Volume 25, Number 2, June 2024 **ISSN 1234-7655**

STATISTICS IN TRANSITION

new series

An International Journal of the Polish Statistical Association and Statistics Poland

IN THIS ISSUE:

ISSN 1234-7655

Scientific Secretary
Marek Cierpiał-Wolan, *Statistical Office in Rzeszów, Rzeszów, Poland, e-mail: m.cierpial-wolan@stat.gov.pl*
Managing Editor
Adriana Nowakowska, *Statistics Poland, Warsaw, e-mail: a.nowakowska3@stat.gov.pl*
Secretary
Patryk Barszcz, *Statistics Poland, Warsaw, Poland, e-mail: p.barszcz@stat.gov.pl, phone number +48 22 – 608 33 66*
Technical Assistant
Rajmund Litkowiec, *Statistical Office in Rzeszów, Rzeszów, Poland, e-mail: r.litkowiec@stat.gov.pl*

Address for correspondence
Statistics Poland, al. Niepodległości 208, 00-925 Warsaw, Poland, tel./fax: +48 22 – 825 03 95

CONTENTS

STATISTICS IN TRANSITION new series, *June 2024*
Vol. 25, No. 2, pp. III

Submission information for Authors

Statistics in Transition new series (SiTns) is an international journal published jointly by the Polish Statistical Association (PTS) and Statistics Poland, on a quarterly basis (during 1993–2006 it was issued twice and since 2006 three times a year). Also, it has extended its scope of interest beyond its originally primary focus on statistical issues pertinent to transition from centrally planned to a market-oriented economy through embracing questions related to systemic transformations of and within the national statistical systems, world-wide.

The SiT*ns* seeks contributors that address the full range of problems involved in data production, data dissemination and utilization, providing international community of statisticians and users – including researchers, teachers, policy makers and the general public – with a platform for exchange of ideas and for sharing best practices in all areas of the development of statistics.

Accordingly, articles dealing with any topics of statistics and its advancement – as either a scientific domain (new research and data analysis methods) or as a domain of informational infrastructure of the economy, society and the state – are appropriate for *Statistics in Transition new series*.

Demonstration of the role played by statistical research and data in economic growth and social progress (both locally and globally), including better-informed decisions and greater participation of citizens, are of particular interest.

Each paper submitted by prospective authors are peer reviewed by internationally recognized experts, who are guided in their decisions about the publication by criteria of originality and overall quality, including its content and form, and of potential interest to readers (esp. professionals).

Manuscript should be submitted electronically to the Editor:
sit@stat.gov.pl,
GUS/Statistics Poland,
Al. Niepodległości 208, R. 296, 00-925 Warsaw, Poland

It is assumed, that the submitted manuscript has not been published previously and that it is not under review elsewhere. It should include an abstract (of not more than 1600 characters, including spaces). Inquiries concerning the submitted manuscript, its current status etc., should be directed to the Editor by email, address above, or w.okrasa@stat.gov.pl.

For other aspects of editorial policies and procedures see the SiT Guidelines on its Web site: https://sit.stat.gov.pl/ForAuthors.

STATISTICS IN TRANSITION new series, *June 2024*
Vol. 25, No. 2, pp. V–VI

Policy Statement

The broad objective of *Statistics in Transition new series* is to advance the statistical and associated methods used primarily by statistical agencies and other research institutions. To meet that objective, the journal encompasses a wide range of topics in statistical design and analysis, including survey methodology and survey sampling, census methodology, statistical uses of administrative data sources, estimation methods, economic and demographic studies, and novel methods of analysis of socio-economic and population data. With its focus on innovative methods that address practical problems, the journal favours papers that report new methods accompanied by real-life applications. Authoritative review papers on important problems faced by statisticians in agencies and academia also fall within the journal's scope.

***

Abstracting and Indexing Databases

Statistics in Transition new series is currently covered in:

BASE – Bielefeld Academic Search Engine

CEEOL – Central and Eastern European Online Library

CEJSH (The Central European Journal of Social Sciences and Humanities)

CNKI Scholar (China National Knowledge Infrastructure)

CNPIEC – cnpLINKer

CORE

Current Index to Statistics

Dimensions

DOAJ (Directory of Open Access Journals)

EconPapers

EconStore

Electronic Journals Library

Elsevier – Scopus

ERIH PLUS (European Reference Index for the Humanities and Social Sciences)

Genamics JournalSeek

Google Scholar

Index Copernicus

J-Gate

JournalGuide

JournalTOCs

Keepers Registry

MIAR

Microsoft Academic

OpenAIRE

ProQuest – Summon

Publons

QOAM (Quality Open Access Market)

ReadCube

RePec

SCImago Journal & Country Rank

TDNet

Technische Informationsbibliothek (TIB) – German National Library of Science and Technology

Ulrichsweb & Ulrich's Periodicals Directory

WanFang Data

WorldCat (OCLC)

Zenodo

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. VII–XI

From the Editor

The June issue presents a set of 12 articles. Their authors, thirty in total, come from nine countries (in the order of their appearance): Poland, Indonesia, India, Morocco, Azerbaijan, Malaysia, Nigeria, Ukraine, and Canada. Such a global scope of cooperation with our partners – authors, readers and reviewers – confirms the validity of the phrase in the title of the journal, "statistics in transition", with respect to both possible interpretations of its meaning. Highlighting the durability of the dynamic development of the discipline itself, it reflects differences in the institutional development profiles on the one hand and research interests of experts in such a broad international environment on the other. It therefore adds to the thematic richness and diversity of the entire statistical community. On behalf of everyone involved in the preparation of this volume, I wish you pleasant reading.

Research articles

In the first paper**,** ***Volatility and models based on the extreme value theory for gold returns,*** **Dominik Krężołek** and **Krzysztof Piontek** focus on the daily gold log-returns process to analyze the quality of forecasting expected shortfalls (ES) using volatility and models based on the extreme value theory (EVT). ES forecasts were calculated for conditional APARCH models formed on the entire distribution of returns, as well as for EVT models. The results of ES forecasts for each model were verified using the backtesting procedure proposed by Acerbi and Szekely. The results show that EVT models provide more accurate one-day ahead ES forecasts compared to the other models. Moreover, the asymmetric theoretical distributions for innovations of EVT models allow the improvement of the accuracy of ES forecasting. The study shows that the process of gold returns is characterized by significant, unpredictable, and heterogeneous volatility (e.g. the apparent effect of the COVID-19 pandemic). Moreover, empirical distributions of gold returns were found to be characterized by clustering of variance, leptokurtosis, asymmetry, and fat tails (compared to a normal distribution).

The article entitled ***Robust Spatial Durbin modelling on tuberculosis data using the MM-estimator method*** by **Ummul Auliyah Syam**, **Siswanto Siswanto**, and **Nurtiti Sunusi** presents the modelling of a Robust Spatial Durbin Model (RSDM) on

tuberculosis data in South Sulawesi Province. The results indicate that RSDM performs better than the Spatial Durbin Model (SDM) in explaining the number of tuberculosis cases in South Sulawesi Province in 2020, as it has a larger Adjusted R value and a smaller MSE value than SDM. Factors that significantly affect the number of tuberculosis cases in South Sulawesi Province in 2020 are population density, the percentage of households with PHBS, the percentage of BCG (Bacillus Calmette-Guérin) immunization, and the percentage of malnutrition. In terms of comparison with other spatial models for tuberculosis or infectious diseases in general, it is important to note that different spatial models may be more appropriate depending on the specific context and available data.

Neha Garg's and **Menakshi Pachori's** manuscript ***Ratio-type estimator of the population mean in stratified sampling based on the calibration approach*** presents an improved ratio-type calibrated estimator that was developed using the logarithmic mean in the calibration constraint for the stratified random sampling scheme. The proposed estimator was extended in the case of stratified double sampling and compared with the estimators given by Tracy et al. (2003) together with its ratio-type estimators, as well as Nidhi et al. (2017) and Khare et al. (2022). A simulation study was also carried out on both a real and artificial dataset in order to evaluate the performance of the proposed estimator compared to the existing estimators. The simulation study conducted on real as well as on artificial populations showed that the proposed estimators have less %RRMSE and their value decrease as the sample sizes increase, which in turn explains their performance better than the existing estimators.

In the paper ***Modeling the impact of an ISO 9001 certified quality management system on the organizational performance of Moroccan services firms*** **Ibtissam El Moury**, **Houda Kacimi**, **Sara Fennane**, and **Adil Echchelh** discuss the consequences of a noticeable absence of such a system (ISO 9001 certified quality management systems) for organizational performance of companies in Moroccan context, where the services sector contributes more than 50% to the national wealth. The authors constructed a causal model to quantify the strength of cause-and-effect relationships among the following key elements: processes within a quality management system and their impact on organizational performance, the influence of management responsibility process on all aspects of the quality management system, and the relationship between the organizational and financial performance. This model is developed using Structural Equation Modeling (SEM) and estimated through the Partial Least Squares (PLS) approach, utilizing the XL-Stat software.

Ibrahim Niftiyev's paper ***Dimensionality reduction analysis of the renewable energy sector in Azerbaijan: nonparametric analyses of large datasets*** analyzes the

renewable energy sector in Azerbaijan using Principal Component Analysis (PCA) and Multiple Correspondence Analysis (MCA). The PCA procedure yielded four distinct principle components reflecting the main macroeconomic variables, renewable energy production, industry-energy relations and natural resource revenues. Meanwhile, the MCA-based cross-country assessment of Azerbaijan's wind, solar and hydropower has struck somewhat pessimistic notes, as the country lags behind neighboring and other post-Soviet countries (e.g. Estonia, Iran, Latvia, Russia) in developing its green energy sector. These findings are of great interest to policymakers, businesses and academics who wish to gain deep insight into the Azerbaijani economy in terms of renewable energy production.

In the next article ***On the Poisson-transmuted exponential distribution and its application to frequency of claim in actuarial science,*** **S. R. M. Sabri** and **A. A. Adetunji** propose a new discrete distribution in the mixed Poisson paradigm to obtain a distribution that provides a better fit to skewed and dispersed count observation with excess zero. The cubic transmutation map is used to extend the exponential distribution, and the obtained continuous distribution is assumed for the parameter of the Poisson distribution. Various moment-based properties of the new distribution are obtained. The Nelder-Mead algorithm provides the fastest convergence iteration under the maximum likelihood estimation technique. The shapes of the proposed new discrete distribution are similar to those of the mixing distribution. Frequencies of insurance claims from different countries are used to assess the performance of the new proposition (and its zero-inflated form). Results show that the new distribution outperforms other competing ones in most cases.

Muzamil Jallal, **Aijaz Ahmed**, and **Rajnee Tripathi** in their paper ***Extended odd Frechet-exponential distribution with applications related to the environment*** attempt to expand the Frechet distribution by employing the T-X family of distributions and named the newly formulated model Extended odd Frechet-exponential distribution (EOFED). Several structural properties, reliability measurements and characteristics were estimated and discussed. The authors present graphs which depict the behavior of the probability density function, cumulative distribution function and the hazard rate function. The adaptability and flexibility of this novel distribution were achieved through the application of real-world data sets. A simulation study was also performed to evaluate and compare the output efficacy of the estimators.

In the next paper, ***Spatial and component structure analysis of the inclusive circular economy: SGICE*** **Oleksandr Osaulenko**, **Alla Shlapak**, **Iryna Zvarych**, **Oksana Brodovska**, and **Kateryna Krysovata** propose a methodology (a conceptualisation) of the global inclusive circular economy, which can be considered as a complex

multidimensional system. The main components of it are the economic, sociological, ecological and circular aspects of the country's life. To achieve this goal, the GNU regression, econometrics and time-series library was used – an applied software package for econometric modeling, a part of the GNU project. Accordingly, the authors define the global inclusive circular economy as the SGICE (Global Inclusive Circular Economy) system, characterized by the vector of functions. For the most exhaustive consideration of the entire range of opportunities of the global inclusive circular economy, the study developed and accordingly analyzed the integrated index of the development of the global inclusive circular economy (IGICE) by ecological, economic, social and circular components with isolated weakly correlated indicators.

Devendra Kumar, **Maneesh Kumar**, **Sapna Yadav**, and **Anju Goyal** present ***A new parameter estimation method for the extended power Lindley distribution based on order statistics with application***. The authors analyze inference procedures for the estimation of parameters by using order statistics and derive some new expressions for single and product moments of the order statistics from the extended power Lindley distribution. A simulation study is conducted for calculating the best linear unbiased estimates (BLUEs) for the location and scale parameters based on Type-II right-censored samples. Finally, one real data set has been used to obtain the MLEs of the model parameters, BLUEs of the parameters and the EPL distribution is also compared with some existing distributions. It was concluded that the EPL distribution provides the best fit among the compared distributions. A real data set is analyzed to illustrate the flexibility and importance of the model.

Katarzyna Kuryło's and **Łukasz Smaga's** article ***Functional repeated measures analysis of variance and its application*** is devoted to a medical problem when the same group of patients with multiple sclerosis are examined at several successive visits (doctor's appointments) and described by fractional anisotropy tract profiles, which can be represented as functions. Since the observations for each patient are dependent random processes, they follow a repeated measures design for functional data. To compare the results for different visits, functional repeated measures analysis of variance is thus considered. A pointwise test statistic is constructed by adapting the classical test statistic for one-way repeated measures analysis of variance to the functional data framework. By integrating and taking the supremum of the pointwise test statistic, two global test statistics are created. In addition to verifying the general null hypothesis of the equality of mean functions corresponding to different objects, authors propose a simple method for ex-post hoc analysis. The authors illustrate the finite sample properties of permutation and bootstrap testing procedures in an extensive simulation study, and analyze a real data example in detail [All methods are implemented in the R package rmfanova, available on CRAN.].

Other articles

XXXXI Multivariate Statistical Analysis 2023, Lodz, Poland. Conference Papers

Błażej Kochański's paper ***The shape of an ROC curve in the evaluation of credit scoring models*** focuses on the AUC, i.e. the area under the receiver operating characteristic (ROC) curve, or its scaled version, the Gini coefficient, which are the standard measures of the discriminatory power of credit scoring. Using binormal ROC curve models, it is shown how the shape of the curves affects the economic benefits of using scoring models with the same AUC. It was proposed next that the shape parameter of the fitted ROC curve be reported alongside its AUC/Gini whenever the quality of a scorecard is discussed. This approach can be useful when discussing the quality of existing or newly created scoring models. Knowing the shape of the curve allows one to assess which tasks the model is best suited for. Measuring the shape of the ROC curve can also be useful when evaluating individual component variables or component scorecards included in the combined master model.

Research Communicates and Letters

In the Research Communicates & Letters section a paper ***Monkeypox obeys the (Benford) law: a dynamic analysis of daily case counts in the United States of America*** by **Leonardo Campanelli** is presented. The article analyses the first-digit distribution of the monkeypox daily cases in the United States of America, from May 17 to September 21, 2022. The overall data follow Benford's law, a conclusion substantiated by eight different statistical tests, including the "Euclidean distance test", which has been designed to specifically check Benford's distribution in data. This result aligns with those of other infectious diseases, such as COVID 19, whose Benfordness has already been confirmed in the literature. Daily counts of monkeypox cases, and in general death and confirmed cases counts for any infectious disease, evolve in time. In order to follow the spread of monkeypox dynamically, the author analyzed the temporal deviation of monkeypox counts from Benford's law. In the case of monkeypox in the USA, no anomalies were detected, with the temporal series of daily cases conforming to Benford's distribution to a remarkably high significance level of about 99.96%. [The Statistical test used for the dynamic analysis was the Euclidean distance test.].

Włodzimierz Okrasa
Editor

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 1–22, https://doi.org/10.59170/stattrans-2024-012
Received – 21.09.2022; accepted – 24.08.2023

Volatility and models based on the extreme value theory for gold returns

Dominik Krężołek[1], **Krzysztof Piontek**[2]

Abstract

In this study, we use daily gold log-returns to analyse the quality of forecasting expected shortfalls (ES) using volatility and models based on the extreme value theory (EVT). ES forecasts were calculated for conditional APARCH models formed on the entire distribution of returns, as well as for EVT models. The results of ES forecasts for each model were verified using the backtesting procedure proposed by Acerbi and Szekely. The results show that EVT models provide more accurate one-day ahead ES forecasts compared to the other models. Moreover, the asymmetric theoretical distributions for innovations of EVT models allow the improvement of the accuracy of ES forecasting.

Key words: expected shortfall, volatility models, EVT, gold returns, backtesting.

1. Introduction

Economic processes observed in the contemporary world are very complex, varied and unpredictable. This is due to the variety of information that flows into the market. Information derives from data, which represent specific facts generated by the reality. An appropriate understanding of the nature of data is crucial in the decision-making process, for investment decisions. The last decade has also seen an increasing interest in other forms of investment of financial assets than those offered by the classical capital market. This is largely due to the uncertainty and unpredictability of the direction of the global economy. The crisis of 2008–2009 caused those investors to decide to transfer capital into other, alternative assets. One of them is gold. The main reason for seeking new markets is the aim to minimize the risk of undertaken investment activity and to hedge positions against adverse trends in the global economy.

[1] University of Economics in Katowice, Poland. E-mail: dominik.krezolek@ue.katowice.pl. ORCID: https://orcid.org/0000-0002-4333-9405.

[2] Wroclaw University of Economics and Business, Poland. E-mail: krzysztof.piontek@ue.wroc.pl. ORCID: https://orcid.org/0000-0001-9197-561X.

Risk is an integral part of any investment activity. By understanding the sources and factors generating risk, it is possible to manage it efficiently and to minimize its adverse consequences. Risk management is related to the decision-making process and the identification of tools allowing for risk reduction, as well as the construction of strategies enabling its monitoring and reporting. The measurement of risk using an appropriate measure is an important part of the risk management process. The Basel Committee on Banking Supervision based on the Basel III document (the Third Basel Accord or Basel Standards) recommends some risk measures that should be used by financial institutions and investors to ensure adequate capital reserves. These include Value-at-Risk proposed by Risk Metrics, but also Expected Shortfall, which is some extension of VaR.

Value-at-Risk, as a measure of risk, was proposed in 1994 by the US financial institution J.P. Morgan, whose analysts developed a risk management system commonly known as RiskMetrics™. VaR is defined as a statistical measure that assesses (in an unambiguous way) the amount of potential loss in the market value of a financial instrument for which the probability of reaching or exceeding it within a specified time horizon is equal to a tolerance level established by the decision maker [Dowd, 1999; Trzpiot, 2004; Doman, Doman, 2009]. The Expected Shortfall (ES), on the other hand, is a risk measure that expresses the expected value of return at a level exceeding VaR. It is commonly known in the literature as Conditional VaR (CVaR). Its advantage over the discussed VaR consists in the fact that VaR determines the minimum loss from an investment in α possible cases, and thus it unambiguously determines a certain threshold of return. In contrast, the ES focuses also on values exceeding VaR, determining the average level of losses in the conditional sense. Additionally, the ES, unlike VaR, satisfies all conditions of a coherent risk measure: the condition of monotonicity, subadditivity, positive homogeneity and translation invariance [Artzner et al. 1999], so it can be used as an evaluation for risk in the case of complex portfolio structures.

Validation of VaR and ES risk measures aims to verify that the estimated risk measure reliably assesses the actual risk under the given assumptions. The methods of such verification are referred to as backtesting methods. Backtesting is a statistical procedure in which actual profits and losses are compared with the corresponding estimates of the risk measure. The backtesting process verifies the hypothesis that the frequency of occurrence of return over a specified period is consistent with an assumed level of significance. Such tests are referred to as unconditional coverage tests. The implementation and application of these tests are generally straightforward, as they do not consider the point at which the significance level is exceeded. However, from a theoretical point of view, a correct model for risk assessment should not only provide

an estimate of the appropriate level of exceedances but also verify that they are uniformly distributed over time, i.e. independent of each other. Clustering of exceedances indicates that the model does not accurately capture market volatility and correlations. Such tests that consider the dynamic aspect of exceedances and their independence are referred to as the conditional coverage tests [Jorion, 2001].

In this article, we attempt to assess the volatility of gold returns. Modeling gold returns can provide valuable insights that help understand the volatility in the market of other financial assets. Investments in gold have many unique characteristics that make them attractive in today's complex financial world. Gold is often seen as a "safe haven" during times of economic and financial uncertainty. When other assets such as stocks or bonds experience significant declines, investors often turn to gold, which usually leads to a price increase. Gold is an asset of particular importance for hedging and diversification of investment portfolios, and therefore it is important to predict future volatility of this asset. Gold is also seen as an effective hedge against inflation. When the value of money decreases, the value of gold often rises, which can help preserve the real value of an investment. Another important fact about gold is related to investment risk. Gold exhibits low correlation with many other asset classes, meaning that its price often behaves differently than other investments. Therefore, adding gold to a portfolio can help diversify risk. Moreover, gold is one of the most universally accepted assets worldwide and can be sold almost anywhere. In addition, gold is a very liquid asset, meaning it can be easily bought and sold [Li et al., 2022].

Given the above, we compare volatility models and EVT models for forecasting extreme risk (ES) and then we apply the approach to backtesting ES proposed by Acerbi and Szekely [Acerbi et al., 2014]. The paper consists of the following sections. Section 2 reviews the literature on VaR and ES risk measures including backtesting. Section 3 describes the methodology used in the study: APARCH model of volatility and some basics of the Extreme Value Theory (EVT). Furthermore, extreme risk measures for volatility and EVT models are defined and the procedure for backtesting ES under the approach of Acerbi and Szekely is described. Section 4 presents the results of the empirical study on the example of gold returns, while in Section 5 all findings are summarized and commented.

2. Literature overview

The problem of risk is a constant object of analysis of researchers around the world. There are many studies that discuss methods of risk measurement using appropriate measures, as well as demonstrate how these methods are used in practice. Risk is related to the issue of volatility, therefore a wide range of methods is based on models describing volatility. In addition, research papers that concern analysis of volatility and

risk mainly focus on assets such as stocks, foreign exchange rates, cryptocurrencies, while the application of these methods in alternative investments, such as gold, is less popular.

Zijing and Zhang [Zijing et al., 2016] analyzed volatility in gold returns using GARCH-type models (GARCH, EGARCH, and TGARCH) with an error term described by GED distribution, while Włodarczyk [Włodarczyk, 2017] analyzed asymmetry and long memory effects on forecasting conditional volatility and risk in the gold and silver market using linear and nonlinear GARCH models. Naeem, Tiwari, Mubashra, and Shahbaz [Naeem et al., 2019] examined volatility on precious metals using Markov-Switching GARCH (MSGARCH) models. They revealed the existence of regime shifts in GARCH models and confirmed the advantages of regime-switched models over classical one-dimensional GARCH models. Mensi, Vinh Vo and Hoon Kang [Mensi et al., 2022] examined the volatility spillovers between the US stock market (S&P500) and both oil and gold before and during the global health crisis. They applied the FIAPARCH-DCC model to the 15-minute intraday data. They results showed negative (positive) conditional correlations between the S&P500 and gold (oil). Moreover, they indicated that gold offers more diversification gains than oil does during the pandemic. Vidal and Kristjanpoller [Vidal et al., 2020] investigated the forecast of gold volatility by combining two deep learning methodologies: short-term memory networks (LSTM) added to convolutional neural networks. They highlighted that these types of hybrid architectures have not been used in time series prediction. Their results showed a substantial improvement when this hybrid model was compared to the GARCH and LSTM models. Kayal and Maheswaran [Kayla et al., 2021] discovered the persistence of excess volatility in gold spot price data that engenders excessive path dependence, whereas it is not the same with silver. They used the extreme value estimator and the VRatio and observed that the strong mean-reverting characteristic in gold makes it a better investment choice than silver. Morales and Andreosso-O'Callaghan [Morales et al., 2021] showed that the daily returns of silver have a standard deviation which is more than twice that of gold. Elsayed, Gozgor and Yarovaya [Elsayed et al., 2022] examined the dynamic connectedness of return and volatility spillovers among cryptocurrency index (CRIX), gold, and uncertainty measures. Apart from traditional uncertainty measures, they also considered two novel uncertainty measures: Cryptocurrency Policy Uncertainty and Cryptocurrency Price Uncertainty indices. They observed that cryptocurrency policy uncertainty was the main transmitter of the return spillovers to other variables. In addition, gold was a net receiver of both the return and the volatility spillovers.

In the literature, there are many papers on risk measurement, however, the vast majority concern the Value-at-Risk proposed by RiskMetrics™. Daníelsson, Jorgensen, Samorodnitsky, Sarma and de Vries [Daníelsson et al., 2013] studied the properties of

VaR. They showed that the VaR measure is subadditive in the corresponding tail region of the return distribution. They also noted that estimating VaR using the historical simulation method may lead to a violation of the subadditivity assumption. They proposed to estimate the VaR risk measure using semi-parametric extreme value theory (EVT) methods. Alexander and Sarabia [Alexander et al., 2012] proposed to estimate risk associated with a value-at-risk model and adjust VaR estimates with respect to estimation and model specification errors. Huang, Huang, Chikobvu, and Chinhamu [Huang et al., 2015] predicted VaR using extreme value theory (EVT) and generalized Pareto distribution (GPD). Other researchers analyzed the forecast quality of VaR estimation using GARCH models. This approach was adopted by Walid, Shawkat and Khuong [Walid et al., 2014] by using nonlinear FIAPARCH models. Yu, Yang, Wei and Lei [Yu et al., 2018] measured VaR using GARCH-type models, extreme value theory (EVT) and copula models. The results of backtesting showed that GARCH-EVT type models and Copula models were able to improve the accuracy of VaR estimation. In contrast, Cheung and Yuen [Cheung et al., 2020] introduced an uncertainty model for the distribution of returns and investigated the impact of this volatility on VaR using a worst-case scenario approach. They showed that the choice of loss model is significant when an uncertainty model is implemented. Fiszeder, Fałdziński and Molnar [Fiszeder et al., 2019] propose some modification of multivariate DCC model to calculate forecasts of VaR. They show that regardless of whether in-sample fit, covariance forecasts or value-at-risk forecasts are considered, the model they propose outperforms not only the standard DCC model, but also an alternative range-based DCC model.

Cheng and Hung [Cheng et al., 2011] evaluated the asymmetry and kurtosis of returns distribution in the crude oil and metals market using skewed Student's *t* distribution and GARCH-type volatility models. The empirical results showed that the predictions of VaR obtained using skewed distribution were more accurate in comparison with a symmetric distribution. Eling [Eling, 2014] applied skewed Student's t distribution to risk analysis in insurance. They have shown that the skewed Student's t distribution is a notably promising distribution for modeling returns on assets such as stocks, bonds, monetary market instruments, and hedge funds. Fernandez-Perez, Frijns, Fuertes, and Miffre [Fernandez-Perez et al., 2018] analyzed the relationship between the skewness of the distribution of commodity futures and expected returns.

When it comes to backtesting risk measures, most of the research papers focus on VaR. Only few works deal with Expected Shortfall. Jalal and Rockinger [Jalal et al., 2008] use a circular block bootstrap to consider the possible dependency among exceedances. Applying the two-step procedure, they found that ES forecasts captured actual shortfalls satisfactorily. Righi and Ceretta [Righi et al., 2015] evaluate unconditional,

conditional and quantile (expectile) regression-based models for ES predictions under the ES backtest approach proposed by McNeil and Frey [McNeil et al., 2000]. Clift, Costanzino and Curran [Clift et al., 2016] apply three approaches recently proposed in the literature for backtesting ES consider a GARCH volatility specification with normal distribution for ES forecasting. Kratz, Lok and McNeil [Kratz et al. 2019] demonstrate that backtests of the forecasting models used to derive ES can be based on a multinomial test of Value-at-Risk exceptions at several levels, using heavy-tailed distributions and GARCH volatility models. Bu, Liao, Shi, and Peng [Bu et al., 2019] propose a new method to capture the dynamics of ES across time horizons using wavelet analysis. Their results confirm that the different frequency components of stock returns exhibit different persistence. Lazar and Zhang [Lazar et al., 2019] propose to measure the model risk of Expected Shortfall as the optimal correction needed to pass several ES backtests. They also investigate properties of proposed measures of model risk using GARCH models. del Brio, Mora-Valencia and Perote [del Brio et al., 2020] apply backtesting techniques for both Value-at-Risk and Expected Shortfall under parametric and semi-nonparametric approaches for modeling commodity ETFs. They recommend the application of leptokurtic distributions and semi-nonparametric techniques to mitigate regulation concerns about global financial stability of commodity business. Argyropoulos and Panopoulou [Argyropoulos et al., 2019] reviews the major VaR and ES forecast evaluation methods and evaluates their performance under a common simulation and financial application framework. They suggest that focusing on specific individual hypothesis tests provides a more reliable alternative than the corresponding conditional coverage ones. Acereda, Leon and Mora [Acereda et al., 2020] calculate ES risk measure for distributions of cryptocurrencies using various error distributions and GARCH-type models. Their results highlight the importance of estimating ES for cryptocurrencies using a generalized GARCH model and a non-normal error distribution with at least two parameters.

The method of modeling volatility using heavy-tailed distributions and Extreme Value Theory, which we present in this paper, is particularly relevant to extreme risk analysis. Many traditional methods of financial analysis assume that returns follow a normal distribution. However, in practice, returns often exhibit "heavy tails" meaning that extremely large gains or losses are more likely than a normal distribution would predict. Heavy-tailed distributions and EVT provide better modeling for these extreme events. On the other hand, analyzing heavy-tailed distributions and EVT can help investors understand the risk associated with potentially large losses. This knowledge can be useful in creating risk management strategies and diversifying portfolios. EVT is specifically designed for modeling and predicting extreme events. This can be especially useful for modeling the risk of a financial crisis or other "black swan" type events. Furthermore, using heavy-tailed distributions and EVT can increase the

credibility of financial modeling and forecasting results. Results based on these methods may be more robust to extreme events.

3. Methodology

Considering vast types of processes observed in the world around us it is possible to detect unexpectable events that generate risk at a level very far from the expected one. That type of risk is called the extreme risk and is associated with events occurring with low probability, but if they do occur, they generate extreme losses [Jajuga, 2008]. If extreme risk analysis is of interest, the theory of extreme events plays a special role. Extreme statistics are used to estimate some characteristics which help to define rare events. These statistics are e.g. quantiles of empirical distributions of examined phenomena or parameters defining periods, in which the analyzed processes take some extreme values [Gumbel, 2004]. During last century we can list many examples of events that have caused various types of drastic changes in the market, for example Black Monday (19 Oct 1987), World Trade Center (11 Sept 2001), recent financial crisis (2008–2009), crisis on crude oil market (2014) or pandemic of COVID-19 (March 2020 until present). All these events had a significant impact on the volatility of market processes and thus also on the level of risk.

3.1. Risk measures for conditional volatility models

In this paper we consider the APARCH model for volatility [Ding et al., 1993]. APARCH, which stands for Asymmetric Power ARCH, is an extension of GARCH class models that introduces certain additional features. APARCH(1,1) is less complex than APARCH(p,q) models with higher orders p and q. Less complex models are usually easier to estimate, interpret, and validate. High model complexity can lead to overfitting, which means that the model fits the training data well but performs poorly on test data or new data. The main advantages of the APARCH model over other GARCH-type models are:

- Modeling asymmetry (leverage effect): The APARCH model allows for the modeling of asymmetry, also known as the leverage effect, which is common in financial data. The leverage effect is the phenomenon where negative price changes have a greater impact on volatility than positive changes of the same magnitude. Most standard GARCH models do not account for this effect.
- Flexibility in modeling extreme events: The APARCH model allows for more flexibility in modeling extreme events (known as 'fat tails'), which are often observed in financial data.
- Powering of variance: The APARCH model allows for the powering of the variance (or conversely, the standard deviation) to any non-necessarily integer

power. This feature is useful in situations where we care about modeling the volatility directly (e.g. percentage volatility), rather than the variance.

Taking into account the characteristics above, we estimate volatility according to APARCH(1,1) model defined by the equation:

$$\sigma_t^{\delta} = \omega + \alpha_1(|\varepsilon_{t-i}| - \gamma_1\varepsilon_{t-i})^{\delta} + \beta_1\left(\sigma_{t-j}\right)^{\delta} \tag{1}$$

where $\omega, \alpha_i, \gamma_i, \beta_j, \delta$ are unknown model parameters.

These parameters play an important role in understanding the concept of the APARCH model:

- ω: This is the so-called "intercept" parameter. It influences the average level of the series' conditional variance.
- α_1: This parameter measures the impact of the squared error from the previous period on today's variance. It determines how much a large (or small) error value in the previous period increases (or decreases) the predicted variance today.
- γ_1: This parameter introduces asymmetry into the model. It determines how different the effects on variance of positive and negative errors from the previous period are.
- β_1: This parameter measures the impact of the predicted variance from the previous period on today's variance. It determines how persistent the effects of shocks on variance are.
- δ: This parameter determines the power to which the conditional standard deviation is raised. It allows for the modeling of the variance (or conversely, the standard deviation) to any non-necessarily integer power.

Moreover, the parameters of any APARCH model must satisfy certain conditions for the model to be well-defined. For instance, for the APARCH(1,1) model, the following conditions must be satisfied: $\omega > 0$, $\alpha_1 \geq 0, \beta_1 \geq$, $\alpha_1 + \beta_1 < 1$, and $\delta > 0$. For error term we consider t-Student, skewed t-Student, GED, and skewed GED distributions defined by the following probability distribution functions:

- Student's t distribution:

$$f(z|v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\Gamma\left(\frac{v}{2}\right)}\left(1 + \frac{z^2}{v}\right)^{-\left(\frac{v+1}{2}\right)} \tag{2}$$

where $\Gamma(\cdot)$ is the gamma function and v is defined as degrees of freedom ($v > 0$).

- Skewed Student's t distribution:

$$f(z|\xi, v) = \frac{2}{\xi + (\xi)^{-1}} s\{g[\xi(sz + m)|v]\mathrm{I}_{(-\infty,0)}(z + ms^{-1}) + g\left[\left(\frac{sz+m}{\xi}\right)\middle|v\right]\mathrm{I}_{(0,+\infty)}(z + ms^{-1})\} \tag{3}$$

where $g(\cdot\,|v)$ is the density of symmetric Student's t distribution, ξ is the skewness parameter defined as $\xi^2 = \frac{P(z\geq 0|\xi)}{P(z<0|\xi)}$ and v is defined as degrees of freedom ($v > 0$). Moreover, two additional parameters m (for mean) and s^2 (for variance) must be defined:

$$m = E(\varepsilon|\xi\,) = M_1(\xi - \xi^{-1}) \tag{4}$$

$$s^2 = Var(\varepsilon|\xi) = (M_2 - M_1^2)(\xi^2 + \xi^{-2}) + 2M_1^2 - M_2 \tag{5}$$

where $M_r = 2\int_0^{+\infty} s^r g(s)ds$ is the absolute moment generating function.

- GED distribution:

$$f(z|v) = \frac{v}{\left[2^{-\left(\frac{2}{v}\right)}\frac{\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}\right]^{\frac{1}{2}} 2^{\left(1+v^{-1}\right)}\Gamma\left(\frac{1}{v}\right)} \exp\left(\frac{z}{2v}\right) \tag{6}$$

where $\Gamma(\cdot)$ is the gamma function and v is defined as degrees of freedom ($v > 0$).

- Skewed GED distribution:

$$f(z|\xi, v) = \frac{2}{\xi + (\xi)^{-1}} s\{g[\xi(sz + m)|v]\mathrm{I}_{(-\infty,0)}(z + ms^{-1}) + g\left[\left(\frac{sz+m}{\xi}\right)\middle|v\right]\mathrm{I}_{(0,+\infty)}(z + ms^{-1})\} \tag{7}$$

where $g(\cdot\,|v)$ is the density of symmetric GED distribution, ξ is the skewness parameter defined as $\xi^2 = \frac{P(z\geq 0|\xi)}{P(z<0|\xi)}$ and v is defined as degrees of freedom ($v > 0$).

All parameters of APARCH(1,1) model have been estimated using Maximum Likelihood Estimation (MLE).

For conditional volatility model we compute VaR and ES risk measures using formulas:

$$VaR_t^{\alpha} = \mu_t + \sigma_t F^{-1}(\alpha) \tag{8}$$

$$ES_t^{\alpha} = \mu_t + \sigma_t\left(\frac{1}{\alpha}\int_0^{\alpha} F^{-1}(s)ds\right) \tag{9}$$

3.2. Risk measures for Extreme Value Theory models (EVT)

Another way to calculate measures of extreme risk is to estimate the conditional quantile using Extreme Value Theory (EVT). Extreme Value Theory is a branch of statistics dealing with statistical methods and the probabilistic and statistical theory related to extreme events, which are often significant in areas such as meteorology,

hydrology, finance, insurance, or even structural engineering. One of the key concepts in EVT is the Generalized Extreme Value (GEV) distribution. It combines three types of extreme value distributions: the Gumbel, Frechet, and Weibull families of distributions. In ETV we deal with some extremes. There are two main types of extremes [Gumbel, 2004]:

- Block maxima: This is the maximum value of a block of data. It is like considering the maximum rainfall recorded every month, for instance.
- Peak over threshold: This considers all values over a certain high threshold, not just the maxima.

In Extreme Value Theory certain theorems play a special role. The first one – the Fisher-Tippett-Gnedenko theorem – is a critical theoretical foundation of EVT, which states that with proper normalization, the maxima of a sequence of random variables converge in distribution to one of the three types of extreme value distributions mentioned above. The Pickands–Balkema–de Haan theorem is another important theoretical foundation of EVT for the peaks-over-threshold approach. This theorem states that above a sufficiently high threshold, the excess distribution over that threshold can be approximated by a Generalized Pareto distribution [Fałdziński, 2014].

EVT provides a rigorous way to make statistical inferences about rare events (extreme events). Because of its ability to predict such events, EVT is increasingly being applied in various fields such as finance, insurance, and environmental science.

The extreme value distribution (EVD) can be described using the following density function [Gumbel, 2004]:

$$EVD_{\gamma}(x) = \begin{cases} exp\left[-(1+\gamma x)^{-1/\gamma}\right],\ 1+\gamma x \geq 0 & for\ \ \gamma \neq 0 \\ exp[-exp(-x)],\ x \in \mathbb{R} & for\ \ \gamma = 0 \end{cases} \tag{10}$$

where γ defines the extreme value index (EVI). It is the most important parameter in this distribution, which measures the thickness of its tail (and thus the probability of an extreme event occurring). The heavier the tail, the higher the EVI value. In the case of a generalized EVD, the tail index is the shape parameter, which is invariant of standardizing the distribution [Németh et al., 2018]. The EVI can be estimated using Hill [Hill, 1975] or Picands estimator [Picands, 1975]. In this paper we use the Peaks-Over-Threshold (POT) approach based on EVT and on the generalized Pareto distribution (GPD), which is the limiting tail distribution for a wide variety of continuous probability distributions.

The POT method is a classical approach used in EVT estimation. It consists of fitting the GPD distribution to the innovations obtained from filtering returns using the model of conditional volatility. For i.i.d. random variable, consider the distribution

function of excesses $Y = u - Z$ for a given threshold u, $F_u(y) = P(Y = u - Z \leq y | Z < u) = \frac{[F(u) - F(u-y)]}{[F(u)]}, y \geq 0$. The excesses over threshold u follow GPD distribution, $Y = u - Z \sim GPD(\xi, \beta)$:

$$F_u(y) \approx GPD_{\xi,\beta}(y) = \begin{cases} 1 - \left(1 + \frac{\xi y}{\beta}\right)^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - \exp\left(-\frac{y}{\beta}\right), & \xi = 0 \end{cases} \tag{11}$$

$GPD_{\xi,\beta}(y)$ provides $y \geq 0$ if $\xi \geq 0$ and $0 \leq y \leq -\frac{\beta}{\xi}$ if $\xi < 0$, where $\beta > 0$ is the scale parameter and ξ is the shape parameter of the tail of the distribution. Consider the following equation for points $z < u$ in the left tail of F as:

$$F(z) = F(u) - F_u(u - z)F(u) = F(u)(1 - F_u(u - z)) \tag{12}$$

Using the proportion of a tailed data $\frac{T_u}{T}$, the tail estimator of GPD is of the form:

$$\hat{F}(z) = \frac{T_u}{T}\left(1 + \hat{\xi}\frac{u-z}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}} \tag{13}$$

All parameters of EVT model have been estimated using Maximum Likelihood Estimation (MLE).

According to the EVT approach, formulas for calculating VaR and ES are presented by:

$$VaR_t^{\alpha} = \mu_t + \sigma_t F_z^{-1}(\alpha) = \mu_t + \sigma_t\left(u + \frac{\beta}{\xi}\left[1 - \left(\frac{\alpha}{\frac{T_u}{T}}\right)^{-\xi}\right]\right) \tag{14}$$

$$ES_t^{\alpha} = \mu_t + \sigma_t\left(\frac{1}{\alpha}\int_0^{\alpha} F_z^{-1}(s)ds\right) = \mu_t + \sigma_t\left[\frac{VaR_t^{\alpha}}{1-\xi} - \left(\frac{\beta + \xi u}{1-\xi}\right)\right] \tag{15}$$

3.3. Backtest for Expected Shortfall

In this paper we focus on Expected Shortfall backtesting procedure proposed by Acerbi and Szekely [Acerbi et al., 2014]. This is one of many possible approaches to testing this risk measure proposed in the literature. They present two non-parametric tests, both free from assumptions about the probability distribution of the returns. The advantage of this approach is that it does not require any particular form of theoretical distribution, only the continuity of the distribution function along with independence of observation in the sample. To estimate p-values an algorithm based on Monte Carlo simulations is used. The first test statistics, for testing ES after VaR, is of the form:

$$Z_1 = \frac{1}{N_T}\sum_{t=1}^{T}\frac{I_t r_t}{ES_t^{\alpha}} - 1 \tag{16}$$

where $N_T = \sum_{t=1}^{T} I_t > 0$ with $I_t = \mathbb{I}_{\{r_t < VaR_t^{\alpha}\}}$ is the indicator function of VaR violations and T is the length of the out-of-sample period.

The null hypothesis says that $P_t^{\alpha} = F_t^{\alpha}$ for all t, where F_t^{α} is the tail of cumulative distribution of forecasts at time t when $r_t < VaR_t^{\alpha,F}$ and P_t^{α} is the tail of the unknown distribution from which the realized returns r_t are drawn. The risk measures VaR and ES under the theoretical and empirical distributions are denoted by $VaR_t^{\alpha,P}$, $ES_t^{\alpha,P}$, $VaR_t^{\alpha,F}$, $ES_t^{\alpha,F}$. The alternative hypothesis says that $ES_t^{\alpha,P} \leq ES_t^{\alpha,F}$ for all t and $ES_t^{\alpha,P} < ES_t^{\alpha,F}$ for some t, together with $VaR_t^{\alpha,P} = VaR_t^{\alpha,F}$ for all t. We can find that the predicted value of VaR^{α} is still correct for alternative hypothesis, according to the idea that this test is subordinate to the initial test of VaR. This test is in fact completely insensitive to an excessive number of exceptions, since it is the average of the exceptions taken over themselves. Assuming these conditions $E_{H_0}[Z_1|N_T > 0] = 0$ and $E_{H_1}[Z_1|N_T > 0] > 0$.

The second test statistics, for testing ES directly, is of the form:

$$Z_2 = \frac{1}{T\alpha}\sum_{t=1}^{T}\frac{I_t r_t}{ES_t^{\alpha}} - 1 \tag{17}$$

provided that $N_T = \sum_{t=1}^{T} I_t > 0$ with $I_t = \mathbb{I}_{\{r_t < VaR_t^{\alpha}\}}$ is the indicator function of VaR violations and T is the length of the out-of-sample period.

The null hypothesis says that $P_t^{\alpha} = F_t^{\alpha}$ for all t, where F_t^{α} is the tail of cumulative distribution of forecasts at time t when $r_t < VaR_t^{\alpha,F}$ and P_t^{α} is the tail of the unknown distribution from which the realized returns r_t are drawn. The alternative hypothesis says that $ES_t^{\alpha,P} \leq ES_t^{\alpha,F}$ for all t and $ES_t^{\alpha,P} < ES_t^{\alpha,F}$ for some t, together with $VaR_t^{\alpha,P} \leq VaR_t^{\alpha,F}$ for all t. As $E_{H_0}[N_T] = T\alpha$ we have $E_{H_0}[Z_2] = 0$ and $E_{H_1}[Z_2] = 0$.

Unlike the Z_1 statistics, the sum of VaR violation event returns is now divided by the expected value. Statistics Z_2 will tend to reject the large number of small VaR violation events. This leads to a difference in alternative hypothesis between the two statistics. Rejecting null hypothesis in Z_2 means rejecting VaR as being correctly defined. The advantage of the proposed approach is its relative computational simplicity. The disadvantage is the requirement of using Monte Carlo simulations to obtain critical values and p-values for the test statistic.

4. Empirical study

The empirical study is based on daily log-returns of gold quoted from January 2015 to December 2021 on the London Metal Exchange. The research period was divided into two sub-periods:

- sub-period of the models' parameter estimation: 2015–2017,
- sub-period of the ES forecast: 2018–2021.

In this research we investigate risk using Expected Shortfall (ES). Expected Shortfall (ES) and Value at Risk (VaR) are both measures used in financial risk management to quantify the level of financial risk within a firm or investment portfolio over a specific time frame. However, they differ in how they approach this risk assessment. VaR measures the maximum loss that will not be exceeded with a certain confidence level. In contrast, ES, also known as Conditional VaR (CVaR), estimates the expected loss given that a loss is greater than the VaR. It provides a more comprehensive view of risk by not only considering the worst-case scenarios but also their potential severity. VaR has been criticized for not adequately capturing tail risk, which refers to extreme events that have low probability but high impact. ES is designed to overcome this shortfall by focusing on the tail of the loss distribution, providing a more accurate measure of potential losses in extreme events. ES is a coherent risk measure, meaning it satisfies properties such as subadditivity. Subadditivity implies that diversifying a portfolio reduces risk, a property that VaR does not have. This can be particularly important for risk management, as it encourages appropriate risk diversification. Following the 2008 financial crisis, many financial regulators have favored ES over VaR as a risk measure. For example, the Basel Committee on Banking Supervision recommended the use of ES for determining regulatory capital requirements due to its ability to better capture tail risk.

The methodology described in Section 3 was used to construct the models. All risk measures were estimated for quantiles of 0.01 and 0.05. Table 1 presents the descriptive statistics of the gold return.

Table 1. Descriptive statistics for gold returns

Statistics	Sub-period 2015–2017	Sub-period 2018–2021
Mean	0.00012	0.00033
Standard error	0.00031	0.00027
Median	0.00007	0.00072
Standard deviation	0.00852	0.00871
Coefficient of variation [%]	7019.87	2647.64
Kurtosis	2.53925	4.80494
Skewness	0.19121	-0.60820
Range	0.08087	0.10212
Minimum	-0.03409	-0.05849
Maximum	0.04678	0.04363
N	755	1032
Kolmogorov-Smirnov test for normality	0.07214	0.06243
p-value	<0.001*	<0.001*

* statistical significance at 0.05.

The average values of the gold return in both sub-periods are similar, but the distributions differ. In the first sub-period (2015–2017), the distribution of gold return is skewed to the right, while in the second sub-period it is skewed to the left. The distributions are characterized by a high level of kurtosis. Moreover, the hypothesis that empirical distributions follow a normal distribution was rejected. Time series and empirical distributions of gold return within both sub-periods and along with the fitted normal distribution are shown in Figure 1–2.

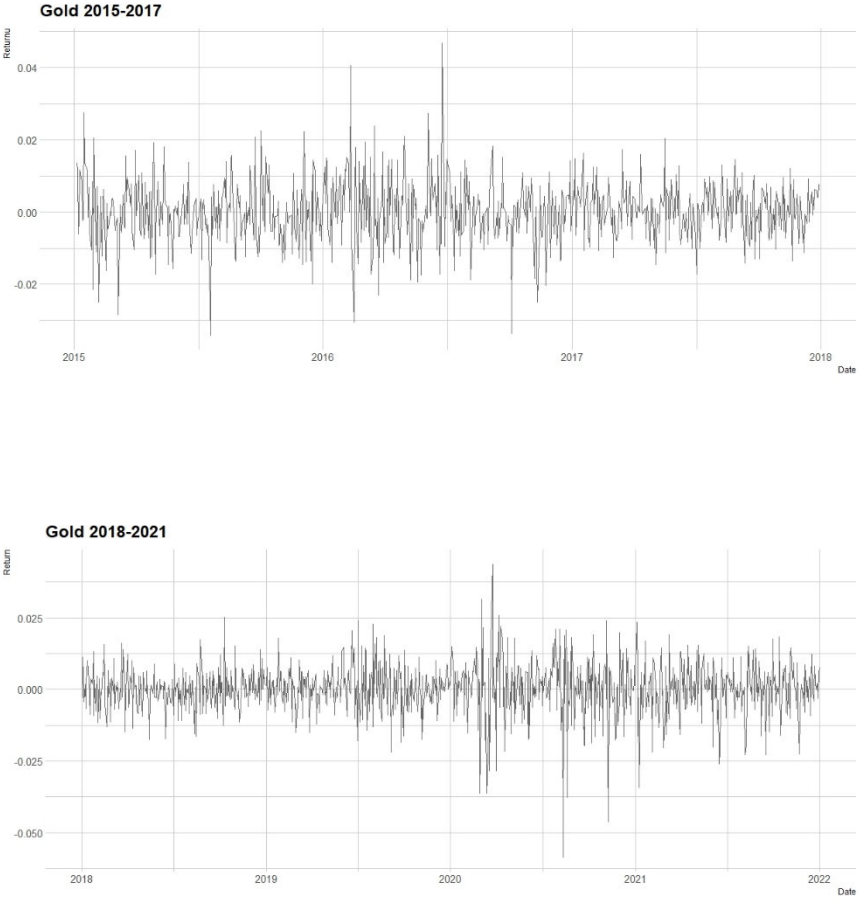


Figure 1: Time series for gold returns (subperiod 2015–2017 — top, subperiod 2018–2021 — bottom).

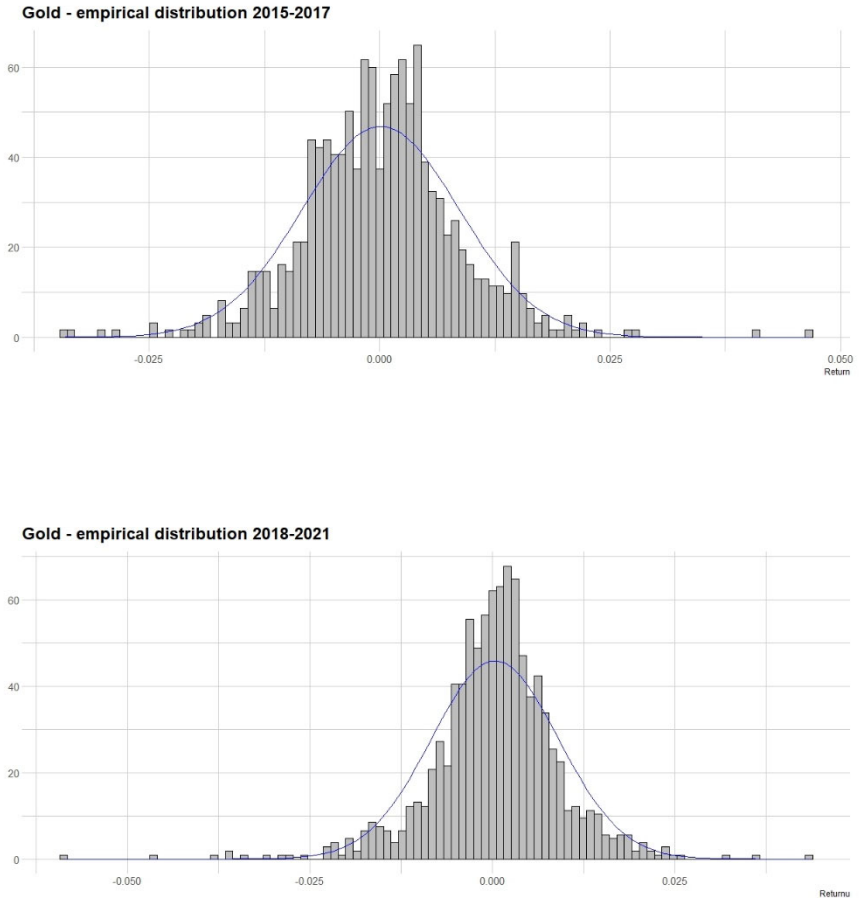


Figure 2: Empirical distributions with fitted normal densities for gold returns (subperiod 2015–2017 — top, subperiod 2018–2021 — bottom).

In the next step of the analysis, the parameters of the conditional volatility and EVT models were estimated. The original study used a variety of volatility models such as GARCH, APARCH, EGARCH, TGARCH, FIGARCH, GARCH-GJR for errors term described by normal, t-Student, skewed t-Student, GED and skewed GED distributions. However, only for APARCH models the largest number of statistically significant parameters and the smallest values of information criterion AIC were obtained. Therefore, only the results for APARCH models are presented in this paper. The results are presented in Table 2.

Table 2: Estimates of model parameters

Model	Parameter	Error~ST		
		Estimates	Standard Error	p-value
APARCH(1,1)	ω	0.0000	0.0000	0.499
	α_1	0.0261	0.0092	0.004*
	γ_1	-0.7735	1.1915	0.516
	β_1	0.9495	0.0445	<0.001*
	δ	1.3537	0.4944	0.006*
EVT	ν	5.6466	1.2022	<0.001*
	ξ	-	-	-
	ξ_{GPD}	0.3431	0.1217	0.001*
	β_{GPD}	0.5543	0.0862	<0.001*
Model	Parameter	Error~ST_{skewed}		
		Estimates	Standard Error	p-value
APARCH(1,1)	ω	0.0000	0.0000	0.495
	α_1	0.0260	0.0089	0.001*
	γ_1	-0.7716	1.1706	0.510
	β_1	0.9498	0.0439	<0.001*
	δ	1.3538	0.4898	0.006*
EVT	ν	5.6524	1.2025	<0.001*
	ξ	0.0095	0.0527	0.856
	ξ_{GPD}	0.3592	0.0621	<0.001*
	β_{GPD}	0.5733	0.0723	<0.001*
Model	Parameter	Error~GED		
		Estimates	Standard Error	p-value
APARCH(1,1)	ω	0.0000	0.0000	0.535
	α_1	0.0185	0.0100	0.006*
	γ_1	-0.4033	0.4030	0.3169
	β_1	0.9776	0.0149	<0.001*
	δ	1.3598	0.3554	<0.001*
EVT	ν	1.3511	0.1000	<0.001*
	ξ	-	-	-
	ξ_{GPD}	0.3318	0.0045	0.001*
	β_{GPD}	0.5982	0.0788	<0.001*
Model	Parameter	Error~GED_{skewed}		
		Estimates	Standard Error	p-value
APARCH(1,1)	ω	0.0000	0.0000	0.538
	α_1	0.0181	0.0098	0.007*
	γ_1	-0.4070	0.4012	0.310
	β_1	0.9781	0.0148	<0.001*
	δ	1.3705	0.3489	<0.001*
EVT	ν	1.3512	0.0995	<0.001*
	ξ	0.0121	0.0527	0.818
	ξ_{GPD}	0.3644	0.1240	<0.001*
	β_{GPD}	0.5417	0.0059	<0.001*

* Statistical significance at 0.05.

Most parameters for the estimated conditional volatility models and all parameters in the EVT models turned out to be statistically significant. A strong long memory effect was observed in the APARCH models (statistically significant, positive values of β_1). Considering the Akaike Information Criterion, the best model is the one in which innovations are described by a t-Student distribution.

Then, the estimated models were used to determine one-day ES forecasts within the period of 2018–2021. The average ES value was computed, and the rate of its violations was calculated for the fixed quantile level. The convergence of the theoretical estimates with the real averaged ES value in the 2018–2021 sub-period was assessed using the RMSE. The p-values for both Z_1 and Z_2 statistics in Acerbi and Szekely approach were also estimated. The results are presented in Table 3.

Table 1: 1-day Ahead average forecasts of ES and p-values for Z_1 and Z_1 statistics

Quantile	Model	$\overline{ES}$	% of violations	p-value Z_1	p-value Z_2	RMSE
0.01	Empirical	-0.0438	0.0100	-	-	-
	APARCH-St	-0.0310	0.0140	0.001*	0.000*	0.0128
	APARCH-St$_{skewed}$	-0.0316	0.0140	0.003*	0.002*	0.0122
	APARCH-GED	-0.0294	0.0160	0.004*	0.000*	0.0144
	APARCH-GED$_{skewed}$	-0.0300	0.0160	0.002*	0.001*	0.0138
	EVT-St	-0.0465	0.0090	**0.895**	**0.970**	0.0027
	EVT-St$_{skewed}$	-0.0474	0.0100	**0.910**	**0.922**	0.0036
	EVT-GED	-0.0441	0.0090	**0.956**	**0.946**	0.0004
	EVT-GED$_{skewed}$	-0.0450	0.0100	**0.944**	**0.970**	0.0012
0.05	Empirical	-0.0275	0.0500	-	-	-
	APARCH-St	-0.0188	0.0410	0.116	0.417	0.0087
	APARCH-St$_{skewed}$	-0.0192	0.0430	0.110	0.398	0.0083
	APARCH-GED	-0.0179	0.0440	0.084	0.786	0.0096
	APARCH-GED$_{skewed}$	-0.0182	0.0440	0.104	0.747	0.0093
	EVT-St	-0.0282	0.0490	**0.851**	**0.922**	0.0007
	EVT-St$_{skewed}$	-0.0288	0.0480	**0.864**	**0.876**	0.0013
	EVT-GED	-0.0268	0.0500	**0.909**	**0.898**	0.0007
	EVT-GED$_{skewed}$	-0.0273	0.0510	**0.896**	**0.922**	0.0001

* Statistical significance at 0.05.

The results show that regardless of the quantile level there is no reason to reject the null hypothesis in Z_1 and Z_2 tests for models based on EVT. The conclusion seems to be logical, because EVT models focus only on values in the tail of the distribution, while conditional volatility models are built using all values of the return. Thereby, the high quality of ES forecasts for EVT models may be confirmed. The averaged ES values for each model calculated in the second sub-period are presented in Figures 3–4.

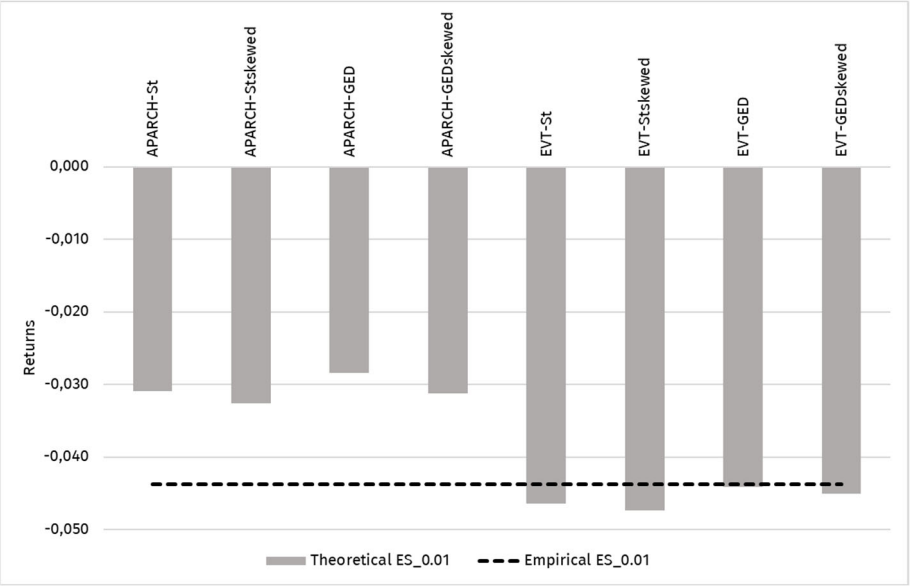


Figure 3: Empirical vs. theoretical averaged ES forecasts for quantile 0.01.

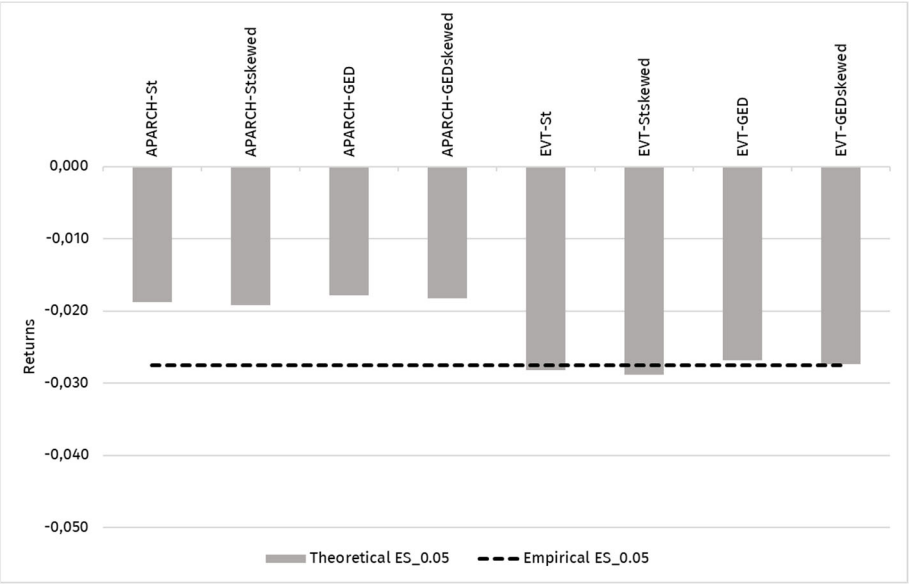


Figure 4: Empirical vs. theoretical averaged ES forecasts for quantile 0.05.

Both figures show that conditional volatility models underestimate ES values, regardless of the quantile level. It is observed that the RMSE values for the EVT models are quite low.

3. Conclusions

In this paper we examined risk using volatility and Extreme Value Theory models. We focus on estimating extreme changes in gold returns for which risk is estimated using ES. There are several reasons why investing in gold is important. First, gold is often seen as a "safe haven" during times of economic and financial uncertainty. When other assets such as stocks or bonds experience significant declines, investors often turn to gold, which usually leads to a price increase. Second, gold is also seen as an effective hedge against inflation. In terms of risk diversification, gold has a low correlation with many other asset classes, which means that its price often behaves differently than other investments. Therefore, adding gold to a portfolio can help diversify risk. The analysis of gold prices is important looking at the global demand. Gold is a unique asset that has both investment and consumption value (e.g. in jewelry). The rise of the middle class in developing countries like China and India may lead to increased demand for gold. The amount of gold in the world is limited, and the extraction process is difficult and costly. This limited supply helps maintain gold's value. Moreover, gold is one of the most universally accepted assets worldwide and can be sold almost anywhere and can be easily bought and sold.

ES is one of the most popular risk measures. As mentioned in Section 4, there are several factors in favour of using this measure instead of VaR. ES is a coherent risk measure and provides a more comprehensive view of risk by not only considering the worst-case scenarios but also their potential severity. Moreover, ES focuses on the tail of the loss distribution, providing a more accurate measure of potential losses in extreme events.

Backtesting provides the means of determining the accuracy of risk forecasts and the corresponding risk model. In this paper we used APARCH conditional volatility models and Extreme Value Theory models to evaluate ES forecasts of gold returns. In addition, we assume different heavy-tailed error distributions. Our study shows that the process of gold returns is characterised by significant, unpredictable, and heterogeneous volatility (e.g. the apparent effect of the COVID-19 pandemic). Moreover, empirical distributions of gold returns were found to be characterized by clustering of variance, leptokurtosis, asymmetry, and fat tails (compared to a normal distribution). A strong long memory effect (statistically significant positive parameter β_1) was observed in the conditional volatility models. We found out that risk estimates based on EVT are closer to real values of ES for EVT models than for APARCH models. The best results were obtained for EVT-GED and EVT-SkewedGED models. ES forecasts obtained for conditional models have p-values close to zero - the null hypothesis is rejected due underestimation of risk. Significant differences in p-value levels were observed between APARCH conditional models and EVT conditional

models. The p-value for the Z_1 statistics in the EVT model means that the average of realised ES exceedances is lower than predicted, while for Z_2 additionally that also the percentage of exceedances is lower. Thus, the validity of conditional EVT models was demonstrated. Nevertheless, these are general conclusions, requiring further in-depth research.

References

Acereda, B., Leon, A., Mora, J., (2020). Estimating the expected shortfall of cryptocurrencies: An evaluation based on backtesting. *Finance Research Letters*, 33, 101181.

Alexander, C., Sarabia, J. M., (2012). Quantile uncertainty and Value-at-Risk model risk. *Risk Anal. Int. J.*, 32 (8), pp. 1293–1308.

Argyropoulos, Ch., Panopoulou, E., (2019). Backtesting VaR and ES under the magnifying glass. *International Review of Financial Analysis*, 64, pp. 22–37.

Artzner, P., Delbaen, F., Eber, J. M., Heath, D., (1999). Coherent Measures of Risk, *Mathematical Finance*, Vol. 9(3), pp. 203–228.

Bu, D., Liao, Y., Shi, J., Peng, H., (2019). Dynamic expected shortfall: A spectral decomposition of tail risk across time horizons. *Journal of Economic Dynamics & Control*, 108, 103753.

Cheng, W-H., Hung, J-C., (2011). Skewness and leptokurtosis in GARCH-typed VaR estimation of petroleum and metal asset returns. *Journal of Empirical Finance*, 18, pp. 160–173.

Cheung, K. C., Yuen F. L., (2020). On the uncertainty of VaR of individual risk. *Journal of Computational and Applied Mathematics*, 367, 112468.

Clift, S. S., Costanzino, N., Curran, M., (2016). Empirical Performance of Backtesting Methods for Expected Shortfall, http://dx.doi.org/10.2139/ssrn.2618345.

Daníelsson, J., Jorgensen, B. N., Samorodnitsky, G., Sarma, M., de Vries, C. G., (2013). Fat tails, VaR and subadditivity. *Journal of Econometrics*, 172, pp. 283–291.

Del Brio, E. B., Mora-Valencia, A., Perote, J., (2020). Risk quantification for commodity ETFs: Backtesting Value-at-Risk and Expected Shortfall. *International Review of Financial Analysis*, 70, 101163.

Ding, Z., Granger, C. W. J., Engle, R. F., (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1, pp. 83–106.

Doman, M., Doman, R., (2009). Modelowanie zmienności i ryzyka. Metody ekonometrii finansowej, Wolters Kluwer Polska, Kraków.

Dowd, K., (1999). Beyond Value at Risk: The New Science of Risk Management, John Wiley & Sons, Chichester.

Eling, M., (2014). Fitting asset returns to skewed distributions: Are the skew-normal and skew-student good models?. *Insurance: Mathematics and Economics*, 59, pp. 45–56.

Elsayed, A. H., Gozgor, G., Yarovaya, L., (2022). Volatility and return connectedness of cryptocurrency, gold, and uncertainty: Evidence from the cryptocurrency uncertainty indices. *Finance Research Letters*, 47, 102732.

Fałdziński, M., (2014). Teoría wartości ekstremalnych w ekonometrii finansowej. *Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika w Toruniu*, Toruń.

Fernandez-Perez, A., Frijns, B., Fuertes, A-M., Miffre J., (2018). The skewness of commodity futures returns. *Journal of Banking and Finance*, 86, pp. 143–158.

Fiszeder, P., Fałdziński, M., Molnár, P., (2019). Range-Based DCC Models for Covariance and Value-at-Risk Forecasting. *Journal of Empirical Finance*, 54, pp. 58–76

Gumbel, E. J., (2004). Statistics of Extremes. *Dover Publications*, Inc, Mineola, New York.

Hill, B. M., (1975). A simple general approach to inference about the tail of the distribution. *The Annals of Statistics*, 3(5), pp. 1163–1174.

Huang, C. K., Huang, C. S., Chikobvu, D., Chinhamu, K., (2015). Extreme risk, Value-at-Risk and Expected Shortfall in the gold market. *Int. Bus. Econ. Res. J.*, 14(1), pp. 91–107.

Jajuga, K., (2008). Zarządzanie ryzykiem. *Polskie Wydawnictwo Naukowe PWN*, Warszawa.

Jalal, A., Rockinger, M., (2018). Predicting tail-related risk measures: The consequences of using GARCH filters for non-GARCH data. *Journal of Empirical Finance*, 15, pp. 868–877.

Jorion, P., (2001). Value-at-Risk, The New Benchmark for Managing Financial Risk, 2nd Edition, McGraw-Hill, New York.

Li, X., Guo Q., Liang Ch., Umar M., (2022). Forecasting gold volatility with geopolitical risk indices. *Research in International Business and Finance*, 64, 101857.

Kayla, P., Maheswaran, S., (2021). A study of excess volatility of gold and silver. *IIMB Management Review*, 33, pp. 133–145.

Kratz, M., Lok, Y. H., McNeil, A. J., (2018). Multinomial VaR backtests: A simple implicit approach to backtesting expected shortfall. *Journal of Banking and Finance*, 88, pp. 393–407.

Lazar, E., Zhang, N., (2019). Model risk of expected shortfall. *Journal of Banking and Finance*, 105, pp. 74–93.

McNeil, A., Frey R., (2000): Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7, pp. 271–300.

Mensi, V., Vinh, Vo X., Hoon Kang, S., (2022). COVID-19 pandemic's impact on intraday volatility spillover between oil, gold, and stock markets. *Economic Analysis and Policy*, 74, pp. 702–715.

Morales, L., Andreosso-O'Callaghan, B., (2011). Comparative analysis on the effects of the Asian and global financial crises on precious metal markets. *Research in International Business and Finance*, 25(2), pp. 203–227.

Naeem, M., Tiwari, A. K., Mubashra, S., Shahbaz, M., (2019). Modeling volatility of precious metals markets by using regime-switching GARCH models. *Resources Policy*, 64, 101497.

Righi M. B., Ceretta P. S., (2015). A comparison of Expected Shortfall estimation models. *Journal of Economics and Business*, 78, pp. 14–47.

Trzpiot, G., (2004). O wybranych własnościach miar ryzyka. *Badania Operacyjne i Decyzje*, no. 3-4, pp. 91–98.

Walid, C., Shawkat, H., Khuong, N. D., (2014). Volatility forecasting and risk management for commodity markets in the presence of asymmetry and long memory. *Energy Econ.*, 41, pp. 1–18.

Włodarczyk, B., (2017). Prognozowanie zmienności stóp zwrotu na rynkach złota i srebra z uwzględnieniem efektu asymetrii i długiej pamięci. *Studia i Prace WNEiZ US*, no. 50, T. 1, pp. 231–247.

Vidal, A., Kristjanpoller, W., (2020). Gold volatility prediction using a CNN-LSTM approach. *Expert Systems with Applications*, 157, 113481.

Yu, W., Yang, K., Wei, Y., Lei, L., (2018). Measuring Value-at-Risk and Expected Shortfall of crude oil portfolio using extreme value theory and vine copula. *Physica A*, 490, pp. 1423–1433.

Zijing, Z., Zhang, H. K., (2016). The dynamics of precious metal markets VaR: A GARCH-EVT approach. *Journal of Commodity Markets*, 4, pp. 14–27.

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 23–38, https://doi.org/10.59170/stattrans-2024-013
Received – 30.03.2023; accepted – 06.11.2023

Robust spatial Durbin modelling on tuberculosis data using the MM-estimator method

Ummul Auliyah Syam[1], **Siswanto Siswanto**[2], **Nurtiti Sunusi**[3]

Abstract

The spatial Durbin model (SDM) is a spatial regression model which shows the existence of spatial dependency on the response variable and predictor variables. However, SDM modelling may sometimes involve problems associated with e.g. the existence of spatial outliers. One way to overcome outliers in the SDM model is to use robust regression in the form of the robust spatial Durbin model (RSDM). This study aims to estimate the parameters of RSDM based on data on tuberculosis (TB) cases recorded in 2020 in the South Sulawesi Province in Indonesia and to identify the factors that affect the number of TB cases in the region. The MM-Estimator robust regression estimation method was used. It is a combination of a method involving a high breakdown value for the S-estimator and a high efficiency of the M-estimator. The results of the analysis show that RSDM can overcome outliers in spatial regression models. This is reflected in the value of the mean square error (MSE) of the RSDM, which is 6,461.734, i.e. smaller than the value of the SDM model, and the adjusted R^2 value of 99.52%, which is greater than that of the SDM model. The factors that influence the number of TB cases in the South Sulawesi Province are population density, the percentage of households leading a healthy lifestyle, the percentage of residents with Bacillus Calmette-Guérin (BCG) immunisation, and the percentage of those suffering from malnutrition.

Key words: spatial regression, RSDM, MM-Estimator, tuberculosis.

1. Introduction

Spatial regression is a statistical analysis that measures the relationship between predictor variables and response variables in regression models by considering spatial effects (Ramadhani et al., 2020). One of the models using spatial regression is the Spatial Durbin Model (SDM). SDM is a spatial regression model that shows the spatial

[1] Department of Statistics, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Indonesia. E-mail: ummulsyam08@gmail.com. ORCID: https://orcid.org/0009-0006-5254-3527.

[2] Corresponding author, Department of Statistics, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Indonesia. E-mail: Siswanto@unhas.ac.id. ORCID: https://orcid.org/0000-0003-1934-5343.

[3] Department of Statistics, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Indonesia. E-mail: nurtitisunusi@unhas.ac.id. ORCID: https://orcid.org/0000-0002-6436-831X.

relationship between response and predictor variables (Feng et al., 2018). However, spatial regression modelling such as SDM sometimes has several problems, one of which is the existence of spatial outliers.

Therefore, to overcome the existence of outliers in spatial regression models, a robust regression method was developed that can generate model estimates that are resistant to outliers (Rahman & Widodo, 2018). Robust regression used to overcome the existence of spatial outliers in the SDM model is called the Robust Spatial Durbin Model (RSDM). Some estimation methods developed in robust regression include M-estimator, Least Median of Square (LMS), Least Trimmed Square (LTS), S-estimator, and MM-estimator. M-estimator has high efficiency, but the breakdown point value is 0. LMS, LTS, and S-estimator have a high breakdown point of 0.5, but the efficiency is very low. Meanwhile, MM-estimation is a combination of the high efficiency of M-estimator with the high breakdown point of S-estimator (Prahutama & Rusgiyono, 2021).

Infectious disease modelling is a part of the larger field of computer modelling, which uses computers to simulate and study the behaviour of complex systems using mathematics, physics, and computer science (Smith et al., 2018). These models contain many variables that characterize the system being studied (Johnson & Jones, 2019). The most basic mathematical models of infectious disease epidemics and endemics form the basis of the necessarily more detailed models currently used by world health organizations (Chasnov, 2022).

Infectious disease transmission models are tools for using epidemiological, biological, statistical, and mathematical techniques to describe how an epidemic may progress (Anderson & May, 1991). Moreover, epidemiology could be considered the first mathematical model applied to the prevention and control of diseases. Epidemic mathematical models have usually been applied to the prediction of outbreaks, epidemics, and transmission, as well as the resurgence of infectious disease (Gestal et al., 2019). These models have been used to inform policy decisions related to health care, public health, and beyond.

Khofifah (2022) has researched the modelling of the open unemployment rate in West Java Province using RSDM with M-estimator. The results showed that RSDM is the best model to explain the open unemployment rate in West Java Province. In addition, RSDM has also been applied to the life expectancy rate of Central Java Province using M-estimator, which shows the result that the best model used to explain the life expectancy rate in Central Java is RSDM (Hakim et al., 2020). The study focused on the application of RSDM in the economic and social fields using M-estimation, so in this study an analysis of RSDM will be carried out in other fields, namely the health sector, in this case the infectious disease tuberculosis (TB), using MM-estimator.

TB is an infectious disease caused by the mycobacterium tuberculosis. Every year, 1.5 million people die from TB, making it the top infectious killer in the world (WHO, 2021). South Sulawesi is ranked 9th out of 34 provinces in Indonesia with the highest number of TB cases (Ministry of Health, 2021). One way to reduce the number of TB cases in South Sulawesi is to know the factors that are expected to affect the number of cases. These factors also allow for variation in the number of TB cases in each district or city.

The diverse conditions of the number of TB cases in South Sulawesi Province can cause some values to have different values from other regions, potentially containing spatial outliers in the data. Based on this, this study focused on the formation of RSDM to overcome the existence of outliers in SDM and identify factors that influence the number of TB cases in South Sulawesi Province. The estimation method used in robust regression is the MM-estimator method. The results of this study can be a reference for the government and community foundations in realizing a TB-free South Sulawesi.

2. Methodology

2.1. Data and Research Variables

Data used in this study are secondary data sourced from the Health Profile of South Sulawesi Province in 2021, published by the South Sulawesi Provincial Health Office. The response variable used is the number of tuberculosis cases in each district or city in South Sulawesi Province. In this study, the predictor variables are population density (X_1), percentage of households with a clean and healthy lifestyle (X_2), percentage of BCG immunization (X_3), percentage of malnutrition (X_4), and percentage of poor population (X_5). The selection of predictor variables was based on information obtained from a literature review of several research studies related to tuberculosis cases and the results of discussions with employees at the South Sulawesi Provincial Health Office during data collection.

2.2. Spatial Regression

Spatial regression is one of the regression methods used to analyse spatial data, or data that has a location effect (spatial effect) (Anselin, 1988). The general model of spatial regression can be written as follows (Yin et al., 2018).

$$\boldsymbol{y} = \rho\mathbf{W}\boldsymbol{y} + \mathbf{X}\boldsymbol{\beta} + \lambda\mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim \mathbf{N}(0, \boldsymbol{\sigma}_{\varepsilon}^2\mathbf{I}_n) \quad (1)$$

where $\boldsymbol{y}$ is the response variable vector sized $n \times 1$, ρ is the spatial lag parameter coefficient of the response variable, $\mathbf{W}$ is the spatial weight matrix sized $n \times n$, $\mathbf{X}$ is the predictor variable matrix sized $n \times (p+1)$, $\boldsymbol{\beta}$ is the regression parameter coefficient vector sized $(p+1) \times 1$, λ is the residual spatial parameter coefficient, $\mathbf{u}$ is the residual vector that has a spatial effect sized $n \times 1$, and $\boldsymbol{\varepsilon}$ is the residual vector sized $n \times 1$.

