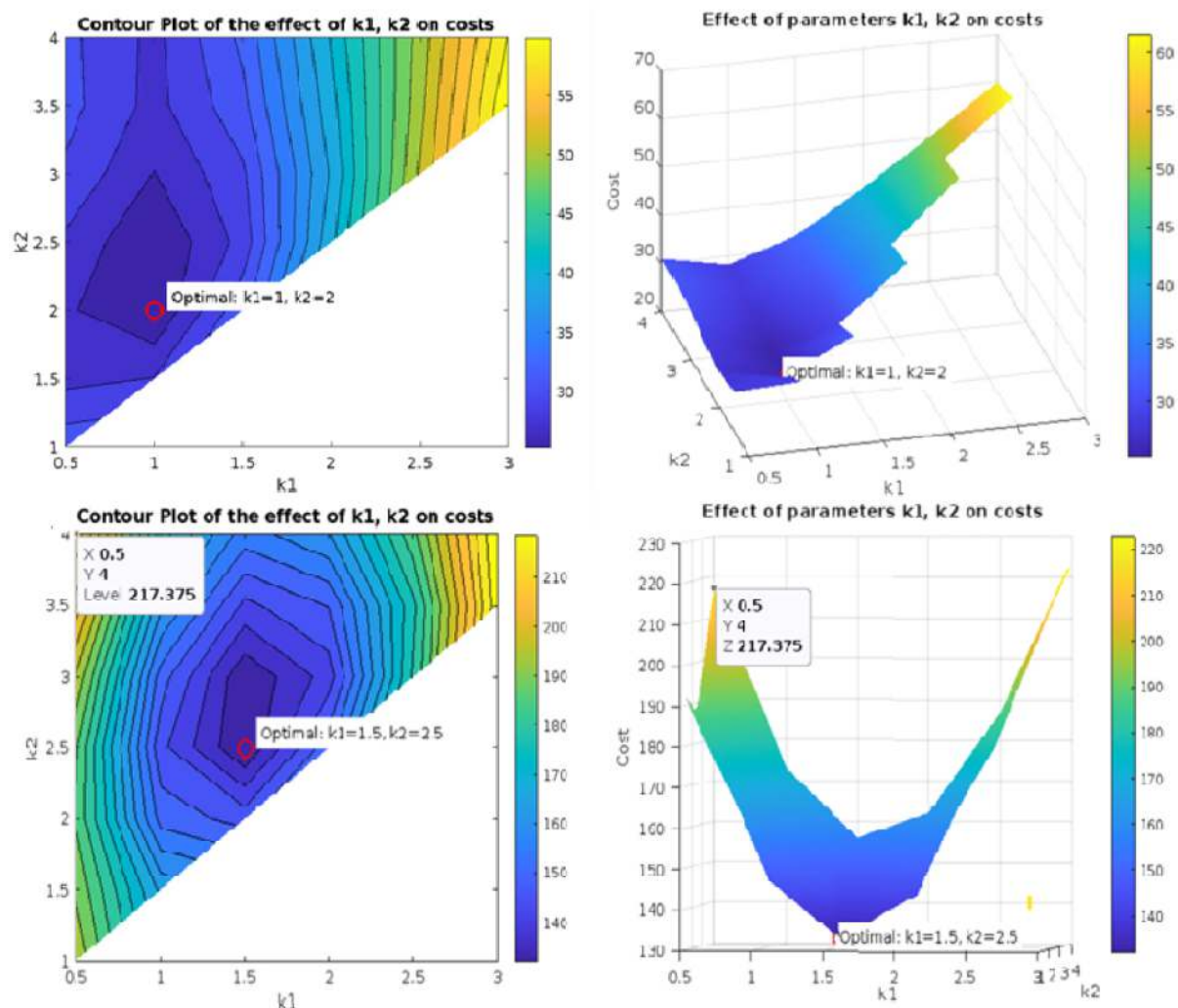


Figure 13.2. *Optimized action limits and cost sensitivity analysis.*

13.8. References

Ben-Daya, M. (1999). Integrated production maintenance and quality model for imperfect processes. *IIE Transactions*, 31, 491–501.

- Cassady, C.R., Bowden, R.O., Liew, L., Pohl, E.A. (2000). Combining preventive maintenance and statistical process control: A preliminary investigation. *IIE Transactions*, 32, 471–478.
- Linderman, K., McKone-Sweet, K.E., Anderson, J.C. (2005). An integrated systems approach to process control and maintenance. *European Journal of Operations Research*, 164, 324–340.
- Nenes, G., Tasias, K.A., Celano, G. (2015). A general model for the economic-statistical design of adaptive control charts for processes subject to multiple assignable causes. *International Journal of Production Research*, 53(7), 2146–2164.
- Tagaras, G. (1988). An integrated cost model for the joint optimization of process control and maintenance. *Journal of the Operational Research Society*, 39(8), 757–766.
- Woodall, W.H. and Faltin, F.W. (2019). Rethinking control chart design and evaluation. *Quality Engineering*, 31(4), 596–605.
- Yeung, T.G., Cassady, C.R., Schneider, K. (2007). Simultaneous optimization of control chart $\bar{X}$ and age-based preventive maintenance policies under an economic objective. *IIE Transactions*, 40(2), 147–159.

Note

Chapter written by Konstantina TSIOTA and Konstantinos A. TASIAS.

14

A Note on Spectral Risk Measures when Systemic Risk is Present

This chapter is devoted to systemic risk (from now on, abbreviated as SR) and its potential depiction in the risk spectrum of spectral risk measures. At first, we propose a fundamental way to quantify the existence of SR. Second, we argue with common practice that suggests that the risk spectrum should solely portray the utility function of the investor–regulator and thus be constructed only according to it. In addition, we present and justify two conditions that the risk spectrum should satisfy in regard of the existence of SR; and, to our knowledge, such conditions have not been suggested by another academic or a practitioner. Moreover, we call the spectral risk measures that have a risk spectrum which satisfies those conditions ‘systemic risk aware’ and we discuss their practicality and applicability.

14.1. Introduction and motivation

Spectral risk measures are considered of high interest in academia. It is widely accepted that the so-called risk spectrum can depict the level of risk aversion of the investor and notable efforts have taken place regarding the portrayal of that factor. For instance, Dowd et al. ([2008](#)) and Leopold ([2015](#)) mentioned some very important classes of risk measures, like the exponential or the power spectral risk measures. Similar to the above is the Wang transform, suggested by Wang ([1996](#), [2000](#)). Apart from the academic interest, there is also the practical aspect which is just as important. As BCBS ([2012](#)) suggests, spectral risk measures are a hopeful generalization of expected shortfall, implying that it may be utilized in the near future.

In addition, there is criticism on the theoretical grounds of spectral risk measures (see, for instance, Balbas et al. ([2009](#))), namely, the properties of completeness (necessary condition), exhaustion and finally, adaptability (necessary and sufficient condition). Apparently,

exponential and power spectral risk measure does not exhibit all of those properties. Also Embrechts et al. ([2014](#)) denoted the robust backtesting problems of spectral risk measures as they are not elicitable. On the other hand, those classes are very easy to construct and the theoretical background leaves room for limitless options, as long as the risk preferences of an investor can be exposed.

We, on the other hand, are motivated by the fact that the utility function of the investor – a concept related to the asset pricing theory (see Cochrane [2000](#); Lofthouse [2000](#)) – and portfolio diversification (see Markowitz [1952](#)) are the factors that determine the risk spectrum of the measure. We argue with this practice and propose that systemic risk should also be included in the depiction. Recall that systemic risk refers to the instability of a financial system that can lead to its entire collapse (Kaufman [2000](#)). In addition, there is academic interest for the implications of systematic risk. For instance, Asimit and Li ([2017](#)) states that the main purpose concerning systemic risk is to evaluate the financial distress of an economy as a consequence of the failure of one of its components. They also point out the importance of the extreme value theory (EVT) in the analysis of systemic risk. In addition, it should be mentioned that their study is based on the framework proposed in Acharya et al. ([2012](#), [2017](#)) in order to capture systemic risk. Moreover, Luciano and Wihlborg ([2015](#)) discussed that the financial institutions can generate systemic risk. Also Acemoglu et al. ([2015](#)) addressed the same topic. In addition, Brownlees and Engle ([2006](#)) introduce SRISK to measure the capital shortfall of a firm conditional on a severe market decline. An interesting comparison would be between our approach and the one of Champervorne ([1953](#)). The author utilizes the income distribution in order to capture how the stochastic process of the income share is taking place among certain groups. We, on the other hand, use the returns' distribution of a certain investment in order to highlight how cautious an investor needs to be due to the risk of the economic environment.

Having the practitioners in mind, we attempt to give a different perspective on how risk spectrum should be comprehended. Rephrasing, instead of attempting to picture an investor or a managers' aversion profile, which by definition is rather subjective,

we suggest the degree of aversion they should have depending on the markets' underlying risk. Moreover, we attempt to suggest not only a theoretically robust, but also an applicable, method.

Furthermore, we are motivated by the fact that situations like the global crisis emerging from late 2007, as a subprime mortgage crisis in United States and LIBOR scandal, triggered the need to reassess the tools a professional or a regulator should employ (see, for instance, Loebnitz and Roorda ([2011](#)) and Loebnitz's ([2011](#)) work on liquidity risk, Lujun et al.'s ([2017](#)) work on worst case scenario and Asimit and Li's ([2017](#)) work on systemic risk). This reassessment is still taking place and a milestone example is the work of BCBS ([2013](#)), as it suggests the expected shortfall instead of the value at risk as a standard risk measure.

Unfortunately, there is criticism on the regulatory logic of using a unique risk assessment tool in order to have a homogeneous benchmark, see, for instance, Embrechts et al. ([2014](#)), as it lacks the notion that every market exhibits certain characteristics that influence its underlying investment risk, which cannot be diversified away. As Acerbi ([2002](#)) suggests, different portfolios' risk can be better detected by different risk measures and a unique risk measure can be a risk generator itself. A notable attempt to personalize the risk measure is the specific risk – a component added to value at risk, due to issuer's specific price movement (McNeil et al. [2005](#)). Our point of view is in compliance with this criticism. Moreover, in this chapter, we denote a way to categorize each market according to the systemic risk it bears. Also, we propose a way to quantify systemic risk via macroeconomic factors as indicators. Consequently, each market will utilize a risk measure that corresponds to its systemic risk.

We mention that this study is in tandem with our previous studies in risk (see, for instance, Konstantinides and Zachos ([2019](#))). The structure of the chapter is the following: the remaining of introduction includes an example that gives further insight into the motivation. Section [14.2](#) contains a literature review on the basic concepts of risk measures and spectral risk measures that we are utilizing. Section [14.3](#) is devoted to the theoretical framework, suggested by Acerbi ([2002](#)). Section [14.4](#) includes our contribution to

the theoretical framework. Section [14.5](#) provides an application that can be empirically validated and finally section [14.6](#) concludes the chapter.

14.1.1. *Motivation*

We will now give a theoretical, motivational example. Consider two markets M_1 , M_2 , and three elements that are responsible for the systemic risk $\{r_1, r_2, r_3\}$. Also, consider that those elements are associated with fundamental macroeconomic factors, for instance, r_1 with the real gross domestic product YoY change in %, r_2 with the unemployment rate and r_3 with the inflation rate (according to Mankiw ([2009](#)), GDP measures everyone's income in an economy, unemployment rate is associated with the fraction of labor capacity that is out of work and inflation measures how fast prices are rising).

We consider that there is a threshold corresponding to each element that the economies acknowledge. Exceeding a threshold for an economy signifies an increase in systemic risk. Let us set that the thresholds are $r_1 = -10\%$, $r_2 = 20\%$ and $r_3 = 5\%$. Moreover, the macroeconomic factors for M_1 , M_2 and for a three-year period are the following:

	M_1		
Year	r_1	r_2	r_3
2015	+5%	1%	3%
2016	+1%	5%	6%
2017	-3%	4%	4%.

	M_2		
Year	r_1	r_2	r_3
2015	+1%	15%	6%
2016	-11%	22%	4%
2017	-12%	21%	3%.

Bear in mind that the elements we have chosen are not at all restrictive. On the contrary, we can choose any element that is associated with the robustness of a market. For instance, the income (see Case and Fair ([2007](#))) for details) which is also of academic interest (see Champernowne [1953](#)) can prove quite useful. We follow the analytical approach of Young ([1993](#)) where we can present for each year the three elements as an ordered triplet of real numbers in a three-dimensional cartesian coordinate system. So, $[r_1, r_2, r_3]$ is a vector for each year and for each market with r_1, r_2, r_3 as the components of the vector. Under that framework, we can represent the macroeconomic factors for each market M_1 and M_2 for 2015–2017 in 3×3 matrices. Specifically:

$$M_1 = \begin{bmatrix} 5 & 1 & 3 \\ 1 & 5 & 6 \\ -3 & 4 & 4 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} 1 & 15 & 6 \\ -11 & 22 & 4 \\ -12 & 21 & 3 \end{bmatrix}$$

where the rows correspond to each year and the columns correspond to each macroeconomic factor. Obviously, there are indications that M_2 bears more systemic risk for all three years compared to M_1 (M_2 has one element that exceeds the threshold in 2015, two in 2016 and two in 2017, while M_1 has one element that exceeds a benchmark only in 2016). A mathematical approach would be to calculate the magnitude of each years' vector (see Young [1993](#), definition 2), then we get:

Vectors' magnitudes		
Year	$ M_1 $	$ M_2 $
2015	5.92	29.61
2016	7.87	24.92

The magnitudes of the vectors representing M_2 are in all occasions much higher than M_1 , which denotes a basic ordering in terms of systemic risk (rephrasing, we should consider that M_2 is more exposed to systemic risk). Since we are having in our example square matrices, we may also utilize the determinants, where $\det M_1 = 15$ and $\det M_2 = 135$. Consider a geometrical approach for the vectors r_1, r_2, r_3 , and assume that they are linearly independent (see Margalit [2018](#), theorem p.216). Then, the absolute value of determinants equals the volume of the parallelepiped that is determined by vectors r_1, r_2, r_3 . Clearly, the value of M_2 parallelepiped is much greater than the value of M_1 , which also denotes in terms of a basic ordering that M_2 is exposed to more systemic risk.

Now, consider a risk-averse investor who is investing to both M_1 and M_2 . Moreover, this investor utilizes a spectral risk measure to dictate his investing attitude. Consequently, the risk spectrum of that measure portrays the aversion toward risk. Under that framework, we propose that apart from the risk aversion of the investor, the risk spectrum should also depict the systemic risk that the market is exposed to. It is obvious that the investor's attitude toward risk needs to be more conservative in case they choose to invest in M_2 (compared to M_1). For instance, if the investor is planning to invest in 2018, then they should take into account the macroeconomic factors of 2017, which suggest that M_2 bears more systemic risk compare to M_1 (two macroeconomic factors exceed the benchmark in M_2 and none in M_1).

With the use of this setting, we observe that the depiction of systemic risk is of major importance. In addition, this depiction makes the risk spectrum aware of the systemic risk underlined in a market. Equally important is the fact that an investor or regulator is able to quantify the systemic risk, as for each macroeconomic factor that exceeds a threshold, the more risk averse the investor will be (more in section [14.4](#)). With this simple example, we provide an insight of

our proposition but in order to mathematically formalize our work, we need to demonstrate the theoretical foundation we utilize.

14.2. Preliminaries

14.2.1. Risk measures

Let Ω be a sample space. In addition, consider an investment over a single period of time, from 0 to T , where $X : \Omega \rightarrow \mathbb{R}$ is the money outcome of the investment. Moreover, consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G}$ is the set of real valued functions on Ω .

DEFINITION 14.1.– A risk measure is a function from the set $\mathcal{G}$ of risks X to the real numbers, $\rho : X \rightarrow \mathbb{R}$.

Furthermore, we take into account the future net worths that are accepted by an investor, namely, the acceptance set $\mathcal{A}$. According to Artzner et al. (1999, definition 2.3), an acceptance set $\mathcal{A}$ associated with a risk measure ρ is denoted by $\mathcal{A}_\rho$ and is defined by:

$$\mathcal{A}_\rho = \{X \in \mathcal{G} | \rho(X) \leq 0\}. \quad [14.1]$$

Moreover, in Artzner et al. (1999, definition 2.2), we find that given the total rate of return r of a reference instrument (for instance, we can consider r as the total return of an investment), the risk measure that is associated with an acceptance set is the mapping from $\mathcal{G}$ to the real numbers. It is denoted by $\rho_{\mathcal{A}, r}$ and is defined by:

$$\rho_{\mathcal{A}, r}(X) = \inf\{m \in \mathbb{R} | m \cdot r + X \in \mathcal{A}\}. \quad [14.2]$$

For the past 20 years up to now, the risk measure concept (based on measure theory) is used widely for regulatory purposes (see, for instance, BCBS (2012)). Such a measure can determine the amount of currency (or other assets) a financial institution should keep in reserve, depending on the financial risks to which it is exposed.

14.2.2. Properties of risk measures

Let us introduce properties of risk measures that are associated with our work:

1. Sub-additivity: $\rho(Z) + \rho(Y) \geq \rho(Z + Y)$ for any $Z, Y \in \mathcal{G}$.
2. Homogeneity: $\rho(\lambda Y) = \lambda \rho(Y)$ for any $\lambda \geq 0$.
3. Monotonicity: for any $Z, Y \in \mathcal{G}$ consider $Z \geq Y$ at all scenarios, then $\rho(Z) \leq \rho(Y)$.
4. Cash or translation invariance: for any $Z \in \mathcal{G}$ and consider M an investment with guaranteed risk free returns m , then $\rho(Z + M) = \rho(Z) - m$.
5. Convexity:

$$\rho[\lambda Y + (1 - \lambda)Z] \leq \lambda \rho(Y) + (1 - \lambda)\rho(Z) \text{ for any } Y, Z \in \mathcal{G} \text{ and any } \lambda \in [0, 1].$$

Interpretations of these properties can be found in Artzner et al. (1999); Follmer and Shied (2004, 2008).

DEFINITION 14.2.– A coherent risk measure is a risk measure satisfying monotonicity, translation invariance, homogeneity and sub-additivity.

The importance of coherency relies in the fact that when a risk measure exhibits them, then the acceptance set associated with this risk measure is closed and satisfies the four axioms suggested by Artzner et al. (1999, from Axiom 2.1 up to 2.4). An in-depth analysis concerning the relation between the axioms of acceptance sets and the properties of coherent risk measures can be found in Artzner et al. (1999, propositions 2.1, 2.2, 2.2 and 2.3) and their proofs.

14.2.3. Important risk measures

Let us introduce the value at risk (VaR_α).

