ISSN 2579-2784 (Print)
ISSN 2538-2788 (Online)

MATHEMATICAL PROBLEMS OF COMPUTER SCIENCE

LXI

Yerevan
2024

Հայաստանի Հանրապետության Գիտությունների ազգային ակադեմիայի
Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտ

Институт проблем информатики и автоматизации Национальной академии наук Республики Армения

Institute for Informatics and Automation Problems of the National Academy of Sciences of the Republic of Armenia

Կոմպյուտերային գիտության մաթեմատիկական խնդիրներ

Математические проблемы компьютерных наук

Mathematical Problems of Computer Science

LXI

ՀՐԱՏԱՐԱԿՎԱԾ Է ՀՀ ԳԱԱ ԻՆՖՈՐՄԱՏԻԿԱՅԻ ԵՎ ԱՎՏՈՄԱՏԱՑՄԱՆ ՊՐՈԲԼԵՄՆԵՐԻ ԻՆՍՏԻՏՈՒՏԻ ԿՈՂՄԻՑ
ОПУБЛИКОВАНО ИНСТИТУТОМ ПРОБЛЕМ ИНФОРМАТИКИ И АВТОМАТИЗАЦИИ НАН РА
PUBLISHED BY THE INSTITUTE FOR INFORMATICS AND AUTOMATION PROBLEMS OF NAS RA

ԵՐԵՎԱՆ 2024 YEREVAN

Կոմպյուտերային գիտության մաթեմատիկական խնդիրներ պարբերականը հրատարակվում է տարեկան երկու անգամ ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտի (ԻԱՊԻ) կողմից: Այն ընդգրկում է տեսական և կիրառական մաթեմատիկայի, ինֆորմատիկայի և հաշվողական տեխնիկայի ժամանակակից ուղղությունները:

Այն ընդգրկված է Բարձրագույն որակավորման հանձնաժողովի ընդունելի ամսագրերի ցանկում:

Տպագրվում է Խմբագրական խորհրդի 2024թ.մայիսի 27-ի N 27-05/1 նիստի որոշման հիման վրա

ԽՄԲԱԳՐԱԿԱՆ ԽՈՐՀՈՒՐԴ

Գլխավոր խմբագիր

Յու. Շուքուրյան *Գիտությունների ազգային ակադեմիա, Հայաստան*

Գլխավոր խմբագրի տեղակալ

Մ. Հարությունյան *ՀՀ ԳԱԱ ԻԱՊԻ, Հայաստան*

Խմբագրական խորհրդի անդամներ

Ս. Աղայան *Նյու Յորքի քաղաքային համալսարան, ԱՄՆ*
Հ. Ավետիսյան *ՌԳԱ Համակարգային ծրագրավորման ինստիտուտ, Ռուսաստան*
Լ. Ասլանյան *ՀՀ ԳԱԱ ԻԱՊԻ, Հայաստան*
Հ. Ասցատրյան *ՀՀ ԳԱԱ ԻԱՊԻ, Հայաստան*
Մ. Դայդե *Թուլուզի համակարգչային գիտությունների հետազոտական համալսարան, Ֆրանսիա*
Ա. Դեգտյարյով *Սանկտ Պետերբուրգի պետական համալսարան, Ռուսաստան*
Ե. Զորյան *Սինոփսիս, Կանադա*
Յու. Հակոբյան *Երևանի պետական համալսարան, Հայաստան*
Գ. Մարգարով *Հայաստանի ազգային պոլիտեխնիկական համալսարան, Հայաստան*
Հ. Մելաձե *Վրաստանի տեխնիկական համալսարան, Վրաստան*
Հ. Շահումյան *Դուբլինի համալսարանական քոլեջ, Իռլանդիա*
Ս. Շուքուրյան *Երևանի պետական համալսարան, Հայաստան*
Է. Պողոսյան *ՀՀ ԳԱԱ ԻԱՊԻ, Հայաստան*
Վ. Սահակյան *ՀՀ ԳԱԱ ԻԱՊԻ, Հայաստան*

Պատասխանատու քարտուղար

Փ. Հակոբյան *ՀՀ ԳԱԱ ԻԱՊԻ, Հայաստան*

ISSN 2579-2784 (Print)
ISSN 2738-2788 (Online)

© Հրատարակված է ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտի կողմից, 2024

Журнал **Математические проблемы компьютерных наук** издается два раза в год Институтом проблем информатики и автоматизации НАН РА. Он охватывает современные направления теоретической и прикладной математики, информатики и вычислительной техники.