2.3. Spatial Weighting Matrix

Each spatial weight, w_{ij}, usually reflects the spatial influence of unit j on unit i (Smith, 2008). The determination of the w_{ij} weight value is determined in two ways, namely based on distance information between locations (distance) and contiguity information (Ren & Matsumoto, 2020). The matrix used is the inverse distance weight, which is based on the centroid distance (d_{ij}) between spatial units i and j, which is generally defined as the Euclidean distance (Varatharajan et al., 2018). The value of the inverse distance location weight is obtained from the inverse of the actual distance, which is then standardized. The value of the inverse distance location weight can be determined using the following formula:

$$w_{ij}^{*} = \frac{1}{d_{ij}} \tag{2}$$

d_{ij} is the distance between the i-th location and the j-th location. The value of the inverse distance weight is then standardized using the following formula:

$$w_{ij} = \begin{cases} \dfrac{w_{ij}^{*}}{\sum_{j=1}^{n} w_{ij}^{*}}, & i \neq j \\ 0 & i = j \end{cases} \tag{3}$$

2.4. Spatial Dependency

Spatial dependence indicates the existence of assemblies between the locations of the research objects (Yasin et al., 2020). The test that can be used to identify the presence of spatial dependencies is the Moran Index (Gong et al., 2020). A positive Moran index indicates that neighbours have similar spatial effect values. Meanwhile, a negative Moran's index indicates that neighbours have different spatial effect values (Kim et al., 2022; Sun et al., 2022). The hypotheses used in this test are as follows:

H_0: $I = 0$ (there is no spatial dependency)

H_1: $I \neq 0$ (there is spatial dependency)

Moran's index can be formulated as follows (Lee & Wong, 2001):

$$I = \frac{n \sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{4}$$

The test statistic used is as follows:

$$Z(I) = \frac{I - E(I)}{\sqrt{Var(I)}} \approx N(0,1) \tag{5}$$

with,

$$E(I) = -\frac{1}{n-1}; Var(I) = \frac{n^2 S_1 - n S_2 + 3S_0^2}{(n^2-1)S_0^2} - [E(I)]^2$$

$$S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\,;\, S_1 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(w_{ij} + w_{ji}\right)^2$$

$$S_2 = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} w_{ij} + \sum_{j=1}^{n} w_{ji}\right)^2$$

where I is Moran's Index, n is the number of locations, x_i is the observation value at the i-th location, x_j is the observation value at the j-th location, $\bar{x}$ is the average observation value, w_{ij} is the standardized weight element between locations i and j, $E(I)$ is the expectation value of I, and $Var(I)$ is the variance value of I. The decision basis for this test is to reject H_0 at the significance level $\alpha = 0.05$ if $|Z(I)| > Z_{\frac{\alpha}{2}}$ or if the p-value $< \alpha$, which means there is spatial dependence.

2.5. Spatial Durbin Model

Spatial Durbin Model is a spatial regression method that has a spatial lag on the response variable (y) and predictor variables (X) (Anselin, 1988). This model has the following equation form (LeSage & Pace, 2009):

$$\boldsymbol{y} = \rho\boldsymbol{W}\boldsymbol{y} + \alpha\mathbf{1}_{\boldsymbol{n}} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \qquad (6)$$

$$\boldsymbol{\varepsilon} \sim N(0, \sigma^2\boldsymbol{I}_{\boldsymbol{n}})$$

or

$$\boldsymbol{y} = \rho\boldsymbol{W}\boldsymbol{y} + \boldsymbol{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \qquad (7)$$

$\boldsymbol{Z} = [\mathbf{1}_{\boldsymbol{n}} \quad \boldsymbol{X} \quad \boldsymbol{WX}]$, $\boldsymbol{\delta} = [\boldsymbol{\alpha} \quad \boldsymbol{\beta} \quad \boldsymbol{\theta}]^T$, α is a constant parameter, and $\boldsymbol{\theta}$ is the spatial lag parameter vector of predictor variables of size $p \times 1$.

2.6. Spatial Outlier

An outlier is an observation that appears different from other observations in the data set (Barnett & Lewis, 1994). Moran's Scatterplot is a graph that shows the relationship between the observed value at a location and the average value of its neighbouring observations. Data plots in the upper left and lower right quadrants indicate the presence of negative spatial autocorrelation, which indicates that low observation areas are surrounded by high observation areas and reversed. Therefore, it can be identified that these points are categorized as spatial outliers.

Besides graphical observation, spatial outlier detection using Moran's Scatterplot can also be identified mathematically based on LISA (Local Indicator of Spatial Autocorrelation). LISA is used to identify autocorrelation locally at each location (Xiao et al., 2018). LISA is defined by the following equation:

$$L_i = \left(Z_i \times \left(\sum\nolimits_j \left(w_{ij}Z_j\right)\right)\right) \qquad (8)$$

$$Z_i = \frac{x_i - \mu}{\sigma}\,;\, Z_j = \frac{x_j - \mu}{\sigma};\, I_i = \sum\nolimits_j \left(\mathrm{w}_{ij}Z_j\right)$$

where Z_i is the standardized observation value at the i-th location, Z_j is the standardized observation value at the j-th location, I_i is the standardized observation value multiplied by the weight matrix element standardized between locations i and j, μ is the average observation value, and σ is the standard deviation of the observation value. If the value of $L_i < 0$, it can be concluded that the point is a spatial outlier (Shekhar et al., 2003).

2.7. Robust MM-Estimator

The MM-estimator is a combination of a method with a high breakdown value (LTS-estimator or S-estimator) and the high efficiency of the M-estimator (Yohai, 1987). The MM-estimator is defined as follows:

$$\hat{\beta}_{MM} = \min \sum_{i=1}^{n} \rho\left(\frac{e_i}{\hat{\sigma}}\right) = \min \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=0}^{p} X_i \beta_j}{\hat{\sigma}}\right) \tag{9}$$

The objective function used as a weighting function is Tukey Bisquare, which is defined as follows:

$$w(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 & \text{for } |u_i| \le c \\ 0 & \text{for } |u_i| > c \end{cases} \tag{10}$$

The value of c in the Tukey Bisquare method is called the tuning constant. The tuning constant value for the Tukey Bisquare weighting function in the S-estimator robust regression analysis is c = 1.547, and the tuning constant for the M-Estimator is c = 4.685 (Fox & Weisberg, 2010).

2.8. Parameter Estimation of Robust Spatial Durbin Regression Model

The robust regression parameter estimation method used is the MM-Estimator (Method of Moment). In general, the MM-Estimator is performed by minimizing the objective function shown in equation (11).

$$\sum_{i=1}^{n} \rho\left(\frac{e_i}{\hat{\sigma}}\right) = \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=0}^{p} X_i \beta_j}{\hat{\sigma}}\right) \tag{11}$$

$\rho' = \psi$ is an influence function so equation (11) can be written as follows:

$$\sum_{i=1}^{n} X_{ij} \psi\left(y_i - \sum_{j=0}^{p} X_i \beta_j\right) = 0 \tag{12}$$

Equation (12) can be written as

$$\sum_{i=1}^{n} X_{ij} w_i \left(y_i - \sum_{j=0}^{p} X_i \beta_j\right) = 0 \tag{13}$$

Equation (13) is solved by Iteratively Reweighted Least Square (IRLS) method, and the initial estimates of coefficients $\hat{\beta}^{(1)}$ and residuals $e_i^{(1)}$ are derived from robust regression with a high breakdown point (S-Estimator) for initial weights $w_i^{(1)}$, hence:

$$\sum_{i=1}^{n} X_{ij} w_i \, y_i - \sum_{i=1}^{n} X_{ij} w_i \sum_{j=0}^{p} X_{ij} \beta_j = 0 \tag{14}$$

Equation (14) in matrix form can be written as follows:

$$\boldsymbol{X'\omega y} = \boldsymbol{X'\omega X\widehat{\beta}_j} \tag{15}$$

For each $\boldsymbol{\omega}$, the estimator $\boldsymbol{\widehat{\beta}_j}$ from equation (15) is obtained as follows:

$$\boldsymbol{\widehat{\beta}}_{MM} = (\boldsymbol{X'\omega X})^{-1}\boldsymbol{X'\omega y} \tag{16}$$

where $\boldsymbol{\omega}$ is an $n \times n$ diagonal elements w_1, w_2, w_3,..., w_n and n is the number of observations, $\boldsymbol{X'}$ is the transpose of $\boldsymbol{X}$, i.e. the $n \times p$ matrix of predictor variables. For a given weight $w^{(m)}$, the estimator $\boldsymbol{\widehat{\beta}}_{MM}^{m+1} = \left(\boldsymbol{X'\omega}^{(m)}\boldsymbol{X}\right)^{-1}\boldsymbol{X'\omega}^{(m)}\boldsymbol{y}$ can be obtained until $\sum_{i=0}^{n}\left|e_i^{(m)}\right|$ converges, i.e. the difference between the values of $\hat{\beta}_j^{(m+1)}$ and $\hat{\beta}_j^{(m)}$ approaches 0, where m is the number of iterations.

RSDM parameter estimation is obtained using Ordinary Least Square (OLS), which is obtained by minimizing the residual sum of squares:

$$\sum_{i=1}^{n} \boldsymbol{\varepsilon}_i^2 = \big((\boldsymbol{I} - \rho\boldsymbol{W})\boldsymbol{y}\big)^T(\boldsymbol{I} - \rho\boldsymbol{W})\boldsymbol{y} - 2\boldsymbol{\delta}^T\boldsymbol{Z}^T\big((\boldsymbol{I} - \boldsymbol{\rho W})\boldsymbol{y}\big) + \boldsymbol{\delta}^T\boldsymbol{Z}^T\boldsymbol{Z\delta} \tag{17}$$

so that the parameter estimate $\boldsymbol{\widehat{\delta}}$ is given as follows:

$$\boldsymbol{\widehat{\delta}}_{OLS} = (\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\boldsymbol{Z}^T(\boldsymbol{I} - \rho\boldsymbol{W})\boldsymbol{y} \tag{18}$$

The estimated value of y is $\boldsymbol{\widehat{y}} = \boldsymbol{X\widehat{\beta}}$, so that by substituting equation (16) into equation (18) we have

$$\boldsymbol{\widehat{\delta}}_{RSDM} = (\boldsymbol{Z}^T\boldsymbol{\omega Z})^{-1}\boldsymbol{Z}^T\boldsymbol{\omega}(\boldsymbol{I} - \rho\boldsymbol{W})\boldsymbol{y} \tag{19}$$

3. Result and Discussion

3.1. Data Exploration

The highest number of tuberculosis cases in South Sulawesi in 2020 was in Makassar City with 5,421 patients, while the lowest number of tuberculosis cases was in Selayar Islands Regency with 220 patients. The average number of TB cases in South Sulawesi in 2020 was 786 patients. The number of tuberculosis cases in the high category is

illustrated in red with 1 district or city, the medium category is illustrated in orange with 2 districts or cities, and the low category is illustrated in green with 21 districts or cities.

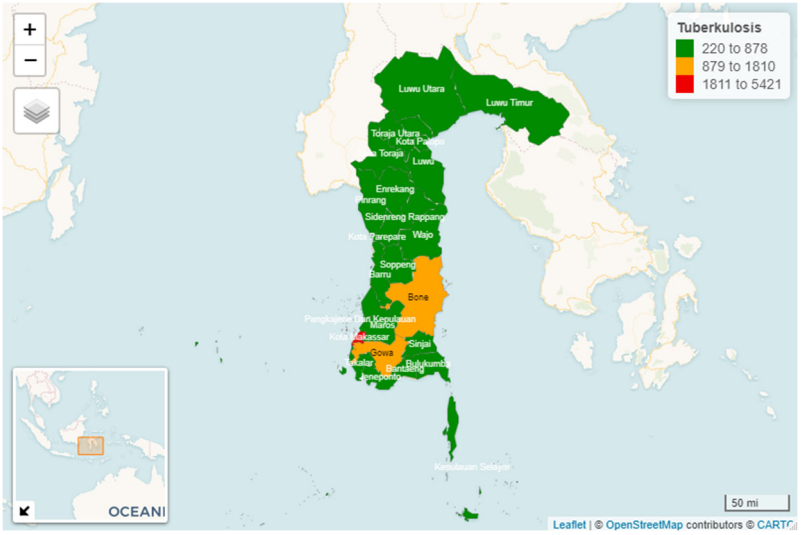


Figure 1: Tuberculosis case distribution map in South Sulawesi

Source: Data processed.

3.2. Spatial Dependency Effects

The tests of global spatial dependency were conducted using Moran's Index with and inverse distance weighting matrix. The results of the Moran Index test on response variables, predictor variables, and regression residuals are presented in Table 1.

Table 1: Spatial Dependence Test Results

Variable	Moran Index	*p-value*	Decision ($\alpha = 0.05$)
Y	0.0059	0.0120	Reject H_0
X_1	-0.0400	0.4001	Accept H_0
X_2	-0.0066	0.1924	Accept H_0
X_3	-0.0051	0.1784	Accept H_0
X_4	-0.0850	0.8489	Accept H_0
X_5	0.2175	0.0000	Reject H_0
Residual	-0.0737	0.7834	Accept H_0

Source: Data processed.

Based on Table 1, it can be observed that the response variable (y) and the percentage of poor population variable (X_5) have a p-value smaller than $\alpha = 0.05$,

which means that there is spatial dependency on the response variable and predictor variables. Therefore, the SDM method can be applied to this study because there is spatial dependency in both the response variable and one of the predictor variables.

3.3. Spatial Outlier Detection

Outlier detection in this study was conducted using Moran's Scatterplot test.

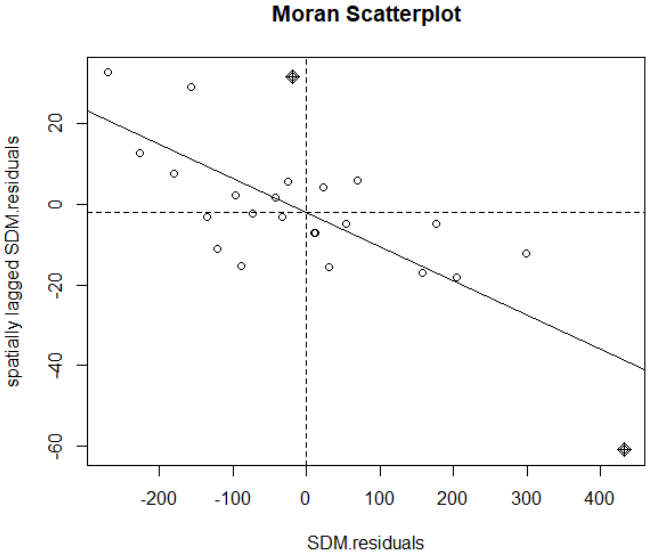


Figure 2: Moran's Scatterplot

Source: Data processed.

Based on Figure 2, it can be concluded that the residual model has 17 observations located in Quadrant II and Quadrant IV, so it can be identified that the observation points can be categorized as spatial outliers. The existence of observations that are indicated as outliers can lead to inaccuracies in the estimation of intercept values and regression parameter coefficients. Therefore, estimation of regression parameter coefficient values that are resistant to outliers using the Robust Spatial Durbin Model (RSDM) is required in this study.

3.4. Robust Spatial Durbin Modelling with MM-Estimator

The preliminary weighting in the estimation uses the residuals from the regression equation model using the S-estimator that has converged. After the weighting process of the S-estimator to get a convergent regression model, the residuals are then used for weighting using the M-estimator method. Weighting using the M-estimator continues until a convergent regression model is obtained. The results of RSDM parameter estimation with MM-Estimator are presented in Table 2.

Table 2. Results of RSDM parameter estimation with MM-Estimator

Parameter	Estimated Value	
	49th Iteration	50th Iteration
$\boldsymbol{\alpha}$	16739.550	16739.550
$\boldsymbol{\beta_1}$	0.833	0.833
$\boldsymbol{\beta_2}$	-13.707	-13.707
$\boldsymbol{\beta_3}$	-14.442	-14.442
$\boldsymbol{\beta_4}$	-94.647	-94.647
$\boldsymbol{\beta_5}$	12.192	12.192
$\boldsymbol{\theta_1}$	4.133	4.133
$\boldsymbol{\theta_2}$	-153.734	-153.734
$\boldsymbol{\theta_3}$	4.055	4.055
$\boldsymbol{\theta_4}$	-279.776	-279.776
$\boldsymbol{\theta_5}$	-503.770	-503.770

Source: Data processed.

Based on the estimation results after 50 iterations, the Robust Spatial Durbin Model using the MM-Estimator is as follows:

$$\hat{y}_i = -3.781\sum_{j=1}^{n} w_{ij}y_j + 16739.550 + 0.833x_{i1} - 13.707x_{i2} - 14.442x_{i3} - 94.647x_{i4} + 12.192x_{i5} + 4.133\sum_{j=1}^{n} w_{ij}x_{i1} - 153.734\sum_{j=1}^{n} w_{ij}x_{i2} + 4.055\sum_{j=1}^{n} w_{ij}x_{i3} - 279.776\sum_{j=1}^{n} w_{ij}x_{i4} - 503.770\sum_{j=1}^{n} w_{ij}x_{i5}$$

$\sum_{j=1}^{n} w_{ij}$ indicates the existence of location effects between the observed district (i) and its neighbours (j).

Based on the calculation results, the $F_{count} = 366.4118$ is greater than the $F_{0.05;5,18} = 2.773$. Then the p-value of $0.0000 < \alpha = 0.05$ was also obtained, so the decision was obtained that H_0 was rejected, which means that RSDM is a suitable model to explain the relationship between response variables and predictor variables.

Based on the parameter significance test using the Wald test, the results show that the intercept α, the response variable spatial lag parameter (ρ), the regression parameters β_1, β_2, β_3, β_4, and β_6, and the predictor variable spatial lag parameters θ_1, θ_2, θ_4, and θ_5 have a significant effect on the model because the Wald value is greater than $\chi^2_{0.05,1} = 3.841$. Meanwhile, the parameters β_5 and θ_3 have no significant effect on the model because the Wald value is smaller than $\chi^2_{0.05,1} = 3.841$.

Table 3: Significance Test of RSDM Parameters

Parameter	Wald	*p-value*	Decision
ρ	11.7246	0.0006**	Reject H_0
α	27.1879	0.0000**	Reject H_0
β_1	41.2240	0.0000**	Reject H_0
β_2	6.4720	0.0110**	Reject H_0
β_3	34.2704	0.0000**	Reject H_0
β_4	93.2526	0.0000**	Reject H_0
β_5	1.0353	0.3089	Accept H_0
θ_1	28.4686	0.0000**	Reject H_0
θ_2	11.9907	0.0005**	Reject H_0
θ_3	0.0238	0.8773	Accept H_0
θ_4	6.8884	0.0087**	Reject H_0
θ_5	20.6021	0.0000**	Reject H_0

**) Significant at level $\alpha = 0.05$

Source: Data processed.

A measure of model goodness is determined based on the R^2_{adj} value and MSE value generated from the SDM model and RSDM.

Table 4: Comparison of Model Goodness Measures

Model	R^2_{adj}	MSE
SDM	0.9686	34618.31
RSDM	0.9952	6461.734

Source: Data processed.

A better model will result in a larger R^2_{adj} value and a smaller MSE value. Based on Table 4, the best spatial regression model for explaining the number of tuberculosis cases in South Sulawesi Province in 2020 is RSDM because it has a larger R^2_{adj} value and a smaller MSE value than SDM.

Outliers in the spatial regression model adversely affect the coefficient value of the spatial regression parameter estimator. Modelling results by applying the principle of the MM-Estimator method to the estimation of RSDM regression parameters is able to accommodate the presence of outlier observations in the Spatial Durbin Model regression model quite effectively. This is indicated by an improvement in the coefficient value of the RSDM parameter estimator and is able to reduce the MSE value and increase the R^2_{adj} value so that the RSDM model is considered preferable for modeling the number of tuberculosis cases in South Sulawesi Province in 2020.

The Robust Spatial Durbin Model (RSDM) and the Spatial Durbin Model (SDM) are both spatial econometric models used to analyse the relationship between a dependent variable and several independent variables, taking into account spatial effects. In the context of modelling the number of tuberculosis cases, these models can help identify spatial dependencies and interactions among different regions. For instance, the number of tuberculosis cases in one region might be influenced by the number of cases in neighbouring regions, as well as other factors such as population density, healthcare facilities, and socioeconomic conditions.

The key difference between RSDM and SDM lies in their robustness to outliers. In SDM, the presence of outlier data can bias the parameter estimates. This means that if there are regions with exceptionally high or low numbers of tuberculosis cases, these could disproportionately influence the model's results, leading to inaccurate predictions. On the other hand, RSDM is designed to be more robust to outliers. This means that it can provide more stable parameter estimates even when outlier data are present. As a result, RSDM might be more suitable for datasets where outliers are expected or unavoidable.

The results of this study indicate that the Robust Spatial Durbin Model (RSDM) performs better than the Spatial Durbin Model (SDM) in explaining the number of tuberculosis cases in South Sulawesi Province in 2020. However, it is important to note that different spatial models may be more suitable depending on the specific context and data available. Different models may have different assumptions and may be more or less suitable depending on factors such as the nature of the spatial dependencies, the quality and granularity of the data, and the specific research questions being addressed.

For instance, a study by Amaral et al. (2022) proposed a spatio-temporal modelling framework to characterize infectious disease dynamics by integrating the SIR compartment and log-Gaussian Cox process (LGCP) models. This approach was used to analyse COVID-19 dynamics and showed better forecasting performance than equivalent models that did not integrate information from the SIR modelling step. Another study conducted a systematic review of spatial studies of TB and found that almost all studies used notification data for spatial analysis. The most common data visualization technique was notification rate mapping, and the use of smoothing techniques was uncommon. Spatial clusters were identified using a range of methods, with the most commonly employed being Kulldorff's spatial scan statistic followed by local Moran's I and Getis and Ord's local Gi (d) tests (Shaweno et al., 2018).

In all included studies, TB showed a heterogeneous spatial pattern at each geographic resolution level examined. This suggests that while RSDM shows better performance than SDM in this study, further research could explore the application and performance of other spatial models for tuberculosis or infectious diseases in general. It is also worth noting that different diseases may require different modelling

approaches. Infectious disease modelling is a complex field that requires the integration of tools from various disciplines such as epidemiology, statistics, ecology, evolutionary biology, sociology, economics, and more. Each disease has unique characteristics and may spread in different ways, requiring different modelling approaches. For instance, a study by Chowell and Rothenberg (2018) noted that modern quantitative computational tools and highly resolved geospatial demographic, epidemiological, and genomic data are enabling actionable insights for public health in near real time. These tools have been used to model various infectious diseases such as influenza, dengue, chikungunya, Zika, and HIV.

4. Conclusions

This study describes the modelling of a Robust Spatial Durbin Model (RSDM) on tuberculosis data in South Sulawesi Province. The results indicate that RSDM performs better than the Spatial Durbin Model (SDM) in explaining the number of tuberculosis cases in South Sulawesi Province in 2020, as it has a larger Adjusted R^2 value and a smaller MSE value than SDM. Factors that significantly affect the number of tuberculosis cases in South Sulawesi Province in 2020 are population density, the percentage of households with PHBS, the percentage of BCG (Bacillus Calmette-Guérin) immunization, and the percentage of malnutrition. This can be a reference for the health department and society to take anticipative steps in reducing the rate of tuberculosis spread in each region in South Sulawesi Province.

In terms of comparison with other spatial models for tuberculosis or infectious diseases in general, it is important to note that different spatial models may be more appropriate depending on the specific context and available data. For instance, some studies have used different types of spatial models such as the Spatial Autoregressive Model (SAR) and various spatio-temporal models. The models also consider various factors and use different methodologies, which could lead to different results. Therefore, while RSDM showed better performance than SDM in this study, it is important to consider other spatial models and methodologies when studying infectious diseases. Further research is needed to compare the performance of other spatial models for Tuberculosis or in different contexts and with different infectious diseases.

Acknowledgement

We are thankful to the South Sulawesi Provincial Health Office for giving us the opportunity to collect data on tuberculosis cases in South Sulawesi.

References

Anderson, R. M., May, R. M., (1991). *Infectious Disease of Humans: Dynamic and Control*. Oxford University Press.

Anselin, L., (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers.

Barnett, V., Lewis, T., (1994). *Outlier In Statistical Data* (3rd ed.). Jhon Wiley & Sons, Inc.

Chasnov, J. R., (2022*). Infectious Disease Modeling*. In: Mathematical Biology. LibreTexts.

Chowell G., Rothenberg R. (2018). Spatial infectious disease epidemiology: on the cusp. *BMC Medicine*, Vol. 16, pp. 1-5.

Feng, Y., Wang, X., Du, W., and Liu, J., (2018). Effects of Air Pollution Control on Urban Development Quality in Chinese Cities Based on Spatial Durbin Model. *International Journal of Environmental Research and Public Health*, Vol. 15, pp. 1–21.

Fox, J., Weisberg, S., (2010). *Robust Regression in R: An Appendix to An R to Applied Regression* (2nd ed.).

Gestal, M. C., Dedloff, M. R., and Torres-Sangiao, E., (2019). Computational Health Engineering Applied to Model Infectious Diseases and Antimicrobial Resistance Spread. *Applied Sciences*, Vol. 9, pp. 1–19.

Gong, Z., Kong, Q., and Tan, J., (2020). Analysis on Multidimensional Poverty Reduction of Industrial Structure Upgrading Based on Provincial Panel Data and Spatial Durbin Model. *International Journal of Sustainable Development and Planning*, Vol. 15, pp. 1197–1204.

Hakim, A. R., Warsito, B., and Yasin, H., (2020). Live Expectancy Modelling using Spatial Durbin Robust Model. *Journal of Physics: Conference Series*, Vol. 1655, pp. 1–11.

Johnson, M., Jones, K., (2019). Variables in Infections Disease Modeling. *Journal of Epidemiology and Public Health*, Vol. 7, pp. 98–105.

Ministry of Health of the Republic of Indonesia, (2021). *Profil Kesehatan Indonesia 2021*. Kementrian Kesehatan Republik Indonesia.

Khofifah, H. N., (2022). Robust Spatial Durbin Model (RSDM) untuk Pemodelan Tingkat Pengangguran Terbuka (TPT) di Provinsi Jawa Barat. *Jurnal Riset Statistika*, Vol. 1, pp. 135–142.

Kim, J. J., Cho, M. J., and Lee, M. H., (2022). An Analysis of the Price Determinants of Multiplex Houses through Spatial Regression Analysis. *Sustainability*, Vol. 14, pp. 1–14.

Lee, J., Wong, D. W. S., (2001). *Statistical Analysis with Arcview GIS*. John Wiley & Sons, Inc.

LeSage, J., Pace, R. K., (2009). *Introduction to Spatial Econometrics*. Taylor & Francis Group.

Prahutama, A., Rusgiyono, A., (2021). Robust Regression with MM-Estimator for Modelling The Number Maternal Mortality of Pregnancy in Central Java, Indonesia. *Journal of Physics: Conference Series*, Vol. 1943, pp. 1–7.

Rahman, M. B., Widodo, E., (2018). Perbandingan Metode Regresi Robust Estimasi Least Trimmed Square, Estimasi Scale, dan Estimasi Method Of Moment. *Prosiding Seminar Nasional Matematika*, pp. 426–433.

Ramadhani, E., Salwa, N., and Mazaya, M. S., (2020). Identifikasi Faktor-Faktor yang Memengaruhi Angka Harapan Hidup di Sumatera Tahun 2018 Menggunakan Analisis Regresi Spasial Pendekatan Area. *Journal of Data Analysis*, Vol. 3, pp. 62–75.

Ren, L., Matsumoto, K., (2020). Effects of Socioeconomic and Natural Factors on Air Pollution in China: A Spatial Panel Data Analysis. *Science of the Total Environment*, Vol. 740, pp. 1–10.

Shaweno D., Karmakar M., Alene K.A., Ragonnet R., Clements A. C. A., Trauer J. M., Denholm J. T., and McBryde, E. S., (2018). Methods used in the spatial analysis of tuberculosis epidemiology: a systematic review. *BMC Medicine*. Vol. 16, pp. 1–18.

Shekhar, S., Lu, C.-T., and Zhang, P., (2003). A Unified Approach to Detecting Spatial Outliers. *GeoInformatica*, Vol. 7, pp. 139–166.

Smith, J., Brown, L., and Green, D. (2018). Computer Modeling in Disease Prevention. *Journal of Health Informatics*, Vol. 12, pp. 201-210.

Smith, T. E. (2008). *Encyclopedia of GIS* (S. Shekhar & H. Xiong, Eds.), Springer US.

Sun, Y., Du, M., Wu, L., Li, C., and Chen, Y., (2022). Evaluating the Effects of Renewable Energy Consumption on Carbon Emissions of China's Provinces: Based on Spatial Durbin Model. *Land*, Vol. 11, pp. 1–17.

Varatharajan, R., Manogaran, G., Priyan, M. K., Balaş, V. E., and Barna, C., (2018). Visual Analysis of Geospatial Habitat Suitability Model Based on Inverse Distance Weighting with Paired Comparison Analysis. *Multimedia Tools and Applications*, Vol. 77, pp. 17573–17593.

WHO, (2021, October 14). *Tuberculosis*. https://www.who.int/news-room/fact-sheets/detail/tuberculosis.

Xiao, G., Hu, Y., Li, N., and Yang, D., (2018). Spatial Autocorrelation Analysis of Monitoring Data of Heavy Metals in Rice in China. *Food Control*, Vol. 89, pp. 32–37.

Yasin, H., Hakim, A. R., and Warsito, B., (2020). Development Life Expectancy Model in Central Java Using Robust Spatial Regression With M-Estimators. *Communications in Mathematical Biology and Neuroscience*, Vol. 2020, pp. 1–16.

Yohai, V. J., (1987). High Breakdown-Point and High Efficiency Robust Estimates for Regression. *The Annals of Statistic*, Vol. 15, pp. 642–656.

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 39–55, https://doi.org/10.59170/stattrans-2024-014
Received – 25.02.2022; accepted – 16.11.2023

Ratio-type estimator of the population mean in stratified sampling based on the calibration approach

Menakshi Pachori[1], **Neha Garg**[2]

Abstract

An improved ratio-type calibrated estimator was developed using the logarithmic mean in the calibration constraint for the stratified random sampling scheme. The proposed estimator was extended in the case of stratified double sampling and compared with the estimators given by Tracy et al. (2003) together its ratio-type estimators, as well as Nidhi et al. (2017) and Khare et al. (2022). A simulation study was also carried out on both a real and artificial dataset in order to evaluate the performance of the proposed estimator compared to the existing estimators.

Key words: auxiliary information, calibration estimation, stratified random sampling, stratified double sampling, ratio estimator.

1. Introduction

Calibration estimation method provides more precise results by making use of the available auxiliary information. Deville and Sarndal (1992) defined it as a procedure of minimizing a distance function subject to calibration constraints. In this direction, various calibration estimators for various sampling schemes for estimating the population parameters by using different calibration constraints based on the available auxiliary information have been developed by Singh (2003), Kim et al. (2007), Singh and Arnab (2011), Koyuncu and Kadilar (2013 and 2016), Mouhamed et al. (2015), Clement and Enang (2017), Nidhi et al. (2017), Garg and Pachori (2019, 2020 and 2021), Alka et al. (2021), Khare et al. (2022), Pachori et al. (2022), Kumar et al. (2023), etc.

It is a well-known fact in the sample surveys that auxiliary information can be used to increase the precision of the estimators of the population parameters. When study variable Y is highly positively correlated with the available auxiliary variable X, the usual

[1] School of Sciences, Indira Gandhi National Open University, New Delhi-110068, India.
E-mail: pachorim07@gmail.com. ORCID: https://orcid.org/0000-0001-5917-9648.
[2] Corresponding author. School of Sciences, Indira Gandhi National Open University, New Delhi-110068, India. E-mail: nehagarg@ignou.ac.in. ORCID: https://orcid.org/0000-0003-3123-1832.

ratio method of estimation is considered. For example, for estimating total/average yield of a crop (Y) and area of cultivation (X); sale of the departmental stores for current year (Y) and sales recorded for past years (X); volume of the tree (Y) and diameter of the tree (X); sugarcane production in the current year (Y) and sugarcane production in past surveys (X), etc. In the case of heterogeneous population, the stratified random sampling is preferred than the simple random sampling. This article deals with the problem of estimation of population mean of the study variable under stratified random sampling utilizing available auxiliary information with the help of the calibration approach.

In this article, we have suggested a new ratio-type calibration estimator using logarithmic mean of the available auxiliary variable as a calibration constraint under stratified and stratified double sampling for estimating the population mean of the study variable. A simulation study has been conducted on a real as well as an artificial dataset in order to support the study and the result has been compared with the estimator suggested by Tracy et al. (2003) and its ratio-type estimator along with Nidhi et al. (2017) and Khare et al. (2022).

1.1. Calibration estimator under stratified sampling

In stratified random sampling, the population of size N is divided into L homogeneous strata consisting of N_h units in the h[th] stratum such that $\sum_{h=1}^{L} N_h = N$. A sample of size n_h is drawn from the h[th] stratum using Simple Random Sampling Without Replacement (SRSWOR), where *n* is the required sample size such that $\sum_{h=1}^{L} n_h = n$.

Suppose y_{hi} and x_{hi} is the i[th] unit of the study and auxiliary variables, respectively, in the h[th] stratum for $i = 1, 2, ..., n_h$ and $h = 1, 2, ..., L$. Let $W_h = \frac{N_h}{N}$ be the stratum weight and $f_h = \frac{n_h}{N_h}$, the sample fraction.

The calibration estimator for population mean under stratified random sampling given by Tracy et al. (2003) is given as:

$$\bar{y}_{tr} = \sum_{h=1}^{L} \Omega_h \bar{y}_h \tag{1}$$

where Ω_h are the new calibrated weights obtained by minimizing the Chi-square type distance measure, subject to the two calibration constraints:

$$\sum_{h=1}^{L} \Omega_h \bar{x}_h = \sum_{h=1}^{L} W_h \bar{X}_h \tag{2}$$

$$\sum_{h=1}^{L}\Omega_h s_{hx}^2 = \sum_{h=1}^{L} W_h S_{hx}^2 \tag{3}$$

where $\bar{x}_h = \frac{\sum_{i=1}^{n_h} x_{hi}}{n_h}$ and $\bar{X}_h = \frac{\sum_{i=1}^{n_h} X_{hi}}{n_h}$ are the h^{th} stratum sample and population means of the auxiliary variable, respectively.

$s_{hx}^2 = \sum_{i=1}^{n_h} \frac{(x_{hi} - \bar{x}_h)^2}{(n_h - 1)}$ and $S_{hx}^2 = \sum_{i=1}^{N_h} \frac{(X_{hi} - \bar{X}_h)^2}{(N_h - 1)}$ are the h^{th} stratum sample and population variances of the auxiliary variable, respectively.

The calibrated weights are obtained as:

$$\Omega_h = W_h + W_h Q_h \bar{x}_h \frac{\left(\sum_{h=1}^{L} W_h Q_h S_{hx}^4\right)\left(\sum_{h=1}^{L} W_h\left(\bar{X}_h - \bar{x}_h\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h S_{hx}^2\right)\left(\sum_{h=1}^{L} W_h\left(S_{hx}^2 - S_{hx}^2\right)\right)\right)}{\left(\sum_{h=1}^{L} W_h Q_h S_{hx}^4\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h S_{hx}^2\right)^2}$$

$$+ W_h Q_h s_{hx}^2 \frac{\left(\sum_{h=1}^{L} W_h\left(S_{hx}^2 - s_{hx}^2\right)\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2 s_{hx}^2\right)\left(\sum_{h=1}^{L} W_h\left(\bar{X}_h - \bar{x}_h\right)\right)}{\left(\sum_{h=1}^{L} W_h Q_h s_{hx}^4\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h s_{hx}^2\right)^2}$$

The calibrated estimator given by Tracy et al. (2003) is

$$\bar{y}_{tr} = \sum_{h=1}^{L} W_h \bar{y}_h + \hat{\beta}_{t1}\left(\sum_{h=1}^{L} W_h\left(\bar{X}_h - \bar{x}_h\right)\right) + \hat{\beta}_{t2}\left(\sum_{h=1}^{L} W_h\left(S_{hx}^2 - s_{hx}^2\right)\right) \tag{4}$$

where

$$\hat{\beta}_{t1} = \left[\frac{\left(\sum_{h=1}^{L} W_h Q_h s_{hx}^4\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h \bar{y}_h\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h s_{hx}^2\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{y}_h s_{hx}^2\right)}{\left(\sum_{h=1}^{L} W_h Q_h s_{hx}^4\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h s_{hx}^2\right)^2}\right]$$

$$\hat{\beta}_{t2} = \left[\frac{\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{y}_h s_{hx}^2\right) - \left(\sum_{h=1}^{L} W_h Q_h s_{hx}^2\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h \bar{y}_h\right)}{\left(\sum_{h=1}^{L} W_h Q_h s_{hx}^4\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h s_{hx}^2\right)^2}\right]$$

To estimate the population mean $\bar{Y}$ under the stratified random sampling, the calibration estimator considered by Nidhi et al. (2017) is:

$$\bar{y}_s = \sum_{h=1}^{L} \Omega_h \bar{y}_h \tag{5}$$

subject to the two calibration constraints:

$$\sum_{h=1}^{L}\Omega_h = 1 \tag{6}$$

$$\sum_{h=1}^{L}\Omega_h \bar{x}_h = \sum_{h=1}^{L} W_h \bar{X}_h \tag{7}$$

The calibrated weight after minimization of Chi-square type distance measure subject to the calibration constraints is given as:

$$\Omega_h = W_h + \left[\frac{\left(W_h Q_h \bar{x}_h\right)\left(\sum_{h=1}^{L} W_h Q_h\right) - \left(W_h Q_h\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h\right)}{\left(\sum_{h=1}^{L} W_h Q_h x_h^2\right)\left(\sum_{h=1}^{L} W_h Q_h\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h\right)^2}\right]\left(\bar{X} - \sum_{h=1}^{L} W_h \bar{x}_h\right)$$

The estimator given by Nidhi et al. (2017) is:

$$\bar{y}_s = \sum_{h=1}^{L} W_h \bar{y}_h + \hat{\beta}_s\left(\bar{X} - \sum_{h=1}^{L} W_h \bar{x}_h\right) \tag{8}$$

where

$$\hat{\beta}_s = \left[\frac{\left(\sum_{h=1}^{L} W_h Q_h\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h \bar{y}_h\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{y}_h\right)}{\left(\sum_{h=1}^{L} W_h Q_h\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h\right)^2}\right]$$

1.2. Calibration Ratio Estimator under Stratified Sampling

If $\bar{y}$ is the sample mean of the study variable Y and $\bar{x}$ is the sample mean of the auxiliary variable X, the usual ratio estimator is given as:

$$\bar{y}_R = \frac{\bar{y}}{\bar{x}}\bar{X}$$

Let $\bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{hi}$ and $\bar{y}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} y_{hi}$ be the sample means of auxiliary variable and study variable, respectively. Now, we can define the ratio-type calibration estimator in stratified random sampling considering the constraints given by Tracy et al. (2003) as:

$$\bar{y}_{r\_tr} = \sum_{h=1}^{L}\Omega_h \bar{y}_{Rh} \tag{9}$$

The calibrated weights are obtained for the estimator defined in equation (9) as:

$$\Omega_h = W_h + W_h Q_h \bar{x}_h \frac{\left(\sum_{h=1}^{L} W_h Q_h S_{hx}^4\right)\left(\sum_{h=1}^{L} W_h\left(\bar{X}_h - \bar{x}_h\right)\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h S_{hx}^2\right)\left(\sum_{h=1}^{L} W_h\left(S_{hx}^2 - S_{hx}^2\right)\right)}{\left(\sum_{h=1}^{L} W_h Q_h S_{hx}^4\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h S_{hx}^2\right)^2}$$

$$+W_h Q_h s_{hx}^2 \frac{\left(\sum_{h=1}^{L} W_h \left(S_{hx}^2 - s_{hx}^2\right)\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h s_{hx}^2\right)\left(\sum_{h=1}^{L} W_h \left(\bar{X}_h - \bar{x}_h\right)\right)}{\left(\sum_{h=1}^{L} W_h Q_h s_{hx}^4\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h s_{hx}^2\right)^2}$$

Thus, the ratio-type calibrated estimator considering the constraints given by Tracy et al. (2003) is:

$$\bar{y}_{r\_tr} = \sum_{h=1}^{L} W_h y_{Rh} + \hat{\beta}_{r\_t1}\left(\sum_{h=1}^{L} W_h \left(\bar{X}_h - \bar{x}_h\right)\right) + \hat{\beta}_{r\_t2}\left(\sum_{h=1}^{L} W_h \left(S_{hx}^2 - s_{hx}^2\right)\right) \quad (10)$$

where

$$\hat{\beta}_{r\_t1} = \left[\frac{\left(\sum_{h=1}^{L} W_h Q_h s_{hx}^4\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h \bar{y}_h\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h s_{hx}^2\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{y}_h s_{hx}^2\right)}{\left(\sum_{h=1}^{L} W_h Q_h s_{hx}^4\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h s_{hx}^2\right)^2}\right]$$

$$\hat{\beta}_{r\_t2} = \left[\frac{\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{y}_h s_{hx}^2\right) - \left(\sum_{h=1}^{L} W_h Q_h s_{hx}^2\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h \bar{y}_h\right)}{\left(\sum_{h=1}^{L} W_h Q_h s_{hx}^4\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right) - \left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h s_{hx}^2\right)^2}\right]$$

The calibration ratio estimator suggested by Khare et al. (2022) considering the constraints given in Nidhi et al. (2017) is:

$$\bar{y}_{r\_s} = \sum_{h=1}^{L} \Omega_h \bar{y}_{Rh} \quad (11)$$

Minimization of the Chi-square distance measure subject to the calibration constraints leads to the calibrated weight given as:

$$\Omega_h = W_h + \left(W_h Q_h \bar{x}_h\right)\left[\frac{\left(\sum_{h=1}^{L} W_h \bar{X}_h - \sum_{h=1}^{L} W_h \bar{x}_h\right)}{\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right)}\right] \quad (12)$$

Thus, the ratio-type calibration estimator given Khare et al. (2022) is

$$\bar{y}_{r\_s} = \sum_{h=1}^{L} W_h \bar{y}_{Rh} + \hat{\beta}_{r\_s}\left(\sum_{h=1}^{L} W_h \bar{X}_h - \sum_{h=1}^{L} W_h \bar{x}_h\right) \quad (13)$$

where $$\hat{\beta}_{r\_s} = \left[\frac{\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h \bar{y}_{Rh}\right)}{\left(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2\right)}\right]$$

2. Proposed Calibration Estimator

2.1. Stratified Random Sampling

We propose a new ratio-type calibration estimator for stratified random sampling as:

$$\bar{y}_{r\_L} = \sum_{h=1}^{L} \Omega_h \bar{y}_{Rh} \tag{14}$$

where the calibration weight Ω_h is so chosen in order to minimize the Chi-square type distance measure given as:

$$\sum_{h=1}^{L} \frac{(\Omega_h - W_h)^2}{W_h Q_h} \tag{15}$$

subject to the following calibration constraints using logarithmic mean of the auxiliary variable:

$$\sum_{h=1}^{L} \Omega_h = \sum_{h=1}^{L} W_h \tag{16}$$

$$\sum_{h=1}^{L} \Omega_h \log \bar{x}_h = \sum_{h=1}^{L} W_h \log \bar{X}_h \tag{17}$$

The Lagrange function is given as:

$$L = \sum_{h=1}^{L} \frac{(\Omega_h - W_h)^2}{Q_h W_h} - 2\lambda_1 \left(\sum_{h=1}^{L} \Omega_h - \sum_{h=1}^{L} W_h \right) - 2\lambda_2 \left(\sum_{h=1}^{L} \Omega_h \log \bar{x}_h - \sum_{h=1}^{L} W_h \log \bar{X}_h \right) \tag{18}$$

where λ_1 and λ_2 are the Lagrange multiplier. To find the optimum value of Ω_h, we differentiate the Lagrange function given in equation (18) with respect to Ω_h and equate it to zero. The calibration weight is derived as:

$$\Omega_h = W_h + W_h Q_h \left(\lambda_1 + \lambda_2 \log \bar{x}_h \right) \tag{19}$$

Here, λ_1 and λ_2 are determined by substituting the value of Ω_h from equation (19) to equations (16) and (17), thus we get the calibrated weight as:

$$\Omega_h = W_h + W_h Q_h \left[\frac{-\left(\sum_{h=1}^{L} W_h \left(\log \bar{X}_h - \log \bar{x}_h \right) \right) \left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h \right)}{\left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h^2 \right) \left(\sum_{h=1}^{L} W_h Q_h \right) - \left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h \right)^2} \right]$$

$$+ W_h Q_h \log \bar{x}_h \left[\frac{\left(\sum_{h=1}^{L} W_h Q_h \right) \left(\sum_{h=1}^{L} W_h \left(\log \bar{X}_h - \log \bar{x}_h \right) \right)}{\left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h^2 \right) \left(\sum_{h=1}^{L} W_h Q_h \right) - \left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h \right)^2} \right] \tag{20}$$

On substituting the value of Ω_h from equation (20) in (14), the proposed calibrated estimator is given as:

$$\bar{y}_{r\_L} = \sum_{h=1}^{L} W_h \bar{y}_{Rh} + \hat{\beta}_{r\_L}\left[\sum_{h=1}^{L} W_h \left(\log \bar{X}_h - \log \bar{x}_h\right)\right] \tag{21}$$

where

$$\hat{\beta}_{r\_L} = \left[\frac{\left(\sum_{h=1}^{L} W_h Q_h\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{y}_h \log \bar{x}_h\right) - \left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h\right)\left(\sum_{h=1}^{L} W_h Q_h \bar{y}_h\right)}{\left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h^2\right)\left(\sum_{h=1}^{L} W_h Q_h\right) - \left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h\right)^2}\right]$$

The mean squared error of the estimator $\bar{y}_{r\_L}$ up to the second order of approximation is given as:

$$\mathrm{MSE}\left(\bar{y}_{r\_L}\right) = \mathrm{E}\left[\bar{y}_{r\_L} - \bar{Y}\right]^2$$

$$= \sum_{h=1}^{L} W_h^2 f_h' \left[\bar{Y}_h^2 C_{yh}^2 + (\bar{Y}_h + \beta_{r\_L})^2 C_{xh}^2 - 2\rho_h \bar{Y}_h (\bar{Y}_h + \beta_{r\_L}) C_{yh} C_{xh}\right]$$

where $\beta_{r\_L} = \left[\frac{\left(\sum_{h=1}^{L} W_h \bar{Y}_h \log \bar{X}_h\right) - \left(\sum_{h=1}^{L} W_h \log \bar{X}_h\right)\left(\sum_{h=1}^{L} W_h \bar{Y}_h\right)}{\left(\sum_{h=1}^{L} W_h \log \bar{X}_h^2\right) - \left(\sum_{h=1}^{L} W_h \log \bar{X}_h\right)^2}\right]$ and

$$f_h' = \left(\frac{1}{n_h} - \frac{1}{N_h}\right)$$

2.2. Stratified Double Sampling

We have also proposed the calibration estimator in the case of stratified double sampling when the stratum population mean of the auxiliary variable is not known. In this practice, a preliminary sample of size m_h units is drawn as a first phase sample using SRSWOR and a subsample of n_h units is drawn from the preliminary sample of size m_h units by SRSWOR. Let $\bar{x}_h^* = \frac{1}{m_h}\sum_{i=1}^{m_h} x_{hi}$ be the first phase sample mean, while $\bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{hi}$ and $\bar{y}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} y_{hi}$ are the second phase sample means of auxiliary and study variables, respectively. The usual ratio estimator in the h^{th} stratum under double sampling is $\bar{y}_{Rh\_d} = \frac{\bar{y}_h}{\bar{x}_h}\bar{x}_h^*$.

Therefore, the proposed calibration estimator in the case of stratified double sampling is given as:

$$\bar{y}_{r\_Ld} = \sum_{h=1}^{L} \Omega_h^* \bar{y}_{Rh\_d} \tag{22}$$

where the calibration weight Ω_h^* is so chosen to minimize the Chi-square type distance measure given as:

$$\sum_{h=1}^{L} \frac{(\Omega_h^* - W_h)^2}{W_h Q_h} \tag{23}$$

subject to the following calibration constraints:

$$\sum_{h=1}^{L} \Omega_h^* = \sum_{h=1}^{L} W_h \tag{24}$$

$$\sum_{h=1}^{L} \Omega_h^* \log \bar{x}_h = \sum_{h=1}^{L} W_h \log \bar{x}_h^* \tag{25}$$

The Lagrange function is given as:

$$L = \sum_{h=1}^{L} \frac{(\Omega_h^* - W_h)^2}{Q_h W_h} - 2\lambda_1 \left(\sum_{h=1}^{L} \Omega_h^* - \sum_{h=1}^{L} W_h \right) - 2\lambda_2 \left(\sum_{h=1}^{L} \Omega_h^* \log \bar{x}_h - \sum_{h=1}^{L} W_h \log \bar{x}_h^* \right) \tag{26}$$

where λ_1 and λ_2 are the Lagrange multipliers. The optimum calibration weight is given as:

$$\Omega_h^* = W_h + W_h Q_h \left[\frac{-\left(\sum_{h=1}^{L} W_h \left(\log \bar{x}_h^* - \log \bar{x}_h \right) \right) \left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h \right)}{\left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h^2 \right) \left(\sum_{h=1}^{L} W_h Q_h \right) - \left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h \right)^2} \right]$$

$$+ W_h Q_h \log \bar{x}_h \left[\frac{\left(\sum_{h=1}^{L} W_h Q_h \right) \left(\sum_{h=1}^{L} W_h \left(\log \bar{x}_h^* - \log \bar{x}_h \right) \right)}{\left(\sum_{h=1}^{L} W_h Q_h \log x_h^2 \right) \left(\sum_{h=1}^{L} W_h Q_h \right) - \left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h \right)^2} \right] \tag{27}$$

Thus, substituting the value of Ω_h^* from equation (27) in (22), the proposed calibrated estimator is obtained as:

$$\bar{y}_{r\_Ld} = \sum_{h=1}^{L} W_h \bar{y}_{Rh\_d} + \left[\hat{\beta}_{r\_Ld} \sum_{h=1}^{L} W_h \left(\log \bar{x}_h^* - \log \bar{x}_h \right) \right] \tag{28}$$

where

$$\hat{\beta}_{r\_Ld} = \left[\frac{\left(\sum_{h=1}^{L} W_h Q_h \right) \left(\sum_{h=1}^{L} W_h Q_h \bar{y}_{Rh\_d} \log \bar{x}_h \right) - \left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h \right) \left(\sum_{h=1}^{L} W_h Q_h \bar{y}_{Rh\_d} \right)}{\left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h^2 \right) \left(\sum_{h=1}^{L} W_h Q_h \right) - \left(\sum_{h=1}^{L} W_h Q_h \log \bar{x}_h \right)^2} \right]$$

For different values of Q_h, we can obtain different forms of calibration estimators.

The mean squared error of the estimator $\bar{y}_{r\_Ld}$ up to the second order of approximation is given as:

$$MSE\left(\bar{y}_{r\_Ld}\right) = E\left[\bar{y}_{r\_Ld} - \bar{Y}\right]^2$$

$$= \sum_{h=1}^{L} W_h^2 \left[f_h' \bar{Y}_h^2 C_{yh}^2 + f_h''(\bar{Y}_h + \beta_{r\_Ld})^2 C_{xh}^2 - 2 f_h'' \rho_h \bar{Y}_h (\bar{Y}_h + \beta_{r\_Ld}) C_{yh} C_{xh} \right]$$

where $\beta_{r\_Ld} = \left[\dfrac{\left(\sum_{h=1}^{L} W_h \log \bar{X}_h \bar{Y}_h\right) - \left(\sum_{h=1}^{L} W_h \log \bar{X}_h\right)\left(\sum_{h=1}^{L} W_h \bar{Y}_h\right)}{\left(\sum_{h=1}^{L} W_h \log \bar{X}_h^2\right) - \left(\sum_{h=1}^{L} W_h \log \bar{X}_h\right)^2} \right]$ and

$$f_h'' = \left(\frac{1}{n_h} - \frac{1}{m_h} \right)$$

3. Simulation Study

3.1. Real Dataset

The real population used here in order to study the performance of the proposed calibrated estimator is MU284 population given in Appendix B of Sarndal et al. (2003). It comprises 284 units, divided into eight strata of varying sizes. The y and x are the study variable and the auxiliary variable respectively. We have selected simple random samples without replacement (SRSWOR) from each stratum using proportional allocation. A simulated study has been conducted by generating 25,000 samples of different sizes in R-software. The description of the population is given as:

X: REV84, Real state value according to 1984 assessment (in millions)
Y: RMT85, Revenues from 1985 municipal taxation (in millions)

X		Y	
Min.	347	Min.	21.00
1st Quartile	1146	1st Quartile	67.75
Median	1854	Median	113.50
Mean	3088	Mean	245.19
3rd Quartile	3345	3rd Quartile	230.25
Max.	59877	Max.	6720.00

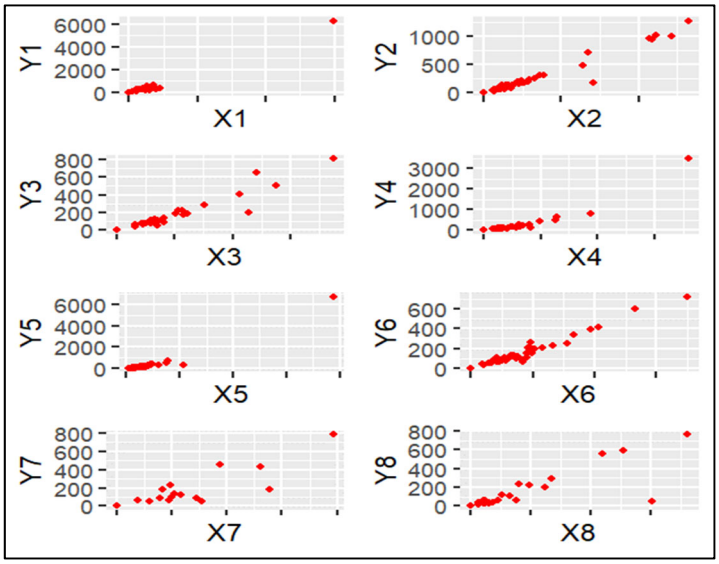


Figure 1: Plot between all strata of the population.

We have computed the empirical percentage absolute relative bias (%ARB) and percentage relative root mean squared error (%RRMSE) of the estimators as:

$$\%\mathrm{ARB}(\bar{y}_\alpha) = \left|\frac{(\bar{y}_\alpha - \bar{Y})}{\bar{Y}}\right| \times 100$$

$$\%\mathrm{RRMSE}(\bar{y}_\alpha) = \sqrt{\frac{1}{25000}\sum_{i=1}^{25000}\left(\frac{(\bar{y}_\alpha - \bar{Y})}{\bar{Y}}\right)^2} \times 100$$

where $\alpha = \mathrm{tr}, \mathrm{s}, \mathrm{r\_tr}, \mathrm{r\_s}, \mathrm{r\_L}, \mathrm{tr.d}, \mathrm{s.d}, \mathrm{r\_tr.d}, \mathrm{r\_s.d}, \mathrm{r\_Ld}$

The computed values of the %ARB and %RRMSE are given in the following tables for stratified sampling and stratified double sampling.

Table 1: Percentage Absolute Relative Bias (%ARB) for MU284 Population under Stratified Sampling

Q_h	**Sample Size**	**%ARB** $(\bar{y}_{tr})$	**%ARB** $(\bar{y}_{r\_tr})$	**%ARB** $(\bar{y}_s)$	**%ARB** $(\bar{y}_{r\_s})$	**%ARB** $(\bar{y}_{r\_L})$
1	**30**	3.004	49.304	5.668	5.969	2.803
	35	3.513	46.094	4.976	5.272	2.596
	40	4.096	42.126	4.505	5.015	2.348
	45	4.937	37.622	3.845	4.320	1.931
	50	5.186	39.019	3.540	3.878	1.821

Table 1: Percentage Absolute Relative Bias (%ARB) for MU284 Population under Stratified Sampling (cont.)

Q_h	**Sample Size**	**%ARB** $(\bar{y}_{tr})$	**%ARB** $(\bar{y}_{r\_tr})$	**%ARB** $(\bar{y}_{s})$	**%ARB** $(\bar{y}_{r\_s})$	**%ARB** $(\bar{y}_{r\_L})$
$\frac{1}{\bar{x}_h}$	**30**	3.374	59.629	4.886	6.367	3.835
	35	3.690	53.318	4.441	5.777	3.564
	40	4.231	47.327	4.077	5.568	3.385
	45	4.762	41.495	3.584	4.918	2.907
	50	4.988	41.971	3.388	4.532	2.795
$\frac{1}{\log \bar{x}_h}$	**30**	3.051	50.355	5.589	6.027	2.952
	35	3.540	46.824	4.922	5.340	2.734
	40	4.125	42.619	4.463	5.094	2.496
	45	4.932	37.982	3.818	4.401	2.067
	50	5.178	39.270	3.525	3.963	1.959

Table 2: Percentage Relative Root Mean Square Error (%RRMSE) for MU284 under Stratified Sampling

Q_h	**Sample Size**	**%RRMSE** $(\bar{y}_{tr})$	**%RRMSE** $(\bar{y}_{r\_tr})$	**%RRMSE** $(\bar{y}_{s})$	**%RRMSE** $(\bar{y}_{r\_s})$	**%RRMSE** $(\bar{y}_{r\_L})$
1	**30**	44.863	105.661	15.356	13.701	12.004
	35	39.069	91.444	14.491	13.389	11.740
	40	37.057	82.211	13.932	12.662	11.082
	45	33.878	74.189	13.424	12.482	10.926
	50	33.319	73.684	13.047	12.098	10.727
$\frac{1}{\bar{x}_h}$	**30**	39.633	106.641	15.124	12.951	11.220
	35	35.028	93.334	14.124	12.709	10.975
	40	33.178	82.878	13.520	11.958	10.318
	45	30.599	73.129	13.003	11.802	10.186
	50	30.383	73.502	12.613	11.513	10.045
$\frac{1}{\log \bar{x}_h}$	**30**	44.124	105.018	15.293	13.575	11.849
	35	38.496	91.088	14.427	13.273	11.596
	40	36.495	81.780	13.866	12.549	10.942
	45	33.411	73.616	13.363	12.371	10.794
	50	32.896	73.266	12.989	11.998	10.605

Table 3: Percentage Absolute Relative Bias (%ARB) for MU284 Population under Stratified Double Sampling

Q_h	**Sample Size**	**%ARB** ($\bar{y}_{tr.d}$)	**%ARB** ($\bar{y}_{r\_tr.d}$)	**%ARB** ($\bar{y}_{s.d}$)	**%ARB** ($\bar{y}_{r\_s.d}$)	**%ARB** ($\bar{y}_{r\_Ld}$)
1	**120;30**	2.418	43.856	4.679	4.847	1.814
	120;35	2.933	40.262	3.881	3.763	1.629
	120;40	2.727	33.652	3.305	3.824	1.405
$\frac{1}{\bar{x}_h}$	**120;30**	2.274	45.533	4.791	5.008	1.941
	120;35	2.749	41.887	4.165	3.917	1.753
	120;40	2.960	36.076	3.505	3.783	1.374
$\frac{1}{\log \bar{x}_h}$	**120;30**	2.834	47.344	4.015	5.123	2.783
	120;35	3.113	41.072	3.441	4.139	2.366
	120;40	2.841	33.352	2.938	4.159	2.197
1	**125;30**	2.707	49.029	4.103	5.260	2.891
	125;35	2.912	43.678	3.734	4.313	2.540
	125;40	3.052	36.093	3.167	4.177	2.288
$\frac{1}{\bar{x}_h}$	**125;30**	2.473	44.109	4.612	4.892	1.948
	125;35	2.962	40.234	3.836	3.815	1.733
	125;40	2.748	33.508	3.268	3.874	1.515
$\frac{1}{\log \bar{x}_h}$	**125;30**	2.329	45.778	4.721	5.050	2.074
	125;35	2.777	41.982	4.122	3.972	1.865
	125;40	2.980	35.963	3.471	3.842	1.499

Table 4: Percentage Relative Root Mean Square Error (%RRMSE) for MU284 under Stratified Double Sampling

Q_h	**Sample Size**	**%RRMSE** ($\bar{y}_{tr.d}$)	**%RRMSE** ($\bar{y}_{r\_tr.d}$)	**%RRMSE** ($\bar{y}_{s.d}$)	**%RRMSE** ($\bar{y}_{r\_s.d}$)	**%RRMSE** ($\bar{y}_{r\_Ld}$)
1	**120;30**	52.709	150.221	21.594	21.904	20.361
	120;35	49.091	133.640	20.869	21.841	19.877
	120;40	43.608	113.620	20.632	21.131	19.725
$\frac{1}{\bar{x}_h}$	**120;30**	47.533	134.851	22.416	20.587	18.853
	120;35	44.435	117.957	21.529	20.658	18.676
	120;40	39.980	97.357	21.084	19.941	18.455
$\frac{1}{\log \bar{x}_h}$	**120;30**	51.970	147.418	21.642	21.704	20.122
	120;35	48.456	131.082	20.906	21.660	19.692
	120;40	43.101	111.061	20.657	20.957	19.536
1	**125;30**	52.495	160.487	20.635	21.585	19.975
	125;35	48.481	135.201	20.205	21.817	19.786
	125;40	43.939	108.609	19.902	20.999	19.510
$\frac{1}{\bar{x}_h}$	**125;30**	47.294	143.024	20.414	20.275	18.495
	125;35	43.954	121.019	19.837	20.549	18.469
	125;40	40.048	93.322	19.502	19.756	18.139
$\frac{1}{\log \bar{x}_h}$	**125;30**	51.742	157.428	20.581	21.383	19.739
	125;35	47.835	132.770	20.147	21.625	19.583
	125;40	43.392	106.078	19.842	20.816	19.305

3.2. Artificial Dataset

A finite population of size N=3000 is generated for 3 strata using SRSWOR within each stratum, where the stratum sizes (N_h) are 1000 each, respectively. The values for the auxiliary variable X are generated considering the exponential distribution with varying values of the parameter for each stratum and the variable of interest Y is generated using the following models:

1st strata: $X_1 = \text{Exponential}(1000; 02)$ and $Y_1 = 100 + (\beta_1 * X_1) + \varepsilon_1$

where, $\beta_1 = 0.25$ and $\varepsilon_1 \sim N(0, 2)$

2nd strata: $X_2 = \text{Exponential}(1000; 03)$ and $Y_2 = 200 + (\beta_2 * X_2) + \varepsilon_2$

where, $\beta_2 = 0.50$ and $\varepsilon_2 \sim N(0, 3)$

3rd strata: $X_3 = \text{Exponential}(1000; 08)$ and $Y_3 = 300 + (\beta_3 * X_3) + \varepsilon_3$

where, $\beta_3 = 0.75$ and $\varepsilon_3 \sim N(0, 4)$

Table 5: Percentage Absolute Relative Bias (%ARB) for Exponential Population under Stratified Sampling

Q_h	**Sample Size**	**%ARB** ($\bar{y}_{tr}$)	**%ARB** ($\bar{y}_{r\_tr}$)	**%ARB** ($\bar{y}_{s}$)	**%ARB** ($\bar{y}_{r\_s}$)	**%ARB** ($\bar{y}_{r\_L}$)
1	**100**	24.082	21.017	1.396	1.690	1.310
	200	12.345	10.586	0.716	0.729	0.593
	300	7.928	6.810	0.415	0.474	0.374
	400	5.867	5.021	0.321	0.322	0.262
	500	4.870	4.294	0.226	0.220	0.180
$\frac{1}{\bar{x}_h}$	**100**	29.162	25.942	1.394	1.851	1.427
	200	14.305	12.355	0.698	0.813	0.645
	300	8.884	7.634	0.401	0.528	0.406
	400	6.437	5.480	0.310	0.361	0.285
	500	5.281	4.617	0.217	0.250	0.196
$\frac{1}{\log \bar{x}_h}$	**100**	20.674	17.744	1.390	1.555	1.247
	200	10.896	9.273	0.730	0.652	0.559
	300	7.196	6.163	0.426	0.424	0.354
	400	5.388	4.618	0.330	0.284	0.246
	500	4.543	4.029	0.234	0.190	0.169

Table 6: Percentage Relative Root Mean Square Error (%RRMSE) for Exponential Population under Stratified Sampling

Q_h	**Sample Size**	**%RRMSE** ($\bar{y}_{tr}$)	**%RRMSE** ($\bar{y}_{r\_tr}$)	**%RRMSE** ($\bar{y}_{s}$)	**%RRMSE** ($\bar{y}_{r\_s}$)	**%RRMSE** ($\bar{y}_{r\_L}$)
1	**100**	103.650	102.677	9.405	8.880	5.517
	200	56.659	57.615	6.328	5.830	3.661
	300	43.196	43.660	5.040	4.582	2.904
	400	35.456	35.661	4.264	3.886	2.462
	500	30.986	30.966	3.756	3.361	2.134
$\frac{1}{\bar{x}_h}$	**100**	116.032	116.691	9.411	8.930	5.854
	200	63.690	65.117	6.267	5.834	3.869
	300	48.085	48.754	4.977	4.578	3.063
	400	39.427	39.731	4.208	3.878	2.594
	500	34.486	34.523	3.701	3.353	2.248
$\frac{1}{\log \bar{x}_h}$	**100**	97.749	96.804	9.513	8.837	5.249
	200	52.652	53.345	6.449	5.831	3.465
	300	40.404	40.746	5.144	4.590	2.742
	400	33.062	33.191	4.348	3.897	2.324
	500	28.890	28.825	3.832	3.372	2.012

Table 7: Percentage Absolute Relative Bias (%ARB) for Exponential Population under Stratified Double Sampling

Q_h	**Sample Size**	**%ARB** ($\bar{y}_{tr.d}$)	**%ARB** ($\bar{y}_{r\_tr.d}$)	**%ARB** ($\bar{y}_{s.d}$)	**%ARB** ($\bar{y}_{r\_s.d}$)	**%ARB** ($\bar{y}_{r\_Ld}$)
1	**1000;100**	21.484	18.734	1.166	1.671	1.254
	1000; 200	9.858	8.318	0.562	0.717	0.552
	1000; 300	5.971	4.980	0.358	0.402	0.321
$\frac{1}{\bar{x}_h}$	**1000; 100**	25.808	22.927	1.157	1.819	1.351
	1000; 200	11.390	9.689	0.549	0.788	0.597
	1000; 300	6.692	5.579	0.348	0.443	0.348
$\frac{1}{\log \bar{x}_h}$	**1000; 100**	18.584	15.921	1.165	1.547	1.202
	1000; 200	8.721	7.289	0.572	0.652	0.523
	1000; 300	5.433	4.522	0.367	0.364	0.303
1	**1200; 100**	22.256	19.293	1.280	1.622	1.241
	1200; 200	10.769	9.255	0.641	0.672	0.547
	1200; 300	6.640	5.666	0.363	0.397	0.317
$\frac{1}{\bar{x}_h}$	**1200; 100**	26.710	23.547	1.276	1.774	1.347
	1200; 200	12.407	10.743	0.624	0.744	0.592
	1200; 300	7.420	6.329	0.352	0.442	0.344
$\frac{1}{\log \bar{x}_h}$	**1200; 100**	19.249	16.413	1.274	1.494	1.180
	1200; 200	9.574	8.162	0.656	0.607	0.518
	1200; 300	6.042	5.147	0.372	0.354	0.299

Table 8: Percentage Relative Root Mean Square Error (%RRMSE) for Exponential Population under Stratified Double Sampling

Q_h	**Sample Size**	**%RRMSE** ($\bar{y}_{tr.d}$)	**%RRMSE** ($\bar{y}_{r\_tr.d}$)	**%RRMSE** ($\bar{y}_{s.d}$)	**%RRMSE** ($\bar{y}_{r\_s.d}$)	**%RRMSE** ($\bar{y}_{r\_Ld}$)
1	**1000;100**	88.753	92.559	8.995	8.514	5.279
	1000; 200	51.137	52.014	5.825	5.404	3.396
	1000; 300	37.607	37.821	4.412	4.006	2.543
$\frac{1}{\bar{x}_h}$	**1000; 100**	101.307	106.522	8.999	8.567	5.583
	1000; 200	57.689	58.891	5.774	5.411	3.578
	1000; 300	42.086	42.396	4.359	4.000	2.676
$\frac{1}{\log \bar{x}_h}$	**1000; 100**	81.993	85.434	9.101	8.473	5.034
	1000; 200	47.245	47.980	5.932	5.405	3.219
	1000; 300	34.956	35.113	4.502	4.018	2.406
1	**1200; 100**	103.234	106.378	9.108	8.618	5.346
	1200; 200	51.810	52.845	5.941	5.474	3.452
	1200; 300	39.073	39.305	4.590	4.209	2.667
$\frac{1}{\bar{x}_h}$	**1200; 100**	115.116	119.277	9.108	8.670	5.654
	1200; 200	58.494	59.851	5.884	5.474	3.641
	1200; 300	43.586	43.981	4.534	4.204	2.807
$\frac{1}{\log \bar{x}_h}$	**1200; 100**	97.379	99.971	9.223	8.575	5.097
	1200; 200	47.916	48.815	6.055	5.480	3.270
	1200; 300	36.419	36.555	4.684	4.219	2.523

4. Conclusion

The ratio-type calibration estimators have been proposed under the stratified random sampling and stratified double sampling. The simulation study conducted on real as well as on artificial populations has shown that the proposed estimators have less %RRMSE and their value decrease as we increase the sample sizes, which in turn explains its performance better than the existing estimators. It is concluded that the proposed ratio-type calibration estimators using logarithmic stratum mean of the auxiliary variable in calibration constraint for the population mean in the case of stratified random sampling and stratified double sampling are found to be more efficient than the estimators given by Tracy et al. (2003), Nidhi et al. (2017) and Khare et al. (2022).

References

Alka, Ekpenyong, E. J., Rai, P. K., Tiwari, S., (2021). Calibration Estimator of Population Mean under Stratified Systematic Sampling Design. *International Journal of Agricultural and Statistical Science*, 17(Sup.1), pp. 2353–2361.

Clement, E. P., Enang, E. I., (2017). On the efficiency of ratio estimator over the regression estimator. *Communications in Statistics-Theory and Methods*, 46(11), pp. 5357–5367.

Deville, J. C., Sarndal, C. E., (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association,* 87(418), pp. 376–382.

Garg, N., Pachori, M., (2019). Calibration estimation of population mean in stratified sampling using coefficient of skewness. *International Journal of Agricultural and Statistical Sciences*, 15(1), pp. 211–219.

Garg, N., Pachori, M., (2020). Use of coefficient of variation in calibration estimation of population mean in stratified sampling. *Communications in Statistics-Theory and Methods*, 49(23), pp. 5842–5852.

Garg, N., Pachori, M., (2021). A Logarithmic Calibration estimator of Population Mean in Stratified Double Sampling. *International Journal of Agricultural and Statistical Science*, 17(Sup.1), pp. 2019–2025.

Khare, B. B., Singh, S., Mishra, M., (2022). Separate Ratio Estimator Using Calibration Approach for the Population mean in Stratified Random Sampling. *Asian Journal of Probability and Statistics*, 20(3), pp. 64–73.

Kim, J-M., Sungur, E. A., Heo T-Y., (2007). Calibration approach estimators in stratified sampling. *Statistics and Probability Letters,* 77, pp. 99–103.

Koyuncu, N., Kadilar, C., (2013). Calibration estimator using different distance measures in stratified random sampling. *International Journal of Modern Engineering Research (IJMER),* 3(1), pp. 415–419.

Koyuncu, N., Kadilar, C., (2016). Calibration weighting in Stratified Random Sampling. *Communications in Statistics-Simulation and Computation,* 45(7), pp. 2267–2275.

Kumar, R., Biswas, A., Singh, D., (2023). Estimation of Ratio in Finite Population using Calibration Approach under Different Calibrated Weights Systems. *Journal of Indian Society of Agricultural Statistics*, 77(2), pp. 225–232.

Mouhamed, A. M., Ei-Sheikh, A. A., Mohamed, H. A., (2015). A new calibration estimator of Stratified random sampling. *Applied Mathematical Sciences,* 9(35), pp. 1735–1744.

Nidhi, Sisodia, B. V. S., Singh, S., Singh, S. K., (2017). Calibration approach estimation of the mean in stratified sampling and stratified double sampling. *Communications in Statistics-Theory and Methods,* 46(10), pp. 4932–4942.

Pachori, P., Garg, N., Sangal, P. K., Rajesh, (2022). Exponential-type Calibration Estimators of Finite Population Mean in Stratified Sampling. *International Journal of Agricultural and Statistical Science*, 18(1), pp. 431–436.

Sarndal, C. E., Swensson, B., Wretman, J., (1992). *Model Assisted Survey Sampling*. Springer Verlag, New York.

Singh, S., (2003). Golden jubilee year 2003 of the linear regression estimator. Working paper at St. Cloud State University, St. Cloud, MN, USA.

Singh, S., Arnab, R., (2011). On Calibration of design weights. *METRON-International Journal of Statistics,* LXIX(2), pp. 185–205.

Tracy, D. S., Singh, S., Arnab, R., (2003). Note on Calibration in Stratified and Double Sampling. *Survey Methodology,* 29(1), pp. 99–104.

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 57–80, https://doi.org/10.59170/stattrans-2024-015
Received – 13.02.2023; accepted – 15.02.2024

Modeling the impact of an ISO 9001 certified quality management system on the organizational performance of Moroccan services firms

Ibtissam El Moury[1], Houda Kacimi[2], Sara Fennane[3], Adil Echchelh[4]

Abstract

Statistical modeling serves as a valuable tool for analyzing and resolving complex issues through the application of mathematical and statistical techniques. In the Moroccan context, where the services sector contributes more than 50% to the national wealth, there is a noticeable absence of such applications to gauge the impact of the ISO 9001 certified quality management systems on the organizational performance of companies.

To address this gap, we have constructed a causal model to quantify the strength of cause-and-effect relationships among the following key elements:

- processes within a quality management system and their impact on organizational performance,
- the influence of management responsibility process on all aspects of the quality management system, and
- the relationship between the organizational and financial performance.

This model is developed using Structural Equation Modeling (SEM) and estimated through the Partial Least Squares (PLS) approach, utilizing the XL-Stat software.

The research involved a sample of 55 Moroccan service firms certified with ISO 9001 across various sectors. Face-to-face surveys were the primary data collection method, engaging multiple executives within the organizations.

The findings of this research validate our conceptual model, revealing that:

- certain processes within the quality management system positively and weakly correlate with the organizational performance,

[1] Electronic Systems, Information Processing, Mechanics and Energetics Laboratory, Faculty of Science, Ibn Tofail University, Morocco. E-mail: ibtissamelmoury@gmail.com. ORCID: https://orcid.org/0009-0000-6175-0380.

[2] Electronic Systems, Information Processing, Mechanics and Energetics Laboratory, Faculty of Science, Ibn Tofail University, Morocco. E-mail: houda.kacimi@uit.ac.ma. ORCID: https://orcid.org/0000-0003-3771-996X.

[3] Electronic Systems, Information Processing, Mechanics and Energetics Laboratory, Faculty of Science, Ibn Tofail University, Morocco. E-mail: sara.fennane@uit.ac.ma. ORCID: https://orcid.org/0000-0003-0953-9056.

[4] Electronic Systems, Information Processing, Mechanics and Energetics Laboratory, Faculty of Science, Ibn Tofail University, Morocco. E-mail: adilechel@gmail.com. ORCID: https://orcid.org/0000-0002-5302-4255.

- the 'Management Responsibility' process has a positive influence on all aspects of the quality management system, and
- the organizational performance strongly and positively impacts the financial performance.

Moreover, our model facilitates the computation of direct and indirect effects among its components, transforming cause-and-effect relationships into structural equations. Our research contributes to enhancing the understanding of the dynamics between quality management systems and the organizational performance in the Moroccan service sector.

Key words: statistical modeling, PLS approach, organizational performance, Quality Management System (QMS), ISO 9001 certification.

1. Introduction

Increased globalization of markets, intensified competition and a turbulent economic environment make the current reality of the world in which companies must operate. Therefore, these firms must be agile enough to seize opportunities and manage risks well. In other words, only 'performing' enterprises have a chance of succeeding in a world where perpetual change is the norm. Performance is hens the mark of excellence that shareholders and managers are seeking. It is this importance that the business world gives to performance that makes the latter a concept worthy of observation and research. Practitioners and theorists have made its definition an object of reflection and we know today, thanks to their efforts, that performance is a complex notion, and difficult to define and measure. One thing is certain: performance, and whatever its nature, is merely a reflection of the company's ability to satisfy its customers and anticipate their needs. As a strong QMS (Quality Management System) can lead to business competitiveness and viability, it is easy to imagine the importance of an ISO 9001 certification while operating in highly competitive environments. [El Moury et al, 2023]

This standard is currently one of the most widely used ISO norms in the world. It is recognized as the most important reference for quality management and allows the establishment of relationships of trust with customers and prospects. However, the analysis of the impact of ISO 9001 certification is relatively little studied regarding the considerable financial and organizational investments made by organizations to meet the requirements of this standard.

The observations mentioned above are the main motivation for this article. The latter is particularly interested in studying: the impact of a certified QMS on the organizational performance of the Moroccan service firm, the effect of Management responsibility on this QMS, and the correlation between organizational and financial performance.

2. Conceptual Framework

The conceptual framework of our research is based on the study of the interrelationships existing between two structures:

- processes of an ISO 9001 certified QMS;
- firm performance: Organizational and financial performance.

2.1. Certification of Quality Management System according to ISO 9001 Standard

The creation of value for stakeholders in service provision can be managed in different ways. One possible strategy is to design and develop a quality management system (QMS) that will direct and control an organization in terms of quality. (Rönnbäck et al., 2009). QMS can be designed to include certain principles, additional practices, and techniques (Dean and Bowen, 1994). It often follows the substance of Deming's plan-do-check-act (PDCA) cycle (Deming, 1986), and provides support to organizations for the assurance and improvement of quality. Within an organization, or business network, there are internal processes that aim to manage and support the operative processes.

ISO defines it as a "management system to direct and control an organization with regard to quality". This system is defined as a formalized system that documents processes, procedures, and responsibilities for achieving quality policies and objectives. A QMS helps coordinate and direct an organization's activities to meet customer and regulatory requirements and improve its effectiveness and efficiency on a continuous basis.

Therefore, we can consider it as a means of controlling processes along with the quality of products (or services) which, in return, allows customer satisfaction and the achievement of the economic objectives pursued by the organization. [El Moury et al., 2020].

The purpose of this system is to establish an organization's policies and to realize the contents of these policies through short and long-term goals (Nilsson, 2000). The substance of a QMS often follows PDCA cycle. The cycle is a continuous quality improvement model consisting of a logical sequence of four repetitive steps for continuous improvement and learning (Deming, 1986). The main purpose is to start by planning and formulating concrete goals for the organization. The next step is to put the action plans or programs into action to reach the goals, check that the goals have been obtained, evaluate, and then further improve the organization's processes (ISO 9001: 2015).

A QMS is made up of correlated and interactive processes that use resources to achieve the desired results and provide value (product, service, etc.)

International Organization for Standardization 'ISO' is an independent, non-governmental international organization, bringing together experts to share knowledge to develop consensus-based, market-relevant, voluntary international standards that support innovation and provide solutions to global challenges.

ISO 9001 is defined as the international standard that specifies requirements for a quality management system (QMS). Organizations use the standard to demonstrate the ability to consistently provide products and services that meet customer and regulatory requirements. It is the most popular standard in the ISO 9000 series and the only standard in the series to which organizations can certify.

According to this Organization implementing a conformant Quality Management System is the source of numerous benefits:

- Allows an overall assessment of the business context and helps determine the expectations that need to be met. The overall result is the ability to define relevant goals and identify new business opportunities.
- Being customer-centric: the customer is placed at the core of any business strategy which guarantees a positive customer experience and hence long-term relationships.
- Work more efficiently because all processes will be aligned and understood by all the staff of the business or organization. The latter will then gain in productivity and efficiency, while reducing its internal costs.
- Being compliant with legal and regulatory requirements.
- Conquer new markets, because in certain sectors and for certain customers, compliance with the ISO 9001 standards is an essential criterion.