DEFINITION 14.3.–

$$VaR_\alpha(X) = \sup\{x \in \mathbb{R} \mid \mathbb{P}(X \geq x) > 1 - \alpha\}.$$

We can describe the VaR_α as a threshold loss value of the investment, that it has probability $1 - \alpha$ to exceed this threshold in a given time horizon (assuming no additional trading). Important advantages of VaR_α include its straight computation while an important drawback is that it is not coherent as its sub-additivity is questionable and depends on the suggested setting. Briefly, we mention that sub-additivity holds when the investment returns are normally distributed so difficulties on that property may arise in cases of fat tails (see Danielsson et al. [2005](#)). On the other hand, a prime example of coherent risks measure is the expected shortfall (ES):

DEFINITION 14.4.– ES is the conditional expectation of loss, given that the loss is greater than the value at risk level. Mathematically, it can be expressed as $ES_\alpha = \frac{1}{\alpha} \int_0^\alpha VaR_p(Y) dp$ for any random variable Y and any real number $0 < \alpha < 1$.

For practical terms, we can perceive it as the expected value of an investment, conditional that the value is less than the quantile p of the worst outcomes. From Leopold ([2015](#)), we have that ES can be expressed in terms of the distribution of the investment's returns $F_X = \mathbb{P}(X \leq x)$:

$$ES_\alpha = -\frac{1}{\alpha} \int_0^\alpha F_X^\leftarrow(p) dp. \quad [14.3]$$

It is a coherent risk measure as it exhibits the four coherency properties (see Acerbi and Tasche [2002](#)). Also, it is conservative in comparison to VaR , in the sense that $ES_\alpha(Y) \geq VaR_\alpha(Y)$.

14.3. Theoretical framework

We rely on the theoretical framework of Acerbi ([2002](#)) for constructing the class of spectral risk measures. Consider that ρ_i are coherent risk measures and $i = 1, 2, 3, \dots, n$. If $\alpha_i \in \mathbb{R}^+$ and $\sum_i \alpha_i = 1$, then any convex combination $\rho = \sum_i \alpha_i \rho_i$ is also a coherent risk measure. Also, if we take that ρ_α represents a one-

parameter family of coherent risk measures, then for any measure $dm(\alpha)$, where $\int dm(\alpha) = 1$, the statistic $\rho = \int \rho_\alpha dm(\alpha)$ is a coherent risk measure.

Taking equation [14.3] and suggesting a measure $dm(\alpha)$, where $\alpha \in [0, 1]$, then:

$$M_m(X) = \int_0^1 \alpha ES_\alpha(X) dm(\alpha) = - \int_0^1 \int_0^\alpha F_X^\leftarrow(p) dp dm(\alpha), \quad [14.4]$$

is a coherent risk measure, provided $\int_0^1 \alpha dm(\alpha) = 1$, which stands as the normalization condition, holds. By interchanging the integrals (Fubini–Tonelli theorem), we get $M_m(X) \equiv M_\phi(X)$ (Acerbi 2002, equation 8).

Also, $dm(\alpha)$ can be expressed by a decreasing and positive $\phi(p)$, namely a risk spectrum, which equals $\int_p^1 dm(\alpha)$ and is normalized, thus $\int_0^1 \phi(p) dp = 1$ (Acerbi 2002, equation 9).

DEFINITION 14.5.– Risk spectrum $\phi(p)$ is a function that equals $\int_p^1 dm(\alpha)$.

Let us define the property of positivity for $\phi(p)$:

DEFINITION 14.6.– $\phi(p)$ exhibits the property of positivity if for every $A \subseteq [0, 1]$, $\int_A \phi(p) dp \geq 0$.

REMARK 14.1.– The property of positivity ensures that the risk spectrum will not give negative weights at some quantiles of the cumulative distribution function of the returns.

Moreover, it is necessary to define the condition that the risk spectrum needs to be endowed with in order to be decreasing:

DEFINITION 14.7.– Condition for $\phi(p)$ to be decreasing is that for every $a \in (0, 1)$ and for every $b > 0$, such that

$$[a - b, a + b] \subset [0, 1], \int_{a-b}^a \phi(p) dp \geq \int_a^{a+b} \phi(p) dp.$$

REMARK 14.2.– The aforementioned definition ensures that the risk spectrum will give more weights as the losses in the cumulative distribution function are greater. In other words, a decreasing function depicts the risk aversion of the investor.

REMARK 14.3.– A risk spectrum $\phi(p) : [0, 1] \rightarrow \mathbb{R}^+$ that is decreasing and satisfies the properties of normalization and positivity gives a coherent risk measure that can be expressed by equation [14.4], where $dm(a) = -d\phi(a)$.

Also, Acerbi (2002) states that the risk spectrum should be presented as an element φ of $L^1([0, 1])$ normed space, where its element is depicted by a class of functions. Different depictions (functions) of the same element φ will define the same measure.

In addition, we consider the real-world risk management applications where Acerbi (2002, equation 13) is well defined and finite. We give the following definition:

DEFINITION 14.8.– If $\varphi(p)$ is positive, decreasing and normalized, then it is an admissible risk spectrum. Moreover, a spectral risk measure endowed with an admissible risk spectrum is a coherent risk measure.

Recall that a basic issue we wish to tackle is the fact that $\varphi(p)$ is connected solely to the utility function of the investor, regardless of anything else. In other terms, it should solely depict the risk aversion of the investor. We, on the other hand, highlight the importance of the systemic risk. In practice, an investment that takes place in a developed market with poor macroeconomic factors will tend to have more systemic risk than an investment in a mature market with solid macroeconomic indexes. In addition, for highlighting the significance of systemic risk, we introduce conditions that $\varphi(p)$ should satisfy.

The connection of $\varphi(p)$ solely with the investors risk preferences is very well documented in many recent studies. Acerbi (2002) states that an investor can demonstrate a personal attitude against risk via $\varphi(p)$, and Dowd et al. (2008) state that spectral risk measures are

the type of risk measures that take into account the risk aversion of the investor.

In addition, Dowd et al. ([2008](#)) highlighted caveats that arise on two classes of spectral risk measures, the power and exponential. Those caveats concern the $\varphi(p)$ (utility function), and suggest that an admissible risk spectrum can lead to a not “well behaved” risk measure. Also, Leopold ([2015](#)) explains that $\varphi(p)$ provides a subjective weight in any quantile of the lost distribution. As losses increase, if the weight is constant, we get a risk neutral attitude and the case of increasing weight suggests risk aversion.

Before proceeding, we should quote some spectral risk measures developed in recent literature where their risk spectrum is admissible under certain conditions:

1. Power: this came to prominence from Dowd et al. ([2008](#)), where they consider utilizing the properties existing in the power utility function. Its depiction is:

$$\phi(p, d) = dp^{d-1} \quad [14.5]$$

where $d \in (0, 1]$.

2. Exponential: initially, it was suggested by Acerbi ([2004](#)) and later represented by Cotter and Dowd ([2006](#)) and Dowd and Blake ([2006](#)) in the following form:

$$\phi(p) = \frac{Ae^{-Ap}}{1 - e^{-A}} \quad [14.6]$$

where $A \geq 0$ stands as the Arrow–Pratt coefficient of absolute risk aversion.

3. Wang transform: this is suggested by Wang ([2000](#)) and despite the fact that Wang transform was developed before the concept of admissible risk spectrum, nevertheless it fulfills its requirements. The Wang transform also fulfills the requirements of Balbas et al. ([2009](#)) in order to be a well defined risk measure.

14.4. Contribution

Considering the example of section [14.1.1](#), it is apparent that whether a macroeconomic factor exceeds the threshold that signifies more systemic risk or not can be perceived as a random experiment. We take that $N_1, N_2, \dots, N_k$ are k economic factors that are responsible for depicting the robustness of an economy (how much systemic risk it bears). In addition, $k \in \mathbb{N}$ and k is of course deterministic. Moreover, we consider that each trial corresponds to each year's figures of those economic factors. Also, $a_1, a_2, \dots, a_k$ are k thresholds as described in section [14.1.1](#), where each a_i for every $i \in \{1, 2, 3, \dots, k\}$ corresponds to the N_i economic factor. For regulation purposes, every a_i could be a threshold that concern macroeconomic indexes while for investing purposes it could be a limit of the capitalization of an index.

Under that perspective, we set that for every i random experiment, there is an X_i random variable that follow the discrete binomial distribution:

$$X_i \sim B(n_i, p_i). \quad [14.7]$$

Each experiment for every year is a Bernoulli trial where n_i is the number of trials and p_i is the probability for success (the probability that the a_i threshold will be exceeded). Also, the probability for failure $q_i = 1 - p_i$ is the complement of the probability for success. Also, its experiment for every year is statistically independent. A similar assumption is met in Champernowne ([1953](#)), where the model's assumption suggests that the distribution will tend toward a unique equilibrium distribution dependent upon the stochastic matrix but not on the initial distribution. Recall that a stochastic matrix describes the transitions of a Markov chain, which is a sequence of memoryless events.

The distribution function of each random variable X_i is:

$$P_{X_i}(x_i) = \binom{n_i}{x_i} p_i^{x_i} (1 - p_i)^{n_i - x_i}, \quad [14.8]$$

where x_i are the number of successes for n_i trials (n years). The sample space $\Omega = \{(N_i > a_i | \prod_{i=1}^k N_i \in \mathbb{R})\}$, and also $\mathbb{P}(\Omega) = 1$, since each factor can either exceed or fail to exceed the threshold that corresponds to it. In addition, X_i 's range is $\{0, 1, 2, \dots, n_i\}$. Following the above setting of our experiment, we also set its probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\mathcal{F}$ is the set of events which includes the subset $\mathcal{X}$ of all x_i succeeded experiments. Also, the indicator function of a subset $\mathcal{X}$ of a set $\mathcal{F}$ is a function $1_{\mathcal{X}} : \mathcal{F} \rightarrow \mathbb{R}_+$ defined as:

$$1_{\mathcal{X}(x_i)} =: \begin{cases} > 0 & \text{if } x_i \in \mathcal{X} \\ = 0 & \text{if } x_i \notin \mathcal{X}. \end{cases} \quad [14.9]$$

Consider the following random variable:

$$Y = 1 + \sum_{i=1}^k 1_{\mathcal{X}(x_i)}, \quad [14.10]$$

where $\sum_{i=1}^k 1_{\mathcal{X}(x_i)}$ is the sum of the aforementioned indicator functions and it is trivial to see that equation [\[14.10\]](#) is non-negative. where

In each experiment, ideally we would get that $N_i \leq a_i$ or that each N_i should have an upper bound a_i in order to be free of systemic risk. Attempting a further discussion on the thresholds, it is obvious that there are many ways to be determined, yet they have to be acknowledged by all of the markets they wish to utilize. For regulation purposes, the limits (macroeconomic factors) that European economies need to reach in order to be inducted in European monetary union can serve as thresholds. For investing purposes, a numeraire can prove quite useful. In any case, these

thresholds will present the ideal and be free of systemic risk economy. In order to further illustrate this matter, we suggest an application in section 14.5. Moreover, equation [14.10] is not determining if all components are equally weighted. While we do that for simplicity, this may not be always the case and it is a topic that can be further addressed.

Consider the power spectral risk measure as depicted in equation [14.5]. Now assume that for the r.v. of X in equation [14.3] (money outcome of the investment), it holds that

$\mathbb{E}[X^+] < \infty$, $\mathbb{E}[X^-] < \infty$, $\sup_{p \in [0,1]} \phi(p) < \infty$ and the observations of the returns of the investment that is represented in equation [14.10] $\rightarrow \infty$. Then, according to Acerbi (2002, theorem 5.4), a spectral risk spectrum with equation [14.5] converges to equation [14.4] with probability 1.

In addition, when the $\phi(p)$ is a known function, then according to Dowd et al. (2008), the equivalent properties for a risk spectrum to be admissible would be:

1. $\phi(p) \geq 0$.
2. $\phi'(p) \leq 0$.
3. $\int_0^1 \phi(p) dp = 1$.

To be precise, Dowd et al. (2008) stated the property of increasingness but this occurs when the distributions in which loss outcomes are given positive values. On the contrary, Acerbi (2002) is implying negative values. Now consider the following function:

$$\psi(Y, p, d) = Y \left[dp^{d-1} \right] - (Y - 1), \text{ where } d \in \left[\frac{Y-1}{Y}, 1 \right] \text{ and } p \in (0, 1]. \quad [14.11]$$

PROPOSITION 14.1.— Note that equation [14.11] is an admissible risk spectrum.

PROOF.– We consider that for $\psi(Y, p, d) \geq 0$, we should get that $Y[dp^{d-1}] - (Y - 1) \geq 0$, which leads to $dp^{d-1} \geq \frac{Y-1}{Y}$.

Moreover, we consider that:

$$\begin{cases} 0 < p \leq 1 \\ d - 1 \leq 0 \end{cases} \quad [14.12]$$

which leads to $p^{d-1} \geq 1 \Rightarrow dp^{d-1} \geq d, \forall p \in (0, 1]$. From the aforementioned, we get that $dp^{d-1} \geq d \geq \frac{Y-1}{Y}$ or if we need $\psi(Y, p, d)$ to be positive, then we should have that

$$d \in \left[\frac{Y-1}{Y}, 1 \right].$$

Now let us prove that $\psi'(Y, p, d) < 0$ (differentiating with respect to p). For that, we get

$$\left[Y dp^{d-1} - (Y - 1) \right] = Y d(d - 1) p^{d-2}, \text{ which is negative}$$

because Y, d and p^{d-2} are positive, while $d - 1$ is negative.

Finally, for normalization, we should get that

$$\int_0^1 \psi(Y, p, d) dp = 1, \text{ which is true because}$$

$$\int_0^1 Y dp^{d-1} - (Y - 1) dp = \int_0^1 Y dp^{d-1} dp - \int_0^1 Y - 1 dp = Y \left[p^d \right]_0^1 - (Y - 1) \left[p \right]_0^1 = Y - Y + 1 = 1$$

.

We refer the reader to the Appendix, where we present a specific example and its proof of the risk spectrum we propose.

Now, we assume that there are always two markets M_a and M_b where market M_a includes more systemic risk, in terms that it has more number of successes x_i . Let consider that Y_a and Y_b are the equation [\[14.10\]](#) functions of M_a and M_b , respectively.

PROPOSITION 14.2.– For every two markets M_a and M_b where market M_a includes more systemic risk compared to M_b , $Y_a > Y_b$ is true.

PROOF.– We have to prove that

$$Y_a > Y_b \Rightarrow \left[1 + \sum_{i=1}^k 1_{\mathcal{X}(x_i)} \right]_a > \left[1 + \sum_{i=1}^k 1_{\mathcal{X}(x_i)} \right]_b \Rightarrow \left[\sum_{i=1}^k 1_{\mathcal{X}(x_i)} \right]_a > \left[\sum_{i=1}^k 1_{\mathcal{X}(x_i)} \right]_b.$$

. We have from the setting that $(x_i)_a > (x_i)_b$, which leads to

$$\left[\sum_{i=1}^k 1_{\mathcal{X}(x_i)} \right]_a > \left[\sum_{i=1}^k 1_{\mathcal{X}(x_i)} \right]_b \text{ and thus } Y_a > Y_b \text{ is true.}$$

Also, let us consider that $\psi_a(Y_a, p, d)$ and $\psi_b(Y_b, p, d)$ are the equation [14.11] functions of M_a and M_b , respectively. In addition, we set that $\psi(Y, p, d) - \phi(p, d) = \epsilon$. Moreover, we can consider the p as an increasing function where $p_{(x)} = x$ and $x \in [0, 1]$. Let us consider the real numbers $\gamma, \delta \in [0, 1]$ such that when $\gamma > \delta$, then $p_\gamma > p_\delta$. Now, we are ready to give the following definition:

DEFINITION 14.9.– A spectral risk measure is systemic risk aware when it satisfies the following conditions:

1. For every $Y_a > Y_b \Rightarrow \psi_a(Y, p, d) > \psi_b(Y, p, d)$.
2. For $Y > 1$, $\psi(Y, p_\gamma, d) - \phi(p_\gamma, d) = \epsilon_\gamma$ and $\psi(Y, p_\delta, d) - \phi(p_\delta, d) = \epsilon_\delta$, $\epsilon_\gamma < \epsilon_\delta$ is true.

REMARK 14.4.– The first condition assures that between two markets, both utilize the same risk measure and both exhibit *SR*; the market that has more *SR* (according to the quantification we propose) will have a greater risk spectrum and will lead the utilizer to be more conservative.

The second condition assures that the magnitude of the difference between a risk spectrum that is not endowed with r.v. Y and the one that is endowed with it is greater when we are considering a worst outcome in the returns' distribution function.

We present a numerical example in the Appendix for the possible understanding of the above remark.

THEOREM 14.1.– For equation [14.11], the conditions of definition 14.9 hold, and equation [14.11] is a systemic risk aware and admissible risk spectrum.

PROOF.– We have from the proof of proposition 14.1 that $\psi(p)$ is an admissible risk spectrum. Consequently, we take as proven that both $\psi(Y, p, d)$ and $\phi(p, d)$ are positive, convex and decreasing functions.

For the first condition, we should get that

$$\begin{aligned} \psi_a(Y, p, d) > \psi_b(Y, p, d) &\Rightarrow Y_a \phi(p, d) - Y_a + 1 > Y_b \phi(p, d) - Y_b + 1 \\ &\Rightarrow Y_a \phi(p, d) - Y_a > Y_b \phi(p, d) - Y_b \Rightarrow Y_a (\phi(p, d) - 1) > Y_b (\phi(p, d) - 1) \\ &\Rightarrow Y_a > Y_b, \text{ which is true and thus the} \\ &\text{second condition is met for equation [14.11].} \end{aligned}$$

For the second condition, we should get that $\epsilon_\gamma < \epsilon_\delta$. So

$$\begin{aligned} \epsilon_\gamma < \epsilon_\delta &\Rightarrow \psi(Y, p_\gamma, d) - \phi(p_\gamma, d) < \psi(Y, p_\delta, d) - \phi(p_\delta, d) \Rightarrow \\ &\psi(Y, p_\gamma, d) + \phi(p_\delta, d) < \psi(Y, p_\delta, d) + \phi(p_\gamma, d) \Rightarrow Y \phi(p_\gamma, d) - Y + 1 + \phi(p_\delta, d) < \\ &Y \phi(p_\delta, d) - Y + 1 + \phi(p_\gamma, d) \Rightarrow Y \phi(p_\gamma, d) + \phi(p_\delta, d) < Y \phi(p_\delta, d) + \phi(p_\gamma, d) \Rightarrow \\ &Y \phi(p_\gamma, d) - \phi(p_\gamma, d) < Y \phi(p_\delta, d) - \phi(p_\delta, d) \Rightarrow \phi(p_\gamma, d) (Y - 1) < \\ &\phi(p_\delta, d) (Y - 1) \Rightarrow \phi(p_\gamma, d) < \phi(p_\delta, d) \text{ which is true as} \\ &\phi(p) \text{ is decreasing and thus third condition is met.} \end{aligned}$$

COROLLARY 14.1.– Any admissible spectral risk spectrum that also exhibits the conditions of definition 14.9 is an admissible and systemic risk-aware spectral risk spectrum.