Он включен в список допустимых журналов Высшей квалификационной комиссии.

Печатается на основании решения N 27-05/1 заседания
Редакционного совета от 27 мая 2024г.

РЕДАКЦИОННЫЙ СОВЕТ

Главный редактор

Ю. Шукурян Национальная академия наук, Армения

Зам. главного редактора

М. Арутюнян Институт проблем информатики и автоматизации, Армения

Члены редакционного совета

А. Аветисян Институт системного программирования РАН, Россия
С. Агаян Городской университет Нью-Йорка, США
Л. Асланян Институт проблем информатики и автоматизации, Армения
Г. Асцатрян Институт проблем информатики и автоматизации, Армения
Ю. Акопян Ереванский государственный университет, Армения
М. Дайде Тулузский научно-исследовательский институт компьютерных наук, Франция
А. Дегтярев Санкт-Петербургский государственный университет, Россия
Е. Зорян Синопсис, Канада
Г. Маргаров Национальный политехнический университет Армении, Армения
Г. Меладзе Грузинский технический университет, Грузия
Э. Погосян Институт проблем информатики и автоматизации, Армения
В. Саакян Институт проблем информатики и автоматизации, Армения
А. Саруханян Институт проблем информатики и автоматизации, Армения
А. Шаумян Дублинский университетский колледж, Ирландия
С. Шукурян Ереванский государственный университет, Армения

Ответственный секретарь

П. Акопян Институт проблем информатики и автоматизации, Армения

ISSN 2579-2784 (Print)
ISSN 2738-2788 (Online)
© Опубликовано Институтом проблем информатики и автоматизации НАН РА, 2024

Mathematical Problems of Computer Science, LXI

The periodical **Mathematical Problems of Computer Science** is published twice per year by the Institute for Informatics and Automation Problems of NAS RA. It covers modern directions of theoretical and applied mathematics, informatics and computer science.

It is included in the list of acceptable journals of the Higher Qualification Committee.

Printed on the basis of decision N 27-05/1 of the session of the Editorial Council dated May 27, 2024.

EDITORIAL COUNCIL

Editor–in–Chief

Yu. Shoukourian — National Academy of Sciences, Armenia

Deputy Editor

M. Haroutunian — Institute for Informatics and Automation Problems, Armenia

Members of the Editorial Council

S. Agaian — City University of New York, USA
A. Avetisyan — Institute for System Programming of the RAS, Russia
L. Aslanyan — Institute for Informatics and Automation Problems, Armenia
H. Astsatryan — Institute for Informatics and Automation Problems, Armenia
M. Dayde — Institute for Research in Computer Science from Toulouse, France
A. Degtyarev — St. Petersburg University, Russia
Yu. Hakopian — Yerevan State University, Armenia
G. Margarov — National Polytechnic University of Armenia, Armenia
H. Meladze — Georgian Technical University, Georgia
E. Pogossian — Institute for Informatics and Automation Problems, Armenia
V. Sahakyan — Institute for Informatics and Automation Problems, Armenia
A. Shahumyan — University College Dublin, Ireland
S. Shoukourian — Yerevan State University, Armenia
E. Zoryan — Synopsys, Canada

Responsible Secretary

P. Hakobyan — Institute for Informatics and Automation Problems, Armenia

ISSN 2579-2784 (Print)
ISSN 2738-2788 (Online)
© Published by the Institute for Informatics and Automation Problems of NAS RA, 2024

CONTENTS

Mathematical Problems of Computer Science 61, 7–14, 2024.
doi: 10.51408/1963-0111

UDC 004.932.2, 519.72

Analyzing the Quality of Distorted Images by the Normalized Mutual Information Measure

Mariam E. Haroutunian [1], David G. Asatryan[1,2] and Karen A. Mastoyan[1]

[1]Institute for Informatics and Automation Problems of NAS RA, Yerevan, Armenia
[2]Russian-Armenian University, Yerevan, Armenia
e-mail: armar@sci.am, dasat@iiap.sci.am, kmastoyan@gmail.com

Abstract

This research explores how different types of distorting algorithms impact the Full-Reference image quality assessment, particularly when subjective quality evaluations are incorporated. We draw upon the TID2013 database, which contains 3000 images distorted by 24 distinct algorithms, in conjunction with Mean Opinion Scores (MOS) for quality ratings. We compare the results of Normalized Mutual Information (NMI) for image quality score with W^2, based on Weibull distribution, the common PSNR similarity measure and MOS. We advocate for integrating of NMI into the repertoire of image quality assessment metrics.
Keywords: Image quality, Distortion types, Evaluation metrics, Normalized mutual information.
Article info: Received 25 March 2024; sent for review 26 March 2024; accepted 16 April 2024.