For [Chebir et al., 2022] among benefit of the ISO 9001 certification is that the company enjoys a competitive advantage that reinforces its position in front of the competition while gaining customers' confidence.

According to [Echour and Nbigui, 2021] the adoption of ISO 9001 is a voluntary approach, certification is a fashion effect for companies that want to be recognized for their quality and resist in an increasingly competitive market. For [Isuf et al., 2016] ISO 9001 standards do not refer to the compliance with a given goal or result. In other words, they are not performance standards measuring the quality of a firm's products or services or a firm's environmental results, rather, they are standards setting out the need to systematize and formalize many corporate processes within a set of procedures, and to document such implementation.

ISO in its ISO/IEC guide defines the standard as "A document, established by consensus and approved by a recognized body, that provides, for common and repeated use, rules, guidelines or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context". [ISO/IEC GUIDE 2, 2004].

ISO defines certificate of conformity as follows: "A document issued under the procedures of a third-Party certification System and a product, or a Service is in conformity with specific Standards or other technical specifications". [ISO/IEC GUIDE 23-1982]. Thus ISO 9001 certification attests that an organization has a quality management system that complies with the ISO 9001 standard. Note that the other standards of the series 9000: vocabulary (ISO 9000), guidelines (ISO 9004) do not contain requirements and cannot be used as a basis for certification.

2.2. Organizational and financial performance

Performance is at the heart of the company's social, economic, and financial balances. In other words, it constitutes the turning point of all decisions (strategic and operational) and must be based on an evaluation criterion to understand its evolution. According to [Renaud and Berland, 2010] performance has been used to assess the company's implementation of sustainable development strategies.

The Axcion lexicon (2019) defines performance as: "The result obtained in a specific area and considered as a successful outcome. It can be the result of human action or a material or process, etc. To be appreciated, it is subject to measurement." Performance measurement focuses mainly on creating value for shareholders. It is therefore not surprising that corporate management is focused on this value creation and the way to manage it. Recent studies show that 200 companies listed by Fortune magazine currently use an indicator based on the value created for shareholders to evaluate performance. [Patrick Jaulient, 2012].

According to [Serhan, 2019] firm performance entails three areas of the company including product market, shareholder return, and financial performance. The improvement of this performance includes business re-engineering activities, processes for continuously improving the business and the quality of services or products offered. To ensure that the organization is efficient it is necessary to analyse the main performance indicators. [Barna, and Roxana, 2021].

For [Rafoi, 2016] a company's performance indicators can be classified as follows: Strategic indicators: market share, turnover, customer satisfaction, return (profit), managerial indicators: availability of resources, costs, budget, operational indicators: individual performance, processes performance, products, efficiency.

For [Moulai Ali, 2012] Organizational performance "is about how the business is organized and managed to achieve its goals". Organizational performance determines the ability of the firm to implement effective processes to reach its operational and strategic projections. [El Moury et al., 2024]. The pillars of this efficiency can only be: The development and respect of a process-oriented approach, relations between the managers of the different compartments of the organization, the quality of the flow of information, and the degree of flexibility of the organization.

Because organizational effectiveness is difficult to express in a concrete formula, [Hadini, 2020] emphasized the need of indicators to keep objective achievements at the forefront of decision making.

Regarding financial performance, [Farrukh el al, 2016] consider it as the extent to which a company's financial health over a period is measured. In other words, financial performance is a composite of an organization's financial health, its ability and willingness to meet its long-term financial obligations, and its commitments to provide services in the foreseeable future. In a broader sense, financial performance refers to the degree to which financial objectives are accomplished. [Ganyam and Ivungu ,2019]. Cost-related performance is measured by quantitative indicators such as return on investment and sales, profitability, productivity, return on assets, efficiency, etc.

3. Research model and hypothesis

3.1. First research construct: Processes of an ISO 9001 certified QMS:

A quality management system (QMS) is the set of activities by which the organization defines, implements, and reviews its quality policy and objectives in accordance with its strategy. An organization's QMS is made up of interrelated and interactive processes that use resources to achieve intended results and deliver value (product, service, etc.). The QMS processes form our first research construct:

- management Responsibility Process: MRP;
- service Realization Process: SRP;
- measurement, Analysis, and Improvement Process: MAIP;
- resource Management Process: RMP.

3.2. Second research construct: Performance of the firm:

- organizational performance: OP;
- financial performance: FP

3.3. Presentation of the research model

The objective of this research work is to test and validate a conceptual model (causal) allowing to measure the impact:

- of the processes of quality management system an 'ISO 9001 certified' on organizational performance;
- organizational performance on financial performance;
- Management Responsibility Processes on these three processes: Service Realization Process 'PRS', Measurement, Analysis, and Improvement Process 'PMAA' and Resource Management Process 'PMR'.

This model is developed using Structural Equation Modeling (SEM) and estimated through the Partial Least Squares (PLS) approach, utilizing the XL-Stat software. The research involved a sample of 55 Moroccan service firms certified with ISO 9001 across various sectors. Face-to-face surveys were the primary data collection method, engaging multiple executives within the organizations.

Our model is based on 6 criteria which are divided into 2 families: 4 criteria refer to the means (QMS process), the other criteria refer to the results (financial and organizational performance), (Figure 1). We assume that there is a causal relationship between the criteria of means and the criteria of results. In other words, the means in place are the causes of the given results. It should be noted that for each causal relationship a hypothesis has been formulated. Since the proposed conceptual model has 8 causal relationships, 8 hypotheses have been formulated.

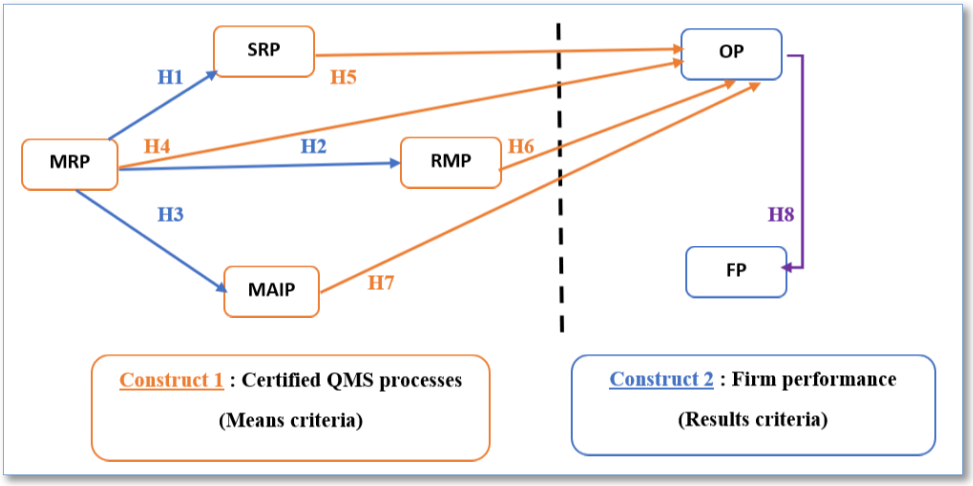


Figure 1: Proposed model

Table 1: Variables codes used in the research model

Model Construct proposed	Code	Title
Certified QMS processes (Means criteria)	MRP	Management Responsibility Process
	SRP	Service Realization Process
	RMP	Resource Management Process
	MAIP	Measurement, Analysis, and Improvement Process:
Firm performance (Results criteria)	OP	Organizational performance
	FP	Financial performance:

Thus, for each causal relationship, we have formulated a hypothesis (total: eight hypotheses).

3.4. Formulated hypotheses:

We intend through our study to validate or invalidate the eight Hypotheses below:

Table 2: List of hypotheses

Hypothesis Number	Causal Relationship	Hypothesis Formulated
H1	MRP→ SRP	We suppose that MRP has a positive impact on SRP
H2	MRP→RMP	We suppose that MRP has a positive impact on RMP
H3	MRP→ MAIP	We suppose that MRP has a positive impact on MAIP
H4	MRP→OP	We suppose that MRP has a positive impact on OP
H5	SRP→ OP	We suppose that SRP has a positive impact on OP
H6	RMP→ OP	We suppose that RMP has a positive impact on OP
H7	MAIP→OP	We suppose that MAIP has a positive impact on OP
H8	OP→FP	We suppose that OP has a positive impact on FP

4. Methodological framework of the research

Our empirical study is conducted via questionnaires, administered face-to-face, to the Directors, quality managers of this enterprises. We sent the questionnaire to 220 companies, and 55 agreed to respond. All these firms are part of the service industry and range from medium-sized to large companies.

4.1. Sampling

This research was conducted on a sample of 55 Moroccan service firms certified ISO 9001 operating in various sectors: passenger transport, goods transport, telecommunication, banking, insurance…

We asked several questions to test the strength and sense of the various causal links between the variables of our research constructs. For instance, questions concerning the contributions of an ISO 9001 certified QMS to organizational performance.

Example: Do you believe that your QMS has a positive impact on:

- The formalization of procedures?
- The standardization of processes?
- The improvement of process understanding?

4.2. Choice of modelling

To analyse our results, we used structural modelling according to the PLS approach. Two families of modelling structural equations were concerned [Gefen et al., 2000]: the methods based on the covariance, represented among other things by LISREL, and the methods based on variance whose PLS approach is the most representative of these techniques. For [Hulland,1999] and [Chenhall, 2005], the PLS approach is particularly suited for the analysis of small samples and when the analysis is exploratory in nature. Note that this approach is recommended when the theory is more approximate, and

the measurements are less well developed as it maximises the possibility of interpreting both the measurement model and the structural model. That is the case for this study. Hence, the choice of the PLS approach.

4.3. Evaluation of the proposed model

There are 6 variables in the research model 4 of which are reserved for QMS process and 2 are reserved for the performance. The variables of our model are operationalized by different items. These items are collected on fifth degree Likert scales: (Strongly disagree, somewhat disagree, moderately agree, somewhat agree, strongly agree). Note that the total number of items is 50 (items/questions).

4.3.1. Reliability of measurements

The reliability of measurements is concerned with the reduction of the random part of the measurement error: if the same phenomenon is measured several times by the same measuring instruments, the results should be as close as possible. To do this, we will use 'Cronbach alpha' and 'rho D.G.' indexes.

Table 3: Reliability of measurements

Latent Variable	*Items*	*Cronbach's alpha*	*D. G Rho*
MRP	9	0.939	0.912
SRP	6	0.829	0.932
RMP	6	0.849	0.958
MAIP	11	0.920	0.952
OP	9	0.923	0.947
FP	5	0.842	0.902

According to the results in Table 3, the Cronbach alpha and rho D.G. indexes calculated for each latent variable are above 0.7 and with reference to the recommendations of [Nunnally and Bernstein ,1994] and to the instructions of [Fornell and Larker,1981], these results are satisfactory (reliable). Similarly, the results in Table 3 also show satisfactory results where rhô D.G. > 0.8 for all measurement models according to the instructions of [Fornell and Larker, 1981].

4.3.2. Evaluation of the measurement model

Note that there are three ways of linking the manifest variables to the latent variables whose scheme can be formative, reflective, or MIMIC. The evaluation of the (external) measurement models then depend on the nature of the selected schema or pattern (formative, reflective or MIMIC) [Jacobowicz, 2007].

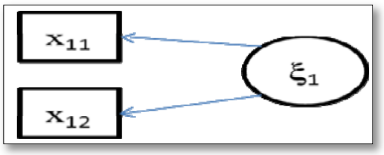


Figure 2: Reflective schema

The same author has confirmed that the reflective type of schema (Figure 2) is the most suitable in most structural equation models and that such a choice is mainly based on the subjectivity of the researcher. This is the scheme that has been chosen and followed in this study. Each manifest variable is connected to a latent variable with simple regression.

The relationship between the latent variable and all the manifest variables that are associated with it, can be written as follows: $x_{kj} = \pi_{kj} * \xi_k + \varepsilon_{kj}$

- x_{kj} vector associated with the *jth* manifest variable of the latent variable ξ_k
- ξ latent variable, k index of latent variables. kj index of manifest variables of block k.
- π loading associated with x_{kj} .ε_{kj} error term (measurement errors of manifest variables)

4.3.2.1. Convergent validity

The convergent validity means that the same items should share more variances with their latent construct than with their measurement errors. To test this validity, we calculate the index of the average variance extracted (AVE) associated with each latent variable:

$$\mathbf{AVE} = \frac{\sum[\gamma_i^2] var(VL)}{\sum[\gamma_i^2] var(VL) + \sum[var\ (\varepsilon_i)]}$$

Along with:

- VL latent variable,
- γ_i^2 Factorial contribution (loadings),
- ε_i variance of errors

Table 4: Quality index of measurement models

Latent variable	*Average variance extracted (AVE)*	*D. G Rho*
MRP	0.729	0.912
SRP	0.665	0.932
RMP	0.792	0.958
MAIP	0.710	0.952
OP	0.665	0.947
FP	0.713	0.902

According to guidelines of [Fornell and Larcker, 1981], we can talk of a good convergent validity if AVE > 0.5. The results in Table 4 clearly confirm the good convergent validity of our measurement model. The D.G. rho coefficient is equally significant for all the measurement models, which confirms again the good convergent validity (D.G. rho > 0.7) [Nunnally and Bernstein, 1994].

The manifest variables form the blocks in goshawks of the latent variables. Given that the measurement models are of the reflective type, the blocks must be one-dimensional to ensure that the manifest variables are the reflection of their latent variable.

Table 5: Eigenvalues latent variables causal model

MRP	SRP	MAIP	RMP	OP	FP
7.285	**4.656**	**7.806**	**4.754**	**5.987**	**2.852**
0.933	0.886	0.875	0.471	1.137	0.565
0.400	0.639	0.684	0.389	0.700	0.394
0.372	0.376	0.417	0.168	0.396	0.190
0.289	0.195	0.344	0.132	0.254	
0.251	0.167	0.236	0.085	0.182	
0.186	0.082	0.194		0.169	
0.154		0.169		0.103	
0.082		0.110		0.071	
0.049		0.088			
		0.076			

To do this, the first eigenvalue for each block must represent at least 50% of the sum of all values in the same block. This is the case for the results depicted in Table 5. This confirms the one-dimensionality of the blocks.

4.3.2.2. Discriminating validity (divergent)

The Items of a construct must be weakly correlated with the items of other constructs. That is to say, the divergent validity is retained only if the items belonging to a single construct do not contribute significantly with the others constructs. This will be tested in accordance with the recommendations of [Chin ,2010]. To do this, we compare the square root of the AVE of each latent variable with the correlations of the different latent variables (two by two). If the square root of the AVE is greater than the correlations between the different dimensions of our model, the divergent validity is ensured. The results of Table 6 show that the square root of the AVE of each latent variable exceeds the correlations between the latent variables (two by two), which confirms the discriminating validity of the external model.

Table 6: The discriminating validity

	MRP	RMP	MAIP	SRP	OP	FP	AVE
MRP	**0.853***						0.729
RMP	0.618	**0.889**					0.792
MAIP	0.614	0.737	**0.842**				0.710
SRP	0.711	0.631	0.678	**0.815**			0.665
OP	0.321	0.274	0.368	0.299	**0.815**		0.665
FP	0.454	0.452	0.460	0.434	0.521	**0.844**	0.713

Note: *the square root.

Convergent validity and divergent validity confirm that our measurement **Model is Valid**.

4.3.3. Validation of the structural model

The structural model defines the nature of the relationships existing between the latent variables. To test the validity of adjusting our internal model, we will resort the following:

4.3.3.1. Goodness of fit index (GoF)

This index considers both the performances of the structural model and the measurement model [Tenenhaus and Esposito Vinzi, 2005]. It is defined by the geometric mean of the average of the communities (or AVE) on all the latent variables (H^2) and the average of R^2 associated with the endogenous latent variables $\bar{R}^2$

$$\mathbf{GoF} = \sqrt{\bar{H}^2 \times \bar{R}^2}$$

According to [Wetzels et al., 2009], the usual values of this index are 0.1, 0.25 and 0.36. They correspond respectively to weak, medium, and large adequacy of the model. So according to the results obtained (Table 7), the research model can be retained in terms of the threshold (GoF > 0.5), and this, according to the instructions of [Wetzels et al.,2009].

Table 7: Adjustment Index

	GoF	GoF (Bootstrap)	Standard error	Critical Ratio (CR)
Absolute	0.670	0.677	0.061	11.071
Relative	0.837	0.813	0.055	15.325
External Model	0.991	0.978	0.041	24.083
Internal Model	0.845	0.830	0.029	28.780

4.3.3.2. The coefficient of determination (R^2)

On the one hand, the usual values of R^2 are 0.67, 0.33 and 0.19. These are respectively considered as substantial, moderate, and weak [Chin, 1998]. On the other hand,

the structural model is retained when $R^2 > 0.67$ [Chin, 1998]. From what preceded, the results of R^2 and R^2-adjusted (Table 8) are substantial to moderate.

The formula of this coefficient is as follows:

$$R^2 = 1 - \frac{SSR}{SST} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

- SSR relative to the sum of squares of the residuals (Residual variance)
- SST relative to the sum of total squares (Total Variance Explained)
- y_i Measurement values
- $\bar{y}$ The average of the measurements
- $\hat{y}_i$ The predicted values.

Table 8: Results R^2 and R^2 – adjusted

	MRP	RMP	MAIP	SRP	OP	FP
R^2	-	0.618	0.768	0.757	0.392	0.655
R^2adjusted	-	0.618	0.764	0.752	0.357	0.627

4.3.3.3. Effect size (f^2)

This index allows us to ensure the validity and magnitude of structural coefficients.

Table 9: Research hypothesis tests

Causal relationship	Path coefficient (β)	Effect size (f^2)	T* student	Hypothesis validation
H1: MRP→ SRP	0.571	0.514	5.169	Valid
H2: MRP→RMP	0.785	1.615	9.252	Valid
H3: MRP→ MAIP	0.284	0.134	2.635	Valid
H4: MRP→OP	0.163	0.028	1.977	Valid
H5: SRP→ OP	0.156	0.000	0.057	Invalid
H6: RMP→ OP	0.150	0.004	-0.459	Invalid
H7: MAIP→OP	0.175	0.079	1.987	Valid
H8: OP→FP	0.457	0.378	4.305	Valid

Its usual values are 0.02, 0.15, and 0.35 corresponding respectively to a weak effect, medium and strong [Cohen, 1988]. According to Table 9, we can conclude that:

- MRP has a strong effect on SRP and RMP;
- MRP has a weak effect on MAIP and OP;
- SRP has a weak effect on OP;
- RMP has a weak effect on OP;
- MAIP has a weak effect on OP;
- OP has a strong effect on FP.

The results of our survey clearly show the validity of the measurement model (external) and the validity of the structural model (internal).

4.4.4. Structural equations of the conceptual model

Our model has a single exogenous variable that is the 'management responsibility process' and owns five endogenous variables. Each endogenous variable is explained by one or more variables and an error term. The internal model is defined by linear equations linking the latent variables between them.

For all ξ_k endogenous, we have a $\xi_k = \sum_{i:\, \xi_i \to \xi_k} \beta_{kj} + \xi_k$ where β_{kj} represents the coefficient associated to the relationship between the variables ξ_k and ξ_i. ξ_k is an error term. $\xi_i \to \xi_k$: ξ_i explains ξ_k in the model. This model has five equations that were tested using the PLS approach through the Xlstat software. The structural equations of the conceptual model are presented as follows:

1) RMP = 0,78590*MRP ;
2) MAIP = 0,28436*MRP ;
3) SRP = 0,57199*MRP ;
4) OP = 0,16327*MRP+0,15063*MRP+0,17575*MAIP+0,15649*SRP;
5) FP = 0,45734*OP.

5. Results analysis

5.1. Hypotheses testing

For each causal relationship, we have formulated a hypothesis and since we have eight causal relationships, we have setup eight hypotheses. These hypotheses were also subjected to tests of confirmation (Table 9).

Table 10: Results of hypothesis testing

Hypothesis	Results
H1: We suppose that MRP positively influences SRP	Valid
H2: We suppose that MRP positively influences RMP	Valid
H3: We suppose that MRP positively influences MAIP	Valid
H4: We suppose that MRP positively influences OP	Valid
H5: We suppose that SRP positively influences OP	Invalid
H6: We suppose that RMP positively influences OP	Invalid
H7: We suppose that MAIP positively influences OP	Valid
H8: We suppose that OP positively influences FP	Valid

According to Table 9, we can confirm the validity of all hypotheses (|T-value| > 1.96) except for Hypotheses 5 and 6: |T-value| = 0.057< 1.96. |T-value| = -0.459< 1.96

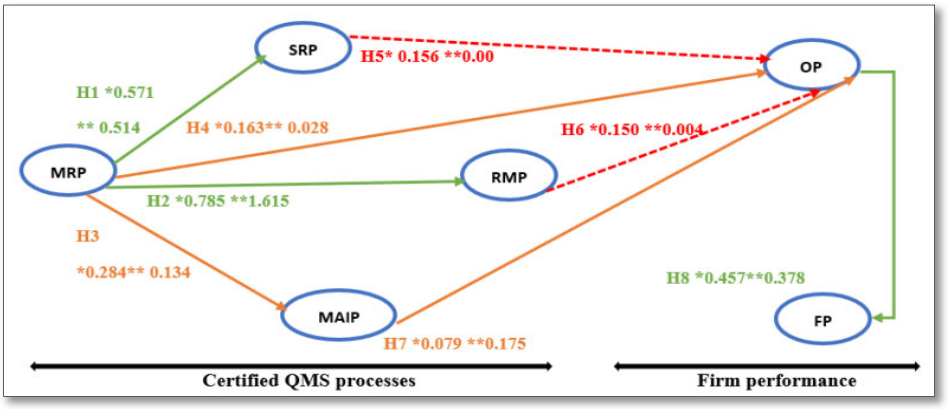


Figure 3: Final model estimated by the PLS

Notes:

- *Structural coefficient (path coefficient)
- **Effect size [0.02: weak, 0.15: medium, 0.35: large (strong)] according to Cohen (1988).
- Large importance of the effect.
- Weak importance of the effect.
- Invalid link.

The final model can be represented as Figure 3.

5.2. Direct effects between the axes of the means criteria and the axes of the results criteria:

Based on the results in Table 9 and Figure 3, we underline the following:

- The exogenous latent variable 'Management Responsibility Process' has positive and statistically significant influences ($T > 1.96$). It has a strong importance of effects on the latent variables: 'Service Realization Process', and 'Resource Management Process'. ($\beta = 0.571$, $f^2 = 0.514$, $t = 5.169$), ($\beta = 0.785$, $f^2 = 1.615$, $t = 9.252$).
- The exogenous latent variable 'Management Responsibility Process' influences weakly 'Measurement, Analysis, and Improvement Process' in a positive way and statistically significant. ($\beta = 0.284$, $f^2 = 0.134$, $t = 2.635$, $p < 0.01$).
- The exogenous latent variable 'Management Responsibility Process' influences weakly 'Organizational performance' in a positive way and statistically significant. ($\beta = 0.163$, $f^2 = 0.028$, $t = 1.977$).
- The latent variable 'Measurement, Analysis, and Improvement Process' influences weakly 'Organizational performance' in a positive way and statistically significant. ($\beta = 0.175$, $f^2 = 0.079$, $t = 1.987$).

- The latent variable 'Service Realization Process' has statistically insignificant influences on 'Organizational performance', proving the invalidity of that relationship. ($\beta = 0.156$, $f^2 = 0.000$, $t = 0.057$).
- The latent variable 'Resource Management Process' has statistically insignificant influences on 'Organizational performance', proving the invalidity of that relationship. ($\beta = 0.150$, $f^2 = 0.004$, $t = -0.459$)
- Finally, the latent variable 'Organizational performance' has positive and statistically significant influences ($T > 1.96$). It has a strong importance of effects on the latent variables 'Financial performance'. ($\beta = 0.457$, $f^2 = 0.378$, $t = 4.305$).

5.3. Indirect effects between the axes of the means criteria and the axes of the results criteria

We can note as well with reference to Table 11:

- The latent variable 'Management Responsibility Process'' has positive indirect effects on 'Measurement, Analysis, and Improvement Process' and 'Service Realization Process'. This exogenous latent variable has 'quite large' indirect effects on the Organizational performance' and 'Financial performance'.
- The latent variable 'Resource Management Process' has a low indirect effect on the Organizational performance' and 'Financial performance'.
- The latent variables 'Measurement, Analysis, and Improvement Process' and 'Service Realization Process' have a low indirect effect on the 'Financial performance'.

Table 11: Indirect effects between latent variables

	MRP	RMP	MAIP	SRP	OP	FP
MRP						
RMP	0.000					
MAIP	0.374	0.000				
SRP	0.346	0.000	0.000			
OP	0.388	0.154	0.000	0.000		
FP	0.532	0.238	0.080	0.072	0.000	

6. Discussion

When improvement is the purpose that governs the thoughts of an entrepreneur or a manager, performance becomes the concept that is constantly referred to. Both theorists and practitioners agree that this concept is multidimensional and often confusing. However, its understanding can only be comprehended based on the context in which it is used: organizational, commercial, or financial performance for instance.

It is agreed that performance, and whatever its nature, is merely a reflection of the company's ability to satisfy its customers and anticipate their needs. As a strong QMS (Quality Management System) can lead to business competitiveness and viability, it is easy to imagine the importance of an ISO 9001 certification while operating in highly competitive environments. Thus, it is very important to analyse and measure the impact of an ISO 9001 certified QMS on organizational performance.

In Canada, researchers have studied the impact of ISO 9001 on organizational performance, [Thaver and Wilcock, 2006]. They concluded that an ISO 9001 certified QMS has a positive impact on organizational performance, and they confirmed that among the benefits of the implementation of this system: reduction of defects and losses, improvement of customer satisfaction and better perception of product quality.

In the Australian context, [Singh, 2008] has shown that the ISO 9001 standard has a positive effect on process control. Again, in the same context, the results of several researchers have proven the positive impact of the ISO 9001 standard on the operational efficiency of the company's processes [Chuong et al., 2010]. For their part, [Terziovski and Guerrero, 2014] demonstrated through a study conducted in 2014 on 220 companies in Australia that ISO 9001 certification statically does not have a significant impact on product innovation, but a significant positive impact on process innovation.

According to [Valmohammadi and Kalantari, 2015] the ISO 9001 certified companies have a more relevant organizational performance than companies not certified. It is a tool to help integrate total quality, develop the quality culture, and improve the internal organization as well as the production process that generates, for example, indirectly the improvement of the quality of the final product. [Katerina et al., 2002].

With regards to the kingdom of Morocco, and from an economic environment perspective, [Ben Ali et al ,2020] measured the impact of an ISO 9001 certified quality approach (case: young manufacturing companies in growth phase, located in the north of Morocco) on the organizational performance, the results showed that there is a positive relationship between these two elements. In addition to this upshot, [Barrijal et al., 2021] added another layer to the building blocks of this study: organizational performance (i.e. process, organizational learning and innovation) may well mediate between quality approach and corporate social responsibility.

However, [Hadini, 2020] proved that the practices of an ISO 9001 certified quality approach (case: multinational firm located in Morocco) have a weak impact on the axes of 'non-financial performance'.

In a business environment where an economic downturn and financial crisis dominates (Greece to be precise). [Evangelos and Dimitrios,2014] concluded that the ISO 9001 certified companies significantly outperform the non-certified regarding product quality, customer satisfaction, operational performance.

[Matheus et al., 2021] have proven that ISO 9001 certified QMS has a positive impact on the organizational performance of several Brazilian companies.

According to our exploratory study, the results show that:

- Although the criteria 'Management responsibility' and 'Measurement, analysis and improvement' displayed a statistically positive influence, our conceptual model has nevertheless suggested a weak effect on organizational performance. To shed light on the meaning of this finding, we take the example of a cross-functional procedure that is of paramount importance to bring an ISO 9001 certification to fruition. This procedure is yet to be crafted and to meet a well-defined deadline; the CEO appointed a task force and:
  - Asked the appointed team (board of directors) to craft a fitting procedure,
  - Appointed his executive secretary to spearhead the coordination and secure a steady follow-up along with a monthly status report.

  First highlight: without the will and strategic thinking of the CEO, this procedure would not have existed. The same could be said about all subsequent control and monitoring tools that his executive was tasked to devise→ This explains why our conceptual model has considered that the influence of the leadership team was 'Positive and statistically significant'.

  Second highlight: If the CEO and his executive team have succeeded in provoking a chain reaction, this involvement can be the guarantor of positive results. In other words, producing success is tightly linked to other levers. For instance, the level of motivation of the company's staff. The latter was indeed quite absent among the sample. --> We can hence understand why the conceptual model considers that the leadership team's involvement produced 'weak impact' on the organizational performance.
- The relationship between the 'Service realization' criterion and organizational performance is invalid.

  Having just procedures does not guarantee an efficient realization of a service. Indeed, other means of support must intervene (e.g.: periodic audits, permanent controls, continuous training, etc.). In short, these are the elements that force the involvement of the human element in the execution of the procedures) → We therefore share the invalidity finding suggested by the conceptual model.
- The relationship between the criterion 'Resource management' and organizational performance is invalid.

  The example of the Executive Secretary appointed to spearhead the coordination and control of procedures requiring technical expertise can only be seen as a real-life example of unfitting profiles instrumental in the execution process. Several interviewers have indeed pointed this surprising fact during our sense-making discussions. It is in this sense that the model considers that the criteria

'Resource Management' and 'Organizational Performance' are not correlated (invalid link).

- Organizational performance has a positive and strong influence on financial performance.

According to [Ataseven et al., 2015], managers must invest their efforts in the proper integration of the requirements of the international standard ISO 9001 and not in obtaining certification. And [Jang and Lin, 2008] confirm that thorough QMS implementation significantly determines firm performance.

The efficiency of the management responsibility process within a firm result from:

- A clear dissemination of the principles of its mission, vision, values, and ethics.
- Disseminating to its team the knowledge to be put into practice to improve the implementation and the quality of the services rendered to clients and other interested parties. [El Moury et al., 2023]

The criterion of Management Responsibility aims at assessing the action of the leaders and the excellence of their behaviour, in the accomplishment of their mission, as well as their vision of the organization through the implementation of values and systems necessary for sustainable success. In accordance with the research of [Calvo-Mora et al., 2005], this criterion has effects on the criteria of resource and process management.

Thus, the results of our study prove that this criterion positively influences in a strong way 'Service realization' and 'Resource management'. However, that impact is feeble when it comes to 'Measurement, Analysis, and Improvement'.

In consideration of the analysis of the observations made by the authors mentioned above and the results of our study, we realize that the commitment of the leaders to manage the processes of the ISO 9001 certified QMS efficiently must be visible, permanent, "proactive" and exist at all levels of management. In addition, managers must invest their efforts in the appropriate integration of the requirements of the international standard ISO 9001 and not just in obtaining certification.

7. Conclusion

Statistical modelling, employing mathematical and statistical techniques, provides a robust framework for analysing and solving complex problems. This article delves into an extensive exploration of structural equation modelling, aiming to assess the strength of cause-and-effect relationships within our research model.

The findings indicate that:

- Positive relationships exist between various components, such as:
  - Management Responsibility Process and Service Realization Process, Management Responsibility Process and Resource Management Process,

Management Responsibility Process and Measurement, Analysis, and Improvement Process,
  - Management Responsibility Process and Organizational Performance, Measurement, Analysis, and Improvement Process and Organizational Performance,
  - Organizational Performance and Financial Performance.
- Some links are found to be invalid, including those between the Service Realization Process and Organizational Performance, as well as the Resource Management Process and Organizational Performance.
- Among the eight cause-and-effect relationships, three exhibit a substantial effect, while five show a weaker effect.

The model facilitates:

- The calculation of direct and indirect effects among the various components, offering a comprehensive understanding of their interplay.
- The transformation of cause-and-effect relationships into structural equations, providing predictive and anticipatory capabilities that support informed decision-making.

This study's results contribute to advancing researchers' knowledge and promoting scientific evolution. Additionally, practitioners gain heightened awareness of the impact of an ISO 9001 certified quality management system on organizational performance, empowering them to enhance managerial practices through the adoption of such an approach.

It is crucial to acknowledge the inherent limitations of research endeavors like ours. For instance, the data collection method using a questionnaire has inherent constraints, primarily allowing for the gathering of subjective data and information based on the perceptions of managers.

References

Ataseven, C., Nair, A. and Prajogo, I., (2015). ISO 9000 Internalization and Organizational Commitment – Implications for Process Improvement and Operational Performance. *IEEE Transactions on Engineering Management*, Vol., 6, No. 1. pp 5–17.

Åsa Rönnbäck, Lars Witell and Bo Enquist, (2009). Quality management systems and value creation, *International Journal of Quality and Service Sciences*, Vol. 1, No: 3 pp 241-254. Doi:10.1108/17566690911004186

Ben Ali Mohamed, Said Rifai, Said Barrijal and Otmane Bouksour, (2020). Proposal of a causal model to measure the impact of quality on industrial performance: a case study from Moroccan context. *Int. J. Productivity and Quality Management*, Vol. 30, No. 2, pp. 143–167.

Barrijal Said, Mohamed Ben Ali, Said Rifai and Otmane Bouksour (2021). 'Quality: from industrial performance to societal responsibility. Case of young industrial enterprises located in the region of Tetouan, Morocco'. International Journal of Industrial and Systems Engineering. Vol. 39, No. 1, pp. 123–150. https://doi.org/10.1504/IJISE.2021.117684

Barna, L., Roxana, D., (2021). The influence of the implementation of ERP systems on the performance of an organization, *Proceedings of the 15th International Conference on Business Excellence*, Vol 15, No 1, pp. 268–279.

Chin, W. W., (1998). The partial least squares approach to structural equation modeling, in Marcoulides, G.A. (Ed.): *Modern Methods for Business Research*, pp. 295–336, Lawrence Erlbaum Associates, Publisher, Mahwah, NJ.

Chenhall, R. H., (2005) 'Integrative strategic performance measurement systems, strategic alignment of manufacturing, learning and strategic outcomes: an exploratory study', Accounting, Organizations and Society, Vol 30, No 5, pp. 395–422.

Chebir Adil, Omar Taouab, El Moury Ibtissam and Adil Echchelh, (2022). Issues of harmonization of ISO 9001 standard and PCI Secure SLC –Electronic Banking and Certification in Morocco: potentials and risks. *Colloquium-journal*, 29(152), pp. 39–47.

Chee Chuong, S., Power, D. and Singh, P. J. (2010). A resource dependence theory perspective of ISO 9000 in managing organizational environment. *Journal of Operations Management*, Vol 29, pp. 49–64. Doi: 10.1016/j.jom.2010.04.002

Chin, W. W., (2010) How to write up and report PLS analyses, in Esposito Vinzi, V., Chin, W.W., Henseler, J. and Wang, H. (Eds.): Handbook of Partial Least Squares, Concepts. *Methods and Applications*, pp. 655–690, Springer, Heidelberg. https://doi.org/10.1007/978-3-540-32827-8_29

Calvo-Mora, A., Leal, A. and Roldan, J., (2005). Relationships between the EFQM Model Criteria: A Study in Spanish Universities. *Total Quality Management*, Vol 16, No 6, pp. 741-770.

Dean, J., Bowen, D., (1994). Management theory and total quality: improving research andpractice through theory development. *Academy of Management Review*, Vol. 19, No. 3, pp. 392–418.

Deming, E. W., (1986). Out of the Crisis, MIT/Center for Advanced Engineering Studies,Cambridge, MA.

El Moury, I., Hadini, M., Chebir, A. and Echchelh, A., (2020). Impact of ISO 9001 Certification on Organizational Performance: *State of the Art'. International Journal of*

Innovation and Applied Studies, Vol 31, No 3, pp. 648–654. http://www.ijias.issr-journals.org/abstract.php?article=IJIAS-20-348-15

El Moury Ibtissam, Mohamed Hadini, Adil Chebir, Mohamed Ben Ali and Adil Echchelh, (2023). Proposal of a causal model measuring the impact of an ISO 9001 certified Quality Management System on financial performance of Moroccan service-based companies. *Statistics in Transition new series*, Vol 24, No 2, pp. 159–184. DOI 10.59170/stattrans-2023-026

El Moury Ibtissam, Adil Chebir, Ben Ali Mohamed and Echchelh Adil, (2023). Modelling the Impact of an ISO9001 Certified Quality Management System on the Commercial Performance of Moroccan Service Firms. *Advances and Applications in Statistic*, Vol. 88, No.1, pp. 49–73. http://dx.doi.org/10.17654/0972361723039

El Moury Ibtissam, Kacimi Houda, Fennane Sara and Echchelh Adil, (2024). Contribution to statistical modeling: the relationship between quality management system processes and financial performance, Advances and Applications in Statistics, Vol. 91, No. 6, pp. 761–780. https://doi.org/10.17654/0972361724040.

Echour Saida, Taibi Nbigui, (2021). Motivations related to the quality management system and benefits of its implementation in the company: state of the art, IEEE 13th International Colloquium of Logistics and Supply Chain Management (LOGISTIQUA), 2-4 Dec. 2020, Fez, Morocco, HST (EST) – Sidi Mohamed Ben Abdellah University. [Viewed 01 April 2021]. Available from: https://ieeexplore.ieee.org/abstract/document/9353877/metrics#metrics

Evangelos, P., Dimitrios, K., (2014). Performance measures of ISO 9001 certified and non-certified manufacturing companies. Benchmarking: An International Journal. Vol. 21, No. 5, pp. 756–774. Doi.org/10.1108/BIJ-04-2012-0028

Farrukh, I. Farah, N. and Faizan, N., (2016). Financial Performance of Firms: Evidence from Pakistan Cement Industry. *Journal of Teaching and Education*, Vol. 5, No. 1, pp. 81–94.

Fornell, C., Larcker, D. F., (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, Vol 18, No. 1, pp. 39–50.

Ganyam, A. L., Ivungu, J. A., (2019). Effect of Accounting Information System on Financial Performance of Firms: A Review of Literature. *Journal of Business and Management*, Vol. 21, No.5, pp 39–49.

Gefen, D., Straub, W. D. and Boudreau, M-C., (2000). Structural equation modelling and regression: guidelines for research practice. *Communications of the Association for Information Systems*, Vol. 4, No.7, pp. 1–79.

Hadini, M., (2020, Feb). The Concept of Industrial Performance: *State of the Art. International Journal of Innovation and Applied Studies*, Vol. 28, No. 3, pp. 726–739. http://www.ijias.issrjournals.org/abstract.php?article=IJIAS-19-323-05

Hulland, J., (1999). Use of partial least squares (PLS) in strategic management research: a review of four recent studies. *Strategic Management Journal*, Vol 20, No 2, pp. 195–204.

Hadini Mohammed, (2020). Doctoral thesis in Industrial Engineering: Change Management through the Quality, Health-Safety & Environment 'QSSE' Approach as a lever for industrial performance management. Case study of a multinational company based in Morocco. Hassan 2 University, Morocco.

Isuf, L., Mane, A., Ilir, K., Remzi, K., (2016). A Literature Review On Iso 9001 Standards. *European Journal of Business, Economics and Accountancy*, 4(2), pp. 81–85.

ISO, (2022). About ISO, March 2022, [online] Available: https://www.iso.org/fr/about-us.html.

Jang, W. Y., Lin, Ch. I. (2008). An integrated framework for ISO 9000 motivations, depth of ISO implementation and firm performance: The case of Taiwan. *Journal of Manufacturing Technology Management*, Vol. 19, No. 2, pp. 194–216.

Jacobowicz, E. (2007) Contributions to Structural Equation Models with Latent Variables, Doctoral Thesis, National Conservatory of Arts and Crafts of Paris.

Katerina, D. Gotzamani, George-D, Tsiotras, (2002). The true motives behind ISO 9000 certification: Their effect on the overall certification benefits and long-term contribution towards TQM. *International Journal of Quality & Reliability Management*, Vol. 19, No. 2, pp 151-169. Doi:10.1108/02656710210413499.

The Axcion lexicon. Edition (2019).

Matheus, B. C, Fabiane, L. L, de Toledo, J. C., (2021). The Impact of ISO 9001 Certification on Brazilian Firms' Performance: Insights from Multiple Case Studies. *International Journal of Economics and Management Engineering*, Vol. 15, No. 8, pp. 677–683. https://publications.waset.org/10012139/the-impact-of-iso-9001-certification-on-brazilianfirms-performance-insights-from-multiple-case-studies

Moulai Ali, M. (2012). PhD thesis in economics: The performance of the national cement industry in the light of contractual theories of organizations, University of Oran, Algeria.

Nunnally, J. C., and Bernstein, I. H., (1994). Psychometric Theory, 3rd ed., McGraw-Hill, New York. O'Donnell, A. (2013), The contribution of networking to small firm marketing. *Journal of Small Business Management*, Vol. 52, No. 1, pp. 164–187.

Nilsson, U., (2000). Att integrera ledning system, SIS Forum, Stockholm, 56 s. Normann, R. and Ramirez, R., (1993), Strategy and the art of reinventing value. *Harvard Business Review*, Vol. 71, No. 5, pp. 50–1.

Jaulient, P., (2012). How do you measure the performance of a company?, Les Echos, viewed 24 May 2021: http://archives.lesechos.fr/archives/cercle/ 2012/12/26/cercle_61804.htm.

Rafoi, A., (2016). Top 5 KPIs for Distribution, Bit software, viewed 30, March 2022: https://info.bitsoftware.eu/blog/bitsoftware-ro/5-indicatori-de-performan ta-importanti-pentru-industria-de-distributie.

Serhan, A. El Hajj, W., (2019). Impact of ERPS on Organizations' Financial Performance. *Proceedings of the 13th International Conference on Business Excellence*, Vol. 13, No. 1, pp. 361–372.

Singh, P. J., (2008). Empirical assessment of ISO 9000 related management practices and performance relationships. *International Journal of Production Economics. Elsevier*, Vol. 113, No.1, pp. 40–59. https://ideas.repec.org/a/eee/proeco/v113y2008i1p40-59.html.

Tenenhaus, M., Esposito Vinzi, V., (2005). PLS regression, PLS path modeling and generalized procrustean analysis: a combined approach for PLS regression, PLS path modeling and generalized multiblock analysis. *Journal of Chemometrics*, 19(3), pp. 145–153.

Thaver, I., Wilcock, A., (2006). Identification of overseas vendor selection criteria used by Canadian apparel buyers: Is ISO 9000 relevant? *Journal of Fashion Marketing and Management*, Vol. 10, No. 1, pp 56–70. https://doi.org/10.1108/13612020610651123.

Terziovski, M., Guerrero, J-L. (2014). ISO 9000 quality system certification and its impact on product and process innovation performance. *International Journal of Production Economics*, Vol 158, pp. 197–207. Doi: 10.1016/j.ijpe.2014.08.011.

Valmohammadin, C., Kalantari, M., (2015). The moderating effect of motivations on the relationship between obtaining ISO 9001 certification and organizational performance. *The TQM Journal*, Vol. 27, No. 5, pp. 503–518. http://dx.doi.org/10.1108/TQM-05-2014-0042.

Wetzels, M., Odekerken-Schroder, G. and Vanopen, C., (2009). Using PLS path modeling for assessing hierarchical construct models: guidelines and empirical illustration. *MIS Quarterly*, Vol. 33, No. 1, pp. 177–195.

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 81–102, https://doi.org/10.59170/stattrans-2024-016
Received – 15.10.2023; accepted – 12.02.2024

Dimensionality reduction analysis of the renewable energy sector in Azerbaijan: nonparametric analyses of large datasets

Ibrahim Niftiyev[1]

Abstract

Although the number of econometric analyses related to the renewable energy sector in Azerbaijan is increasing, studies on nonparametric dimensionality reduction are rather sparse. Principal component analysis (PCA) and multiple correspondence analysis (MCA) were chosen to fill this apparent research gap. As a result, a large dataset including the renewable energy sector and selected key macroeconomic indicators was evaluated. The PCA procedure yielded four distinct principle components reflecting the main macroeconomic variables, renewable energy production, industry-energy relations and natural resource revenues. The PCA method offers the possibility to examine the precise correlations and the underlying patterns between the displayed clusters of variables. Meanwhile, the MCA-based cross-country assessment of Azerbaijan's wind, solar and hydropower has struck somewhat pessimistic notes, as the country lags behind neighbouring and other post-Soviet countries (e.g. Estonia, Iran, Latvia, Russia) in developing its green energy sector. These findings are of great interest to policymakers, businesses and academics who wish to gain deep insight into the Azerbaijani economy in terms of renewable energy production. The practical value of the present work also lies in the fact that it analyses a multidimensional and relatively longitudinal dataset (1990–2022), which is an example of a methodological application of two nonparametric approaches.

Key words: Azerbaijani economy, energy transition, multiple correspondence analysis (MCA), nonparametric analysis, renewable energy, principal component analysis (PCA).

1. Introduction

Although developed nations lead the renewable energy, oil-rich countries are also keen to diversify their energy mix despite adoption barriers, including hydrocarbon subsidies, low electricity tariffs, fragmented policies, lack of renewable energy regulation and controlled electricity markets (Al-Sarihi and Mansouri, 2022). Azerbaijan is one of the oil-rich developing countries where renewable energy activities

[1] Department of International Economics and Business, Azerbaijan State University of Economics (UNEC), Baku, Azerbaijan. E-mail: ibrahim.niftiyev@unec.edu.az. ORCID: https://orcid.org/0000-0003-3437-9824.

have recently increased due to new international agreements on the introduction of renewable energy technologies (Mehdiyev, 2022). Only hydropower generation is a well-founded field of alternative energy, which developed during the years of the Soviet Union, while solar and wind energy have experienced an upswing. As oil resources dwindle and global pressure for an energy transition increases, it is becoming increasingly important to understand the dynamics of the renewable energy sector. For this reason, the development of the renewable energy sector could provide additional opportunities to break away from oil dependence as the country's population grows and non-oil industrialization increases. Azerbaijan promises to be an interesting case study where there is tremendous potential for renewable energy development, while both the economic, political and institutional factors are unclear.

The objective of the current research is to examine the relationships and patterns between the renewable energy sector and various macroeconomic variables using principal component analysis (PCA) on the example of the Azerbaijani economy. The secondary research objective is to position Azerbaijan in comparison with selected Eurasian countries (both neighbouring and non-neighbouring) with respect to specific green energy sectors (e.g. wind, solar and hydropower) using multiple correspondence analysis (MCA). Both PCA and MCA allow for the identification of clusters of variables that share similar degrees of variance, a deeper understanding of the large dataset collected and a descriptive exploration of the qualitative nature of the energy transition. The research questions of this study are as follows: With which macroeconomic variables do the indicators for the renewable energy sector in Azerbaijan exhibit longitudinal covariance? What is the relative position of Azerbaijan in terms of specific renewable energy generation compared to the selected Eurasian countries (e.g. solar, wind, hydro)? Although the number of studies analysing the renewable energy sector in Azerbaijan is increasing, there are still research gaps that should be methodologically and empirically analysed in terms of the country's key macroeconomic variables and its position relative to other countries.

This paper contributes to the literature by systematically analysing the relationship between renewable energy and key macroeconomic variables in Azerbaijan between 1990 and 2022. By uncovering common trends and identifying clusters of variables, it provides insights into the complex dynamics and enables a comprehensive understanding of the dataset. It also clarifies which macroeconomic indicators correlate with renewable energy and assesses Azerbaijan's position on specific green energy sources. Furthermore, it provides a methodological exercise that serves as a benchmark study for future research with PCA and MCA.

The next section contains a literature review and the theoretical framework of the study. Section 3 explains the data and methodology, while Section 4 presents the results. Finally, conclusions are drawn in Section 5.

2. Literature review and theoretical framework

In general, scholars argue that Azerbaijan has many opportunities to improve its renewable electricity generation per capita (Vidadili et al., 2017), but this potential remains largely untapped due to slow technological improvements and unclear legal regulations (Mustafayev et al., 2022). With the adoption of the "State Program for the Utilization of Alternative and Renewable Energy Sources in the Republic of Azerbaijan," the Azerbaijani government has also shown its willingness to develop green energy and achieve a sustainable energy transition away from the fossil fuel-based economy (Hasanov, 2023). However, there are still numerous problems and challenges in the form of policy barriers, underdeveloped institutional mechanisms, weak regulations, limited storage capacity and an imbalance between supply and demand.

A large strand of emerging empirical studies (mainly parametric econometric methods) related to the renewable energy sector in Azerbaijan has been conducted by Mukhtarov et al. (2020), Huseynli and Huseynli (2022), Mukhtarov (2022), Kalyoncu et al. (2013), Huseynli (2022) and Hasanov et al. (2023). The general idea behind the above works is that it is possible to identify a casual relationship between economic growth, employment, production, financial development and renewable energy in Azerbaijan, indicating their importance for the national economic development agenda. Similarly, Guliyeva (2023) argued that a linear relationship exists between variables such as CO_2 emissions, energy consumption and economic growth in Azerbaijan using the ARDL method. Based on the same method, Acar et al. (2023) showed that economic growth and ecological footprint of Azerbaijan have an inverted U-shaped relationship with the environmental Kuznets curve. Acar et al. (2023) also claimed that financial development reduced the ecological footprint in Azerbaijan between 1996 and 2017. In addition, Mukhtarov and Aliyev (2022) claimed that institutional quality (represented by the variable government effectiveness) does not have a statistically significant impact on renewable energy production. The authors also pointed out that CO_2 emissions have a statistically insignificant relationship with renewable energy production.

Some studies, such as Ibadoghlu (2023; 2023), give a pessimistic outlook for the Azerbaijani renewable energy sector. They point to a decline in its share of total electricity generation and the reluctance of consumers and industry to adopt green energy technologies despite falling prices (Ibadoghlu, 2022). Although Azerbaijan is an energy exporter, the country is focusing on renewable energies and wants to increase their share of total energy production to 30% by 2030. This transition is in line with efforts to combat climate change and promote sustainable development (Mammadli, 2022).

Azerbaijan exhibits significant potential in renewable energy. The International Energy Agency (IEA, 2023) emphasizes its high potential for sustainable power

generation and natural gas exports. Power generation capacity stood at 16% in 2018, expected to reach 30% by 2030 with energy market reforms. Technical potential is estimated at 3,000 MW for wind, 23,040 MW for solar, 380 MW for bio/waste energy, and 520 MW for small hydropower (IRENA, 2019).

Sound empirical research should be based on a solid, contemporary and effective theoretical foundation. For this reason, the theoretical framework of the present study is based on the metatheoretical framework proposed by Cherp et al. (2018), inspired by the contributions of Elionor Ostrom (2007, 2008). In simple terms, Cherp et al. (2018) claim that national energy transformations emerge through the interaction of three types of systems: techno-economic, socio-technical and political (see Figure 1). Techno-economic factors include energy markets that coordinate energy production, transformation, and ultimately production and consumption. Socio-technical aspects include the necessary knowledge, practices and networks that can be used in the energy transition. Finally, the political factors include the national energy-related policies that regulate and stimulate the energy transition. Cherp et al. (2018) point out that all of these three key aspects develop together and eventually lead to the energy transition. Therefore, capturing variables that cover these aspects both directly and indirectly (at least in the form of proxy variables) will provide unique insight into the research question at hand.

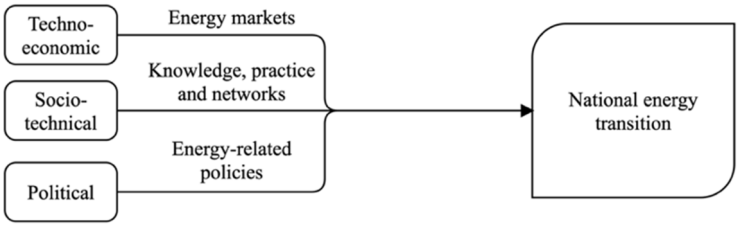


Figure 1: Theoretical framework of the current research

Source: Adapted from Cherp et al. (2018).

In the research design of the present study, due to the limited data available in the case of the Azerbaijani economy, it was not possible to consider every single aspect as precisely as Cherp et al. (2018) describe in their metatheoretical framework. Since the main objective of the work is to evaluate green energy from a multidimensional perspective, there were two major groups of variables, namely those related to renewable energy and economic variables. The latter were mainly formed considering the theoretical framework. For example, variables such as GDP, resource rents and oil rents were considered as techno-economic factors. Manufacturing value added, trade openness, etc. stood for the socio-technical dimension. Finally, institutional quality stands for the political aspects and is itself a composite index constructed for this study.

While additional details on the variables can be found in the Data and Methodology section, the theoretical model of the present study, which is guided by the literature review and theoretical framework, is based on the following motivation and rationale for the selection of variables: Following Mukhtarov et al. (2020), Huseynli and Huseynli (2022), Mukhtarov (2022), Kalyoncu et al. (2013), Huseynli (2022) and Hasanov et al. (2023), Ibadoghlu (2022; 2023), the variables of production, consumption, economic growth and industrial production were selected to measure their covariance with the renewable energy indicators. Similarly, following Acar et al. (2023) and other international publications (Jebli et al., 2020), CO_2 emission variables were collected to capture the environmental dimensions of the process, and institutional variables were developed as composite indices according to the discussion of Mukhtarov and Aliyev (2022) and Aydin (2019). Moreover, technological challenges, especially those related to technology transfer, highlights the inherent complexity associated with renewable energy technologies. By examining the role of technological change, the analysis can provide insights into overcoming barriers and transitioning to a more diversified and sustainable energy portfolio in Azerbaijan. For this reason, the index of technological change was also developed and included in this study.

Due to their immense importance and role in the Azerbaijani economy, oil and natural gas prices as well as economic structure variables (e.g. manufacturing, services, agriculture) and resource rents were also considered to examine their covariation with other variables in the PCA. Finally, the representation of the demand variable in its natural values typically refers to the total population of a given country. Therefore, both the total and urban population variables were included in the main list of variables.

3. Data and methodology

The phases of data analysis included data collection, data interpretation, data trimming, data analysis, clarification of the meaning of the results and overview of the PCA-based index values (see Figure 2). The data were collected from the World Bank, the State Statistical Committee of the Republic of Azerbaijan, etc. First, their descriptive statistics, normality and correlation analyses were performed to analyse the dataset descriptively (descriptive statistics is available upon request). Second, missing values were replaced and outlier values were trimmed (more on this later in this section). Third, PCA and MCA procedures were performed. Finally, the PCA-based index values were visualized. For this purpose, a total of 27 variables of interest were analysed, of which five are renewable energy related variables and 22 are macroeconomic variables covering the period between 1990 and 2022 for PCA. The period for MCA covers 2016–2022. The complete list of variables of interest, their definition, sources and abbreviations can be found in Table 1.

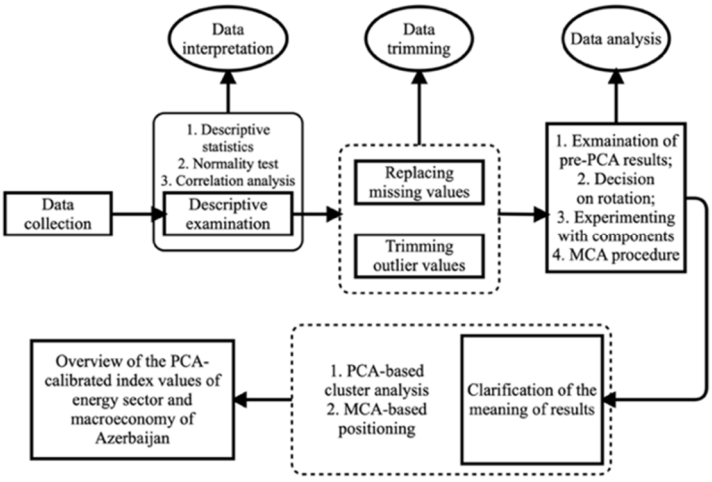


Figure 2: Conceptual framework of the data analysis process

Source: Author's own construction based on the research design.

Variables numbered from 1 to 4 represent renewable energy variables obtained from the Our World in Data platform. Starting from variable 6 onwards, the dataset comprises macroeconomic variables primarily sourced from the World Bank. Notably, two of these variables, namely TechCh and InsQ, are composite index variables specifically constructed for this study.

Table 1: List of the variables of interest used in principal component analysis

No.	Variable name	Abbreviation	Definition and measurement
1	Share of renewable energy in primary energy production	RenEn	Hydropower, solar energy, wind energy, geothermal energy, bioenergy, wave energy and tidal energy are categorized as renewable energy sources. Biofuels are not encompassed within these sources, in %.
2	Per capita electricity generated from renewable sources	RenElPc	Includes hydropower, solar energy, wind energy, geothermal energy, biomass, wave energy and tidal energy, constitutes the total of renewable electricity production, in kilowatt-hours.
3	Per capita consumption of renewable energy	ConRePc	
4	Share of electricity from renewables	RenElShr	Production of electricity from hydropower, solar, wind, biomass and waste, geothermal, wave and tidal sources, in %.

Table 1: List of the variables of interest used in principal component analysis (cont.)

No.	Variable name	Abbreviation	Definition and measurement
5	Low-carbon energy consumption	LoCaEn	Low-carbon energy encompasses the aggregate of nuclear and renewable sources. Among the renewable sources are hydropower, solar energy, wind energy, geothermal energy, wave and tidal energy and bioenergy. Conventional biofuels are excluded from this categorization, in TWh – equivalent.
6	Oil prices	OilPr	USD per barrel, Brent trademark.
7	Natural gas prices	NatGasPr	Average German import price, in US dollars per megawatt-hour (MWh).
8	Total population	TotPop	The total population is determined using the de facto population definition, which includes all residents irrespective of their legal status or citizenship, in persons.
9	Urban population	UrPop	The total size of population living in urban areas, in persons.
10	Agriculture, forestry and fishing, value added	Agr	Economic structure variables. In other words, the percentage share of agriculture, forestry and fishing, industry, services and manufacturing value added in GDP, in %.
11	Industry (including construction), value added	Ind	
12	Services, value added	Serv	
13	Manufacturing, value added	Man	
14	Oil rents	OilRent	The difference between the value of total crude oil and production cost, in % of GDP.
15	Arable land	ArLand	Arable land refers to agricultural land that is suitable and utilized for growing crops. In hectares per person.
16	Consumer price index	CPI	in index values, 2010 = 100.
17	Foreign direct investment	FDI	net inflows, in % of GDP.
18	Gross Domestic Product	GDP	in current USD.
19	General government final consumption expenditure	GovExp	in % of GDP.
20	Gross capital formation	GrCapFor	
21	Trade openness	TrOpp	
22	Total natural resources rents	ResRents	

Table 1: List of the variables of interest used in principal component analysis (cont.)

No.	Variable name	Abbreviation	Definition and measurement
23	Energy use per person	EnUsPerPer	Energy use encompasses more than just electricity; it extends to various sectors of use, encompassing transportation, heating and cooking as well, in kilowatt-hours.
24	Per capita CO_2 emissions	CO_2Emm	Emissions of carbon dioxide (CO_2) originating from fossil fuels and industrial activities, excluding any effects resulting from changes in land use, in tons.
25	Primary energy consumption	PrEnCon	in terawatt-hours (TWh).
26	Technological change	TechCh	Principal component-based index value generated from the aggregation of four information communication variables, namely fixed telephone subscriptions, mobile cellular subscriptions, number of internet users and 4G network coverage.
27	Institutional quality	InsQ	Principal component-based index value generated from the aggregation of five variables, namely Human Rights Scores, Physical Integrity Rights Index, Private Civil Liberties Index, Political Civil Liberties and Individual Liberties and Equality Before the Law Index.

In general, there were only 12 missing values in the dataset, representing only 1.4% of the total observations (879 observations or data points in the 1990 and 2022 range from 27 variables). Some missing values were replaced by the mean of the series, while the missing values for the other variables were replaced by linearly interpolated values. The decision to use different methods to replace missing values depended on the relative position of the missing values—the beginning, middle or end of the time series.

The next part of the preliminary data analysis involved clarification of outlier values. The treatment of outliers in time series data is important because outliers can significantly distort the statistical properties and interpretation of the data in the PCA technique (Jolliffe and Cadima, 2016). Almost half of the 27 variables in the dataset had at least one outlier value in this study. To maintain data integrity, preserve statistical assumptions, minimize bias, achieve robustness, various methods and approaches (e.g. imputation of means, quantile-based flooring and capping, winsorization) were used to

handle outliers. The decision was based on the location and frequency of outliers in time series.

The two main methods used in this work are PCA and MCA, both dimensionality reduction techniques that can be used to represent multivariate, large and longitudinal data in fewer—usually two or more—dimensions (Jolliffe, 2002). PCA is a multivariate technique that analyses datasets where observations are described by multiple correlating quantitative variables. Its goal is to extract the important information from the dataset, represent it as a set of new orthogonal variables called principal components (PCs), and display the similarity patterns of the observations and the variables (Hamid et al., 2016). PCA is a useful technique to capture the covariance in the original data and represent them as less correlated PCs. On the other hand, MCA can be generalized as PCA to deal with qualitative variables and as multiple factor analysis to deal with heterogeneous sets of variables (Van Rijckevorsel and De Leeuw, 1988). MCA is thus an exploratory method for representing multivariate data in two dimensions and, like PCA, is preferred for large datasets (Baxter et al., 1990).

Two main sources were used to conduct the PCA: Sarstedt and Mooi (2014) and Joint Research Center-European Commission (2008) "Handbook on constructing composite indicators: methodology and user guide." Both sources explain comprehensively and step by step how to apply PCA. For the MCA, the basic and theoretical suggestions of Kroonenberg and Greenacre (2004) were considered. Both PCA and MCA were conducted using SPSS version 26 software.

PCA was initially applied by preparing the dataset, which consisted of 27 variables associated with 891 observations (or 33 years for 27 variables). This process involved linearly projecting the data matrix to maximize variance. The significance of the correlation matrix was highlighted, and the optimization problem was addressed by determining eigenvalues along with their corresponding eigenvectors. Based on Wingdes et al. (2021), the formula for PCA used in this study is given below:

$$\max var(\gamma_1^T x) = \gamma_1^T \Sigma \gamma_1 \quad (1)$$
$$s.t. \quad \gamma_1^T \gamma_1 = 1$$

Lagrange multipliers approach was used to maximize $\gamma_1^T \Sigma \gamma_1$ subject to $\gamma^T \gamma = 1$

$$\gamma_1^T \Sigma \gamma_1 - \lambda(\gamma_1^T \gamma_1 - 1) \quad (2)$$

where λ is a Lagrange multiplier. The following is gained after differentiation with respect to γ_1:

$$(\Sigma - \lambda I_P)\gamma_1 = 0 \quad (3)$$

where Σ should be treated as covariance matrix and I_P is another form of $\Sigma\gamma_1 - \lambda\alpha_1 = 0$. Essentially, I_P is the $(p \times p)$ identity matrix. The maximization of the quantity happens as follows

$$\gamma_1^T \Sigma \gamma_1 = \gamma_1^T \lambda \gamma_1 = \lambda \gamma_1^T \gamma_1 = \lambda \quad (4)$$

In order to maximize λ, it is essential to make the corresponding eigenvector as large as possible. The process for determining the second component follows similar steps, as it is associated with the second-largest eigenvalue.

PCA and MCA entail preliminary analyses to adhere to the required methodological criteria. Firstly, the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) and Bartlett's sphericity test assess the suitability of the dataset for PCA. Typically, a KMO threshold of 0.500 is considered appropriate (Kalayci, 2005; Tastan and Yilmaz, 2008), with some dependency on variable types and sample size. In this study, 0.500 served as the focal criterion, aligning well with the study's nature. Secondly, the scree plot, communalities and total explained variance guided PCA progression. Variables with commonality values below 0.750 were excluded. Thirdly, Brown's (2009) recommendations dictated the choice of PCA rotation. To ensure robustness, Direct Oblimin rotation was preferred initially, checking for component correlations above 0.320. If met, oblique rotation was employed. This pragmatic approach aligned with Brown's (2009) guidance, ensuring practical and efficient outcomes.

Finally, MCA was performed at two levels: MCA of 2022 values and MCA of average annual percentage changes between 2014 and 2022. With this approach, we can see not only the current static values for a single year, but also how each country has accelerated its renewable energy production, including Azerbaijan.

4. Results

Throughout the PCA procedure, the significant loadings are those that exceed 0.300 (based on Brown, 2009). The MCA included attributes and countries. The attributes are specific renewable energy sources such as wind, solar and hydropower, while the countries were selected as Eurasian countries to position Azerbaijan in a relative and comparative manner.

4.1. Results of Dimensionality Reduction Analysis (or Principal Component Analysis)

The initial phase of PCA involved a meticulous examination of KMO values, Bartlet's sphericity test outcomes, and extraction results. In situations with large datasets, PCA often necessitates an iterative approach to potentially reduce variables, thus enhancing KMO values and cumulative variance. To attain optimal dimension reduction, a three-stage PCA was conducted. Initially, all variables were included. Subsequently, GovExp, GrCapFr, TrOpp and CO2Emm were excluded due to their low extraction values (see Table 2). To further optimize the process, in the third stage, FDI and Ser were also removed, as their extraction values were the lowest at 0.715 and 0.753, respectively. Consequently, KMO values exhibited incremental improvements, initially

rising from 0.486 to 0.628 and then to 0.689. These transformations indicate both rapid and subsequent moderate enhancements in the dataset's suitability for PCA (see Figure 3, panel *a*).

Table 2: Communalities of the three-stage dimension reduction analysis

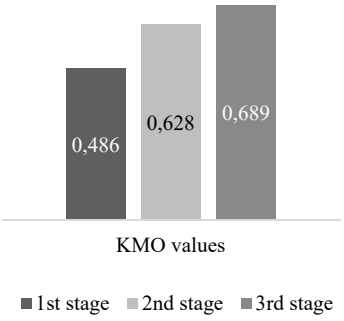
Variable	Extraction	Variable	Extraction	Variable	Extraction
RenEn	0.918	RenEn	0.927	RenEn	0.925
RenElP	0.801	RenElPc	0.881	RenElPc	0.911
ConRenPc	0.821	ConRenPc	0.892	ConRenPc	0.919
LoCaEn	0.921	LoCaEn	0.945	LoCaEn	0.947
RenElShr	0.874	RenElShr	0.909	RenElShr	0.933
NatGasPr	0.819	NatGasPr	0.832	NatGasPr	0.814
OilPr	0.910	OilPr	0.915	OilPr	0.896
TotPop	0.936	TotPop	0.928	TotPop	0.921
FDI	0.836	FDI	0.715		
GDP	0.925	GDP	0.922	GDP	0.913
GovExp	0.749				
GrCapFor	0.727				
TrOpp	0.626				
ResRents	0.849	ResRents	0.892	ResRents	0.950
EnUsPerPer	0.891	EnUsPerPer	0.917	EnUsPerPer	0.938
CO2Emm	0.675				
PrEnCon	0.934	PrEnCon	0.884	PrEnCon	0.887
TechCh	0.961	TechCh	0.959	TechCh	0.956
InsQ	0.725	InsQ	0.788	InsQ	0.783
UrbPop	0.930	UrbPop	0.924	UrbPop	0.915
Agr	0.957	Agr	0.953	Agr	0.948
Ind	0.940	Ind	0.946	Ind	0.925
Ser	0.757	Ser	0.753		
Man	0.896	Man	0.916	Man	0.915
OilRent	0.876	OilRent	0.914	OilRent	0.960
ArLand	0.762	ArLand	0.785	ArLand	0.789
CPI	0.890	CPI	0.882	CPI	0.874

Source: Author's own calculations based on the collected dataset.

Notes: Extraction Method: Principal Component Analysis.

Figure 3, panel *b*, illustrates the alterations in cumulative variance across PCA stages. The exclusion of variables with limited contributions to PCs led to enhancements in both individual and overall cumulative variances. In essence, during the third stage, the primary component elucidated 44.0% of the dataset's main variance, as opposed to 36.9% in the initial stage. Simultaneously, the fourth component attained a cumulative variance of 90.6%, marking a noteworthy increase of 5.8 percentage points relative to the initial stage outcomes.

a. Changes in the KMO values according to PCA stages.



b. Changes in cumulative variance according to PCA stages, in %.

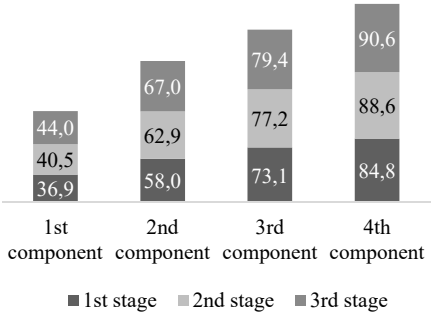


Figure 3: Graphical representation of the pre-PCA procedure of three-stage dimension reduction analysis

Source: Author's calculations based on the collected dataset.

Direct Oblimin rotation showed component correlations below 3.200 (see Table 3), prompting the choice of the Varimax method for rotation. Furthermore, all stages, guided by scree plot values, suggested retaining four PCs, supported by statistically significant results from Bartlett's test of sphericity, denoting substantial correlations among the variables of interest (see Table 4).

Table 3: Component correlation matrix

Component	1	2	3	4
1	1			
2	0.183	1		
3	0.127	–0.092	1	
4	0.106	–0.268	0.027	1

Source: Author's own calculations based on the collected dataset.

Notes: Extraction method: principal component analysis; rotation method: Oblimin with Kaiser normalization.

Table 4: Results of Bartlett's test of sphericity and scree plot analysis

Specification	Bartlett's test of sphericity			Scree plot
	Approximate Chi-square	df	Significance	
1st stage	1907.65	351	0.000	4
2nd stage	1697.45	253	0.000	4
3rd stage	1567.36	210	0.000	4

Source: Author's own calculations based on the collected dataset.

Table 5 presents the results of the third stage PCA, dividing the dataset into four PCs. The first PC exhibits noteworthy positive loadings from TechCh (0.941), GDP (0.925), TotPop (0.917), UrbPop (0.895), CPI (0.880), OilPr (0.872), NatGasPr (0.807) and Ind (0.793). Notably, TechCh, GDP and population-related variables hold substantial loadings above 0.800. Conversely, negative loadings appear on the first PC, including Agr (–0.933), AgLand (–0.872), Man (–0.834), InsQ (–0.693) and EnUsPerPer (–0.321). This reflects a significant decline in the agricultural and manufacturing sectors within Azerbaijan's economy, resulting in negative correlations with the positively loaded variables. A similar trend is observed for agricultural lands and institutional quality, although energy use per person exhibits a weaker association. In essence, the first PC encompasses key macroeconomic variables of Azerbaijan.

Table 5: Rotated component matrix of the analysis

Specification	Components			
	1	2	3	4
TechCh	0.941	–0.228	0.052	–0.121
Agr	–0.933	–0.046	0.274	–0.036
GDP	0.925	–0.222	0.018	0.082
TotPop	0.917	–0.261	–0.029	–0.100
UrbPop	0.895	–0.303	0.089	–0.121
CPI	0.880	–0.284	–0.060	–0.122
OilPr	0.872	0.018	–0.032	0.365
ArLand	–0.872	0.089	–0.112	0.097
Man	–0.834	0.077	0.454	–0.085
NatGasPr	0.807	0.060	–0.097	0.388
Ind	0.793	0.299	–0.259	0.374
InsQ	–0.693	–0.150	–0.522	0.087
LoCaEn	0.017	0.954	0.049	0.185
RenElPc	–0.161	0.928	0.022	0.153
ConRenPc	–0.237	0.917	0.019	0.145
RenElShr	–0.203	0.897	–0.181	0.232
RenEn	–0.029	0.818	–0.482	0.147
PrEnCon	0.057	–0.188	0.919	–0.056
EnUsPerPer	–0.321	–0.050	0.904	0.122
ResRents	0.029	0.301	0.033	0.926
OilRent	0.051	0.380	–0.009	0.902

Source: Author's own calculations based on the collected dataset.

Notes: Extraction method: Principal Component Analysis; Rotation Method: Varimax with Kaiser Normalization; Rotation converged in 5 iterations.

The second PC predominantly represents the renewable energy sector, with LoCaEn (0.954), RenElPc (0.928), ConRenPc (0.917) and RenElShr (0.897) as primary, positively correlated and highly loaded variables. They also exhibit positive correlations with ResRents (0.301) and OilRent (0.380), which aligns with Azerbaijan's economic realities, where resource revenues heavily influence nearly everything, including renewable energy development. UrbPop (–0.303) is the sole negative loading variable in the second PC, indicating an adverse correlation with the aforementioned renewable and non-renewable energy variables. Interestingly, this suggests that while urban population increased, renewable energy sector growth did not follow the same trajectory.

The third PC displays significant loadings from PrEnCon (0.919), EnUsPerPer (0.904) and Man (0.454), all contributing positively. Conversely, InsQ (–0.522) and RenEn (–0.482) exhibit negative loadings on this component. It appears that the third PC primarily represents industrial energy usage, as the main loaders are associated with energy consumption and the manufacturing sector. In other words, this PC signifies the interplay between industry and energy, contrasted by a negative correlation with institutional quality and the proportion of renewable energy in primary energy production. In essence, both consumer-focused renewable energy generation and industrial energy production appear insufficient to meet the demand.

The fourth PC elucidates the role of non-renewable resource rents in Azerbaijan's economy. Notably, all significant loadings are positive, with ResRents (0.926) and OilRent (0.902) as the primary contributors. Additionally, OilPr (0.365), NatGasPr (0.388) and Ind (0.374) exhibit moderate positive loadings on the fourth PC. Given Azerbaijan's oil-centric economy, these findings align with expectations, essentially categorizing the fourth component as "natural resource rents."

To aid visual comprehension, PCA was repeated with a fixed number of PCs set to two, resulting in Figure 4. This representation offers clearer clustering of key economic and renewable energy sector variables compared to the three or four-dimensional counterparts. The first PC features NatGasPr, OilPr, GDP, TechCh, CPI and TotPop as major positive loaders, negatively correlated with ArLand, InsQ, Man, Agr and EnUsPerPer. Notably, Ind loads significantly on both the first and second PCs, highlighting its importance as it bridges both renewable and non-renewable energy sectors within the industry.

In the second PC, RenElShr, RenElPc, ConRenPc, LoCaEn, RenEn, OilRent and ResRents load positively and significantly, whereas PrEnCon is the sole negative and significant loader. This finding aligns with the declining trends in renewable energy-related variables since 2006, contrasted with the rising trend in primary energy consumption. Interestingly, PrEnCon exhibited no significant negative loading on the PC in the initial analysis stage. It appears that the second PC in the two-component solution mirrors the second PC in the initial approach.

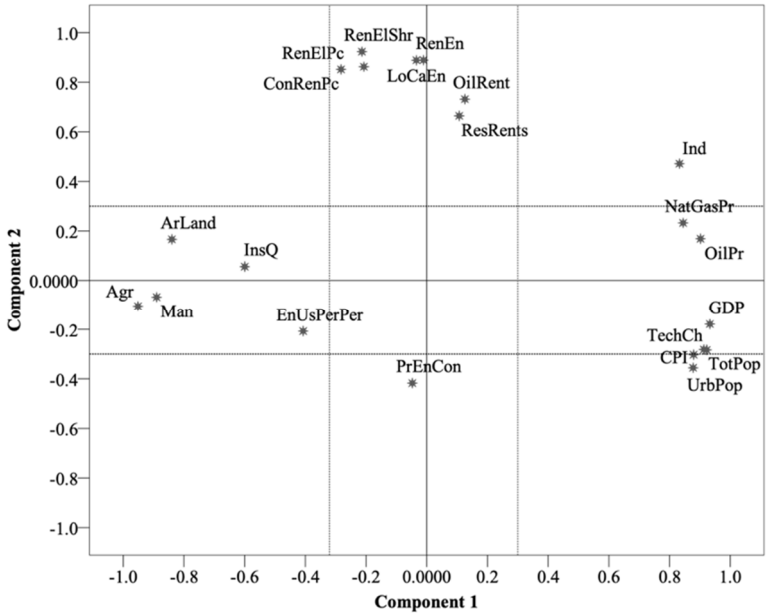


Figure 4: Two-dimensional component loadings, 1990–2022

Source: Author's calculations based on the collected dataset.

4.2. Results of Multivariate Correspondence Analysis

Figure 5 displays the initial MCA findings, encompassing four attributes tied to specific renewable energy sources: solar, wind, hydropower and nuclear. The analysis encompasses 24 countries, allowing a comparative assessment of Azerbaijan's standing. Despite considerable renewable energy potential, Azerbaijan exhibits subpar performance in hydropower, wind and solar energy. For instance, neighbouring countries like Russia and Iran surpass Azerbaijan in hydropower production, and post-Soviet nations such as Ukraine, Estonia and Belarus outperform Azerbaijan in solar energy production. Hungary stands as a notable leader in this regard since the Hungarian government has been proactively advocating the advancement of solar energy through diverse initiatives and policies. In fact, as of the conclusion of 2022, Hungary possessed slightly over 4,000 MW of photovoltaic capacity, marking a substantial surge compared to the levels recorded a decade earlier (Eurobserv-er, 2023). In fact, Hungary's commitment to a clean energy transition has even been recognized by the International Energy Agency (2022). Additionally, Lithuania leads in wind power production, with Mongolia, Türkiye and Poland displaying strong results, while Azerbaijan's wind power generation in 2022 remains limited. Lituania's leading position has been attributed to its geographical position and government's support (Evwind, 2023).

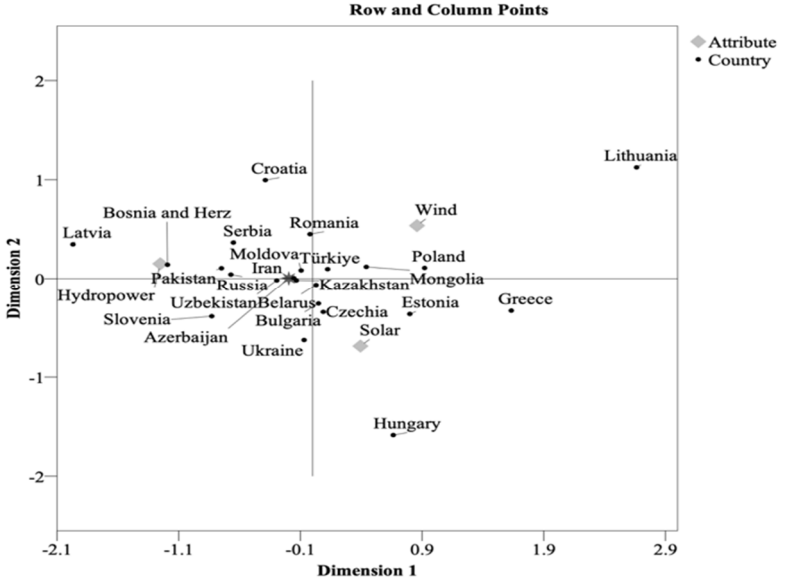


Figure 5: Multiple correspondence analysis of the renewable energy production by source across selected countries, only in 2022.*Source: Author's calculations based on the collected dataset*

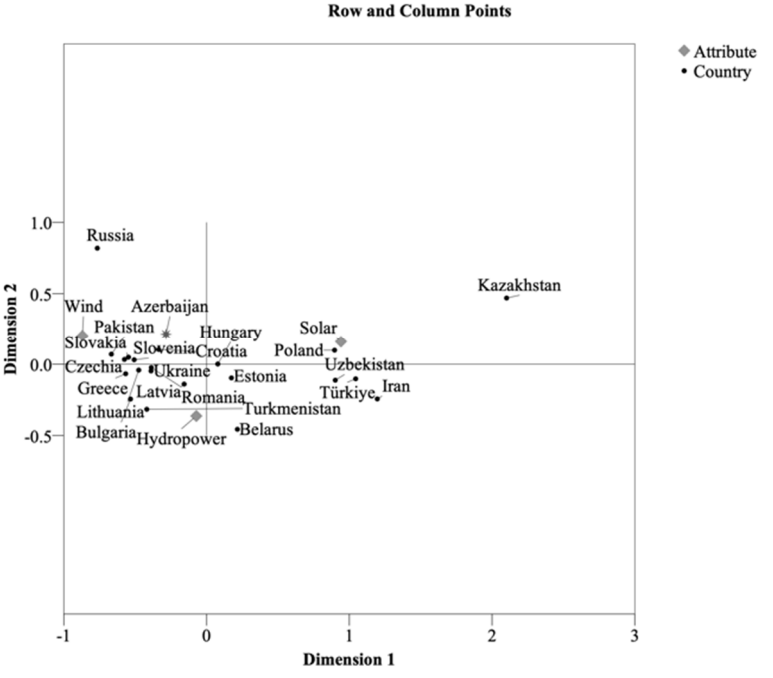


Figure 6: Multiple correspondence analysis of the renewable energy production by source across selected countries, based on annual average percentage changes, 2014–2022

Source: Author's calculations based on the collected dataset.