PROOF.– Assume that there is an admissible risk spectrum other than equation [14.11] and a variable other than equation [14.10]. In addition, assume that this variable preserves for the risk spectrum the properties of definition 14.9. With those assumptions and setting, we observe that the admissible risk spectrum is systematic risk-aware.

So far, we have proposed three desirable conditions that the risk spectrum needs to satisfy once there is systemic risk. We have also found a way to quantify the systemic risk and embody it to the risk

spectrum with the utilization of a variable. In tandem, we preserved the properties of the risk spectrum needed in order to be admissible. In terms of practicality preliminaries, we use basics of set theory (see, for instance, Halmos (1974)) to ensure that the set of coherent and systemic risk aware measures is not the null set. Consider that $\mathcal{C}$ stands as the set of all coherent risk measures, then we have the following.

LEMMA 14.1.– A spectral risk measure $M_\phi(X)$ that has an admissible risk spectrum ϕ is a subset of coherent risk measures, or $M_\phi(X) \subset \mathcal{C}$.

PROOF.– Consider an $M_\phi(X)$ that is a spectral risk measure and has an admissible risk spectrum $\phi(p)$. We know that $M_\phi(X) \neq \emptyset$, as we already mentioned power spectral risk measures, exponential spectral risk measures and Wang transform, which are spectral risk measures that have an admissible risk spectrum. Consequently, since $M_\phi(X)$ is a coherent risk measure, $M_\phi(X) \subseteq \mathcal{C}$.

In addition, we are also aware of ES_a , which is a risk measure that is coherent, yet its risk spectrum is not admissible. This suggests that $M_\phi(X) \neq \mathcal{C}$ and so $M_\phi(X) \subset \mathcal{C}$.

LEMMA 14.2.– A spectral risk measure $M_\psi(X)$ that has an admissible risk spectrum $\psi(p)$ that satisfies the conditions of definition 14.9 is a coherent risk measure and so $M_\psi(X) \subset \mathcal{C}$. In addition, $M_\psi(X) \neq \emptyset$.

PROOF.– The previous lemma suggests that a $M_\psi(X)$ is coherent and so $M_\psi(X) \subseteq \mathcal{C}$. Also, $ES_a \subset \mathcal{C}$ and does not have an admissible risk spectrum. Consequently, $M_\psi(X) \neq \mathcal{C}$ and thus $M_\psi(X) \subset \mathcal{C}$.

For proving $M_\psi(X) \neq \emptyset$, we need to suggest at least one coherent spectral risk measure that has a risk spectrum that satisfies conditions of definition 14.9. For that, we consider equation [14.11], where the property of admissibility holds, the risk measure is coherent and $M_\psi(X) \neq \emptyset$.

14.5. A suggested application

As aforementioned, the limits (macroeconomic factors) that European economies need to reach in order to be inducted in European monetary union can serve as thresholds. To that end, if a regulator needs to quantify the systemic risk that exists in a country, they can utilize the five convergence criteria (European Commission n.d.; European Union [2008](#)) that needs to be met in order to join the eurozone:

1. First criterion is interested in the stability of an economy. An economy meets that criterion if its inflation is not more than 1.5% above the rate of the three best performing member states.
2. Second criterion examines the robustness of the public finance of the economy. The criterion is met if the government deficit is not more than 3 of the GDP.
3. Third criterion is about the sustainability of public finances. If an economy needs to meet the criteria, the debt of the government should not exceed 60% of the GDP.
4. Fourth criterion is interested in the durability of convergence. To that end, the criterion is that the long-term interest rate is not more than 2% above the rate of the three best performing Member States.
5. Fifth and final criterion is the stability of the exchange rate. The criterion for that is that the inductee economy is participating in ERM II for at least two years without severe tensions.

Of course, those criteria can serve directly as N factors for an investor who wishes to invest in economies either in the eurozone or planning to adopt the euro. Its elementary to see that $k = 5$ and there are five a corresponding to each criterion. Now we can directly proceed by utilizing equation [\[14.10\]](#) of the economy they are interested to invest. Afterward, they can calculate the systemic risk aware risk spectrum that is proposed from equation [\[14.11\]](#) and utilize it accordingly for their investing strategy. The apparent benefit is a spectral risk measures that is dictated not only by the risk aversion of the investor but also by the quantified systemic risk

underlined in its economy. Moreover, if someone chooses to invest in a country outside the European union, then they simply have to exclude the fifth criterion and set $k = 4$.

14.6. Conclusion

In this chapter, we tackle the issue that risk spectrum of a spectral risk measure is portraying solely the preferences of an investor/regulator on their risk tolerance. We argue by suggesting that also the systemic risk of market should be embodied in the risk spectrum. In line with the above suggestion, we attempt to solve that issue by suggesting two conditions that a risk spectrum should satisfy. We gave an example where we demonstrated how systemic risk can be quantified from one market to another. In the same example, fundamental macroeconomic indexes prove quite useful as they can expose in a sense “how much” systemic risk a market bears. In addition, we proposed a risk spectrum class that satisfy those conditions. Summing up, we mention that there are limitations concerning the spectral risk measures we propose. For instance, we need to reconsider if those factors should be equally weighted. We need to rethink that depending on the extent of the poor performance of a factor, maybe it should be weighted more than the others. Also, we consider that we could use the determinants that we mentioned in section [14.1.1](#) for calculating the r.v. Y of each market. In this way, the random experiment could be avoided but some other conditions should be utilized. Although it is beyond the scope of this study to address such matters, we believe that there should be a study with respect to those topics in the near future. We also believe that an empirical validation of the proposed application should take place with investments within economies that are either members of the eurozone or scheduling to join.

14.7. References

Acemoglu, D., Ozdaglar, A., Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, 105(2), 564–608.

- Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *The Journal of Banking and Finance*, 26(7), 1505–1518.
- Acerbi, C. (2004). Coherent representation of subjective risk aversion. In *Risk Measures for the 21st Century*, Szego, G. (ed.). Wiley, Hoboken, NJ.
- Acerbi, C. and Tasche, D. (2002). On the coherence of expected shortfall. *The Journal of Banking and Finance*, 26(7), 1487–1503.
- Acharya, V.V., Engle, R.F., Richardson, M.P. (2012). Capital shortfall: A new approach to ranking and regulating systemic risks. *The American Economic Review*, 102(3), 59–64.
- Acharya, V.V., Pedersen, L.H., Philippon, T., Richardson, M.P. (2017). Measuring systemic risk. *The Review of Financial Studies*, 30(1), 2–47.
- Artzner, P., Delbaen, F., Eber, J.M., Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203–228.
- Asimit, V.A. and Li, J. (2017). Systemic risk: An asymptotic evaluation. *SSRN*. doi: 10.2139/ssrn.2911586.
- Balbas, A., Garrido, J., Mayoral, S. (2009). Properties of distortion risk measures. *Methodology and Computing in Applied Probability*, 11, 385–399.
- BCBS (2012). Fundamental review of the trading book: Basel committee on banking supervision. Report, Basel Committee of Banking Supervision.
- BCBS (2013). Fundamental review of the trading book: A revised market risk framework. Report, Basel Committee of Banking Supervision.
- Brownlees, C. and Engle, R. (2006). SRISK: A conditional capital shortfall measure of systemic risk. *The Review of Financial Studies*, 30(1), 48–79.

- Case, K. and Fair, R. (2007). *Principles of Economics*. Pearson Education, Upper Saddle River, NJ.
- Champernowne, D. (1953). A model of income distribution. *The Economic Journal*, 63(250), 318–351.
- Cochrane, J.H. (2000). *Asset Pricing*. Princeton University Press, Princeton, NJ.
- Cotter, J. and Dowd, K. (2006). Extreme spectral risk measures: An application to futures clearinghouse margin requirements. *Journal of Banking and Finance*, 30(12), 3469–3485.
- Danielsson, J., Jorgensen, B.N., Samorodnitsky, G., Sarma, N., de Vries, C.G. (2005). Subadditivity re-examined: The case for value-at-risk. Technical Report, Cornell University, London.
- Dowd, K. and Blake, D. (2006). After VaR: The theory, estimation, and insurance applications of quantile-based risk measures. *Journal of Risk and Insurance*, 73(2), 193–229.
- Dowd, K., Cotter, J., Sorwar, G., Darkiewicz, G. (2008). Spectral risk measures: Properties and limitations. *Journal of Financial Services Research*, 34(1), 61–75.
- Embrechts, P., Puccetti, G., Ruschendorf, L., Wang, R., Beleraz, A. (2014). An academic response to Basel 3.5. *Risks*, 2, 25–48. doi: 10.3390/risks2010025.
- European Commission (n.d.). European Commission [Online] Available at: <https://commission.europa.eu/>.
- European Union (2008). Consolidated version of the treaty on the functioning of the European Union. *Official Journal of the European Union*, 115, 47–388.
- Follmer, H. and Shied, A. (2004). *Stochastic Finance: An Introduction in Discrete Time*. de Gruyter, Berlin.
- Follmer, H. and Shied, A. (2008). Convex and coherent measures of risk. Humboldt University, Berlin.

- Halmos, R. (1974). *Naive Set Theory*. Springer, New York. doi: 10.1007/978-1-4757-1645-0.
- Kaufman, G. (2000). Banking and currency crises and systemic risk: Lessons from recent events. *Economic Perspectives*, 24(3), 9–28.
- Konstantinides, D. and Zachos, G. (2019). Exhibiting abnormal returns under a risk averse strategy. *Methodology and Computing in Applied Probability*, 21(2), 551–566. doi: 10.1007/s11009-018-9673-9.
- Leopold, M. (2015). *Value-at-Risk and Other Risk Measures*, 1st edition. University of Zurich and the Swiss Finance Institute and Oxford University Press, New York.
- Loebnitz, K. (2011). Liquidity risk meets economic capital and RAROC. A framework for measuring liquidity risk in banks. PhD Thesis, University of Twente. doi: 10.3990/1.9789036532990.
- Loebnitz, K. and Roorda, B. (2011). Liquidity risk meets economic capital and RAROC. *SSRN*. doi: 10.2139/ssrn.1853233.
- Lofthouse, S. (2000). *Investment Management*, 2nd edition. Wiley, Hoboken, NJ.
- Luciano, E. and Wihlborg, C. (2015). Complex financial institutions and systemic risk [Online]. Available at: <https://www.semanticscholar.org/paper/Complex-financial-institutions-and-systemic-risk-Luciano-Wihlborg>.
- Lujun, L., Hui, S., Ruodu, W., Jingping Hui, Y. (2017). Worst-case range value-at-risk with partial information. *SSRN* [Online]. Available at: <https://ssrn.com/abstract=2920334>.
- Mankiw, G.N. (2009). *Macroeconomics*, 9th edition. Worth Publishers, New York.
- Margalit, D., Rabinoff, J., Rolen, L. (2018). Interactive linear algebra. Georgia Institute of Technology [Online]. Available at: <https://textbooks.math.gatech.edu/ila/ila.pdf>.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.

McNeil, J.A., Frey, R., Embrechts, P. (2005). *Quantitative Risk Management, Concepts, Techniques and Tools*. Princeton University Press, Princeton, NJ.

Wang, S.S. (1996). Premium calculation by transforming the layer premium density. *ASTIN Bulletin*, 26(1), 71–92. doi: 10.2143/AST.26.1.563234.

Wang, S.S. (2000). A class of distortion operators for financial and insurance risks. *The Journal of Risk and Insurance*, 67(1), 15–36.

Young, C.E. (1993). *Vector and Tensor Analysis*, 2nd edition. Marcel Dekker, Inc., New York.

14.8. Appendix A

EXAMPLE 14.1.– Let us take equation [14.11] and set that $d = 0.9$ and $k \leq 89$ or $\psi(Y, p) = Y [0.9p^{-0.1}] - (Y - 1)$. Now we can prove that it is admissible:

PROOF.–

$$\psi(Y, p) = ((dY)p^{d-1} - (Y - 1)) \Rightarrow ((0.9Y)p^{-0.1} - (Y - 1)) \Rightarrow \frac{9Y}{10 \sqrt[10]{p}} - Y + 1 \quad [14.13]$$

Calculating for the first derivative with respect to p , we get:

$$\psi(Y, p)' = \frac{d}{dp} \left[\frac{9Y}{10 \sqrt[10]{p}} - Y + 1 \right] = \frac{9Y}{10} \frac{d}{dp} \left[\frac{1}{\sqrt[10]{p}} \right] + \frac{d}{dp} [-Y] + \frac{d}{dp} [1] = \quad [14.14]$$

$$\frac{9 \left(-\frac{1}{10} \right) p^{-\frac{1}{10}-1} f(p_i)}{10} = -\frac{9Y}{100 p^{\frac{11}{10}}}, \quad [14.15]$$

where $100p^{\frac{11}{10}} > 0$ as $p \in (0, 1]$ and $9Y > 0$ as equation [14.10] is non-negative. Also, $k \leq 89 \Rightarrow 1 \leq Y \leq 9.9$. Recall that from the setting, we observe that Y has range $[1, 9.9]$ and so

$d \in \left[\frac{Y-1}{Y}, 1 \right]$ is satisfied. Summing up,

$\psi(Y, p)' < 0 \Rightarrow \psi(Y, p)$ is decreasing in $(0, 1]$ and second condition is met for $\psi(Y, p)$ to be admissible. Calculating for the second derivative, we get:

$$\psi(p)'' = \frac{d}{dp} \left[-\frac{9Y}{100p^{\frac{11}{10}}} \right] = -\frac{9Y}{100} \frac{d}{dp} \left[\frac{1}{p^{\frac{11}{10}}} \right] = -\frac{9(-\frac{11}{10})p^{-\frac{11}{10}-1}Y}{100} = \frac{99Y}{1,000p^{\frac{21}{10}}}, \quad [14.16]$$

where $1,000p^{\frac{21}{10}} > 0$ as $p \in (0, 1]$, $99Y > 0$ as $\psi(Y, p)$ is non-negative. Concluding $\psi(Y, p)'' > 0$ and thus $\psi(p)$ is convex. Moreover, for proving the positivity, we consider the decreasing property and we take $\psi(0^+)$:

$$\psi(0^+) = (Y0.9)0^{+(-0.1)} - Y + 1, \quad [14.17]$$

and we consider that

$$\psi(0^+) > 0 \Rightarrow (Y0.9)0^{+(-0.1)} - Y + 1 > 0 \Rightarrow (Y0.9)0^{+(-0.1)} > Y - 1$$

, which is true since $(Y0.9)0^{+(-0.1)} = \infty$ and $Y \leq 9.9$. For the latter, recall that we set $k < 89$ or there is a limit of macroeconomic factors that are responsible for generating systemic risk. In case all experiments succeed, then $Y = 9.9$. Moreover,

$$\psi(1) = (Y0.9)1^{-0.1} - Y + 1 \Rightarrow 0.9Y - Y + 1 > 0 \Rightarrow 0.9Y - Y > -1 \Rightarrow -0.1Y > -1 \Rightarrow 0.1Y < 1 \Rightarrow \frac{Y}{10} < 1 \Rightarrow Y < 10,$$

, which is true. Since $\psi(0^+) > 0$ and $\psi(1) > 0$, the positivity condition is proven. Concerning normalization condition, we should have that

$$\int_0^1 \psi(Y, p) dp = 1 \text{ and so:}$$

$$\int_0^1 \psi(Y, p) dp = \int_0^1 \left[\frac{9Y}{10 \sqrt[10]{p}} - Y + 1 \right] dp = \frac{9Y}{10} \int_0^1 \frac{1}{\sqrt[10]{p}} + 1 - Y dp = \quad [14.18]$$

$$\left[(1 - Y)p + Y p^{\frac{9}{10}} + C \right]_0^1 = \left[p - Y p + Y p^{\frac{9}{10}} + C \right]_0^1 = \quad [14.19]$$

$$\left[1 - Y + Y \right] - \left[0 + 0 + 0 \right] = 1 \quad [14.20]$$

normalization condition holds and thus is an admissible risk spectrum.

14.9. Appendix B

EXAMPLE 14.2.– Consider the equation in the above example. If we set $p = 0.001$, the greater losses are, then $\varphi_{0.001} = 2.01640$ while $\psi_{0.001} = 2.21968$. On the other side of the distribution, where $p = 0.999$, we get that $\varphi_{0.999} = 1.01069$ while $\psi_{0.999} = 1.01282$. We take into account the properties of $\varphi(p)$ and $\psi(p)$ and we observe that when systemic risk emerges, then $\psi(p) > \varphi(p)$ for almost the whole p interval and the first condition is met. Also, the second condition is met as $2.21968 - 2.01640 > 1.01282 - 1.01069$ is true.

Moreover, compared to $\varphi(p)$, $\psi(p)$ guides a practitioner to be more conservative in the presence of systemic risk as losses become greater. On the contrary, when losses diminish, $\psi(p)$ grants similar weight like $\varphi(p)$. Trivially, we observe that Y forces $\psi(p)$ to give different weight to $F_X^{\leftarrow}(p)$ depending on the existing systemic risk of each market and thus it quantifies systemic risk. Also, from the above setting, we observe that $\psi(p)$ grants more weight to the increasing losses of $F_X^{\leftarrow}(p)$ as more systemic risk occurs.

Note

Chapter written by Georgios C. ZACHOS.

15

Mutual Information and Pearson's Correlation in Analysis of Statistical Tests Suites: A Case Study

In this chapter, we present a case study in which a statistical battery is analyzed using two different methodologies. The study focuses on the composition of statistical test batteries; a current line of research aimed at ensuring their efficiency and reliability. To work efficiently with such batteries, it is necessary that the tests composing them are independent, that is, they do not assess the same property in the same way, and that they are computationally efficient. If these requirements are not met, the time required to obtain results may increase significantly and the outcomes may even be masked.

15.1. Introduction

There are several areas of knowledge in the literature in which the use of sequences of (pseudo-)random numbers is of fundamental importance. We can highlight, for example, applications in mathematical simulation, statistics, operations research or cryptography, among others.

The sequences used must verify a series of conditions such as randomness and uniformity, as well as unpredictability if the application is focused on the area of cryptography. Generally, in order to verify the goodness of fit of the generated sequences, hypothesis tests are used, such as the χ^2 goodness-of-fit test or the runs test, among others. In practice, sets of tests are usually used to check the properties of the sequences from different points of view; these sets are known as suites or batteries.

Among the best known batteries, NIST STS 800-22 (Rukhin et al. [2010](#)), Dieharder (Brown et al. [2014](#)), TestU01 (L'Ecuyer and Simard [2007](#)) or ENT (Walker [2008](#)) are worth mentioning. For a complete

development on the different batteries and the tests that compose them, see Almaraz Luengo and García Villalba ([2021](#)).