1. Introduction

Assessing the quality of images is a crucial process for various applications, including pattern recognition, classification, restoration, and others. The definition of quality, however, lacks an unambiguous formal consensus, leading to the requirement of specific interpretations of image quality and respective methods for its estimation. Three key methodologies exist for evaluating image quality. *Full-Reference* methods are based on the comparison between a distortion-free reference image and a test image, which is a distorted version of the original. The level of distortion serves as an indicator of the quality of the test image. The change in quality may either indicate a decrease or an increase, depending on the result of the distortion process [1, 2].

On the other hand, *No-Reference* methods assess the quality solely based on the analysis of the test image, taking into account its structural characteristics and other properties. *Reduced-Reference* methods fall in the middle, employing partial information about the original image in the assessment process.

There is abundant research literature focusing on these three types of image quality evaluation methods. These techniques can be bifurcated into two classes: *objective* and *subjective*. Objective methods utilize formal image theory and image processing techniques, whereas subjective methods rely on human visual system (HVS), based on expert quality assessments. The average of these subjective assessment results is the MOS [3, 4].

Several primary factors influence the assessment of image quality. The first one is the inherent quality of the test image and its depiction. The second are the distortions introduced into the image content through processes like image acquisition, visualization, transmission, etc. The third are the changes made in the image structure and parameters during image processing using various mathematical or computational methods.

Given the diversity of these factors, universal methods for assessing the quality of any image do not exist. Decisions on quality must consider the unique properties of the tested images and the employed methods, in combination with the available subjective assessments. Thus, there is continuous need for developing new quality criteria, similarity assessment methods, and methods for analyzing and comparing different approaches.

Several image databases with MOS assessments exist that have been collected through experimental procedures involving a large number of experts. For instance, 40 such databases are critically examined in [4]. The literature provides extensive references to studies on quality assessment through both objective and subjective methods [5]. In [6], the regularities of influence of the type of distorting algorithm on the result of evaluating the image quality by the Full-Reference method in the presence of subjective quality assessments were studied. As an example, the TID2013 database [7] with 3000 images distorted by 24 types of algorithms and subjective MOS quality ratings was used. An image quality score based on the Weibull distribution model and the usual Peak Signal-to-Noise Ratio (PSNR) similarity measure was applied. It was shown that the applied distorting algorithms are classified into two types - normal, leading to results consistent with the HVS, and "anomalous", the corresponding quality estimates of which are disordered or chaotic.

In this research, we investigate another approach to image quality assessment using the concept of NMI, which was introduced and studied in [8]. Its theoretical grounding in information theory [9] provides a robust and well-defined basis for measuring image similarity. Additionally, NMI's scale invariance makes it versatile and applicable to images of diverse resolutions. Furthermore, its non-parametric nature eliminates the need for prior assumptions about the image data, enhancing its adaptability to various image types. NMI quantifies the amount of information shared between the reference and the distorted images. This metric has shown potential between the reference and the distorted images [10].

As research on NMI and its applications in image quality assessment continues to evolve, exciting possibilities emerge. The development of deep learning-based NMI variants, for instance, holds promises for further enhancing accuracy and robustness in complex scenarios.

Within the scope of our investigation, we aim to rigorously examine and evaluate the performance of the NMI metric across a diverse array of image distortion types and levels. Our endeavor is directed towards discerning NMI's nuanced impact and effectiveness in capturing the similarities between datasets that have undergone different manifestations and intensities of image distortion. The research article employs a structured approach, encompassing research methodology and experimental results.

The paper is organized as follows. The next section introduces the considered measures. In Section 3 experimental results on the TID2013 database are discussed. The paper concludes in Section 4, summarizing key findings and advocating for a balanced consideration

of both NMI and subjective evaluation methods.

2. Description of Considered Measures

First, we consider **MOS**, which is a subjective measure that represents the average opinion of human observers. It is useful for the evaluation of other measures.

- It quantifies the perceived quality of an image based on human evaluation.
- MOS scores are typically obtained through subjective experiments, where human observers rate the quality of images on a scale.
- Higher MOS scores indicate better perceived image quality.
- MOS is commonly used as a benchmark for objective image quality assessment algorithms.