Azerbaijan's average annual growth rates in solar, wind and hydropower are concerning, showing a considerable gap in solar energy development compared to Kazakhstan, Uzbekistan, Iran and Türkiye. In this aspect, Azerbaijan lags behind Estonia, Hungary and Belarus but surpasses Slovakia, Pakistan, Latvia, Lithuania, Bulgaria and others (see Figure 6). Meanwhile, Azerbaijan's wind energy development improved from 2014 to 2022, akin to Croatia and Russia but behind Pakistan, Czechia, Slovakia and Greece. However, hydropower development lags behind compared to Belarus, Turkmenistan, Bulgaria, Romania and Ukraine, with Azerbaijan outperforming only Russia in this regard.

4.3. Index of Renewable Energy Sector in Azerbaijan

Figure 7 summarizes the dimension reduction analysis of Azerbaijan's economy concerning the renewable energy sector. Factor 1 predominantly represents macroeconomic variables, reflecting overall economic growth and per capita energy use. Factor 2 mainly loads renewable energy variables and resource rents, showing volatility and an overall decline. Factor 3 indicates industry-energy relations, displaying a positive upward trend since 2010 after disruptions in the 1990s. Factor 4, reflecting natural resource rents, is volatile with peaks in 2007 and 2008, lacking a clear trend due to unpredictable oil and natural gas prices and the associated rents in Azerbaijan.

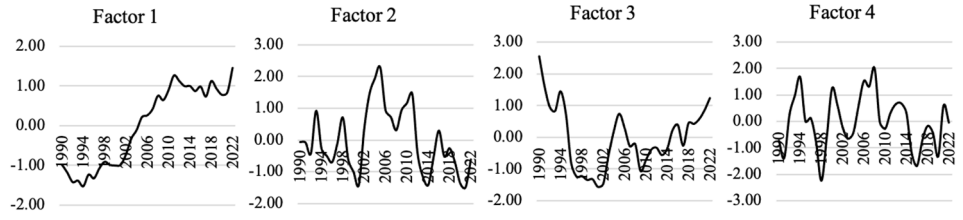


Figure 7: Principal component-based factor scores, in index values, 1990–2022

Source: Author's own calculation based on the collected dataset.

5. Conclusions

Using PCA and MCA approaches as nonparametric dimensionality reduction methods, this study examined two aspects of the Azerbaijani economy: first, what type of underlying relationships (i.e., covariance) exists between a number of macroeconomic variables and variables from the renewable energy sector; second, how Azerbaijan's relative performance in specific renewable energy sectors such as wind, solar and hydropower compares to other selected Eurasian countries (e.g. Kazakhstan, Latvia).

The PCA results effectively segmented the initial dataset of 22 variables (originally 27) into four principal components representing distinct aspects: macroeconomic

indicators (PC 1), the renewable energy sector (PC 2), industry-energy dynamics (PC 3) and natural resource rents (PC 4). PC 1 and PC 3 exhibited consistent positive trends between 1990 and 2020, as indicated by their index values. In contrast, PC 2 and PC 4 displayed volatility without discernible trends. Notably, certain economic metrics like FDI, trade openness, services value added, gross capital formation, government expenditure and CO_2 emissions held limited significance in the PCA, and their exclusion notably enhanced the quality of the analysis by elevating KMO values and cumulative variance explained. Moreover, the MCA results showed that Azerbaijan's abundant renewable energy potential is currently underutilized and lags behind that of other similar post-Soviet countries. These analyses illustrated the methodological exercise of dimensionality reduction in a relatively longitudinal dataset using a small and resource-rich economy from the perspective of renewable energy sector development.

The PCA results unveil several key findings. Firstly, as the Azerbaijani economy recuperated from transition shocks of the 1990s and witnessed an oil boom, energy consumption per capita declined, institutional quality deteriorated, and the contributions of manufacturing and agriculture to value-added diminished (evident from the significant loadings on PC 1). Secondly, there is a strong connection between renewable and non-renewable energy variables, indicating interdependence, where funding for the renewable energy sector often stems from the non-renewable sector's revenue generation capability. Thirdly, renewable energy struggles to keep pace with escalating energy demands, particularly in industrial sectors, amidst declining institutional quality. Lastly, PC 4 demonstrates that even within a dataset encompassing 22 variables, oil and gas-related factors significantly cluster around industrial activities, reflecting the nation's heavy reliance on oil-based export revenues and domestic value added.

The theoretical implications of the current study are as follows. First, economic diversification and reliance on non-energy sectors might contribute to more sustainable energy consumption patterns in small and dependent economies such as Azerbaijan. Second, the strong connection between renewable and non-renewable energy variables underscores the interdependence in Azerbaijan's energy landscape. Third, the MCA results, revealing Azerbaijan's underutilized renewable energy potential compared to other Eurasian countries, have profound theoretical and policy implications. The comparative perspective emphasizes that countries with more or less similar geographical and natural conditions differ in their renewable energy performance, which underlines the role of institutional and political factors in theory building.

The study yields several crucial policy implications for Azerbaijan, a country with substantial untapped potential in solar, wind and hydro energy and a commitment to

greener energy sources. While optimism surrounds this transition, challenges persist that foreign investments alone cannot resolve. First, actionable policies, not just rhetoric, are needed to accelerate renewable energy development. Setting specific targets for renewable energy in production and consumption is vital. Additionally, a revised electricity pricing policy can incentivize solar energy adoption. Second, amidst Azerbaijan's role as a gas supplier to the EU, a comprehensive, long-term energy policy is crucial, considering diversification and a favourable business environment to reduce commodity dependency. Third, the energy sector's vulnerability to international commodity price fluctuations necessitates broader economic diversification. Addressing economic, legal and technical obstacles promptly is essential for sustainable renewable energy growth. Lastly, fostering environmental awareness among key stakeholders and authorities is critical to transition from fossil fuels to green energy. Effective communication on green energy, recycling and environmental risks is vital for progress.

The generalizability of the results obtained in this study is underpinned by the methodological rigor and versatility of the analytical techniques used, namely PCA and MCA. These methods can be applied to countries where the renewable energy sector is still at an early stage of development and where there is a lack of understanding of the sector in terms of the complex socioeconomic relationships.

There are some limitations and areas that should be explored in the future. Nonparametric methods provide insights but do not allow causal inferences due to limited functional specifications, control and precision. Context-specific results limit generalizability beyond Azerbaijan. Endogeneity problems remain, especially in energy economic analyses, which complicates policy analysis. Future research should integrate parametric and nonparametric approaches, examine longer time periods and consider policy implications. Expert interviews could complement quantitative analysis and provide nuanced perspectives on renewable energy dynamics and policy impacts.

References

Acar, S., Altıntaş, N., Haziyev, V., (2023). The effect of financial development and economic growth on ecological footprint in Azerbaijan: An ARDL bound test approach with structural breaks. *Environmental and Ecological Statistics*, 30(1), pp. 41–59.

Al-Sarihi, A., Mansouri, N., (2022). Renewable energy development in the Gulf cooperation council countries: Status, barriers, and policy options. *Energies*, 15(5), 1923.

Aydin, U., (2019). Energy insecurity and renewable energy sources: Prospects and challenges for Azerbaijan. ADBI Working Paper Series No. 992.

Baxter, M. J., Cool, H. E. M., Heyworth, M. P., (1990). Principal component and correspondence analysis of compositional data: some similarities. *Journal of Applied Statistics*, 17(2), pp. 229–235.

Brown, J. D., (2009). Choosing the right type of rotation in PCA and EFA. *JALT Testing & Evaluation SIG Newsletter*, 13(3), pp. 20–25.

Butenko, V., (2022). EU external energy policy: The case of Azerbaijan. *Przegląd Europejski*, 1, pp. 61–71.

Cherp, A., Vinichenko, V., Jewell, J., Brutschin, E., Sovacool, B., (2018). Integrating techno-economic, socio-technical and political perspectives on national energy transitions: A meta-theoretical framework. E*nergy Research & Social Science*, 37, pp. 175–190.

Eichberger, S., Hajiyeva, G. E., Leow, W. J., Damianova, A. J., Golub, E. S., Govorukha, K., … Bagirov, F., (2022). Azerbaijan: Towards green growth-issues Note. World Bank Group, USA.

Eurobserv-er., (2023). Photovoltaic barometer 2023. *Eurobserv-er*, Available at: https://www.eurobserv-er.org/photovoltaic-barometer-2023/ [Accessed 22.01.2024].

Evwind., (2023). Exploring the Untapped Potential of Wind Energy in Lithuania's Energy Market. *REVE*, Available at: https://www.evwind.es/2023/07/05/exploring-the-untapped-potential-of-wind-energy-in-lithuanias-energy-market/92629 [Accessed 22.01.2024].

Guliyeva, S., (2023). Energy consumption, economic growth and CO2 emissions in Azerbaijan. *Multidisciplinary Science Journal*, 5(4), 2023052.

Hamid, H., Aziz, N., Huong, P. N. A., (2016). Variable extractions using principal component analysis and multiple correspondence analysis for large number of mixed variables classification problems. *Global Journal of Pure and Applied Mathematics*, 12(6), pp. 5027–5038.

Hamidova, L., (2018). Diversification of the economy of Azerbaijan: How to overcome resource dependence. In *Sosyoekonomik Boyutlarıyla İnovasyon*, Manisa Celal Bayar Üniversitesi Yayınları, No. 34, pp. 11–18.

Ibadoghlu, G., (2022). Problems and prospects of transition to alternative energy in Azerbaijan. *SSRN Electronic Journal*, Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4249068 [Accessed 26.09.2023].

Ibadoghlu, G., (2023). Is the increase in gas exports from Azerbaijan to Europe an illusion or a reality?. *SSRN Electronic Journal*, Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4361366 [Accessed 26.09.2023].

International Energy Agency–IEA., (2023). Azerbaijan Energy Profile. IEA Publications. Available at: https://iea.blob.core.windows.net/assets/0528affc-d2ba-49c9-ac25-17fc4e8724f7/AzerbaijanEnergyProfile_2023.pdf [Accessed 22.01.2024].

International Renewable Energy Agency., (2019). Renewables Readiness Assessment Republic of Azerbaijan. Available at: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2019/Dec/IRENA_RRA_Azerbaijan_2019.PDF [Accessed 22.01.2024].

Jebli, M. B., Farhani, S., Guesmi, K., (2020). Renewable energy, CO2 emissions and value added: Empirical evidence from countries with different income levels. *Structural Change and Economic Dynamics*, 53, pp. 402–410.

Joint Research Centre-European Commission., (2008). *Handbook on constructing composite indicators: Methodology and user guide*. OECD publishing.

Jolliffe, I. T., (2002). *Principal component analysis for special types of data*. Springer, New York, pp. 338–372.

Jolliffe, I. T., Cadima, J., (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202.

Kalaycı, Ş., (2005). Faktör analizi. In Kalaycı, Ş. (Ed.), *SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri.* Asil Yayın Dağıtım Ltd. Şti., Ankara, Türkiye.

Kroonenberg, P. M., Greenacre, M. J. (2004). *Correspondence analysis*. Encyclopedia of statistical sciences, Wiley, Chichester.

Mammadli, N., (2022). *Azerbaijan Pushes Ahead with Plan to Increase Renewables Share in Energy Mix to 30% by 2030*. Available at: https://caspiannews.com/news-detail/azerbaijan-pushes-ahead-with-plan-to-increase-renewables-share-in-energy-mix-to-30-by-2030-2022-6-21-0/ [Accessed 01.10.2023].

Mehdiyev, M., (2022). *Azerbaijan Pushes for Renewables Expansion with Major Projects*. Available at: https://caspiannews.com/news-detail/azerbaijan-pushes-for-renewables-expansion-with-major-projects-2022-2-15-0/ [Accessed 01.10.2023].

Mukhtarov, S., Aliyev, J., Maharramli, S., (2022). Does institutional quality affect renewable energy in oil-rich developing countries? Evidence from Azerbaijan. In: *Circular economy and the energy market: Achieving sustainable economic development through energy policy*, pp. 173–184, Cham: Springer International Publishing.

Ostrom, E., (2007). A diagnostic approach for going beyond panaceas, *National Academy of Sciences*, 104(39), pp. 15181–15187,

Ostrom, E., (2008). Doing institutional analysis: Digging deeper than markets and hierarchies. In: Claude Ménard, Mary M. Shirley (Eds.), *Handbook of New Institutional Economics*, Springer, Berlin, Heidelberg, pp. 819–848.

Sarstedt, M., Mooi, E., (2014). *A concise guide to market research. The Process, Data, and Methods Using IBM SPSS Statistics*. Springer, Berlin, Heidelberg.

Tastan, M., Yilmaz, K., (2008). Organizational citizenship and organizational justice scales' adaptation to Turkish. *Egitim ve Bilim*, 33(150), pp. 87–96.

Van Rijckevorsel, J. L., De Leeuw, J., (1988). *Component and correspondence analysis. Dimension reduction by functional approximation*. Wiley Series in Probability and Mathematical Statistics, Wiley, Chichester.

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 103–120, https://doi.org/10.59170/stattrans-2024-017
Received – 05.05.2023; accepted – 19.02.2024

On the Poisson-transmuted exponential distribution and its application to frequency of claim in actuarial science

Shamsul Rijal Muhammad Sabri[1], **Ademola Abiodun Adetunji**[2]

Abstract

This study proposes a new discrete distribution in the mixed Poisson paradigm to obtain a distribution that provides a better fit to skewed and dispersed count observation with excess zero. The cubic transmutation map is used to extend the exponential distribution, and the obtained continuous distribution is assumed for the parameter of the Poisson distribution. Various moment-based properties of the new distribution are obtained. The Nelder-Mead algorithm provides the fastest convergence iteration under the maximum likelihood estimation technique. The shapes of the proposed new discrete distribution are similar to those of the mixing distribution. Frequencies of insurance claims from different countries are used to assess the performance of the new proposition (and its zero-inflated form). Results show that the new distribution outperforms other competing ones in most cases. It is also revealed that the natural form of the new distribution outperforms its zero-inflated version in many cases despite having observations with excess zero counts.

Key words: mixed Poisson-exponential distribution, skewed count data, dispersed observation, zero-inflated model, claim frequency.

1. Introduction

Distributions that were well regarded in the past have been improved recently, especially with the advent of software and programming languages that make mathematical computations easier. The situation has provided an enabling platform for introducing more complex distributions that better-fit observations.

In compounding baseline distributions, different techniques have been developed. Among them are: quadratic transmutation (Shaw & Buckley, 2007); beta-extended generalized (Alexander *et al*., 2012); exponentiated generalized (Yousof *et al*., 2015); transmuted exponentiated generalized (Shehata *et al*., 2021); alpha-power transformation (Mahdavi & Kundu, 2017); and various forms of cubic rank

[1] Corresponding author. School of Mathematical Sciences, Universiti Sains Malaysia, 11800, Penang, Malaysia. E-mail: rijal@usm.my.

[2] Department of Statistics, Federal Polytechnic, 35101, Ile-Oluji, Nigeria.

transmutations (Al-kadim, 2018; Aslam *et al*., 2018; Granzotto *et al*., 2017; Rahman *et al*., 2019).

The classical Poisson distribution is popular among distributions for discrete observations. The distribution assumes equidispersion (equality of mean and variance). In many cases, count data are overdispersed (Nikoloulopoulos & Karlis, 2008), especially the frequency of claims in actuary science (Adcock *et al*., 2015; Adetunji & Sabri, 2021; Omari *et al*., 2018) with attending excess zero counts. Assuming the Poisson for such observation may lead to model misspecification. This has led to the development of several techniques to model skewed and overdispersed count data, such as the negative binomial. The negative binomial distributions come from the mixed Poisson paradigm when the gamma distribution is assumed for the parameter of the Poisson (Greenwood & Yule, 1920). Several other continuous distributions with positive supports have been assumed for the Poisson parameter (Al-Awadhi & Ghitany, 2001; Bhati et al., 2015, 2017; Das et al., 2018; Gómez-Déniz & Calderín-Ojeda, 2016; Mahmoudi & Zakerzadeh, 2010).

If a discrete random variable $X \sim Poisson(\lambda)$ and $\lambda \sim \pi(\lambda)$ where $\pi(\lambda)$ is the probability distribution function (PDF) with positive supports, a mixed Poisson distribution (with $\pi(\lambda)$ as the mixing distribution) is obtained when the unconditional distribution for $\boldsymbol{X}$ is expressed from equation (1).

$$P_x = \int_0^\infty f(x|\lambda)\,\pi(\lambda)\,d\lambda \tag{1}$$

With a scale parameter θ, the distribution function (CDF) of a random variable λ with the classical exponential distribution is defined in equation (2) as:

$$G(\lambda) = 1 - e^{-\theta\lambda} \tag{2}$$

Leveraging on the skewness and unimodality properties of transmuted exponential distribution (which also characterize observations in claim frequency (Karlis & Xekalaki, 2005)), this research extends the classical exponential distribution using the cubic transmutation map (Al-kadim, 2018). If the baseline CDF is as given in equation (2) and ρ is the transmutation parameter, the CDF of the cubic rank transmutation map due to (Al-kadim, 2018) is defined in equation (3) as:

$$\pi(\lambda) = (1+\rho)G(\lambda) - 2\rho\big(G(\lambda)\big)^2 + \rho\big(G(\lambda)\big)^3,\ \lambda > 0, |\rho| \le 1 \tag{3}$$

The classical exponential distribution serves as baseline distribution in diverse fields of study. In probability theory and applications, extended exponential distribution has enjoyed patronage (Yang et al., 2021). This spans different fields of study, including environmental, economic, reliability, and industrial (Aguilar *et al*., 2019; Rasekhi *et al*., 2017).

2. Materials and Methods

Cubic Rank Transmuted Exponential Distribution (CRTED)

The distribution is obtained when the transmutation map in equation (3) is used to extend the exponential distribution given in equation (2). This gives the Cubic Rant Transmuted Exponential Distribution (CRTED) with the CDF and PDF given in equations (4) and (5) as:

$$F(\lambda) = 1 - e^{-\theta\lambda} + \rho e^{-2\theta\lambda} - \rho e^{-3\theta\lambda} \tag{4}$$

$$f(\lambda) = \theta e^{-\theta\lambda}\left(1 - 2\rho e^{-\theta\lambda} + 3\rho e^{-2\theta\lambda}\right) \tag{5}$$

The corresponding survival and hazard rate functions are respectively given in equation (6) and (7) as:

$$S(\lambda) = e^{-\theta\lambda} - \rho e^{-2\theta\lambda} + \rho e^{-3\theta\lambda} \tag{6}$$

$$h(\lambda) = \frac{\theta\left(1-2\rho e^{-\theta\lambda}+3\rho e^{-2\theta\lambda}\right)}{1-\rho e^{-\theta\lambda}+\rho e^{-2\theta\lambda}} \tag{7}$$

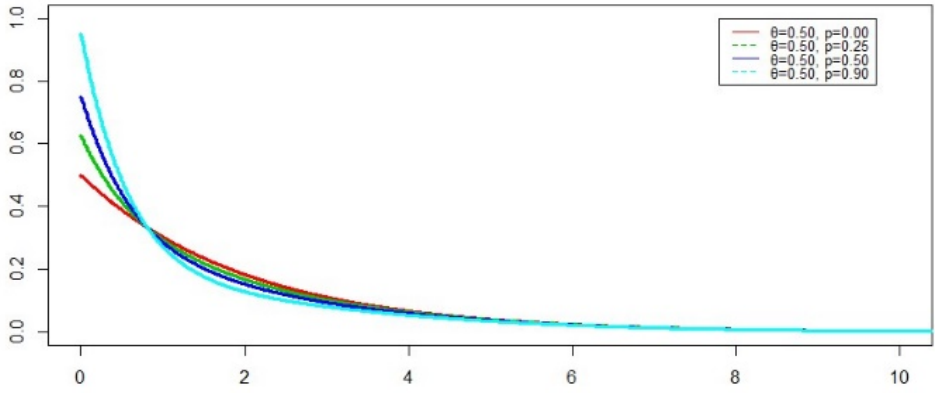


Figure 1: Shapes of the PDF CRTED

Figure 1 shows the shapes of the PDF of the CRTED for different values of θ and ρ. The figure shows that the PDF is a monotonically decreasing function.

r^{th} Moment of the CRTED

Proposition 1: If a random variable λ has a CRTED, the r^{th} moment is defined in equation (8) as:

$$E(\lambda^r) = \left(1 - \frac{\rho}{2^r} + \frac{\rho}{3^r}\right)\frac{r!}{\theta^r} \tag{8}$$

Proof:

$$E(\lambda^r) = \int_0^\infty \lambda^r f(\lambda)\, d\lambda = \int_0^\infty \lambda^r \left(\theta e^{-\theta\lambda} - 2\rho\theta e^{-2\theta\lambda} + 3\rho\theta e^{-3\theta\lambda}\right) d\lambda$$
$$= \theta \int_0^\infty \lambda^r e^{-\theta\lambda}\, d\lambda - 2\rho\theta \int_0^\infty \lambda^r e^{-2\theta\lambda}\, d\lambda + 3\rho\theta \int_0^\infty \lambda^r e^{-3\theta\lambda}\, d\lambda$$
$$= \frac{r!}{\theta^r} - \frac{\rho r!}{(2\theta)^r} + \frac{\rho r!}{(3\theta)^r} = \left(1 - \frac{\rho}{2^r} + \frac{\rho}{3^r}\right)\frac{r!}{\theta^r}$$

***Mo**ment Generating Function* **of the CRTED**

Proposition 2: If a random variable λ has a CRTED, the MGF is defined in equation (9) as:

$$E\left(e^{t\lambda}\right) = \frac{\theta}{\theta - t} - \frac{2\rho\theta}{2\theta - t} + \frac{3\rho\theta}{3\theta - t} \tag{9}$$

Proof:

$$E\left(e^{t\lambda}\right) = \int_0^\infty e^{t\lambda} \left(\theta e^{-\theta\lambda} - 2\rho\theta e^{-2\theta\lambda} + 3\rho\theta e^{-3\theta\lambda}\right) d\lambda$$
$$= \int_0^\infty \theta e^{-(\theta - t)\lambda} - 2\rho\theta e^{-(2\theta - t)\lambda} + 3\rho\theta e^{-(3\theta - t)\lambda}\, d\lambda$$
$$= \theta \int_0^\infty e^{-(\theta - t)\lambda}\, d\lambda - 2\rho\theta \int_0^\infty e^{-(2\theta - t)\lambda}\, d\lambda + 3\rho\theta \int_0^\infty e^{-(3\theta - t)\lambda}\, d\lambda$$
$$= \frac{\theta}{\theta - t} - \frac{2\rho\theta}{2\theta - t} + \frac{3\rho\theta}{3\theta - t}$$

3. Poisson Crted

PMF of the Poisson CRTED

Proposition 3: If a discrete random variable $X \sim Poisson(\lambda)$ and $\lambda \sim CRTED(\rho, \theta)$ then the PMF of ***X*** has a Poisson-CRTED if its PMF is defined in equation (10) as:

$$P_x = \frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}} \tag{10}$$

Proof

Since $f(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$, $\pi(\lambda) = \theta e^{-\theta\lambda}\left(1 - 2\rho e^{-\theta\lambda} + 3\rho e^{-2\theta\lambda}\right)$

$$P_x = \int_0^\infty f(x|\lambda)\, \pi(\lambda)\, d\lambda$$
$$= \int_0^\infty \frac{\lambda^x e^{-\lambda}}{x!} . \theta e^{-\theta\lambda}\left(1 - 2\rho e^{-\theta\lambda} + 3\rho e^{-2\theta\lambda}\right) d\lambda$$
$$= \frac{\theta}{x!} \int_0^\infty \lambda^x e^{-(1+\theta)\lambda}\left(1 - 2\rho e^{-\theta\lambda} + 3\rho e^{-2\theta\lambda}\right) d\lambda$$
$$= \frac{\theta}{x!} \int_0^\infty \left(\frac{1}{(1+\theta)^{x+1}} u^x e^{-u} - \frac{2\rho}{(1+2\theta)^{x+1}} u^x e^{-u} + \frac{3\rho}{(1+3\theta)^{x+1}} u^x e^{-u}\right) du$$
$$= \frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}}$$

Shapes of the PMF of the Poisson CRTED in Figure 2 show that the distribution is unimodal with the ability to model positively skewed observation with excess zero counts.

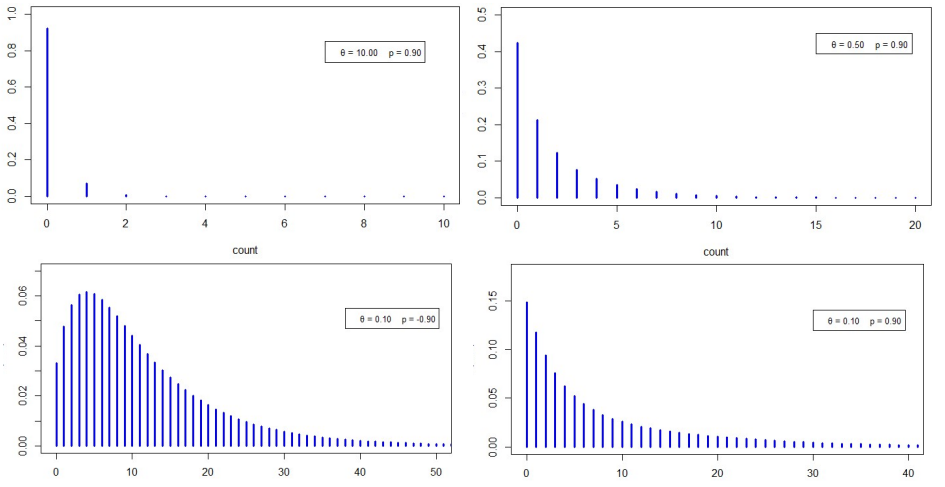


Figure 2: Shapes of the PMF of the Poisson CRTED

Mathematical Properties of the Poisson CRTED

3.1. Probability Generating Function: The PGF of a random variable $\boldsymbol{X}$ with the Poisson-CRTED is obtained in equation (11).

$$P_x(z) = \int_0^\infty e^{\lambda(z-1)}\,\pi(\lambda)d\lambda$$

$$= \int_0^\infty e^{\lambda(z-1)}\theta e^{-\theta\lambda}\left(1 - 2\rho e^{-\theta\lambda} + 3\rho e^{-2\theta\lambda}\right)d\lambda$$

$$= \theta\int_0^\infty \left(e^{-(1-z+\theta)\lambda} - 2\rho e^{-(1-z+2\theta)\lambda} + 3\rho e^{-(1-z+3\theta)\lambda}\right)d\lambda$$

Therefore,

$$P_x(z) = \frac{\theta}{1+\theta-z} - \frac{2\rho\theta}{1+2\theta-z} + \frac{3\rho\theta}{1+3\theta-z} \tag{11}$$

3.2. Moment Generating Function: The MGF is obtained in equation (12) by replacing $\boldsymbol{z}$ with e^t in (11).

$$M_X(t) = \frac{\theta}{1+\theta-e^t} - \frac{2\rho\theta}{1+2\theta-e^t} + \frac{3\rho\theta}{1+3\theta-e^t} \tag{12}$$

3.3. Mean and Variance: The mean and variance of the Poisson-CRTED are respectively obtained in equations (13) and (14) as:

$$E(X) = \frac{6-\rho}{6\theta} \tag{13}$$

$$Var(X) = \frac{36+2\rho-\rho^2-6\rho\theta+36\theta}{36\theta^2} \tag{14}$$

The Coefficient of Variation and the Dispersion Index are given in equations (15) and (16).

$$CV(X) = \frac{\sqrt{36+2\rho-\rho^2-6\rho\theta+36\theta}}{6-\rho} \tag{15}$$

$$DI(X) = \frac{36+2\rho-\rho^2-6\rho\theta+36\theta}{6\theta(6-\rho)} \tag{16}$$

3.4. Skewness and Kurtosis: The first four raw moments for a random variable with the Poisson-CRTED are presented in equations (17) to (20).

$$E(X) = \frac{6-\rho}{6\theta} \tag{17}$$

$$E(X^2) = \frac{36-5\rho+18\theta-3\rho\theta}{18\theta^2} \tag{18}$$

$$E(X^3) = \frac{216-19\rho+(216-30\rho)\theta+(36-6\rho)\theta^2}{36\theta^3} \tag{19}$$

$$E(X^4) = \frac{1296-65\rho+(54-9\rho)\theta^3+(756-105\rho)\theta^2+(1944-171\rho)\theta}{54\theta^4} \tag{20}$$

The skewness and kurtosis for the Poisson CRTED are given in equations (21) and (22) as:

$$S_k(X) = \frac{2(216+33\rho+3\rho^2-\rho^3+(324+18\rho-9\rho^2)\theta+(108-18\rho)\theta^2)}{(36+2\rho-\rho^2+(36-6\rho)\theta)^{\frac{3}{2}}} \tag{21}$$

$$K(X) = \frac{1296(1+\theta)(9+9\theta+\theta^2)-3\rho^4+12\rho^3(1-3\theta)+24\rho^2(2-3\theta-6\theta^2)+8\rho(201-27\theta^3-99\theta^2+189\theta)}{(\rho^2+(6\theta-2)\rho-36\theta-36)^2} \tag{22}$$

Tables 1–3 show simulated Skewness, Kurtosis, and Dispersion Index for selected parameters of the Poisson CRTED

Table 1: Skewness for some parameters of the Poisson-CRTED

Specification	$\theta = 0.1$	$\theta = 2.0$	$\theta = 10$
$\rho = -0.9$	1.920	2.056	3.300
$\rho = -0.5$	1.938	2.162	3.433
$\rho = 0.0$	2.002	2.309	3.618
$\rho = 0.5$	2.108	2.478	3.827
$\rho = 0.9$	2.222	2.632	4.016

Table 2: Kurtosis for some parameters of the Poisson-CRTED

Specification	$\theta = 0.1$	$\theta = 2.0$	$\theta = 10$
$\rho = -0.9$	8.928	8.908	15.366
$\rho = -0.5$	8.877	9.479	16.476
$\rho = 0.0$	9.009	10.333	18.091
$\rho = 0.5$	9.360	11.387	20.028
$\rho = 0.9$	9.813	12.415	21.879

Table 3: Dispersion Index for some parameters of the Poisson-CRTED

Specification	$\theta = 0.1$	$\theta = 2.0$	$\theta = 10$
$\rho = -0.9$	9.065	1.403	1.081
$\rho = -0.5$	9.910	1.446	1.089
$\rho = 0.0$	11.000	1.500	1.100
$\rho = 0.5$	12.136	1.557	1.111
$\rho = 0.9$	13.088	1.604	1.121

Remarks:

i. For fixed ρ, both skewness and kurtosis increase as θ increases.
ii. For fixed ρ, the dispersion index decreases as θ increases.
iii. For fixed θ, both skewness and kurtosis increase as ρ increases.
iv. For fixed θ, the dispersion index slowly increases as ρ increases.

Maximum Likelihood Estimation of the Poisson-CRTED

Assuming $x_1, x_2, \dots, x_n$ are a random sample of size n from the Poisson-CRTED, the log-likelihood function of the distribution is obtained in equation (23) as:

$$\mathcal{L} = \prod_{i=1}^{n} P_{(x_i)} = \prod_{i=1}^{n} \left(\frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}} \right)$$
$$\ell = log\, \mathcal{L} = \sum_{i=1}^{n} log \left(\frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}} \right) \quad (23)$$

Equation (23) gives a non-linear model that can be solved numerically with different algorithms from the optimr (Nash *et al.*, 2019) from the R language (R-Core Team, 2020).

4. Zero-Inflated Poisson-Crted

Proposition 4: If a discrete random variable $\boldsymbol{X}$ has a Poisson-CRTED with PMF denoted by P_x, if the inflation parameter is denoted with τ, the PMF of the zero-inflated Poisson-CRTED is given in equation (24).

$$P_x^{ZI} = \begin{cases} \tau + (1-\tau)P_0, & x = 0 \\ (1-\tau)P_x, & x = 1,2,3, \dots \end{cases}$$

This is obtained as:

$$P_x^{ZI} = \begin{cases} \tau + (1-\tau)\left(\frac{\theta}{(1+\theta)} - \frac{2\rho\theta}{(1+2\theta)} + \frac{3\rho\theta}{(1+3\theta)}\right), & x = 0 \\ (1-\tau)\left(\frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}}\right), & x = 1,2,3,\ldots \end{cases} \tag{24}$$

Proof

Since the PMF of the Poisson CRTED is given as: $P_x = \frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}}$ and $P_0 = \frac{\theta}{(1+\theta)} - \frac{2\rho\theta}{(1+2\theta)} + \frac{3\rho\theta}{(1+3\theta)}$, hence the result.

Mathematical Properties of the Zero-Inflated Poisson CRTED

The PGF of the Zero-Inflated Poisson CRTED denoted by $P_x^{ZI}(z)$ is obtained using: $P_x^{ZI}(z) = (1-\tau)P_x(z)$ where $P_x(z)$ is the PGF of the Poisson CRTED in equation (11). Hence, the PGF of the ZI-Poisson CRTED is given in equation (25) as:

$$P_x^{ZI}(z) = (1-\tau)\left(\frac{\theta}{1+\theta-z} - \frac{2\rho\theta}{1+2\theta-z} + \frac{3\rho\theta}{1+3\theta-z}\right) \tag{25}$$

The MGF is therefore expressed in equation (26) as:

$$M_X^{ZI}(t) = (1-\tau)\left(\frac{\theta}{1+\theta-e^t} - \frac{2\rho\theta}{1+2\theta-e^t} + \frac{3\rho\theta}{1+3\theta-e^t}\right) \tag{26}$$

If the r^{th} moment of the Poisson CRTED is denoted by $E(X^r)$, then the r^{th} moment of the ZI-Poisson CRTED is defined as:
$m_r = E(X_{ZI}^r) = (1-\tau)E(X^r)$

Hence, the first four moments of the ZI-Poisson CRTED are given in equations (27) – (30).

$$m_1 = (1-\tau)\frac{6-\rho}{6\theta} \tag{27}$$

$$m_2 = (1-\tau)\frac{36-5\rho+18\theta-3\rho\theta}{18\theta^2} \tag{28}$$

$$m_3 = (1-\tau)\frac{216-19\rho+(216-30\rho)\theta+(36-6\rho)\theta^2}{36\theta^3} \tag{29}$$

$$m_4 = (1-\tau)\frac{1296-65\rho+(54-9\rho)\theta^3+(756-105\rho)\theta^2+(1944-171\rho)\theta}{54\theta^4} \tag{30}$$

MLE of the Parameters of the ZI-Poisson CRTED

If a random variable $\boldsymbol{X}$ is assumed to follow the ZI-Poisson CRTED with its PMF P_x indexed with (θ, ρ), and τ as the zero-inflation parameter, then the likelihood function is defined as:

$$\mathcal{L}(\tau, \Theta) = \prod_{n_0}\left[(\tau + (1-\tau)P_0\right] \prod_{n_1}\left[\left((1-\tau)P(x>0)\right)\right.$$

where n_0 is the frequency of zero counts in the dataset; n_1 is the frequency of non-zero counts; $n = (n_0 + n_1)$, P_0 is the realization of at $x = 0$. The log-likelihood function is obtained as follows:

$$\ell = n_0 \ln(\tau + (1-\tau)P_0) + n_1 \ln(1-\tau) + \left(\sum_{n_1} \ln(P_x)\right)$$

$$= n_0 \ln\left(\tau + (1-\tau)\left(\frac{\theta}{(1+\theta)} - \frac{2\rho\theta}{(1+2\theta)} + \frac{3\rho\theta}{(1+3\theta)}\right)\right) + n_1 \ln(1-\tau) + \sum_{n_1} \ln\left(\left(\frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}}\right)\right)$$

$$\frac{\partial \ell}{\partial \tau} = \frac{n_0(1-P_0)}{\tau + (1-\tau)P_0} - \frac{n_1}{(1-\tau)}$$

$$\hat{\tau} = \frac{n_0}{n_1} - \frac{n_1}{n}\left(\frac{P_0}{1-P_0}\right)$$

The MLE for parameters (τ, θ, ρ) is obtained numerically by solving $\frac{\partial \ell}{\partial \tau} = 0$, $\frac{\partial \ell}{\partial \theta} = 0$ and $\frac{\partial \ell}{\partial \rho} = 0$. Among competing algorithms for optimizations that come with the optimr packages (Nash *et al*., 2019) in R language (R-Core Team, 2020), the Nelder-Mead provides the fastest iterations for convergence and the least log-likelihood values.

Competing Distributions

The new propositions are assessed by comparing their performances with the (i) Poisson, (ii) Zero Inflated Poisson (ZIP), (iii) Negative Binomial (Neg. Bin.), and (iv) Zero Inflated Negative Binomial (ZI-Neg. Bin.) distributions.

5. Simulation Studies

The new proposition is assessed in this study by simulating count observations with similar characteristics to the claim frequency in actuaries. The simulated data are based on the assumption that a policyholder may not report more than 4 claims in a reference period (usually a year). The algorithm utilized for the simulation is:

i. Specify the proportion of zero in each sample (P_0 = 0.5, 0.7, 0.9)
ii. Simulate a random sample of size n (n = 50, 250, 500, and 1000)
iii. Obtain parameter estimates and log-likelihood for each assumed distribution.

Table 4: Results for simulated data when $P_0 = 0.50$

Sample Size	*Distribution*	*Parameter Estimates*	*–LL*
$n = 50$	*PCRTED*	$\hat{\theta} = 1.341;\ \hat{\rho} = -0.771$	62.259
	ZI-PCRTED	$\hat{\theta} = 20.060;\ \hat{\rho} = 5.867;\ \hat{\tau} = -716.5$	65.594
	Poisson	$\hat{\theta} = 0.840$	65.213
	ZI-Poisson	$\hat{\theta} = 1.247;\ \hat{\tau} = 0.327$	61.232
	Neg. Bin.	$\hat{\theta} = 1.543;\ \hat{\rho} = 0.647$	63.034
	ZI-Neg. Bin.	$\hat{\theta} = 0.00004;\ \hat{\rho} = 0.357;\ \hat{\tau} = -117.2$	63.188
			65.079

Table 4: Results for simulated data when $P_0 = 0.50$ (cont.)

Sample Size	*Distribution*	*Parameter Estimates*	*–LL*
n = 250	*PCRTED*	$\hat{\theta} = 1.209;\ \hat{\rho} = -0.484$	323.713
	ZI-PCRTED	$\hat{\theta} = 1.256;\ \hat{\rho} = 0.696;\ \hat{\tau} = 0.372$	387.764
	Poisson	$\hat{\theta} = 0.892$	339.839
	ZI-Poisson	$\hat{\theta} = 1.379;\ \hat{\tau} = 0.353$	324.602
	Neg. Bin.	$\hat{\theta} = 1.357;\ \hat{\rho} = 0.603$	326.427
	ZI-Neg. Bin.	$\hat{\theta} = 0.763;\ \hat{\rho} = 0.518;\ \hat{\tau} = 0.377$	387.993
n = 500	*PCRTED*	$\hat{\theta} = 1.330;\ \hat{\rho} = -0.647$	628.953
	ZI-PCRTED	$\hat{\theta} = 1.338;\ \hat{\rho} = -0.006;\ \hat{\tau} = 0.677$	718.784
	Poisson	$\hat{\theta} = 0.832$	653.206
	ZI-Poisson	$\hat{\theta} = 1.244;\ \hat{\tau} = 0.331$	631.018
	Neg. Bin.	$\hat{\theta} = 1.414;\ \hat{\rho} = 0.630$	629.484
	ZI-Neg. Bin.	$\hat{\theta} = 1.072;\ \hat{\rho} = 0.588;\ \hat{\tau} = 0.681$	718.719
n = 1000	*PCRTED*	$\hat{\theta} = 1.377;\ \hat{\rho} = -1.220$	1284.447
	ZI-PCRTED	$\hat{\theta} = 1.365;\ \hat{\rho} = -0.852;\ \hat{\tau} = 0.838$	1404.054
	Poisson	$\hat{\theta} = 0.873$	1322.324
	ZI-Poisson	$\hat{\theta} = 1.192;\ \hat{\tau} = 0.268$	1293.749
	Neg. Bin.	$\hat{\theta} = 1.785;\ \hat{\rho} = 0.672$	1285.411
	ZI-Neg. Bin.	$\hat{\theta} = 1.562;\ \hat{\rho} = 0.651;\ \hat{\tau} = 0.838$	1403.110

Table 5: Results for simulated data when $P_0 = 0.70$

Sample Size	*Distribution*	*Parameter Estimates*	*–LL*
n = 50	*PCRTED*	$\hat{\theta} = 1.891;\ \hat{\rho} = 0.788$	43.395
	ZI-PCRTED	$\hat{\theta} = 5.929;\ \hat{\rho} = 5.435;\ \hat{\tau} = -25.544$	43.421
	Poisson	$\hat{\theta} = 0.460$	47.622
	ZI-Poisson	$\hat{\theta} = 0.926;\ \hat{\tau} = 0.503$	45.397
	Neg. Bin.	$\hat{\theta} = 0.760;\ \hat{\rho} = 0.623$	45.425
	ZI-Neg. Bin.	$\hat{\theta} = 0.0003;\ \hat{\rho} = 0.449;\ \hat{\tau} = -10.40$	46.220
n = 250	*PCRTED*	$\hat{\theta} = 1.896;\ \hat{\rho} = 0.314$	236.621
	ZI-PCRTED	$\hat{\theta} = 2.039;\ \hat{\rho} = 2.445;\ \hat{\tau} = -0.047$	287.128
	Poisson	$\hat{\theta} = 0.500$	248.799
	ZI-Poisson	$\hat{\theta} = 0.912;\ \hat{\tau} = 0.452$	239.519
	Neg. Bin.	$\hat{\theta} = 0.911;\ \hat{\rho} = 0.646$	238.655
	ZI-Neg. Bin.	$\hat{\theta} = 0.362;\ \hat{\rho} = 0.552;\ \hat{\tau} = -0.032$	287.579
n = 500	*PCRTED*	$\hat{\theta} = 1.910;\ \hat{\rho} = 0.574$	461.388
	ZI-PCRTED	$\hat{\theta} = 1.969;\ \hat{\rho} = 1.559;\ \hat{\tau} = 0.560$	541.627
	Poisson	$\hat{\theta} = 0.474$	483.392
	ZI-Poisson	$\hat{\theta} = 0.919;\ \hat{\tau} = 0.484$	462.576
	Neg. Bin.	$\hat{\theta} = 0.822;\ \hat{\rho} = 0.634$	462.559
	ZI-Neg. Bin.	$\hat{\theta} = 0.539;\ \hat{\rho} = 0.587;\ \hat{\tau} = 0.564$	542.152
n = 1000	*PCRTED*	$\hat{\theta} = 1.879;\ \hat{\rho} = 0.776$	910.515
	ZI-PCRTED	$\hat{\theta} = 1.907;\ \hat{\rho} = 1.246;\ \hat{\tau} = 0.791$	1021.272
	Poisson	$\hat{\theta} = 0.464$	958.230
	ZI-Poisson	$\hat{\theta} = 0.938;\ \hat{\tau} = 0.506$	912.204
	Neg. Bin.	$\hat{\theta} = 0.755;\ \hat{\rho} = 0.619$	913.020
	ZI-Neg. Bin.	$\hat{\theta} = 0.617;\ \hat{\rho} = 0.597;\ \hat{\tau} = 0.792$	1022.125

Table 6: Results for simulated data when $P_0 = 0.90$

Sample Size	*Distribution*	*Parameter Estimates*	*–LL*
n = 50	*PCRTED*	$\hat{\theta} = 3.055;\ \hat{\rho} = 3.282$	21.126
	ZI-PCRTED	$\hat{\theta} = 11.610;\ \hat{\rho} = 5.759;\ \hat{\tau} = -34.25$	25.230
	Poisson	$\hat{\theta} = 0.160$	24.740
	ZI-Poisson	$\hat{\theta} = 1.027;\ \hat{\tau} = 0.844$	21.039
	Neg. Bin.	$\hat{\theta} = 0.144;\ \hat{\rho} = 0.473$	21.880
	ZI-Neg. Bin.	$\hat{\theta} = 0.00005;\ \hat{\rho} = 0.392; \hat{\tau} = -14.41$	26.255
n = 250	*PCRTED*	$\hat{\theta} = 1.811;\ \hat{\rho} = 4.072$	122.289
	ZI-PCRTED	$\hat{\theta} = 4.097;\ \hat{\rho} = 5.615;\ \hat{\tau} = -11.987$	125.321
	Poisson	$\hat{\theta} = 0.208$	156.709
	ZI-Poisson	$\hat{\theta} = 1.493;\ \hat{\tau} = 0.861$	121.234
	Neg. Bin.	$\hat{\theta} = 0.106;\ \hat{\rho} = 0.338$	122.435
	ZI-Neg. Bin.	$\hat{\theta} = 0.00001;\ \hat{\rho} = 0.303; \hat{\tau} = -69.32$	126.856
n = 500	*PCRTED*	$\hat{\theta} = 1.922;\ \hat{\rho} = 4.033$	238.213
	ZI-PCRTED	$\hat{\theta} = 3.752;\ \hat{\rho} = 5.471;\ \hat{\tau} = -5.372$	261.363
	Poisson	$\hat{\theta} = 0.200$	302.900
	ZI-Poisson	$\hat{\theta} = 1.440;\ \hat{\tau} = 0.861$	236.515
	Neg. Bin.	$\hat{\theta} = 0.108;\ \hat{\rho} = 0.350$	240.191
	ZI-Neg. Bin.	$\hat{\theta} = 0.022;\ \hat{\rho} = 0.321;\ \hat{\tau} = -2.541$	266.676
n = 1000	*PCRTED*	$\hat{\theta} = 2.061;\ \hat{\rho} = 3.918$	462.265
	ZI-PCRTED	$\hat{\theta} = 2.500;\ \hat{\rho} = 4.501;\ \hat{\tau} = 0.311$	543.589
	Poisson	$\hat{\theta} = 0.193$	587.124
	ZI-Poisson	$\hat{\theta} = 1.372;\ \hat{\tau} = 0.859$	464.867
	Neg. Bin.	$\hat{\theta} = 0.112;\ \hat{\rho} = 0.367$	473.291
	ZI-Neg. Bin.	$\hat{\theta} = 0.066;\ \hat{\rho} = 0.351;\ \hat{\tau} = 0.325$	541.587

When 50% of simulated data is zero, Table 4 reveals that the new proposition (PCRTED) provides the best fit in most cases across different sample sizes. The ZI-Poisson has a relatively better fit for small samples. It is also observed that the natural forms of the PCRTED and Neg. Bin. provide better fits than their respective zero-inflated forms.

Table 5 shows various statistics when the simulated data have a 70% proportion of zero counts. The ZI-Poisson has a relatively better fit for small samples. It is also observed that the natural forms of the PCRTED and Neg. Bin. provide better fits than their respective zero-inflated forms.

Table 5 shows various statistics when the simulated data have 70% proportion of zero counts. The PCRTED, ZI-Poisson, and Neg. Bin. have relatively better fits than other competing distributions.

When 90% of the simulated dataset are zero counts, the ZI-Poisson better fits smaller samples, while the PCRTED best fits larger samples, as shown in Table 6.

Generally, the –LL statistics increase as the sample size increases for a different proportion of zero counts, while it reduces as the proportion of zero counts increases.

6. Applications

Four sets of claim frequency from different countries are assessed for model comparisons in the study. The first dataset represents claim frequency from automobile injuries from the General Insurance Association of Singapore in 1993 (Frees, 2010; Frees & Valdez, 2008). The second observation is the frequency of third-party claims for Australian vehicle owners (De Jong & Heller, 2008). The third dataset is the insurance claim from Belgium in 1993 (Denuit, 1997; Zamani & Ismail, 2014). The last set of observations considered in the study represents the claim frequency of 10,814 policyholders for the automobile portfolio in a Turkish insurance company between 2012 and 2014 (Meytrianti *et al.*, 2019).

All datasets considered are positively skewed and dispersed with varying proportions of zero counts that often characterize the frequency of claims in actuarial science, as shown in Table 7. This is suggestive that the classical Poisson distribution may not provide an adequate fit.

Table 7: Descriptive Statistics of the Claim Frequency

Dataset	*Dispersion Index*	*Skewness*	*Kurtosis*	*% of Zero*
Dataset I	1.09	4.27	20.88	93.49
Dataset II	1.06	4.07	18.50	93.19
Dataset III	1.09	3.52	14.59	90.33
Dataset IV	1.26	2.56	7.71	79.01

7. Results and Discussion

The results from analyzing the four datasets assuming the new proposition (and zero-inflated form) and the competing distributions are presented in Tables 8 to 11.

Table 8: Parameter estimates for claim frequency from Singapore (dataset I)

Obser.	*Freq.*	*PCRTED*	*ZI-PCRTED*	*Poisson*	*ZI-Poisson*	*Neg. Bin.*	*ZI-Neg. Bin.*
0	6996	6996.26	7000.79	6977.80	6995.92	6996.71	7006.80
1	455	452.95	449.63	487.75	452.76	452.52	444.59
2	28	31.42	31.25	17.05	32.69	31.38	28.85
3	4	2.20	1.30	0.40	1.57	2.22	2.49
	$\hat{\theta}$	13.52	58.41	0.07	0.14	0.87	0.01
	$\hat{\tau}$	0.33	5.68		0.52	0.93	0.87
	$\hat{\rho}$		-75.60				-73.71
–LL		1931.40	1939.36	1941.18	1933.17	1932.38	1938.20
Chi-Square		1.05	5.98	41.96	4.43	1.79	1.80

From Table 8, the new proposition (PCRTED) provides the best fit judging by the smallest –LL and chi-square statistics values. It is also observed that the PCRTED and Neg. Bin. provide better fit to the datasets when compared with their respective zero-inflated forms.

Table 9: Parameter estimates for claim frequency from Australia (dataset II)

Obser.	*Freq.*	*PCRTED*	*ZI-PCRTED*	*Poisson*	*ZI-Poisson*	*Neg. Bin.*	*ZI-Neg. Bin.*
0	63232	63232.31	63227.36	63091.61	63230.49	63230.60	63317.89
1	4333	4330.09	4321.38	4593.07	4325.83	4330.57	4252.49
2	271	275.14	298.19	167.19	286.59	276.48	261.98
3	18	17.29	8.91	4.06	12.66	17.22	21.45
4	2	1.09	0.15	0.07	0.42	1.06	1.97
	$\hat{\theta}$	14.62	85.73	0.07	0.13	1.16	0.01
	$\hat{\tau}$	-0.38	5.79		0.45	0.94	0.88
	$\hat{\rho}$		-179.55				-76.06
–LL		18048.63	18081.29	18101.50	18052.20	18049.68	18105.58
Chi-Square		0.75	34.56	177.66	9.07	0.98	2.51

Table 9 shows that PCRTED is the best fit for the dataset. The negative binomial distribution follows this, while the Poisson distribution does not fit well.

Table 10: Parameter estimates for claim frequency from Belgium (dataset III)

Obser.	*Freq.*	*PCRTED*	*ZI-PCRTED*	*Poisson*	*ZI-Poisson*	*Neg. Bin.*	*ZI-Neg. Bin.*
0	57178	57180.35	57196.18	56949.76	57177.48	57188.34	57249.63
1	5617	5594.29	5584.82	6019.59	5584.80	5581.31	5558.90
2	446	480.24	497.42	318.14	504.87	485.28	438.37
3	50	40.40	20.12	11.21	30.43	40.47	45.91
4	8	3.42	0.46	0.30	1.38	3.30	5.40
	$\hat{\theta}$	10.52	63.36	0.11	0.18	1.28	0.01
	$\hat{\tau}$	-0.67	5.78		0.42	0.92	0.84
	$\hat{\rho}$		-179.11				-71.13
–LL		22063.75	22123.32	22150.54	22075.30	22064.31	22136.57
Chi-Square		1.97	174.23	413.84	51.55	12.33	2.44

The PCRTED also best fits the third dataset with the least value of *–LL* (22063.75) and the chi-square statistic (1.97) from Table 10. The performance of the PCRTED is significantly better than its zero-inflated form (ZI-PCRTED).

Table 11: Parameter estimates for claim frequency from Turkey (dataset IV)

Obser.	*Freq.*	*PCRTED*	*ZI-PCRTED*	*Poisson*	*ZI-Poisson*	*Neg. Bin.*	*ZI-Neg. Bin.*
0	8544	8544.62	8544.30	8292.42	8544.19	8543.47	8561.78
1	1796	1793.77	1765.70	2201.64	1759.23	1795.62	1807.66
2	370	375.96	428.09	292.27	430.75	375.71	331.89
3	81	78.77	69.38	25.87	70.31	78.50	81.03
4	23	16.50	6.07	1.72	8.61	16.39	22.23
	$\hat{\theta}$	3.77	16.19	0.27	0.49	1.01	0.01
	$\hat{\tau}$	-0.01	5.67		0.46	0.79	0.63
	$\hat{\rho}$		-76.17				-82.43
–LL		7028.72	7066.88	7153.16	7038.91	7029.71	7057.06
Chi-Square		2.72	57.60	484.41	35.02	2.84	4.51

With the lowest –LL (7028.72) and chi-square (2.72), the PCRTED best fits the fourth dataset. Both PCRTED and Neg. Bin. provide a better fit for the dataset than their zero-inflated versions.

8. Conclusion

This study proposed a new continuous lifetime distribution with positive support using a one-parameter cubic transmutation map to extend the exponential distribution. The new distribution is assumed for the Poisson parameter in the mixed Poisson paradigm. A new mixed Poisson distribution (the Poisson Cubic Rank Transmuted Exponential Distribution, PCRTED) is proposed along with its zero-inflated version. Various moment-based mathematical properties of the PCRTED are obtained. The shapes of the new discrete distribution proposed are similar to that of the continuous mixing distribution.

With a focus on the frequency of claims in insurance, datasets with varying proportions of zero counts are simulated at different sample sizes. The performance of the new proposition is compared with the classical Poisson and negative binomial distributions (with their zero-inflated forms). The real-life application of the new proposition is assessed on claim frequency from different countries. Results show that the new proposition better fits various datasets than competing distributions using both –LL and chi-square goodness of fit statistics as the selection criteria. It is also found that the natural form of the new distribution outperforms its zero-inflated version in many cases despite having observations with higher-than-expected frequencies of zero counts.

Data Availability

The details of the data used for this study have been discussed in the article.

Declaration of Interest

The authors have no conflicts of interest to report.

References

Adcock, C., Eling, M., & Loperfido, N., (2015). Skewed distributions in finance and actuarial science: a review. *European Journal of Finance*, *21*(13–14), pp. 1253–1281. https://doi.org/10.1080/1351847X.2012.720269

Adetunji, A. A., Sabri, S. R. M., (2021). Modelling Claim Frequency in Insurance Using Count Models. *Asian Journal of Probability and Statistics*, *14*(4), pp. 14–20. https://doi.org/10.9734/ajpas/2021/v14i430334

Aguilar, G. A. S., Moala, F. A., & Cordeiro, G. M., (2019). Zero-Truncated Poisson Exponentiated Gamma Distribution : Application and Estimation Methods. *Journal of Statistical Theory and Practice*. https://doi.org/10.1007/s42519-019-0059-2

Al-Awadhi, S. A., Ghitany, M. E., (2001). Statistical Properties of Poisson–Lomax Distribution and its Application to Repeated Accidents Data. *Journal of Applied Statistical Science*, *10*(4), pp. 365–372.

Al-kadim, K. A., (2018). Proposed Generalized Formula for Transmuted Distribution. *Journal of Babylon University, Pure and Applied Sciences*, *26*(4), pp. 66–74.

Alexander, C., Cordeiro, G. M., Ortega, E. M. M., & Sarabia, J. M., (2012). Generalized Beta-Generated Distributions. *Computational Statistics & Data Analysis*, *56*, pp. 1880–1897.

Aslam, M., Hussain, Z., & Asghar, Z., (2018). Cubic Transmuted-G family of distributions and its properties. *Stochastic and Quality Control, De Gruyte*, *33*(2), pp. 103–112. https://doi.org/10.1515/eqc-2017-0027

Bhati, D., Kumawat, P., & Gómez–Déniz, E., (2017). A New Count Model Generated from Mixed Poisson Transmuted Exponential Family with an application to Health Care Data. *Communications in Statistics - Theory and Methods*, *46*(22), pp. 11060–11076. https://doi.org/10.1080/03610926.2016.1257712

Bhati, D., Sastry, D. V. S., & Maha Qadri, P. Z., (2015). A New Generalized Poisson-Lindley Distribution: Applications and Properties. *Austrian Journal of Statistics*, *44*(4), pp. 35–51. https://doi.org/10.17713/ajs.v44i4.54

Das, K. K., Ahmed, I., & Bhattacharjee, S., (2018). A New Three-Parameter Poisson-Lindley Distribution for Modelling Over-dispersed Count Data. *International Journal of Applied Engineering Research*, *13*(23), pp. 16468–16477. http://www.ripublication.com

De Jong, P., Heller, G. Z., (2008). *Generalized Linear Models for Insurance Data*. Cambridge University Press.

Denuit, M., (1997). A New Distribution of Poisson-Type for the Number of Claims. *ASTIN Bulletin*, *27*(2), pp. 229–242. https://doi.org/10.2143/AST.27.2.542049

Frees, E. W., (2010). *Regression Modeling with Actuarial and Financial Applications*. Cambridge University Press. https://doi.org/10.1017/CBO9780511814372

Frees, E. W., Valdez, E. A., (2008). Hierarchical insurance claims modeling. *Journal of the American Statistical Association*, *103*(484), pp. 1457–1469. https://doi.org/10.1198/016214508000000823

Gómez-Déniz, E., Calderín-Ojeda, E., (2016). The Poisson-Conjugate Lindley Mixture Distribution. *Communications in Statistics - Theory and Methods*, *45*(10), pp. 2857–2872. https://doi.org/10.1080/03610926.2014.892134

Granzotto, D. C. T., Louzada, F., & Balakrishnan, N., (2017). Cubic Rank Transmuted Distributions: Inferential Issues and Applications. *Journal of Statistical Computation and Simulation*, *87*(14), pp. 2760–2778. https://doi.org/10.1080/00949655.2017.1344239

Greenwood, M., Yule, G. U., (1920). An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks of Disease or of Repeated Accidents. *Journal of the Royal Statistical Society*, *83*(2), p. 255. https://doi.org/10.2307/2341080

Karlis, D., Xekalaki, E., (2005). Mixed Poisson distributions. *International Statistical Review*, *73*(1), pp. 35–58. https://doi.org/10.1111/j.1751-5823.2005.tb00250.x

Mahdavi, A., Kundu, D., (2017). A new Method for Generating Distributions with an Application to Exponential Distribution. *Communications in Statistics - Theory and Methods*, *46*(13), pp. 6543–6557. https://doi.org/10.1080/03610926. 2015.1130839

Mahmoudi, E., Zakerzadeh, H., (2010). Generalized Poisson-Lindley Distribution. *Communications in Statistics - Theory and Methods*, *39*(10), pp. 1785–1798. https://doi.org/10.1080/03610920902898514

Meytrianti, A., Nurrohmah, S., & Novita, M., (2019). An Alternative Distribution for Modelling Overdispersion Count Data: Poisson Shanker Distribution. *ICSA - International Conference on Statistics and Analytics 2019*, *1*, pp. 108–120. https://doi.org/10.29244/icsa.2019.pp108-120

Nash, J. C., Varadhan, R., & Grothendieck, G., (2019). *optimr Package: A Replacement and Extension of the "optim" Function*.

Nikoloulopoulos, A. K., & Karlis, D., (2008). On modeling count data: A comparison of some well-known discrete distributions. *Journal of Statistical Computation and Simulation*, *78*(3), pp. 437–457. https://doi.org/10.1080/10629360601010760

Omari, C. O., Nyambura, S. G., & Mwangi, J. M. W., (2018). Modeling the Frequency and Severity of Auto Insurance Claims Using Statistical Distributions. *Journal of Mathematical Finance*, *8*(1), pp. 137–160. https://doi.org/10.4236/jmf.2018.81012

R-Core Team, (2020). *R: A language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/

Rahman, M. M., Al-Zahrani, B., Shahbaz, S. H., & Shahbaz, M. Q., (2019). Cubic Transmuted Uniform Distribution: An Alternative to Beta and Kumaraswamy Distributions. *European Journal of Pure and Applied Mathematics*, *12*(3), 1106–1121. https://doi.org/10.29020/nybg.ejpam.v12i3.3410

Rasekhi, M., Alizadeh, M., Altun, E., Hamedani, G., Afify, A. Z., & Ahmad, M., (2017). The Modified Exponential Distribution with Applications. *Pakistan Journal of Statistics*, *33*(5), pp. 383–398.

Shaw, W. T., Buckley, I. R. C., (2007). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *Research Report*.

Shehata, W. A. M., Yousof, H., & Aboraya, M., (2021). A Novel Generator of Continuous Probability Distributions for the Asymmetric Left-skewed Bimodal Real-life Data with Properties and Copulas. *Pakistan Journal of Statistics and Operation Research*, *17*(4), pp. 943–961. https://doi.org/10.18187/pjsor.v17i4.3903

Yang, Y., Tian, W., & Tong, T., (2021). Generalized Mixtures of Exponential Distribution and Associated Inference. *Mathematics*, *9*(12), pp. 1–22. https://doi.org/10.3390/math9121371

Yousof, H. M., Afify, A. Z., Alizadeh, M., Butt, N. S., Hamedani, G. . G., & Ali, M. M., (2015). The Transmuted Exponentiated Generalized-G Family of Distributions. *Pakistan Journal of Statistics and Operation Researchakis*, *11*(4), pp. 441–464. https://doi.org/10.18187/pjsor.v11i4.1164

Zamani, H., Ismail, N., (2014). Functional form for the Zero Inflated Generalized Poisson Regression Model. *Communications in Statistics - Theory and Methods*, *43*, pp. 515–529.

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 121–136, https://doi.org/10.59170/stattrans-2024-018
Received – 13.02.2023; accepted – 03.10.2023

Extended odd Frechet-exponential distribution with applications related to the environment

Muzamil Jallal[1], **Aijaz Ahmed**[2], **Rajnee Tripathi**[3]

Abstract

In this paper, we attempted to expand the Frechet distribution by employing the T-X family of distributions and named the newly formulated model Extended odd Frechet-exponential distribution (EOFED). Several structural properties, reliability measurements and characteristics were estimated and discussed. The study presents graphs which depict the behaviour of the probability density function, cumulative distribution function and the hazard rate function. The adaptability and flexibility of this novel distribution were achieved through the application of real-world data sets. A simulation study was performed to evaluate and compare the output efficacy of the estimators.

Key words: Frechet distribution, moments, quantile function, Renyi entropy, maximum likelihood estimation, order statistics.

1. Introduction

The Frechet distribution is a subset of the generalized extreme value distribution named after a French mathematician Maurice Rene Frechet (1878–1973), who defined it as the limiting potential distribution for sequences of extremes and published a related work in 1927. Later, Fisher and Tippett in 1928 and Gumbel in 1958 completed even more research. Rosin and Rammler employed it in 1933 to match a particle size distribution. By employing various techniques and methods, many researchers have done a lot of work by extending the Frechet distribution and thus showing its importance in various fields like Social, Medical and Engineering Sciences (such as work done by Haq et al. (2017), Pelumi et al. (2019), Penson et al. (2014), Hamed M S (2020) and so on).

[1] Bhagwant University, Ajmer, India. E-mail: muzamiljallal@gmail.com.
[2] Bhagwant University, Ajmer, India. E-mail: ahmadaijaz4488@gmail.com.
[3] Bhagwant University, Ajmer, India. E-mail: rajneetripathi8@gmail.com. ORCID: https://orcid.org/0000-0002-0301-9845.

The cdf of the extended Frechet distribution is given by

$$F(t) = (1 + \beta t^{-\alpha})^{\frac{-1}{\beta}} \quad ; t > 0, \alpha, \beta > 0 \tag{1.1}$$

$$f(t) = \alpha t^{-\alpha-1}(1 + \beta t^{-\alpha})^{\frac{-1}{\beta}-1} \; ; t > 0, \alpha, \beta > 0 \tag{1.2}$$

The exponential distribution is the continuous analogue of a geometric distribution. The exponential distribution is one parametric continuous distribution and has various attractive statistical characteristics like memoryless property. Historically, the first lifetime model for which statistical methods were extensively developed was the exponential distribution. In particular, it has its applications in various fields like biomedicine, engineering, economics, medical sciences, etc.; see (Tomitaka, Kawasaki, Ide, Akutagawa, Yamada and Furukawa (2017).

$$G(x, \theta) = 1 - e^{-\theta x} \, ; x \, , \theta > 0 \tag{1.3}$$

$$\bar{G}(x, \theta) = e^{-\theta x} \, ; x \, , \theta > 0 \tag{1.4}$$

The related pdf of the exponential distribution is

$$g(x, \theta) = \theta e^{-\theta x} \, ; x \, , \theta > 0 \tag{1.5}$$

Transformed-Transformer (T-X) family of distributions (Alzaatreh et al. (2013)) is given by

$$F(x) = \int_0^{W[G(x)]} f(t)dt \tag{1.6}$$

where f(t) is the probability density function of a random variable T and $W[G(x)]$ is a function of cumulative density function of random variable X.

Suppose $[G, \xi]$ denotes the baseline cumulative distribution function, which depends on parameter vectorξ. Now, using T-X approach, the cumulative distribution function of the extended odd Frechet-exponential distribution (EOFED) can be derived by replacing f(t) in equation (1.6) by equation (1.2) and $W[G(x)] = \frac{G(x,\xi)}{\bar{G}(x,\xi)}$,where $\bar{G}(x, \xi) = 1 - G(x, \xi)$, which follows

$$F(x, \xi) = \int_0^{\frac{G(x,\xi)}{\bar{G}(x,\xi)}} \alpha t^{\alpha-1}(1 + \beta t^{-\alpha})^{\frac{-1}{\beta}-1} dt$$

$$F(x, \xi) = \left[1 + \beta \left(\frac{G(x,\xi)}{\bar{G}(x,\xi)}\right)^{-\alpha}\right]^{\frac{-1}{\beta}} \tag{1.7}$$

The corresponding pdf of (1.7) becomes

$$f(x, \xi) = \alpha g(x, \xi) \frac{(G(x,\xi))^{-\alpha-1}}{(\bar{G}(x,\xi))^{-\alpha+1}} \left[1 + \beta \left(\frac{G(x,\xi)}{\bar{G}(x,\xi)}\right)^{-\alpha}\right]^{\frac{-1}{\beta}-1} \tag{1.8}$$

2. Mixture Form

From equation (1.8) we have

$$f(x, \xi) = \alpha g(x, \xi) \frac{(G(x, \xi))^{-\alpha-1}}{(\bar{G}(x, \xi))^{-\alpha+1}} \left[1 + \beta \left(\frac{G(x, \xi)}{\bar{G}(x, \xi)}\right)^{-\alpha}\right]^{\frac{-1}{\beta}-1}$$

Using binomial expansion

$$(1+x)^{-a} = \sum_{k=0}^{\infty} \binom{-a}{k} x^k |x| < 1$$

$$f(x,\xi) = \alpha g(x,\xi) \frac{(G(x,\xi))^{-\alpha-1}}{(\bar{G}(x,\xi))^{-\alpha+1}} \sum_{p=0}^{\infty} \binom{-\frac{1}{\beta}-1}{p} \beta^p \left(\frac{G(x,\xi)}{\bar{G}(x,\xi)}\right)^{-p\alpha}$$

Again, using binomial expansion we get

$$f(x,\xi) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty} (-1)^q \binom{-\frac{1}{\beta}-1}{p} \binom{(p+1)\alpha-1}{q} \alpha\beta^p g(x,\xi)(G(x,\xi))^{q-(p+1)\alpha-1}$$

Using (1.3) and (1.5) in the above expression we get

$$f(x,\xi) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{s=0}^{\infty} (-1)^q \binom{-\left(\frac{1}{\beta}+1\right)}{p} \binom{(p+1)\alpha-1}{q} \binom{q-(p+1)\alpha-1}{s} (-1)^s \alpha\beta^p \theta e^{-\theta x}\left(e^{-\theta x s}\right)$$

$$f(x,\xi) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{s=0}^{\infty} \varsigma_{pqs} \alpha\beta^p \theta e^{-\theta x(s+1)} \tag{2.1}$$

where $\varsigma_{pqs} = (-1)^q(-1)^s \binom{-\left(\frac{1}{\beta}+1\right)}{p} \binom{(p+1)\alpha-1}{q} \binom{q-(p+1)\alpha-1}{s}$

3. Formation of Extended Odd Frechet-Exponential Distribution

The extended odd Frechet-exponential distribution (EOFED) is developed by using the T-X family of distributions as described by Alzaatreh et al. (2013). We formulated the distribution as follows.

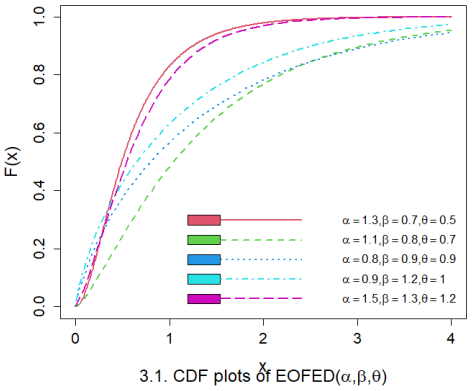
Substituting equation (1.3) and (1.4) in equation (1.7), we obtain the cdf of formulated distribution, which follows

$$F(x,\alpha,\beta,\theta) = \left[1+\beta\left(e^{\theta x}-1\right)^{-\alpha}\right]^{\frac{-1}{\beta}}; x>0, \alpha,\beta,\theta>0 \tag{3.1}$$

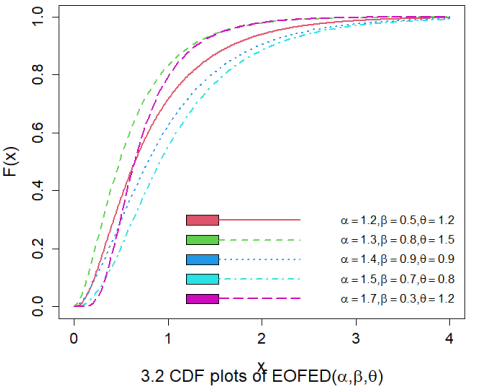
The related pdf of (3.1) is given by

$$f(x,\alpha,\beta,\theta) = \alpha\theta e^{-\theta x}\left(e^{-\theta x}-1\right)^{-\alpha-1}\left[1+\beta\left(e^{-\theta x}-1\right)^{-\alpha}\right]^{\frac{-1}{\beta}-1}; x>0, \alpha,\beta,\theta>0 \tag{3.2}$$

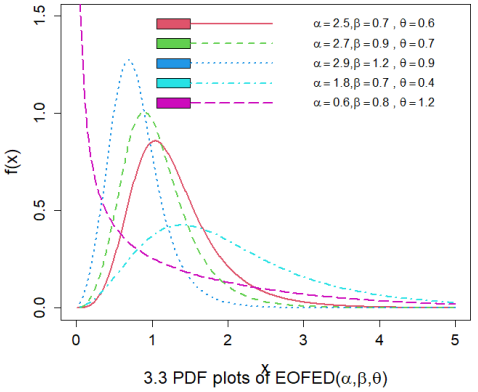
Figures (3.1), (3.2), (3.3) and (3.4) expound some of possible contours of cdf and pdf for a distinct choice of parameters respectively.



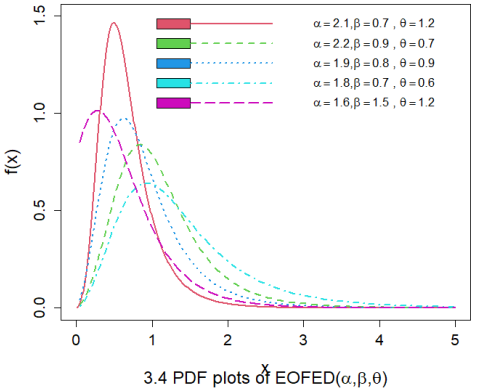
3.1. CDF plots of EOFED(α,β,θ)



3.2 CDF plots of EOFED(α,β,θ)



3.3 PDF plots of EOFED(α,β,θ)



3.4 PDF plots of EOFED(α,β,θ)

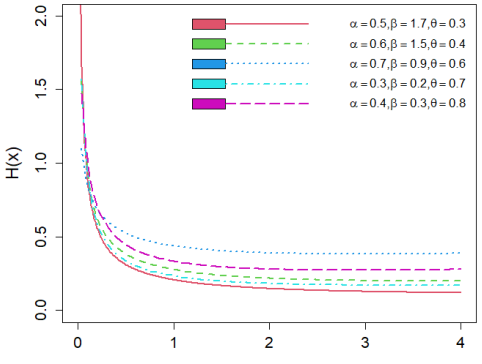
4. Reliability Measures

Survival function, hazard function and reverse hazard function of extended odd Frechet-exponential distribution (EOFED) are given by

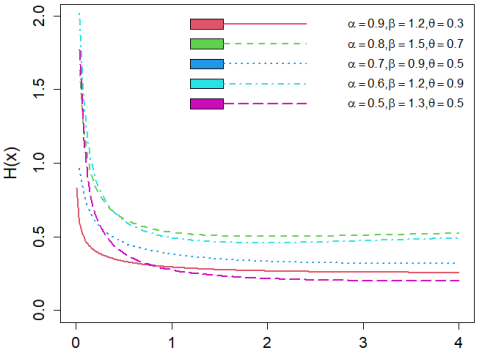
$$S_x(x) = 1 - \left[1 + \beta\left(e^{\theta x} - 1\right)^{-\alpha}\right]^{\frac{-1}{\beta}}$$

$$H(x) = \frac{\alpha\theta e^{\theta x}\left(e^{\theta x}-1\right)^{-\alpha-1}\left(1+\beta\left(e^{\theta x}-1\right)^{-\alpha}\right)^{\frac{-1}{\beta}-1}}{1-\left\{1+\beta\left(e^{\theta x}-1\right)^{-\alpha}\right\}^{\frac{-1}{\beta}}}$$

Figure (4.1) and (4.2) shows some of possible shapes of extended odd Frechet-exponential distribution's hazard rate function for distinct choice of parameters respectively.



4.1. HRF plots of EOFED(α,β,θ)



4.2. HRF plots of EOFED(α,β,θ)

5. Structural properties of extended odd Frechet-exponential distribution (EOFED)

Theorem 1: *If* $\rightleftarrows x \sim E\,OFED(\alpha,\beta,\theta)$ then its r[th] moment is given by

$$E(x)^r = \mu_r{}' = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{s=0}^{\infty} \varsigma_{pqs}\frac{\Gamma(r+1)}{((s+1)\theta)^{r+1}}$$

Proof: We know that r[th] moment about origin is given by

$$\mu_r{}' = E(x^r) = \int_0^{\infty} x^r f(x,\alpha,\beta,\theta)dx$$

$$= \int_0^{\infty} x^r \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{s=0}^{\infty}(-1)^q \binom{-\left(\frac{1}{\beta}+1\right)}{p}\binom{(p+1)\alpha-1}{q}\binom{q-(p+1)\alpha-1}{s}(-1)^s\alpha\beta^p\theta e^{-\theta x}\left(e^{-\theta xs}\right)dx$$

$$= \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{s=0}^{\infty}(-1)^q(-1)^s \binom{-\left(\frac{1}{\beta}+1\right)}{p}\binom{(p+1)\alpha-1}{q}\binom{q-(p+1)\alpha-1}{s}\alpha\beta^p\theta \int_0^{\infty} x^r\, e^{-\theta x}\left(e^{-\theta xs}\right)dx$$

After solving the above integral, we get the following expression

$$\mu_r{}' = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{s=0}^{\infty}\varsigma_{pqs}\frac{\Gamma(r+1)}{((s+1)\theta)^{r+1}} \tag{5.1}$$

where $\varsigma_{pqs} = (-1)^q(-1)^s \binom{-\left(\frac{1}{\beta}+1\right)}{p}\binom{(p+1)\alpha-1}{q}\binom{q-(p+1)\alpha-1}{s}\alpha\beta^p\theta$

Putting r=1,2,3,4 in the above equation we get first four moments about origin.

Theorem 2: *If* $\rightleftarrows x \sim E\,OFED(\alpha,\beta,\theta)$, then it shows that the moment generating function of the extended odd Frechet-exponential distribution (EFOED) is given by

$$M_x(t) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{s=0}^{\infty}\sum_{r=0}^{\infty}\frac{t^r}{r!}\varsigma_{pqs}\frac{\Gamma(r+1)}{((s+1)\theta)^{r+1}}$$

Proof: We know that the moment generating function is given by

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f(x;\alpha,\beta,\theta)dx$$

Applying Taylor's theorem

$$M_x(t) = E(e^{tx}) = \int_0^\infty \left\{1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \ldots\right\} f(x;\alpha,\beta,\theta)dx$$
$$= \sum_{r=0}^\infty \frac{t^r}{r!} \sum_{p=0}^\infty \sum_{q=0}^\infty \sum_{s=0}^\infty \varsigma_{pqs} \frac{\Gamma(r+1)}{((s+1)\theta)^{r+1}}$$

Using (5.1)

$$M_x(t) = \sum_{p=0}^\infty \sum_{q=0}^\infty \sum_{s=0}^\infty \sum_{r=0}^\infty \frac{t^r}{r!} \varsigma_{pqs} \frac{\Gamma(r+1)}{((s+1)\theta)^{r+1}} \tag{5.2}$$

The characteristic function of the extended odd Frechet-exponential distribution (EOFED) is obtained by replacing $t$by itin equation (5.2) as follows:

$$\Phi_x(it) = E(e^{itx}) = \int_0^\infty e^{itx} f(x;\alpha,\beta,\theta)dx$$
$$\Phi_x(it) = \sum_{p=0}^\infty \sum_{q=0}^\infty \sum_{s=0}^\infty \sum_{r=0}^\infty \frac{(it)^r}{r!} \varsigma_{pqs} \frac{\Gamma(r+1)}{((s+1)\theta)^{r+1}} \tag{5.3}$$

6. Quantile Function

The quantile function of the above distribution can be put in the following form:

$$Q_n(u) = X_q = F^{-1}(u)$$

where $Q_n(u)$ represents the quantile function of $F(x)$ for $u \in (0,1)$.