One of the fundamental aspects of practical data management is to perform these verifications in a complete and efficient manner. That is to say that the desirable properties for the sequences are verified in a statistically valid way and that such verification is performed in a reasonable computational time. This is why a current line of research is focused on the correct design of a suite of statistical tests and on the analysis of the characteristics of the existing ones. The fundamental objective is that a test suite should contain the necessary and sufficient number of tests to allow conclusions to be drawn reliably and that this should not take an excessive amount of time to verify. For this reason, the tests that constitute a suite must be focused on the non-redundant checking of properties, since if, for example, there were two tests in a suite that were measuring or verifying the same property, the final results could be masked and the execution time would be longer.

The analysis of the design of a hypothesis tests suite has traditionally been performed by studying the p -values derived from the application of the tests on the sequences. The most commonly used approach is to measure the interrelation of the tests through the sample Pearson's correlation coefficient¹, being the samples composed by the p -values (see, for example, Hellekalek and Wegenkittl ([2003](#)), Turan et al. ([2008](#)), Doganaksoy et al. ([2008](#)), Georgescu et al. ([2017](#)) or Burciu and Simion ([2019](#)), among others). There is also the possibility that the samples are formed by elements other than the p -values, such as samples of results of values of statistics and measures, as they do in Hernández-Castro and Barrero ([2017](#)), although this is not the most widespread approach. In that case, the authors justified the sample selection to avoid a possible loss that could result from not considering the information contained in the ENT measures. However, that approach has the disadvantage of not being comparable with other related studies and, in addition, the lack of a p -value prevents the user from determining when or under which threshold a statistical test should be considered passed or failed.

Continuing with the p -value study approach, it is also possible to analyze measures such as the fail-fail ratio (FFR), pass-fail ratio (PFR) or pass-pass ratio (PPR). This approach has been developed in Sulak et al. (2017). In this study, the authors investigate the dependencies among the NIST tests, focusing not on the overall correlation values, but rather on the correlations at the extreme values to determine test dependencies. In Dog˘anaksoy et al. (2008), the study of correlations is also performed by analyzing the sequences of p -values, but the analysis of transformed sequences is added. This approach has also been used in Dog˘anaksoy et al. (2017) applied to the NIST battery. Other less used possibilities are the factor analysis used in NIST or the analysis through densities (see Fan et al. (2013)).

More recently, in Karell-Albo et al. (2020), the discrete version of the mutual information measure² has been proposed to address the analysis of interdependencies (not exclusively linear as analyzed by Pearson's correlation) between the tests in a suite. In fact, it can be shown (see Cover and Thomas (2006)) that two random variables X and Y are independent if and only if $I(X, Y) = 0$. In concrete, if $U = \{u_1, \dots, u_n\}$ is the set of sequences on which the tests are evaluated, and T_1, T_2 two tests, tests T_1 and T_2 are applied on the elements in U obtaining two sets of p -values (one for each test)

$P_1 = \{p_1^1, \dots, p_n^1\}$ and $P_2 = \{p_1^2, \dots, p_n^2\}$. When the set of sequences is sufficiently large, the resulting lists of p -values can be regarded as random variables that carry meaningful information about each test run. Therefore, the mutual information between P_1 and P_2 , $I(P_1, P_2)$, can give a measure on the (possible) interrelationship there might be between T_1 and T_2 .

This more novel approach has been applied to the analysis of the NIST (Karell-Albo et al. 2020) battery and more recently to analyze the FIPS (Evans et al. 2002) tests suite in Almaraz Luengo et al. (2022c).

This chapter will analyze the possible interdependencies presented by the ENT battery from the perspective of (sample) Pearson's correlation coefficient and mutual information measure.

15.2. ENT battery analysis

The ENT battery consists of four measures (entropy, arithmetic mean, Monte–Carlo estimation of π value and serial correlation coefficient) and a statistic value (χ^2). It is a battery widely used in cryptography for sequence verification as it is extremely computationally efficient, unlike other suites that perform highly expensive computations. This can be especially critical when the measurements themselves are performed not on desktop computers, but on smaller devices with hardware limitations. For example, to process a 10 MB file on a computer with an Intel® Core i7-5500U processor and 8 GB of RAM, NIST SP 800-22 takes more than 1 min to complete, while ENT processes it in 0.3 s.

In ENT, the sequence is analyzed at the byte level by default (with an optional mode for bit-level analysis). This means the data are processed 8 bits at a time, each interpreted as a non-negative integer ranging from 0 to 255, inclusive. It was designed by John Walker in 2008 and its main applications are those in which the information (entropy, hence its name) of a file is of interest, such as encryption, statistical sampling or compression algorithms. The ENT battery can be found on their website (<https://www.fourmilab.ch/random/>); however, this implementation is not entirely useful as it only returns the resulting statistics, not the p -values. However, after an in-depth study of the definition of the measures that constitute the battery, it is possible to determine p -values associated with them as recently developed by Almaraz Luengo et al. ([2022b](#)).

There are not too many works in the literature that have focused on the analysis of the design of this battery; we can mention Hernández-Castro and Barrero ([2017](#)) in which the correlation results between the measures provided by ENT are analyzed in raw form, that is, without calculating associated p -values, justifying as we mentioned previously that no information is lost by doing so. The problem with this approach is that it does not allow us to analyze the statistical significance of the results by not providing an underlying distribution of each measure (with the exception of the χ^2 statistic) and, therefore, not being able to calculate a p -value that would allow comparisons to be made given a predetermined significance level.

For example, if we are told that the arithmetic mean of a sequence is 127'9, how do we know if this is significantly far from 127'5?

Obviously it is not the same to obtain that mean with a sequence of only a hundred bytes as with a sequence of a million bytes. That is, rather than being at the user's discretion to decide whether the arithmetic mean is close enough to the expected one, regardless of the length of the sequence, it would be preferable to have a p -value in the style of that provided in the χ^2 test. It is then desired to obtain the probability that a random sequence of the same length as the one under study will obtain the observed results or more extreme ones.

A more recent analysis has been made in Almaraz Luengo et al. (2022a) showing its weaknesses and revealing that it is possible to design "degenerate" sequences that can pass the tests of this battery.

In this chapter, an analysis of the ENT suite (battery) is performed through the p -values associated with the battery measures from Pearson's correlation and mutual information measure approaches. For this purpose, the tests are run on 10^4 sequences of the same size. The experiments have been done for different sizes: 10^5 , 10^6 and 10^7 bits and for different generators. In this chapter, as the results obtained have been similar for the different sizes and generators, the results are shown for the case of the dev/urandom generator and a size of 10^7 bits. Note that 100 different tests for each measure have been carried out, and on each set of p -values a Kolmogorov–Smirnov (K-S) test with significance level $\alpha = 0.001$ has been performed.

In the following tables, MC π represents Monte Carlo estimation of π and SC represents serial correlation.

[Table 15.1](#) shows the Pearson's correlation results obtained on the sample p -values associated with the ENT measures showing a very high correlation between the p -values associated with the entropy and χ^2 measures. The rest of the couples do not have a high correlation.

Table 15.1. *Correlation matrix of samples of p-values*

	Entropy	χ^2	Mean	MC π	SC
Entropy	–	–0.999549	–0.047	–0.019	–0.0022
χ^2	–0.999549	–	0.047	0.019	0.0022
Mean	–0.047	0.047	–	0.11	0.0011
MC π	–0.019	0.019	0.11	–	–0.00042
SC	–0.0022	0.0022	0.0011	–0.00042	–

[Table 15.2](#) shows the results corresponding to the application of the K-S test. The test validates the above results.

Table 15.2. *Significance matrix obtained from the applied K-S test (case of Pearson's correlation)*

	Entropy	χ^2	Mean	MC π	SC
Entropy	–	0	0.22	0.29	0.28
χ^2	0	–	0.18	0.26	0.40
Mean	0.22	0.18	–	1.6×10^{-7}	0.61
MC π	0.29	0.26	1.6×10^{-7}	–	0.34
SC	0.28	0.40	0.61	0.34	–

In relation to the mutual information measure, the results are very similar to those shown by Pearson's correlation coefficient, as we can see in [Table 15.3](#), only the entropy- χ^2 pair turns out to be significant.

With regard to the analysis of significance, we show [Table 15.4](#). The results of the K-S test are like those of the Pearson's correlation.

Table 15.3. *Mutual information results*

	Entropy	χ^2	Mean	MC π	SC
Entropy	–	0.88	0.00084	0.00024	0.00023
χ^2	0.88	–	0.00084	0.00025	0.00022
Mean	0.00084	0.00084	–	0.0046	0.00016
MC π	0.00024	0.00025	0.0046	–	0.0002
SC	0.00023	0.00022	0.00016	0.0002	–

Table 15.4. *Significance matrix obtained from the applied K-S test (case mutual information measure)*

	Entropy	χ^2	Mean	MC π	SC
Entropy	–	0	0.59	0.047	0.32
χ^2	0	–	0.22	0.26	0.23
Mean	0.59	0.22	–	2.5×10^{-7}	0.22
MC π	0.047	0.26	2.5×10^{-7}	–	0.80
SC	0.32	0.23	0.22	0.80	–

Finally, in [Figure 15.1](#) the distribution of the p -values is represented. This graph can be an intuitive help, especially for high correlations, but on many occasions, it is not able to distinguish possible dependencies. Again, it is possible to see the almost perfect correlation between entropy and χ^2 . However, the rest of the correlations are no longer distinguished.

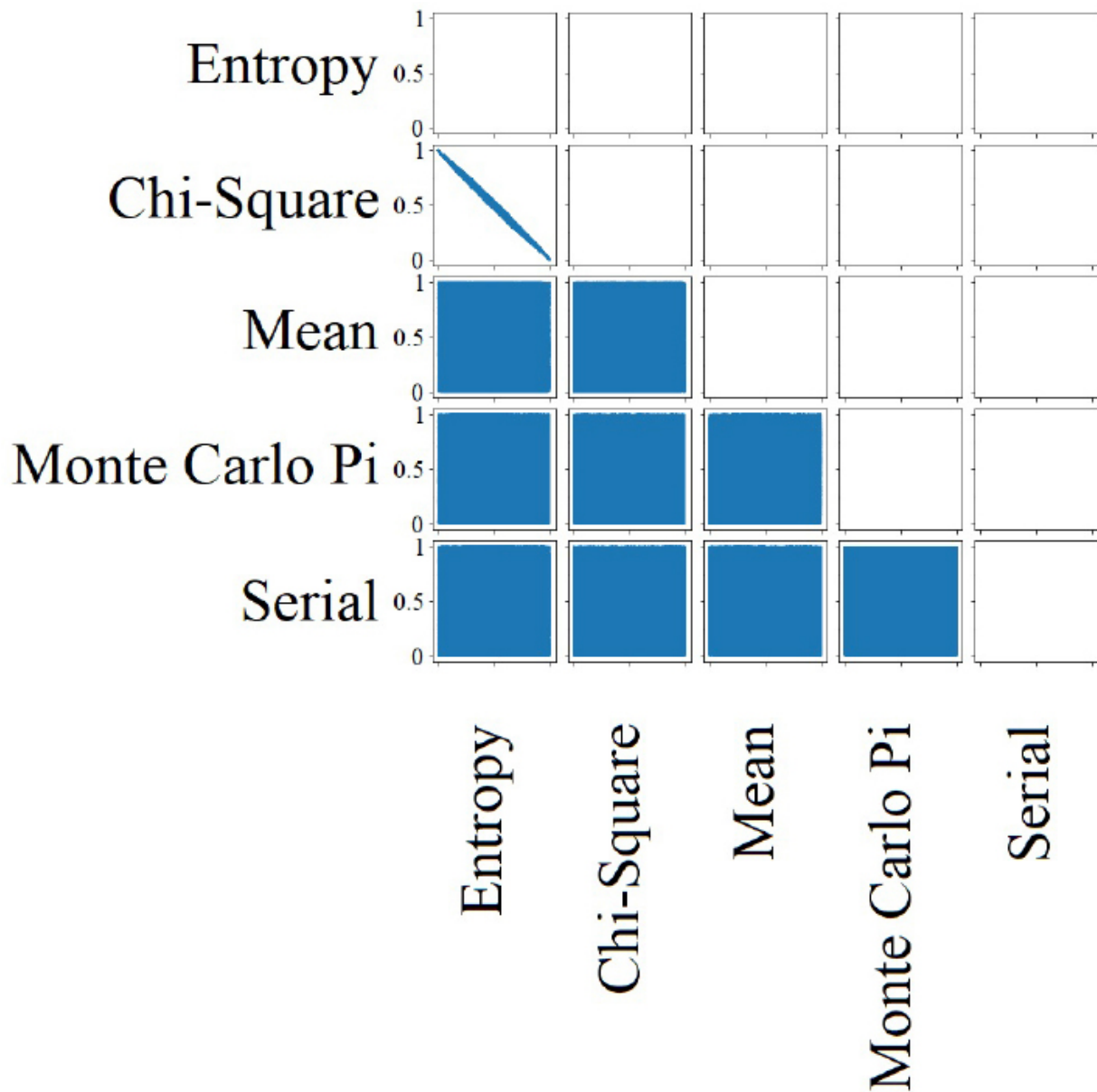


Figure 15.1. *Dispersion matrix of the p-values*

15.3. Conclusion

Observing the results, it is possible to affirm the existence of a strong interrelation between the entropy tests and χ^2 . This relationship is very evident in the Pearson's correlation matrix, indicating that this dependence is of a linear type. The measure of mutual information, in the specific analysis of this battery, does not provide additional significant results, so that no nonlinear dependencies between the tests are observed.

In view of the design of a battery and the need to eliminate redundancies between tests, the entropy and χ^2 tests are the ones that should be carefully analyzed when considering the elimination of one of them. From a computational point of view, both tests have the same complexity: in both cases, the occurrences of each value between 0 and 255 are calculated ($O(n)$, linear in the size n of the sequence), and then operations are performed on these 256 frequencies ($O(1)$). The choice between the two will depend on whether the user wants the value of a statistic (with the possibility of obtaining a p -value) or a measure of entropy. If the use is for purely statistical purposes, the choice of the χ^2 statistic would be appropriate.

Unlike the work of Karell-Albo et al. ([2020](#)), in which the use of the mutual information measure was able to detect relationships not previously detected by the Pearson's correlation between tests of the NIST battery, for the case analyzed in the present work, in the ENT battery the mutual information measure has not detected relationships other than those detected by Pearson's correlation coefficient, thus confirming that the detected relationship is linear.

15.4. References

- Almaraz Luengo, E. and García Villalba, L.J. (2021). Recommendations on statistical randomness test batteries for cryptographic purposes. *ACM Computing Surveys*, 54(4). doi: 10.1145/3447773.
- Almaraz Luengo, E., Alaña, B., García Villalba, L.J., Hernández-Castro, J. (2022a). Weaknesses in ENT battery design. *Applied Sciences*, 12(9), 1–7. doi: 10.3390/app12094230.

- Almaraz Luengo, E., Alaña, B., García Villalba, L.J., Hernández-Castro, J., Hurley-Smith, D. (2022b). StringENT test suite: ENT battery revisited for efficient p-value calculation. *Journal of Cryptographic Engineering*, 13, 235–249. doi: 10.1007/s13389-023-00313-5.
- Almaraz Luengo, E., Leiva, M., García Villalba, L.J., Hernández-Castro, J., Hurley-Smith, D. (2022c). Critical analysis of hypothesis tests in Federal Information Processing Standard (140-2). *Entropy*, 24(5), 1–16. doi: 10.3390/e24050613.
- Brown, R.G., Eddelbuettel, D., Bauer, D. (2014). Dieharder: A random number test suite (version 3.31.1) [Online]. Available at: <https://webhome.phy.duke.edu/~rgb/General/dieharder.php>.
- Burciu, P. and Simion, E. (2019). A systematic approach of NIST statistical tests dependencies. *Journal of Electrical Engineering, Electronics, Control and Computer Science*, 5(1), 1–6 [Online]. Available at: <https://jeeccs.net/index.php/journal/article/view/113>.
- Cover, T.M. and Thomas, J.A. (2006). *Elements of Information Theory*, 2nd edition. John Wiley & Sons, Hoboken, NJ.
- Doğanaksoy, A., Ege, B., Muş, K. (2008). Extended results for independence and sensitivity of NIST randomness tests. In *Proceedings of the Information Security and Cryptography Conference*, Ankara.
- Doğanaksoy, A., Sulak, F., Uguz, M., Seker, O., Akcengiz, Z. (2017). Mutual correlation of NIST statistical randomness tests and comparison of their sensitivities on transformed sequences. *Turkish Journal of Electrical Engineering & Computer Sciences*, 25, 655–665.
- Evans, D.L., Bond, P., Bement, A. (2002). FIPS Pub 140-2: Security Requirements for Cryptographic Modules. Federal Information Processing Standards Publication.
- Fan, L., Chen, H., Gao, S. (2013). A general method to evaluate the correlation of randomness tests. In *14th International Workshop*

on Information Security Applications, 52–62, Springer, Berlin, Heidelberg. doi: 10.1007/978-3-319-05149-9_4.

- Georgescu, C., Simion, E., Nita, A.P., Toma, A. (2017). A view on NIST randomness tests (in) dependence. In *Proceedings of the IEEE 2017 9th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, Targoviste.
- Hellekalek, P. and Wegenkittl, S. (2003). Empirical evidence concerning AES. *ACM Transactions on Modeling and Computer Simulation*, 13(4), 322–333.
- Hernández-Castro, J. and Barrero, D.F. (2017). Evolutionary generation and degeneration of randomness to assess the independence of the ENT test battery. In *2017 IEEE Congress on Evolutionary Computation (CEC)*, San Sebastián.
- Karell-Albo, J.A., Legón-Pérez, C.M., Madarro-Capó, E.J., Rojas, O., Sosa-Gómez, G. (2020). Measuring independence between statistical randomness tests by mutual information. *Entropy*, 22(7), 741.
- L’Ecuyer, P. and Simard, R. (2007). Testu01: A c library for empirical testing of random number generators. *ACM Transactions on Mathematical Software*, 33(4), 22:1–22:40. doi: 10.1145/1268776.1268777
- Rukhin, A.L., Soto, J., Nechvatal, J.R., Smid, M.E., Barker, E., Leigh, S., Levenson, M., Vangel, M., Banks, D., Heckert, A. et al. (2010). SP 800-22 Rev. 1a. A statistical test suite for random and pseudorandom number generators for cryptographic applications. Technical Report, Gaithersburg, MD.
- Sulak, F., Uguz, M., Koçak, O., Dog̃anaksoy, A. (2017). On the independence of statistical randomness tests included in the nist test suite. *Turkish Journal of Electrical Engineering & Computer Sciences*, 25, 3673–3683.
- Turan, M.S. and Dog̃anaksoy, A.S.B. (2008). On independence and sensitivity of statistical randomness tests. In *Sequences and Their*

Applications - SETA 2008, Golomb, S.W., Parker, M.G., Pott, A., Winterhof, A. (eds). Springer, New York.