PSNR is a widely used metric in image quality assessment, commonly applied in image processing and compression. It quantifies the fidelity of an image by comparing the maximum signal power (original image) to the noise power (introduced during representation, often as Gaussian noise). The key points are:

- (Objective Measure) PSNR provides an objective numerical assessment of image quality, enabling quantitative comparisons between different algorithms.
- (Decibel Scale) The use of decibels ensures a perceptually relevant representation of quality ratios.
- (Higher Values, Better Quality) Higher PSNR values signify better image quality with minimal noise interference [2, 13, 14].

While PSNR offers simplicity and objectivity in evaluating signal quality, it has limitations in accurately reflecting human perception. It may not be suitable for all types of signals or compression techniques. It is essential to consider its advantages and disadvantages carefully when using it for quality assessment in image and video processing applications.

$\mathbf{W^2}$ is an image quality metric. It measures the structural similarity between the original image and the image with additive Gaussian noise. W^2 values range from 0 to 1, where 1 indicates perfect structural similarity. Higher W^2 values suggest better image quality and preservation of structural information. W^2 is commonly used for image restoration and enhancement [15]. This image quality estimation is based on a Weibull distribution model of image gradient magnitude, the density of which is given by the formula

$$f(x;\lambda,\eta) = \frac{\eta}{\lambda}\left(\frac{x}{\lambda}\right)^{\eta-1}\exp\left[-\left(\frac{x}{\lambda}\right)^{\eta}\right], x \geq 0,$$

where $\eta > 0$ is the shape parameter, $\lambda > 0$ is the scale parameter. Distribution parameters are estimated from the totality of all gradient magnitudes using the Sobel operator. The similarity (proximity) of two images is estimated by the proximity of the estimates of the parameters of the Weibull distribution by the formula

$$W^2 = \frac{min\,(\eta_1, \eta_2)\, min\,(\lambda_1, \lambda_2)}{max\,(\eta_1, \eta_2)\, max\,(\lambda_1, \lambda_2)}.$$

The research in [6] presented that this measure is sensitive to those types of distortions that affect the structure and content of the image.

NMI is a measure of the distance between two images based on their joint probability distribution. It quantifies the amount of information shared between two images, considering both their individual distributions and their joint distribution. Higher NMI values indicate greater distance, whereas lower NMI values indicate less distance.

Mutual information is a fundamental concept in information theory. Given two random variables X and Y, the mutual information (MI) is defined as

$$I(X;Y) = H(X) \ + H(Y) - H(X,Y),$$

where $H(X)$ is the well-known notion of entropy [9]. MI is a non-negative quantity and can be used as a similarity or distance measure depending on various applications.

We consider the following **normalized** version of MI

$$NMI = 1 - \frac{I(X;Y)}{\max H\,(X)\,, H(Y)},$$

which is a **distance measure**. It was proved in [11] that this measure satisfies metric properties, in other words, it adheres to the criteria of a true metric, encompassing positive definiteness, symmetry, and triangle inequality. At its core, the metric property aligns with our intuitive understanding of distance, providing a foundational framework for quantifying spatial relationships. NMI values range from 0 to 1, with 0 indicating perfect similarity and 1 indicating no similarity at all. Beside from information theory, NMI is widely used also in image registration, image segmentation, and other applications [10], [12]. NMI is often used to evaluate clustering algorithms or comparing different clusterings of the same data [11]. NMI is based on the principles of information theory, which makes it theoretically grounded and well-suited for various applications in fields such as machine learning, pattern recognition, and data mining.

3. Experimental Results

The selected database is TID2013 [7]. This database contains 3000 images obtained from 25 originals, distorted by 24 different types of five levels each (for example, see Fig. 1). The authors of the database conducted an extensive experiment on the visual assessment of the quality of database images using a point system by a large number of people from different countries. As a result of processing these data, each of the 3000 images is assigned a numerical MOS score.

All necessary quantities are calculated using the developed software system, and the results are entered into Excel tables. The base data are the results related to the original and five distorted samples of a particular image.

For each such data set, three evaluation methods were employed: NMI, PSNR, W^2, and compared with MOS. These methods were utilized to assess and analyze the quality of the images in the database.

Fig. 1. The first, third, and fifth levels of distorted images from Image Number 10 in the TID 2013 database using the 'Sparse Sampling and Reconstruction' method.

In numerous scenarios, NMI demonstrated similar sensitivity in comparison with alternative measures, for example on Fig. 2 five levels of Additive noise in color distortion are demonstrated. The values of all measures for each level are given in Table 1.