Suppose we put $F(x) = \left[1 + \beta\left(e^{\theta x} - 1\right)^{-\alpha}\right]^{\frac{-1}{\beta}} = u$ (6.1)

After solving equation (6.1), we obtain the quantile function of the extended odd Frechet-exponential distribution as given by

$$Q_n(u) = X_q = \frac{1}{\theta} log\left[1 + \left(\frac{\mu^{-\beta} - 1}{\beta}\right)^{\frac{-1}{\alpha}}\right]$$

7. Renyi Entropy

If x is a continuous random variable following EOFED with pdf $f(x;\alpha,\beta,\theta)$, then

$$R_e = \frac{1}{1-\rho} log\left\{\int_0^\infty f^\rho(x,\alpha,\beta,\theta)dx\right\}$$

$$R_e = \frac{1}{1-\rho} log\left\{\sum_{p=0}^\infty \sum_{q=0}^\infty \binom{-\rho\left(\frac{\beta+1}{\beta}\right)}{p} \binom{(p+\rho)\alpha - 1}{q} \alpha^\rho \beta^p \int_0^\infty \left(\theta e^{-\theta x}\right)^\rho \left(1 - e^{-\theta x}\right)^{q-(\rho+p)\alpha-1} dx\right\}$$

$$R_e = \frac{1}{1-\rho} log \sum_{p=0}^\infty \sum_{q=0}^\infty \sum_{s=0}^\infty (-1)^q (-1)^s \binom{-\rho\left(\frac{\beta+1}{\beta}\right)}{p} \binom{(p+\rho)\alpha - 1}{q} \binom{q - (p+\rho)\alpha - 1}{s} \alpha^\rho \beta^p \int_0^\infty \left(\theta e^{-\theta x}\right)^\rho e^{-\theta x s} dx$$

$$R_e = \frac{1}{1-\rho} log\left\{\sum_{p=0}^\infty \sum_{q=0}^\infty \sum_{s=0}^\infty \varsigma_{pqs} \frac{\theta^\rho}{(s+\rho)\theta}\right\} \tag{7.1}$$

8. Tsallis Entropy

If x is a continuous random variable following EOFED with pdf $f(x;\alpha,\beta,\theta)$, then

$$s_e = \frac{1}{\rho-1}\left\{1-\int_0^{\infty} f^{\rho}(x,\alpha,\beta,\theta)dx\right\}$$

$$s_e = \frac{1}{\rho-1}\left\{1-\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{s=0}^{\infty}\varsigma_{pqs}\frac{\theta^{\rho}}{(s+\rho)\theta}\right\} \tag{8.1}$$

9. Incomplete Moments

We know that

$$I_q(x) = \int_0^s x^s f(x;\alpha,\beta,\theta)dx$$

$$I_q(x) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}(-1)^q(-1)^r\binom{-\left(\frac{\beta+1}{\beta}\right)}{p}\binom{(p+1)\alpha-1}{q}\binom{q-(p+1)\alpha-1}{r}\alpha\beta^p\theta\int_0^s x^s e^{-\theta x(s+1)}dx$$

$$I_q(x) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\varsigma_{pqs}\int_0^s x^s e^{-\theta x(s+1)}dx$$

After solving the above integral, we will get the following equation:

$$I_q(x) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\varsigma_{pqs}\frac{\gamma((s+1),\theta(s+1)s)}{(\theta(s+1))^{s+1}} \tag{9.1}$$

10. Mean Residual Function

We know that

$$m_r(x) = \frac{1}{S_x(x)}\int_x^{\infty} t\,f(t;\alpha,\beta,\theta)dt - x$$

$$m_r(x) = \frac{1}{S_x(x)}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}(-1)^q\binom{-\left(\frac{\beta+1}{\beta}\right)}{p}\binom{(p+1)\alpha-1}{q}\alpha\beta^p\int_x^{\infty} t\left(\theta e^{-\theta t}\right)\left(1-e^{-\theta t}\right)^{q-(p+1)\alpha-1}dt - x$$

After solving the above equation, we get

$$m_r(x) = \frac{1}{1-\left[1+\beta\left(e^{\theta x}-1\right)^{-\alpha}\right]^{\frac{-1}{\beta}-1}}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\varsigma_{pqs}\frac{\Gamma(2,\theta(s+1)x)}{\{\theta(s+1)\}^2} - x \tag{10.1}$$

11. Mean deviation from Mean

We know that

$$D(\mu) = E(|x-\mu|)$$

$$D(\mu) = 2\mu F(\mu) - 2\int_0^{\mu} xf(x)dx \tag{11.1}$$

We know that

$$\int_0^{\mu} xf(x)dx = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}(-1)^q(-1)^r\binom{-\left(\frac{\beta+1}{\beta}\right)}{p}\binom{(p+1)\alpha-1}{q}\binom{q-(p+1)\alpha-1}{r}\alpha\beta^p\theta\int_0^{\mu}xe^{-(s+1)\theta x}dx$$

$$\int_0^{\mu} xf(x)dx = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\varsigma_{pqs}\frac{\gamma(2,\theta(s+1)\mu)}{(\theta(s+1))^2} \tag{11.2}$$

Now, putting the above value in equation (11.1), we get the required equation as given below:

$$D(\mu) = 2\mu[1+\beta(e^{\mu}-1)^{-\alpha}]^{\frac{-1}{\beta}} - 2\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\varsigma_{pqs}\frac{\gamma(2,\theta(s+1)\mu)}{(\theta(s+1))^2} \tag{11.3}$$

12. Mean deviation from Median

We know that

$$D(M) = E(|x-M|)$$

$$D(M) = \mu - 2\int_0^{M} xf(x)dx \tag{12.1}$$

We have

$$\int_0^{M} xf(x)dx = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}(-1)^q(-1)^r\binom{-\left(\frac{\beta+1}{\beta}\right)}{p}\binom{(p+1)\alpha-1}{q}\binom{q-(p+1)\alpha-1}{r}\alpha\beta^p\theta\int_0^{M}e^{(s+1)\theta x}dx$$

$$\int_0^{M} xf(x)dx = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\varsigma_{pqs}\frac{(\gamma(2,\theta(s+1)M))}{(\theta(s+1))^2} \tag{12.2}$$

Now, putting the above value in equation (12.1), we get the required equation as given below:

$$D(M) = \mu - 2\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\varsigma_{pqs}\frac{(\gamma(2,\theta(s+1)M))}{(\theta(s+1))^2} \tag{12.3}$$

13. Order Statistics

Let $x_{(1)}, x_{(2)}, x_{(3)}, \ldots, x_{(n)}$ denote the order statistics of n random samples drawn from the extended odd Frechet-exponential distribution. Then, the pdf of $x_{(k)}$is given by

$$f_{x(k)}(x;\theta) = \frac{n!}{(k-1)!\,(n-k)!}f_X(x)[F_X(x)]^{k-1}[1-F_X(x)]^{n-k}$$

$$f_{x(k)}(x;\alpha,\beta,\theta) = \left[\begin{array}{c}\frac{n!}{(k-1)!\,(n-k)!}\alpha\theta e^{-\theta x}\left(e^{-\theta x}-1\right)^{-\alpha-1}\left[1+\beta\left(e^{-\theta x}-1\right)^{-\alpha}\right]^{\frac{-1}{\beta}-1} \\ *\left\{\left[1+\beta\left(e^{\theta x}-1\right)^{-\alpha}\right]^{\frac{-1}{\beta}}\right\}^{k-1}\left(1-\left[1+\beta\left(e^{\theta x}-1\right)^{-\alpha}\right]^{\frac{-1}{\beta}-1}\right)\end{array}\right]^{n-k}$$

Then, the pdf of first order $X_{(1)}$ of the extended odd Frechet-exponential distribution is given by

$$f_{x(1)}(x;\alpha,\beta,\theta) = n\alpha\theta e^{-\theta x}\left(e^{-\theta x}-1\right)^{-\alpha-1}\left[1+\beta\left(e^{-\theta x}-1\right)^{-\alpha}\right]^{\frac{-1}{\beta}-1}\left(1 - \left[1+\beta\left(e^{\theta x}-1\right)^{-\alpha}\right]^{\frac{-1}{\beta}-1}\right)^{n-1}$$

and the pdf of nth order $X_{(n)}$ of the extended odd Frechet-exponential distribution is given by

$$f_{x(n)}(x;\alpha,\beta,\theta) = n\alpha\theta e^{-\theta x}\left(e^{-\theta x}-1\right)^{-\alpha-1}\left[1 + \beta\left(e^{-\theta x}-1\right)^{-\alpha}\right]^{\frac{-1}{\beta}-1}\left\{\left[1+\beta\left(e^{\theta x}-1\right)^{-\alpha}\right]^{\frac{-1}{\beta}}\right\}^{n-1}$$

14. Maximum Likelihood Estimation

Let $x_{(1)}, x_{(2)}, x_{(3)}, \ldots, x_{(n)}$ be n random samples from the extended odd Frechet-exponential distribution, and then its likelihood function is given by

$$l = \prod_{i=1}^{n} f(x,\alpha,\beta,\theta)$$

$$l = (\alpha\theta)^n e^{\sum_{i=1}^{n}\theta x}\prod_{i=1}^{n}\left(e^{\theta x}-1\right)^{-\alpha-1}\left(1+\beta\left(e^{\theta x}-1\right)^{-\alpha}\right)^{-\frac{1}{\beta}-1}$$

Taking log on both sides we get

$$log\, l = n\, log\,\alpha + n\, log\,\theta + \sum_{i=1}^{n}\theta x_i - (\alpha+1)\sum_{i=1}^{n} log\left(e^{\theta x_i}-1\right) - \left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n} log\left(1+\beta\left(e^{\theta x_i}-1\right)^{-\alpha}\right)$$

Differentiating w.r.t α,β and θ the above equation we get

$$\frac{\partial\, log\, l}{\partial\alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} log\left(e^{\theta x_i}-1\right) + (\beta+1)\sum_{i=1}^{n}\frac{\left(e^{\theta x_i}-1\right)^{-\alpha}}{1+\beta(e^{\theta x_i}-1)^{-\alpha}} log\left(e^{\theta x_i}-1\right)$$

$$\frac{\partial\, log\, l}{\partial\beta} = -\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{\left(e^{\theta x_i}-1\right)^{-\alpha}}{1+\beta(e^{\theta x_i}-1)^{-\alpha}} + \left(\frac{1}{\beta^2}\right)\sum_{i-1}^{n} log\left(1+\beta\left(e^{\theta x_i}-1\right)^{-\alpha}\right)$$

$$\frac{\partial\, log\, l}{\partial\theta} = \frac{n}{\theta} + \sum_{i-1}^{n} x_i - (\alpha+1)\theta\sum_{i=1}^{n}\frac{e^{\theta x_i}}{e^{\theta x_i}-1} + \alpha(\beta+1)\theta\sum_{i=1}^{n}\frac{e^{\theta x_i}\left(e^{\theta x_i}-1\right)^{-\alpha-1}}{1+\beta(e^{\theta x_i}-1)^{-\alpha}}$$

The above mentioned equations are non-linear equations that cannot be expressed in compact form, and solving them directly for $\alpha, \beta$and θ is difficult. The MLE of the

parameters labelled as $\hat{\varsigma}(\hat{\alpha},\hat{\beta},\hat{\theta})$of $\varsigma(\alpha,\beta,\theta)$ may be derived by employing iterative methods such as Newton–Raphson method, secant method, regula–falsi method, and so on.

15. Application

In this segment, the efficacy of the newly developed distribution has been assessed using two realistic sets of data. As the new distribution is compared to new modified Weibull distribution (NMWD), extended Frechet distribution (EFD), Frechet distribution (FD), exponential distribution (ED), inverse Weibull Burr distribution (IWBD) and Kumarswamy power Frechet distribution (KPFD). It is revealed that the newly developed distribution offers an appropriate fit.

Various criteria including the AIC (Akaike information criterion), CAIC (Consistent Akaike information criterion), BIC (Bayesian information criterion), HQIC (Hannan-Quinn information criteria) and KS (Kolmogorov-Smirnov) are used to compare the fitted models. The p-value of each model is also recorded.

Data Set 1: The following observations contain 30 successive march precipitations (in inches). These observations were studied by Hankley 1977. The observations are as follows: 0.77, 1.74, 0.81, 1.20, 1.95, 1.20, 0.47, 1.43, 3.37, 2.20, 3.00, 3.09, 1.51, 2.10, 0.55, 1.62, 1.31, 0.32, 0.59, 0.81, 2.81, 1.87, 1.18, 1.35, 4.75, 2.48, 0.96, 1.89, 0.90, 2.05

Table 15.1: Descriptive statistics of data set 1

Min.	Q_1	Median	Mean	Q_3	Max.	S.D	Skew.	Kurt.
0.320	0.915	1.470	1.676	2.087	4.750	0.999	1.092	4.216

Table 15.2: The ML Estimates of the unknown parameters for data set 1

Model	Estimates				Standard Error			
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\lambda}$
EOFED	1.743	0.695	0.501	-----	0.729	0.806	0.117	----
NMWD	0.126	0.193	1.631	2.551	0.089	0.144	0.473	0.789
EFD	1.899	0.241	-----	------	0.384	0.256	-----	-----
FD	0.802	1.162	-----	-----	0.164	0.177	------	-----
ED	0.597	-----	-----	-----	0.108	-----	-----	-----
IWBD	0.314	1.813	1.000	-----	0.164	0.372	0.290	-----
KPFD	0.569	0.106	66.183	48.659	0.442	0.418	264.9	158.9

Table 15.3: Measures of goodness-of-fit statistics for the data set 1

Model	EOFED	NMWD	EFD	FD	ED	IWBD	KPFD
-2logl	76.501	76.789	82.398	108.26	90.985	84.303	76.044
AIC	80.501	84.789	86.398	112.26	92.985	83.003	84.033
CAIC	82.945	86.389	86.842	112.71	93.429	98.117	85.633
HQIC	81.397	86.583	87.294	113.16	95.881	96.336	85.827
BIC	86.704	90.394	89.200	115.07	94.386	89.001	89.638

Table 15.4: The K-S and p-value for data set 1

Model	EOFED	NMWD	EFD	FD	ED	IWBD	KPFD
K-S value	0.0680	0.0612	0.2102	0.2160	0.5159	0.1881	0.0608
p-value	0.9999	0.9991	0.1409	0.1216	2.319 e-07	0.2393	0.9998

Data Set 2: The data is obtained from Hinkley (1977) and consists of 30 successive values of March precipitation (in inches) in Minneapolis/St. Paul. The data sets are: 2.2, 3.37, 1.43, 0.74, 1.2, 1.95, 1.2, 0.81, 1.74, 0.77, 0.81, 0.59, 0.32, 1.31, 1.62, 0.52, 2.1, 1.51, 3.09, 3.00, 2.05, 0.9, 1.89, 0.96, 2.48, 4.75, 1.35, 1.18, 1.87, 2.81.

Table 15.5: Descriptive statistics of data set 2

Min.	Q_1	Median	Mean	Q_3	Max.	S.D	Skew.	Kurt.
0.320	0.915	1.470	1.684	2.087	4.750	0.9905	1.1260	4.2882

Table 15.6: The ML Estimates of the unknown parameters for data set 2

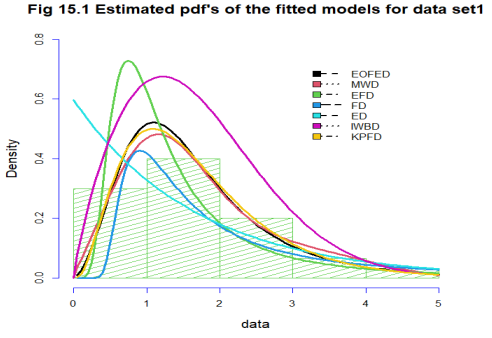
Model	Estimates				Standard Error			
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\lambda}$
EOFED	1.710	0.589	0.508	-----	0.640	0.682	0.107	----
NMWD	0.117	0.192	1.682	2.683	0.083	0.133	0.474	0.816
EFD	1.941	0.216	-----	------	0.385	0.235	-----	-----
FD	0.855	1.185	-----	-----	0.173	0.178	------	-----
ED	0.594	-----	------	------	0.108	-----	-----	-----
IWBD	0.305	1.843	1.000	-----	0.159	0.372	0.287	-----
KPFD	0.621	0.151	43.21	35.01	0.623	1.005	291.9	123.8

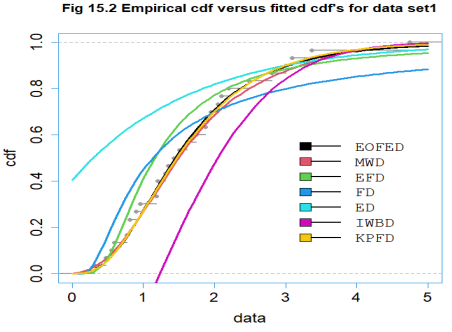
Table 15.7: Measures of goodness-of-fit statistics for the data set 2

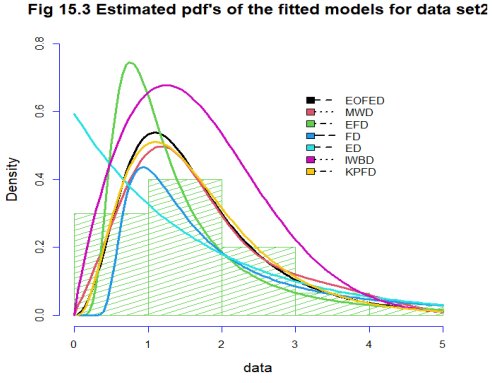
Model	EOFED	NMWD	EFD	FD	ED	IWBD	KPFD
-2logl	75.552	76.039	81.584	105.42	91.270	76.039	76.223
AIC	81.552	84.039	85.584	109.43	93.270	84.040	83.207
CAIC	81.996	85.639	86.028	109.87	93.715	85.639	84.807
HQIC	80.448	85.833	86.480	110.32	96.167	85.833	84.999
BIC	85.755	89.645	88.386	112.23	94.672	89.645	88.812

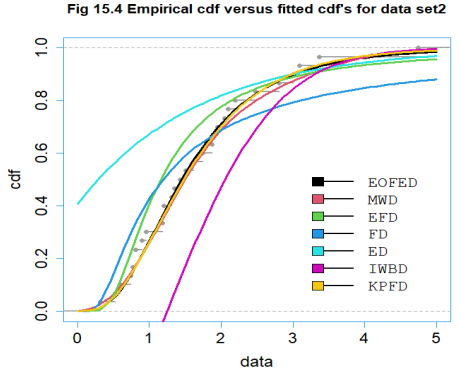
Table 15.8: The K-S and p-value for data set 2

Model	EOFED	NMWD	EFD	FD	ED	IWBD	KPFD
K-S value	0.0696	0.0612	0.2097	0.1953	0.5307	0.2817	0.0640
p-value	0.9999	0.9986	0.1423	0.2025	9.2e-08	0.01713	0.9997









Figures (15.1), (15.2), (15.3) and (15.4) represent the estimated densities and cdfs of the fitted distributions to data set 1 and 2.

16. Simulation Study

In this section, we study the performance of ML estimators for different sample sizes (n=150, 250,500, 800). We have employed the inverse CDF technique for data simulation for EOFED distribution using R software. The process was repeated 1000 times for calculation of bias, variance and MSE. It is evident from the tables that a decreasing trend is being observed in bias, variance and MSE as we increase the sample size. Hence, the performance of ML estimators is quite well, consistent in the case of the extended odd Frechet-exponential distribution.

Table 16.1: sThe Bias, Variance and MSEs of 1,000 simulations of EOFED for parameter values $\alpha = 1.0\ \beta = 0.9$ and $\theta = 2.3$

Sample Size	Parameters	Bias	Variance	MSE
150	α	-0.064	7.41e-05	0.0212
	β	-0.125	8.11e-05	0.0157
	θ	-0.216	7.07e-03	0.0531
250	α	-0.062	7.23e-05	0.0210
	β	-0.123	6.81e-05	0.0156
	θ	-0.213	5.24e-05	0.0497
500	α	-0.061	5.43e-05	0.0200
	β	-0.122	3.11e-05	0.0152
	θ	-0.209	3.15e-05	0.0470
800	α	-0.059	4.22e-05	0.0199
	β	-0.121	2.98e-05	0.0149
	θ	-0.209	1.33e-03	0.0453

Table 16.2: The Bias, Variance and MSEs of 1,000 simulations of EOFED for parameter values $\alpha = 1.1\ \beta = 0.8$ and $\theta = 2.0$

Sample Size	Parameters	Bias	Variance	MSE
150	α	-0.156	6.88e-o5	0.0266
	β	-0.029	6.04e-05	0.0009
	θ	-0.199	6.59e-03	0.0459
250	α	-0.155	5.29e-05	0.0253
	β	-0.027	5.05e-05	0.0008
	θ	-0.191	5.33e-03	0.0399
500	α	-0.149	2.38e-05	0.0222
	β	-0.026	2.70e-05	0.0007
	θ	-0.187	2.74e-03	0.0376
800	α	-0.148	2.09e-05	0.0219
	β	-0.026	1.89e-05	0.0007
	θ	-0.185	1.23e-03	0.0352

17. Result and Discussion

In this paper, a new distribution named "extended odd Frechet-exponential distribution", which is obtained by T-X method, is introduced. Several mathematical quantities for the newly developed distribution are derived. The method of maximum likelihood estimation is used to estimate the unknown parameters of the established model. It is obvious from Tables 15.3, 15.4, 15.7 and 15.8 that the extended odd Frechet-exponential distribution has smaller values for AIC, AICC, BIC, HQIC and K-S statistics as well as higher p-value when compared with other competitive models. Accordingly, we arrive at the conclusion that the extended odd Frechet-exponential distribution provides a more adequate fit than the compared ones.

References

Ahmed, T. F., Dina, A. R., Eldesouky, B. S., (2023). Statistical inference of modified Frechet-exponential distribution with applications to real life data. *Applied mathematics and Information sciences, an international journal*. Vol. 17, No. 1, pp. 109-124.

Aisha, F., Tahir, M. H., Algarni, A., Imran, M., Jamal, F., (2022). A New Useful Exponential Model with Applications to Quality Control and Actuarial Data. Hindawi, *Computational Intelligence and Neuroscience*, Vol. 2022, Article ID 2489998,1-27

Ali, M., Khalil, A., Ijaz, M., Saeed, N., (2021). Alpha-Power Exponentiated Inverse Rayleigh Distribution and its applications to real and simulated data. PLoS ONE 16(1), e0245253.

Ali, M., Khalil, A., Mashwani, W. K., Alrajhi, S., Al-marzouki, S., Shah, K., (2022). A novel Frechet-type probability distribution: its properties and applications. Hindawi, *Mathematical problems in engineering*. Vol. 2022, article ID 2537332, pp. 1-18.

Alsadat, N., Ahmad, A., Jallal, M., Gemeay, A. M., (2023). The novel Kumaraswamy power Frechet distribution with data analysis related to diverse scientific areas. *Alexandria engineering journal*, Vol. 70, pp. 651-644.

Alshanbari, H. M., Gemeay, A. M., El-bagoury, A. A. H., Khosa, S. K., Hafez, E. H., Muse, A. H., (2022). A novel extension of Frechet distribution:application on real data and simulation. *Alexandria engineering journal*, Vol. 61, issue 10, pp. 7917-7938.

Alzaatreh, A., Lee, C., Famoye, F., (2013). A new method for generating families of distributions. *Metron*, 71, pp. 63–79.

Fisher, R. A., Tippett, L. H. C., (1928). Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Proc Cambridge Philosophical Society*, 24(2), pp. 180–190.

Frechet, M., (1927). Sur la loi de probabilite de lecart maximum. *Ann. Soc. Polon. de Math.*, Cracovie, 6, pp. 93–116.

Gumbel, E. J., (1958). Statistics of extremes. New York: *Columbia university press*. OCLC 180577.

Hamed, M. S., (2020). Extended Poisson-Frechet distribution: Mathematical properties and applications to survival and repair times. *Journal of Data Science*, 18(2), pp. 319–342.

Haq, M. A., Hashmi, S., Yousuf, H. M., (2017). A new five-parameter Frechet model for extreme values. *Pakistan journal of Statistics and Operation Research*, Vol(3), pp. 617–632

Klakattawi, H. S., Khormi, A. A., Baharith, L. A., (2023). The new generalized exponentiated Frechet-Weibull distribution: properties, applications and regression model. Hindawi(Wiley) , *Complexity*, Vol. 2023, article ID 2196572, pp. 1–23.

Klakattawi, H. S., Alsulami, D., Elaal, M. A., Dey, S., Baharith, L., (2022). A new generalized family of distribution based on combining marshal-olkin transformation with TX family. *Plos One*, Vol. 17, No. 2, Article ID e0263673.

Ocloo, S. K., Brew, L., Nasiru, S., Odoi, B., (2022). Harmonic mixture Frechet distribution: properties and applications to lifetime data. Hindawi, *International journal of mathematics and mathematical sciences*, Vol. 2022, article ID 6460362, pp. 1–20.

Oguntunde, P. E., Khaleel, M. A., Ahmed, M. T., Okagbue, H. I., (2019). The Gompertz-Frechet distribution, properties and applications. *Congent Mathematics and Statistic*, 6, 1, 1568662.

Penson, K. A., Gorska, K., (2014). On the Laplace transform of the Frechet distribution. Journal of Mathematical Physics, 55,093501.

Rosin, P., Rammler, E., (1933). The laws governing the Fineness of Powdered Coal. *Journal of the Institute of Fuel,* 7, pp. 29–36.

Tomitaka, S., Kawasaki, Y., Ide, K., Akutagawa, M., Yamada, H., Furukawa, T. A., (2017). Exponential distribution of total depressive symptom scores in relation to exponential latent trait and item threshold distributions: a simulation study. BMC Research Notes, 10(1), doi:10.1186/s 13104-017-2937-6.

Zubir, S., Ali, M., Hamraz, M., Khan, D. M., Khan, Z., EL-Morshedy, M., Al-Bossly, A., Almaspoor, Z., (2022). A new member of T-X Family with applications in different sectors. Hindawi. *Journal of Mathematics*, Vol. 2022, Article ID, 1453451, pp. 1–15.

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 137–165, https://doi.org/10.59170/stattrans-2024-019
Received – 13.02.2023; accepted – 20.03.2024

Spatial and component structure analysis of the inclusive circular economy: SGICE

Oleksandr Osaulenko[1], Alla Shlapak[2], Iryna Zvarych[3], Oksana Brodovska[4], Kateryna Krysovata[5]

Abstract

Modern statistics and scientific research consider the circular economy rather one-sidedly, essentially as recycling and reuse of resources. The article proposes methodology (concept) of the global inclusive circular economy, which can be considered as a complex multidimensional system, the main components of which are the economic, sociological, ecological and circular aspects of the country's life. To achieve this goal, the GNU regression, econometrics and time-series library was used – an applied software package for econometric modeling, a part of the GNU project. Accordingly, we will define the global inclusive circular economy as the SGICE (Global Inclusive Circular Economy) system, characterized by the vector of functions. For the most exhaustive consideration of the entire range of opportunities of the global inclusive circular economy, the study developed and accordingly analyzed the integrated index of the development of the global inclusive circular economy (IGICE) by ecological, economic, social and circular components with isolated weakly correlated indicators. The ranking of countries was carried out separately by the components of the index, on the basis of which circular cores were distinguished: the social component (Belgium, the Czech Republic, USA, China, France, Greece, Austria, Australia), the ecological component (Japan, Denmark), the economic component (Germany, China), which became the basis for the model of formation of global inclusive circular chains. The conducted cluster analysis based on the component indicators of the index of the global inclusive circular economy confirmed the formation of a large circular gap (gap) in the ecological and circular components. This indicator is introduced for the first time and makes it possible to comprehensively analyze the country and highlight additional effects that arise from the moment of implementation

[1] National Academy of Statistics, Accounting and Audit, Kiev, Ukraine. E-mail: O.Osaulenko@nasoa.edu.ua. ORCID: http://orcid.org/0000-0002-7100-7176.

[2] Department of International Economics, Kyiv National Economic University named after Vadym Hetman, Ukraine. E-mail: allashlapak@gmail.com. ORCID: http://orcid.org/0000-0001-8697-7039.

[3] International Economics Department, West Ukrainian National University, Ukraine. E-mail: irazvarych@gmail.com. ORCID: http://orcid.org/0000-0001-5155-540X.

[4] Limited Liability Company Financial Company MAGNAT, West Ukrainian National University, Ukraine. E-mail: ksenyabrodovska@gmail.com. ORCID: https://orcid.org/0000-0003-4828-528X.

[5] Taxes and Fiscal Policy Department, West Ukrainian National University, Ukraine. E-mail: k.krysovata@wunu.edu.ua. ORCID: https://orcid.org/0000-0003-1147-8811.

of circular business projects and the inclusion of the country (on the basis of firms, corporations) in global circular chains of added value. After all, the main problematic aspects are the illegal trade in waste and growing smuggling, which cause serious negative social consequences and actualize the inclusive component in the justification of the global inclusive circular economy paradigm. Waste reduction combined with wise use of resources has the potential to address the gap resulting from the scarcity of natural resources and a growing global population or consumption. The formation of circular trade will contribute to: the determination of priority materials for trade and the required level of processing capacity, coordination of material quality standards, promoting demand for used goods and secondary raw materials, removing unnecessary regulatory barriers and avoiding environmentally harmful activities such as non-compliance, poor regulation and informal recovery.

Key words: Global Inclusive Circular Economy, circular gap, structural components.

JEL Classification: A20, F02, F23.

1. Introduction

The prerequisites for the formation of a circular economy as a business model are resources and their pricing, the growth of middle-class consumers, «big data», changes in legislation, globalization of management, the transition from «agreement» to «relationships». To ensure the effective functioning of this system, it is necessary to take into account the following methods of effective functioning of circular inclusive economies: inclusiveness in the informal sector, the role of women in society and the economy, the active position of citizens, authorized ecosystem players, political interventions.

The ways of implementing the circular economy are: encouraging innovation, «smart» regulation, promoting goods and services with a shorter life cycle, preventing resource shortages, defining circular policy goals for all economic entities, monitoring the use of land, materials, water, and emissions greenhouse gases. At the same time, measures to introduce industry into the circular process are defined as: fines, cleaning of the territory, coordination of activities between monitoring bodies, liquidation of illegal and unmanaged landfills, environmental monitoring based on the EU methods, alternatives to plastic, inter-sectoral agreements to reduce emissions, investments in circular solutions, cross-border cooperation regarding disposal.

2. Methodology

According to the proposed concept of the global inclusive circular economy, it can be considered as a complex multidimensional system, the main components of which are the economic, sociological, ecological and circular aspects of the country's life. To achieve the goal, the GNU Regression, Econometrics and Time-series Library was used – an applied software package for econometric modeling, part of the GNU project.

Accordingly, we will define the global inclusive circular economy as the S_{GICE} (Global Inclusive Circular Economy) system, characterized by the vector of functions:

$$\overrightarrow{S_{GICE}} = \left(y(\vec{x}_1), y(\vec{x}_2), y(\vec{x}_3), y(\vec{x}_4)\right), \quad (1)$$

where $\vec{x}_1$– vector of indicators (indicators) describing the economic component, $\vec{x}_2$ – vector of social component indicators, $\vec{x}_3$ – vector of environmental component indicators, $\vec{x}_4$ —vector of circular component system indicators.

Let us consider a set of indicators for each component of the global inclusive circular economy.

Economic component: $\vec{x}_1 = \{x_{1i}\}, i = 1, \dots, 6$

where x_{11} – GDP, million $;

x_{12}– employment, %;

x_{13} – real GDP per capita, $;

x_{14} – GDP growth, %;

x_{15} – taxes related to the environment, % of total tax revenues;

x_{16}.– subsidies for the development of eco-technologies related to the environment, % of total assistance.

Social component: $\vec{x}_2 = \{x_{2i}\}, i = 1, \dots, 5$

where x_{21} – welfare costs from premature death from exposure to lead, equivalent to GDP;

x_{22}– population with access to improved sanitation, % of the total population;

x_{23} – population with access to purified spring drinking water, % of the total population;

x_{24} – population connected to water supply networks, % of the total population;

x_{25} – population connected to water supply networks with preventive disinfection, % of the total population.

Ecological component: $\vec{x}_3 = \{x_{3i}\}, i = 1, \dots, 10$

where x_{31} – greenhouse gas emissions, thousand tons of CO2 equivalent; x_{32}– land use, km^2;

x_{33} – renewable energy, %;

x_{34} – internal consumption of material resources, million tons;

x_{35} – state budget for R&D related to the environment;

x_{36} – ecologically adjusted multifactor productivity growth;

x_{37}– municipal waste is processed or composted, % of processed waste;

x_{38}– productivity of CO2, $/kg;

x_{39} – adjustment to reduce pollution, %;

x_{310} – development of environmental technologies, % of all technologies.

Circulation component: $\vec{x}_4 = x_{4i}, i = 1, \dots, 4$

where x_{41} – percentage of garbage recovery, %;
x_{42} – percentage of recycling garbage, %;
x_{43} – treated household waste, million tons;
x_{44} – recycling garbage, million tons.

To build an integrated indicator of the global inclusive circular economy, we will apply one of the most well-known methods of artificial intelligence - the method of principal component analysis (PCA). The main advantages of this method in our case are the possibility of taking into account a large number of performance indicators of various components of the global inclusive circular economy and bringing them to the main or main component, which will reflect the system as a whole with «maximum» reliability.

The main task of the method of principal components is to replace the original data with some aggregated values in a new space, while solving two tasks - the first of which consists in combining the most important (from the point of view of minimizing the mean square error) values into a smaller number of parameters, but the most informative (reducing the dimensionality of the data space), and the second is to reduce the noise in the data. To solve this problem, PCA looks for a space that best represents the variance of the data. The direction with the largest predicted variance is called the first principal component. The orthogonal direction that captures the second largest predicted variance is called the second principal component, and so on. Note that the direction, maximizing the variance, minimizes the root mean square error.

Finding the main components is reduced to the calculation of eigenvectors and eigenvalues of the covariance matrix of the initial data. Sometimes the method of principal components is called the Karhunen-Loev transformation or the Hotelling transformation.

So, as an indicator, we will use a generalized indicator:

$$P_{GICE} = \frac{1}{N}\sum_{i=1}^{N} \lambda_{y(\vec{x}_i)}, \quad (2)$$

where $\lambda_{y(\vec{x}_i)}$ – s the eigenvalue of the correlation matrix, which characterizes the absolute contribution (significance) of the corresponding main component for the function $y(\vec{x}_i)$to the total variance of the vector of indicators of the system $\vec{x}_i$, N – is the dimension of the vector $\vec{S}$, in our case $N = 4$.

The function $y(\vec{x}_i)$ takes the following form:

$$y(\vec{x}_i) = \sum_{j=1}^{p} w_{kj} \cdot x_{ij}, \quad (3)$$

where w_{kj} – are the elements of the eigenvector of the correlation matrix for the *κ-th* principal component, p – the dimension of the vector $\vec{x}_i$.

For ease of operation and presentation, the integral indicator can be reduced to a scale from 0 to 1. For this, we will use the following formula:

$$Indeks\left(P_{GICE}^{k}\right) = \frac{P_{GICE}^{k} - \min\left(P_{GICE}\right)}{\max(P_{GICE}) - \min(P_{GICE})}, \quad (4)$$

where P_{GICE}^{k} – the indicator of the k-th country, $\max\left(P_{GICE}\right)$, $\min\left(P_{GICE}\right)$ – the maximum and minimum value of the indicator among the studied countries, respectively.

The set of indicators of the global inclusive circular economy for the social component will look like this:

$$\vec{x}_2 = \{x_{2i}\}, i = 1, \dots, 5$$

where x_{21} welfare costs from premature death from exposure to lead, equivalent to GDP;

x_{22} – population with access to improved sanitation, % of the total population;

x_{23} – population with access to purified spring drinking water, % of the total population;

x_{24} – population connected to water supply networks, % of the total population;

x_{25} – population connected to water supply networks with preventive disinfection, % of the total population.

To rank countries according to a set of indicators of the social component, we will use a partial case for the integrated indicator N =1, accordingly we will get:

$$P_S^i = \lambda_{y(\vec{x}_2^i)}$$

where $\lambda_{y(\vec{x}_2^i)}$– the eigenvalue of the correlation matrix, which characterizes the absolute contribution (significance) of the corresponding main component for the function $y\left(\vec{x}_2^i\right)$ to the total variance of the indicators of the social component system for the i-th country.

The regression function $y(\vec{x}_\text{i})$ takes the following general form:

$$y(\vec{x}_2) = \sum_{j=1}^{5} w_{1j} \cdot x_{2j},$$

where w_{1j} – elements of the eigenvector of the correlation matrix for the first principal component. In the course of calculations, some indicators from the set $\vec{x}_2$may turn out to be uncorrelated and not included in the general regression equation.

To rank countries according to a set of indicators of the economic component, we will use the partial case for the integral indicator N =1, accordingly we will get:

$$P_E^i = \lambda_{y(\vec{x}_1^i)}$$

where $\lambda_{y(\vec{x}_1^i)}$– the eigenvalue of the correlation matrix, which characterizes the absolute contribution (significance) of the corresponding main component for the function $y\left(\vec{x}_1^i\right)$ to the total variance of the indicators of the economic component system for the i-th country.

The function $y(\vec{x}_1)$ takes the following form:

$$y(\vec{x}_1) = \sum_{j=1}^{6} w_{1j} \cdot x_{1j},$$

where w_{1j} – the elements of the eigenvector of the correlation matrix for the first principal component. In the course of calculations, some indicators from the set $\vec{x}_1$ may turn out to be uncorrelated and not included in the general regression equation.

3. Empirical results and discussion

According to the proposed concept of the global inclusive circular economy, it can be considered as a complex multidimensional system, the main components of which are the economic, sociological, ecological and circular aspects of the country's life. To fulfill the task, the indicators of the relevant countries and statistics of the Organization for Economic Cooperation and Development for 28 member countries of the organization from 1995 to 2020 were used: Australia, Austria, Belgium, Great Britain, Greece, Denmark, Estonia, Israel, Canada, China, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, Germany, New Zealand, South Africa, South Korea, Poland, USA, Turkey, Hungary, Ukraine, Finland, France, Czech Republic, Japan. The analysis was carried out on the basis of the GNU Regression, Econometrics and Time-series Library (Library for regressions, econometrics and time series) - an applied software package for econometric modeling, part of the GNU project.

Thus, Table 1 shows the results of the economic component of the indicator, in particular, the regression functions and uncorrelated indicators are indicated, respectively, for each country under study. Therefore, on the basis of the analysis carried out for Australia, the proper value of the correlation matrix (i.e. the characteristic number at which a solution can be obtained to describe the global inclusive circular economy), which characterizes the absolute contribution (significance) of the corresponding main component, is equal to λ_y(x⃗_1) =3,6582 and describes the solution of the given problem for 73% of all processes, which indicates its sufficient adequacy. (see the appendix).

Table 1: Results for the economic component of the indicator

Country	Regression function	% comp.	Weakly correlated indicators
Australia	$y(\vec{x}_1) = 0{,}485 \cdot x_{11} + 0{,}475 \cdot x_{12} + 0{,}5 \cdot x_{13} - 0{,}454 \cdot x_{15} + 0{,}284 \cdot x_{16},$	73%	x_{14}
Austria	$y(\vec{x}_1) = -0{,}5 \cdot x_{11} - 0{,}364 \cdot x_{12} - 0{,}493 \cdot x_{13} - 0{,}494 \cdot x_{15} + 0{,}36 \cdot x_{16},$	76%	
Belgium	$y(\vec{x}_1) = 0{,}483 \cdot x_{11} + 0{,}473 \cdot x_{12} + 0{,}5 \cdot x_{13} + 0{,}362 \cdot x_{14} + 0{,}4 \cdot x_{15},$	76%	x_{16}
Greece	$y(\vec{x}_1) = 0{,}494 \cdot x_{11} + 0{,}483 \cdot x_{12} + 0{,}382 \cdot x_{13} + 0{,}437 \cdot x_{14} + 0{,}431 \cdot x_{15}$	73%	x_{16}

Table 1: Results for the economic component of the indicator (cont.)

Country	Regression function	% comp.	Weakly correlated indicators
Denmark	$y(\vec{x}_1) = 0{,}508 \cdot x_{11} + 0{,}504 \cdot x_{12} + 0{,}499 \cdot x_{13} + 0{,}177 \cdot x_{15} + 0{,}455 \cdot x_{16}$	76%	x_{14}
Great Britain	$y(\vec{x}_1) = 0{,}517 \cdot x_{11} + 0{,}52 \cdot x_{12} + 0{,}521 \cdot x_{13} + 0{,}437 \cdot x_{16}$	89%	x_{14}, x_{15}
Estonia	$y(\vec{x}_1) = 0{,}513 \cdot x_{11} + 0{,}467 \cdot x_{12} + 0{,}514 \cdot x_{13} + 0{,}316 \cdot x_{14} + 0{,}393 \cdot x_{16}$	72%	x_{15}
Israel	$y(\vec{x}_1) = -0{,}587 \cdot x_{11} - 0{,}52 \cdot x_{12} - 0{,}52 \cdot x_{13} - 0{,}276x_{14} + 0{,}339x_{15}$	67%	x_{16}
Canada	$y(\vec{x}_1) = -0{,}538 \cdot x_{11} - 0{,}551 \cdot x_{13} + 0{,}558 \cdot x_{15} + 0{,}307 \cdot x_{16}$	97%	x_{12}, x_{14}
China	$y(\vec{x}_1) = -0{,}452 \cdot x_{11} + 0{,}443 \cdot x_{12} - 0{,}462 \cdot x_{13} + 0{,}455 \cdot x_{14} + 0{,}423 \cdot x_{15}$	92%	x_{16}
Latvia	$y(\vec{x}_1) = 0{,}482 \cdot x_{11} + 0{,}482 \cdot x_{12} + 0{,}483 \cdot x_{13} + 0{,}16 \cdot x_{14} + 0{,}327 \cdot x_{15} + 0{,}411 \cdot x_{16},$	66%	
Lithuania	$y(\vec{x}_1) = -0{,}516 \cdot x_{11} + 0{,}514 \cdot x_{12} + 0{,}520 \cdot x_{13} - 0{,}428 \cdot x_{15} + 0{,}122 \cdot x_{16},$	72%	x_{14}
Luxembourg	$y(\vec{x}_1) = -0{,}476 \cdot x_{11} - 0{,}4 \cdot x_{12} - 0{,}477 \cdot x_{13} - 0{,}054 \cdot x_{14} + 0{,}459 \cdot x_{15} + 0{,}416 \cdot x_{16}$	72%	
Mexico	$y(\vec{x}_1) = 0{,}493 \cdot x_{11} + 0{,}461 \cdot x_{12} + 0{,}499 \cdot x_{13} + 0{,}483 \cdot x_{15} + 0{,}248 \cdot x_{16},$	77%	x_{14}
Netherlands	$y(\vec{x}_1) = 0{,}435 \cdot x_{11} + 0{,}444 \cdot x_{12} + 0{,}432 \cdot x_{13} + 0{,}467 \cdot x_{14} + 0{,}457 \cdot x_{15},$	88%	x_{16}
Germany	$y(\vec{x}_1) = -0{,}446 \cdot x_{11} - 0{,}423 \cdot x_{12} - 0{,}442 \cdot x_{13} - 0{,}299 \cdot x_{14} + 0{,}344 \cdot x_{15} + 0{,}38 \cdot x_{16},$	83%	
New Zealand	$y(\vec{x}_1) = -0{,}462 \cdot x_{11} - 0{,}499 \cdot x_{12} - 0{,}425 \cdot x_{13} - -0{,}427 \cdot x_{15} + 0{,}417 \cdot x_{16},$	75%	x_{14}
South Africa	$y(\vec{x}_1) = -0{,}556 \cdot x_{11} + 0{,}132 \cdot x_{13} + 0{,}576 \cdot x_{14} + 0{,}585 \cdot x_{15}$	67%	x_{12}, x_{16}
South Korea	$y(\vec{x}_1) = 0{,}516 \cdot x_{11} + 0{,}524 \cdot x_{12} + 0{,}521 \cdot x_{13} + 0{,}433 \cdot x_{15},$	87%	x_{14}, x_{16}
Poland	$y(\vec{x}_1) = 0{,}453 \cdot x_{11} + 0{,}455 \cdot x_{12} + 0{,}451 \cdot x_{13} + 0{,}435 \cdot x_{14} - 0{,}438 \cdot x_{15},$	79%	x_{16}
Turkey	$y(\vec{x}_1) = -0{,}536 \cdot x_{11} - 0{,}529 \cdot x_{12} - 0{,}531 \cdot x_{13} + 0{,}389 \cdot x_{15},$	85%	x_{14}, x_{16}
Hungary	$y(\vec{x}_1) = -0{,}468 \cdot x_{11} - 0{,}474 \cdot x_{12} - 0{,}466 \cdot x_{13} - 0{,}474 \cdot x_{14} + 0{,}34 \cdot x_{15}$	76%	x_{16}
Finland	$y(\vec{x}_1) = 0{,}536 \cdot x_{11} + 0{,}492 \cdot x_{13} + 0{,}478 \cdot x_1 + 0{,}379 \cdot x_{15} + 0{,}315 \cdot x_{16},$	62%	x_{12}
France	$y(\vec{x}_1) = -0{,}522 \cdot x_{11} + 0{,}112 \cdot x_{12} - 0{,}514 \cdot x_{13} - 0{,}497 \cdot x_{14} - 0{,}014 \cdot x_{15} + 0{,}451 \cdot x_{16}$	79%	
Czech Republic	$y(\vec{x}_1) = -0{,}433 \cdot x_{11} - 0{,}435 \cdot x_{12} - 0{,}436 \cdot x_{13} - 0{,}364 \cdot x_{14} + 0{,}436 \cdot x_{15} + 0{,}334 \cdot x_{16}$	80%	
Japan	$y(\vec{x}_1) = -0{,}489 \cdot x_{11} - 0{,}503 \cdot x_{12} - 0{,}476 \cdot x_{13} + 0{,}205 \cdot x_{14} + 0{,}49 \cdot x_{15}$	76%	x_{16}

Source: constructed by author.

x_{11} – GDP, million \$;
x_{12} – employment, %;
x_{13} – real GDP per capita, \$;

x_{14} – GDP growth, %;
x_{15} – taxes related to the environment, % of total tax revenues;
x_{16}.– subsidies for the development of eco-technologies related to the environ ment, % of total assistance.

Similarly, research was conducted for all proposed countries (Table 1).

Weakly correlated indicators for the economic component turned out to be the following:

- *GDP growth, %:*
  Australia, Denmark, Great Britain, Canada, Lithuania, Mexico, New Zealand, South Korea and Turkey.
- *taxes related to the environment, % of total tax revenues:*
  Great Britain, Estonia.
- *subsidies for the development of eco-technologies related to the environment, % of total assistance:*
  Belgium, Greece, Israel, China, Netherlands, South Africa, South Korea, Poland, Turkey, Hungary, Japan.
- *employment, %:*
  Canada, Finland.

The decrease in the influence of the economic component in the overall integrated indicator of the global inclusive circular economy in the following countries is caused by the growth of the following indicators.

This means that an increase in the weight of the displayed indicators affects the main component among the set of indicators of the economic component, which generally frees up space for other components in the integrated indicators of the global inclusive circular economy, for example, for circular, ecological or social (or inclusive).

Table 2: Indicators that lead to a decrease in the economic component of GICE

Countries	Indicators
Australia	taxes related to the environment, % of total tax revenues
Austria	GDP, million $; employment, %; taxes related to the environment, % of total tax revenues
Israel	GDP, million $; employment, %; GDP growth, %.
Canada	GDP, million $; real GDP per capita, $.
China	GDP, million $;

Table 2: Indicators that lead to a decrease in the economic component of GICE (cont.)

Countries	Indicators
Lithuania	GDP, million $; taxes related to the environment, % of total tax revenues
Luxembourg	GDP, million $; employment, %; real GDP per capita, $.
Germany	GDP, million $; employment, %;
New Zealand	GDP, million $; employment, %;
	taxes related to the environment, % of total tax revenues
Turkey	GDP, million $; employment, %;
Hungary	GDP, million $; employment, %; GDP growth, %;
France	real GDP per capita, $; taxes related to the environment, % of total tax revenues
Czech Republic	GDP, million $; employment, %; real GDP per capita, $.
Japan	GDP, million $; employment, %; real GDP per capita, $

Source: constructed by author.

The table shows the rating of the countries of the world for which research was conducted, according to the economic component of the global inclusive circular economy.

Table 3: Rating of the countries of the world according to the economic component of the global inclusive circular economy

Attitude	Country	Indicator, $\lambda_{y(\vec{x}_1)}$	Normalized index
1	Belgium	4.8185	**1**
2	Czech Republic	4.76	**0.9807**
3	USA	4.7321	**0.9714**
4	China	4.6212	**0.9348**
5	France	4.6083	**0.9305**
6	Greece	4.5131	**0.8991**

Table 3: Rating of the countries of the world according to the economic component of the global inclusive circular economy (cont.)

Attitude	Country	Indicator, $\lambda_{y(\vec{x}_1)}$	Normalized index
7	Austria	4.4996	**0.8946**
8	Australia	4.38	**0.8551**
9	Lithuania	3.9858	0.7248
10	South Korea	3.9389	0.7093
11	Israel	3.9173	0.7022
12	Mexico	3.8922	0.6939
13	Estonia	3.8876	0.6924
14	United Kingdom	3.8749	0.6882
15	Poland	3.8676	0.6857
16	Luxembourg	3.8424	0.6774
17	Latvia	3.7489	0.6465
18	Hungary	3.7426	0.6444
19	Netherlands	3.6259	0.6059
20	Denmark	3.5104	0.5677
21	Japan	3.4447	0.5460
22	Canada	3.319	0.5044
23	Germany	3.2441	0.4797
24	Finland	3.1894	0.4616
25	Ukraine	**3.137**	**0.4443**
26	New Zealand	3.002	0.3997
27	Turkey	2.2761	0.1598
28	South Africa	1.7926	0

Source: constructed by author.

Germany and China are the absolute leaders in the group of the economic component of the indicator, according to the analysis of these countries, GDP and employment in Germany and GDP in China determine this component (Germany is the strongest economically in Europe, and one of the leading in the world, and China – the leader in world exports and imports). The lion's share of production and, accordingly, employment is occupied by fossil fuels and, accordingly, global supply chains.

The transformation of supply chains into global multi-stage production networks has occurred in a benign environment of falling trade barriers and an implicit willingness to accept growing interdependence and associated risks. But over the last decade, we have had a number of events caused by «black swans». Although such cases should be quite rare, there have been several in the world over the past decade: China's introduction of export quotas for critical resources in 2010; earthquake and tsunami in 2011; the flood in Thailand later that year; the US-China trade war. After the 2011 episodes, some companies made adjustments and created alternative sources, particularly for semiconductors coming from Japan's Naka region. Even in the wake of the recent trade war, many companies have reverted to quota status, believing it will be nearly impossible to replace their key suppliers in China.

Relying on suppliers with unique capabilities causes serious problems, consider semiconductors as an example. Taiwan represents about 22% of the world's semiconductor integrated circuit manufacturing capacity and more than half of its foundry capacity. One company, Taiwan Semiconductor Manufacturing Company (TSMC), accounts for about 67% of Taiwan's capacity and by far the largest share of the global market for the most advanced chip manufacturing processes. Companies like Apple and Qualcomm are completely dependent on this company for their most advanced chips. Accordingly, TSMC has geographically diversified its capacity across three science parks on the island, but the company as a whole still depends on a single Dutch supplier for its advanced lithography systems. This supplier, ASML, in turn depends on a single plant in Germany for its optical engine. Developing an alternative source of supply goes beyond chip design companies. Addressing this critical dependency is actually one of the pillars of the Chinese government's Made in China 2025 initiative (underscoring the scale of the challenge).

The decentralized, informal economy makes up a significant portion of the global economy, employing an estimated 61% of all workers. In fact, including agriculture, the informal sector accounts for more than 90% of total employment in a number of African and Asian countries. Despite the low cost of labor characteristic of the informal sector, they can cope with highly efficient sorting and reverse logistics techniques at the «last mile» or the last point of the value chain, involve poor working conditions, insufficient income, and as a result problems of related to health. However, this decentralized, distributed nature of the informal economy modifies the economic structure into a very organic and flexible organism. Therefore, the driver of social justice is the requirement to improve the qualifications and organization of informal workers, ensuring that their working conditions will improve. Education is also a necessary element to facilitate some processes and tasks in the informal economy, which are a kind of tools that will help to implement and prolong the concept of circularity and inclusiveness. Education is an absolutely essential element for young entrepreneurs, designers and

engineers on the technical characteristics of circularity and employment (i.e. job vacancies) among the informal sector and communities in recycling waste into new products. In this way, we are talking about filling newly created niches in the labor market, creating and providing circular workplaces of the future.

This is especially true for those entrepreneurs who solve visible environmental problems and provide jobs for as many people as possible. Such interests are important for continuous growth in the areas of social and sustainable entrepreneurship. Coupled with this, digital technologies across the informal economy can also help secure supply and redefine business models by promoting consumer confidence. Digital access and engagement can not only stimulate markets for new products, but also services that do not use assets, secondary materials and human capital.

Therefore, the cooperation of local authorities and the provision of adequate institutional representation of the informal sector in the development of local and national policies through associations, chambers of commerce and unions are key development paths to achieve this, as well as improving working conditions, which are key aspects in the formation of the social and inclusive component of the GICE indicator.

The set of global inclusive circular economy indicators for the environmental component will look like this: $\vec{x}_3 = \{x_{3i}\}, i = 1, \dots, 10$

where x_{31} – greenhouse gas emissions, thousand tons of CO2 equivalent;
x_{32} – land use, km²;
x_{33} – renewable energy, %;
x_{34} – internal consumption of material resources, million tons;
x_{35} – state budget for R&D related to the environment;
x_{36} – ecologically adjusted multifactor productivity growth;
x_{37} – household waste is processed or composted, % of processed waste;
x_{38} – productivity of CO2, $/kg;
x_{39} – adjustment to reduce pollution, %;
x_{310} – development of environmental technologies, % of all technologies.

To rank countries according to the set of indicators of the environmental component, we will have a partial case for the integrated indicator N =1:

$$P_S^i = \lambda_{y(\vec{x}_3^i)} \tag{5}$$

where $\lambda_{y(\vec{x}_3^i)}$ – the eigenvalue of the correlation matrix, which characterizes the absolute contribution (significance) of the corresponding main component for the function $y(\vec{x}_3^i)$ to the total variance of the indicators of the ecological component system for the i-th country.

The regression function $y(\vec{x}_i)$ takes the following general form:

$$y(\vec{x}_3) = \sum_{j=1}^{10} w_{1j} \cdot x_{3j}, \tag{6}$$

where w_{1j} – the elements of the eigenvector of the correlation matrix for the first principal component. In the course of calculations, some indicators from the set $\vec{x}_3$may turn out to be uncorrelated and not included in the general regression equation.

Table 4: The results of the study on the ecological component of the indicator

Country	Regression function	% comp.	Weakly correlated indicators
Australia	$y(\vec{x}_3) = -0{,}322 \cdot x_{31} - 0{,}392 \cdot x_{33} - 0{,}408 \cdot x_{34} + 0{,}42 \cdot x_{37} - 0{,}391 \cdot x_{38} + 0{,}277 \cdot x_{39} + 0{,}413 \cdot x_{310}$,	78%	$x_{32}x_{35}x_{36}$
Austria	$y(\vec{x}_3) = -0{,}317 \cdot x_{34} + 0{,}411 \cdot x_{35} + 0{,}43 \cdot x_{36} + 0{,}393 \cdot x_{37} - 0{,}433 \cdot x_{38} + 0{,}433 \cdot x_{310}$,	84%	$x_{31}x_{32}x_{33}$
Belgium	$y(\vec{x}_3) = -0{,}453 \cdot x_{31} + 0{,}406 \cdot x_{33} - 0{,}441 \cdot x_{35} + 0{,}378 \cdot x_{36} - 0{,}339 \cdot x_{37} + 0{,}421 \cdot x_{38}$,	70%	$x_{32}x_{34}x_{39}x_{310}$
Greece	$y(\vec{x}_3) = 0{,}427 \cdot x_{31} - 0{,}386 \cdot x_{33} + 0{,}422 \cdot x_{34} - 0{,}414 \cdot x_{38} + 0{,}39 \cdot x_{39} + 0{,}409 \cdot x_{310}$,	86%	$x_{32}x_{35}x_{36}x_{37}$
Denmark	$y(\vec{x}_3) = 0{,}358 \cdot x_{31} + 0{,}317 \cdot x_{32} - 0{,}36 \cdot x_{33} + 0{,}385 \cdot x_{36} - 0{,}376 \cdot x_{37} - 0{,}368 \cdot x_{38} + 0{,}334 \cdot x_{39} + 0{,}37 \cdot x_{310}$,	85%	$x_{34}x_{35}$
Great Britain	$y(\vec{x}_3) = 0{,}425 \cdot x_{31} - 0{,}42 \cdot x_{33} + 0{,}353 \cdot x_{35} - 0{,}371 \cdot x_{37} - 0{,}41 \cdot x_{38} + 0{,}293 \cdot x_{39} + 0{,}356 \cdot x_{310}$,	77%	$x_{32}x_{34}x_{36}$
Estonia	$y(\vec{x}_3) = -0{,}372 \cdot x_{31} + 0{,}378 \cdot x_{32} + 0{,}498 \cdot x_{33} - 0{,}438 \cdot x_{36} + 0{,}209 \cdot x_{37} + 0{,}485 \cdot x_{38}$,	86%	$x_{34}x_{35}x_{39}x_{310}$
Israel	$y(\vec{x}_3) = 0{,}36 \cdot x_{31} - 0{,}447 \cdot x_{33} - 0{,}305 \cdot x_{37} - 0{,}454 \cdot x_{38} - 0{,}417 \cdot x_{39} + 0{,}444 \cdot x_{310}$,	79%	$x_{32}x_{34}x_{35}x_{36}$
Canada	$y(\vec{x}_3) = 0{,}234 \cdot x_{33} - 0{,}324 \cdot x_{34} - 0{,}019 \cdot x_{35} + 0{,}425 \cdot x_{36} - 0{,}426 \cdot x_{37} - 0{,}457 \cdot x_{38} - 0{,}234 \cdot x_{39} + 0{,}463 \cdot x_{310}$,	83%	$x_{31}x_{32}$
China	$y(\vec{x}_3) = -0{,}425 \cdot x_{33} - 0{,}407 \cdot x_{34} + 0{,}432 \cdot x_{35} + 0{,}393 \cdot x_{36} - 0{,}426 \cdot x_{38} + 0{,}363 \cdot x_{310}$,	89%	$x_{31}x_{32}$
Latvia	$y(\vec{x}_3) = -0{,}124 \cdot x_{31} + 0{,}091 \cdot x_{34} + 0{,}404 \cdot x_{35} + 0{,}444 \cdot x_{36} + 0{,}467 \cdot x_{37} + 0{,}488 \cdot x_{38} + 0{,}401 \cdot x_{310}$,	79%	x_{37}
Lithuania	$y(\vec{x}_3) = -0{,}415 \cdot x_{32} + 0{,}39 \cdot x_{33} \mp 0{,}379 \cdot x_{34} - 0{,}426 \cdot x_{35} + 0{,}421 \cdot x_{37} + 0{,}416 \cdot x_{38}$,	88%	$x_{31}x_{36}x_{310}$
Luxembourg	$y(\vec{x}_3) = -0{,}385 \cdot x_{31} + 0{,}403 \cdot x_{33} + 0{,}408 \cdot x_{34} + 0{,}342 \cdot x_{36} + 0{,}361 \cdot x_{37} + 0{,}392 \cdot x_{38} + 0{,}349 \cdot x_{39}$,	82%	$x_{32}x_{35}x_{310}$

Table 4: The results of the study on the ecological component of the indicator (cont.)

Country	Regression function	% comp.	Weakly correlated indicators
Mexico	$y(\vec{x}_3) = -0{,}408 \cdot x_{31} + 0{,}324 \cdot x_{32} - 0{,}411 \cdot x_{33} - 0{,}357 \cdot x_{34} + +0{,}274 \cdot x_{35} - 0{,}433 \cdot x_{38} + 0{,}409 \cdot x_{310},$	74%	$x_{36}x_{37}x_{39}$
Netherlands	$y(\vec{x}_3) = -0{,}469 \cdot x_{33} + 0{,}338 \cdot x_{35} - 0{,}173 \cdot x_{36} - 0{,}461 \cdot x_{37} - -0{,}432 \cdot x_{38} - 0{,}111 \cdot x_{39} + 0{,}475 \cdot x_{310},$	83%	$x_{31}x_{32}x_{34}$
Germany	$y(\vec{x}_3) = 0{,}337 \cdot x_{31} - 0{,}419 \cdot x_{32} - 0{,}421 \cdot x_{33} - 0{,}417 \cdot x_{37} - -0{,}364 \cdot x_{38} + 0{,}271 \cdot x_{39} + 0{,}392 \cdot x_{310},$	83%	$x_{34}x_{35}x_{36}$
New Zealand	$y(\vec{x}_3) = -0{,}431 \cdot x_{33} - 0{,}414 \cdot x_{34} + 0{,}313 \cdot x_{35} - 0{,}315 \cdot x_{36} - -0{,}418 \cdot x_{37} - 0{,}387 \cdot x_{38} + 0{,}347 \cdot x_{310},$	76%	$x_{31}x_{32}x_{39}$
South Africa	$y(\vec{x}_3) = -0{,}378 \cdot x_{31} - 0{,}522 \cdot x_{35} + 0{,}238 \cdot x_{36} - 0{,}512 \cdot x_{39} + 0{,}516 \cdot x_{310},$	72%	$x_{32}x_{33}x_{34}x_{37}x_{38}$
South Korea	$y(\vec{x}_3) = -0{,}371 \cdot x_{31} - 0{,}421 \cdot x_{33} - 0{,}426 \cdot x_{35} - 0{,}373 \cdot x_{37} - - 0{,}426 \cdot x_{38} + 0{,}428 \cdot x_{310}$	90%	$x_{32}x_{34}x_{36}x_{39}$
Poland	$y(\vec{x}_3) = -0{,}353 \cdot x_{33} + 0{,}491 \cdot x_{34} + 0{,}462 \cdot x_{35} - -0{,}208 \cdot x_{36} + 0{,}422 \cdot x_{37} + 0{,}446 \cdot x_{38},$	76%	$x_{31}x_{32}x_{39}x_{310}$
Turkey	$y(\vec{x}_3) = 0{,}442 \cdot x_{31} + 0{,}391 \cdot x_{33} + 0{,}449 \cdot x_{34} + 0{,}222 \cdot x_{36} + +0{,}355 \cdot x_{37} + 0{,}363 \cdot x_{38} + 0{,}378 \cdot x_{310},$	69%	$x_{32}x_{35}x_{39}$
Hungary	$y(\vec{x}_3) = 0{,}369 \cdot x_{31} - 0{,}348 \cdot x_{34} - 0{,}411 \cdot x_{35} + 0{,}482 \cdot x_{37} - -0{,}489 \cdot x_{38} + 0{,}319 \cdot x_{310}$	93%	$x_{32}x_{33}x_{36}x_{39}$
Finland	$y(\vec{x}_3) = 0{,}419 \cdot x_{31} - 0{,}387 \cdot x_{32} - 0{,}424 \cdot x_{33} - 0{,}38 \cdot x_{35} - -0{,}407 \cdot x_{37} + 0{,}430 \cdot x_{310},$	84%	$x_{34}x_{36}x_{38}x_{39}$
France	$y(\vec{x}_3) = 0{,}447 \cdot x_{31} - 0{,}173 \cdot x_{34} + 0{,}216 \cdot x_{35} - 0{,}451 \cdot x_{37} - -0{,}441 \cdot x_{38} + 0{,}359 \cdot x_{39} + 0{,}443 \cdot x_{310}$	83%	$x_{32}x_{33}x_{36}$
Czech Republic	$y(\vec{x}_3) = 0{,}339 \cdot x_{31} - 0{,}414 \cdot x_{33} - 0{,}405 \cdot x_{34} - 0{,}319 \cdot x_{35} + +0{,}224 \cdot x_{36} - 0{,}402 \cdot x_{37} - 0{,}397 \cdot x_{38} + 0{,}28 \cdot x_{310}$	72%	$x_{32}x_{39}$
Japan	$y(\vec{x}_3) = 0{,}342 \cdot x_{31} - 0{,}368 \cdot x_{32} - 0{,}362 \cdot x_{33} + 0{,}371 \cdot x_{34}0{,}261 \cdot x_{35} - 0{,}374 \cdot x_{37} - 0{,}365 \cdot x_{38} + 0{,}372 \cdot x_{310}$	86%	$x_{36}x_{39}$

Source: constructed by author.

Leader countries in the economic and social component of the GICE indicator have a strong influence on the global economic arena. They produce 66% of GDP, with only 20% of the world's population, plus their material consumption per capita is 10 times that of developing countries. The material influence of such countries lies in the high share of employed services, which provide 71% of the EU's GDP and 80% of the USA's. In contrast, agriculture provides only 4% of employment in developed countries – compared to 25% in countries such as Pakistan and 18% in India. Imports and export volumes account for 68% of world trade, illustrating the extent to which countries that will initiate change in the implementation of a global inclusive circular economy are at the stage.

Per capita waste production is close to low enough to indicate effective waste avoidance and management in advanced GICE countries (according to the constituents studied) and compensates for the disproportionately large amount of incoming materials. The leading countries of GICE are faced with a difficult choice, which in some cases obliges them to make a costly «correction» of the current situation. In the energy transition, for example, such countries are faced with accumulated assets, demanding to get rid of dependence on coal or nuclear power plants. Trade restrictions imposed by developing countries have prompted some countries to change, in particular, to take control of their own waste that needs processing. China actually imports with restrictions on certain types of waste, prompting other countries to take similar measures (to impose trade restrictions on other countries and regions), accelerating the development of a national circular direction, such as the economic program in Australia and other measures at the local, regional and national levels.

In a world with limited resources, «heavyweights» - consumers are a problem both for themselves and for others. It is precisely such developed countries, which are already in the spotlight regarding their climate responsibilities and post-industrial status, that are under special pressure and under the influence of the development of Industry 4.0. in the direction of their digitization and building a smart society.

Based on the above, it is possible to distinguish the following methods of consumption in the «heavyweight» countries:

- firstly, to reduce the consumption of goods due to the extension of their service life;
- secondly, to increase the efficiency of the use of materials with the help of new technologies and circular design;
- thirdly, to reduce the total number and volume of necessary goods, through the promotion and adoption of business models of joint use (proliferation of sharing platforms).

Since such countries consume a disproportionate amount of global resources, they bear the entire burden of bringing packaging to a circular design. Such a circular design

in the future will allow stockpiling towards renewable resources during construction or production, optimizing utility during the use phase (perhaps in combination with approaches reoriented to circular business models) and, for as long as possible, maintaining and preserving what is already a consequence of reuse, recycling and extension of their service life.

Table 5: Indicators that lead to a decrease in the ecological component

Countries	Indicators
1	2
Australia	• emissions of greenhouse gases, thousand tons of CO2 equivalent; • renewable energy, %; • internal consumption of material resources, million tons; • productivity of CO2, $/kg;
Austria	• internal consumption of material resources, million tons; • productivity of CO2, $/kg;
Belgium	• emissions of greenhouse gases, thousand tons of CO2 equivalent; • the state budget for R&D related to the environment; • municipal waste is processed or composted, % of processed waste;
Greece	• renewable energy, %; • x_38– productivity of CO2, $/kg
Denmark	• renewable energy, % • municipal waste is processed or composted, % of processed waste; • productivity of CO2, $/kg
Great Britain	• renewable energy, % • productivity of CO2, $/kg
Estonia	• emissions of greenhouse gases, thousand tons of CO2 equivalent; • environmentally adjusted multifactor productivity growth;
Israel	• renewable energy, %; • municipal waste is processed or composted, % of processed waste; • regulation to reduce pollution, %;
Canada	• internal consumption of material resources, million tons; • the state budget for R&D related to the environment; • municipal waste is processed or composted, % of processed waste; • productivity of CO2, $/kg; • regulation to reduce pollution, %;
China	• renewable energy, %; • internal consumption of material resources, million tons; • productivity of CO2, $/kg;
Latvia	• greenhouse gas emissions, thousand tons of CO2 equivalent

Table 5: Indicators that lead to a decrease in the ecological component (cont.)

Countries	Indicators
1	2
Luxembourg	• greenhouse gas emissions, thousand tons of CO2 equivalent
Mexico	• greenhouse gas emissions, thousand tons of CO2 equivalent • renewable energy, %; • internal consumption of material resources, million tons • productivity of CO2, $/kg
Germany	• land use, km^2 • municipal waste is processed or composted, % of processed waste • productivity of CO2, $/kg • renewable energy, %;
South Africa	• emissions of greenhouse gases, thousand tons of CO2 equivalent; • the state budget for R&D related to the environment • regulation to reduce pollution, %;
Poland	• renewable energy, %; • environmentally adjusted multifactor productivity growth
Finland	• land use, km^2; • renewable energy, %; • the state budget for R&D related to the environment; • municipal waste is processed or composted, % of processed waste
France	• internal consumption of material resources, million tons • municipal waste is processed or composted, % of processed waste • productivity of CO2, $/kg
Czech Republic	• renewable energy, %; • internal consumption of material resources, million tons; • the state budget for R&D related to the environment • municipal waste is processed or composted, % of processed waste; • productivity of CO2, $/kg
Japan	• land use, km^2; • renewable energy, %; • the state budget for R&D related to the environment • municipal waste is processed or composted, % of processed waste; • productivity of CO2, $/kg

Source: constructed by author.

The table shows the rating of the countries of the world for which research was conducted, according to the environmental component of the global inclusive circular economy.

Table 6: Rating of the countries of the world according to the ecological component of the global inclusive circular economy

Position	Country	Indicator, $\lambda_{y(\vec{x}_3)}$	Normalized index
1	Japan	6.9147	**1**
2	Denmark	6.7992	**0.9682**
3	Germany	5.9637	*0.7379*
4	Czech Republic	5.7241	*0.6718*
5	Luxembourg	5.7176	*0.6701*
6	Australia	5.4443	0.5947
7	United Kingdom	5.4076	0.5846
8	South Korea	5.3953	0.5812
9	China	5.3548	0.5701
10	Lithuania	5.3035	0.5559
11	Mexico	5.1803	0.5220
12	Greece	5.1587	0.5160
13	USA	5.0768	0.4934
14	Finland	5.0639	0.4899
15	Austria	5.0472	0.4853
16	Turkey	4.8309	0.4257
17	Israel	4.7633	0.4070
18	New Zealand	4.6616	0.3790
19	Netherlands	4.2680	0.2705
20	Canada	4.1352	0.2339
21	Belgium	4.1084	0.2265
22	Poland	3.9752	0.1898
23	France	3.9200	0.1746
24	Latvia	3.9133	0.1728
25	Hungary	3.8993	0.1689
26	South Africa	3.5843	0.0821
27	Estonia	3.4134	0.0350
28	Ukraine	**3.2865**	**0**

Source: constructed by author.

So, the rating of the countries of the world according to the ecological component of the global inclusive circular economy with the derived normative index shows the highest rating in Japan and Denmark. The group of countries with an indicator from 0 to 0.2 includes Ukraine and Estonia, South Africa, Hungary, Latvia, France, Poland. Unlike the social component, the ecological component is not represented by countries with high data, which would allow the indicator to be brought to a higher level.

Modern statistics and scientific research consider the circular economy rather one-sidedly, essentially as recycling and reuse of resources. The article proposed methodology (concept) of the global inclusive circular economy. It can be considered as a complex multidimensional system, the main components of which are the economic, sociological, ecological and circular aspects of the country's life. To achieve the goal, the GNU Regression, Econometrics and Time-series Library was used - an applied software package for econometric modeling, part of the GNU project. Accordingly, we will define the global inclusive circular economy as the SGICE (Global Inclusive Circular Economy) system, characterized by the vector of functions. For the most exhaustive consideration of the entire range of opportunities of the global inclusive circular economy, the work developed and accordingly analyzed the integrated index of the development of the global inclusive circular economy (Igice) by ecological, economic, social and circular components with isolated weakly correlated indicators, and the ranking of countries was carried out separately by the components of the index, on the basis of which circular cores are distinguished: social component (Belgium, Czech Republic, USA, China, France, Greece, Austria, Australia); ecological component (Japan, Denmark); economic component (Germany, China), which became the basis for the model of formation of global inclusive circular chains. The conducted cluster analysis based on the component indicators of the index of the global inclusive circular economy confirmed the formation of a large circular gap (gap) in the ecological and circular components. This indicator is introduced for the first time and makes it possible to comprehensively analyze the country and highlight additional effects that arise from the moment of implementation of circular business projects and the inclusion of the country (on the basis of firms, corporations) in global circular chains of added value. After all, the main problematic aspects are the illegal trade in waste and growing smuggling, which cause serious negative social consequences and actualize the inclusive component in the justification of the global inclusive circular economy paradigm. Waste reduction combined with resource wise use has the potential to address the gap resulting from the scarcity of natural resources and a growing global population or consumption. The formation of circular trade will contribute to: determination of priority materials for trade and the required level of processing capacity; coordination of material quality standards; promoting demand for used goods and secondary raw materials; removing

unnecessary regulatory barriers and avoiding environmentally harmful activities such as non-compliance, poor regulation and informal recovery.

The next step is the ranking of countries according to the circular component. Accordingly, the set of indicators of the global inclusive circular economy for the circular component will look like this:

$$\vec{x}_4 = x_{4i}, i = 1, \dots, 4$$

where x_{41} – percentage of garbage recovery, %;
x_{42} – percentage of recycling garbage, %;
x_{43} – treated household waste, million tons;
x_{44} – recycling garbage, million, tons.

To rank countries according to a set of indicators of the circular component, we will use the partial case for the integrated indicator N =1, accordingly we will get:

$$P_{\text{C}}^{i} = \lambda_{y(\vec{x}_4^i)} \tag{7}$$

where $\lambda_{y(\vec{x}_4^i)}$– eigenvalue of the correlation matrix, which characterizes the absolute contribution (significance) of the corresponding main component for the function $y(\vec{x}_4^i)$ to the total variance of the indicators of the circular component system for the ith country.

The function $y(\vec{x}_{\text{i}})$ takes the following form:

$$y(\vec{x}_4) = \sum_{j=1}^{6} w_{1j} \cdot x_{4j}, \tag{8}$$

where w_{1j} – elements of the eigenvector of the correlation matrix for the first principal component. In the course of calculations, some indicators from the set $\vec{x}_4$may turn out to be uncorrelated and not included in the general regression equation.

Countries are of absolute importance in shaping the circular economy. However, we are talking about a world that is only 8.6% circular, which indicates the corrosive effects and negative trends of the tradition of accepting waste.

Table 7: Results of analysis by circular component

Country	Regression function	% comp.	Weakly correlated indicators
Australia	$y(\vec{x}_4) = -0{,}506 \cdot x_{41} + 0{,}66 \cdot x_{43} + 0{,}556 \cdot x_{44},$	76%	x_{42}
Austria	$y(\vec{x}_4) = 0{,}536 \cdot x_{41} + 0{,}463 \cdot x_{42} + 0{,}477 \cdot x_{43} + 0{,}52 \cdot x_{44}.$	86%	
Belgium	$y(\vec{x}_4) = 0{,}543 \cdot x_{41} - 0{,}691 \cdot x_{43} + 0{,}476 \cdot x_{44}$	76%	x_{42}
Greece	$y(\vec{x}_4) = 0{,}568 \cdot x_{41} + 0{,}581 \cdot x_{43} + 0{,}582 \cdot x_{44}$	96%	x_{42}
Denmark	$y(\vec{x}_4) = 0{,}499 \cdot x_{41} + 0{,}5 \cdot x_{42} + 0{,}494 \cdot x_{43} + 0{,}506 \cdot x_{44}.$	96%	
Great Britain	$y(\vec{x}_4) = 0{,}474 \cdot x_{41} + 0{,}463 \cdot x_{42} + 0{,}481 \cdot x_{43} + 0{,}575 \cdot x_{44}.$	72%	

Table 7: Results of analysis by circular component (cont.)

Country	Regression function	% comp.	Weakly correlated indicators
Estonia	$y(\vec{x}_4) = 0{,}54 \cdot x_{41} + 0{,}603 \cdot x_{43} + 0{,}589 \cdot x_{44},$	86%	x_{42}
Israel	$y(\vec{x}_4) = 0{,}553 \cdot x_{41} - 0{,}383 \cdot x_{42} + 0{,}56 \cdot x_{43} + 0{,}484 \cdot x_{44}.$	79%	
Canada	$y(\vec{x}_4) = 0{,}494 \cdot x_{41} + 0{,}5 \cdot x_{42} + 0{,}504 \cdot x_{43} + 0{,}502 \cdot x_{44}.$	98%	
China	$y(\vec{x}_4) = -0{,}531 \cdot x_{41} + 0{,}625 \cdot x_{43} + 0{,}572 \cdot x_{44}$	78%	x_{42}
Latvia	$y(\vec{x}_4) = 527 \cdot x_{41} + 0{,}661 \cdot x_{43} + 0{,}534 \cdot x_{44}$	74%	x_{42}
Lithuania	$y(\vec{x}_4) = 0{,}557 \cdot x_{41} + 0{,}561 \cdot x_{42} - 0{,}289 \cdot x_{43} + 0{,}54 \cdot x_{44}.$	78%	
Luxembourg	$y(\vec{x}_4) = 0{,}56 \cdot x_{41} - 0{,}229 \cdot x_{42} + 0{,}591 \cdot x_{43} + 0{,}533 \cdot x_{44}.$	72%	
Mexico	$y(\vec{x}_4) = 0{,}571 \cdot x_{41} + 0{,}619 \cdot x_{43} + 0{,}54 \cdot x_{44}$	74%	x_{42}
Netherlands	$y(\vec{x}_4) = 0{,}525 \cdot x_{41} + 0{,}555 \cdot x_{42} - 0{,}446 \cdot x_{43} + 0{,}466 \cdot x_{44}$	77%	
Germany	$y(\vec{x}_4) = 0{,}486 \cdot x_{41} + 0{,}504 \cdot x_{42} + 0{,}494 \cdot x_{43} + 0{,}516 \cdot x_{44}$	93%	
New Zealand	$y(\vec{x}_4) = 0{,}585 \cdot x_{41} + 0{,}561 \cdot x_{43} + 0{,}586 \cdot x_{44},$	91%	x_{42}
South Africa	$y(\vec{x}_4) = 0{,}585 \cdot x_{41} + 0{,}561 \cdot x_{43} + 0{,}586 \cdot x_{44}$	91%	x_{42}
South Korea	$y(\vec{x}_4) = 0{,}46 \cdot x_{41} + 0{,}465 \cdot x_{42} + 0{,}539 \cdot x_{43} + 0{,}531 \cdot x_{44}.$	83%	
Poland	$y(\vec{x}_4) = 0{,}575 \cdot x_{41} + 0{,}582 \cdot x_{43} + 0{,}575 \cdot x_{44}$	98%	x_{42}
Turkey	$y(\vec{x}_4) = 0{,}425 \cdot x_{41} + 0{,}523 \cdot x_{42} + 0{,}52 \cdot x_{43} + 0{,}524 \cdot x_{44}$	80%	
Hungary	$y(\vec{x}_4) = 0{,}51 \cdot x_{41} + 0{,}539 \cdot x_{42} - 0{,}416 \cdot x_{43} + 0{,}525 \cdot x_{44}$	84%	
Finland	$y(\vec{x}_4) = 0{,}461 \cdot x_{41} + 0{,}526 \cdot x_{42} + 0{,}487 \cdot x_{43} + 0{,}523 \cdot x_{44}$	87%	
France	$y(\vec{x}_4) = 0{,}576 \cdot x_{41} + 0{,}578 \cdot x_{43} + 0{,}578 \cdot x_{44}$	99%	x_{42}
Czech Republic	$y(\vec{x}_4) = 0{,}476 \cdot x_{41} + 0{,}515 \cdot x_{42} + 0{,}494 \cdot x_{43} + 0{,}515 \cdot x_{44}$	92%	
Japan	$y(\vec{x}_4) = -0{,}527 \cdot x_{41} + 0{,}443 \cdot x_{42} + 0{,}498 \cdot x_{43} + 0{,}535 \cdot x_{44}.$	82%	

Source: constructed by author.