Walker, J. (2008). ENT: A pseudorandom number sequence test program [Online]. Available at:
<https://www.fourmilab.ch/random/>.

Notes

Chapter written by Elena ALMARAZ LUENGO.

1 Pearson's correlation coefficient of two random variables X and Y is $\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$ where σ_X and σ_Y are the standard

deviations of X and Y , and $Cov(X, Y)$ is the covariance between X and Y . $\rho_{X,Y} \in [-1, 1]$, the greater the absolute value the more linear dependence between X and Y . Its sign determines whether this dependence is positive or negative. This measure can also be applied to random samples $X = \{x_1, \dots, x_n\}$ and

$Y = \{y_1, \dots, y_n\}$, in this case, the sample Pearson's correlation coefficient is

$$r_{X,Y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}},$$
 where $\bar{x}$ and $\bar{y}$ are the arithmetic means of X and Y , respectively.

2 Mutual information of two random variables X and Y taking values in S_x and S_y , respectively, is defined as

$$I(X, Y) = \int_{S_x} \int_{S_y} p(x, y) \log \left(\frac{p(x, y)}{p(x)p(y)} \right) dx dy,$$
 with $p(x, y)$ being the joint density and $p(x)$ and $p(y)$ the marginals, and, in its discrete form as:
$$I(X, Y) = \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \log \left(\frac{p(x_i, y_j)}{p(x_i)p(y_j)} \right).$$

16

Binary Regression and Non-monotonic Link Functions

Binary regression models play a crucial role in analyzing data where the outcome variable has two possible states. Traditional logistic and probit models employ monotonic link functions, but many real-world datasets exhibit non-monotonic relationships between predictors and binary outcomes, violating the assumptions of these standard approaches. To overcome this limitation, this chapter presents a novel non-monotonic link function for maximum likelihood estimation (MLE). The proposed function is designed to capture complex, curvilinear relationships while retaining interpretability and computational efficiency. The parameters of the proposed link function are discussed in detail, with each parameter providing unique control over the shape and behavior of the probability response. This flexibility allows the model to adapt to diverse datasets with non-monotonic structures, addressing the limitations of traditional monotonic link functions and nonparametric approaches. The advantages of the proposed method are demonstrated through comprehensive comparisons with traditional logistic regression, the cumulative distribution function mixture model (CDFMM) and local logistic regression on widely used benchmark datasets. The proposed method is particularly well suited for applications in fields such as healthcare, social sciences and ecology, where non-monotonic relationships are prevalent. This chapter not only introduces a new methodological tool for binary regression but also provides a robust framework for future research on non-monotonic link functions, including potential extensions to high-dimensional data and integration with machine learning frameworks. The proposed method represents a significant advancement in binary regression modeling, offering a practical and interpretable solution for analyzing complex, non-monotonic relationships in real-world data.

16.1. Introduction

Binary regression models are widely used in statistics and machine learning to analyze binary outcomes, such as success/failure or presence/absence. These models typically rely on monotonic link functions, such as those used in logistic or probit regression, which assume that the probability of the outcome consistently increases or decreases with predictor values. While effective in many cases, these models face challenges when dealing with non-monotonic relationships, where the probability of the outcome varies in more complex ways, for example, increasing up to a certain point and then decreasing, or exhibiting U-shaped or multi-peaked patterns (Gheno [2025](#)). Such non-monotonic patterns frequently occur in real-world data (Nottingham et al. [2000](#); Gheno [2025](#)). In healthcare, for instance, both low and high insulin levels may correspond to a higher risk of diabetes, producing a U-shaped relationship. Similarly, in ecology, species often thrive only within specific environmental ranges. In these contexts, monotonic models may produce poor fits and misleading results due to their rigid structure and reliance on strong parametric assumptions. To overcome these limitations, non-monotonic link functions provide a more flexible alternative (Nottingham et al. [2000](#); Gheno [2025](#)). These functions enable models to capture curvilinear and asymmetric relationships between predictors and outcomes, enhancing fit and predictive accuracy – particularly in datasets with skewed or heavy-tailed distributions. By adapting more effectively to the underlying data structure and avoiding assumptions of a uniform direction of change, non-monotonic link functions better accommodate complex patterns (Nottingham et al. [2000](#); Gheno [2025](#)). Real-world datasets often exhibit such complex relationships, where outcome probabilities increase up to a point before decreasing or display U-shaped or S-shaped trends. Traditional models may misestimate probabilities under these circumstances, leading to poor performance and incorrect conclusions. Flexible link functions reduce bias and improve generalization by accounting for skewness and outliers commonly present in data. Although various methods, such as the Robbins–Monro procedure (Nottingham et al. [2000](#); Gheno [2025](#)), kernel regression and local logistic regression (called also local logit regression), have been proposed to model non-monotonic data, these

approaches have notable drawbacks. They can be computationally inefficient, prone to boundary issues, susceptible to overfitting and impractical for large datasets due to the need to fit multiple local models. For example, local logistic regression often struggles near data boundaries and requires fitting many individual models, limiting its scalability. Non-monotonic link functions therefore offer a promising direction for improving the modeling of complex binary outcome data across diverse fields, including healthcare, economics, social sciences and ecology, where relationships between predictors and outcomes frequently deviate from simple monotonic patterns. This chapter analyzes a new non-monotonic link function that overcomes these limitations. The new non-monotonic link function proposed by Gheno ([2025](#)) captures complex patterns, such as inflection points and U-shapes, while remaining interpretable. The proposed approach is evaluated against standard logistic regression, the CDFMM model (a previously proposed non-monotonic method) (Nottingham et al. [2000](#), [2001](#)) and local logistic regression using the Cancer dataset. The results demonstrate the method's superior ability to model complex binary outcomes, highlighting its potential for broader application in fields where non-monotonicity is prevalent. Other examples are in Gheno's paper (Gheno [2025](#)).

16.2. Proposed method

The method proposed by Gheno ([2025](#)) introduces a non-monotonic link function that addresses the limitations of traditional approaches. By offering greater flexibility, it effectively captures complex relationships between predictors and binary outcomes, including curvilinear trends and inflection points. Unlike monotonic link functions, this approach accommodates nonlinear patterns while maintaining interpretability. Additionally, it enhances predictive performance across diverse datasets. The non-monotonic link function is defined by the following equations:

$$f_{1i} = \frac{1}{1 + \gamma |x_i|} e^{a + \beta x_i} \quad [16.1]$$

$$f_{2i} = \frac{\gamma |x_i|}{1 + \gamma |x_i|} \left(\frac{2x_i + \omega}{x_i^2 + \tau + 1} \right)^2 \quad [16.2]$$

$$p_i = \frac{f_{1i} + f_{2i}}{1 + f_{1i} + f_{2i}} \quad [16.3]$$

with $i = 1, \dots, n$; $\tau > 0$ and $\gamma > 0$. The function takes values within the interval $(0,1)$. The exponential part of f_{1i} is monotonic with respect to x . The square part of f_{2i} is not monotonic with respect to x . If γ increases, the weight of the non-monotonic part increases, while if γ decreases, the weight of the monotonic part increases. When $x \rightarrow \pm\infty$ and γ are finite, the limit of the function f_{2i} is 0, so the only function that determines p_i is f_{1i} and, in particular, β . If β is positive, p_i goes from 0 to 1 when x increases; if β is negative, p_i goes from 1 to 0 when x increases. For this reason, when γ is not very large, the U-shape of the probability function may only appear locally, as observed in the diabetes dataset analyzed by Gheno (2025). If both β and ω are equal to 0, the shape of p_i may become symmetric with respect to 0, while still remaining non-monotonic. If only β is equal to 0, when $x \rightarrow \pm\infty$, p_i tends to 0. The parameters of the proposed non-monotonic link function, denoted as $\theta = (\alpha, \beta, \gamma, \tau, \omega)$, are estimated using the MLE method. While parameter estimation is stable and relatively insensitive to initial conditions, certain types of data (such as sinusoidal datasets) can lead to instability (Gheno 2025). In such cases, stability can often be improved by introducing additional constraints on model parameters (Gheno 2025). For instance, enforcing a higher bound on the parameter gamma (e.g. $\gamma > 0.1$ or higher) may be necessary. Stability problems can occur for very high estimated γ or very small estimated γ . To rigorously assess estimation stability, it is important to examine both the Hessian matrix and the Fisher Information Matrix, as discussed in section 16.4. The Fisher information matrix is commonly approximated using the Hessian of the negative log-likelihood, which corresponds to the negative of the Hessian of the log-likelihood (Nottingham et al. 2000; Gheno 2025). This approximation is frequently used in

practice because it is computationally tractable, especially when the exact Fisher information is difficult to compute.

The first derivative is also analyzed to understand the role of each parameter. The behavior of parameter α is characterized by $\partial p_i / \partial \alpha > 0$, meaning that the derivative is always positive. The constant positivity of the derivative with respect to α means that α acts as a sort of intercept that uniformly raises the probability p_i for all values of x . This has several practical implications in real-world applications. This result also holds for the logit link function. For the new link function and for the logit link function, $p_i = p(x_i = 0) = \exp(\alpha) / [1 + \exp(\alpha)]$. Of course, if $\alpha \rightarrow +\infty$, p_i tends to 1. Because p_i represents a probability, α could control a baseline risk or a starting probability independent of other variables. In a medical context, for example, an increase in α could reflect a higher baseline risk for a certain condition, regardless of variables such as age or level of exposure (represented by x). Another example could be in a sales model, where a higher value of α could indicate a higher baseline sales probability, which grows regardless of the type of product or market.

The parameter β influences p_i directionally: it increases the probability when $x > 0$ and decreases it when $x < 0$. This is typical of linear terms in regression models. This result also holds for the logit link function. The behavior of the derivative of p_i with respect to ω shows that $\partial p_i / \partial \omega > 0$ if $\omega > -2x$, and $\partial p_i / \partial \omega < 0$ otherwise. This is an interesting behavior because ω seems to influence the function critically according to an inequality condition with respect to x . The effect of ω changes according to this relationship, suggesting that ω has a shift effect on the shape of the function. When $\omega > -2x$, an increase in ω shifts the curve toward higher values, increasing the probability. If $\omega < -2x$, an increase in ω reduces the probability, possibly triggering an inversion of the function's shape. From a practical perspective, ω could model a transition parameter that changes the behavior of the probability according to the position of x . In a binary decision context, such as risk analysis or financial choices, ω could represent a critical factor that modifies the decision outcome according to a threshold value (e.g. an approval policy that changes according to a determined risk level). It is particularly useful

when trying to model phenomena with threshold effects or where the probability changes abruptly according to specific conditions. The behavior of p_i with respect to τ is characterized by $\partial p_i / \partial \tau < 0$, meaning that the derivative is always negative. It appears that the parameter τ consistently reduces the probability regardless of the value of x . This behavior suggests that τ has a “smoothing” or “flattening” effect on the curve. An increase in τ makes the probability p_i less sensitive to variations in x . This could be associated with a reduction in the influence of x on the probability. In an applied model, τ could be used to control the “grain” of the model. By increasing τ , one obtains a smoother curve that could be desirable when wanting to avoid peaks or excessive valleys. A good application example could be in the context of long-term forecasting systems where one wants to reduce the importance of sharp fluctuations in the data. The behavior of p_i with respect to γ shows that the derivative depends on the sign of a complex difference that involves exponential and rational terms: $[(2x + \omega)/(x^2 + \tau + 1)]^2 - \exp(\alpha + \beta x)$. The fact that the derivative depends on a difference of terms involving both rational and exponential components suggests that γ models the non-monotonicity of the function. The effect of γ will be influenced by the complex combination of these terms. This implies that γ does not act in a simple way but has a behavior that can amplify or dampen peaks and valleys in the curve, making it more complex and non-monotonic.

From a practical standpoint, γ can be useful for modeling complex dynamics. This parameter is important because it provides an indication of the model’s complexity, specifically the degree of nonlinearity involved. For instance, when comparing results between two groups, the group associated with a higher γ value demonstrates a more complex relationship between the variable x and the probability p than the group with a lower γ value.

In summary, it is important to give an interpretation of the parameters:

- Alpha: acts as an intercept term that shifts the entire probability curve upward or downward. It uniformly increases the probability π across all values of x .

- Beta: influences the direction of the curve, indicating whether the relationship between x and the probability is positive or negative.
- Gamma: modifies the overall shape of the curve, enabling more complex patterns such as U-shaped trends instead of a simple monotonic relationship.
- Tau: controls the sharpness of transitions in the probability curve as x varies, modulating how abruptly or smoothly the probability responds to changes in x .
- Omega: acts as a timing shift. Negative values delay the curve (shifting it toward higher x values), while positive values advance it (shifting toward lower x values), helping to capture lag or early-response effects in real-world phenomena.

Together, these parameters enable the model to represent rich, nonlinear dynamics in data more accurately. The parameter behaviors are coherent with the theoretical intuition for a nonlinear or non-monotonic model. Each parameter has a specific role in determining the shape of the curve p_i . Because the new link function seeks to model complex and dynamic phenomena, these parameters are useful for capturing directional effects, thresholds, curvature and non-monotonicity.

To apply asymptotic theory in MLE, certain regularity conditions must be satisfied (section [16.3](#)). One important consideration involves parameter constraints. When parameters such as γ and τ are restricted to be positive, estimating them directly can place their values on the boundary of the parameter space, which can violate the assumptions required for standard asymptotic inference. To address this, reparameterization is often used, transforming γ and τ into unconstrained parameters γ' and τ' : $\gamma = \exp(\gamma')$ and $\tau = \exp(\tau')$. This maps the constrained parameter space onto the real line, ensuring that the estimators lie in the interior of the domain and allowing the use of asymptotic results. Asymptotic theory can be applied when the true parameter values lie in the interior of the parameter space, the likelihood function is sufficiently smooth and differentiable, and the parameters are identifiable from the model. In addition, the Fisher information must be finite and nonsingular, and the data must

satisfy basic assumptions such as independence or appropriate dependence structures. When these conditions hold and the sample size is sufficiently large, the MLE is consistent and asymptotically normal. This permits the use of standard errors derived from the observed Fisher information matrix and enables valid statistical inference. Therefore, the transformations of γ and τ are essential not only for numerical stability and interpretability but also for ensuring that the regularity conditions required for asymptotic inference are met.

16.3. Methodology

The proposed link function is specifically designed to model binary outcome probabilities while capturing non-monotonic relationships between predictors and responses. Unlike traditional models that assume a consistent directional effect, this approach allows for more flexible, data-driven patterns (including inflection points and nonlinear trends). To assess its effectiveness, the method is compared against logistic regression, the CDFMM model (Chang and Bodt [1997](#); Nottingham et al. [2001](#)) and local logistic regression techniques (Nottingham et al. [2000](#), [2001](#)). Performance is evaluated using standard metrics, including the area under the curve (AUC).

Logistic regression, the most widely used model for binary classification, estimates outcome probabilities through a monotonic logistic link function and relies on MLE. It assumes a consistent increase or decrease in the probability of the outcome as predictor values change. The logistic function, which ensures this monotonic relationship, is defined as:

$$p_i^{Logit} = \frac{e^{\alpha_L + \beta_L x_i}}{1 + e^{\alpha_L + \beta_L x_i}} \quad [16.4]$$

For observations $i = 1, \dots, n$, the model estimates parameters by maximizing the log-likelihood function using iterative optimization techniques such as gradient ascent or the Newton–Raphson method. These parameter estimates define how the predictors influence the

outcome probability, constrained by the monotonicity inherent in the logistic function.

The CDFMM was the first method proposed to explicitly model non-monotonic relationships between a binary response variable and its predictors within the framework of Bernoulli regression. In traditional Bernoulli (or logistic) regression, the link function (typically the logit function) assumes a strictly monotonic relationship, meaning the probability of a positive outcome consistently increases or decreases with the predictor values. This assumption limits the model's ability to capture more complex, real-world patterns.

The CDFMM approach addresses this limitation by introducing a non-monotonic link function, constructed as a mixture of cumulative distribution functions (CDFs). This formulation allows the model to flexibly approximate a wide range of response patterns, including those with turning points or U-shaped trends, without sacrificing probabilistic interpretation of the model. By combining multiple CDFs, the CDFMM approach creates a richer and more adaptable functional form for the link, making it well suited for datasets where the effect of predictors on the outcome is not strictly linear or monotonic. Its link function is:

$$p_i^{CDFMM} = (1 - F_{Ti}) F_{Ii} + F_{Ti} F_{Si} \quad [16.5]$$

where

$$F_{wi} = \frac{1}{1 + e^{-\frac{x_i - \mu_w}{\sigma_w}}} \quad [16.6]$$

with $w = T, S, I$ and $i = 1..n$.

The parameters of the new logit regression method and the CDFMM approach are estimated using the MLE method. Logit regression respects the regularity conditions, but it is necessary to check if the new method also respects them. To ensure the validity of asymptotic results for MLE, standard regularity conditions must be satisfied. In the proposed model, the log-likelihood function is constructed from

a binary response model with a smooth, non-monotonic link function:

$$\ell(\theta) = \sum_{i=1}^n \left[y_i \log p(x_i, \theta) + (1 - y_i) \log (1 - p(x_i, \theta)) \right] \quad [16.7]$$

where $p(x_i, \theta)$ is a composition of exponential and rational functions. For the new method, the probability function $p(x_i, \theta)$ is strictly bounded between 0 and 1 for all values of x and admissible parameter values, ensuring that the logarithmic terms are always well defined. Furthermore, in the new method, the function $p(x_i, \theta) = p_i$ is continuously differentiable (infinitely differentiable in fact) with respect to all parameters, as it involves only smooth transformations of θ . To preserve these properties and to enforce positivity constraints, the parameters γ and τ are reparametrized as: $\gamma = \exp(\gamma')$ and $\tau = \exp(\tau')$, so that the parameter space remains open and differentiable. Both in the original and reparametrized models, the support of θ does not depend on θ , allowing the interchange of differentiation and integration required by regularity theory. These conditions imply that the log-likelihood function is continuously differentiable with respect to the parameters, and the expected score function satisfies:

$$E_{\theta_0} [\nabla \ell(\theta)] = 0 \quad [16.8]$$

allowing the use of Fisher information and standard asymptotic theory for inference.

Local logistic regression for binary outcomes adapts the concept of local regression to classification tasks by estimating the probability of class membership rather than predicting continuous values. Unlike standard logistic regression, which fits a single global model using a monotonic link function, this method constructs localized logistic models around each target point in the dataset. Specifically, it performs weighted logistic regression using a subset of nearby data points, where the weights are assigned through a kernel function that emphasizes closer observations. By estimating probabilities locally,

the method captures complex and non-monotonic relationships between predictors and the binary response, allowing for a flexible and adaptive decision boundary. This flexibility helps overcome the structural limitations of traditional logistic regression, particularly in settings where the effect of predictors changes across different regions of the input space. While the method retains interpretability and typically reduces bias near regions of high data density, its performance is highly sensitive to the choice of bandwidth and kernel parameters. Poor parameter selection can lead to overfitting or excessive smoothing. Nevertheless, local logistic regression is especially advantageous in scenarios where global monotonic models fail to capture nonlinear or region-specific patterns in the data.

16.4. Comparison

The new method is compared with the other methods using the Cancer dataset (Kaggle, license: public domain). The Cancer dataset contains diagnostic measurements related to breast cancer. Its sample size is 569. It includes features computed from digitized images of fine needle aspirates (FNA) of breast masses, along with a binary diagnosis indicating whether the tumor is malignant (M) or benign (B). Each row represents a distinct sample, with an identifier (id) and a target variable (diagnosis). The dataset comprises 30 numerical features derived from cell nuclei characteristics, categorized into three groups: mean values, standard errors (SE) and worst (largest observed) values. These features include measurements such as radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry and fractal dimension. For instance, radius_mean represents the average distance from the nucleus center to its perimeter, while fractal_dimension_worst quantifies the complexity of the most irregular cell boundaries observed in the sample. The relationship between the predicted probabilities of malignancy and the scaled fractal dimension (worst) can be interpreted as follows. The fractal_dimension_worst feature reflects the highest level of boundary irregularity among the observed nuclei. In medical diagnostics, more complex and irregular structures are often associated with malignant tumors. When a predictive model is

applied to this dataset, the output may include probabilities that estimate the probability of a tumor being malignant based on its features.

16.4.1. Sensitivity analysis of starting values

To demonstrate the stability of the parameter estimates, we examine how small changes in the starting values affect the final solutions. Specifically, we re-estimate the model many times, each time seeding the optimizer with slightly perturbed initial parameters (random noise added for variability). This procedure is applied to two competing link functions (the proposed non-monotonic link and the CDFMM link) using the L-BFGS-B algorithm (Byrd et al. [1995](#)) to minimize the negative log-likelihood under identical bounds and convergence criteria. The resulting estimates, summarized in Table 16.1, reveal a clear contrast: the CDFMM estimates fluctuate markedly with different starts, confirming the sensitivity previously reported (Nottingham et al. [2001](#); Gheno [2025](#)), whereas the proposed link yields virtually identical solutions across runs. These findings suggest that the CDFMM optimizer often settles at local optima, while the new link function consistently converges to (or is very near) the global maximum, underscoring its robustness to initialization choices. The stability of the new method is also evaluated by analyzing the Fisher information matrix (FIM) and the Hessian matrix of the negative log likelihood (Gheno [2025](#)). Numerical differentiation is employed to compute the gradient and FIM, a common approach in parameter estimation supported by numerous studies. The eigenvalues of the Fisher information matrix are [86.4596, 59.9501, 11.0486, 5.0726, 0.6257], with a determinant of 180,347.25, indicating a well-conditioned matrix with no near-singularity (smallest eigenvalue > 0.001). This ensures that all parameters are identifiable, with low estimation uncertainty. Similar findings are observed for the Hessian matrix, which is full rank (5), with eigenvalues [85.5363, 58.4830, 11.6552, 6.3363, 1.3207] and a determinant of 487,916.3. The proximity of these matrices suggests that the likelihood is well approximated by a quadratic form, validating the numerical stability and reliability of the estimation procedure. Overall, these results demonstrate that the estimation procedure is both stable and reliable.

Table 16.1. *Estimated parameters using different starting values*

New model					
	1	2	3	4	5
Alpha	−0.46	−0.46	−0.46	−0.46	−0.46
Beta	0.96	0.96	0.96	0.96	0.96
Gamma	0.72	0.72	0.72	0.72	0.72
Omega	0.59	0.59	0.59	0.59	0.59
Tau	0.36	0.36	0.36	0.36	0.36
CDFMM					
	1	2	3	4	5
muT	0.86	0.26	0.25	0.26	0.26
sigmaT	3.09	0.37	0.37	0.36	0.35
muI	7,705.74	538.62	4,980.76	31,111.06	8,756.99
sigmaI	11,369.44	416.20	3,828.55	24,526.42	6,962.03
muS	0.23	−1.50	−1.39	−1.34	−1.34
sigmaS	0.38	3.59	3.48	3.45	3.45

16.4.2. Nonlinearity

[Figure 16.1](#) illustrates the predicted probabilities of a binary outcome as a function of the scaled fractal dimension (worst), comparing four different models: the logit model, the local logistic model, the CDFMM model and a newly proposed model. The logit model, which uses a monotonic link function, produces a smooth and strictly increasing probability curve, reflecting its assumption that the likelihood of the outcome consistently rises with the predictor. However, this model fails to capture more complex structures in the data, particularly in regions where the relationship is non-monotonic. In contrast, the local logistic model shows a highly flexible curve that adjusts to local variations in the data. It captures a dip in predicted probabilities at lower values of the predictor, indicating sensitivity to non-monotonic behavior. While this flexibility allows the model to follow intricate patterns, it may also introduce instability near the data boundaries, potentially leading to

overfitting. The CDFMM model introduces a non-monotonic link function and captures some degree of curvature, although its response is smoother and less adaptive than that of the local logistic model. It underrepresents variation in regions where the predictor–response relationship is more dynamic, possibly due to limitations in its estimation procedure. The newly proposed model demonstrates a well-balanced performance. It reflects the non-monotonic trend present in the data with greater stability and smoother transitions than the local logistic model while offering more flexibility than the logit and CDFMM models. This allows it to adapt to complex patterns without overreacting to local noise. The data points shown along the top and bottom margins of the plot indicate the observed binary outcomes, supporting the presence of nonlinear and non-monotonic relationships between the predictor and the outcome variable.

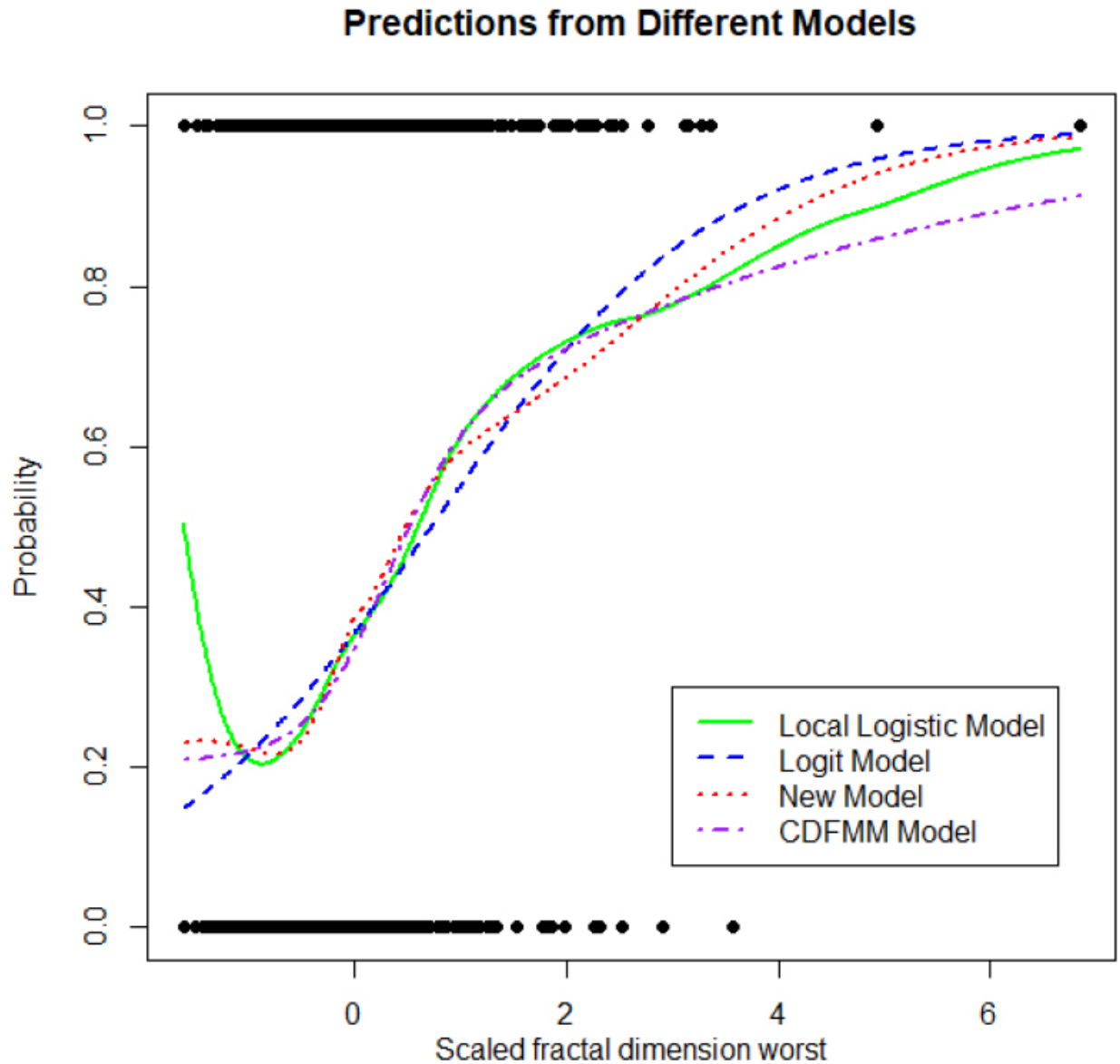


Figure 16.1. *Different shapes*

16.4.3. ROC curve comparison

[Figure 16.2](#) displays the receiver operating characteristic (ROC) curves for four classification models: the logit model, the new model, the CDFMM model and the local logistic model. The ROC curve represents the relationship between the true positive rate (sensitivity) and specificity across different classification thresholds, providing a visual evaluation of model performance in binary classification. In this representation, the x-axis displays specificity, decreasing from left to right, while the y-axis shows sensitivity. The

conventional interpretation remains valid, with the top-left corner indicating optimal performance where both sensitivity and specificity are maximized. The diagonal line maintains its representation of random classification performance. This presentation is mathematically equivalent to the standard ROC curve plotting sensitivity against one-specificity, with only the x-axis orientation differing. All performance metrics, including the area under the curve, retain their standard interpretations in this format. All models perform better than random classification, as indicated by their curves lying above the diagonal reference line (which represents the performance of a classifier with no discriminative ability). Among the models, the New Model (red curve) achieves the best overall performance, consistently maintaining a higher true positive rate across most specificity levels. This suggests that it is more effective at identifying positive cases without sacrificing specificity. The local logistic (green curve) closely reflects its locally adaptive nature, which may offer advantages in certain regions of the predictor space but can lead to less stable global performance. The CDFMM Model (purple curve) and Logit model (blue curve) trail slightly behind the others, particularly at smaller specificity. Their lower sensitivity at comparable specificity indicates reduced discriminative power relative to the other models. Overall, the ROC analysis supports the superior performance of the New Model, which appears to offer a better balance between sensitivity and specificity across a range of thresholds.

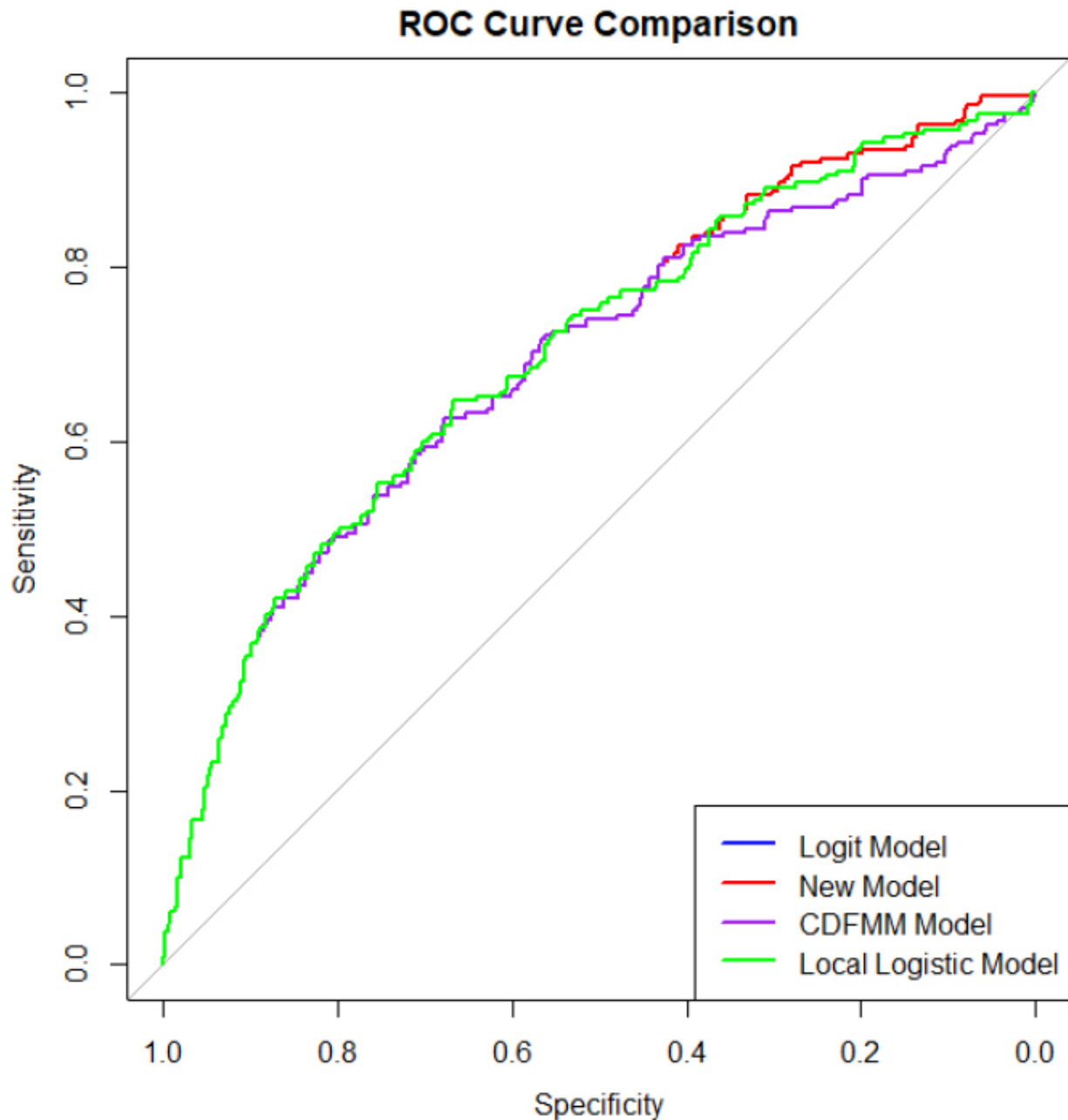


Figure 16.2. *ROC curve comparison*

16.4.4. AUC area comparison

In evaluating classification models, the area under the receiver operating characteristic curve (AUC) is a widely used metric to assess how well a model can distinguish between positive and negative classes. Higher AUC values indicate better discriminative ability, with a value of 0.5 representing random chance and 1 representing

perfect classification. The following results compare four different models based on their AUC scores.

- The logit model and the CDFMM model both achieved an AUC of approximately 0.686, indicating moderate and similar classification performance.
- The New Model achieved the highest AUC of 0.699, demonstrating superior ability in distinguishing between classes compared to the other models.
- The local logistic model also performed well, with an AUC of 0.698, which is very close to that of the new model, indicating strong predictive power.

Although the numerical differences between the top two models are small, the New Model still shows the best overall performance based on this metric. Overall, the New Model is the best performing model in terms of classification accuracy, followed closely by the local logistic model. The logit and CDFMM models show similar but comparatively lower predictive accuracy.

16.4.5. Inference

Table 16.2. Inference

New model					
Parameter	Estimate	SE	t value	p value	
Alpha	−0.46	0.21	−2.19	0.03	*
Beta	0.96	0.22	4.33	1.52E-05	***
Gamma'	−0.32	0.93	−0.35	0.73	
Omega	0.59	0.36	1.63	0.1	
Tau'	−1	1.62	−0.62	0.54	
Logit					
alpha_L	−0.55	0.09	−6.00	2.43E-09	***
beta_L	0.75	0.10	7.13	9.93E-13	***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.					

The results in [Table 16.2](#) provide insights into the statistical significance and performance of the parameters in both the proposed non-monotonic link function (New Model) and the traditional logistic regression (Logit). For the New Model, the parameters α and β are statistically significant, with p values of 0.03 and 1.52E-05, respectively. The parameter β is positive (0.96), indicating a generally positive relationship: an increase in the fractal dimension in the worst-case scenario tends to increase the probability of cancer. This finding is consistent with theoretical expectations and supported by the logistic regression result ($\beta_L > 0$). It should be noted that “tends to” refers to a general tendency rather than a strictly monotonic relationship. In contrast, ω does not reach statistical significance (p value > 0.05), indicating that its influence on the model may be negligible. If the estimates for γ' and τ' are not significantly different from zero (as suggested by their large p values in [Table 16.2](#)), this implies that their exponentiated forms, γ and τ , are not significantly different from 1. Since γ controls the weight of the non-monotonic component in the link function, a value close to 1 suggests a contribution from both the monotonic (exponential) and non-monotonic (quadratic) parts of the model. The parameter τ influences the smoothness and inflection points of the probability curve. A value near 1 suggests moderate curvature. While the logit model’s parameters are highly significant, its rigid monotonicity limits its ability to capture nonlinear trends. The New Model with $\gamma \approx 1$ and $\tau \approx 1$ offers a more flexible structure, as evidenced by its superior AUC (ROC analysis).

16.5. Conclusion

This study introduces a novel non-monotonic link function for binary regression, designed to address the limitations of traditional monotonic models in capturing complex relationships between predictors and outcomes. The proposed method demonstrates flexibility in modeling curvilinear and asymmetric patterns while maintaining computational stability and interpretability. The analysis reveals that the model performs robustly even when key parameters, such as γ and τ , are close to their default values, suggesting inherent adaptability to diverse datasets. Comparative

evaluations against logistic regression, the CDFMM and local logistic regression highlight the method's superior performance, particularly in scenarios where non-monotonicity is present but not extreme. The higher AUC score and stable parameter estimates further support its practical utility. While the current results indicate that the model does not require highly tuned parameters for effective performance, future research could explore its behavior in datasets with stronger non-monotonic signals or under different regularization schemes. Extensions to high-dimensional settings or integration with machine learning frameworks may also broaden its applicability. In summary, this work advances binary regression methodology by introducing a flexible, interpretable and efficient tool for analyzing complex real-world data. Its relevance is particularly notable in fields such as healthcare, ecology, and social sciences, where nonlinear relationships are common. The proposed link function offers a promising direction for future developments in non-monotonic modeling.