The linear economy and its legacy are embedded deep within our society. These problems are highlighted by a widening circular gap. It is driven by entrepreneurs, businesses and communities coming together. So, sensing this acute urgency and parallel emerging opportunity, a growing number of countries and national governments are now beginning to shape their strategies to support investment in a sustainable and spe-

cific circular economy agenda. In addition to facilitating national and local action, national governments are also key agents and organizers of the global coordination or global coalition that together effectively establish new circular rules. In turn, the dynamics of this world trade is reflected correctly through national, regional and local markets. This can be seen in global supply chains that often cross continents - leaving footprints wherever they go in terms of resource and material extraction, processing, production, consumption and waste management. Until such joint governance systems exist, countries will remain tempted to regulate the quality of imports and exports by acting unilaterally, shirking their environmental and social responsibilities.

Table 8: Weakly correlated indicators of the circular component

Country	Indicator
Australia	• debris recovery percentage
Belgium	• processed household waste
Israel	• percentage of recycling waste, %
China	• debris recovery percentage
Lithuania	• processed household waste
Luxembourg	• percentage of recycling waste, %
Netherlands	• processed household waste, million, tons
Hungary	• processed household waste, million, tons
Japan	• debris recovery percentage, %

Source: constructed by author.

The table shows the rating of the countries of the world for which research was conducted according to the *circular component* of the global inclusive circular economy.

Table 9: Ranking of the countries of the world by the circular component of the global inclusive circular economy

Position	Country	Indicator, $\lambda_{y(\vec{x}_1)}$	Normalized index
1	**Canada**	**3.9411**	**1**
2	**Denmark**	**3.8456**	**0.9497**
3	**Germany**	**3.7145**	**0.8806**
4	**Czech Republic**	**3.71**	**0.8782**
5	*Finland*	3.4847	0.7595
6	*Austria*	3.4446	0.7384
7	*Hungary*	3.3892	0.7092
8	*Japan*	3.2734	0.6482

Table 9: Ranking of the countries of the world by the circular component of the global inclusive circular economy (cont.)

Position	Country	Indicator, $\lambda_{y(\vec{x}_1)}$	Normalized index
9	*Israel*	3.1874	0.6029
10	*Lithuania*	3.1303	0.5728
11	*Netherlands*	3.0689	0.5405
12	France	2.9826	0.4950
13	Poland	2.9334	0.4691
14	Greece	2.899	0.4510
15	United Kingdom	2.8922	0.4474
16	Belgium	2.7625	0.3791
17	New Zealand	2.743	0.3688
18	USA	2.743	0.3688
19	South Africa	2.743	0.3688
20	Luxembourg	2.6456	0.3175
21	Estonia	2.5277	0.2554
22	China	2.35	0.1617
23	South Korea	2.3458	0.1595
24	Australia	2.2956	0.1331
25	Mexico	2.2375	0.1025
26	Latvia	2.2219	0.0943
27	Turkey	2.218	0.0922
28	Ukraine	2.043	0

Source: constructed by author.

According to the ranking of the countries of the world according to the circular component of the global inclusive circular economy, the leaders are Canada, Denmark, Germany and the Czech Republic. The group of countries with a normalized index (0–0.2) consists of China, South Korea, Australia, Mexico, Latvia, Turkey and Ukraine. Based on the obtained indicators in the ratings by components, an integral index of the global inclusive circular economy was determined for the countries under study. The results of calculations are given in Table 10.

Table 10: A generalized indicator of the global inclusive circular economy

Position	Country	Ecological	Economic	Social	Circular	iGICE
1	Germany	0.7379	1.0000	0.4797	0.8806	0.7745
2	Czech Republic	0.6718	0.4718	0.9807	0.8782	0.7506
3	Denmark	0.9682	0.4815	0.5677	0.9497	0.7418
4	Japan	1.0000	0.4781	0.5460	0.6482	0.6681
5	Austria	0.4853	0.4968	0.8946	0.7384	0.6538
6	USA	0.4934	0.7230	0.9714	0.3688	0.6392
7	China	0.5701	0.8350	0.9348	0.1617	0.6254
8	Luxembourg	0.6701	0.6893	0.6774	0.3175	0.5886
9	Greece	0.5160	0.4178	0.8991	0.4510	0.5710
10	Canada	0.2339	0.5310	0.5044	1.0000	0.5673
11	Lithuania	0.5559	0.3936	0.7248	0.5728	0.5618
12	Netherlands	0.2705	0.7479	0.6059	0.5405	0.5412
13	Belgium	0.2265	0.5202	1.0000	0.3791	0.5314
14	United Kingdom	0.5846	0.3853	0.6882	0.4474	0.5264
15	Hungary	0.1689	0.4899	0.6444	0.7092	0.5031
16	Australia	0.5947	0.4185	0.8551	0.1331	0.5003
17	Israel	0.4070	0.2738	0.7022	0.6029	0.4965
18	France	0.1746	0.3735	0.9305	0.4950	0.4934
19	Finland	0.4899	0.1794	0.4616	0.7595	0.4726
20	Mexico	0.5220	0.5121	0.6939	0.1025	0.4576
21	South Korea	0.5812	0.3518	0.7093	0.1595	0.4505
22	Poland	0.1898	0.3667	0.6857	0.4691	0.4278
23	New Zealand	0.3790	0.4676	0.3997	0.3688	0.4038
24	Latvia	0.1728	0.5323	0.6465	0.0943	0.3615
25	Estonia	0.0350	0.3947	0.6924	0.2554	0.3443
26	Turkey	0.4257	0.3054	0.1598	0.0922	0.2458
27	Ukraine	0.0000	0.0836	0.4443	0.0001	0.1320
28	South Africa	0.0821	0.0001	0.0001	0.3688	0.1127

Source: constructed by author.

According to the results of calculations, the integral index was normalized according to formula (4), which is shown in Table 10. The table shows that Germany has the best indicator of global inclusive circular economy from the given list of countries $P_{GICE}^{1} = 4{,}612$, which took, accordingly, the first position in the rating. Among the specified countries, Ukraine occupies the last position in terms of the value of the integrated indicator $P_{GICE}^{5} = 2{,}013$.

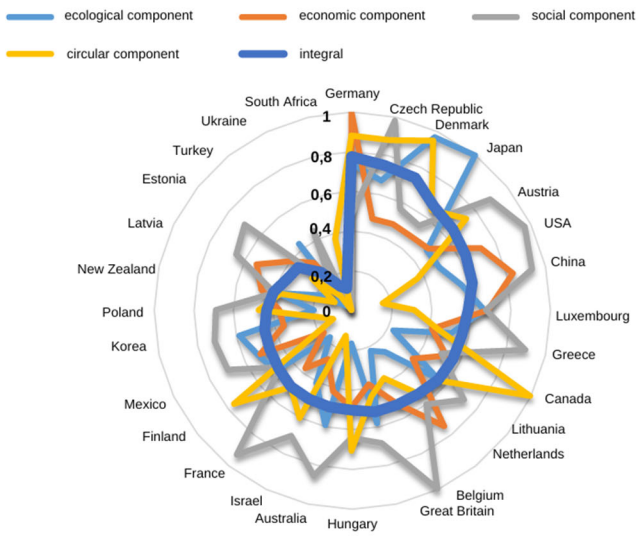


Figure 1: Integral index of the global inclusive circular economy

Source: constructed by author.

So, for the leading countries of Germany, the Czech Republic, and Denmark, in which iGICE is within 0.74-0.77, the highest values are shown by the circular component (0.94-0.88). In addition, it is important to note that Denmark has a high ecological component of the indicator (0.97). Thus, in the potential formation of global circular chains of added value, these countries will play the role of circular cores. In addition, the cluster analysis (Fig. 3.14) based on the component indicators of the global inclusive circular economy index revealed a huge circular gap (gap-circularity) between this core and other countries (1).

Thus, Figures 1 and 2 make it possible to formulate separate clusters, on the basis of which the cores of the global inclusive circular economy with isolated gaps (gap-circularity) between them are distinguished.

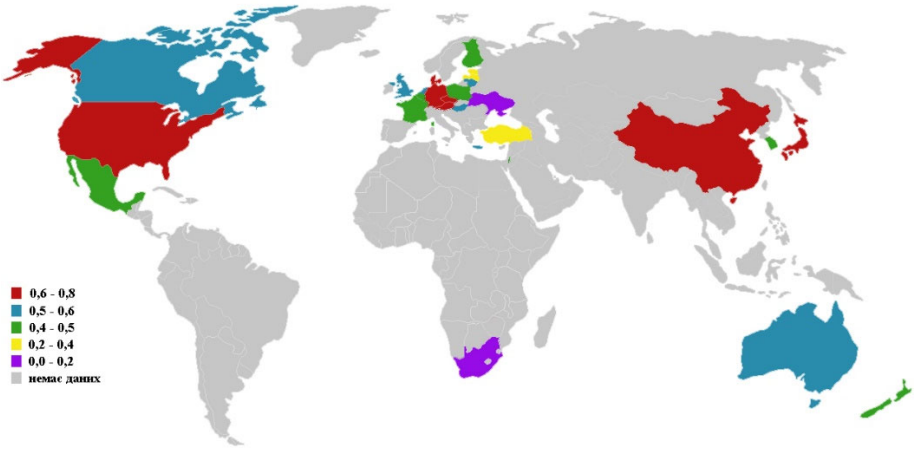


Figure 2: Mapping the integral index of the global inclusive circular economy
Source: constructed by the author.

The second group with iGICE 0.62-0.67 consists of the following countries: Japan, Austria, USA, China. This group is special in that it includes the leading countries in international trade and the international economy in general - these are the USA and China. Accordingly, for them, the most important component of the index is social and economic. In addition, based on the sectoral analysis of international trade in waste and scrap for 2016, the USA, China, Germany and Canada are the leaders in export and import, which confirms their place in the first group according to the index. The third group is the most numerous (9 out of 28 studied countries) iGICE within 0.5: Luxembourg, Greece, Canada, Lithuania, the Netherlands, Belgium, Great Britain, Hungary, Australia. The most important component is the social one, along with the average level of the circular component. In addition, the social vector is the core that unites all the studied countries. This confirms the socially inclusive vector, which is actively implemented in all countries and reflects the inseparability in the positioning of the global inclusive circular economy, where inclusiveness is at the priority level. The fourth group of countries (0.40-0.49) consists of seven countries with the highest index of the social component, with the exception of Finland, where the maximum is the circular. A low level of iGICE (0.24-0.36) is noted in Latvia, Estonia and Turkey.

4. Conclusion

The work analyzed the integral index of the development of the global inclusive circular economy indicator (Igice) according to environmental, economic, social and circular components with isolated weakly correlated indicators and ranked countries separately according to the components of the index, on the basis of which the circular

cores were isolated: social component (Belgium, Czech Republic, USA, China, France, Greece, Austria, Australia); ecological component (Japan, Denmark); economic component (Germany, China), which became the basis for the model of formation of global inclusive circular chains. Leader countries in the economic and social component of the GICE indicator have strong influence on the global economic arena. They produce 66% of GDP, having only 20% of the world's population, plus their material consumption per capita the population is 10 times larger than in developing countries. Material the influence of such countries lies in the high share of employed services, providing 71% GDP of the EU and 80% of the US. In contrast, agriculture provides only 4% employment in developed countries, while in countries like Pakistan this figure is 25%, and in India - 18%. The import and export amounts are 68% of world trade, illustrating the degree to which countries are which will initiate changes in the implementation of the global inclusive circular economy.

The production of waste per capita is close enough which indicates effective waste avoidance and management in leading GICE countries (according to the studied components) and compensates for the disproportionately large amount of incoming materials. The leading countries of GICE are facing a difficult task of a choice that in some cases obliges them to make a costly "adjustment" to the current situation. In the energy transition, for example, such countries face each other with accumulated assets, demanding to get rid of dependence on coal or nuclear power plants. Trade restrictions imposed by countries that developing countries encouraged some to change, in particular, to take control of their own waste that needs processing. China actually imports with restrictions on certain types of waste, encouraging other countries to take similar measures (introduce trade restrictions by other countries and regions), accelerating development in national circular direction, such as the economic program in Australia and other events at local, regional and national levels.

In a world with limited resources, "heavyweights" - consumers are a problem for themselves and for others. It is such developed countries, which are already in the center of attention regarding their climatic responsibilities and post-industrial situation, that are under special pressure and under the influence of the development of Industry 4.0. in their direction digitization and building a smart society.

Based on the above, the following methods of consumption can be identified in "heavyweight" countries: first, to reduce the consumption of goods due to the extension of their term exploitation; secondly, to increase the efficiency of the use of materials with the help of new technologies and circular design; thirdly, to reduce the total number and volume of necessary goods, through promotion and adoption of sharing business models (proliferation of sharing platforms).

Because such countries consume disproportionately global resources, they bear the entire burden of bringing packaging to circularity design. Such a circular design will

allow for stacking in the future stocks towards renewable resources during construction or production, optimizing utility during the use phase (perhaps in combination with approaches that are reoriented into circular business models) and, as long as possible, to maintain and preserve what is already a consequence of the repeated use, processing and extension of their service life. The conducted cluster analysis based on the component indicators of the index of the global inclusive circular economy confirmed the formation of a large circular gap (gap) in the ecological and circular components. Graphical visualization of the data clearly showed the circular circle that the index forms.

References

Benner, C., Pastor, M., (2018). Inclusive Economy Indicators. *The Rockefeller Foundation*. https://www.rockefellerfoundation.org/wp-content/uploads/Inclusive-Economies-Indicators-Exec-Summary.pdf.

Butler, B., (2018). *What Is An Inclusive Economy?*. https://accessventures.org/blog/what-is-an-inclusive-economy/.

Craig, N., (2021). Knowing the goal: an inclusive economy that can address the public health challenges of our time, *BMJ Journals*. https://jech.bmj.com/content/jech/75/11/1129.full.pdf.

European Commission, (2020). A strategy for smart, sustainable and inclusive growth. Brussels. http://ec.europa.eu/europe2020/index_en.htm.

GDP, The Guardian. *GDP is not a good measure of wellbeing – it's too materialistic*. https://www.theguardian.com/business/2018/dec/03/gdp-wellbeing-health-education-environment-joseph-stiglitz.

Krysovatyy, A., Zvarych, R., Zvarych, I., Krysovatyy, I., Krysovata, K., (2020). Methodological architectonics of inclusive circular economy for eco-security of society under pandemic. *Economic Annals-XXI*, 184(7-8), 4-15. doi: https://doi.org/10.21003/ea.V184-01.

Krysovatyy, A., Mokiy, A., Zvarych, R., Zvarych, I., (2018). Alterglobalization via the inclusive circular economy paradigm. *Economic Annals-XXI*, 174(11-12), 4–9. doi: https://doi.org/10.21003/ea.V174-0.

McCartney, G., (2021). Knowing the goal: an inclusive economy that can address the public health challenges of our time, *BMJ Journals*. https://jech.bmj.com/content/jech/75/11/1129.full.pdf.

McGregor, P., (2021). Knowing the goal: an inclusive economy that can address the public health challenges of our time, *BMJ Journals*. https://jech.bmj.com/content/jech/75/11/1129.full.pdf.

Reznikova, N., Zvarych, R., Zvarych, I., Shnyrkov, O., (2019). Global circular e-chain in overcoming the global waste. *Procedia Environmental Science, Engineering and Management*, 6(4), 641–647.http://www.procedia-esem.eu/pdf/issues/2019/no4/72_Reznikova_19.pdf.

Roy, G., (2021). Knowing the goal: an inclusive economy that can address the public health challenges of our time, *BMJ Journals*. https://jech.bmj.com/content/jech/75/11/1129.full.pdf.

Sarica, S., (2021). Knowing the goal: an inclusive economy that can address the public health challenges of our time, *BMJ Journals*. https://jech.bmj.com/content/jech/75/11/1129.full.pdf.

Scobie, G., Katikireddi, S. V., (2021). Knowing the goal: an inclusive economy that can address the public health challenges of our time, *BMJ Journals*. https://jech.bmj.com/content/jech/75/11/1129.full.pdf.

Shipton, D., (2021). Knowing the goal: an inclusive economy that can address the public health challenges of our time, *BMJ Journals*. https://jech.bmj.com/content/jech/75/11/1129.full.pdf.

The World Economic Forum, (2018). The Inclusive Development Index. https://www.weforum.org/reports/the-inclusive-development-index-2018.

Zvarych, I., (2019). Hlobalna tsyrkuliarna ekonomika: «ekonomika kovboiv» vs «ekonomika kosmichnoho korablia». *Ternopil: VPTs «Ekonomichna dumka TNEU*, 337.

STATISTICS IN TRANSITION new series, June 2024
Vol. 24, No. 2, pp. 167–184, https://doi.org/10.59170/stattrans-2024-020
Received – 10.07.2019; accepted – 28.03.2023

A new parameter estimation method for the extended power Lindley distribution based on order statistics, with application

Devendra Kumar[1], **Maneesh Kumar**[2], **Sapna Yadav**[3], **Anju Goyal**[4]

Abstract

In this paper, we propose inference procedures for the estimation of parameters by using order statistics. First, we derive some new expressions for single and product moments of the order statistics from the extended power Lindley distribution. We then use these moments to obtain the best linear unbiased estimates (BLUEs) of the location and scale parameters based on Type-II right-censored samples. A real data set is analysed to illustrate the flexibility and importance of the model.
Key words: extended power Lindley distribution, order statistics, moments, best linear unbiased estimator.

1. Introduction

The Lindley distribution was proposed by Lindley (1958) in the context of fiducial and Bayesian statistics, which can be seen as a mixture of $exp(\xi)$ and $gamma(2,\xi)$ distributions. Later, Ghitany et al. (2008) studied the statistical and mathematical properties of this distribution and also showed that this distribution performs better than the well-known exponential distribution in many ways. The Lindley distribution has only one scale parameter and is capable of modelling data with monotonic increasing failure rates and due to this, the Lindley distribution does not provide enough flexibility for analysing different types of lifetime data. To increase the flexibility for modeling purposes it will be useful to consider further alternatives of this distribution. Also, Bouchahed and Zeghdoudi (2018) and Zeghdoudi et al. (2018) generalized the Lindley distribution. Recently, the three parameter extended power Lindley distribution was proposed by Alkarni (2015) for the flexibility of purpose. The extended power Lindley (EPL) distribution is specified by the following probability density function (pdf):

$$f(x;\tau,\xi,\delta) = \frac{\tau\xi^2}{\xi+\delta}(1+\delta x^{\tau})\, x^{\tau-1}\, e^{-\xi x^{\tau}}, \quad x>0; \quad \tau>0,\xi>0,\delta>0 \tag{1}$$

[1]Department of Statistics, Central University of Haryana and Department of Statistics, Faculty of Mathematical Sciences, University of Delhi, Delhi, India. E-mail: devendrastats@gmail.com. ORCID: https://orcid.org/0000-0001-5831-3315.
[2]Department of Statistics, Central University of Haryana, India. E-mail: maneeshjindal0307@gmail.com. ORCID: https://orcid.org/0000-0002-4186-5962.
[3]Department of Statistics, Central University of Haryana, India. E-mail: mailto:yadavsapna8622@gmail.com. ORCID: https://orcid.org/0009-0000-3143-3827.
[4]Department of Statistics, Panjab University, Chandigarh, India. E-mail: mailto:anju.statistics@gmail.com. ORCID: https://orcid.org/0000-0002-0171-3163.

and the cumulative distribution function (cdf) is

$$F(x;\tau,\xi,\delta) = 1 - \left(1 + \frac{\delta\xi}{\xi+\delta}x^{\tau}\right)e^{-\xi x^{\tau}}, \quad x > 0; \quad \tau > 0, \xi > 0, \delta > 0. \tag{2}$$

For $\delta = 1$ and $\delta = 1, \ \tau = 1$, the EPL distribution reduces to the power Lindley (PL) and Lindley distributions respectively.

Let us consider the pdf of the scale-parameter EPL distribution as

$$f(x;\tau,\delta,\xi,\sigma) = \frac{\tau\xi^2}{\sigma(\xi+\delta)}\left[1+\delta\left(\frac{x}{\sigma}\right)^{\tau}\right]\left(\frac{x}{\sigma}\right)^{\tau-1} e^{-\xi\left(\frac{x}{\sigma}\right)^{\tau}}, \ x > 0; \ \tau > 0, \delta > 0, \xi > 0, \sigma > 0, \tag{3}$$

and the pdf of the location-scale of the EPL distribution is

$$\begin{aligned} f(x;\tau,\delta,\xi,\sigma) &= \frac{\tau\xi^2}{\sigma(\xi+\delta)}\left[1+\delta\left(\frac{x-\mu}{\sigma}\right)^{\tau}\right]\left(\frac{x-\mu}{\sigma}\right)^{\tau-1} \\ &\times \ e^{-\xi\left(\frac{x-\mu}{\sigma}\right)^{\tau}}, \ x > \mu; \ \tau > 0, \delta > 0, \xi > 0, \sigma > 0, \mu > 0, \end{aligned} \tag{4}$$

Order statistics play an important role in a wide range of theoretical and practical problems such as estimation of the problems, reliability analysis, quality control and strength of materials, for example, Balakrishnan and Cohan (1991), Sanmel and Thomas (1997), Balakrishnan and Chandramouleeswaran, (1996), Sultan et al. (2000), Sultan and Balakrishnan (2000), Mahmoud et al. (2005), Jabeen et al. (2013), Sultan and AL-Thubyani (2016), Kumar et al. (2017), Kumar and Dey (2017a, 2017b), Ahsanullah and Alzaatreh (2018), Kumar and Goyal (2019a, 2019b), Kumar et al. (2020a, 2020b) and the references cited therein.

The rest of the paper is organized as follows. In Section 2, we introduce some lemmas on the EPL distribution and we discuss some expressions for single and double moments of order statistics. We use these moments to obtain the BLUEs for μ and σ in Section 3. In Section 4, we performed a numerical study using R software and we see that these expressions provide precise numerical evaluations. A real data application is presented in Section 5. Finally, we conclude the paper in Section 6.

2. Moments of Order Statistics

In this section, we provide and prove some lemmas and some new expressions for single and product moments of order statistics for the given a random sample $X_1, X_2, \ldots, X_n$ from the EPL distribution. Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote the order statistics corresponding to $X_1, X_2, \ldots, X_n$. The pdf of the rth order statistic is given by Arnold et al. (2003) and David and Nagaraja (2003) as

$$f_{X_{(r:n)}}(x) = C_{r:n}\, F^{r-1}(x)\,[1-F(x)]^{n-r} f(x) \tag{5}$$

for $x > 0$. The joint pdf of the rth and sth order statistics is given by Arnold et al. (2003) and David and Nagaraja (2003) as

$$f_{X_{(r:n)},X_{(s:n)}}(x,y) = C_{r,s:n}\, F^{r-1}(x)\,[F(y)-F(x)]^{s-1-r}\,[1-F(x)]^{n-s} f(x) f(y) \qquad (6)$$

for $0 < x < y$, where $C_{r:n} = \frac{n!}{(r-1)!(n-r)!}$ and $C_{r,s:n} = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}$.

2.1. Some Lemmas

Here, we provide and prove two lemmas.

Lemma 1 *Let $f(x)$ and $F(x)$ be given by (1) and (2), respectively. For $a > 0$, $b > 0$ and $p > 0$, let*

$$K(a,b,p) = \int_0^\infty x^p F^a(x)[1-F(x)]^b f(x)dx.$$

Then,

$$\begin{aligned} K(a,b,p) &= \frac{\xi^2}{(\xi+\delta)^{b+1}} \sum_{i_1=0}^{a} \sum_{i_2=0}^{i_1+b} \sum_{i_3=0}^{i_2+1} (-1)^{i_1} \frac{\xi^{i_2}\delta^{b+i_1-i_2+i_3}}{(\xi+\delta)^{i_1}} \binom{a}{i_1}\binom{i_1+b}{i_2}\binom{i_2+1}{i_3} \\ &\times \frac{\Gamma\left(\frac{p+\tau(i_3+1)}{\tau}\right)}{[\xi(i_1+b+1)]^{\frac{p+\tau(i_3+1)}{\tau}}}. \end{aligned}$$

Proof. By using binomial expansion in (5) and then from (1) and (2), we get

$$\begin{aligned} K(a,b,p) &= \sum_{i_1=0}^{a} (-1)^{i_1} \binom{a}{i_1} \int_0^\infty x^p [1-F(x)]^{i_1+b} f(x)dx. \\ &= \frac{\tau\xi^2}{(\xi+\delta)} \sum_{i_1=0}^{a} (-1)^{i_1} \binom{a}{i_1} \int_0^\infty x^{p+\tau-1}(1+\delta x^\tau) e^{-\xi(i_1+b+1)x^\tau} \\ &\times \left[\left(1+\frac{\delta\xi}{\xi+\delta}x^\tau\right)\right]^{i_1+b} dx \\ &= \frac{\tau\xi^2}{(\xi+\delta)^{b+1}} \sum_{i_1=0}^{a} \sum_{i_2=0}^{i_1+b} \sum_{i_3=0}^{i_2+1} (-1)^{i_1} \frac{\xi^{i_2}\delta^{b+i_1-i_2+i_3}}{(\xi+\delta)^{i_1}} \binom{a}{i_1}\binom{i_1+b}{i_2}\binom{i_2+1}{i_3} \\ &\times \int_0^\infty x^{p+\tau(i_3+1)-1}\, e^{-\xi(i_1+b+1)x^\tau} dx \\ &= \frac{\xi^2}{(\xi+\delta)^{b+1}} \sum_{i_1=0}^{a} \sum_{i_2=0}^{i_1+b} \sum_{i_3=0}^{i_2+1} (-1)^{i_1} \frac{\xi^{i_2}\delta^{b+i_1-i_2+i_3}}{(\xi+\delta)^{i_1}} \binom{a}{i_1}\binom{i_1+b}{i_2}\binom{i_2+1}{i_3} \\ &\times \frac{1}{[\xi(i_1+b+1)]^{\frac{p+\tau(i_3+1)}{\tau}}} \int_0^\infty z^{\frac{p+\tau i_3}{\tau}}\, e^{-z} dz, \end{aligned}$$

where $z = \xi(i_1+b+1)x^\tau$.

Lemma 2 *Let $f(x)$ and $F(x)$ be given by (1) and (2), respectively. For $a>0$, $b>0$, c, $p>0$ and $q>0$, let*

$$L(a,b,c,p,q) = \int_0^\infty \int_x^\infty x^p y^q [F(x)]^a [F(y)-F(x)]^b [1-F(y)]^c f(x) f(y) dy dx.$$

Then,

$$\begin{aligned} L(a,b,c,p,q) &= \frac{\xi^4}{(\xi+\delta)^{b+c+2}} \sum_{i_1=0}^{a} \sum_{i_2=0}^{b} \sum_{i_3=0}^{i_1+b-i_2} \sum_{i_4=0}^{i_2+c} \sum_{i_5=0}^{i_3+1} \sum_{i_6=0}^{i_4+1} \sum_{i_7=0}^{\frac{q+\tau i_6}{\tau}} (-1)^{i_1+i_2} \binom{a}{i_1} \binom{b}{i_2} \\ &\times \binom{i_1+b-i_2}{i_3} \binom{i_2+c}{i_4} \binom{i_3+1}{i_5} \binom{i_4+1}{i_6} \frac{\xi^{i_3+i_4} \delta^{i_1+b-i_3-i_4+i_5+i_6}}{(\xi+\delta)^{i_1}} \\ &\times \frac{\left(\frac{q+\tau i_6}{\tau}\right)!}{[\xi(i_2+c+1)]^{\frac{q+\tau(i_6+1)}{\tau}}} \frac{[\xi(i_2+c+1)]^{i_7}}{i_7!} \frac{\Gamma\left(\frac{p+\tau(i_5+i_7+1)}{\tau}\right)}{[\xi(i_1+b+c+2)]^{\frac{p+\tau(i_5+i_7+1)}{\tau}}}. \end{aligned}$$

Proof. By using binomial expansion in (6) and then from (1) and (2), we get

$$\begin{aligned} L(a,b,c,p,q) &= \sum_{i_1=0}^{a} \sum_{i_2=0}^{b} (-1)^{i_1+i_2} \binom{a}{i_1} \binom{b}{i_2} \\ &\times \int_0^\infty \int_x^\infty x^p y^q [1-F(x)]^{i_1+b-i_2} [1-F(y)]^{i_2+c} f(x) f(y) dy dx \\ &= \frac{\tau^2 \xi^4}{(\xi+\delta)^2} \sum_{i_1=0}^{a} \sum_{i_2=0}^{b} (-1)^{i_1+i_2} \binom{a}{i_1} \binom{b}{i_2} \int_0^\infty \int_x^\infty x^{p+\tau-1} y^{q+\tau-1} \\ &\times (1+\delta x^\tau)(1+\delta y^\tau) e^{-\xi(i_1+b-i_2+1)x^\tau} e^{-\xi(i_2+c+1)y^\tau} \\ &\times \left[\left(1+\frac{\xi\delta}{\xi+\delta} x^\tau\right)\right]^{i_1+b-i_2} \left[\left(1+\frac{\xi\delta}{\xi+\delta} y^\tau\right)\right]^{i_2+c} dx dy \\ &= \frac{\tau^2 \xi^4}{(\xi+\delta)^{b+c+2}} \sum_{i_1=0}^{a} \sum_{i_2=0}^{b} \sum_{i_3=0}^{i_1+b-i_2} \sum_{i_4=0}^{i_2+c} \sum_{i_5=0}^{i_3+1} \sum_{i_6=0}^{i_4+1} (-1)^{i_1+i_2} \\ &\times \frac{\xi^{i_3+i_4} \delta^{i_1+b-i_3-i_4+i_5+i_6}}{(\xi+\delta)^{i_1}} \binom{a}{i_1} \binom{b}{i_2} \binom{i_1+b-i_2}{i_3} \binom{i_2+c}{i_4} \binom{i_3+1}{i_5} \\ &\times \binom{i_4+1}{i_6} \frac{1}{[\xi(i_2+c+1)]^{\frac{q+\tau(i_6+1)}{\tau}}} \int_0^\infty x^{p+\tau(i_5+1)-1} e^{-\xi(i_1+b-i_2+1)x^\tau} \\ &\times \Gamma\left(\frac{q+\tau(i_6+1)}{\tau},\ \xi(i_2+c+1)x^\tau\right) dx, \end{aligned}$$

By using the relation

$$\Gamma(p,\ y) = (p-1)!\, e^{-y} \sum_{l=0}^{p-1} \frac{y^l}{l!}.$$

$$
\begin{aligned}
L(a,b,c,p,q) &= \frac{\tau^2\xi^4}{(\xi+\delta)^{b+c+2}}\sum_{i_1=0}^{a}\sum_{i_2=0}^{b}\sum_{i_3=0}^{i_1+b-i_2}\sum_{i_4=0}^{i_2+c}\sum_{i_5=0}^{i_3+1}\sum_{i_6=0}^{i_4+1}\sum_{i_7=0}^{\frac{q+\tau i_6}{\tau}}(-1)^{i_1+i_2}\binom{a}{i_1}\binom{b}{i_2}\\
&\times \frac{\xi^{i_3+i_4}\delta^{i_1+b-i_3-i_4+i_5+6}}{(\xi+\delta)^{i_1}}\binom{i_1+b-i_2}{i_3}\binom{i_2+c}{i_4}\binom{i_3+1}{i_5}\binom{i_4+1}{i_6}\\
&\times \frac{\left(\frac{q+\tau i_6}{\tau}\right)!}{[\xi(i_2+c+1)]^{\frac{q+\tau(i_6+1)}{\tau}}}\frac{[\xi(i_2+c+1)]^{i_7}}{i_7!}\\
&\times \int_0^{\infty}x^{p+\tau(i_5+i_7+1)-1}e^{-\xi(i_1+b+c+2)x^{\tau}}dx\\
&= \frac{\xi^4}{(\xi+\delta)^{b+c+2}}\sum_{i_1=0}^{a}\sum_{i_2=0}^{b}\sum_{i_3=0}^{i_1+b-i_2}\sum_{i_4=0}^{i_2+c}\sum_{i_5=0}^{i_3+1}\sum_{i_6=0}^{i_4+1}\sum_{i_7=0}^{\frac{q+\tau i_6}{\tau}}(-1)^{i_1+i_2}\binom{a}{i_1}\binom{b}{i_2}\\
&\times \frac{\xi^{i_3+i_4}\delta^{i_1+b-i_3-i_4+i_5+i_6}}{(\xi+\delta)^{i_1}}\binom{i_1+b-i_2}{i_3}\binom{i_2+c}{i_4}\binom{i_3+1}{i_5}\binom{i_4+1}{i_6}\\
&\times \frac{\left(\frac{q+\tau i_6}{\tau}\right)!}{[\xi(i_2+c+1)]^{\frac{q+\tau(i_6+1)}{\tau}}}\frac{[\xi(i_2+c+1)]^{i_7}}{i_7!}\\
&\times \frac{1}{[\xi(i_1+b+c+2)]^{\frac{p+\tau(i_5+i_7+1)}{\tau}}}\int_0^{\infty}z^{\frac{p+\tau(i_5+i_7)}{\tau}}e^{-z}dz,
\end{aligned}
$$

where $z=\xi(i_1+b+c+2)x^{\tau}$.

2.2. Single moments

Here, we present expressions for the single moments of rth order statistics, $E\left(X_{r:n}^{(p)}\right)=\mu_{r:n}^{(p)}$ from the EPL distribution because these are very important to calculate the variance and draw the inferential techniques.

Theorem 1. *For the EPL distribution given in (1) and for,* $1\le r\le n$

$$
\begin{aligned}
\mu_{r:n}^{(p)} &= \frac{\xi^2\,C_{r:n}}{(\xi+\delta)^{n-r+1}}\sum_{i_1=0}^{r-1}\sum_{i_2=0}^{i_1+n-r}\sum_{i_3=0}^{i_2+1}(-1)^{i_1}\binom{r-1}{i_1}\binom{i_1+n-r}{i_2}\binom{i_2+1}{i_3}\\
&\times \frac{\xi^{i_2}\delta^{i_1+n-r-i_2+i_3}}{(\xi+\delta)^{i_1}}\frac{\Gamma(\frac{p+\tau(i_3+1)}{\tau})}{[\xi(i_1+n-r+1)]^{\frac{p+\tau(i_3+1)}{\tau}}}. \qquad (7)
\end{aligned}
$$

Proof. By using (5), we get

$$
\mu_{r:n}^{(p)}=C_{r:n}\int_0^{\infty}x^p\,F^{r-1}(x)\,[1-F(x)]^{n-r}\,f(x)dx. \qquad (8)
$$

The result follows from applying lemma 1.

Special Cases

i For $\tau = 1$ and $\delta = 1$ in (7), we obtain relation for order statistic of the Lindley distribution

$$\begin{aligned}\mu_{r:n}^{(p)} &= \frac{\xi^2\, C_{r:n}}{(\xi+1)^{n-r+1}} \sum_{i_1=0}^{r-1} \sum_{i_2=0}^{i_1+n-r} \sum_{i_3=0}^{i_2+1} (-1)^{i_1} \binom{r-1}{i_1} \binom{i_1+n-r}{i_2} \binom{i_2+1}{i_3} \\ &\times \frac{\xi^{i_2}}{(\xi+1)^{i_1}} \frac{\Gamma(p+i_3+1)}{[\xi(i_1+n-r+1)]^{p+i_3+1}},\end{aligned}$$

as obtained by Sultan and AL-Thubyani (2016).

ii For $\delta = 1$ in (7), we obtain relation for order statistic of the power Lindley distribution

$$\begin{aligned}\mu_{r:n}^{(p)} &= \frac{\xi^2\, C_{r:n}}{(\xi+1)^{n-r+1}} \sum_{i_1=0}^{r-1} \sum_{i_2=0}^{i_1+n-r} \sum_{i_3=0}^{i_2+1} (-1)^{i_1} \binom{r-1}{i_1} \binom{i_1+n-r}{i_2} \binom{i_2+1}{i_3} \\ &\times \frac{\xi^{i_2}}{(\xi+1)^{i_1}} \frac{\Gamma\frac{(p+\tau(i_3+1))}{\tau}}{[\xi(i_1+n-r+1)]^{\frac{p+\tau(i_3+1)}{\tau}}},\end{aligned}$$

as obtained by Kumar and Goyal (2019a).

iii If $p = n = r = 1$ in (7), we unordered mean of the extended power Lindley distribution

$$\mu_{1:1}^{(1)} = \frac{\xi^2}{(\xi+\delta)} \sum_{i_3=0}^{1} \frac{\delta^{i_3}\, \Gamma(\frac{\tau(i_3+1)+1}{\tau})}{\xi^{\frac{\tau(i_3+1)+1}{\tau}}} = E(X),$$

as obtained by Alkarni (2015).

In particular, the first moment (mean) of the r-th order statistic is

$$\mu_{r:n} = \mu_{r:n}^{(1)} = \vartheta(\tau,\xi,\delta,r,n,1),$$

where

$$\begin{aligned}\vartheta(\tau,\xi,\delta,r,n,k) &= \frac{\xi^2\, C_{r:n}}{(\xi+\delta)^{n-r+1}} \sum_{i_1=0}^{r-1} \sum_{i_2=0}^{i_1+n-r} \sum_{i_3=0}^{i_2+1} (-1)^{i_1} \binom{r-1}{i_1} \binom{i_1+n-r}{i_2} \binom{i_2+1}{i_3} \\ &\times \frac{\xi^{i_2}\delta^{i_1+n-r-i_2+i_3}}{(\xi+\delta)^{i_1}} \frac{\Gamma(\frac{k+\tau(i_3+1)}{\tau})}{[\xi(i_1+n-r+1)]^{\frac{k+\tau(i_3+1)}{\tau}}}.\end{aligned}$$

In addition, the variance of $X_{r:n}$ is found to be

$$\sigma_{r:n}^2 = \mu_{r:n}^{(2)} - [\mu_{r:n}^{(1)}]^2 = \vartheta(\tau,\xi,\delta,r,n,2) - [\vartheta(\tau,\xi,\delta,r,n,1)]^2.$$

2.3. Double moments

Here, we provide relation for the double moment of rth and sth order statistics, $E\left(X_{r,s:n}^{(p,q)}\right)=\mu_{r,s:n}^{(p,q)}$ from the EPL distribution, which are useful to calculate the variances, covariances, BLUEs and other inferential techniques.

Theorem 2. *For the EPL distribution given in (1) and for* $1\leq r<s\leq n$

$$\begin{aligned}\mu_{r,s:n}^{(p,q)} &= C_{r,s:n}\frac{\xi^4}{(\xi+\delta)^{n-r+1}}\sum_{i_1=0}^{r-1}\sum_{i_2=0}^{s-r-1}\sum_{i_3=0}^{i_1+s-r-i_2-1}\sum_{i_4=0}^{i_2+n-s}\sum_{i_5=0}^{i_3+1}\sum_{i_6=0}^{i_4+1}\sum_{i_7=0}^{\frac{q+\tau i_6}{\tau}}(-1)^{i_1+i_2}\\ &\times \frac{\xi^{i_3+i_4}\delta^{i_1+s-r-1-i_3-i_4+i_5+i_6}}{(\xi+\delta)^{i_1}}\binom{r-1}{i_1}\binom{s-r-1}{i_2}\binom{i_1+s-r-i_2-1}{i_3}\\ &\times \binom{i_2+n-s}{i_4}\binom{i_3+1}{i_5}\binom{i_4+1}{i_6}\frac{\left(\frac{q+\tau i_6}{\tau}\right)!}{[\xi(i_2+n-s+1)]^{\frac{q+\tau(i_6+1)}{\tau}}}\\ &\times \frac{[\xi(i_2+n-s+1)]^{i_7}}{i_7!}\frac{\Gamma\left(\frac{p+\tau(i_5+i_7+1)}{\tau}\right)}{[\xi(i_1+n-r+1)]^{\frac{p+\tau(i_5+i_7+1)}{\tau}}}.\end{aligned}$$

Proof. By using (6), we obtain

$$\mu_{r,s:n}^{(p,q)}=C_{r,s:n}\int_0^\infty\int_x^\infty x^p\,y^q\,F^{r-1}(x)\,[F(y)-F(x)]^{s-1-r}\,[1-F(y)]^{n-s}\,f(x)f(y)dxdy \qquad (9)$$

The result follows from applying lemma 2.

Special Cases

i For $\tau=1$ and $\delta=1$ in Theorem 2, we obtain the relation for order statistic of the Lindley distribution

$$\begin{aligned}\mu_{r,s:n}^{(p,q)} &= \frac{\xi^4 C_{r,s:n}}{(\xi+1)^{n-r+1}}\sum_{i_1=0}^{r-1}\sum_{i_2=0}^{s-r-1}\sum_{i_3=0}^{i_1+s-r-i_2-1}\sum_{i_4=0}^{i_2+n-s}\sum_{i_5=0}^{i_3+1}\sum_{i_6=0}^{i_4+1}\sum_{i_7=0}^{q+i_6}(-1)^{i_1+i_2}\frac{\xi^{i_3+i_4}}{(\xi+1)^{i_1}}\\ &\times \binom{r-1}{i_1}\binom{s-r-1}{i_2}\binom{i_1+s-r-i_2-1}{i_3}\binom{i_2+n-s}{i_4}\binom{i_3+1}{i_5}\binom{i_4+1}{i_6}\\ &\times \frac{(q+i_6)!}{[\xi(i_2+n-s+1)]^{q+i_6+1}}\frac{[\xi(i_2+n-s+1)]^{i_7}}{i_7!}\frac{\Gamma(p+i_5+i_7+1)}{[\xi(i_1+n-r+1)]^{p+i_5+i_7+1}}.\end{aligned}$$

as obtained by Sultan and AL-Thubyani (2016).

ii For $\delta = 1$ in Theorem 2, we get the explicit expression for order statistic of the power Lindley distribution

$$\begin{aligned}\mu_{r,s:n}^{(p,q)} &= \frac{\xi^4 C_{r,s:n}}{(\xi+1)^{n-r+1}} \sum_{i_1=0}^{r-1} \sum_{i_2=0}^{s-r-1} \sum_{i_3=0}^{i_1+s-r-i_2-1} \sum_{i_4=0}^{i_2+n-s} \sum_{i_5=0}^{i_3+1} \sum_{i_6=0}^{i_4+1} \sum_{i_7=0}^{\frac{q+\tau i_6}{\tau}} (-1)^{i_1+i_2} \frac{\xi^{i_3+i_4}}{(\xi+1)^{i_1}} \\ &\times \binom{r-1}{i_1}\binom{s-r-1}{i_2}\binom{i_1+s-r-i_2-1}{i_3}\binom{i_2+n-s}{i_4}\binom{i_3+1}{i_5}\binom{i_4+1}{i_6} \\ &\times \frac{\left(\frac{q+\tau i_6}{\tau}\right)!}{[\xi(i_2+n-s+1)]^{\frac{q+\tau(i_6+1)}{\tau}}} \frac{[\xi(i_2+n-s+1)]^{i_7}}{i_7!} \frac{\Gamma\left(\frac{p+\tau(i_5+i_7+1)}{\tau}\right)}{[\xi(i_1+n-r+1)]^{\frac{p+\tau(i_5+i_7+1)}{\tau}}}.\end{aligned}$$

as obtained by Kumar and Goyal (2019a).

For simplicity, we denote the (1, 1)-th moment of $X_{r:n}$ and $X_{s:n}$, which are also called the simple product moment of these order statistics, by $\mu_{r,s:n}$. The simple product moments are used for evaluating the covariances, in other words

$$\sigma_{r,s:n} = Cov(X_{r:n},\, X_{s:n}) = \mu_{r,s:n} - \mu_{r:n}\mu_{s:n}.$$

3. Location and scale parameter estimation based on Type-II right-censored sample

In this section we have obtained the BLUEs of the location and scale parameters to approximate population parameters. Let $Y_{1:n} \le Y_{2:n} \le \cdots \le Y_{n-c:n},\ c = 0, 1, \ldots, n-1$ denote Type-II right-censored sample from the location-scale parameter extended power Lindley distribution in Equation (1). Let us denote $Z_{r:n} = (Y_{r:n} - \mu)/\sigma,\ E(Z_{r:n}) = \mu_{r:n}^{(1)},\ 1 \le r \le (n-c)$, and $Cov(Z_{r:n},\, Z_{s:n}) = \sigma_{r,s:n} = \mu_{r,s:n}^{(1,1)} - \mu_{r:n}^{(1)}\mu_{s:n}^{(1)},\ 1 \le r < s \le (n-c)$. We shall use the following notations:

$$\mathbf{Y} = (Y_{1:n}, Y_{2:n}, \ldots, Y_{(n-c):n})^T,$$

$$\mu = (\mu_{1:n}, \mu_{2:n}, \ldots, \mu_{(n-c):n})^T,$$

$$\mathbf{1} = \underbrace{(1,\ 1,\ \ldots,\ 1)^T}_{n-c},$$

and

$$\sum = ((\sigma_{r,s})),\ \ 1 \le r,\ \ s \le n-c.$$

The BLUEs of μ and σ are given by Arnold *et al*. (2003)

$$\mu^* = \left\{ \frac{\mu^{\mathbf{T}}\Sigma^{-1}\mu\mathbf{1}^{\mathbf{T}}\Sigma^{-1} - \mu^{\mathbf{T}}\Sigma^{-1}\mathbf{1}\mu^{\mathbf{T}}\Sigma^{-1}}{(\mu^{\mathbf{T}}\Sigma^{-1}\mu)(\mathbf{1}^{\mathbf{T}}\Sigma^{-1}\mathbf{1}) - (\mu^{\mathbf{T}}\Sigma^{-1}\mathbf{1})^2} \right\} \; Y_{r:n} = \sum_{r=1}^{n-c} a_r Y_{r:n}, \tag{10}$$

$$\sigma^* = \left\{ \frac{\mathbf{1}^{\mathbf{T}}\Sigma^{-1}\mathbf{1}\mu^{\mathbf{T}}\Sigma^{-1} - \mathbf{1}^{\mathbf{T}}\Sigma^{-1}\mu\mathbf{1}^{\mathbf{T}}\Sigma^{-1}}{(\mu^{\mathbf{T}}\Sigma^{-1}\mu)(\mathbf{1}^{\mathbf{T}}\Sigma^{-1}\mathbf{1}) - (\mu^{\mathbf{T}}\Sigma^{-1}\mathbf{1})^2} \right\} \; Y_{r:n} = \sum_{r=1}^{n-c} b_r Y_{r:n}. \tag{11}$$

Furthermore, the variances and covariance of these BLUEs are given by Arnold *et al*. (2003)

$$Var(\mu^*) = \sigma^2 \left\{ \frac{\mu^{\mathbf{T}}\Sigma^{-1}\mu}{(\mu^{\mathbf{T}}\Sigma^{-1}\mu)(\mathbf{1}^{\mathbf{T}}\Sigma^{-1}\mathbf{1}) - (\mu^{\mathbf{T}}\Sigma^{-1}\mathbf{1})^2} \right\} \; \sigma^2 V_1, \tag{12}$$

$$Var(\sigma^*) = \sigma^2 \left\{ \frac{\mathbf{1}^{\mathbf{T}}\Sigma^{-1}\mathbf{1}}{(\mu^{\mathbf{T}}\Sigma^{-1}\mu)(\mathbf{1}^{\mathbf{T}}\Sigma^{-1}\mathbf{1}) - (\mu^{\mathbf{T}}\Sigma^{-1}\mathbf{1})^2} \right\} \; \sigma^2 V_2, \tag{13}$$

and

$$Cov(\mu^*,\, \sigma^*) = \sigma^2 \left\{ \frac{-\mu^{\mathbf{T}}\Sigma^{-1}\mathbf{1}}{(\mu^{\mathbf{T}}\Sigma^{-1}\mu)(\mathbf{1}^{\mathbf{T}}\Sigma^{-1}\mathbf{1}) - (\mu^{\mathbf{T}}\Sigma^{-1}\mathbf{1})^2} \right\} \; \sigma^2 V_3. \tag{14}$$

4. Numerical Results

The relations obtained in the preceding sections allow us to evaluate the expected values, second moments, variances, product moment and covariances of order statistics from samples of sizes up to 10 for various values of the parameters. The relation in (7) can be used to compute the expected values, second order moments and variances of all order statistics for sample sizes $n = 1, 2, 3, \ldots, 10$. In Tables 1, we have presented expected values, second order moments, variances, skewness and kurtosis of the rth order statistics from the EPL distribution for $n = 1, 2, 3, \ldots, 10$ and $\tau = 2$, $\delta = 0.5$ and $\xi = 5$. One can see that the means and variances are decreasing both with respect to n and ξ. In Table 2 we have reported the double moments and covariances of the rth and sth order statistics from the EPL distribution for $n = 1, 2, 3, \ldots, 10$ and $\tau = 2$, $\delta = 0.5$ and $\xi = 5, 10$. From Table 3, one can observe that product moments are decreasing with respect to n.

The coefficients of the BLUEs for Type-II right-censored samples of sample sizes $n =$ 7, 10, $\tau = 2$, $\delta = 0.5$ and $\xi = 5$ and different censoring cases $c = 0(1)([n/2] - 1)$ are displayed in Tables 3 and 4. Also, the variances and covariances of the BLUEs are provided in Table 5.

Table 1: Moments, variances skewness and kurtosis of order statistic from EPL distribution $\tau = 2$, $\delta = 0.5$ and $\xi = 5$

r	n	$E(X)$	$E(X^2)$	$V(X)$	δ_1	δ_2	γ_1	γ_2
1	1	0.414348	0.218182	0.046498	0.379571	3.207079	0.616093	0.207079
	2	0.293422	0.109504	0.023408	0.387238	3.219388	0.622284	0.219388
	3	0.239709	0.073106	0.015646	0.389946	3.229529	0.624457	0.229529
	4	0.207652	0.054870	0.011751	0.391543	3.234472	0.625734	0.234472
	5	0.185762	0.043916	0.009408	0.392307	3.236857	0.626344	0.236857
	6	0.169597	0.036608	0.007845	0.394587	3.229235	0.628161	0.229235
	7	0.157029	0.031385	0.006727	0.395586	3.221588	0.628956	0.221588
	8	0.146897	0.027467	0.005888	0.395741	3.229525	0.629079	0.229525
	9	0.138503	0.024418	0.005235	0.394610	3.235330	0.628180	0.235330
	10	0.131400	0.021978	0.004712	0.396862	3.226521	0.629970	0.226521
2	2	0.535273	0.326860	0.040343	0.241593	3.215795	0.491521	0.215795
	3	0.400850	0.182301	0.021620	0.189444	3.115636	0.435251	0.115636
	4	0.335877	0.127814	0.015001	0.174957	3.088871	0.418279	0.088871
	5	0.295213	0.098686	0.011535	0.170039	3.070969	0.412358	0.070969
	6	0.266589	0.080456	0.009386	0.167111	3.066828	0.408792	0.066828
	7	0.245001	0.067944	0.007919	0.165394	3.065728	0.406687	0.065728
	8	0.227956	0.058815	0.006851	0.164916	3.061962	0.406099	0.061962
	9	0.214051	0.051856	0.006038	0.164358	3.069852	0.405411	0.069852
	10	0.202423	0.046373	0.005398	0.164208	3.048848	0.405226	0.048848
3	3	0.602485	0.399139	0.036151	0.225634	3.273408	0.475009	0.273408
	4	0.465823	0.236788	0.019797	0.145361	3.125628	0.381262	0.125628
	5	0.396873	0.171505	0.013997	0.123096	3.072745	0.350850	0.072745
	6	0.352462	0.135146	0.010917	0.112742	3.054180	0.335770	0.054180
	7	0.320559	0.111736	0.008978	0.108337	3.051939	0.329145	0.051939
	8	0.296135	0.095332	0.007636	0.105760	3.035578	0.325207	0.035578
	9	0.276626	0.083171	0.006649	0.105379	3.008423	0.324621	0.008423
	10	0.260564	0.073785	0.005891	0.104265	3.025598	0.322900	0.025598
4	4	0.648039	0.453256	0.033301	0.229320	3.321976	0.478874	0.321976
	5	0.511790	0.280311	0.018382	0.130464	3.147984	0.361198	0.147984
	6	0.441284	0.207865	0.013133	0.102281	3.090069	0.319815	0.090069
	7	0.394999	0.166359	0.010335	0.090404	3.056551	0.300673	0.056551
	8	0.361267	0.139076	0.008562	0.084264	3.040724	0.290282	0.040724
	9	0.335151	0.119652	0.007326	0.081505	3.010495	0.285491	0.010495
	10	0.314105	0.105073	0.006411	0.078213	3.023921	0.279667	0.023921
5	5	0.682101	0.496492	0.031230	0.237912	3.362265	0.487762	0.362265
	6	0.547043	0.316534	0.017278	0.124926	3.165975	0.353448	0.165975
	7	0.475998	0.238994	0.012420	0.092794	3.093716	0.304621	0.093716
	8	0.428730	0.193642	0.009833	0.078295	3.075215	0.279812	0.075215
	9	0.393912	0.163355	0.008188	0.071074	3.053238	0.266597	0.053238
	10	0.366721	0.141522	0.007038	0.067297	3.019226	0.259417	0.019226
6	6	0.709112	0.532483	0.029643	0.247927	3.393989	0.497922	0.393989
	7	0.575461	0.347549	0.016394	0.123722	3.179939	0.351742	0.179939
	8	0.504359	0.266206	0.011828	0.087950	3.098604	0.296564	0.098604
	9	0.456584	0.217871	0.009402	0.072183	3.067347	0.268669	0.067347
	10	0.421103	0.185187	0.007859	0.063867	3.048346	0.252719	0.048346
7	7	0.731387	0.563306	0.028379	0.257507	3.423281	0.507451	0.423281
	8	0.599162	0.374664	0.015668	0.123781	3.197735	0.351825	0.197735
	9	0.528246	0.290374	0.011330	0.081133	3.116019	0.284838	0.116019
	10	0.480238	0.239660	0.009031	0.067515	3.077528	0.259836	0.077528
8	8	0.750277	0.590255	0.027339	0.267491	3.444852	0.517195	0.444852
	9	0.619424	0.398746	0.015060	0.125692	3.209343	0.354531	0.209343
	10	0.548821	0.312108	0.010904	0.083717	3.116898	0.289339	0.116898
9	9	0.766633	0.614193	0.026467	0.276388	3.468726	0.525726	0.468726
	10	0.637074	0.420406	0.014543	0.127287	3.217234	0.356773	0.217234
10	10	0.781029	0.635725	0.025719	0.285348	3.482843	0.534180	0.482843

Table 2: Covariances of order statistics

			$\tau = 2, \xi = 5, \delta = 0.5$				$\tau = 2, \xi = 10, \delta = 0.5$			
s	r	n	$\mu_{r,s:n}$	$E(X_{r:n})$	$E(X_{s:n})$	$\sigma_{r,s:n}$	$\mu_{r,s:n}$	$E(X_{r:n})$	$\mu_{r,s:n}$	$\sigma_{r,s:n}$
2	1	2	0.125100	0.293422	0.535273	-0.03196	0.059192	0.202969	0.370875	-0.01608
		3	0.073220	0.239709	0.400850	-0.02287	0.034402	0.165748	0.277412	-0.01158
		4	0.052133	0.207652	0.335877	-0.01761	0.024407	0.143552	0.232333	-0.00895
		5	0.040552	0.185762	0.295213	-0.01429	0.018943	0.128403	0.204150	-0.00727
		6	0.033203	0.169597	0.266589	-0.01201	0.015488	0.117219	0.184323	-0.00612
		7	0.028119	0.157029	0.245001	-0.01035	0.013102	0.108526	0.169376	-0.00528
		8	0.024389	0.146897	0.227956	-0.00910	0.011355	0.101518	0.157579	-0.00464
		9	0.021534	0.138503	0.214051	-0.00811	0.010020	0.095714	0.147956	-0.00414
		10	0.019279	0.131400	0.202423	-0.00732	0.008966	0.090803	0.139911	-0.00374
3	1	3	0.146966	0.239709	0.602485	0.002545	0.066242	0.165748	0.417607	-0.00298
		4	0.100791	0.207652	0.465823	0.004062	0.044251	0.143552	0.322490	-0.00204
		5	0.079630	0.185762	0.396873	0.005906	0.034146	0.128403	0.274609	-0.00112
		6	0.067164	0.169597	0.352462	0.007387	0.028186	0.117219	0.243803	-0.00039
		7	0.058882	0.157029	0.320559	0.008545	0.024224	0.108526	0.221690	0.000165
		8	0.052962	0.146897	0.296135	0.009461	0.021392	0.101518	0.204769	0.000604
		9	0.048512	0.138503	0.276626	0.010199	0.019262	0.095714	0.191258	0.000956
		10	0.045043	0.131400	0.260564	0.010805	0.017601	0.090803	0.180137	0.001244
3	2	3	0.155112	0.400850	0.602485	-0.08639	0.076932	0.277412	0.417607	-0.03892
		4	0.087823	0.335877	0.465823	-0.06864	0.044543	0.232333	0.322490	-0.03038
		5	0.059381	0.295213	0.396873	-0.05778	0.031058	0.204150	0.274609	-0.02500
		6	0.043331	0.266589	0.352462	-0.05063	0.023522	0.184323	0.243803	-0.02142
		7	0.032948	0.245001	0.320559	-0.04559	0.018679	0.169376	0.221690	-0.01887
		8	0.025656	0.227956	0.296135	-0.04185	0.015295	0.157579	0.204769	-0.01697
		9	0.020245	0.214051	0.276626	-0.03897	0.012794	0.147956	0.191258	-0.01550
		10	0.016067	0.202423	0.260564	-0.03668	0.010868	0.139911	0.180137	-0.01434
4	1	4	0.157123	0.207652	0.648039	0.022556	0.067875	0.143552	0.449313	0.003375
		5	0.116281	0.185762	0.511790	0.021210	0.048633	0.128403	0.354410	0.003126
		6	0.096719	0.169597	0.441284	0.021879	0.039214	0.117219	0.305414	0.003413
		7	0.084955	0.157029	0.394999	0.022929	0.033425	0.108526	0.273287	0.003766
		8	0.077118	0.146897	0.361267	0.024049	0.029472	0.101518	0.249893	0.004104
		9	0.071578	0.138503	0.335151	0.025159	0.026598	0.095714	0.231790	0.004413
		10	0.067511	0.131400	0.314105	0.026238	0.024416	0.090803	0.217207	0.004693
	2	4	0.172829	0.335877	0.648039	-0.04483	0.084966	0.232333	0.449313	-0.01942
		5	0.112133	0.295213	0.511790	-0.03895	0.055696	0.204150	0.354410	-0.01666
		6	0.082851	0.266589	0.441284	-0.03479	0.041888	0.184323	0.305414	-0.01441
		7	0.064709	0.245001	0.394999	-0.03207	0.033554	0.169376	0.273287	-0.01273
		8	0.052012	0.227956	0.361267	-0.03034	0.027892	0.157579	0.249893	-0.01149
		9	0.042428	0.214051	0.335151	-0.02931	0.023756	0.147956	0.231790	-0.01054
		10	0.034802	0.202423	0.314105	-0.02878	0.020579	0.139911	0.217207	-0.00981
	3	4	0.179898	0.465823	0.648039	-0.12197	0.089110	0.322490	0.449313	-0.05579
		5	0.104111	0.396873	0.511790	-0.09900	0.052452	0.274609	0.354410	-0.04487
		6	0.071720	0.352462	0.441284	-0.08382	0.036947	0.243803	0.305414	-0.03751
		7	0.053410	0.320559	0.394999	-0.07321	0.028192	0.221690	0.273287	-0.03239
		8	0.041644	0.296135	0.361267	-0.06534	0.022532	0.204769	0.249893	-0.02864
		9	0.033502	0.276626	0.335151	-0.05921	0.018569	0.191258	0.231790	-0.02576
		10	0.027592	0.260564	0.314105	-0.05425	0.015641	0.180137	0.217207	-0.02349
5	1	5	0.165311	0.185762	0.682101	0.038602	0.068527	0.128403	0.473039	0.007787
		6	0.128166	0.169597	0.547043	0.035389	0.051311	0.117219	0.378909	0.006896
		7	0.110052	0.157029	0.475998	0.035306	0.042581	0.108526	0.329508	0.006820
		8	0.099086	0.146897	0.428730	0.036107	0.037070	0.101518	0.296682	0.006952
		9	0.091832	0.138503	0.393912	0.037274	0.033237	0.095714	0.272522	0.007153
		10	0.086817	0.131400	0.366721	0.038630	0.030414	0.090803	0.253665	0.007380
	2	5	0.173400	0.295213	0.682101	-0.02796	0.085162	0.204150	0.473039	-0.01141
		6	0.119750	0.266589	0.547043	-0.02609	0.059439	0.184323	0.378909	-0.01040
		7	0.092013	0.245001	0.475998	-0.02461	0.046480	0.169376	0.329508	-0.00933
		8	0.073911	0.227956	0.428730	-0.02382	0.038296	0.157579	0.296682	-0.00846

Table 2: Continued.

			$\tau = 2, \xi = 5, \delta = 0.5$				$\tau = 2, \xi = 10, \delta = 0.5$			
s	r	n	$\mu_{r,s:n}$	$E(X_{r:n})$	$E(X_{s:n})$	$\sigma_{r,s:n}$	$\mu_{r,s:n}$	$E(X_{r:n})$	$\mu_{r,s:n}$	$\sigma_{r,s:n}$
		9	0.060674	0.214051	0.393912	-0.02364	0.032550	0.147956	0.272522	-0.00777
		10	0.050264	0.202423	0.366721	-0.02397	0.028244	0.139911	0.253665	-0.00725
	3	5	0.202606	0.396873	0.682101	-0.06810	0.099405	0.274609	0.473039	-0.03050
		6	0.133799	0.352462	0.547043	-0.05901	0.065762	0.243803	0.378909	-0.02662
		7	0.100903	0.320559	0.475998	-0.05168	0.049799	0.221690	0.329508	-0.02325
		8	0.080899	0.296135	0.428730	-0.04606	0.040135	0.204769	0.296682	-0.02062
		9	0.067311	0.276626	0.393912	-0.04166	0.033574	0.191258	0.272522	-0.01855
		10	0.057471	0.260564	0.366721	-0.03808	0.028798	0.180137	0.253665	-0.01690
	4	5	0.197590	0.511790	0.682101	-0.15150	0.097898	0.354410	0.473039	-0.06975
		6	0.115810	0.441284	0.547043	-0.12559	0.058351	0.305414	0.378909	-0.05737
		7	0.080302	0.394999	0.475998	-0.10772	0.041419	0.273287	0.329508	-0.04863
		8	0.059934	0.361267	0.428730	-0.09495	0.031757	0.249893	0.296682	-0.04238
		9	0.046659	0.335151	0.393912	-0.08536	0.025457	0.231790	0.272522	-0.03771
		10	0.037334	0.314105	0.366721	-0.07786	0.021014	0.217207	0.253665	-0.03408
6	1	6	0.173540	0.169597	0.709112	0.053277	0.069056	0.117219	0.491865	0.011400
		7	0.138937	0.157029	0.575461	0.048573	0.053325	0.108526	0.398669	0.010059
		8	0.122004	0.146897	0.504359	0.047915	0.045185	0.101518	0.349204	0.009735
		9	0.111775	0.138503	0.456584	0.048537	0.039958	0.095714	0.316010	0.009712
		10	0.105094	0.131400	0.421103	0.049761	0.036274	0.090803	0.291379	0.009816
	2	6	0.169539	0.266589	0.709112	-0.01950	0.083624	0.184323	0.491865	-0.00704
		7	0.121311	0.245001	0.575461	-0.01968	0.060716	0.169376	0.398669	-0.00681
		8	0.095139	0.227956	0.504359	-0.01983	0.048693	0.157579	0.349204	-0.00633
		9	0.077345	0.214051	0.456584	-0.02039	0.040853	0.147956	0.316010	-0.00590
		10	0.063859	0.202423	0.421103	-0.02138	0.035208	0.139911	0.291379	-0.00556
	3	6	0.206018	0.352462	0.709112	-0.04392	0.100332	0.243803	0.491865	-0.01959
		7	0.145146	0.320559	0.575461	-0.03932	0.070483	0.221690	0.398669	-0.01790
		8	0.114479	0.296135	0.504359	-0.03488	0.055437	0.204769	0.349204	-0.01607
		9	0.095232	0.276626	0.456584	-0.03107	0.045952	0.191258	0.316010	-0.01449
		10	0.081919	0.260564	0.421103	-0.02781	0.039327	0.180137	0.291379	-0.01316
	4	6	0.223267	0.441284	0.709112	-0.08965	0.110002	0.305414	0.491865	-0.04022
		7	0.148164	0.394999	0.575461	-0.07914	0.073256	0.273287	0.398669	-0.03569
		8	0.111855	0.361267	0.504359	-0.07035	0.055703	0.249893	0.349204	-0.03156
		9	0.089461	0.335151	0.456584	-0.06356	0.045006	0.231790	0.316010	-0.02824
		10	0.073972	0.314105	0.421103	-0.05830	0.037697	0.217207	0.291379	-0.02559
	5	6	0.211616	0.547043	0.709112	-0.17630	0.104759	0.378909	0.491865	-0.08161
		7	0.125475	0.475998	0.575461	-0.14844	0.063092	0.329508	0.398669	-0.06827
		8	0.087690	0.428730	0.504359	-0.12854	0.045094	0.296682	0.349204	-0.05851
		9	0.065827	0.393912	0.456584	-0.11403	0.034744	0.272522	0.316010	-0.05138
		10	0.051487	0.366721	0.421103	-0.10294	0.027951	0.253665	0.291379	-0.04596
7	1	7	0.182405	0.157029	0.731387	0.067556	0.069687	0.108526	0.507398	0.014621
		8	0.149550	0.146897	0.599162	0.061535	0.055063	0.101518	0.415157	0.012917
		9	0.133567	0.138503	0.528246	0.060403	0.047410	0.095714	0.365801	0.012398
		10	0.124002	0.131400	0.480238	0.060899	0.042442	0.090803	0.332431	0.012256
	2	7	0.163819	0.245001	0.731387	-0.01537	0.081635	0.169376	0.507398	-0.00431
		8	0.119737	0.227956	0.599162	-0.01685	0.060918	0.157579	0.415157	-0.00450
		9	0.094876	0.214051	0.528246	-0.01820	0.049741	0.147956	0.365801	-0.00438
		10	0.077366	0.202423	0.480238	-0.01984	0.042276	0.139911	0.332431	-0.00423
	3	7	0.205092	0.320559	0.731387	-0.02936	0.099057	0.221690	0.507398	-0.01343
		8	0.150629	0.296135	0.599162	-0.02680	0.072334	0.204769	0.415157	-0.01268
		9	0.122355	0.276626	0.528246	-0.02377	0.058356	0.191258	0.365801	-0.01161
		10	0.104276	0.260564	0.480238	-0.02086	0.049294	0.180137	0.332431	-0.01059
	4	7	0.227234	0.320559	0.731387	-0.00722	0.111556	0.273287	0.507398	-0.02711
		8	0.160109	0.296135	0.599162	-0.01732	0.078663	0.249893	0.415157	-0.02508
		9	0.125850	0.276626	0.528246	-0.02028	0.062003	0.231790	0.365801	-0.02279
		10	0.103939	0.260564	0.480238	-0.02119	0.051445	0.217207	0.332431	-0.02076

Table 2: Continued.

			$\tau = 2, \xi = 5, \delta = 0.5$				$\tau = 2, \xi = 10, \delta = 0.5$			
s	r	n	$\mu_{r,s:n}$	$E(X_{r:n})$	$E(X_{s:n})$	$\sigma_{r,s:n}$	$\mu_{r,s:n}$	$E(X_{r:n})$	$\mu_{r,s:n}$	$\sigma_{r,s:n}$
	5	7	0.240059	0.475998	0.731387	-0.10808	0.118411	0.329508	0.507398	-0.04878
		8	0.160346	0.428730	0.599162	-0.09653	0.079330	0.296682	0.415157	-0.04384
		9	0.121653	0.393912	0.528246	-0.08643	0.060576	0.272522	0.365801	-0.03911
		10	0.097711	0.366721	0.480238	-0.07840	0.049095	0.253665	0.332431	-0.03523
	6	7	0.223156	0.575461	0.731387	-0.19773	0.110362	0.398669	0.507398	-0.09192
		8	0.133614	0.504359	0.599162	-0.16858	0.067043	0.349204	0.415157	-0.07793
		9	0.094015	0.456584	0.528246	-0.14717	0.048208	0.316010	0.365801	-0.06739
		10	0.070923	0.421103	0.480238	-0.13131	0.037307	0.291379	0.332431	-0.05956
8	1	8	0.192134	0.146897	0.750277	0.081920	0.070480	0.101518	0.520575	0.017632
		9	0.160463	0.138503	0.619424	0.074671	0.056695	0.095714	0.429259	0.015609
		10	0.145243	0.131400	0.548821	0.073128	0.049442	0.090803	0.380103	0.014928
	2	8	0.156885	0.227956	0.750277	-0.01415	0.079554	0.157579	0.520575	-0.00248
		9	0.116038	0.214051	0.619424	-0.01655	0.060572	0.147956	0.429259	-0.00294
		10	0.092204	0.202423	0.548821	-0.01889	0.050124	0.139911	0.380103	-0.00306
	3	8	0.203197	0.296135	0.750277	-0.01899	0.097193	0.204769	0.520575	-0.00940
		9	0.153848	0.276626	0.619424	-0.01750	0.072961	0.191258	0.429259	-0.00914
		10	0.127820	0.260564	0.548821	-0.01518	0.059979	0.180137	0.380103	-0.00849
	4	8	0.225773	0.361267	0.750277	-0.04528	0.110449	0.249893	0.520575	-0.01964
		9	0.165186	0.335151	0.619424	-0.04241	0.080817	0.231790	0.429259	-0.01868
		10	0.133147	0.314105	0.548821	-0.03924	0.065255	0.217207	0.380103	-0.01731
	5	8	0.245110	0.428730	0.750277	-0.07656	0.120610	0.296682	0.520575	-0.03384
		9	0.173431	0.393912	0.619424	-0.07057	0.085377	0.272522	0.429259	-0.03161
		10	0.136810	0.366721	0.548821	-0.06445	0.067482	0.253665	0.380103	-0.02894
	6	8	0.253982	0.504359	0.750277	-0.12443	0.125347	0.349204	0.520575	-0.05644
		9	0.170589	0.456584	0.619424	-0.11223	0.084412	0.316010	0.429259	-0.05124
		10	0.129956	0.421103	0.548821	-0.10115	0.064702	0.291379	0.380103	-0.04605
	7	8	0.232942	0.504359	0.750277	-0.14547	0.115084	0.415157	0.520575	-0.10104
		9	0.140652	0.456584	0.619424	-0.14217	0.070427	0.365801	0.429259	-0.08660
		10	0.099571	0.421103	0.548821	-0.13154	0.050909	0.332431	0.380103	-0.07545
9	1	9	0.202844	0.138503	0.766633	0.096663	0.071447	0.095714	0.531990	0.020528
		10	0.171939	0.131400	0.637074	0.088228	0.058309	0.090803	0.441548	0.018215
	2	9	0.148835	0.214051	0.766633	-0.01526	0.077490	0.147956	0.531990	-0.00122
		10	0.110573	0.202423	0.637074	-0.01839	0.059909	0.139911	0.441548	-0.00187
	3	9	0.201458	0.276626	0.766633	-0.01061	0.095239	0.191258	0.531990	-0.00651
		10	0.156264	0.260564	0.637074	-0.00973	0.073008	0.180137	0.441548	-0.00653
	4	9	0.222547	0.335151	0.766633	-0.03439	0.108528	0.231790	0.531990	-0.01478
		10	0.167173	0.314105	0.637074	-0.03294	0.081537	0.217207	0.441548	-0.01437
	5	9	0.244147	0.393912	0.766633	-0.05784	0.119779	0.272522	0.531990	-0.02520
		10	0.179218	0.366721	0.637074	-0.05441	0.087879	0.253665	0.441548	-0.02413
	6	9	0.260024	0.456584	0.766633	-0.09001	0.128154	0.316010	0.531990	-0.03996
		10	0.184607	0.421103	0.637074	-0.08367	0.091033	0.291379	0.441548	-0.03762
	7	9	0.265867	0.528246	0.766633	-0.13910	0.131234	0.365801	0.531990	-0.06337
		10	0.179461	0.480238	0.637074	-0.12649	0.088775	0.332431	0.441548	-0.05801
	8	9	0.241423	0.619424	0.766633	-0.23345	0.119156	0.429259	0.531990	-0.10921
		10	0.146845	0.548821	0.637074	-0.20279	0.073382	0.380103	0.441548	-0.09445
10	1	10	0.214615	0.131400	0.781029	0.111988	0.072587	0.090803	0.542039	0.023369
	2	10	0.139577	0.202423	0.781029	-0.01852	0.075465	0.139911	0.542039	-0.00037
	3	10	0.200369	0.260564	0.781029	-0.00314	0.093369	0.180137	0.542039	-0.00427
	4	10	0.218699	0.314105	0.781029	-0.02663	0.106389	0.217207	0.542039	-0.01135
	5	10	0.241199	0.366721	0.781029	-0.04522	0.117950	0.253665	0.542039	-0.01955
	6	10	0.259491	0.421103	0.781029	-0.06940	0.127622	0.291379	0.542039	-0.03032
	7	10	0.272860	0.480238	0.781029	-0.10222	0.134607	0.332431	0.542039	-0.04558
	8	10	0.276212	0.548821	0.781029	-0.15243	0.136336	0.380103	0.542039	-0.06969
	9	10	0.248896	0.637074	0.781029	-0.24868	0.122731	0.441548	0.542039	-0.11661

Table 3: Coefficients of the BLUEs of the location parameter for $\tau = 2$, $\delta = 0.5$.

ξ	n	c	$a_i,\ i = 1,2,3,\ldots,(n-c)$						
5	4	0	0.60056	0.52054	-0.0249	-0.0962			
		1	1.14951	-0.0022	-0.1473				
		2	1.46179	-0.4618					
	8	0	0.94978	-0.0079	-0.0355	0.11756	0.04068	-0.0357	-0.0211
			-0.0079						
		1	0.99500	-0.0542	-0.0168	0.13125	0.02120	-0.0520	-0.0245
		2	1.10008	-0.1932	0.02847	0.16721	-0.0256	-0.0769	
		3	-0.1442	3.22795	-1.6960	-1.0836	0.69588		
		4	1.22957	-5E-05	-0.2371	0.00762			
		5	1.20093	0.03644	-0.2374				
		6	1.55875	-0.5587					
	10	0	0.84633	0.46665	-0.1943	0.02331	0.04968	-0.1587	-0.0944
			0.02940	0.02612	0.00600				
		1	0.87109	0.40252	-0.1787	0.02622	0.03714	-0.1427	-0.0747
			0.03345	0.02564					
		2	0.92580	0.23990	-0.1220	0.04079	0.00905	-0.0958	-0.0266
			0.02881						
		3	1.14784	-0.0594	-0.0458	0.05472	-0.0716	-0.0521	0.02645
		4	0.97357	0.11809	-0.0719	0.05459	-0.0127	-0.0617	
		5	1.28460	-0.3778	0.12364	0.10493	-0.1354		
		6	1.19440	0.06579	-0.2574	-0.0028			

Table 4: Coefficients of the BLUEs of the scale parameter for $\tau = 2$, $\delta = 0.5$.

ξ	n	c	$b_i,\ i = 1,2,3,\ldots,(n-c)$						
5	4	0	-0.19072	0.15597	0.04744	-0.01269			
		1	-0.11835	0.08705	0.03130				
		2	-0.18472	0.18472					
	8	0	-0.11319	0.09364	0.00620	-0.02865	0.01231	0.02046	0.00619
			0.00305						
		1	-0.13061	0.11148	-0.00102	-0.03392	0.01981	0.02676	0.00750
		2	-0.16282	0.15410	-0.01488	-0.04495	0.03417	0.03438	
		3	0.39367	-1.37593	0.75633	0.51444	-0.28851		
		4	-0.17592	-0.03759	0.15149	0.06202			
		5	-0.40913	0.25959	0.14954				
		6	-0.63456	0.63456					
	10	0	-0.05656	-0.13237	0.10417	0.01245	-0.02274	0.06273	0.04769
			-0.00559	-0.00908	-0.00070				
		1	-0.05946	-0.12487	0.10235	0.01211	-0.02127	0.06086	0.04538
			-0.00606	-0.00902					
		2	-0.07871	-0.06765	0.08238	0.00698	-0.01139	0.04436	0.02846
			-0.00443						
		3	-0.11285	-0.02163	0.07068	0.00484	0.00101	0.03764	0.02030
		4	-0.24665	0.11466	0.05071	0.00474	0.04621	0.03032	
		5	-0.39958	0.35850	-0.04542	-0.02001	0.10651		
		6	-0.32861	0.00947	0.25437	0.06478			

Table 5: Variances and covariance of the BLUEs when $\tau = 0.5$, $\delta = 2$ and $\mu = 0$ and $\sigma = 1$.