16.6. References

- Byrd, R.H., Lu, P., Nocedal, J., Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5), 1190–1208.
- Chang, A.L. and Bodt, B.A. (1997). JTCG/AS interlaboratory ballistic test program—final report. US Army Research Laboratory Technical Report No. ARL-TR-1577.
- Gheno, G. (2025). Non-monotonic link function: A new method for binary regression in maximum likelihood estimation (MLE). In *Artificial Intelligence Applications and Innovations. AIAI 2025. IFIP Advances in Information and Communication Technology*, Maglogiannis, I., Iliadis, L., Andreou, A., Papaleonidas, A. (eds). Springer, Cham.
- Guo, S. and Spall, J.C. (2019). Relative accuracy of two methods for approximating observed Fisher information. In *Data-Driven Modeling, Filtering and Control: Methods and Applications*, Novara, C. and Formentin, S. (eds). IET Press, London, 189–211.

Nottingham, Q.J., Birch, J.B., Bodt, B.A. (2000). Local logistic regression: An application to army penetration data. *Journal of Statistical Computation and Simulation*, 66(1), 35–50.

Nottingham, Q.J., Birch, J.B., Bodt, B.A. (2001). Modeling non-monotonic dose–response curves. Army Research Laboratory.

Note

Chapter written by Gloria GHENO.

List of Authors

Elena ALMARAZ LUENGO

Group of Analysis, Security and Systems (GASS)

and

Department of Statistics and Operational Research

Faculty of Mathematical Science Complutense University of Madrid

and

Interdisciplinary Mathematics

Institute (IMI)

Spain

Volodymyr ANISIMOV

Center for Design & Analysis

Amgen

London

UK

Dominika Sónak BALLOVÁ

Department of Mathematics and Descriptive Geometry

Faculty of Civil Engineering

Slovak University of Technology

Bratislava

Slovakia

James R. BOZEMAN

American University of Malta

Malta

Michaela ČERVEŇANSKÁ

Department of Hydraulic Engineering

Faculty of Civil Engineering

Slovak University of Technology

Bratislava

Slovakia

Mathieu CHALIFOUR

Department of Mathematics and Statistics

MacEwan University

Edmonton
Alberta
Canada

Tomáš FIALA
Department of Demography
Faculty of Economics and Statistics
Prague University of Economics and Business
Czechia

Matthew FISCHER
Department of Computing Science
University of Alberta
Edmonton
Canada

Christos FLOROS
Department of Accounting and Finance
Hellenic Mediterranean University
Heraklion
Crete
Greece

Jesús E. GARCÍA
Department of Statistics
University of Campinas
Sergio Buarque de Holanda
Brazil

Gloria GHENO
Ronin Institute 2.0
Sacramento
USA

V.A. GONZÁLEZ-LÓPEZ
Department of Statistics
University of Campinas
Sergio Buarque de Holanda
Brazil

Leo GORTZAK
Center for Design & Analysis
Amgen

London
UK

Giseon HEO
School of Dentistry
University of Alberta
Edmonton
Canada

Jelena JOCKOVIĆ
Faculty of Mathematics
University of Belgrade
Serbia

Jana KALICKÁ
Institute of Computer Science and Mathematics
Faculty of Electrical Engineering and Information Technology
Slovak University of Technology
Bratislava
Slovakia

Emmanuel KARAPIDAKIS
Electrical and Computer Engineering Department
Hellenic Mediterranean University
Heraklion
Greece

Adam B. KASHLAK
Department of Mathematical and Statistical Sciences
University of Alberta
Edmonton
Canada

Aleksei N. KURBATSKII
Moscow School of Economics
Department of Econometrics and Mathematic Methods of
Economics
Lomonosov Moscow State University
Russia

Jitka LANGHAMROVÁ
Department of Demography
Faculty of Economics and Statistics

Prague University of Economics and Business
Czechia

George MATALLIOTAKIS
Department of Accounting and Finance
Hellenic Mediterranean University
Heraklion
Crete
Greece

Catherine MICHALOPOULOU
Department of Social Policy
Panteion University of Social and Political Sciences
Athens
Greece

Marina V. MIRONENKOVA
Moscow School of Economics
Department of Econometrics and Mathematic Methods of
Economics
Lomonosov Moscow State University
Russia

Maria Manuela NEVES
CEAUL & Instituto Superior de Agronomia
Universidade de Lisboa
Portugal

Marios NIKOLOGIANIS
Business Administration and Tourism Department
Hellenic Mediterranean University
Heraklion
Greece

Neir Antunes PAES
Health and Decision Modelling Postgraduate Program
Federal University of Paraíba
Brazil

Fotis PAPATHEOFANOUS
Department of Accounting and Finance
Hellenic Mediterranean University
Heraklion

Crete
Greece

Rozane PEREIRA DE SOUSA
Health and Decision Modelling
Postgraduate Program
Federal University of Paraíba
Brazil

Vivian Conceição Alves Leite
PEREIRA DO LAGO
Health and Decision Modelling Postgraduate Program
Federal University of Paraíba
Brazil

Dora PRATA GOMES
Center for Mathematics and Applications (NOVA Math) and
Department of Mathematics
NOVA School of Science and Technology (NOVA FCT)
Portugal

D.S. SANTOS
Department of Statistics
University of Campinas
Sergio Buarque de Holanda
Brazil

Christos H. SKIADAS
Data Analysis and Forecasting Laboratory
Technical University of Crete
Greece

Konstantinos A. TASIAS
Department of Mechanical Engineering
University of Western Macedonia
Greece

Bojana TODIĆ
Faculty of Mathematics
University of Belgrade
Serbia

Konstantina TSIOTA
Department of Mechanical Engineering
University of Western Macedonia
Greece

Jana VRABCOVÁ
Department of Demography Faculty of Economics and Statistics
Prague University of Economics and Business
Czechia

Aggeliki YFANTI
Research Centre for Greek Society
Academy of Athens
Greece

Georgios C. ZACHOS
Department of Accounting and Finance
School of Economics and Administrative Sciences
University of Ioannina
Greece

Index

A, B

age-based PM, [204](#)
assignable causes, [205](#), [208](#)
Bayes factor(s), [32](#), [34](#), [38](#), [39](#)
BESS hybrid, [92](#), [97](#), [101](#)
bidding behaviors, [92](#)

C, D

clustering, [171](#)
 quality, [171](#)
completed cohort fertility, [7](#), [10](#)
country
 -specific patterns, [197](#)
 donor, [190](#)
demographic analysis, [200](#)
DFA, [110](#), [112–117](#), [122](#), [123](#), [128–143](#)
 Hurst, [112](#), [113](#), [115–117](#), [122](#), [123](#), [128–142](#)
 Multifractal (MF-DFA), [110](#)
dynamic relationship, [126](#)

E, F

electroencephalogram (EEG), [33](#)

EPU

UK, [111–117](#), [122](#), [123](#), [136–143](#)

US, [103–105](#), [110](#), [112](#), [113](#), [115–121](#), [124–126](#)

equitable normal pension age (ENPA), [3](#)

European Social Survey (ESS), [55](#)

datasets, [55](#)

extreme precipitation, [155](#), [160](#)

forecasting, [71–73](#)

G, H

government measures, [56](#), [120](#), [123](#)

groundwater-centric, [163](#)

high RES penetration, [92](#), [100](#)

human development index (HDI), [19](#), [23](#)

Hurst exponent(s), [103–110](#), [112–117](#), [122](#), [123](#), [129–134](#), [136–142](#)

I, L

indicators

demographic, [189](#)

quality-of-life, [190](#)

logistic

regression, [245](#), [246](#), [250](#), [252](#), [258](#)

technique, [250](#)

long memory, [103–109](#), [113](#), [116](#), [122](#), [126](#)

M, N

Markov

chain(s), [44–47](#), [175](#), [177](#), [178](#), [183](#), [187](#)

decision process, [204](#), [209](#), [215](#)

techniques, [177](#)

max-stable processes, [147](#), [149–151](#)

maximum likelihood estimation (MLE), [245](#)

metric(s), [43](#), [44](#), [46](#), [47](#), [51](#), [52](#)

model selection, [32](#), [38](#)

Bayesian, [31](#)

mutual information measure, [239–242](#)

negative binomial distribution(s), [34–37](#), [39](#)

O, P

old age, [18](#), [26](#)

adjusted (AOADR), [3](#), [6](#)

dependency ratio, [7](#), [12–14](#)

path analysis, [55](#), [60](#), [65](#)

model, [55](#), [60](#)

possible stopping time, [83](#)

pre-screening, [71–75](#), [77](#), [78](#), [80–83](#), [86](#)

preventive maintenance, [203](#), [213](#)

interval, [213](#)

R, S

R software, [147](#), [154](#)

reset(s) partial, [178](#), [182](#), [184](#), [185](#), [187](#), [188](#)

risk measure(s), [217–219](#), [221–225](#), [227](#), [229–233](#)

coherent, [223](#), [224](#), [230](#)

concept, [222](#)

statistical

battery, [237](#)

process control (SPC), [203](#)

significance, [240](#)

T, W, Z

time series, [34](#), [93](#), [103–113](#), [116](#), [122](#), [123](#), [147](#), [148](#), [155](#), [173](#)

trend(s), [163–168](#), [171](#), [172](#)

waiting time, [175](#), [177](#), [182–188](#)

Z-DNA conformation(s), [45](#), [51](#)

Other titles from



in

Innovation, Entrepreneurship and Management

2026

AOUINAÏT Camille, UZUNIDIS Dimitri

Testing Innovation Digital Modes and Models

(Innovation in Engineering and Technology Set – Volume 10)

GARIDI Sarah, SEBAI Jihane

The Patient Pathway: Building Partnerships in Healthcare

(Health and Innovation Set – Volume 9)

GILBERT Patrick, LE BOULAIRE Martine

The Hidden Faces of Innovation

(Technological Changes and Human Resources Set – Volume 6)

ORUEZABALA Gwenaëlle, ABBAD Hicham

Digital Strategies for Supply Chains: From Design to Deployment

SCHNEBELIN Eléonore

*Ecologization and Digitalization of Agriculture: Convergence or
Opposition?*

(Smart Innovation Set – Volume 45)

SEBAI Jihane, SZOSTAK Bérangère L.

*Social Innovation Through Design in Healthcare 1: Thinking,
Understanding and Implementing*

(Health and Innovation Set – Volume 7)

*Social Innovation Through Design in Healthcare 2: Rethinking and
Designing the Future*

(Health and Innovation Set – Volume 8)

2025

CHABAUD Didier, EYNAUD Philippe, RAULET-CROSET Nathalie
Reinventing Territories through Solidarity
(Solidarity-based Management and Organizations Set – Volume 1)

DIMOTIKALIS Yannis, SKIADIS Christos H.
Data Analysis and Related Applications 5: Models, Methods and Techniques
(Big Data, Artificial Intelligence and Data Analysis Set – Volume 13)

GORIA Stéphane
Serious Games and Innovation Gains: Think Innovation through Games
(Innovation in Engineering and Technology Set – Volume 9)

GRENIER Corinne, FRANKLIN-JOHNSON Elizabeth, CAJAIBA-SANTANA Giovany
Promoting a More Inclusive Society for Dependent or Disabled People: New Paradigms
(Health and Innovation Set – Volume 6)

KLESZCZOWSKI Julien, RAULET-CROSET Nathalie
Social Impact in Solidarity-based Organizations: Values and Assessment Practices
(Solidarity-based Management and Organizations Set – Volume 2)

2024

AÏT-EL-HADJ Smaïl
Environmental Transition and Technological Change
(Smart Innovation Set – Volume 43)

ALBERT-CROMARIAS Anne, ALBERTINI Thérèse, TERRAMORSI Patrice
Collective Dynamics and Territories: 9 Issues for Competitive, Innovative and Sustainable Territories
(Territorial Entrepreneurship and Innovation Set – Volume 3)

BRASSEUR Martine, BARTOLI Annie, CHABAUD Didier, GROUIEZ Pascal, ROUET Gilles

*Inclusive Territories 2: Role of Institutions and Local Actors
(Territorial Entrepreneurship and Innovation Set – Volume 2)*

COURIE-LEMEUR Aline

*Innovation, Collective Intelligence and Resiliency in Healthcare Organizations
(Health and Innovation Set – Volume 4)*

DARTIGUEPEYROU Carine, SALOFF-COSTE Michel

*Futures in Action: Strategic Anticipations and Deployments in Organizations Facing the Future
(Innovation and Technology Set – Volume 19)*

DELÉCOLLE Thierry, JACOB Florence, PRIM-ALLAZ Isabelle
New Frontiers of Customer Strategy: Managing Sustainable, Environmental and Ethical Transitions

DIMOTIKALIS Yannis, SKIADIS Christos H.

*Data Analysis and Related Applications 3: Theory and Practice – New Approaches
(Big Data, Artificial Intelligence and Data Analysis Set – Volume 11)*

*Data Analysis and Related Applications 4: New Approaches
(Big Data, Artificial Intelligence and Data Analysis Set – Volume 12)*

FLORA Luigi, GRENIER Corinne, PONISGNON Frédéric

*Experience in Healthcare Innovation: Fad or New Paradigm?
(Health and Innovation Set – Volume 5)*

MERLIN-BROGNIART Céline

*Functional Economy: Economic Models, Challenges and Innovation Dynamics
(Smart Innovation Set – Volume 44)*

PARMENTIER Guy

*Organizational Creative Capabilities: Management Factors, Processes and Devices
(Organizational Creativity Set – Volume 1)*

RUIZ Émilie

*The Creative Capabilities of Open Organizations
(Organizational Creativity Set – Volume 2)*

VERNIER Marie-France

*Eco-Design and Ecological Transition: Questioning the Economic Model
(Smart Innovation Set – Volume 42)*

2023

BOLLINGER Sophie

*Management Control and Creativity: Challenges of Managing Innovation Processes
(Smart Innovation Set – Volume 41)*

BOUVIER-PATRON Paul

*Frugal Innovation and Innovative Creation
(Smart Innovation Set – Volume 40)*

BRASSEUR Martine, BARTOLI Annie, CHABAUD Didier, GROUIEZ Pascal, ROUET Gilles

*Inclusive Territories 1: Role of Enterprises and Organizations
(Territorial Entrepreneurship and Innovation Set – Volume 1)*

CASADELLA Vanessa, UZUNIDIS Dimitri

*Agri-Innovations and Development Challenges: Engineering, Value Chains and Socio-economic Models
(Innovation in Engineering and Technology Set – Volume 8)*

DARTIGUEPEYROU Carine, SALOFF-COSTE Michel

*Futures: The Great Turn
(Innovation and Technology Set – Volume 18)*

PEYROUX Élisabeth, RAIMOND Christine, VIEL Vincent, LAVIE Émilie

Development and Territorial Restructuring in an Era of Global Change: Theories, Approaches and Future Research Perspectives

SAULAIS Pierre

Knowledge and Ideation: Inventive Knowledge Analysis for

Ideation Stimulation
(*Innovation and Technology Set – Volume 17*)

2022

AOUINAÏT Camille
Open Innovation Strategies
(*Smart Innovation Set – Volume 39*)

BOUCHÉ Geneviève
Productive Economy, Contributory Economy: Governance Tools for the Third Millennium
(*Innovation and Technology Set – Volume 15*)

BRUYÈRE Christelle
Caring Management in Health Organizations: A Lever for Crisis Management
(*Health and Innovation Set – Volume 3*)

ENJOLRAS Manon
Innovation and Export: The Joint Challenge of the Small Company
(*Smart Innovation Set – Volume 37*)

FLEURY Sylvain, RICHIR Simon
Immersive Technologies to Accelerate Innovation: How Virtual and Augmented Reality Enables the Co-Creation of Concepts
(*Smart Innovation Set – Volume 38*)

HELLER David
Valuation of the Liability Structure by Real Options
(*Modern Finance, Management Innovation and Economic Growth Set – Volume 5*)

MATHIEU Valérie
A Customer-oriented Manager for B2B Services: Principles and Implementation

MORALES Lucía, DZEVER Sam, TAYLOR Robert
Asia-Europe Industrial Connectivity in Times of Crisis
(*Innovation and Technology Set – Volume 16*)

NOËL Florent, SCHMIDT Géraldine
Employability and Industrial Mutations: Between Individual

Trajectories and Organizational Strategic Planning
(Technological Changes and Human Resources Set – Volume 4)

DE SAINT JULIEN Odile

The Innovation Ecosystem as a Source of Value Creation: A Value Creation Lever for Open Innovation
(Diverse and Global Perspectives on Value Creation Set – Volume 4)

SALOFF-COSTE Michel

Innovation Ecosystems: The Future of Civilizations and the Civilization of the Future
(Innovation and Technology Set – Volume 14)

SOULÉ Bastien, HALLÉ Julie, VIGNAL Bénédicte, BOUTROY Éric, NIER Olivier

Innovation in Sport: Innovation Trajectories and Process Optimization
(Smart Innovation Set – Volume 35)

VALLIER Estelle

Innovation in Clusters: Science–Industry Relationships in the Face of Forced Advancement
(Smart Innovation Set – Volume 36)

VAYRE Emilie

Digitalization of Work: New Spaces and New Working Times
(Technological Changes and Human Resources Set – Volume 5)

ZAFEIRIS Konstantinos N, SKIADIS Christos H, DIMOTIKALIS Yannis, KARAGRIGORIOU Alex, KARAGRIGORIOU-VONTA Christina

Data Analysis and Related Applications 1: Computational, Algorithmic and Applied Economic Data Analysis
(Big Data, Artificial Intelligence and Data Analysis Set – Volume 9)
Data Analysis and Related Applications 2: Multivariate, Health and Demographic Data Analysis
(Big Data, Artificial Intelligence and Data Analysis Set – Volume 10)

2021

ACH Yves-Alain, RMADI-SAÏD Sandra
Financial Information and Brand Value: Reflections, Challenges and Limitations

ARCADE Jacques
Strategic Engineering
(*Innovation and Technology Set – Volume 11*)

BÉRANGER Jérôme, RIZOULIÈRES Roland
The Digital Revolution in Health
(*Health and Innovation Set – Volume 2*)

BOBILLIER CHAUMON Marc-Eric
Digital Transformations in the Challenge of Activity and Work: Understanding and Supporting Technological Changes
(*Technological Changes and Human Resources Set – Volume 3*)

BUCLET Nicolas
Territorial Ecology and Socio-ecological Transition
(*Smart Innovation Set – Volume 34*)

DIMOTIKALIS Yannis, KARAGRIGORIOU Alex, PARPOULA Christina, SKIADIS Christos H
Applied Modeling Techniques and Data Analysis 1: Computational Data Analysis Methods and Tools
(*Big Data, Artificial Intelligence and Data Analysis Set - Volume 7*)
Applied Modeling Techniques and Data Analysis 2: Financial, Demographic, Stochastic and Statistical Models and Methods
(*Big Data, Artificial Intelligence and Data Analysis Set – Volume 8*)