ξ	n	c	$Var(\mu^*)$	$Var(\sigma^*)$	$Cov(\mu^*, \sigma^*)$
0.5	4	0	17.65257	1.410936	-0.566364
		1	14.02154	1.347831	-1.045046
		2	4.033713	0.896757	1.077511
	8	0	2.337863	0.846369	-0.520900
		1	2.167946	0.821161	-0.455453
		2	1.696970	0.776917	-0.311100
		3	20.75401	4.588457	-8.833814
		4	4.587047	1.809407	-2.130921
		5	4.594606	2.310717	-2.069361
		6	0.286717	0.600882	0.644635
	10	0	2.750427	1.010938	-1.077970
		1	2.628375	1.009270	-1.063700
		2	2.206944	0.957081	-0.915396
		3	1.727925	0.945757	-0.841744
		4	1.848615	1.016895	-0.749085
		5	0.544039	0.701471	-0.107607
		6	3.260254	2.383097	-2.244815
		7	3.261429	3.000961	-2.271760
		8	0.140384	0.562046	0.487219

5. Application with example

Here, we illustrate the applicability of the EPL distribution among a set of classical and recent models containing Pareto, generalized Lindley (GL) and type-II exponentiated log-logistic distribution (TIIELL), based on a set of goodness-of-fit statistics. We compare the goodness-of-fit of the models with the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) goodness-of-fit statistics. Further, we get the Kolmogorov-Smirnov (K-S) statistics with their corresponding p-value. In general, the smaller the values of these statistics and the larger value of the p-value, the better the fit to the data.

Description of the data is as follows: The data set refers to a study on the vinyl chloride from clean upgradient monitoring wells in mg/L, which was studied by Bhaumik et al. (2009). The data are:

```
5.1 1.2 1.3 0.6 0.5 2.4 0.5 1.1 8.0 0.8 0.4 0.6 0.9 0.4 2.0 0.5 5.3
3.2 2.7 2.9 2.5 2.3 1.0 0.2 0.1 0.1 1.8 0.9 2.0 4.0 6.8 1.2 0.4 0.2
```

In this data set, we select a random sample of size 10 and data are: 0.1, 1.1, 0.9, 2.3, 1.3, 2.5, 0.4, 2.0, 0.5, 3.2. By using the EPL distribution in Eq. (1) for the given sample, we have the maximum likelihood estimate of $\tau_{ML} = 1.2945$, $\delta_{ML} = 0.5514$ and $\xi_{ML} = 0.8099$. For this data set, a comparison of the EPL distribution with other classical and recent distributions is done as follows. We obtain the K-S, p-value, AIC and BIC statistics. The obtained results are reported in Table- 6. From this table, the smallest value of K-S, AIC and BIC and the largest value of p-value are obtained for the EPL distribution. Hence, we conclude

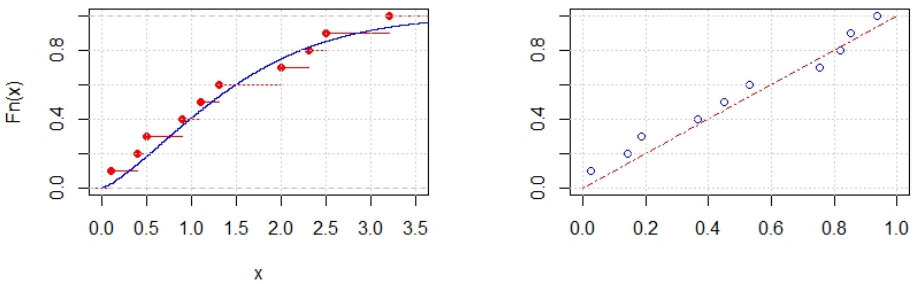


Figure 1: ECDF and Q-Q plot of the EPL distribution for the data set

that the EPL distribution provides the best fit among the compared distributions. Figure 1 shows Q-Q plot of the sample.

By using Tables 4 and 5, we have

$$\mu^* = \sum_{r=1}^{n} a_r X_{r:n} = 0.20081 \quad \text{and} \quad \sigma^* = \sum_{r=1}^{n} b_r X_{r:n} = 0.36501.$$

Table 6: The values of K-S, p- value, AIC and BIC statistics for some models fitted to the vinyl chloride data.

Distribution	K-S	p-value	AIC	BIC
Pareto	0.19982	0.1332	121.2704	122.7968
GL	0.11628	0.7474	116.2222	119.2747
TIIELL	0.10401	0.8556	116.1526	119.2053
EPL	0.08219	0.9757	116.0253	120.6044

6. Conclusion

In this paper, we have obtained the explicit expression for single and double moments of order statistics. Also, these results are reduced for special cases of order statistics of the power Lindley distribution and the Lindley distribution, which is obtained by Kumar and Goyal (2019a) and Sultan and AL-Thubyani (2016) respectively. The single and double moments of order statistics can be used to obtain BLUEs of the location and scale parameters. A simulation study is conducted for calculating BLUEs for the location and scale parameters based on Type-II right-censored samples. Finally, one real data set has been used to obtain the MLEs of the model parameters, BLUEs of the parameters and the EPL distribution is also compared with some existing distributions. We conclude that the EPL distribution provides the best fit among the compared distributions.

Acknowledgments

The authors would like to express thanks to the Editor and anonymous referees for useful suggestions and comments which have improved the presentation of the manuscript.

References

Alkarnil, S. H., (2015). Extended power Lindley distribution: a new statistical model for non-monotone survival data. *European Journal of Statistics and Probability*, 3, pp. 19–34.

Arnold, B. C., Balakrishnan, N., Nagaraja, H. N., (2003). *A First Course in Order Statistics*, New York: John Wiley.

Ahsanullah, M., Alzaatreh, A., (2018). Parameter estimation for the log-logistic distribution based on order statistics. *REVSTAT Statistical Journal*, 16, pp. 429–443.

Balakrishnan, N., Cohan, A. C., (1991). *Order Statistics and Inference: Estimation Methods*, San Diego: Academic Press.

Balakrishnan, N., Chandramouleeswaran, M. P., (1996). BLUEs of location and scale parameters of Laplace distribution based on Type-II censored samples and associated. *Microelectron Reliab*, 36, pp. 371–374.

Balakrishnan N., Melas VB, Ermakove S., (2000). *Advances in stochastic simulation and methods*. Boston, MA: Birkhauser, pp. 245–282.

Bhaumik, D. K., Kapur, K., Gibbons, R. D., (2009). Testing parameters of a gamma distribution for small samples. *Technometrics*, 51, pp. 326–334.

Bouchahed, L., Zeghdoudi, H., (2018). A new and unified approach in generalizing the Lindley's distribution with applications. *Statistics in Transition new series*, 19(1), pp. 61–74.

David, H. A., Nagaraja, H. N., (2003). *Order Statistics. Third edition*. New York: John Wiley.

Ghitany, M., Atieh, B., Nadarajah, S., (2008), Lindley distribution and its application. *Mathematics and Computers in Simulation* , 78(4), pp. 493–506.

Jabeen, R., Ahmad, A., Feroze, N. and Gilani, G. M., (2013). Estimation of location and scale parameters of Weibull distribution using generalized order statistics under Type

II singly and doubly censored data. *Int J Adv Sci Technol.*, 55, pp. 67–80.

Kumar, D., Dey, S., Nadarajah, S., (2017). Extended exponential distribution based on order statistics. *Communication in Statistics-Theory and Methods*, 46(18), pp. 9166–9184.

Kumar, D., Dey, S., (2017a). Power generalized Weibull distribution based on order statistics. *Journal of Statistical Research*, 51(1), pp. 61–78.

Kumar, D., Dey, S., (2017b). Relations for Moments of Generalized Order Statistics from Extended Exponential Distribution. *American Journal of Mathematical and Management Sciences*. 36, pp. 378–400.

Kumar, D., Goyal, A., (2019a). Order Statistics from the power Lindley distribution and associated inference with application. *Annals of data Science*, 6, pp. 153–177.

Kumar, D., Goyal, A., (2019b). Generalized Lindley Distribution Based on Order Statistics and Associated Inference with Application. *Annals of data Science*, 6, pp. 707–736.

Kumar, D., Kumar, M., Joorel, J. P. S., (2020a). Estimation with modified power function distribution based on order statistics with application to evaporation data. *Annals of data Science*, 9, pp. 723–748.

Kumar, D., Kumar, M., Dey, S., (2020b). Inferences for the type-II exponentiated log-logistic distribution based on order statistics with application. *Journal of Statistical Theory and Methods*, 13(3), pp. 352–367.

Lindley, D., (1958). Fiducial distributions and bayes theorem. *J. R. Statist. Soc. Ser. B*, 20, pp. 102–107.

Mahmoud, M. A. W., Sultan, K. S., Moshref, M. E., (2005). Inference based on order statistics from the generalized Pareto distribution and application. *Commun. Stat. Simul. Comput.*, 34, pp. 267–282.

Sanmel, P., Thomas, P. Y., (1997). Estimation of location and scale parameters of U-shaped distribution. *J Int. Soc. Agric. Stat.*, 50, pp. 75–94.

Sultan K. S, Childs A, Balakrishnan, N., (2000). Higher order moments of order statistics from the power function distribution and Edgeworth approximate inference.

Sultan, K. S., Balakrishnan, N., (2000). Higher order moments of order statistics from the power function distribution and edgeworth approximate inference. *Advances in Stochastic Simulation Methods*, pp. 245–282 , Springer, New York.

Sultan, K. S., Al-Thubyani, W. S., (2016). Higher order moments of order statistics from the Lindley distribution and associated inference. *Journal of Statistical Computation and Simulation*, 86, pp. 3432–3445.

Zeghdoudi, H., Nouar, L., Yahia, D., (2018). Lindley Pareto Distribution. *Statistics in Transition new series*, 19(4), pp. 671–692.

STATISTICS IN TRANSITION new series, June 2024
Vol. 24, No. 2, pp. 185–204, DOI 10.59170/stattrans-2024-021
Received – 21.02.2023; accepted – 20.10.2023

Functional repeated measures analysis of variance and its application

Katarzyna Kuryło[1], **Łukasz Smaga**[2]

Abstract

This paper is inspired by medical studies in which the same patients with multiple sclerosis are examined at several successive visits (doctor's appointments) and described by fractional anisotropy tract profiles, which can be represented as functions. Since the observations for each patient are dependent random processes, they follow a repeated measures design for functional data. To compare the results for different visits, we thus consider functional repeated measures analysis of variance. For this purpose, a pointwise test statistic is constructed by adapting the classical test statistic for one-way repeated measures analysis of variance to the functional data framework. By integrating and taking the supremum of the pointwise test statistic, we create two global test statistics. In addition to verifying the general null hypothesis of the equality of mean functions corresponding to different objects, we also propose a simple method for post hoc analysis. We illustrate the finite sample properties of permutation and bootstrap testing procedures in an extensive simulation study. Finally, we analyze a real data example in detail. All methods are implemented in the R package rmfanova, available on CRAN.

Key words: analysis of variance, bootstrap, functional data analysis, permutation method, post hoc analysis, repeated measures.

1. Introduction

Functional data analysis (FDA) is a branch of statistics which analyses observations treated as functions, curves, or surfaces. To represent the data in such a way, one needs only to measure a variable over time or space, which is a scenario encountered in many fields. Then the discrete data observed at the so-called design time points can be transformed into functional data. Such a representation allows us to avoid many problems of classical multivariate statistical methods, e.g. the curse of dimensionality and missing data. Therefore, in the last two decades, numerous methods have been developed for classification, clustering, dimension reduction, regression, and statistical hypothesis testing for functional data. We refer to the following key textbooks in FDA for methodology, real data examples, and computational aspects of these methods: Ferraty and Vieu (2006), Horváth and Kokoszka (2012), Ramsay and Silverman (2005), Ramsay et al. (2009), and Zhang (2013).

[1]Corresponding author. Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland. E-mail: katkur7@st.amu.edu.pl. ORCID: https://orcid.org/0000-0002-5907-6128.

[2]Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland. E-mail: ls@amu.edu.pl. ORCID: https://orcid.org/0000-0002-2442-8816.

In the literature and in software, many tests can be found for standard functional analysis of variance (FANOVA), where the samples corresponding to different objects are independent (see, Cuevas et al., 2004; Mrkvička et al., 2020; Pini et al., 2018; Smaga and Zhang, 2020; Zhang, 2013; Zhang et al., 2019 and references therein). However, the independence assumption does not always hold for functional samples. In this paper, we consider a motivating real data example where the same patients with multiple sclerosis are examined using the magnetic resonance imaging technique at several successive visits. For each visit, the results of the examination are curves observed at some design time points, which allows them to be treated as functional data. It is of interest to determine whether or not the examined functional variable changes at the different visits. Examples of similar problems appear in many fields, including medicine, education, social sciences, and psychology. Due to the dependence of functional samples corresponding to different visits, we have to consider repeated measures analysis of variance (ANOVA) in the framework of functional data. In this case, the literature is not so rich as for standard FANOVA. The repeated measures ANOVA for functional data was apparently first considered by Martínez-Camblor and Corral (2011). Developing an idea of Cuevas et al. (2004), they proposed tests based on the integral of the square of differences between sample mean functions. In the case of two groups, Smaga (2019) constructed a test based on the same test statistic, but to approximate its null distribution he used the Box-type approximation method (Box, 1954). This approach resulted in a test which is much less time-consuming than the tests of Martínez-Camblor and Corral (2011), while being comparable to them in terms of type I error level control and power. Note that these statistical tests do not take into account information about variance (see Section 2 for details). This is not an isolated case; for example, the L^2-norm-based test statistic for the standard FANOVA problem is also constructed based on the distance between sample mean functions only (Zhang, 2013). However, using information about variance can improve the test, as was shown in particular by Smaga (2020) for the paired two-sample problem for functional data. In a short note, Smaga (2021) extended these results from the two-sample case to the strict ANOVA case, i.e. when the number of groups is greater than two. In the present paper, the ideas presented in that note are developed in greater detail. First, we present the methodology at length, showing how the construction of tests for the global null hypothesis for functional data is based on classic ANOVA for real data. Moreover, we propose a simple post hoc method, which appears to be powerful (see Sections 4 and 5). To obtain tests with good finite sample properties even for smaller sample sizes, we study different bootstrap and permutation procedures. These demonstrate good qualities in many models (see, e.g. Amro et al., 2021; Ditzhaus et al., 2021; Konietschke et al., 2015; Konietschke and Pauly, 2014; Smaga and Zhang, 2020). However, this is not an invariable rule, as we also note in the extensive simulation study, where we check the control of type I error level and power of the new tests in comparison with those of Martínez-Camblor and Corral (2011) and their modifications. The simulation study is conducted for artificial and real data. The latter case explains the results of the hypothesis testing for a diffusion tensor imaging data set. This real data example demonstrates the applicability of the presented methods in an important problem. The public implementation of the methods considered in this paper is given in the R package rmfanova, which is available on CRAN (Kuryło and Smaga, 2023).

The rest of the paper is organized as follows: In Section 2 we recall the repeated measures ANOVA for random variables and present its counterpart for functional data. Moreover, we consider different pointwise test statistics. Section 3 presents the global test statistics, permutation and bootstrap tests based on them, and the simple post hoc procedure for a functional repeated measures ANOVA. The type I error level control and power of these tests are investigated in the simulation study in Section 4. Section 5 presents a real data application for diffusion tensor imaging data. Finally, Section 6 concludes the paper. The remarks about the numerical implementation of new methods and all results of the simulation study are contained in the supplement, which is available from the corresponding author upon request.

2. Functional repeated measures ANOVA model and pointwise test statistics

In this section, we first recall the repeated measures ANOVA problem for real data, and then we present the counterpart of it for functional data. Finally, we define new pointwise test statistics that will be used later to construct statistical tests.

2.1. Repeated measures ANOVA model for real data

Suppose that there are n subjects (e.g. patients), measured for $\ell \geq 2$ experimental conditions or time points (e.g. before, during, and after treatment). Assume that Y_{ij} is an observation concerning the i-th object (repetition of an experiment) applied to the j-th subject, $i = 1,\ldots,\ell$, $j = 1,\ldots,n$. Let $N = n\ell$ be the number of all observations. In repeated measures ANOVA, the model $Y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$ is used, where μ is the general mean, α_i is the fixed effect of the i-th object, β_j is a random effect of the j-th subject, and e_{ij} is a random error. There are usual assumptions about the distributions of random elements of this model, but we do not specify them, as they will not be needed for our tests. The main aim is to test the null hypothesis $\mathscr{H}_0 : \alpha_1 = \cdots = \alpha_\ell = 0$ asserting the non-significance of the objects. The F-type test statistic for this hypothesis is of the form:

$$F = \frac{\mathrm{SSA}/(\ell-1)}{\mathrm{SSR}/((\ell-1)(n-1))}, \tag{1}$$

where $\mathrm{SSA} = n\sum_{i=1}^{\ell}(\bar{Y}_{i\cdot} - \bar{Y})^2$ and $\mathrm{SSR} = \sum_{i=1}^{\ell}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y})^2$ are the sum of squares due to the hypothesis and residuals respectively, $\bar{Y}_{i\cdot} = n^{-1}\sum_{j=1}^{n} Y_{ij}$, $\bar{Y}_{\cdot j} = \ell^{-1}\sum_{i=1}^{\ell} Y_{ij}$, and $\bar{Y} = N^{-1}\sum_{i=1}^{\ell}\sum_{j=1}^{n} Y_{ij}$. The null hypothesis is rejected for large values of F. The F test statistic can be constructed in functional form and used for repeated measures analysis for functional data, as we describe in the next sections.

2.2. Repeated measures ANOVA model for functional data

Now, we describe the functional repeated measures ANOVA. It is a little different from that for real variables, which follows from the special characteristics of the functional data

framework. Similarly as above, we have n subjects subjected to $\ell \geq 2$ (possibly) different conditions, but the results of the experiments are functional observations. Let the subjects be represented by a functional sample consisting of independent stochastic processes $Y_1, \ldots, Y_n$ defined on $[0,\ell]$. We assume that these data satisfy the following model of Martínez-Camblor and Corral (2011):

$$Y_j(t) = \mu(t) + e_j(t),\ j = 1, \ldots, n,\ t \in [0,\ell], \tag{2}$$

where μ is a fixed mean function, and e_j is a random process with zero mean function and covariance function $\gamma(s,t)$, $s,t \in [0,\ell]$. In this notation, $t \in [0,1]$ corresponds to the first experimental condition, $t \in [1,2]$ to the second, and so on. Thus, in this model we ignore the possible time periods between repetitions of the experiment, but this does not mean that they do not exist. In the model (2), it is interesting to test the equality of ℓ mean functions corresponding to experimental conditions; namely, the global null hypothesis is as follows:

$$\mathscr{H}_0 : \mu(t) = \mu(t+1) = \cdots = \mu(t+(\ell-1)) \ \ \forall t \in [0,1]. \tag{3}$$

By rejecting $\mathscr{H}_0$ we determine the presence of significant differences in the mean functions corresponding to the experimental conditions. However, we do not know which conditions are significantly different and which are not. To solve this problem, one needs to perform a post hoc analysis. This is of particular practical interest.

2.3. Pointwise test statistics

For the global null hypothesis (3), the tests given by Martínez-Camblor and Corral (2011) and Smaga (2019) used the pointwise sum of squares due to the hypothesis:

$$\mathrm{SSA}_{point}(t) = n\sum_{i=1}^{\ell}(\bar{Y}_{i\cdot}(t) - \bar{Y}(t))^2,\ t \in [0,1], \tag{4}$$

where $\bar{Y}_{i\cdot}(t) = n^{-1}\sum_{j=1}^{n} Y_j(t+(i-1))$, $\bar{Y}(t) = N^{-1}\sum_{i=1}^{\ell}\sum_{j=1}^{n} Y_j(t+(i-1))$, $i = 1, \ldots, \ell$. This pointwise test statistic takes into account "between-group variability" only. In contrast, the following pointwise counterpart of the F-type test statistic (1) for $\mathscr{H}_0$ in (3) also uses "within-group variability":

$$F_{point}(t) = \frac{\mathrm{SSA}_{point}(t)/(\ell-1)}{\mathrm{SSR}_{point}(t)/((\ell-1)(n-1))},\ t \in [0,1], \tag{5}$$

where $\mathrm{SSR}_{point}(t) = \sum_{i=1}^{\ell}\sum_{j=1}^{n}(Y_j(t+(i-1)) - \bar{Y}_{i\cdot}(t) - \bar{Y}_{\cdot j}(t) + \bar{Y}(t))^2$ is the pointwise adaptation of SSR in (1), and $\bar{Y}_{\cdot j}(t) = \ell^{-1}\sum_{i=1}^{\ell} Y_j(t+(i-1))$, $j = 1, \ldots, n$. Using such additional information usually results in more powerful testing procedures. This was observed for standard ANOVA (Zhang et al., 2019; Smaga and Zhang, 2020) and for a paired two-sample test (Smaga, 2020) for functional data.

Using the pointwise F-type test statistic (5), we can construct the pointwise F-type test for (3) similarly to the pointwise F-type test in Ramsay and Silverman (2005) for stan-

dard ANOVA for functional data. Such a test rejects the null hypothesis when it rejects the hypothesis for any fixed $t \in [0,1]$. Unfortunately, the pointwise F-type test may be time-consuming, since it must be performed for all $t \in [0,1]$. Moreover, this test does not guarantee that the null hypothesis in (3) is significant overall for a given significance level even when the pointwise F-type test is significant for all $t \in [0,1]$ at the same significance level (see, Zhang, 2013). Thus, in the next section, we propose global tests for (3), which overcome this difficulty.

3. Statistical tests

In this section, we construct tests based on the pointwise F-type test statistic (5). To construct tests, we use different bootstrap and permutation methods.

3.1. Test statistics

To obtain global test statistics for $\mathscr{H}_0$ in (3), one has to aggregate the pointwise statistics (4) and (5) into one real variable. In FDA, the natural idea is integration, since the common assumption is that the elements of a functional sample $Y_1, \ldots, Y_n$ belong to the Hilbert space $L_2(T)$ of square integrable functions on the interval T. For instance, to verify the null hypothesis $\mathscr{H}_0$ in (3), Martínez-Camblor and Corral (2011) and Smaga (2019) integrated SSA_{point} as given in (4), i.e.

$$\mathscr{C}_n(\ell) = \int_0^1 \mathrm{SSA}_{point}(t)dt.$$

Another less common idea in functional data analysis is to use a supremum. For example, Zhang et al. (2019) and Smaga and Zhang (2020) developed this idea in the case of standard ANOVA for functional data. In this paper, we use both of these ideas, but in contrast to $\mathscr{C}_n(\ell)$, we apply the appropriate operations to the pointwise F-type test statistic (5) instead of the pointwise sum of squares due to hypothesis (4). Namely, we propose the following test statistics:

$$\mathscr{D}_n(\ell) = \int_0^1 F_{point}(t)dt, \quad \mathscr{E}_n(\ell) = \sup_{t\in[0,1]} F_{point}(t).$$

For the paired two-sample problem for functional data, these test statistics reduce to those considered by Smaga (2020). Thus, the results of the present paper are in fact extensions of the results obtained for the case $\ell = 2$ by Smaga (2020) to repeated measures analysis for $\ell \geq 2$ groups.

3.2. Permutation and bootstrap tests

Now, we discuss the methods of approximation of the null distributions of the test statistics $\mathscr{D}_n(\ell)$ and $\mathscr{E}_n(\ell)$, which lead to new tests for $\mathscr{H}_0$ in (3). For $\ell = 2$, Smaga (2020) proposed several methods for this purpose. Unfortunately, they do not perform equally well,

as was shown in a simulation study. Namely, the tests based on methods of approximation of the asymptotic null distributions of the test statistics (e.g. the parametric bootstrap approach) were usually too liberal for small and moderate values of the number n of subjects. Thus, we do not consider such methods in this paper, but concentrate on permutation and bootstrap approaches. In general, these methods do not need particular assumptions about the distribution of the data and a large number of observations to give a good approximation. Of course, not all resampling methods work well, as we will see in the simulation study.

As the first method, we consider the permutation approach considered by Martínez-Camblor and Corral (2011) and Smaga (2020). Let $\mathscr{T}_n(\ell)$ denote one of the test statistics $\mathscr{D}_n(\ell)$ and $\mathscr{E}_n(\ell)$ (or possibly $\mathscr{C}_n(\ell)$) and let B be a large number (e.g. $B = 1000$). The steps of this procedure are as follows:

1. Compute $\mathscr{T}_n(\ell)$ for the original data $Y_1(t),\dots,Y_n(t)$, $t \in [0,\ell]$.

2. Create a permutation sample $Y_1^b(t),\dots,Y_n^b(t)$, $t \in [0,\ell]$ from the given data in the following way: for each $j = 1,\dots,n$ separately, randomly permute the observations $Y_j(t), Y_j(t+1),\dots,Y_j(t+(\ell-1))$, $t \in [0,1]$ corresponding to ℓ experimental conditions, and use them to form the permuted observation $Y_j^b(t)$.

3. Repeat step 2 B times to obtain B independent permutation samples.

4. For each permutation sample, compute the value of the test statistic $\mathscr{T}_n(\ell)$. Denote these by $\mathscr{T}_{n,b}(\ell)$, $b = 1,\dots,B$.

5. The final p-value of the permutation test is defined by $B^{-1}\sum_{b=1}^{B} I(\mathscr{T}_{n,b}(\ell) > \mathscr{T}_n(\ell))$, where $I(A)$ stands for the usual indicator function on a set A.

We refer to this permutation test as $\mathscr{P}_1$ for short. The following approaches follow the same steps, but use different methods of constructing new samples in step 2 and possibly modified forms of the test statistics in step 4.

In the second permutation approach, step 2 is performed as follows: We draw $Y_1^b(t),\dots,Y_n^b(t)$ for $t \in [0,1]$ randomly without replacement from the set

$$\mathscr{A} = \{Y_1(t),\dots,Y_n(t), Y_1(t+1),\dots,Y_n(t+1),\dots,Y_1(t+(\ell-1)),\dots,Y_n(t+(\ell-1))\}$$

for $t \in [0,1]$ containing all N observations; after that we draw $Y_1^b(t),\dots,Y_n^b(t)$ for $t \in [1,2]$ randomly without replacement from the remaining elements in $\mathscr{A}$, and so on. We refer to this procedure as the $\mathscr{P}_2$ method. It is similar to the procedure considered by Konietschke and Pauly (2014, method (I)).

Let us turn now to three bootstrap approaches. The first is the nonparametric bootstrap approach considered by Martínez-Camblor and Corral (2011) and Smaga (2020). We refer to it as the $\mathscr{B}_1$ method for short. In this approach, the second step of the above procedure selects independent bootstrap samples $Y_1^b(t),\dots,Y_n^b(t)$, $t \in [0,\ell]$ drawn with replacement from the original sample $Y_1(t),\dots,Y_n(t)$, $t \in [0,\ell]$. Moreover, we need to modify step 4; namely, the pointwise sum of squares due to the hypothesis SSA_{point} for bootstrap samples

is calculated as follows:

$$\mathrm{SSA}^b_{point}(t) = n\sum_{i=1}^{\ell}(\bar{Y}^b_{i\cdot}(t) - \bar{Y}_{i\cdot}(t) - \bar{Y}^b(t) + \bar{Y}(t))^2,\ t \in [0,1],$$

where $\bar{Y}^b_{i\cdot}(t)$ and $\bar{Y}^b(t)$ are the sample means computed on the bootstrap sample.

In step 2 of the second nonparametric bootstrap approach, we first have to center the observations. Namely, let $Y_{1,c}(t) = Y_1(t) - \bar{Y}_\bullet(t), \ldots, Y_{n,c}(t) = Y_n(t) - \bar{Y}_\bullet(t)$ for $t \in [0,\ell]$, where $\bar{Y}_\bullet(t) = n^{-1}\sum_{j=1}^n Y_j(t)$ is the sample mean function for the original sample $Y_1, \ldots, Y_n$. Second, $Y^b_1(t), \ldots, Y^b_n(t)$ for $t \in [0,1]$ are randomly drawn with replacement from $Y_{1,c}(t), \ldots, Y_{n,c}(t)$, $t \in [0,1]$; independently, $Y^b_1(t), \ldots, Y^b_n(t)$ for $t \in [1,2]$ are randomly drawn with replacement from $Y_{1,c}(t), \ldots, Y_{n,c}(t)$, $t \in [1,2]$, and so on. We refer to this method for short as the $\mathscr{B}_2$ method. It is an extension of the methods considered by Konietschke and Pauly (2014) and Smaga (2020).

The last method considered is the parametric bootstrap approach (the $\mathscr{B}_3$ test for short), which uses bootstrap samples that are constructed in a much different way than in the above methods. The idea of this approach is similar to the parametric bootstrap proposed by Konietschke et al. (2015). In step 2, we generate the bootstrap samples $Y^b_1(t), \ldots, Y^b_n(t)$, $t \in [0,\ell]$ from the Gaussian process with zero mean function and covariance function equal to the sample covariance function

$$\hat{\gamma}(s,t) = \frac{1}{n-1}\sum_{j=1}^{n}(Y_j(s) - \bar{Y}_\bullet(s))(Y_j(t) - \bar{Y}_\bullet(t)),\ s,t \in [0,\ell], \tag{6}$$

which is the unbiased estimator of the covariance function $\gamma(s,t)$ (Zhang, 2013, page 108). Such bootstrap samples satisfy the null hypothesis $\mathscr{H}_0$ in (3), since the mean function is the same for each Y^b_j. Moreover, the estimator of the covariance function used mimics the given covariance structure of the original observations, and it does this quite well, as we will see in the simulation study. Note that this parametric bootstrap method is different from the parametric bootstrap test in Martínez-Camblor and Corral (2011) and Smaga (2020), which was used to approximate the asymptotic distributions of the test statistics $\mathscr{C}_n(\ell)$, $\mathscr{D}_n(2)$, and $\mathscr{E}_n(2)$.

3.3. Post hoc analysis

After rejecting the global null hypothesis (3), it is of practical interest to perform post hoc analysis. The aim is to determine which experimental conditions give significantly different mean functions of the dependent variable Y, and which do not result in such differences. Namely, we would like to test the family of null hypotheses:

$$\begin{cases} \mathscr{H}^{rs}_0 : \mu(t+(r-1)) = \mu(t+(s-1))\ \forall t \in [0,1], \\ \mathscr{H}^{rs}_1 : \mu(t+(r-1)) \neq \mu(t+(s-1)) \text{ for some } t \in [0,1], \end{cases} \tag{7}$$

for $r,s = 1,\ldots,\ell$, $r \neq s$. These hypotheses are also named pairwise comparisons. To test the family of local hypotheses (7), we propose the following simple procedure:

1. Test each of the hypotheses in (7) using the data for the r-th and s-th objects, i.e. $Y_1(t),\ldots,Y_n(t)$ for $t \in [r-1,r]$ and $t \in [s-1,s]$ respectively, and the chosen test from Section 3.2. Let p_{rs} denote the p-values obtained.
2. Make a final decision using the Bonferroni method, i.e. reject the null hypothesis $\mathscr{H}_0^{rs}$ if $p_{rs} \leq \alpha/m$, where $m = \#\{p_{rs} : r,s = 1,\ldots,\ell, r \neq s\}$ is the number of null hypotheses considered.

Step 2 of the above procedure can also be conducted by comparing the corrected p-values $p_{rs}^{Bonf} = m \cdot p_{rs}$ with the significance level α, i.e. we reject $\mathscr{H}_0^{rs}$ if $p_{rs}^{Bonf} \leq \alpha$.

One of the common measures of the quality of multiple testing procedures is the family-wise error rate (FWER), namely the probability of making at least one type I error in the family of hypotheses (Tukey, 1953). It is known that the Bonferroni method controls the FWER at level α, i.e. FWER $\leq \alpha$, if the individual tests control the type I error at level α.

4. Simulation study

In this section, we investigate the finite sample properties of the considered tests in a Monte Carlo simulation study. We consider the control of the type I error level and the power of fifteen tests: the $\mathscr{P}_1$, $\mathscr{P}_2$, $\mathscr{B}_1$, $\mathscr{B}_2$, and $\mathscr{B}_3$ procedures based on the test statistics $\mathscr{C}_n(\ell)$, $\mathscr{D}_n(\ell)$, and $\mathscr{E}_n(\ell)$. We include the tests based on $\mathscr{C}_n(\ell)$ to provide a fair comparison between our tests and those based on the test statistic proposed by Martínez-Camblor and Corral (2011).

In simulation experiments, we consider $\ell = 3$ repeated samples with different distributions, sample sizes, and amounts of correlation. Simulation studies in the case $\ell = 2$ were conducted by Smaga (2020). In this case, the $\mathscr{P}_1$ tests based on the test statistics $\mathscr{D}_n(2)$ and $\mathscr{E}_n(2)$ performed best. For $\ell > 2$, the situation may change, as we will describe. We consider testing of the global null hypothesis (3) as well as the pairwise comparisons (7).

4.1. Simulation design

The $\ell = 3$ repeated samples were generated according to model (2). We consider the following six main models specifying the mean functions, which directly indicate the truth or falsity of the null hypothesis ($t \in [0,1]$):
M1: $\mu(t) = \mu(t+1) = \mu(t+2) = (\sin(2\pi t^2))^5$,
M2: $\mu(t) = \mu(t+1) = (\sin(2\pi t^2))^5$ and $\mu(t+2) = (\sin(2\pi t^2))^7$,
M3: $\mu(t) = \mu(t+1) = (\sin(2\pi t^2))^5$ and $\mu(t+2) = (\sin(2\pi t^2))^3$,
M4: $\mu(t) = \mu(t+1) = \mu(t+2) = \sqrt{6t/\pi}\exp(-6t)$,
M5: $\mu(t) = \mu(t+1) = \sqrt{6t/\pi}\exp(-6t)$ and $\mu(t+2) = \sqrt{13t/(2\pi)}\exp(-13t/2)$,
M6: $\mu(t) = \mu(t+1) = \sqrt{6t/\pi}\exp(-6t)$ and $\mu(t+2) = \sqrt{11t/(2\pi)}\exp(-11t/2)$.

These models are very similar to the models considered in the simulation study by Martínez-Camblor and Corral (2011). In models M1 and M4 (respectively M2, M3, M5,

and M6), the null hypothesis $\mathscr{H}_0$ in (3) is true (respectively false), and hence we investigate the type I error level control (respectively power). We also considered two settings for the distribution of the data, namely normal and lognormal. In a normal setting, the errors were generated as follows: $e_j(t) = \xi B_{j1}(t)$, $e_j(t+1) = \rho e_j(t) + \xi(1-\rho^2)^{1/2} B_{j2}(t)$ and $e_j(t+2) = \rho e_j(t+1) + \xi(1-\rho^2)^{1/2} B_{j3}(t)$ for $t \in [0,1]$ and $\xi = 0.5, 0.05$ for models M1–M3 and M4–M6 respectively. Here B_{j1}, B_{j2}, and B_{j3} are the independent standard Brownian bridges, $j = 1,\ldots,n$. We set $\rho = 0, 0.25, 0.5, 0.75$ for zero, small, moderate, and large correlation respectively. The lognormal errors were the (appropriately centred) exponentially transformed errors of the normal setting, i.e. $\exp(e_j(t))$, $j = 1,\ldots,n$, $t \in [0,3]$. The sample sizes were set to $n = 35, 50, 75, 100$. Since functional data are not usually continuously observed in practice, the trajectories of the observations $Y_1(t),\ldots,Y_n(t)$, $t \in [0,3]$ were discretized at the design time points $t_1,\ldots,t_{101}, t_1+1,\ldots,t_{101}+1, t_1+2,\ldots,t_{101}+2$, where t_k, $k = 1,\ldots,101$ were equispaced in the interval $[0,1]$.

The significance level α was set to 5%. The empirical sizes and powers of the tests were estimated as the proportion of rejections of the null hypothesis based on 1000 Monte Carlo runs. The p-values were calculated based on $B = 1000$ bootstrap and permutation samples. The R environment (R Core Team, 2022) was used to conduct the simulation study and the real data analysis described in Section 5.

4.2. Simulation results

All simulation results are contained in the supplement, which is available from the corresponding author upon request. Here, to save space, we present only selected results, which are nevertheless representative and given in Tables 1–5 and Figure 1.

Let us first discuss the results of the simulation experiments for the fifteen tests given in Section 3.2 for testing the global null hypothesis (3). The results are presented in Tables 1–2 and Figure 1. The type I error control is studied in models M1 (Table 1 and Figure 1) and M4. Since the results for model M4 are similar to those for model M1, they are available in the supplement. For both normal and lognormal distributions, the empirical sizes have similar behaviour. Although the $\mathscr{P}_1$ test with $\mathscr{E}_n(2)$ performed very well for $\ell = 2$, it is too liberal when the number of groups is greater than two, especially for larger correlation ($\rho = 0.5, 0.75$). The same problem is observed for two other tests based on $\mathscr{E}_n(3)$, namely $\mathscr{P}_2$ and $\mathscr{B}_2$. Moreover, this bad behaviour seems to diminish very slowly as the number of observations n increases. This means that these three tests should be used with great care. On the other hand, for any amount of correlation, the $\mathscr{C}_n(3)$-based $\mathscr{P}_2$ and $\mathscr{B}_2$ tests have small empirical sizes, which means that these testing procedures are conservative. They become more conservative with increasing correlation. Of course, such tests control the type I error level, but may have a larger type II error rate, or in other words, a smaller power. The remaining tests exhibit much better performance, although the $\mathscr{B}_1$ and $\mathscr{B}_3$ tests based on the test statistics $\mathscr{D}_n(3)$ and $\mathscr{E}_n(3)$ may have a slightly conservative character.

For completeness, we consider all tests in the investigation of power. However, let us note that the results for the $\mathscr{P}_1$, $\mathscr{P}_2$, and $\mathscr{B}_2$ tests based on $\mathscr{E}_n(3)$ may be better than for the other tests due to the former tests' being too liberal, making such a comparison unfair. The empirical powers of selected cases are presented in Table 2 and Figure 1. For

Table 1: Empirical sizes (as percentages) of all tests obtained in model M1.

	$\mathscr{C}_n(3)$					$\mathscr{D}_n(3)$					$\mathscr{E}_n(3)$				
ρ	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$
	normal														
	$n=35$														
0.00	4.9	5.4	5.0	5.5	4.9	5.1	5.5	3.6	4.4	4.5	4.1	4.2	2.9	3.4	3.3
0.25	4.8	1.7	4.8	2.1	4.6	5.7	5.3	3.4	4.6	4.0	6.1	6.1	3.4	4.4	3.7
0.50	5.3	0.5	4.8	0.8	4.5	5.2	4.5	3.7	4.7	4.0	6.7	7.1	3.7	6.0	3.9
0.75	4.7	0.0	3.9	0.0	4.2	5.6	5.2	3.3	5.1	4.0	8.2	8.1	3.7	7.4	4.3
	$n=50$														
0.00	4.7	4.6	4.6	5.4	4.2	5.1	5.0	3.9	4.3	4.2	4.7	4.9	3.9	4.0	3.4
0.25	5.3	2.8	5.3	2.6	4.8	5.7	5.7	4.7	5.2	4.9	5.6	5.9	4.5	5.4	4.4
0.50	5.7	0.7	5.6	0.6	5.0	6.0	5.8	4.5	5.3	4.8	6.8	6.9	4.3	6.3	4.3
0.75	5.8	0.0	5.3	0.0	5.0	5.8	5.5	4.8	5.2	5.0	9.4	9.1	5.3	8.9	5.2
	$n=75$														
0.00	5.0	5.2	4.5	5.3	5.0	4.6	5.1	4.3	4.7	4.6	4.7	5.1	3.9	4.2	3.4
0.25	5.6	1.9	4.9	2.2	4.4	5.1	5.5	4.1	4.5	4.2	5.4	5.9	4.3	4.9	3.8
0.50	5.6	0.4	5.1	0.4	5.0	5.7	5.3	4.8	5.1	4.8	6.5	6.4	4.3	6.6	3.9
0.75	5.7	0.0	5.3	0.0	5.2	5.9	5.7	4.6	6.2	4.8	9.5	9.8	4.7	9.4	4.2
	$n=100$														
0.00	5.4	5.7	5.8	5.5	5.5	5.6	5.3	4.8	4.9	5.0	4.8	5.2	3.9	4.5	3.4
0.25	5.3	2.2	5.1	2.2	5.2	5.5	5.4	5.0	5.5	4.9	5.6	5.6	4.0	4.9	3.4
0.50	5.1	0.6	4.7	0.5	4.7	5.2	5.4	4.6	5.1	4.4	7.2	6.7	4.6	7.0	4.7
0.75	6.1	0.0	5.3	0.0	5.7	5.5	5.8	4.7	5.3	5.1	9.4	9.4	4.1	9.5	4.1
	lognormal														
	$n=35$														
0.00	4.4	4.8	3.4	4.4	3.5	4.4	4.7	3.0	3.7	3.5	5.4	6.0	3.3	4.3	4.1
0.25	4.6	1.8	3.5	1.7	3.8	4.7	5.2	3.3	4.0	3.7	5.4	6.0	3.5	4.4	4.4
0.50	5.0	0.1	3.6	0.1	3.4	5.4	5.3	3.0	4.4	3.5	7.2	7.1	3.7	5.7	4.6
0.75	5.6	0.0	4.0	0.0	3.4	5.7	5.6	2.9	4.3	3.5	9.3	9.1	4.0	7.8	5.0
	$n=50$														
0.00	4.8	5.0	4.4	4.6	4.1	4.8	5.3	3.6	4.0	3.9	5.1	5.2	3.7	4.1	4.4
0.25	6.0	2.1	4.3	1.9	4.8	5.4	5.5	3.8	4.4	4.3	5.9	5.9	4.1	5.0	4.5
0.50	6.5	0.3	5.1	0.4	5.0	6.3	6.0	4.4	5.1	4.6	7.9	7.4	5.0	6.5	5.9
0.75	5.8	0.0	4.2	0.0	4.0	5.6	5.7	3.2	5.4	3.9	9.6	9.0	4.2	8.6	4.9
	$n=75$														
0.00	5.4	5.9	4.9	5.0	4.4	5.3	5.3	4.2	4.4	4.2	5.4	5.5	4.3	4.7	4.5
0.25	5.7	2.2	4.2	2.3	5.1	5.8	5.6	4.1	5.1	4.8	5.0	5.3	3.8	4.3	3.7
0.50	6.2	0.4	5.3	0.4	5.3	6.3	6.5	4.8	5.5	5.0	8.2	7.8	4.7	7.3	5.4
0.75	6.3	0.0	5.0	0.0	5.2	6.4	5.9	4.9	6.3	5.0	9.6	8.9	4.6	8.9	5.4
	$n=100$														
0.00	5.3	5.9	4.8	5.7	5.1	5.3	5.6	4.4	5.1	4.8	5.3	5.4	4.4	4.9	4.8
0.25	5.2	1.7	4.8	1.6	5.1	5.6	5.7	4.7	5.2	5.0	5.9	6.2	4.6	5.8	5.1
0.50	3.5	0.1	3.2	0.0	3.0	3.7	3.7	3.1	3.4	2.9	6.2	5.9	3.4	5.4	3.5
0.75	4.5	0.0	4.1	0.0	3.9	4.5	4.3	3.9	4.3	3.7	7.5	7.9	3.2	7.5	3.6

the other scenarios, the results are similar or close to 100%. Naturally, the power increases when we have more data, i.e. as n increases. Moreover, for almost all tests, this also holds for increasing correlation between consecutive samples, i.e. as ρ increases. However, for the $\mathscr{C}_n(3)$-based $\mathscr{P}_2$ and $\mathscr{B}_2$ testing procedures, we usually observe the reverse behaviour, which can be explained by their extreme conservative character. Furthermore, these two tests are the least powerful among all of the procedures considered when $\rho > 0$. The remaining three tests based on $\mathscr{C}_n(3)$ perform much better, but they are usually less powerful than the $\mathscr{D}_n(3)$ procedures. In most cases, the permutation method outperforms the bootstrap approach for the $\mathscr{D}_n(3)$ tests. The same seems to be true for the $\mathscr{C}_n(3)$ tests. Finally, the most powerful methods are the $\mathscr{E}_n(3)$ tests. This last observation was made in almost all of our experiments, but we note that this is not always necessarily the case, as we will show below for pairwise comparisons.

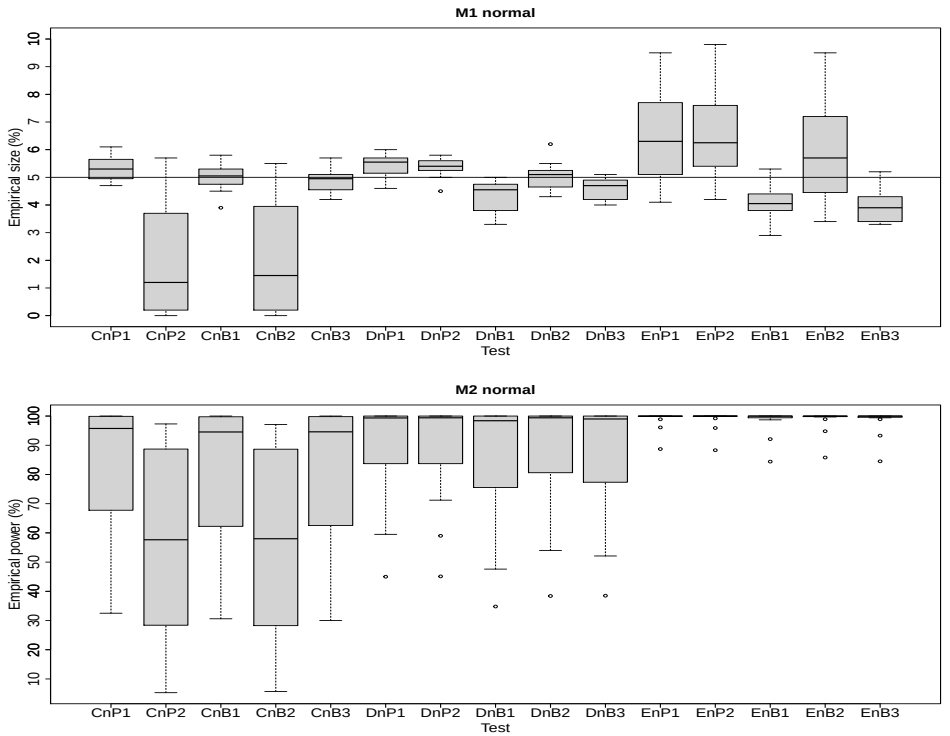


Figure 1: Box-and-whisker plots for empirical sizes and powers (as percentages) of all tests obtained in models M1 and M2.

Let us now turn to verifying the pairwise comparisons (7) (Tables 3–5). For the cases where all null hypotheses are true (models M1 and M4), we estimated the FWER (see Section 3.3). It should be controlled at level α, i.e. FWER $\leq \alpha$. The results for model M1 are given in Table 3; those for model M4 are again similar, and hence are available in the supplement. Since we use just two groups in each comparison and the Bonferroni correction is applied, the FWER is controlled at the significance level by all tests considered. However, some of them again have a conservative character. This is evident for the $\mathscr{C}_n(3)$-based $\mathscr{P}_2$ and $\mathscr{B}_2$ testing procedures, but the $\mathscr{D}_n(3)$-based $\mathscr{P}_2$, $\mathscr{B}_1$, and $\mathscr{B}_2$ tests and $\mathscr{E}_n(3)$-based $\mathscr{B}_1$ and $\mathscr{B}_2$ tests are also slightly conservative. This results in loss of power (Table 4), especially for the $\mathscr{P}_2$ and $\mathscr{B}_2$ $\mathscr{C}_n(3)$ tests, except in the independent case ($\rho = 0$). The empirical powers are presented for two comparisons, between the first and third samples and the second and third samples (denoted by 1-3 and 2-3 respectively), since in these cases the mean functions are different. The empirical powers show a good ability of the post hoc tests to detect differences in the mean functions for particular pairs of samples. In model M2 (Table 4) as well as models M3, M5, and M6 (data available in the supplement), we observe that the $\mathscr{E}_n(3)$ tests usually outperform the $\mathscr{D}_n(3)$ procedures, which are more powerful than the $\mathscr{C}_n(3)$ tests in most cases. However, we note that the opposite situation can sometimes

Table 2: Empirical powers (as percentages) of all tests obtained in selected cases in models M2, M3, M5, and M6.

	$\mathscr{C}_n(3)$					$\mathscr{D}_n(3)$					$\mathscr{E}_n(3)$				
ρ	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$
Model 2 normal															
	$n=35$														
0.00	32.5	32.1	30.6	32.0	30.0	45.0	45.1	34.8	38.4	38.5	88.7	88.3	84.4	85.8	84.5
0.25	42.9	24.7	40.6	24.5	40.2	59.5	59.0	47.6	54.0	52.1	96.1	95.9	92.1	94.8	93.3
0.50	65.6	13.7	59.3	14.8	59.6	82.1	81.9	72.3	78.8	74.0	99.9	100	98.9	99.7	99.4
0.75	97.8	5.3	96.6	5.7	96.2	99.9	99.8	98.7	99.7	99.5	100	100	100	100	100
	$n=50$														
0.00	52.6	53.1	51.0	53.4	50.9	71.3	71.2	64.7	68.3	66.8	98.9	99.2	98.7	98.9	98.9
0.25	69.9	43.0	65.2	45.0	65.5	85.3	85.5	78.8	82.4	80.7	99.9	99.9	99.9	99.9	99.8
0.50	91.0	33.5	87.1	33.4	87.9	98.3	97.9	96.2	97.8	96.7	100	100	100	100	100
0.75	100	16.4	99.9	18.2	99.9	100	100	100	100	100	100	100	100	100	100
	$n=75$														
0.00	82.8	82.2	80.6	82.7	80.5	94.8	95.5	94.0	94.8	94.6	100	100	100	100	100
0.25	94.0	78.1	92.5	77.0	93.0	98.9	99.0	98.1	99.1	98.5	100	100	100	100	100
0.50	99.5	70.6	99.5	71.2	99.7	100	100	100	100	100	100	100	100	100	100
0.75	100	62.2	100	62.6	100	100	100	100	100	100	100	100	100	100	100
	$n=100$														
0.00	97.5	97.3	96.9	97.1	97.4	99.8	99.8	99.8	99.8	99.9	100	100	100	100	100
0.25	99.8	96.5	99.6	96.8	99.6	100	100	100	100	100	100	100	100	100	100
0.50	100	95.7	100	95.2	100	100	100	100	100	100	100	100	100	100	100
0.75	100	95.2	100	94.6	100	100	100	100	100	100	100	100	100	100	100
Model 2 lognormal															
	$n=35$														
0.00	86.8	87.3	82.9	86.4	82.3	86.7	86.7	79.5	83.4	81.7	95.5	95.7	92.4	93.6	94.4
0.25	94.5	84.1	91.3	82.5	92.0	94.2	94.1	89.2	91.8	91.4	98.7	98.8	97.7	98.6	98.5
0.50	99.1	78.3	98.7	77.1	98.6	99.1	99.1	98.3	99.1	98.5	99.9	100	99.8	100	99.8
0.75	100	68.7	100	69.6	100	100	100	100	100	100	100	100	100	100	100
	$n=50$														
0.00	98.6	98.6	97.7	98.5	97.8	98.4	98.4	97.1	97.7	97.5	99.7	99.6	99.6	99.6	99.7
0.25	99.5	97.8	99.5	97.8	99.4	99.5	99.5	99.1	99.6	99.3	100	100	100	100	100
0.50	100	97.3	100	97.3	100	100	100	100	100	100	100	100	100	100	100
0.75	100	97.5	100	97.3	100	100	100	100	100	100	100	100	100	100	100
Model 3 normal															
	$n=35$														
0.00	92.9	93.6	91.1	93.1	91.5	99.3	99.7	98.2	98.9	99.2	100	100	100	100	100
0.25	98.3	91.3	97.6	91.2	97.5	99.8	99.8	99.4	99.8	99.7	100	100	100	100	100
0.50	99.8	90.9	99.7	91.0	99.8	100	100	100	100	100	100	100	100	100	100
0.75	100	88.4	100	89.3	100	100	100	100	100	100	100	100	100	100	100
Model 5 normal															
	$n=35$														
0.00	55.1	54.4	54.1	54.2	54.1	63.2	63.4	55.2	60.0	60.0	91.0	91.1	86.7	88.7	88.2
0.25	61.3	47.5	61.8	49.2	61.8	72.4	72.2	66.2	70.6	69.1	96.6	96.8	94.1	96.2	93.8
0.50	77.6	40.6	76.7	42.2	75.8	89.0	89.2	83.0	87.5	85.6	98.9	99.0	97.9	98.6	98.3
0.75	97.3	34.3	97.0	37.8	97.1	99.8	99.8	99.4	99.8	99.6	100	100	100	100	100
	$n=50$														
0.00	72.5	72.6	73.3	73.0	73.6	81.4	82.2	78.8	80.2	80.1	98.9	99.0	98.4	98.6	98.3
0.25	82.3	69.9	81.1	70.7	81.1	91.4	91.1	88.7	89.7	88.9	99.7	99.7	99.6	99.5	99.5
0.50	92.4	66.9	92.2	69.0	92.0	97.5	97.3	96.7	97.2	97.0	100	100	100	100	100
0.75	100	63.9	99.9	67.0	99.9	100	100	100	100	100	100	100	100	100	100
Model 6 normal															
	$n=35$														
0.00	67.5	68.4	67.1	68.3	67.5	75.4	75.9	70.7	72.9	72.6	96.7	96.4	93.7	95.1	94.5
0.25	71.9	60.7	72.5	62.9	71.9	82.1	82.3	77.2	80.7	79.4	98.6	99.0	98.0	98.4	97.8
0.50	87.5	57.9	86.2	60.7	86.5	95.2	94.8	92.0	94.0	93.2	100	99.9	99.6	99.9	99.9
0.75	98.8	52.7	98.6	56.2	98.9	100	100	100	100	100	100	100	100	100	100
	$n=50$														
0.00	84.3	84.4	83.6	84.7	83.4	91.7	91.4	89.1	90.4	90.8	99.9	99.9	99.9	99.7	99.7
0.25	89.1	81.6	88.2	82.2	88.1	95.3	95.0	94.1	94.8	94.7	99.9	99.9	99.9	99.9	99.9
0.50	96.2	82.1	95.7	82.6	96.2	99.4	99.8	98.9	99.3	99.2	100	100	100	100	100
0.75	100	83.3	100	85.3	100	100	100	100	100	100	100	100	100	100	100

Table 3: Empirical FWER (as percentages) of all post hoc tests obtained in model M1.

	$\mathscr{C}_n(3)$					$\mathscr{D}_n(3)$					$\mathscr{E}_n(3)$				
ρ	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$
	normal														
	$n=35$														
0.00	4.9	4.7	5.9	5.3	5.5	4.5	4.5	4.0	3.9	4.6	5.3	5.4	3.5	4.6	5.6
0.25	4.4	2.3	5.4	2.4	4.9	4.2	3.8	3.3	3.6	3.6	4.1	4.3	2.4	3.6	4.1
0.50	4.0	0.5	4.9	0.8	4.9	3.8	3.7	3.3	3.7	3.9	3.8	3.8	2.5	3.7	3.7
0.75	5.4	0.1	6.4	0.1	6.4	5.2	5.9	4.5	4.8	5.6	5.4	5.3	3.3	5.1	5.0
	$n=50$														
0.00	4.8	4.4	5.4	5.6	6.0	4.9	4.2	3.7	4.1	4.0	4.8	4.3	3.4	3.4	4.3
0.25	5.0	2.3	5.4	2.7	5.6	5.1	4.8	4.0	4.4	5.1	5.8	5.7	4.0	5.3	5.6
0.50	4.2	0.2	5.3	0.6	5.9	4.1	4.5	4.0	4.1	4.2	4.5	4.7	3.1	4.5	4.2
0.75	5.3	0.0	6.1	0.0	6.2	5.4	4.4	4.4	4.6	5.2	4.3	4.3	3.1	3.8	3.8
	$n=75$														
0.00	4.7	3.8	4.3	3.6	4.2	4.1	3.8	3.5	3.6	3.9	4.1	3.8	3.5	3.7	4.2
0.25	5.6	1.8	6.2	2.3	6.3	5.7	5.4	5.5	5.1	5.6	4.3	4.4	3.4	4.0	4.4
0.50	5.2	0.4	5.6	0.7	6.1	5.2	4.8	5.1	5.3	5.4	4.8	5.2	4.8	4.7	5.2
0.75	3.6	0.0	4.4	0.0	4.5	4.4	3.9	3.7	3.9	3.8	3.9	4.3	4.2	4.3	3.8
	$n=100$														
0.00	4.8	5.1	5.1	4.8	4.7	5.1	5.0	5.0	4.9	5.0	4.8	4.5	4.7	4.7	4.6
0.25	6.2	2.9	6.6	3.4	6.3	6.1	6.7	6.1	6.3	6.2	4.9	5.2	5.1	4.8	4.9
0.50	5.7	0.9	5.7	1.3	6.2	5.2	5.2	5.4	5.2	5.5	5.1	4.8	4.4	5.5	4.6
0.75	4.6	0.0	5.8	0.0	5.6	4.8	4.1	4.8	4.7	4.5	5.0	5.2	4.4	5.1	5.3
	lognormal														
	$n=35$														
0.00	5.2	5.0	5.4	4.9	5.1	5.6	5.6	4.1	4.7	4.8	5.0	5.4	3.2	3.5	5.2
0.25	5.2	2.7	4.7	2.1	4.6	4.5	4.6	3.2	3.6	4.7	4.2	4.3	2.4	2.8	4.3
0.50	5.1	0.2	4.7	0.4	4.3	4.9	4.6	3.4	3.6	3.9	5.2	5.2	2.7	3.9	4.7
0.75	4.3	0.0	4.1	0.0	3.7	4.0	3.9	3.0	3.7	3.8	4.5	4.5	2.6	3.6	4.4
	$n=50$														
0.00	3.8	4.4	4.3	4.4	3.7	3.6	4.4	3.3	4.0	3.6	4.4	4.1	2.5	3.4	4.2
0.25	4.1	2.3	4.9	1.9	4.6	4.1	4.9	3.3	3.8	4.1	4.7	4.7	3.2	3.4	4.5
0.50	5.7	0.6	5.7	0.5	5.7	6.4	5.7	4.9	5.6	6.3	6.2	5.8	3.7	5.0	6.3
0.75	4.1	0.0	4.7	0.0	3.9	4.7	4.6	4.3	3.8	4.2	5.0	4.7	3.1	4.2	4.6
	$n=75$														
0.00	4.2	4.2	4.4	4.1	4.2	4.0	4.2	3.8	3.5	4.0	4.9	5.1	3.8	4.2	4.7
0.25	4.9	2.2	4.6	2.4	4.0	4.7	4.8	4.4	4.5	4.6	5.4	5.7	4.3	5.0	5.0
0.50	5.2	0.4	4.7	0.3	5.1	4.9	4.6	4.0	4.5	4.8	4.7	4.9	4.0	4.4	4.4
0.75	4.5	0.0	3.7	0.0	3.9	4.3	4.7	3.2	4.0	3.7	4.7	4.1	3.5	3.8	4.2
	$n=100$														
0.00	4.7	4.7	4.4	4.7	4.6	4.7	4.5	4.4	4.4	4.2	4.3	4.5	4.2	4.2	5.0
0.25	5.3	2.7	5.3	2.5	5.4	5.2	5.3	4.8	5.0	5.2	4.9	4.6	4.0	4.4	4.8
0.50	3.7	0.5	3.3	0.3	3.3	4.0	3.7	3.2	3.7	3.2	5.3	5.4	4.5	5.1	5.4
0.75	4.3	0.0	4.6	0.0	4.6	4.5	4.7	3.8	4.6	4.6	6.1	5.9	4.9	5.3	5.3

apply. This is shown in Table 5, where the $\mathscr{E}_n(3)$ tests are less powerful than the remaining tests in a few cases. Except for the extremely conservative tests, the results for the 1-3 comparison are usually worse than those for the 2-3 comparison. This can be explained by the different amounts of correlation: the observations in samples 1 and 3 are less correlated than in samples 2 and 3, and as we observed in the investigation of powers for the global null hypothesis, greater correlation leads to larger power. The latter observation is also usually true for pairwise comparisons.

To sum up, the best type I error control and power are achieved by the $\mathscr{E}_n(3)$-based $\mathscr{B}_1$ and $\mathscr{B}_3$ tests. However, the tests based on the test statistic $\mathscr{D}_n(3)$ (especially permutation tests) also have good finite sample properties.

Table 4: Empirical powers (as percentages) of all post hoc tests obtained in selected cases in model M2.

		$\mathscr{C}_n(3)$					$\mathscr{D}_n3)$					$\mathscr{E}_n(3)$				
ρ	Comp	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$
normal																
$n=35$																
0.00	1-3	12.8	12.2	12.7	13.1	12.6	17.1	16.6	12.1	14.1	15.2	62.9	62.3	55.9	59.8	63.8
	2-3	14.7	14.7	15.0	14.7	14.5	20.7	19.3	15.9	16.4	18.7	63.5	62.7	56.0	60.1	64.5
0.25	1-3	16.2	11.9	16.3	12.1	16.5	22.8	20.9	16.0	17.5	18.9	65.5	66.7	59.5	63.2	67.6
	2-3	22.8	9.0	21.7	8.8	22.4	33.1	30.6	24.5	26.7	28.6	80.3	80.4	73.1	78.6	81.8
0.50	1-3	23.9	9.1	23.8	10.0	23.5	35.2	32.6	27.3	29.3	30.7	81.5	81.9	73.9	80.1	82.5
	2-3	49.4	5.3	47.1	6.4	46.6	65.4	62.5	55.3	58.7	60.1	97.0	96.5	94.9	96.5	97.1
0.75	1-3	57.4	3.6	55.6	4.2	55.6	74.3	72.8	66.1	68.5	69.4	98.4	98.5	96.9	98.5	98.6
	2-3	94.3	1.0	93.0	1.6	92.8	99.1	98.7	97.5	98.4	98.0	100	100	100	100	100
$n=50$																
0.00	1-3	26.7	25.9	26.9	25.9	26.1	39.1	36.0	32.7	33.8	35.0	90.3	90.1	87.5	89.3	90.8
	2-3	24.6	23.8	24.5	23.4	23.9	38.4	35.9	32.7	33.1	34.3	89.2	89.0	87.7	88.9	89.9
0.25	1-3	25.3	19.9	25.6	21.1	25.5	39.7	38.7	31.4	35.2	36.8	91.8	91.1	89.8	91.4	92.0
	2-3	41.8	18.1	40.4	17.5	40.7	57.9	56.5	52.8	54.3	56.3	98.0	98.1	96.9	97.9	98.1
0.50	1-3	41.1	18.3	40.7	18.8	41.5	59.7	57.5	53.3	55.1	56.9	96.8	96.6	95.9	96.4	97.0
	2-3	76.3	9.7	74.4	10.3	74.7	91.5	90.8	86.5	89.2	89.0	99.9	99.9	99.9	99.9	99.9
0.75	1-3	86.0	8.9	84.1	9.4	85.0	96.9	96.7	94.8	96.5	95.6	100	100	100	100	100
	2-3	100	4.5	99.9	5.0	99.8	100	100	100	100	100	100	100	100	100	100
$n=75$																
0.00	1-3	49.3	47.5	48.7	47.6	48.4	73.1	70.8	68.9	69.1	70.6	99.3	99.1	99.4	99.0	99.3
	2-3	50.4	49.6	48.1	48.4	48.5	72.9	72.2	68.5	68.1	67.8	99.4	99.4	99.4	99.5	99.5
0.25	1-3	55.1	48.2	54.5	48.3	54.2	77.4	77.1	73.2	75.2	75.7	99.8	99.7	99.6	99.6	99.8
	2-3	75.1	44.2	74.2	43.5	73.4	92.7	91.9	89.5	90.6	89.9	100	100	100	100	100
0.50	1-3	78.6	43.4	77.1	45.0	78.5	93.4	94.1	91.8	92.7	93.0	99.9	99.9	99.9	99.9	99.9
	2-3	98.1	37.4	97.5	38.4	97.7	99.6	99.6	99.5	99.6	99.6	100	100	100	100	100
0.75	1-3	99.4	36.4	99.5	36.5	99.4	100	99.9	99.9	99.9	99.9	100	100	100	100	100
	2-3	100	31.5	100	33.0	100	100	100	100	100	100	100	100	100	100	100
$n=100$																
0.00	1-3	75.9	78.0	75.8	77.2	75.4	92.9	93.5	92.2	92.0	92.1	100	100	100	100	100
	2-3	77.1	76.9	75.3	75.9	74.6	93.8	93.1	92.1	91.4	92.6	100	100	100	100	100
0.25	1-3	80.9	74.9	79.8	75.2	79.4	95.5	95.6	94.0	94.6	94.5	100	100	100	100	100
	2-3	94.5	73.8	94.2	73.8	94.4	99.5	99.7	99.6	99.6	99.7	100	100	100	100	100
0.50	1-3	96.3	76.6	96.1	77.0	95.7	100	99.9	99.9	99.7	100	100	100	100	100	100
	2-3	100	76.9	100	74.6	100	100	100	100	100	100	100	100	100	100	100
0.75	1-3	100	75.4	100	74.2	100	100	100	100	100	100	100	100	100	100	100
	2-3	100	73.9	100	74.8	100	100	100	100	100	100	100	100	100	100	100
lognormal																
$n=35$																
0.00	1-3	64.5	66.7	61.0	64.5	60.7	63.2	62.8	53.7	57.3	59.0	78.6	79.2	69.0	73.6	78.7
	2-3	65.6	64.5	62.1	63.2	62.0	64.1	63.1	55.3	58.1	59.3	80.1	80.0	70.2	74.7	79.8
0.25	1-3	68.5	63.5	63.8	60.1	63.1	67.0	66.8	56.6	59.4	61.2	82.9	82.5	72.5	77.9	83.8
	2-3	83.7	62.1	79.8	61.0	79.6	82.9	82.4	75.4	77.9	79.2	92.2	92.4	86.2	89.7	92.3
0.50	1-3	83.3	61.9	80.5	60.6	80.2	82.4	80.7	75.2	78.2	79.0	91.4	91.4	85.5	89.1	91.6
	2-3	98.1	58.6	97.6	57.1	97.5	97.8	98.3	95.5	97.0	97.2	99.3	99.2	98.6	99.1	99.2
0.75	1-3	99.4	58.1	98.7	55.3	98.8	99.5	98.9	97.9	98.9	98.9	99.6	99.7	99.1	99.4	99.7
	2-3	100	50.9	100	50.4	100	100	100	100	100	100	100	100	100	100	100
$n=50$																
0.00	1-3	88.8	90.2	87.5	88.2	87.9	88.8	89.1	85.9	86.7	87.3	95.6	95.7	94.3	95.0	95.4
	2-3	88.2	89.5	87.4	88.3	87.7	88.5	88.0	85.0	85.3	86.3	95.6	95.8	94.8	94.9	95.7
0.25	1-3	92.2	88.8	90.8	88.1	90.3	92.4	91.3	88.4	88.7	90.1	97.1	96.6	95.5	96.3	96.9
	2-3	97.7	89.0	96.9	87.9	97.1	97.2	97.0	95.4	96.2	96.0	99.0	99.1	98.3	98.8	98.7
0.50	1-3	97.5	88.6	96.4	86.9	96.7	97.8	97.7	95.3	96.5	96.5	99.4	99.6	99.0	99.4	99.4
	2-3	99.9	87.6	99.9	87.6	99.9	99.9	100	100	99.9	99.9	100	100	100	100	100
0.75	1-3	100	88.5	100	87.4	100	100	100	100	100	100	100	100	100	100	100
	2-3	100	90.1	100	89.8	100	100	100	100	100	100	100	100	100	100	100

5. Real data application

In this section, we present an illustrative real data example and a simulation study based on a diffusion tensor imaging (DTI) data set. This example shows the practical application

Table 5: Empirical powers (as percentages) of all post hoc tests obtained in selected cases in model M5 under lognormal distribution.

		$\mathscr{C}_n(3)$					$\mathscr{D}_n(3)$					$\mathscr{E}_n(3)$				
ρ	Comp	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$
$n = 35$																
0.00	1-3	100	100	99.8	100	99.7	99.9	99.8	98.4	99.4	99.6	89.9	90.0	84.1	87.7	91.3
	2-3	99.9	100	99.7	99.8	99.7	99.8	99.8	98.9	98.8	99.2	89.4	89.5	82.9	86.6	89.6
0.25	1-3	99.9	100	99.8	99.9	99.9	99.9	99.9	99.4	99.7	99.6	92.6	92.4	87.8	91.0	93.5
	2-3	100	100	100	99.9	100	100	100	100	100	100	98.2	97.7	96.7	97.4	97.9
0.50	1-3	100	99.9	100	99.8	100	100	100	100	100	100	97.9	98.1	95.8	97.6	98.6
	2-3	100	100	100	100	100	100	100	100	100	100	99.9	99.9	99.8	99.8	99.9
0.75	1-3	100	100	100	99.9	100	100	100	100	100	100	100	100	99.9	100	100
	2-3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

of the tests considered.

5.1. Analysis of DTI data set

The DTI data were collected at Johns Hopkins University and the Kennedy Krieger Institute, and they are available in the R package `refund` (Goldsmith et al., 2022). The DTI is a magnetic resonance imaging technique providing different measures of water diffusivity along brain white matter tracts. It is used in diseases affecting the brain white matter tissue, such as multiple sclerosis (MS). In the experiment, DTI brain scans were recorded for MS patients at several visits (from 2 to 8). The aim was to assess the effect of neurodegeneration on disability. The DTI was used to determine the fractional anisotropy (FA) tract profiles for the corpus callosum.

We first delete incomplete data in the DTI data set. For illustrative purposes, we then select the observations for $n = 17$ patients (subjects) with $\ell = 4$ successive visits (objects). Since the trajectories for the FA profiles are given at 93 design time points for each patient and each visit, we treat these data as functional data on $[1, 93]$. They are presented in Figure 2. It is of interest to check whether the FA profiles change significantly at different visits. Thus, we would like to test the equality of mean functions of the FA at different visits. First, we test the global null hypothesis (3). The sample mean functions and the pointwise test statistics (4) and (5) are also depicted in Figure 2. We observe that larger (respectively lower) values of these statistics correspond to larger (respectively smaller) differences between the sample mean functions. For the global null hypothesis (3), the p-values of the $\mathscr{P}_2$ and $\mathscr{B}_2$ $\mathscr{C}_{17}(4)$ tests and the $\mathscr{B}_1$ $\mathscr{D}_{17}(4)$ test are 0.346, 0.285 and 0.001 respectively, while the p-values of the remaining tests considered in Section 3 are equal to zero. The inadequate p-values of the $\mathscr{P}_2$ and $\mathscr{B}_2$ $\mathscr{C}_{17}(4)$ tests can be explained by the fact that these tests may be extremely conservative, which was observed in the simulation study based on the DTI data set presented in the next section (see also Section 4). The other tests reject the null hypothesis, indicating significant differences in the mean functions of the FA at different visits. From Figure 2, it seems that the values of mean functions increase at subsequent visits. Since most of the tests rejected the global null hypothesis, we know that there are significant differences between visits, but we do not know where exactly they occur. Thus, we perform tests for pairwise comparisons. We use Tukey's contrasts (Tukey,

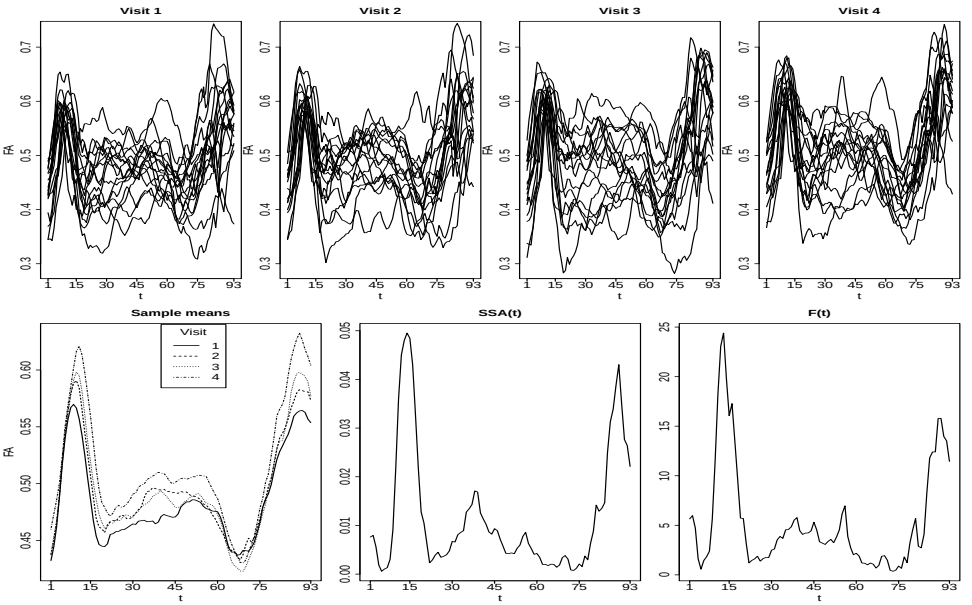


Figure 2: Trajectories for the FA profiles at four visits (top panels); sample mean functions of the FA at different visits and pointwise test statistics (bottom panels).

Table 6: P-values (as percentages) of all tests for pairwise comparisons for the DTI data set.

	$\mathscr{C}_{17}(4)$					$\mathscr{D}_{17}(4)$					$\mathscr{E}_{17}(4)$				
Comp	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$
1-2	6.0	100	6.6	100	5.4	8.4	29.4	19.2	24.0	10.8	20.4	15.6	43.8	9.6	13.8
1-3	30.0	100	34.8	100	33.0	32.4	31.2	48.6	39.6	36.6	2.4	1.2	30.0	1.8	6.0
1-4	0.0	19.2	0.0	15.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.0	0.0
2-3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
2-4	1.2	100	0.0	100	0.0	1.2	0.0	2.4	2.4	0.6	0.6	0.6	17.4	0.6	0.0
3-4	6.0	100	3.6	100	3.0	8.4	8.4	12.6	10.2	7.8	28.2	19.2	64.2	24.6	33.6

1953), i.e. 1 vs. 2, 1 vs. 3, 1 vs. 4, 2 vs. 3, 2 vs. 4, and 3 vs. 4, where 1, 2, 3, and 4 are the numbers of the visits. Based on the procedure presented in Section 3.3, for each pair of visits, we apply the tests proposed in Section 3.2, and correct the obtained p-values using the Bonferroni method. The corrected p-values are given in Table 6. Most of the tests (except the conservative ones) detect significant differences between visits 1 and 4 and between visits 2 and 4. Thus, it seems that the FA profiles for the last visit have much greater values than those for the first two visits. For the other comparisons, there are almost no significant differences. The rejections of equality of mean functions for 1 vs. 3 and 3 vs. 4 follow from the larger power of the corresponding tests (Section 5.2). The above observations seem to be in line with the expectation that the disease progresses over time.

5.2. Simulation study based on DTI data set

To check the correctness of the above results for the DTI data set, we conduct an additional simulation study based on these data. To mimic the observation given in the data example, we generated the simulation data in the following way:

- we used $\ell = 4$ repeated samples with $n = 17$ functional observations at 93 design time points;
- the covariance function $\gamma(s,t)$, $s,t \in [0,\ell]$ was equal to the sample covariance function (6) for the given data set;
- for checking the type I error control, in each group, the mean function was equal to the sample mean function for the pooled sample of $\ell \cdot n = 4 \cdot 17 = 68$ observations collected from four original samples;
- for the power study, the mean function in the i-th repeated sample was equal to the sample mean function for the i-th sample from the data set, $i = 1,\ldots,\ell$.

For given ℓ and n, the 93-dimensional data were generated from the normal (N), Student (t_5), and chi-square (χ^2_{10}) distributions having expected value and covariance matrix equal to the above sample mean functions and sample covariance function respectively. Thus, we consider symmetric, heavily tailed, and skewed distributions.

For the global null hypothesis (Table 7), we observe that the $\mathscr{P}_2$ and $\mathscr{B}_2\ \mathscr{C}_{17}(4)$ tests are extremely conservative, which was also noted in Section 4 for larger correlation. This conservativeness results in significant loss of power for these tests – the largest empirical power among them is 20.8%, while the other tests have power equal to 100% or slightly smaller. These facts explain the testing results for the global null hypothesis in Section 5.1. Note, however, that some of the tests do not control the type I error level, namely $\mathscr{P}_1\ \mathscr{C}_{17}(4)$; $\mathscr{P}_1\ \mathscr{D}_{17}(4)$; $\mathscr{P}_1$, $\mathscr{P}_2$, and $\mathscr{B}_2\ \mathscr{E}_{17}(4)$. This is due to a smaller number of observations ($n = 17$) and four groups ($\ell = 4$).

Let us now turn to the pairwise comparisons (Table 7). All tests control the FWER at level 5%, but the following tests have a conservative character: $\mathscr{P}_2$ and $\mathscr{B}_2\ \mathscr{C}_{17}(4)$; $\mathscr{P}_2$, $\mathscr{B}_1$, and $\mathscr{B}_2\ \mathscr{D}_{17}(4)$; $\mathscr{B}_1\ \mathscr{E}_{17}(4)$. Once again, the first two of these are extremely conservative. This usually results in a lower power compared with the other tests. These observations are consequences of applying the tests to just two groups and using the Bonferroni correction. For the power study, we consider six cases corresponding to the six pairwise comparisons. The empirical power varies for different comparisons. The largest powers appear for 1-4 and 2-4, which explains the detection of significant differences for these comparisons in the analysis of the DTI data set (Section 5.1). On the other hand, for 2-3, all tests have a smaller power, which is a result of the fact that visits 2 and 3 have the most similar FA profiles. For 1-3 and 3-4, the $\mathscr{P}_1$, $\mathscr{P}_2$, and $\mathscr{B}_2\ \mathscr{E}_{17}(4)$ tests and the $\mathscr{B}_1$ and $\mathscr{B}_3\ \mathscr{C}_{17}(4)$ tests respectively have the largest power; this explains the rejections of equality of mean functions in these cases noted in Section 5.1.

6. Conclusions and future work

We have considered one-way repeated measures ANOVA for functional data, which aims to find significant differences in mean functions representing successive experimental

Table 7: Empirical sizes, FWER and powers (as percentages) of all tests for the global null hypothesis (3) and the pairwise comparisons obtained in the simulation study based on the DTI data set.

	$\mathscr{C}_{17}(4)$					$\mathscr{D}_{17}(4)$					$\mathscr{E}_{17}(4)$				
	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{P}_1$	$\mathscr{P}_2$	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$
Global null hypothesis															
Empirical sizes															
N	9.1	0.0	4.2	0.0	3.8	8.9	5.6	2.4	4.5	3.3	9.1	12.0	2.1	9.0	4.1
t_5	10.9	0.0	3.4	0.0	3.5	11.3	6.1	1.9	5.0	3.2	11.2	13.5	2.1	9.1	3.7
χ^2_{10}	9.6	0.0	4.9	0.0	4.3	9.8	6.1	2.9	5.4	3.7	8.6	10.7	3.1	8.3	4.2
Empirical powers															
N	100	8.1	100	11.4	100	100	100	99.9	100	100	100	100	100	100	100
t_5	99.8	17.8	99.2	20.8	99.0	100	99.8	97.5	99.7	99.1	100	100	98.6	100	99.9
χ^2_{10}	100	8.4	100	11.1	99.9	100	100	99.8	99.9	99.9	100	100	99.9	100	100
Pairwise comparisons															
Empirical FWER															
N	4.7	0.0	6.1	0.0	5.2	3.7	1.6	2.3	1.3	4.4	3.8	5.8	0.6	4.8	4.3
t_5	4.7	0.0	4.5	0.0	4.5	4.2	1.2	1.0	1.4	3.5	4.0	5.2	0.3	3.2	2.6
χ^2_{10}	4.2	0.0	5.3	0.0	5.1	2.8	0.7	1.5	0.8	3.9	4.2	6.6	0.5	4.7	4.5
Empirical powers															
N															
1-2	55.4	0.0	55.3	0.0	53.7	59.0	37.0	36.8	33.1	53.3	56.2	63.5	18.7	59.8	57.6
1-3	30.4	0.0	31.4	0.1	30.5	33.6	23.3	19.3	19.9	30.6	69.3	74.7	35.5	69.7	72.2
1-4	100	20.7	100	28.5	100	100	99.9	99.9	99.9	100	100	100	97.8	100	100
2-3	3.5	0.0	4.0	0.0	3.8	3.5	1.1	1.2	1.1	3.4	9.9	12.6	1.4	11.6	10.3
2-4	90.6	0.8	92.4	1.1	91.4	92.6	85.0	82.5	82.6	90.8	96.9	98.0	73.2	97.4	97.7
3-4	59.3	0.0	67.4	0.0	66.9	57.0	48.4	48.7	46.9	60.5	56.6	62.8	21.2	56.9	56.6
t_5															
1-2	64.6	0.0	54.8	0.0	56.3	71.2	47.4	42.0	42.7	60.7	75.4	78.4	23.9	71.7	71.2
1-3	38.4	0.2	34.2	0.3	34.0	48.6	33.7	25.0	27.1	39.3	80.7	82.7	44.1	77.0	79.2
1-4	99.0	33.4	98.4	35.8	98.7	100	99.7	98.3	99.2	99.8	99.9	99.9	94.4	99.9	99.9
2-3	5.2	0.0	3.9	0.0	3.4	5.7	2.2	1.4	2.0	3.6	19.9	23.4	3.6	18.1	16.4
2-4	92.7	1.5	88.7	2.8	90.8	95.0	89.8	83.9	87.3	91.8	98.4	99.2	74.1	97.1	97.6
3-4	64.5	0.2	68.5	0.7	68.4	65.4	59.4	52.9	55.6	64.6	70.6	74.1	26.1	67.0	66.3
χ^2_{10}															
1-2	58.5	0.0	58.2	0.0	57.1	62.1	37.6	40.3	34.2	55.1	58.6	67.6	19.3	63.0	60.9
1-3	32.2	0.1	34.2	0.2	32.4	39.1	26.5	20.9	22.1	35.3	73.6	77.7	38.4	73.7	74.7
1-4	100	20.2	100	28.8	100	100	99.8	99.9	99.9	100	100	100	97.7	100	99.9
2-3	4.3	0.0	4.6	0.0	4.6	4.4	2.1	1.8	1.7	4.0	12.5	16.0	2.6	14.0	12.8
2-4	93.4	0.8	93.8	1.0	93.9	94.7	89.2	85.4	86.4	93.6	98.4	98.9	77.5	98.0	98.4
3-4	58.5	0.0	68.7	0.1	67.5	56.4	50.4	49.5	47.8	60.5	58.2	64.4	22.6	60.1	58.9

conditions applied to the same subjects. Despite the evident interest in this analysis, it is rarely considered in the literature. We have investigated tests for any number of repeated samples, not just two as in most previous studies. Moreover, the procedures take into account both mean and variance, which usually results in more powerful tests. In addition to the procedures for testing the global null hypothesis, we also propose a simple post hoc method, which is of great interest since practitioners usually wish to know in which cases the significant differences appear, not just that such differences exist. To approximate the null distribution of test statistics, two permutation and three bootstrap methods were used. These are distribution-free and have achieved promising levels of performance in previous studies. However, for the present problem, they do not perform equally well, and it is not easy to suggest the best test for each case. The simulation results nevertheless indicated that the parametric bootstrap $\mathscr{E}_n(\ell)$ test seems to have very good properties in terms of type I error level control and power. This test also usually performs much better than those based

on the test statistic of Martínez-Camblor and Corral (2011). Nevertheless, the choice of the best test for a given practical problem can be assisted by a simulation study based on available data, as we showed in Section 5, where we applied the tests to an important medical problem. Our methods are implemented in the R package rmfanova, available on CRAN.

There are many possible directions for future work. One is to extend the methodology to more complicated designs than the one-way classification. A second direction is the development of more sophisticated post hoc tests, which may be more powerful than our simple procedure. Moreover, from practical point of view, the present problem with different number of repeated measurements for patients is interesting and warranted. Perhaps, it will be possible to solve this problem by adopting an approach similar to Amro et al. (2021).

Acknowledgements

A part of the calculations for the simulation study was carried out at the Poznań Supercomputing and Networking Center (grant no. 617). We thank the two referees for helpful suggestions and comments for the improvement of this manuscript.

References

Amro, L., Pauly, M., Ramosaj, B., (2021). Asymptotic-Based Bootstrap Approach for Matched Pairs with Missingness in a Single Arm. *Biometrical Journal*, 63, pp. 1389–1405.

Box, G. E. P., (1954). Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification. *The Annals of Mathematical Statistics*, 25, pp. 290–302.

Cuevas, A., Febrero, M., Fraiman, R., (2004). An ANOVA Test for Functional Data. *Computational Statistics & Data Analysis*, 47, pp. 111–122.

Ditzhaus, M., Genuneit, J., Janssen, A., Pauly, M., (2021). CASANOVA: Permutation Inference in Factorial Survival Designs. *Biometrics*, pp. 1–13.

Ferraty, F., Vieu, P., (2006). *Nonparametric Functional Data Analysis: Theory and Practice*. New York: Springer.

Goldsmith, J., Scheipl, F., Huang, L., Wrobel, J., Di, C., Gellar, J., Harezlak, J., Mclean, M.W., Swihart, B., Xiao, L., Crainiceanu, C., Reiss, P. T., (2022). refund: Regression with Functional Data. R Package version 0.1-28, `https://CRAN.R-project.org/package=refund`.

Horváth, L., Kokoszka, P., (2012). *Inference for Functional Data with Applications*. New York: Springer.

Konietschke, F., Bathke, A. C., Harrar, S. W., Pauly, M., (2015). Parametric and Nonparametric Bootstrap Methods for General MANOVA. *Journal of Multivariate Analysis*,

140, pp. 291–301.

Konietschke, F., Pauly, M., (2014). Bootstrapping and Permuting Paired t-Test Type Statistics. *Statistics and Computing*, 24, pp. 283–296.

Kuryło K., Smaga ł. (2023). rmfanova: Repeated Measures Functional Analysis of Variance. R package Version 0.1.0, `https://CRAN.R-project.org/package=rmfanova`.

Martínez-Camblor, P., Corral, N., (2011). Repeated Measures Analysis for Functional Data. *Computational Statistics & Data Analysis*, 55, 3244–3256.

Mrkvička, T., Myllymäki, M., Jílek, M., Hahn, U., (2020). A One-Way ANOVA Test for Functional Data with Graphical Interpretation. *Kybernetika*, 56, pp. 432–458.

Pini, A., Vantini, S., Colosimo, B. M., Grasso, M., (2018). Domain-Selective Functional Analysis of Variance for Supervised Statistical Profile Monitoring of Signal Data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 67, pp. 55–81.

Ramsay, J. O., Hooker, G., Graves, S., (2009). *Functional Data Analysis with R and MATLAB*. Berlin: Springer.

Ramsay, J. O., Silverman, B.W., (2005). *Functional Data Analysis*, 2nd Edition. New York: Springer.

R Core Team, (2022). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. `https://www.R-project.org/`

Smaga, Ł., (2019). Repeated Measures Analysis for Functional Data Using Box-Type Approximation – with Applications. *REVSTAT – Statistical Journal*, 17, pp. 523–549.

Smaga, Ł., (2020). A Note on Repeated Measures Analysis for Functional Data. *AStA Advances in Statistical Analysis*, 104, pp. 117–139.

Smaga, Ł., (2021). One-Way Repeated Measures ANOVA for Functional Data. In: *Data Analysis and Rationality in a Complex World. IFCS 2019. Studies in Classification, Data Analysis, and Knowledge Organization*. T. Chadjipadelis, B. Lausen, A. Markos, T. R. Lee, A. Montanari, R. Nugent (eds.): Springer, Cham., pp. 243–251.

Smaga, Ł., Zhang, J. T., (2020). Linear Hypothesis Testing for Weighted Functional Data with Applications. *Scandinavian Journal of Statistics*, 47, pp. 493–515.

Tukey, J. W., (1953). The Problem of Multiple Comparisons. Princeton University.

Zhang, J. T., (2013). *Analysis of Variance for Functional Data*. London: Chapman & Hall.

Zhang, J. T., Cheng, M. Y., Wu, H. T., Zhou, B., (2019). A New Test for Functional One-Way ANOVA with Applications to Ischemic Heart Screening. *Computational Statistics & Data Analysis*, 132, pp. 3–17.

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 205–218, https://doi.org/10.59170/stattrans-2024-022
Received – 04.08.2023; accepted – 06.11.2023

The shape of an ROC curve in the evaluation of credit scoring models

Błażej Kochański[1]

Abstract

The AUC, i.e. the area under the receiver operating characteristic (ROC) curve, or its scaled version, the Gini coefficient, are the standard measures of the discriminatory power of credit scoring. Using binormal ROC curve models, we show how the shape of the curves affects the economic benefits of using scoring models with the same AUC. Based on the results, we propose that the shape parameter of the fitted ROC curve is reported alongside its AUC/Gini whenever the quality of a scorecard is discussed.

Key words: credit scoring, receiver operational characteristic, AUC.

1. Introduction

Credit scoring is probably the best known application of statistical methods in finance. Usually based primarily on a customer's credit history and application data, typically built using logistic regression, and less often using neural networks or other machine learning methods, a credit scoring model (a scorecard) returns a numerical score that allows the ranking of potential credit customers. Low scores typically indicate increased risk of default on the loan, while high scores imply lower credit risk.

In the past, the primary goal of a credit scorecard was to provide a single binary classification – to inform the decision maker when to accept or reject a loan. Today, the banking industry uses scorecards to rank customers for many other purposes. Credit institutions use scorecards to differentiate loan terms, such as interest rates (a practice referred to as "risk-based pricing"), loan amounts and periods. Credit limits are set and reset based on credit scores. Models drive cross-sell, up-sell and collection strategies. Credit scores feed into the capital adequacy and loss provisioning models.

The receiver operating characteristic (ROC) curve is a graphical plot that shows how well a binary classifier performs at various levels of the discrimination threshold. It is widely used in many domains, including signal processing, medical statistics, psychology, finance and machine learning applications in general. The ROC curve is

[1] Department of Statistics and Econometrics, Faculty of Management and Economics, Gdańsk University of Technology, Gdańsk, Poland. E-mail: blakocha@pg.edu.pl. ORCID: https://orcid.org/0000-0001-8502-931X.

plotted in the unit square with the false positive proportion (FPP) on the X-axis and the true positive proportion (TPP) on the Y-axis. It represents combinations of FPP and TPP for many possible thresholds (cut-off points). Credit scorecards are set up to identify potential non-payers ("defaults" or "bad" customers), therefore in this context, FPP is the share of non-defaulters ("goods") who scored below a given cut-off among all goods, and TPP is the share of bads who scored below the same given cut-off among all bads.

The AUC (area under the ROC curve) or its rescaled version, the Gini coefficient, are typical ways of assessing the discriminatory power of a credit scorecard. The AUC can be thought of as the probability of correctly assigning a worse score to a bad case, given one random good case and one random bad case (Hanley and McNeil, 1982). The AUC can vary from 0 to 1, where an uninformative scorecard has an AUC of 0.5, while a perfect scorecard has an AUC of 1. Credit scoring practitioners often prefer the Gini coefficient (equivalent to Somers' D statistic), which is a linearly scaled version of the AUC: uninformative scoring has a Gini of 0, while perfect scorecards have a Gini of 1.

Although the AUC is widely used by both practitioners and academics to assess the discriminatory power of credit scoring models, it has several serious drawbacks. Lobo, Jiménez-Valverde and Real (2008) summarise the drawbacks from a biostatistical perspective. From a credit scoring standpoint, the most important disadvantage of AUC can be described as "lift inconsistency". As Idczak (2019) and Řezáč and Koláček (2012) have shown, scorecards with the same AUC/Gini can have different efficiencies in reducing the default rate once a percentage of the worst applicants is rejected. This is related to Hand's (2009) broader claim that AUC is "fundamentally incoherent in terms of misclassification costs".

To address the shortcomings of the AUC, alternative measures are sometimes proposed. Řezáč and Koláček (2012) suggest using a summary of the lift curve at all possible cut-off points. Hand (2009) introduces an h-measure, which evaluates the model based on the distribution of likely values of the classification cost. Verbeke et al. (2012), Verbraken, Verbeke and Baesens (2013) and Garrido, Verbeke and Bravo (2018) introduce profit-based measures.

In this paper, we propose to take a different route. Often, if not always, practitioners will use the same credit scoring model for multiple purposes. Some objectives are better served by one scorecard, while other objectives are better served by another. Rather than deciding priorities for potential users or ranking different objectives, we suggest summarising the scoring model and the associated ROC curve by two numbers instead of one: (1) the AUC/Gini, which measures the area under the curve, and (2) a measure that summarises the shape of the curve. When discussing the quality of the model, two measures could be provided.

The remainder of the paper is organised as follows: in the next section, we present a motivating example of two intersecting ROC curves representing models that excel at different tasks. Then, a binormal ROC curve model is introduced, which produces

smooth, idealised ROC curves with only two parameters: AUC/Gini and the shape/symmetry parameter. The minimal distance fitting of binormal ROC curves to empirical ROC data is then described. This approach is subsequently demonstrated using two empirical examples from the literature. Finally, we discuss our conclusions.

2. When two ROC curves intersect – a motivating example

Several authors (e.g. Adams and Hand, 1999; Řezáč and Koláček, 2012) note that the problem with comparing the two AUCs becomes most apparent when the ROC curves of two scorecards intersect. Idczak (2019) and Tang and Chi (2015) provide empirical illustrations.

In Figure 1, we present a motivating (hypothetical) example of two scorecards (labelled I and II) with identical areas under their ROC curves (Gini = 0.7, AUC = 0.85) but different shapes. The intersecting ROC curves have been simulated using the binormal ROC formulas introduced in the next section.

We will now evaluate the scorecards based on how much the proportion of the customers who fail to repay their loans ("bad rate") in the loan portfolio will be lowered when the cut-off point is introduced and loan applications below cut-off are rejected. We assume here and in the examples below that the bad rate in the total scored population of loan applicants is 8%, which is a typical bad rate in the consumer finance market. However, analogous results can be obtained for population bad rates at other levels.

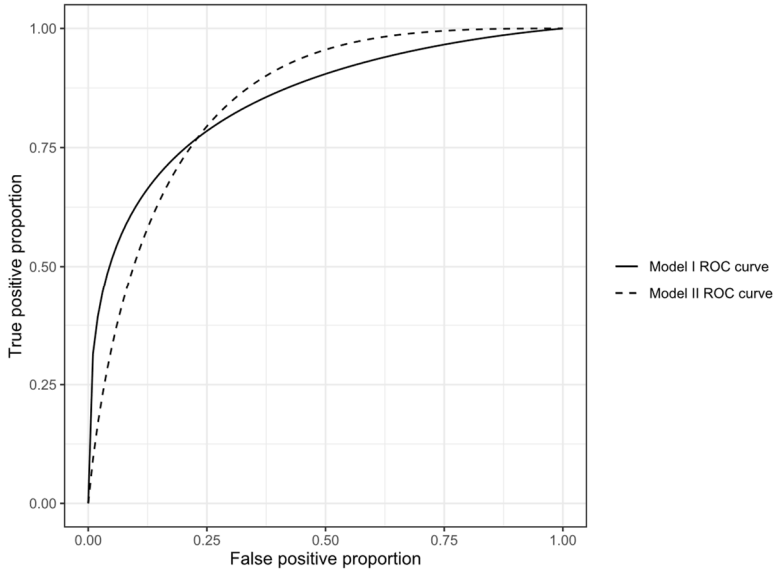


Figure 1: ROC curves for two scoring models with Gini=0.7. Model I is an example of a TPP-asymmetric ROC, and model II of a TNP-asymmetric ROC

If our strategy was to reject 10% of customers with the lowest scores, Model I would be preferred. It would have reduced the bad rate by 49.7% (after the introduction of the cut-off, and rejecting 10% of the applications, the share of defaulted customers would fall from 8% to 4.02%). Model II would have performed much worse in this setup, with a reduction of only 35.7% (from 8% to 5.15%).

The situation would be quite different if the portfolio manager were to select the top 10% of customers and offer them, for example, an increase in their credit limit. In such a case, Model II would prove to be much more effective. If selected using Model II, the top 10% of clients would have a default rate of only 0.04% (~200 times lower than the average). In comparison, the top 10% selected using Model I would have a default rate of 0.95% (only ~8 times lower than the average).

The difference between the curves is apparent on visual inspection. The ROC curve for Model I is more inclined to the OY axis, and the ROC curve for Model II is more inclined to the top line of the unit square (whose equation is y=1). The inclination indicates where the model performs best. Model I excels at identifying the worst of the bad cases (inclination to the OY axis), while Model II wins at picking the best of the good customers (inclination to the y=1 line). Outside of the field of credit scoring, such curves are referred to as asymmetric ROC curves (Killeen and Taylor, 2004, Bhattacharya and Hughes, 2015). Model I is an example of a TPP-asymmetric ROC and Model II of a TNP-asymmetric ROC, where TNP stands for "true negative proportion".

As this simple motivating example shows, two scorecards with the same AUC may be far from equivalent. Which of the models is most effective depends on the intended use in a particular case. Since credit scorecards have more than one specific application nowadays, they need to be assessed at many thresholds for various purposes. It would be virtually impossible to devise a single metric summarising their effectiveness. The h-measure – a popular alternative to AUC – is not a step forward in this case. The Monte Carlo simulation with R and the h-measure package (using the package's default assumptions about the distribution of possible cost values) showed that Model I has an h-measure of 0.325, while Model II has an h-measure of 0.193. The h-measure seems to indicate that Model II is inferior to Model I, which, as we have seen, is not necessarily true.

3. Modelling ROC curves with a binormal model

ROC curve models are mathematical formulae that allow idealised smooth ROC curves to be plotted that approximate real data. Examples of ROC curve models include binormal (Hanley, 1996), bilogistic (Walsh, 1997), bibeta (Chen and Hu, 2016) or bigamma (Dorfman et al., 1997) curves. Such models have 1-4 parameters that determine the course of the curve. ROC curve models are widely used in biostatistics

and signal processing, but their use in credit scoring research has so far been limited. Some prominent exceptions are Blöchlinger and Leippold (2006), Satchell and Xia (2008) and Kürüm et al. (2012), who applied ROC curve models to credit scoring problems. Kochański (2022) evaluates ROC curve models from a credit scoring perspective and examines how well they fit empirical data. He finds that the binormal model is the best fit for the largest number of empirical data sets and can be accommodated to have the Gini coefficient as a parameter. Following this result, we choose the binormal function from among the many ROC curve models. It can be expressed as the formula y=f(x) with two parameters: one representing the area under the ROC curve (or the Gini coefficient) and the other describing the shape (symmetry/inclination) of the curve.

The name "binormal" refers to the assumption that the scores of both good and bad customers are normally distributed (perhaps after a monotone transformation). The standard formula for the binormal ROC curve is as follows:

$$y = \Phi\left(a + b\Phi^{-1}(x)\right), \tag{1}$$

where x is a false positive proportion, y is a true positive proportion, Φ is the CDF of a standard normal variable, Φ^{-1} is its inverse, and a and b are parameters that can be interpreted as functions of the means and variances of the normally distributed scores of goods and bads (Bandos, Guo and Gur, 2017).

It can be shown that the AUC in this case is:

$$AUC = c = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right), \tag{2}$$

which leads to an equivalent alternative parameterisation:

$$y = \Phi\left(\Phi^{-1}(c)\sqrt{1+b^2} + b\Phi^{-1}(x)\right) \tag{3}$$

or, using Gini g = $2c - 1$:

$$y = Bin_{b,g}(x) = \Phi\left(\Phi^{-1}\left(\frac{g+1}{2}\right)\sqrt{1+b^2} + b\Phi^{-1}(x)\right) \tag{4}$$

The ROC curve generated by the binormal model has a shape that depends on the parameter b. In the context of the binormal model construction, it can be thought of as the ratio between the standard deviations of the good and bad scores (which are assumed to follow normal distributions). Curves with b < 1 are TPP-asymmetric, curves with b > 1 exhibit TNP asymmetry. When b = 1, the ROC curve is symmetric about the negative diagonal of the unit square, i.e. the line connecting points (0,1) and (1,0).

Figure 2 illustrates how the parameter b affects the shape of the ROC curve in practice. The figure displays curves with two different levels of Gini and three

different values of the shape parameter *b*. The curves were generated using the binormal model as described by equation 4. Note that the curves in Figure 1 were also drawn using the binormal formula with parameters g = 0.7 (both models), b = 0.77 (Model I), b = 1.3 (Model II).

4. Shape of the ROC curves and efficiency of scoring models

As the motivating example showed, the reduction in the bad rate on accepted loans at the 10% cut-off appears to vary significantly depending on the shape of the ROC curve. Figure 3 extends this example by showing the reduction in the bad rate for a range of values of the shape parameter while maintaining Gini at g = 0.7. Since the empirical ROC curves have shapes corresponding to the range of 0.7 to 1.4 (Kochański, 2022), this graph shows this range. It appears that a 10% cut-off can reduce the bad rate on approved loans from 34% to 52%, depending on the asymmetry of the ROC curve.

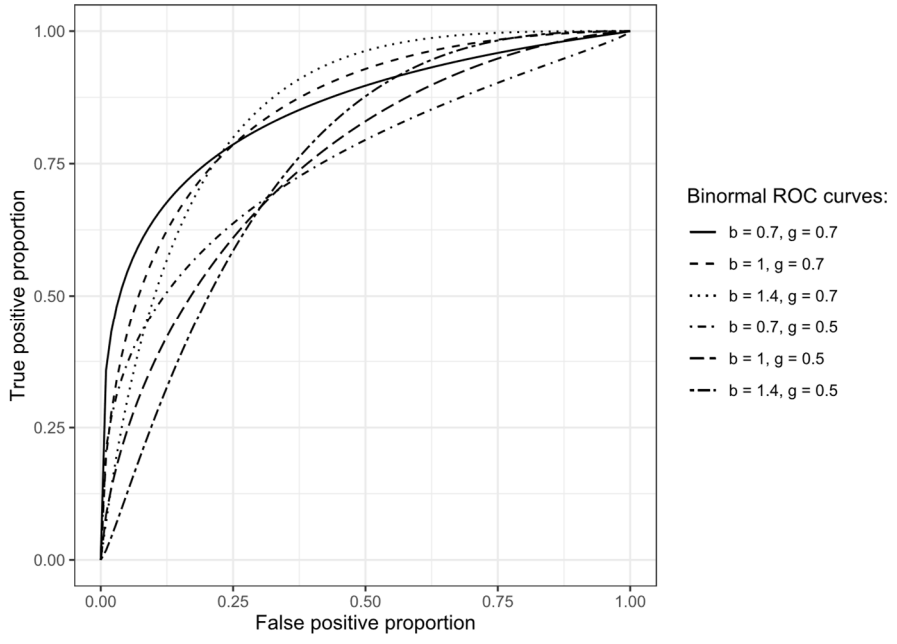


Figure 2: Binormal ROC curves with two Gini levels (0.5 and 0.7) and three values of b parameter

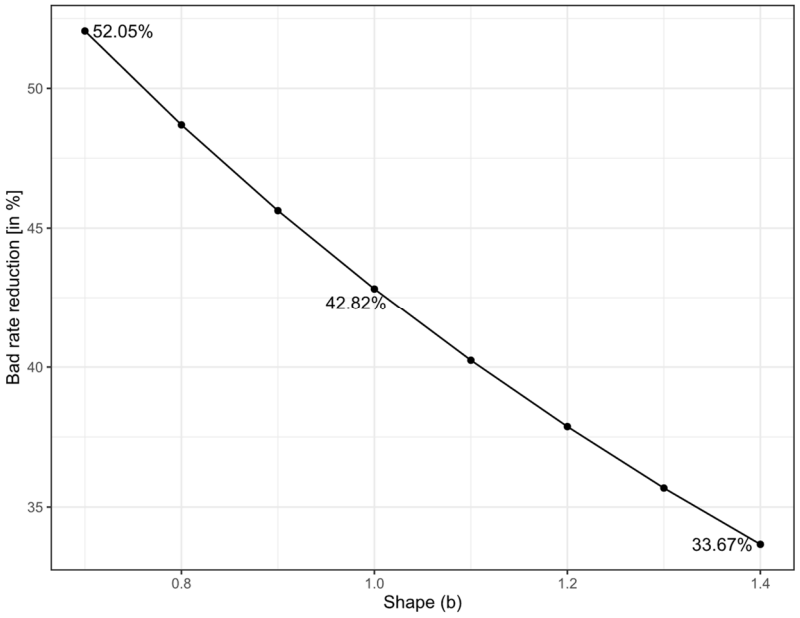


Figure 3: Bad rate reduction at 10% cut-off (90% approval rate) for a Gini=0.7 scorecard and various shapes of the ROC curve as modelled by the binormal model

Figure 4 illustrates this issue from a different perspective. We now fix the desired reduction in the bad rate (40% at the 10% cut-off) and see which Gini/shape combinations produce a reduction at this level. Again, the binormal model was used to answer this question. It turns out that, depending on the shape of the curve, such a reduction is possible with a Gini ranging from 0.56 to 0.75. Three such ROC curves are shown in Figure 5. As can be seen, they all intersect at one point, which corresponds to the cut-off point.

5. Minimal distance fitting of the shape parameter

The simulations presented show that information about the shape of the ROC curve can be crucial for credit decision makers and users of credit scoring. Therefore, we suggest that in business and research practice, in addition to reporting the Gini/AUC value, analysts should report a measure that describes the shape of the curve.

We suggest that ROC curve models could be used to describe the shape of the curve. To apply such a model, one must fit the empirical ROC data to a theoretical curve and find the parameters that give the best fit.

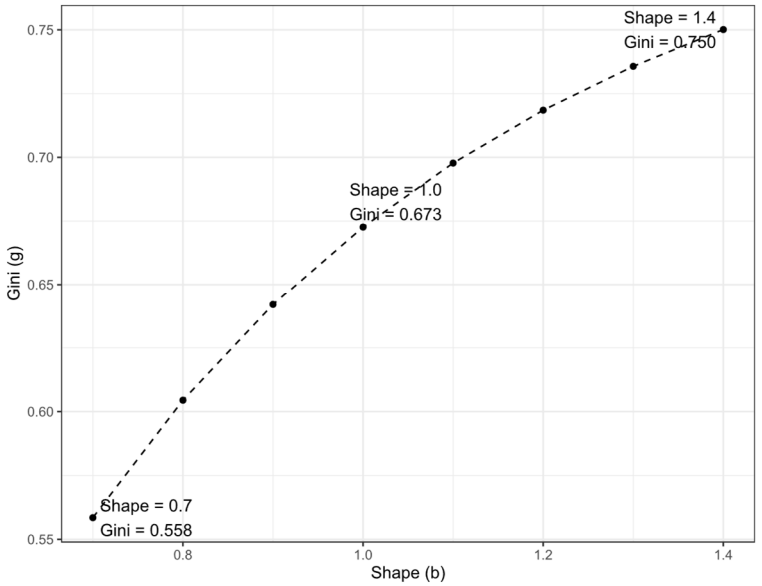


Figure 4: Shape (b) and Gini (g) combinations resulting in a bad rate reduction from 8.0% to 4.8% modelled with binormal ROC curves

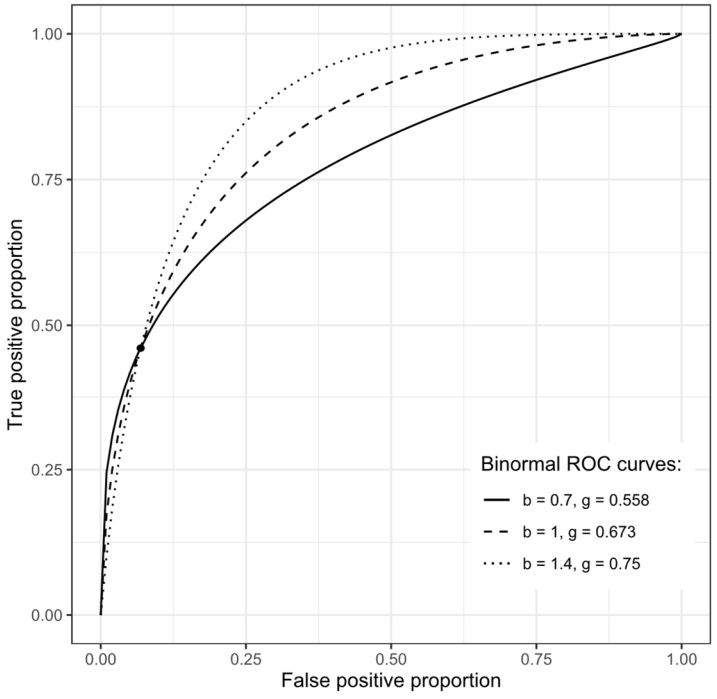


Figure 5: Three binormal ROC curves that reduce the bad rate by 40% when the cut-off point is set to reject 10% of loan applicants

The literature suggests many fitting methods, but for the problem at hand, minimal distance fitting, as described by Hsieh and Turnbull (1996), Davidov and Nov (2012), and Jokiel-Rokita and Topolnicki (2019), seems to be the best approach. For the purposes of this paper, we use the following procedure to fit an ROC curve where the two parameters are the Gini/AUC and the shape parameter:

(1) Empirical ROC data points are collected representing the TPP and FPP for a sufficient set of cut-off points on the ROC graph,

(2) The empirical ROC curve function is derived as a polygon (piecewise linear function) connecting the empirical ROC data points,

(3) The Gini coefficient (or AUC) is calculated using the usual trapezoid method,

(4) The shape parameter of the best fitting ROC curve is found after fixing the AUC.

Step (4) is performed using an appropriately adapted "minimum distance estimation" method. The aim is to minimise the L_2 distance between the empirical ROC curve y=ROC(x) and a theoretical ROC curve function $y = Bin_{b,g}(x)$, when g is already given, and the only value sought is the shape parameter b that gives the best-fitting curve.

Using the notation of Jokiel-Rokita and Topolnicki (2019), the minimised objective function is:

$$\|\xi(b)\| = \int_0^1 \xi^2(b,t)dt$$

where

$$\xi^2(b,t) = \left(\mathrm{ROC}(t) - Bin_{b,g}(t)\right)^2$$

6. Describing the shape of an ROC curve – examples

Rezáč and Rezáč (2011) and Hahm and Lee (2011) are rare examples of papers that provide data on empirical credit-scoring ROC curves in a form of a table. We use these examples to demonstrate the usefulness of describing the shape of ROC curves in assessing the quality of scoring models.

Figure 6 shows data points from Rezáč and Rezáč (2011) with polygon interpolation (dotted line) and fitted binormal function (Gini = 0.451, shape = 1.045). The shape parameter b is close to 1, which means that the curve is almost perfectly symmetric; it is similarly effective in both the high and low score ranges – a result that is usually desirable for a credit scorecard used for risk-based pricing.

The fit is quite good. The square root of $\|\xi(b)\|$ is 0.009, which means that the average (or, more precisely, the root mean square) vertical difference between the empirical and fitted ROC curves is less than one percentage point.

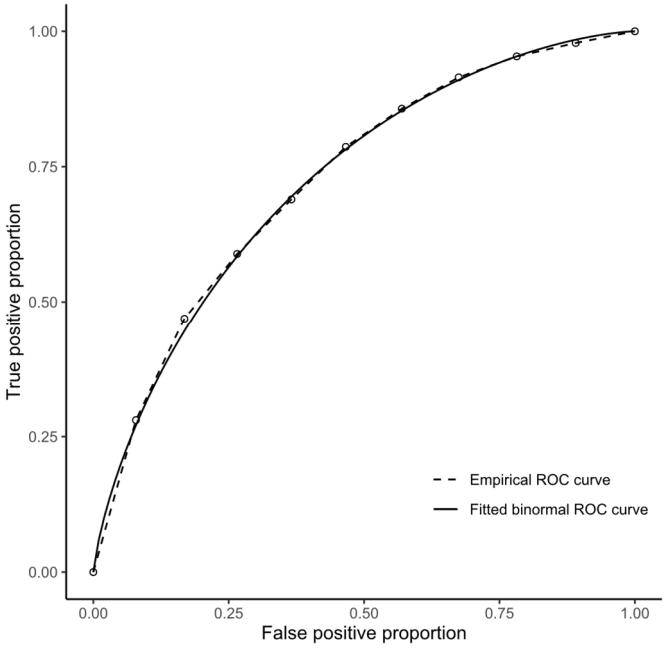


Figure 6: Rezáč and Rezáč (2011) data points with polygon interpolation (dotted line) and fitted binormal function ($g = 0.451$, $b = 1.045$, $\sqrt{\|\xi(b)\|}$=0.009)

The empirical and fitted ROC curves for models from Hahm and Lee (2011), who discuss scorecards developed using credit bureau data in a Korean bank, are shown in Figure 7. The fit of model A is quite good, but visibly worse than the previous case; the fit of model B is slightly better. Model A is a model that uses only negative credit bureau data (only default and late repayment data from other banks), while Model B is a model that also includes positive data (data on timely loan repayments). The inclusion of positive data increases the discriminatory power of the scorecard: AUC increases from 0.841 to 0.869 (Gini from 0.682 to 0.739). The shape parameter of the curve indicates that Model B is less asymmetric than Model A ($b = 0.785$ for Model A and 0.935 for Model B). The change in the shape of the curve implies that the inclusion of positive credit bureau data significantly improves the scorecard's ability to identify the best customers. Indeed, calculations show that both models reduce the default rate by about half when the 8% worst customers are rejected. If the rejection rate is set at 60%, banks can achieve a fivefold reduction using model A and about a tenfold reduction using model B. The apparent advantage of positive credit bureau data is that it allows the best customers to be identified much more easily, while negative data only filters out the worst customers.

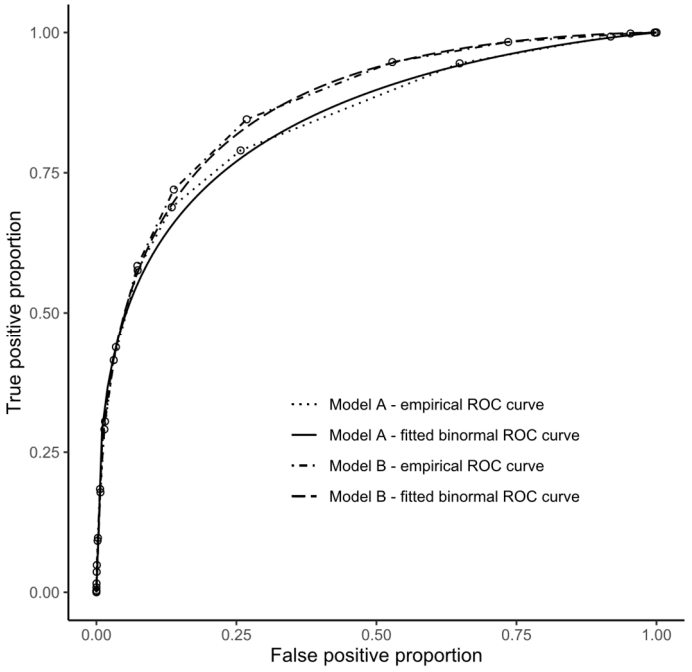


Figure 7: Hahm and Lee (2011) data points for models A and B with polygon interpolation and fitted binormal curves (model A: $g = 0.682$, $b = 0.785$, $\sqrt{\|\xi(b)\|} = 0.014$, model B: $g = 0.739$, $b = 0.935$, $\sqrt{\|\xi(b)\|} = 0.007$).

7. Conclusions

As it has been shown in the examples presented above, scorecards with the same AUC have different effectiveness in different score ranges. In this paper, we propose that when discussing the quality of a scorecard, we should not only provide the AUC/Gini, but also a figure that summarises the shape of the curve.

This approach can be useful when discussing the quality of existing or newly created scoring models. Knowing the shape of the curve allows one to assess which tasks the model is best suited for. Measuring the shape of the ROC curve can also be useful when evaluating individual component variables or component scorecards included in the combined master model. Reporting the shape of the ROC curve allows analysts to assess whether an individual variable or component scorecard owes its separation strength to its ability to find the best or worst customers, or – if the curve is symmetrical – it is equally efficient at both tasks.

In the current practice of lending businesses, the use of scoring models is multifaceted – they are used not only to identify the worst customers, but also to set the prices and terms, demand collateral, drive customer relationship and collection actions,

support loan provisioning procedures and capital adequacy calculations. Using only one number (AUC/Gini) or focusing only on one cut-off is simplistic and may not be sufficient to assess the separation power of a model. However, descriptive measures are needed to summarise the strength of the scorecard. Presenting two numbers instead of one seems to be a useful compromise between synthesis and attention to detail.

References

Adams, N. M., Hand, D. J., (1999). Comparing classifiers when the misallocation costs are uncertain. *Pattern Recognition*, Vol. 32(7), pp. 1139–1147.

Bandos, A. I., Guo, B., Gur, D., (2017). Estimating the area under ROC curve when the fitted Binormal Curves Demonstrate Improper Shape. *Academic Radiology*, Vol. 24(2), pp. 209–219.

Blöchlinger, A., Leippold, M., (2006). Economic benefit of powerful credit scoring. *Journal of Banking & Finance*, 30(3), pp. 851–873.

Chen, W., Hu, N., (2016). Proper bibeta ROC model: algorithm, software, and performance evaluation, in: *Medical Imaging 2016: Image Perception, Observer Performance, and Technology Assessment. Medical Imaging 2016: Image Perception, Observer Performance, and Technology Assessment*, SPIE, pp. 97–104.

Davidov, O., Nov, Y., (2012). Improving an estimator of Hsieh and Turnbull for the binormal ROC curve. *Journal of Statistical Planning and Inference*, Vol. 142(4), pp. 872–877.

Dorfman, D. D., Berbaum, K. S., Metz, C. E., Lenth, R. V., Hanley, J. A., Dagga, H. A., (1997). Proper receiver operating characteristic analysis: The bigamma model. *Academic Radiology*, 4(2), pp. 138–149.

England, W. L., (1988). An exponential model used for optimal threshold selection on ROC curves. *Medical Decision Making*, Vol. 8(2), pp. 120–131.

Garrido, F., Verbeke, W., Bravo, C., (2018). A Robust profit measure for binary classification model evaluation. *Expert Systems with Applications*, Vol. 92, pp. 154–160.

Hahm, J.-H., Lee, S., (2011). Economic effects of positive credit information sharing: the case of Korea. *Applied Economics*, Vol. 43(30), pp. 4879–4890.

Hand, D. J., (2009). Measuring classifier performance: A coherent alternative to the area under the ROC curve. *Machine Learning*, Vol. 77(1), pp. 103–123.

Hand, D. J., Anagnostopoulos, C., (2013). When is the area under the receiver operating characteristic curve an appropriate measure of classifier performance? *Pattern Recognition Letters*, Vol. 34(5), pp. 492–495.

Hanley, J. A., (1996). The Use of the "Binormal" Model for Parametric ROC Analysis of Quantitative Diagnostic Tests. *Statistics in Medicine*, 15(14), pp. 1575–1585.

Hanley, J. A., McNeil, B. J., (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, Vol. 143(1), pp. 29–36.

Hsieh, F., Turnbull, B. W., (1996). Nonparametric and semiparametric estimation of the receiver operating characteristic curve. *The Annals of Statistics*, Vol. 24(1), pp. 25–40.

Hughes, G., Bhattacharya, B., (2013). Symmetry Properties of Bi-Normal and Bi-Gamma Receiver Operating Characteristic Curves are Described by Kullback-Leibler Divergences. *Entropy*, Vol. 15(4), pp. 1342–1356.

Idczak, A. P., (2019). Remarks on statistical measures for assessing quality of scoring models. *Acta Universitatis Lodziensis. Folia Oeconomica*, Vol. 4(343), pp. 21–38.

Jokiel-Rokita, A., Topolnicki, R., (2019). Minimum distance estimation of the binormal ROC curve. *Statistical Papers*, Vol. 60(6), pp. 2161–2183.

Killeen, P. R., Taylor, T. J., (2004). Symmetric receiver operating characteristics. *Journal of Mathematical Psychology*, Vol. 48(6), pp. 432–434.

Kochański, B., (2022). Which curve fits best: fitting ROC curve models to empirical credit-scoring data. *Risks*, Vol. 10(10), p. 184.

Kürüm, E., Yildirak, K., Weber, G.-W., (2012). A classification problem of credit risk rating investigated and solved by optimisation of the ROC curve. *Central European Journal of Operations Research*, 20(3), pp. 529–557.

Lobo, J. M., Jiménez-Valverde, A., Real, R., (2008). AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, Vol. 17(2), pp. 145–151.

Řezáč, M., Koláček, J., (2012). Lift-based quality indexes for credit scoring models as an alternative to Gini and KS. *Journal of Statistics: Advances in Theory and Applications*, Vol. 7(1), pp. 1–23.

Řezáč, M., Řezáč, F., (2011). How to Measure the Quality of Credit Scoring Models. *Czech Journal of Economics and Finance (Finance a úvěr)*, Vol. 61(5), pp. 486–507.

Satchell, S. Xia, W., (2008). Analytic models of the ROC Curve: Applications to credit rating model validation, in G. Christodoulakis and S. Satchell (eds). *The Analytics of Risk Model Validation*. Burlington: Academic Press (Quantitative Finance), pp. 113–133.

Tang, T.-C., Chi, L.-C., (2005). Predicting multilateral trade credit risks: comparisons of Logit and Fuzzy Logic models using ROC curve analysis. *Expert Systems with Applications*, Vol. 28(3), pp. 547–556.

Verbeke, W., Dejaeger, K., Martens, D., Hur, J., Baesens, B., (2012). New insights into churn prediction in the telecommunication sector: A profit driven data mining approach, Vol. 218(1), pp. 211–229.

Verbraken, T., Verbeke, W., Baesens, B., (2013). A Novel Profit Maximizing Metric for Measuring Classification Performance of Customer Churn Prediction Models. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 25(5), pp. 961–973.

Walsh, S. J., (1997). Limitations to the Robustness of Binormal Roc Curves: Effects of Model Misspecification and Location of Decision Thresholds on Bias, Precision, Size and Power. *Statistics in Medicine*, 16(6), pp. 669–679.

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 219–227, https://doi.org/10.59170/stattrans-2024-023
Received – 27.09.2022; accepted – 09.02.2024

Monkeypox obeys the (Benford) law: a dynamic analysis of daily case counts in the United States of America

Leonardo Campanelli[1]

Abstract

We analyze, for the first time, the first-digit distribution of the monkeypox daily cases in the United States of America, from May 17 to September 21, 2022. The overall data follow Benford's law, a conclusion substantiated by eight different statistical tests, including the "Euclidean distance test", which has been designed to specifically check Benford's distribution in data. This result aligns with those of other infectious diseases, such as COVID 19, whose Benfordness has already been confirmed in the literature. Daily counts of monkeypox cases, like any other disease evolve in time. For this reason, we analyzed the temporal deviation of monkeypox counts from Benford's law to check for possible anomalies in the temporal series of cases. The dynamic analysis was performed by means of the Euclidean distance test. This is because, to our best knowledge, that is the only statistically valid, Benford-specific test whose underlying estimator has a cumulative distribution function with known analytical properties, and is applicable to small and large samples. This is the case in dynamic analyses, where the number of data points usually starts from small values and then increases in time. No anomalies were detected, which indicates that no (fraudulent) alterations or errors in data gathering took place.

Key words: Benford's law, monkeypox, Euclidean distance statistic, dynamic analysis

1. Introduction

Monkeypox is a viral zoonotic infectious disease caused by a virus in the genus Orthopoxvirus. An ongoing outbreak started on May 6, 2022 in London, United Kingdom. From May 18 onwards, cases were reported worldwide in more than about 100 countries. This is the first time monkeypox has spread outside Central (Congo Basin Clade) and West Africa (West African Clade), where the disease is endemic (WHO, 2022).

There is evidence that the spread of infectious diseases conforms to Benford's law. Indeed, Sambridge et al. (2010) found that the total numbers of cases of 18 infectious diseases reported to the World Health Organization (WHO) by 193 countries worldwide in 2007 follow a Benford's distribution. Recently, Benford's law has been applied to the study of COVID-19 data, in particular to daily, weekly, and cumulative case and death counts of various countries [see, e.g., Sambridge and Jackson (2020), Farhadi (2021), and Campanelli (2022a).] The general result is that the Benford's distribution well describes the first-digit distributions of COVID-19 data for most of the countries and, then, it can be used to flag "anomalies" in the data of specific countries.

[1] All Saints University School of Medicine, Canada. E-mail: leonardo.s.campanelli@gmail.com.
ORCID: https://orcid.org/0000-0002-7200-9990.

Benford's law (Benford, 1938) is an empirical statistical law according to which the probability $P_B(d)$ of occurrence of the first significant digit d in "particular" data sets is

$$P_B(d) = \log\left(1 + \frac{1}{d}\right). \tag{1}$$

Although it is now known that some distributions satisfy Benford's law [see, e.g. Morrow (2014) and references therein] and that particular principles lead to the emergence of the Benford's phenomenon in data (Hill, 1995a, 1995b, and 1995c), no general criteria has be found that fully explain when and why Benford's law holds for a "generic" set of data.

Although much work is still needed to understand the theoretical basis of the law, the number of its applications has grown in the last few decades [for theoretical insights and general applications of Benford's law see, e.g. Miller (2015)]. Probably, the most famous applications are to detecting tax (see e.g. Nigrini, 1996), campaign finance (see e.g., Cho and Gaines, 2007), and election (see e.g. Roukema, 2013) frauds. Other interesting applications are in image processing (Pérez-González et al., 2007), where Benford's law can be used to test whether or not the image has been compressed, in natural sciences, where the law has been shown to hold for geophysical observables such as the depths of earthquakes (Sambridge et al., 2010), and in cryptology, where it can be used to examine the truthfulness of undeciphered numerical codes (Wase, 2021, Campanelli, 2022b). [2]

The aim of this paper is to assert if the data relative to the monkeypox daily counts in the United States of America (USA) comply or not with Benford's law. The motivation behind this is that, as already noticed, there are already sufficient indications that the number counts of (confirmed and/or death) cases for other infectious diseases (notably COVID 19) follow a Benford distribution. Therefore, a departure of monkeypox data from Benford's law could signal an anomaly, and eventually a fraud, in the data.

2. Method

It is well known that the compliance of data sets to Benford's law improves as the range of the data increases. Daily and cumulative death cases by country are then not appropriate when checking for the compliance of the monkeypox first-digit distributions to Benford's law because there have been only few tens of deaths worldwide since the start of the outbreak (WHO, 2022). Another possibility would be the use of cumulative confirmed case counts. The disadvantage of using this type of data is that as cumulative case numbers begin to flatten (e.g. after a monkeypox "wave" has passed), first digits tend to become all the same, thus distorting relative digit frequencies. In order to overcome this problem, we will only analyze the data on daily confirmed cases by country. However, the only country with daily case numbers which extend on a statistically appreciable range is the USA: Here, the data cover about three orders of magnitude, while in all the other countries affected

[2] The number of articles, books, and other resources related to Benford's law is enormous. Only in 2021 (the year preceding the writing of the present article), for example, the total number of articles, preprints, proceedings, research reports, books chapters, bachelor, master and PhD theses directly connected to Benford's law was (greater than) 131 (Berger et al., 2009). The interested reader can refer the "Benford online bibliography" (Berger et al., 2009) for an up-to-date collection of Benford's-law-related works.

by monkeypox they extend at most on two (WHO, 2022). Accordingly, we will focus our analysis on the daily case counts from the USA.

3. Results

In order to check the conformance of monkeypox data to Benford's law, we will use the "Euclidean distance test", which has been recently proposed by the author to specifically quantify the goodness of fit of a data sample to Benford's law (Campanelli, 2022c). The reasons behind this choice are discussed in Section 4.

3.1. Overall analysis

The Euclidean distance test is based on the Euclidean distance estimator d_N^*, first introduced by Cho and Gaines (2007) and then analyzed by Morrow (2014),

$$d_N^* = \sqrt{N \sum_{d=1}^{9} [P(d) - P_B(d)]^2}, \tag{2}$$

where $P(d)$ is the observed first-digit frequency distribution of a sample of size N. The (empirical) cumulative distribution function (CDF) of the Euclidean distance statistic found by the author (Campanelli, 2022c) allow us to evaluate p values as $p = 1 - \mathrm{CDF}(d_N^*)$.

Data of the 2022 USA monkeypox outbreak are from the Centers for Disease Control and Prevention (CDC, 2022) and are updated to September 21, 2022. They are the confirmed daily cases reported to the CDC since May 17, 2022, the start of the response to the current outbreak. They include either the positive laboratory test report date, CDC call center reporting date, or the case data entry date into CDC's emergency response common operating platform.

In Table 1, we show the range of daily cases, $[\min, \max]$, the number of days, N, the Euclidean distance, d_N^*, and the corresponding p value. In the left panel of Figure 1, instead, we show the observed first-digit frequency distribution of daily case counts superimposed to Benford's law. As it is clear from the table and figure, the data comply with Benford's law at a high level of significance.

Table 1: The Euclidean distance d_N^* in Equation (2) and its corresponding p value for the first-digit distribution of the monkeypox daily case counts in the USA. Also indicated are the range of cases, $[\min, \max]$, and the number of days, N. Counts are from the CDC (2022) and are updated to September 21, 2022. The last three columns show the reduced χ^2 score, $\chi^2_{\mathrm{red}} = \chi^2/\nu$, the number ν of degrees of freedom, and the p value, $p(\chi^2)$, of the χ^2 statistic defined in Equation (4).

Range	N	d_N^*	p	χ^2_{red}	ν	$p(\chi^2)$
$[1,916]$	125	1.0031	0.284	0.5462	76	0.9996

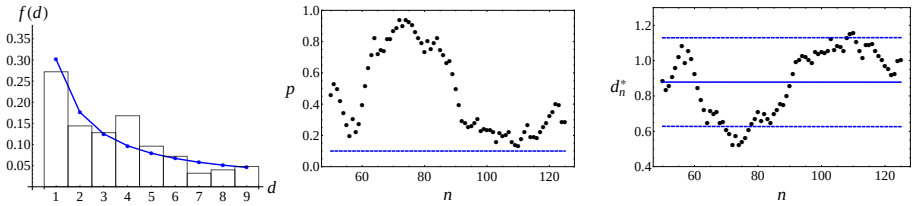


Figure 1: *Left panel.* Observed first-digit frequencies of the monkeypox daily case counts in the USA. The (blue) continuous line represents Benford's law. *Middle panel.* p values of the Euclidean distance statistic d_n^* as a function of the number of data points n (number of days). The (blue) dashed line is $p = 0.10$. *Right panel.* The Euclidean distance statistic as a function of n. The (blue) continuous line represents the expected value of d_n^* for a Benford's distribution, while the (blue) dashed lines show the corresponding one-sigma interval.

3.2. Dynamic analysis

Since the daily counts relative to monkeypox, as well as to any other infectious disease, evolve in time, it is interesting and statistically befitting to quantify the deviation of the timeline of those counts from Benford's law. Indeed, a dynamic data analysis of the chronology of the counts better captures the statistical properties of the spread of a disease.

Such a dynamic analysis can be performed by considering the following χ^2 statistic:

$$\chi^2 = \sum_{n=N'}^{N} \left(\frac{d_n^* - \overline{d_n^*}}{\sigma_n} \right)^2 . \tag{3}$$

Here, $\overline{d_n^*}$ and σ_n are the expected value and standard deviation of the Euclidean distance statistic for the Benford's distribution (Campanelli, 2022c), while d_n^* is the value of the observed Euclidean distance statistic for n data points (the ordinal number of days in our case).

As already noticed, the compliance to Benford's law improves as the range of the data increases. For this reason, we let the sum in Equation (4) to begin from N', the day starting from which the data range extends at least on two orders of magnitude (in the case at hand, $N' = 50$). The number ν of degrees of freedom for the χ^2 statistic is then $\nu = N - N'$.

In the middle and right panels of Figure 1 we show, respectively, the p values and scores of the Euclidean distance statistic d_n^* as a function of n. The (blue) continuous line in the right panel represents the expected value of d_n^* for a Benford's distribution, while the (blue) dashed lines show the corresponding one-sigma interval.

As it is clear from the figure, the null hypothesis of conformance to Benford's law can never be rejected at a 10% level of significance. Moreover, the values of the observed d_n^* are relatively close to the ones expected for a Benford's distribution. This closeness can be quantified by using Equation (4), which gives $\chi^2 = 41.51$ for 76 degrees of freedom. This corresponds to a reduced χ^2 score as low as $\chi^2_{\text{red}} = \chi^2/\nu = 0.5462$ and to a p value as large as $p(\chi^2) = 0.9996$. These values for the reduced χ^2 and p values, reported in Table 1 for

convenience, show that the temporal series of the monkeypox daily case counts in the USA conforms to the Benford's distribution to a very high significance level.

4. Discussion

The most common test in use for testing whether a numerical sample satisfies Benford's law is the Pearson's χ^2 (with 8 degrees of freedom), whose estimator is

$$\chi^2_8 = N \sum_{d=1}^{9} \frac{[P(d) - P_B(d)]^2}{P_B(d)}. \tag{4}$$

The well-known problem with this statistic is the low power at small N ($N < 100$) (see, e.g., Morrow, 2014).

To overcome this problem, other more powerful test statistics like the Kolmogorov-Smirnov (Kolmogorov, 1933) and Kuiper (Kuiper, 1960) statistics have been used. However, these statistics have been constructed for continuous distributions and are generally conservative when testing discrete distributions (see, e.g., Morrow, 2014). Only recently, Benford-specific asymptotic test values have been found by Morrow (2014). The Kolmogorov-Smirnov and Kuiper statistics, D_N^* and V_N^* respectively, are defined as follows. Let D_N^+ and D_N^- be $D_N^{\pm} = \max_d[\pm Z(d)]$, where $Z(d) = S(d) - T(d)$, $S(d) = \sum_{d'=1}^{d} P(d')$, and $T(d) = \sum_{d'=1}^{d} P_B(d')$. Then,

$$D_N^* = (\sqrt{N} + \delta_N) \max[D_N^+, D_N^-] \tag{5}$$

and

$$V_N^* = (\sqrt{N} + \nu_N)(D_N^+ + D_N^-). \tag{6}$$

Here, the correction terms $\delta_N = 0.12 + 0.11/\sqrt{N}$ and $\nu_N = 0.155 + 0.24/\sqrt{N}$ have been introduced by Stephens (1970) for the case of continuous distributions to produce accurate test statistics regardless of the sample size. Asymptotic test values for these statistics are given in Table 2. However, since both the asymptotic and the small-sample form of the CDF for the case of a Benford's distribution are still unknown, p values cannot be calculated. A dynamic analysis of monkeypox cases is then not feasible in this case.

Another statistic used to test Benford's law is the so-called "max statistic". It is defined by

$$m_N^* = \sqrt{N} \max_d [P(d) - P_B(d)]. \tag{7}$$

Introduced by Leemis et al. (2000) to specifically test Benford' law, the statistical properties of the max estimator were subsequently analyzed by Morrow (2014), who provided asymptotic test values (reported in Table 2). As for the above two statistics, the Benford-specific form of the CDF is unknown.

Test statistics based on the Cramér–von Mises statistic have been introduced in the literature to test discrete distributions, as the Benford one. Following Lockhart et al. (2007),

we consider the following forms of the Cramér–von Mises statistic,

$$W_N^2 = N \sum_{d=1}^{8} Z^2(d)t(d), \tag{8}$$

$$U_N^2 = N \sum_{d=1}^{9} (Z(d) - \overline{Z})^2 t(d), \tag{9}$$

$$A_N^2 = N \sum_{d=1}^{8} \frac{Z^2(d)t(d)}{T(d)[1-T(d)]}, \tag{10}$$

where $t(d) = [P_B(d) + P_B(d+1)]/2$ with $P_B(10) \doteq P_B(1)$ and $\overline{Z} = \sum_{d=1}^{8} Z(d)t(d)$. The main problem of the above statistics, whose asymptotic test values are presented in Table 2, is that only their asymptotic distributions are known. Therefore, their application to small data samples, which is our case, cannot be completely trusted. [A minor point is that their asymptotic distributions are not known in closed form, although a precise method to find them is known and is based on numerical integration (see, e.g., Lockhart et al., 2007)].

In Table 2, we report the test values for the above test statistics for the first-digit distribution of the monkeypox daily case counts in the USA. The null hypothesis of conformance to Benford's law, for the overall monkeypox data, cannot be excluded at 90% confidence level. This result corroborates our finding that monkeypox data comply with Benford's law at a high statistically significance level.

Table 2: Test values for different test statistics for the first-digit distribution of the monkeypox daily case counts in the USA. Also indicated is the asymptotic critical values at 90% confidence level for each statistic and the corresponding reference.

Test	test val.	crit. val.	Ref.
χ_8^2	9.877	13.362	—
D_N^*	0.691	1.012	Morrow, 2014
V_N^*	1.091	1.191	Morrow, 2014
m_N^*	0.795	0.851	Morrow, 2014
W_N^2	0.162	0.351	Lesperance et al., 2016
U_N^2	0.127	0.163	Lesperance et al., 2016
A_N^2	0.732	1.743	Lesperance et al., 2016

Other goodness-of-fit tests could be used, in principle, to test monkeypox data against Benford's law, such as the Goodman's rule of thumb (Goodman, 2016, Campanelli, 2022d), simultaneous confidence intervals for multinomial probabilities [see Lesperance et al. (2016) for a discussion of seven different simultaneous confidence intervals tests], the mean absolute deviation (MAD) criterion (see, e.g. Nigrini, 2012), and the Friedmann test (Giles, 2013), just to cite a few.

However, it is important to stress the fact that, at least to our knowledge, the only Benford-specific test statistic with known analytical expression for the CDF, valid for ei-

ther small and large data samples, is the Euclidean distance. This makes the Euclidean distance test the ideal one for studying Benford's law in dynamic data, where the number of data points is variable and usually starts from small values.

5. Conclusions

There is an increasing evidence that the number of counts of both death and confirmed cases due to infectious diseases conforms to Benford's law. This is the case, for example, of COVID 19, where extensive analyses have been performed in the last few years. The aim of this paper was to analyze the first-digit distribution of the daily case counts for the ongoing 2022 monkeypox outbreak in the USA.

A global analysis of the data was performed by using 8 different statistical tests, including the "Euclidean distance test", which has been proposed by the author elsewhere to specifically quantify the goodness of fit of a data sample to Benford's law. Our results show that the data comply with Benford's law at a high significance level. This suggests that no manipulations or errors in data collection occurred.

Daily counts of monkeypox cases, and in general death and confirmed cases counts for any infectious disease, evolve in time. In order to follow the spread of monkeypox dynamically, we analyzed the temporal deviation of monkeypox counts from Benford's law. Indeed, a dynamic data analysis of the chronology of the counts could not only flag anomalies but also could frame an anomaly temporally. In the case of monkeypox in the USA, no anomalies were detected, with the temporal series of daily cases conforming to the Benford's distribution to a remarkably high significance level of about 99.96%. The statistical test we used for the dynamic analysis was the Euclidean distance test. The motivation was that, as far as we know, this is the only Benford-specific test with known analytical expression for the cumulative distribution function of the underlying estimator which is valid for either small and large data samples. This last property is strongly required when testing Benford's law in dynamic data, since the number of data points usually starts from small values and than grows in time.

A similar analysis to the one presented in this paper could be applied to both monkeypox counts from other countries when sufficient data become available and/or to future infectious diseases to flag anomalies and fraudulent manipulations either globally or temporally.

References

Benford, F., (1938). The Law of Anomalous Numbers. *Proceedings of the American Physical Society*, 78, pp. 551–572.

Berger A., Hill, T., Rogers, E., (2009). Benford online bibliography. https://www.benfordonline.net/ (accessed on 2023-09-29).

Campanelli, L., (2022a). Breaking Benford's law: A statistical analysis of Covid-19 data using the Euclidean distance statistic. *Statistics in Transition new series*, 24(2), pp. 201–215.

Campanelli, L., (2022b). A Statistical Cryptanalysis of the Beale Ciphers. *Cryptologia*, 47(5), pp. 466–473.

Campanelli, L., (2022c). On the Euclidean Distance Statistic of Benford's Law. *Communications in Statistics - Theory and Methods.* DOI: 10.1080/03610926.2022.2082480.

Campanelli, L., (2022d). Testing Benford's Law: from small to very large data sets. *Spanish Journal of Statistics*, 4, pp. 41–54.

CDC, (2022). U.S. Monkeypox Case Trends Reported to CDC. https://www.cdc.gov/poxvirus/monkeypox/response/2022/ (accessed on 2022-09-24).

Cho, W. K. T., Gaines, B. J., (2007). Breaking the (Benford) Law: Statistical Fraud Detection in Campaign Finance. *Am. Stat.*, 61, pp. 218–223.

Farhadi, N., (2021). Can we rely on COVID-19 data? An assessment of data from over 200 countries worldwide. *Sci. Prog.* ,104, pp. 1–19.

Giles, D. E., (2013). Exact asymptotic goodness-of-fit testing for discrete circular data, with applications. *Chilean Journal of Statistics*, 4(1), pp. 19–34.

Goodman, W., (2016). The promises and pitfalls of Benford's law. *Significance*, 13, pp. 38–41.

Hill, T. P., (1995a). The significant-digit phenomenon. *Am. Math. Mon.*, 102, pp. 322 -327.

Hill, T. P., (1995b). Base-invariance implies Benford's law. *Proc. Am. Math. Soc.*, 123, pp. 887–895.

Hill, T. P., (1995c). A statistical derivation of the significant-digit law. *Stat. Sci.*, 10, pp. 354–363.

Lockhart, R. A., Spinelli, J. J., Stephens, M. A., (2007). Cramér–von Mises statistics for discrete distributions with unknown parameters. *The Canadian Journal of Statistics*, 35, 1, pp. 125–133.

Kolmogorov, A., (1933). Sulla determinazione empirica di una legge di distribuzione. *G. Ist. Ital. Attuari*, 4, pp. 83–91.

Kuiper, N. H., (1960). Tests concerning random points on a circle. *Proc. Koninkl. Nederl. Akad. Van Wettenschappen*, A63, pp. 38–47.

Miller, S. J. (ed.), (2015). *Benford's Law: Theory and Applications*, Princeton: Princeton University Press.

Leemis, L. M., Schmeiser, B. W., Evans, D. L., (2000). Survival Distributions Satisfying Benford's Law.
Am. Stat., 54, pp. 236–241.

Lesperance, M., Reed, J. W., Stephens, M. A., Tsao, C., Wilton, B., (2016). Assessing Conformance with Benford's Law: Goodness-Of-Fit Tests and Simultaneous Confidence Intervals. *PLos ONE* 11(3): e0151235. DOI:10.137/journal.pone.0151235.

Morrow, J., (2014). Benford's Law, Families of Distributions and a Test Basis. *Centre for Economic Performance.* London.

Nigrini, M. J., (1996). A taxpayer compliance application of Benford's law. *Journal of the American Taxation Association*, 18, pp. 72–91.

Nigrini, M. J., (2012). *Benford's Law. Applications for Forensic Accounting, Auditing, and Fraud Detection*, Hoboken, New Jersey: John Wiley & Sons.

Pérez-González, F., Abdallah, C. T., Heileman, G. L., (2007). Benford's Law in Image Processing. IEEE International Conference on Image Processing, pp. 405–408.

Roukema, B. F., (2013). A first-digit anomaly in the 2009 Iranian presidential election. *J. Appl. Stat.*, 41(1), pp. 164–199.

Sambridge, M., Jackson, A., (2020). National COVID numbers - Benford's law looks for errors. *Nature*, 581, p. 384.

Sambridge, M., Tkalčić, H., Jackson, A., (2010). Benford's law in the natural sciences. *Geophys. Res. Lett.*, 37, L22301.

Stephens, M. A., (1970). Use of the Kolmogorov-Smirnov, Cramér–von Mises and Related Statistics Without Extensive Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 32, pp. 115–122.

Wase, V., (2021). Benford's law in the Beale ciphers. *Cryptologia*, 45(3), pp. 282–286.

WHO, (2022). https://www.who.int/emergencies/situations/monkeypox-outbreak-2022 (accessed on 2022-09-24).

STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 229–234

About the Authors

Brodovska Oksana, PhD, is a Doctoral Student of International Economics Department, West Ukrainian National University, Ternopil, Ukraine. She is President of Limited Liability Company Financial Company MAGNAT. The main areas of interests include: inclusive economy, social policy, finance, financial and debt policy, FDI policy.

Campanelli Leonardo is a Lecturer at All Saints University School of Medicine in Canada, where he teaches statistics and physics. He obtained his PhD in Theoretical Physics and Mathematical Methods for Physics in 2004 at the University of Bari (Italy). Dr. Campanelli has published more than 50 research papers in international journals. His primary interest in statistics is data analysis.

Echchekh Adil started his career as a teacher-researcher at the University Louis Pasteur in Strasbourg (1992), the University of Limoges and finally the Ibn Tofaïl University, where he participated in the work of the establishment as a member elected to the University Council through the Pedagogical Commission, the Research Commission, and the management council as a member of the said commissions. In the case of international influence, he participated as a member of the work of the International Association of University Pedagogy, of the French mechanical society, of the French Society of Process Engineering, and an expert member of the CNRST project. Today, he focuses on priority research areas such as water, environment, health, energy, transport, road safety, and artificial intelligence. On the educational level, he is responsible for three fields of study.

El Moury Ibtissam holds a PHD degree in Applied Mathematics and Statistics at the Faculty of Science, Ibn Tofail University, Kénitra, Morocco. As a seasoned researcher, she has contributed significantly to the realms of mathematical and statistical modeling. Mrs. Elmoury's expertise extends across various domains including quality management, project management, logistics, and econometric modeling, underscoring her multidisciplinary approach to problem-solving. With her academic prowess and practical experience, she continues to be an asset in driving innovation and excellence in both academic and industrial spheres.

Fennane Sara is s PhD student specializing in physical sciences, with a master's degree in energy mechanics and fluids from Ibn Tofail University. She also holds a bachelor's degree in physics, which she obtained in 2018 from the same university. With a strong academic background and excellent research skills, Sara is dedicated to making valuable

contributions to her chosen field. Her passion for the subject and analytical mindset drives her pursuit of meaningful insights in her research.

Garg Neha is currently working in the School of Sciences, Indira Gandhi National Open University, New Delhi, India since 2010. Her research interests are in the area of sampling techniques and estimation theory. She has 19 years of teaching and research experience. She has published many research papers in national and international refereed journals. She has reviewed research papers for several national and international journals.

Goyal Anju completed her MSc, MPhil. and PhD in statistics at Panjab University. She is working as an Assistant Professor in Department of Statistics, Panjab University, Chandigarh since March, 2013. She is the Ex-Chairperson of the department of Statistics. Dr. Anju has various profound publication and have presented various papers.

Kacimi Houda, a PhD student in physics, received his baccalaureate degree in physics with a good mention in 2015. In 2017 he passed his university diploma in technology (DUT) at EST Meknes. At the same school, he obtained the professional license in the renewable energetic and energetic efficacy (LP) in 2018 and during his three years at the school of technology he completed internships in many companies. In 2021, he graduated with a master's degree in energetic and fluid mechanics from Ibn Tofail University in Kenitra.

Kochański Błażej is an Assistant Professor at the Department of Statistics and Econometrics within the Faculty of Management and Economics at Gdańsk University of Technology. His expertise lies in banking risk, and he has a professional background that includes roles as a risk specialist, planning and analysis manager, chief risk officer, supervisory board member, and management consultant in the banking sector. During his tenure, he developed credit risk management models, effectively managed credit portfolios, and crafted strategies for mitigating risk among retail clients. Now, as part of his academic journey, Błażej is eager to share his experience, knowledge, and insights through consulting, training, and scholarly publications.

Krężołek Dominik is an Associate Professor at the Department of Demography and Economic Statistics, Faculty of Informatics and Communication, University Economics in Katowice. Simultaneously he holds the position of a senior specialist in the Śląski Centre for Regional Survey at the Statistical Office in Katowice. His main areas of interest include financial risk modelling (mainly extreme risk), portfolio optimization, multivariate statistical analysis, time series analysis, structural modelling, analysis of metals market. He is the author of over fifty scientific publications in Polish and foreign scientific journals. He is also a member of the Polish Statistical Association and Sekcja Klasyfikacji i Analizy danych.

Krysovata Kateryna, PhD in Economics, Associate Professor at the S.I. Yuriy Department of Finance, West Ukrainian National University (Ternopil, Ukraine). The main areas of interests include: international finance, the budget system, customs policy, foreign economic activity, sustainable development of the economy.

Kumar Devendra received his BSc, MSc, MPhil, and PhD degrees in Statistics from Aligarh Muslim University, India. He is currently working as an Associate Professor in the Department of Statistics, Faculty of Mathematical Sciences, at the University of Delhi, India. Before joining Delhi University, Dr. Kumar was working as an Associate Professor in the Department of Statistics at the Central University of Haryana, India. His research interests include order statistics, statistical inference, Bayesian estimation, and distribution theory.

Kumar Maneesh completed his PhD in Statistics from the Central University of Haryana under the supervision of Dr. Devendra Kumar. He is currently working as an Assistant director in Ministry of Statistics and Programme and Implementation, New Delhi, India.

Kuryło Katarzyna is a student in the Data Analysis and Processing field of study at the Faculty of Mathematics and Computer Science at Adam Mickiewicz University in Poznań. Simultaneously, she holds the position of a data scientist at the Supercomputing and Networking Center affiliated to the Institute of Bioorganic Chemistry PAS in Poznań. Her main areas of interest are statistical methods, especially functional data analysis, and applications of machine learning in practical issues.

Niftiyev Ibrahim PhD in Economics, is a full-time lecturer at the Department of International Economics and Business and a research associate at the Research Center for European Economies at the Azerbaijan State University of Economics (UNEC). Dr. Niftiyev's main research interests include natural resource economics, de-industrialization, agricultural economics, industrial policy and renewable energy. His research approaches are both quantitative and qualitative, paying particular attention to statistical and econometric methods such as linear programming, principal component analysis, multiple correspondence analysis and other non-parametric methods. He has authored 26 journal articles, ten conference papers, one book chapter and several working papers. Dr. Niftiyev is a member of the editorial board of the journal SocioEconomics, published by Sumy State University.

Osaulenko Oleksandr is Rector of the National Academy of Statistics, Accounting and Audit, Doctor of Public Administration, Professor, Academician of the National Academy of Science of Ukraine. During 1996-2014, he headed the national statistical office of Ukraine. Author of over 260 scientific papers, including 34 monographs and handbooks on the problems of statistics, public administration and information security. Head of the Specialized Scientific Council of the National Academy of

Statistics, Accounting and Audit. Editor-in-chief of two editorial boards: Statistics of Ukraine and Scientific Bulletin of National Academy of Statistics, Accounting and Audit. His research results became the theoretical and methodological basis for the preparation of long-term programs for development of national statistics and improvement of statistical support management at national and regional levels, and received practical implementation in a number of legislative and other normative legal acts.

Pachori Menakshi has completed her PhD (Statistics) from School of Sciences, Indira Gandhi National Open University, New Delhi, India. She is currently working in the Sanskriti University, Mathura, UP, India. Her research interests are in the area of sampling and medical statistics. She has published many research papers in national and international refereed journals.

Piontek Krzysztof is an Associate Professor of Finance at Wrocław University of Economics and Business, Poland. He is a member of the Department of Financial Investments and Risk Management. He holds a master's degree from Wroclaw University of Science and Technology and a bachelor's, a doctoral and a habilitation degree ("Construction and Verification of Quantile-Based Market Risk Measurement Models") from Wrocław University of Economics and Business. He specializes in using mathematical methods in finance, focusing on time series analysis. Member of The Classification and Data Analysis Section (SKAD) of The Polish Statistical Association (PTS). Chairman of the Board of DB Energy SA and a member of many business organizations.

Shlapak Alla is the Doctor of Economics, the Associate Professor, the Honoured Worker of Education, the Head of the Department of International Economics at Borys Grinchenko Kyiv Metropolitan University, Ukraine. She is the author of about 100 scientific and educational works, including scientific monographs and textbooks on the development of international economic relations. On the initiative and under the leadership of Alla Shlapak, a significant number of departmental and university events were held, including guest lectures by well-known domestic and foreign scholars and practitioners, summer economic schools for young scholars and students of Kyiv secondary schools, international and national scientific conferences, etc. Alla Shlapak is an active participant in national and international social movements. She is a member of the International Anti-Terrorist Unity, an international NGO with the status of a member of the Council of Europe. She is the Honorary President of the Social Justice NGO SOVIST', which has implemented hundreds of projects in the social and humanitarian spheres.

Siswanto Siswanto is an Associate Professor at the Department of Statistics, Faculty of Mathematics and Natural Sciences, University of Hasanuddin. His research field is

spatial statistics, especially spatial in the fields of health sciences, economics and geostatistics. To date, he has written 40 articles. He is also active in participating in statistics forum activities, namely FORSTAT, which is an adequate forum for institutions providing higher education in the field of statistics, in order to improve the quality of statistics providers in Indonesia.

Smaga Łukasz is the Associate Professor at the Department of Mathematical Statistics and Data Analysis at the Faculty of Mathematics and Computer Science of the Adam Mickiewicz University in Poznań. He works scientifically in the fields of statistical hypothesis testing, functional data analysis, and experimental theory. He also deals with the applications of statistical methods and machine learning in practical issues. He cooperates with scientists from the National University of Singapore, TU Dortmund University, Otto von Guericke University, Texas Tech University, Shiga University, Poznań University of Technology, and the Poznań University of Economics.

Sunusi Nurtiti is a Full Professor of Statistics. Her research interest are stochastic processes, statistical inference, probability models, and statistical Process Control. Professor Nurtiti has published 73 research papers in international/national, and conferences. She also published two books. Professor Nurtiti is an active member of Indonesian Mathematical Society.

Syam Ummul Auliyah is a master student of Statistics and Data Science at IPB University. Her research interests are spatial statistical analysis, statistical modelling, and predictive analysis. She actively participates in research with lecturers and has published one research paper in national journal related to spatial regression analysis.

Tripathi Rajnee is an Assistant Professor in University Institute of Sciences, the Department of Mathematics, Chandigarh University, Gharuan, Mohali, Punjab, India. She received the MPhil degree from Awadhesh Pratap Singh University, Rewa (MP), India in 2012 and MSc degree from Awadhesh Pratap Singh University, Rewa (MP), India in 2009. She has completed her PhD degree from Jaypee University of Engineering and Technology, Guna (MP), India in 2018. She has published 30 research papers in various research areas. Her current research area of interest covers Laplace homotopy perturbation method for solving linear and non-linear differential equations (PDEs and ODEs), fuzzy information measures and intuitionistic fuzzy information measures, statistical distributions theory, methods of estimations and simulations, generalization of probability distributions and statistical properties of probability distributions.

Yadav Sapna completed her MSc in Statistics from the Central University of Haryana, India.

Zvarych Iryna, DSc (Economics), Professor, Head of International Economics Department, West Ukrainian National University, Ternopil, Ukraine. The sphere of

her scientific interests today is the global inclusive circular economy; global (international) chains of added value, international production networks. Author of more than 100 scientific publications. She was owner of Research Grant of Ukrainian Ministry of Education and Science, Co-author and Performer of fundamental state-budget scientific research on the topic "National concept of eco-security of society and inclusion of the circular economy in the conditions of the pandemic" (2020–2023). Scholarship holder of the Cabinet of Ministers of Ukraine for young scientists 2020-2022. Executor of Erasmus+ Jean Monnet Higher Education Teaching and Research Project "European inclusive circular economy: post-war & post-pandemic module for Ukraine" 101085640 — EICEPPMU — ERASMUS-JMO-2022-HEI-TCH-RSCH. Foreign internship (international scientific and pedagogical internship): VARNA FREE UNIVERSITY OF CHRONORIZETS HRABRA CERTIFICATE No. C – 9824/ 15.08.2019. In 2023 she passed the e-tutor qualification courses in TU Dresden "Ukraine Digital I: Ensuring academic success in times of crisis". Iryna Zvarych is a Distance Learning Program Local Instructor CERGE-EI Foundation (USA), Academic Year 2022-2023 course "Climate change economics".

GUIDELINES FOR AUTHORS

We will consider only original work for publication in the Journal, i.e. a submitted paper must not have been published before or be under consideration for publication elsewhere. Authors should consistently follow all specifications below when preparing their manuscripts.

Manuscript preparation and formatting

The Authors are asked to use *A Simple Manuscript Template (Word or LaTeX) for the Statistics in Transition Journal (published on our web page:* https://sit.stat.gov.pl/ForAuthors.

- ***Title and Author(s)***. The title should appear at the beginning of the paper, followed by each author's name, institutional affiliation and email address. Centre the title in **BOLD CAPITALS**. Centre the author(s)'s name(s). The authors' affiliation(s) and email address(es) should be given in a footnote.
- ***Abstract.*** After the authors' details, leave a blank line and centre the word **Abstract** (in bold), leave a blank line and include an abstract (i.e. a summary of the paper) of no more than 1,600 characters (including spaces). It is advisable to make the abstract informative, accurate, non-evaluative, and coherent, as most researchers read the abstract either in their search for the main result or as a basis for deciding whether or not to read the paper itself. The abstract should be self-contained, i.e. bibliographic citations and mathematical expressions should be avoided.
- ***Key words***. After the abstract, Key words (in bold) should be followed by three to four key words or brief phrases, preferably other than used in the title of the paper**.**
- ***Sectioning***. The paper should be divided into sections, and into subsections and smaller divisions as needed. Section titles should be in bold and left-justified, and numbered with **1.**, **2.**, **3.**, etc.
- ***Figures and tables***. In general, use only tables or figures (charts, graphs) that are essential. Tables and figures should be included within the body of the paper, not at the end. Among other things, this style dictates that the title for a table is placed above the table, while the title for a figure is placed below the graph or chart. If you do use tables, charts or graphs, choose a format that is economical in space. If needed, modify charts and graphs so that they use colours and patterns that are contrasting or distinct enough to be discernible in shades of grey when printed without colour.
- ***References.*** Each listed reference item should be cited in the text, and each text citation should be listed in the References**.** Referencing should be formatted after the Harvard Chicago System – see http://www.libweb.anglia.ac.uk/referencing/harvard.htm. When creating the list of bibliographic items, list all items in alphabetical order. References in the text should be cited with authors' name and the year of publication. If part of a reference is cited, indicate this after the reference, e.g. (Novak, 2003, p.125).