DISPAS Christophe, KAYANAKIS Georges, SERVEL Nicolas, STRIUKOVA Ludmila
Innovation and Financial Markets
(*Innovation between Risk and Reward Set – Volume 7*)

GAUDIN Thierry, MAUREL Marie-Christine, POMEROL Jean-Charles
Chance, Calculation and Life

GASMI Nacer
Corporate Innovation Strategies: Corporate Social Responsibility and Shared Value Creation
(*Smart Innovation Set – Volume 33*)

GIORGINI Pierre

The Contributory Revolution

(Innovation and Technology Set – Volume 13)

GOGLIN Christian

Emotions and Values in Equity Crowdfunding Investment Choices

1: Transdisciplinary Theoretical Approach

Emotions and Values in Equity Crowdfunding Investment Choices

2: Modeling and Empirical Study

GRENIER Corinne, OIRY Ewan

Altering Frontiers: Organizational Innovations in Healthcare

(Health and Innovation Set – Volume 1)

GUERRIER Claudine

Security and Its Challenges in the 21st Century

(Innovation and Technology Set – Volume 12)

HELLER David

Performance of Valuation Methods in Financial Transactions

(Modern Finance, Management Innovation and Economic Growth Set – Volume 4)

LEHMANN Paul-Jacques

Liberalism and Capitalism Today

LIMA Marcos

Entrepreneurship and Innovation Education: Frameworks and Tools

(Smart Innovation Set – Volume 32)

MACHADO Carolina, DAVIM J. Paulo

Sustainable Management for Managers and Engineers

MEUNIER François-Xavier

Dual Innovation Systems: Concepts, Tools and Methods

(Smart Innovation Set – Volume 31)

UZUNIDIS Dimitri, KASMI Fedoua, ADATTO Laurent

Innovation Economics, Engineering and Management Handbook 1: Main Themes

Innovation Economics, Engineering and Management Handbook 2: Special Themes

2020

ANDREOSSO-O'CALLAGHAN Bernadette, DZEVER Sam,
JAUSSAUD Jacques, TAYLOR Robert
Sustainable Development and Energy Transition in Europe and Asia
(*Innovation and Technology Set – Volume 9*)

BEN SLIMANE Sonia, M'HENNI Hatem
Entrepreneurship and Development: Realities and Future Prospects
(*Smart Innovation Set – Volume 30*)

CHOUTEAU Marianne, FOREST Joëlle, NGUYEN Céline
Innovation for Society: The P.S.I. Approach (*Smart Innovation Set – Volume 28*)

CORON Clotilde
Quantifying Human Resources: Uses and Analysis
(*Technological Changes and Human Resources Set – Volume 2*)

CORON Clotilde, GILBERT Patrick
Technological Change (*Technological Changes and Human Resources Set – Volume 1*)

CERDIN Jean-Luc, PERETTI Jean-Marie
The Success of Apprenticeships: Views of Stakeholders on Training and Learning
(*Human Resources Management Set – Volume 3*)

DELCHET-COCHET Karen
Circular Economy: From Waste Reduction to Value Creation
(*Economic Growth Set – Volume 2*)

DIDAY Edwin, GUAN Rong, SAPORTA Gilbert, WANG Huiwen
Advances in Data Science (Big Data, Artificial Intelligence and Data Analysis Set – Volume 4)

DOS SANTOS PAULINO Victor
Innovation Trends in the Space Industry
(*Smart Innovation Set – Volume 25*)

GUILHON Bernard
Venture Capital and the Financing of Innovation
(*Innovation Between Risk and Reward Set – Volume 6*)

JURCZENKO Emmanuel

Machine Learning for Asset Management: New Developments and Financial Applications

LATOUCHE Pascal

Open Innovation: Human Set-up

(Innovation and Technology Set – Volume 10)

MAKRIDES Andreas, KARAGRIGORIOU Alex, SKIADAS Christos H.

Data Analysis and Applications 3: Computational, Classification, Financial, Statistical and Stochastic Methods

(Big Data, Artificial Intelligence and Data Analysis Set – Volume 5)

Data Analysis and Applications 4: Financial Data Analysis and Methods

(Big Data, Artificial Intelligence and Data Analysis Set – Volume 6)

MASSOTTE Pierre, CORSI Patrick

Complex Decision-Making in Economy and Finance

MICHAUD Thomas

Science Fiction and Innovation Design

(Innovation in Engineering and Technology Set – Volume 6)

MONINO Jean-Louis

Data Control: Major Challenge for the Digital Society

(Smart Innovation Set – Volume 29)

MORLAT Clément

Sustainable Productive System: Eco-development versus Sustainable Development

(Smart Innovation Set – Volume 26)

SAULAIS Pierre, ERMINE Jean-Louis

Knowledge Management in Innovative Companies 2:

Understanding and Deploying a KM Plan within a Learning Organization

(Smart Innovation Set – Volume 27)

UZUNIDIS Dimitri

Systemic Innovation: Entrepreneurial Strategies and Market Dynamics

AMENDOLA Mario, GAFFARD Jean-Luc
Disorder and Public Concern Around Globalization

BARBAROUX Pierre
Disruptive Technology and Defence Innovation Ecosystems
(*Innovation in Engineering and Technology Set – Volume 5*)

DOU Henri, JUILLET Alain, CLERC Philippe
Strategic Intelligence for the Future 1: A New Strategic and Operational Approach
Strategic Intelligence for the Future 2: A New Information Function Approach

FRIKHA Azza
Measurement in Marketing: Operationalization of Latent Constructs

FRIMOUSSE Soufyane
Innovation and Agility in the Digital Age
(*Human Resources Management Set – Volume 2*)

GAY Claudine, SZOSTAK Bérangère L.
Innovation and Creativity in SMEs: Challenges, Evolutions and Prospects
(*Smart Innovation Set – Volume 21*)

GORIA Stéphane, HUMBERT Pierre, ROUSSEL Benoît
Information, Knowledge and Agile Creativity
(*Smart Innovation Set – Volume 22*)

HELLER David
Investment Decision-making Using Optional Models
(*Economic Growth Set – Volume 2*)

HELLER David, DE CHADIRAC Sylvain, HALAOUI Lana, JOUVET Camille
The Emergence of Start-ups
(*Economic Growth Set – Volume 1*)

HÉRAUD Jean-Alain, KERR Fiona, BURGER-HELMCHEN Thierry
Creative Management of Complex Systems
(*Smart Innovation Set – Volume 19*)

LATOUCHE Pascal

Open Innovation: Corporate Incubator

(Innovation and Technology Set – Volume 7)

LEHMANN Paul-Jacques

The Future of the Euro Currency

LEIGNEL Jean-Louis, MÉNAGER Emmanuel, YABLONSKY Serge

Sustainable Enterprise Performance: A Comprehensive Evaluation Method

LIÈVRE Pascal, AUBRY Monique, GAREL Gilles

Management of Extreme Situations: From Polar Expeditions to Exploration-Oriented Organizations

MILLOT Michel

Embarrassment of Product Choices 2: Towards a Society of Well-being

N'GOALA Gilles, PEZ-PÉRARD Virginie, PRIM-ALLAZ Isabelle

Augmented Customer Strategy: CRM in the Digital Age

NIKOLOVA Blagovesta

The RRI Challenge: Responsibilization in a State of Tension with Market Regulation

(Innovation and Responsibility Set – Volume 3)

PELLEGRIN-BOUCHER Estelle, ROY Pierre

Innovation in the Cultural and Creative Industries

(Innovation and Technology Set – Volume 8)

PRIOLON Joël

Financial Markets for Commodities

QUINIOU Matthieu

Blockchain: The Advent of Disintermediation

RAVIX Joël-Thomas, DESCHAMPS Marc

Innovation and Industrial Policies

(Innovation between Risk and Reward Set – Volume 5)

ROGER Alain, VINOT Didier

Skills Management: New Applications, New Questions

(Human Resources Management Set – Volume 1)

SAULAIS Pierre, ERMINE Jean-Louis
*Knowledge Management in Innovative Companies 1:
Understanding and Deploying a KM Plan within a Learning
Organization*
(Smart Innovation Set – Volume 23)

SERVAJEAN-HILST Romaric
*Co-innovation Dynamics: The Management of Client-Supplier
Interactions for Open Innovation*
(Smart Innovation Set – Volume 20)

SKIADAS Christos H., BOZEMAN James R.
*Data Analysis and Applications 1: Clustering and Regression,
Modeling-estimating, Forecasting and Data Mining*
(Big Data, Artificial Intelligence and Data Analysis Set – Volume 2)
*Data Analysis and Applications 2: Utilization of Results in Europe
and Other Topics*
(Big Data, Artificial Intelligence and Data Analysis Set – Volume 3)

VIGEZZI Michel
*World Industrialization: Shared Inventions, Competitive
Innovations and Social Dynamics*
(Smart Innovation Set – Volume 24)

2018

BURKHARDT Kirsten
*Private Equity Firms: Their Role in the Formation of Strategic
Alliances*

CALLENS Stéphane
Creative Globalization
(Smart Innovation Set – Volume 16)

CASADELLA Vanessa
*Innovation Systems in Emerging Economies: MINT – Mexico,
Indonesia, Nigeria, Turkey*
(Smart Innovation Set – Volume 18)

CHOUTEAU Marianne, FOREST Joëlle, NGUYEN Céline
Science, Technology and Innovation Culture
(Innovation in Engineering and Technology Set – Volume 3)

CORLOSQUET-HABART Marine, JANSSEN Jacques
Big Data for Insurance Companies
(*Big Data, Artificial Intelligence and Data Analysis Set – Volume 1*)

CROS Françoise
Innovation and Society
(*Smart Innovation Set – Volume 15*)

DEBREF Romain
Environmental Innovation and Ecodesign: Certainties and Controversies
(*Smart Innovation Set – Volume 17*)

DOMINGUEZ Noémie
SME Internationalization Strategies: Innovation to Conquer New Markets

ERMINE Jean-Louis
Knowledge Management: The Creative Loop
(*Innovation and Technology Set – Volume 5*)

GAUDIN Thierry, LACROIX Dominique, MAUREL Marie-Christine, POMEROL, Jean-Charles
Life Sciences, Information Sciences

GILBERT Patrick, BOBADILLA Natalia, GASTALDI Lise, LE BOULAIRE Martine, LELEBINA Olga
Innovation, Research and Development Management

VAN HOOREBEKE Delphine
The Management of Living Beings or Emo-management

IBRAHIMI Mohammed
Mergers & Acquisitions: Theory, Strategy, Finance

LEMAÎTRE Denis
Training Engineers for Innovation

LÉVY Aldo, BEN BOUHENI Faten, AMMI Chantal
Financial Management: USGAAP and IFRS Standards
(*Innovation and Technology Set – Volume 6*)

LOISY Catherine, COULET Jean-Claude
Competence and Program-based Approach in Training: Tools for

*Developing Responsible Activities Innovation and Responsibility Set
– Volume 1)*

MILLOT Michel

Embarrassment of Product Choices 1: How to Consume Differently

PANSERA Mario, OWEN Richard

*Innovation and Development: The Politics at the Bottom of the
Pyramid*

(Innovation and Responsibility Set – Volume 2)

RICHEZ Yves

Corporate Talent Detection and Development

SACHETTI Philippe, ZUPPINGER Thibaud

New Technologies and Branding

(Innovation and Technology Set – Volume 4)

SAMIER Henri

Intuition, Creativity, Innovation

TEMPLE Ludovic, COMPAORÉ SAWADOGO Eveline M.F.W.

*Innovation Processes in Agro-Ecological Transitions in Developing
Countries*

(Innovation in Engineering and Technology Set – Volume 2)

UZUNIDIS Dimitri

Collective Innovation Processes: Principles and Practices

(Innovation in Engineering and Technology Set – Volume 4)

2017

AÏT-EL-HADJ Smaïl

The Ongoing Technological System

(Smart Innovation Set – Volume 11)

D'ANDRIA Aude, GABARRET Inés

Building 21st Century Entrepreneurship

(Innovation and Technology Set – Volume 2)

BAUDRY Marc, DUMONT Béatrice

Patents: Prompting or Restricting Innovation?

(Smart Innovation Set – Volume 12)

BÉRARD Céline, TEYSSIER Christine
Risk Management: Lever for SME Development and Stakeholder Value Creation

CHALENÇON Ludivine
Location Strategies and Value Creation of International Mergers and Acquisitions

CHAUVEL Danièle, BORZILLO Stefano
The Innovative Company: An Ill-defined Object
(Innovation between Risk and Reward Set – Volume 1)

CORSI Patrick
Going Past Limits To Growth

DAIDJ Nabyla
Cooperation, Coopetition and Innovation
(Innovation and Technology Set – Volume 3)

FERNEZ-WALCH Sandrine
The Multiple Facets of Innovation Project Management
(Innovation between Risk and Reward Set – Volume 4)

FOREST Joëlle
Creative Rationality and Innovation
(Smart Innovation Set – Volume 14)

GUILHON Bernard
Innovation and Production Ecosystems
(Innovation between Risk and Reward Set – Volume 2)

HAMMOUDI Abdelhakim, DAIDJ Nabyla
Game Theory Approach to Managerial Strategies and Value Creation
(Diverse and Global Perspectives on Value Creation Set – Volume 3)

LALLEMENT Rémi
Intellectual Property and Innovation Protection: New Practices and New Policy Issues
(Innovation between Risk and Reward Set – Volume 3)

LAPERCHE Blandine
Enterprise Knowledge Capital
(Smart Innovation Set – Volume 13)

LEBERT Didier, EL YOUNSI Hafida
International Specialization Dynamics
(*Smart Innovation Set – Volume 9*)

MAESSCHALCK Marc
Reflexive Governance for Research and Innovative Knowledge
(*Responsible Research and Innovation Set – Volume 6*)

MASSOTTE Pierre
Ethics in Social Networking and Business 1: Theory, Practice and Current Recommendations
Ethics in Social Networking and Business 2: The Future and Changing Paradigms

MASSOTTE Pierre, CORSI Patrick
Smart Decisions in Complex Systems

MEDINA Mercedes, HERRERO Mónica, URGELLÉS Alicia
Current and Emerging Issues in the Audiovisual Industry
(*Diverse and Global Perspectives on Value Creation Set – Volume 1*)

MICHAUD Thomas
Innovation, Between Science and Science Fiction
(*Smart Innovation Set – Volume 10*)

PELLÉ Sophie
Business, Innovation and Responsibility
(*Responsible Research and Innovation Set – Volume 7*)

SAVIGNAC Emmanuelle
The Gamification of Work: The Use of Games in the Workplace

SUGAHARA Satoshi, DAIDJ Nabyla, USHIO Sumitaka
Value Creation in Management Accounting and Strategic Management: An Integrated Approach
(*Diverse and Global Perspectives on Value Creation Set – Volume 2*)

UZUNIDIS Dimitri, SAULAIS Pierre
Innovation Engines: Entrepreneurs and Enterprises in a Turbulent World
(*Innovation in Engineering and Technology Set – Volume 1*)

2016

BARBAROUX Pierre, ATTOUR Amel, SCHENK Eric
Knowledge Management and Innovation
(Smart Innovation Set – Volume 6)

BEN BOUHENI Faten, AMMI Chantal, LEVY Aldo
Banking Governance, Performance And Risk-Taking: Conventional Banks Vs Islamic Banks

BOUTILLIER Sophie, CARRÉ Denis, LEVRATTO Nadine
Entrepreneurial Ecosystems
(Smart Innovation Set – Volume 2)

BOUTILLIER Sophie, UZUNIDIS Dimitri
The Entrepreneur
(Smart Innovation Set – Volume 8)

BOUVARD Patricia, SUZANNE Hervé
Collective Intelligence Development in Business

GALLAUD Delphine, LAPERCHE Blandine
Circular Economy, Industrial Ecology and Short Supply Chains
(Smart Innovation Set – Volume 4)

GUERRIER Claudine
Security and Privacy in the Digital Era
(Innovation and Technology Set – Volume 1)

MEGHOUAR Hicham
Corporate Takeover Targets

MONINO Jean-Louis, SEDKAOUI Soraya
Big Data, Open Data and Data Development (Smart Innovation Set – Volume 3)

MOREL Laure, LE ROUX Serge
Fab Labs: Innovative User
(Smart Innovation Set – Volume 5)

PICARD Fabienne, TANGUY Corinne
Innovations and Techno-ecological Transition
(Smart Innovation Set – Volume 7)

2015

CASADELLA Vanessa, LIU Zeting, DIMITRI Uzunidis
Innovation Capabilities and Economic Development in Open Economies
(*Smart Innovation Set – Volume 1*)

CORLOSQUET-HABART Marine, GEHIN William, JANSSEN Jaques, MANCA Raimondo
Asset and Liabilities Management for Banks and Insurance Companies

CORSI Patrick, MORIN Dominique
Sequencing Apple's DNA

CORSI Patrick, NEAU Erwan
Innovation Capability Maturity Model

FAIVRE-TAVIGNOT Bénédicte
Social Business and Base of the Pyramid

GODÉ Cécile
Team Coordination in Extreme Environments

MAILLARD Pierre
Competitive Quality and Innovation

MASSOTTE Pierre, CORSI Patrick
Operationalizing Sustainability

MASSOTTE Pierre, CORSI Patrick
Sustainability Calling

2014

DUBÉ Jean, LEGROS Diègo
Spatial Econometrics Using Microdata

LESCA Humbert, LESCO Nicolas
Strategic Decisions and Weak Signals

2013

HABART-CORLOSQUET Marine, JANSSEN Jacques, MANCA Raimondo

VaR Methodology for Non-Gaussian Finance

2012

POMEROL Jean-Charles
Decision-Making and Action

SZYLAR Christian
UCITS Handbook

2011

LESCA Nicolas
Environmental Scanning and Sustainable Development

LESCA Nicolas, LESCOA Humbert
Weak Signals for Strategic Intelligence: Anticipation Tool for Managers

MERCIER-LAURENT Eunika
Innovation Ecosystems

2010

SZYLAR Christian
Risk Management under UCITS III/IV

2009

COHEN Corine
Business Intelligence

ZANINETTI Jean-Marc
Sustainable Development in the USA

2008

CORSI Patrick, DULIEU Mike
The Marketing of Technology Intensive Products and Services

DZEVER Sam, JAUSSAUD Jacques, ANDREOSSO Bernadette
Evolving Corporate Structures and Cultures in Asia: Impact of Globalization

2007

AMMI Chantal
Global Consumer Behavior

2006

BOUGHZALA Imed, ERMINE Jean-Louis
Trends in Enterprise Knowledge Management

CORSI Patrick, CHRISTOFOL Hervé, RICHIR Simon, SAMIER Henri
Innovation Engineering: the power of intangible networks

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley's ebook EULA.