Fig. 2. 5 Levels of Additive Noise in Color Distortion from the TID2013 Database: Image Number 15.

Table 1. Experimental results for the 15th image from TID2013 database (Additive noise in color)

NMI	PSNR	W^2	MOS
0.14	42.33	0.87	6.09
0.17	39.45	0.78	5.82
0.22	36.47	0.66	5.64
0.27	33.61	0.53	4.89
0.34	31.39	0.38	4.64

In some cases, NMI demonstrates higher efficiency. For example, in the case of distortion with the Non eccentricity pattern noise method, the W^2 values are close to one, which means that it performs poorly in terms of human evaluation and human understanding, and the NMI values are close to human evaluation (Table 2, Fig. 3).

Fig. 3. Reference image and level 5 distorted image (Non eccentricity pattern noise)

Table 2. Experiment results for the 8th image from TID2013 database (Non-eccentricity pattern noise).

NMI	PSNR	W^2	MOS
0.06	43.33	1	5.65
0.10	41.30	1	5.43
0.16	39.08	0.99	4.87
0.20	37.82	0.99	4.75
0.24	36.92	0.99	4

4. Conclusion

Our experimental results revealed interesting insights into the performance of metrics across different types of distortions. We found that NMI, being a normalized distance measure, showed promising results for various distortion types. Particularly, it exhibited close alignment with human subjective evaluations in almost all cases, indicating its potential as an effective image quality assessment metric.

Moreover, NMI's theoretical foundation in information theory and its versatility in capturing differences between images of diverse resolutions contribute to its robustness and applicability in image quality assessment tasks. NMI consistently demonstrated its efficacy for a wide range of distortions.

In conclusion, our findings advocate for integrating NMI into the repertoire of image quality assessment metrics, complementing traditional measures like PSNR and W^2. By leveraging NMI's inherent advantages and considering its performance in conjunction with subjective evaluations, we can enhance the accuracy and reliability of image quality assessment methods, catering to diverse application scenarios in image processing, computer vision, and beyond.

References

[1] A. George and S. J. Livingston, "A survey on full reference image quality assessment algorithms", *International Journal of Research in Engineering and Technology*, vol.2, no.12, pp. 303-307, 2013.

[2] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity", *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.

[3] T.-J. Liu, Y.-C. Lin, W. Lin and C.-C.J. Kuo, "Visual Quality Assessment: Recent Developments, Coding Applications and Future Trends", *APSIPA Trans. on Signal and Information Processing*, vol. 2, 2013. DOI: https://doi.org/10.1017/ATSIP.2013.5

[4] G. Zhai and X. Min, "Perceptual image quality assessment: a survey", *Science China Information Sciences*, vol. 63, no. 11, 211301, 2020. DOI: 10.1007/s11432-019-2757-1

[5] J. Wang, Z. Wang, L. Lu and A. C. Bovik, "Subjective Quality Assessment of Stereoscopic Omnidirectional Image", *Lecture Notes in Computer Science, in book Advances in Multimedia Information Processing PCM*, pp. 589-599, 2018.

[6] D. Asatryan, M. Haroutunian, G. Sazhumyan and G. Hakobyan, "Procedure for analyzing the quality, structure and subjective rating of distorted images by the Full-Reference technique", *Intern. Scientific Journals of Scientific Technical Union of Mechanical Engineering "Industry 4.0", Mathematical Modeling*, vol. 6, no. 4, pp. 100-102, 2022.

[7] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, Battisti F., et al. "Image database tid2013", *Signal Processing: Image Communication*, vol. 30, pp. 57-77, 2015.

[8] A. Kraskov, H. Stogbauer, R. G. Andrzejak and P. Grassberger, "Estimating mutual information", *Physical Review*, E 69, 066138, 2004.

[9] T. M. Cover and J. A Thomas, *Elements of Information Theory*, Second Edition, Wiley-Interscience, 2006.

[10] F. E. Ruiz, P. S. Pérez and B. I. Bonev, "Information theory in computer vision and pattern recognition", *Springer Science and Business Media*, 2009.

[11] N. X. Vinh, J. Epps and J. Bailey, "Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance", *Journal of Machine Learning Research*, vol. 11, pp. 2837-2854, 2010.

[12] J. P. W. Pluim, J. B. A. Maintz and M. A. Viergever, "Mutual-information-based registration of medical images: A survey", *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986-1004, Aug. 2003.

[13] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, 3rd edition, Pearson Education, 2008.

[14] Z. Wang, E. P. Simoncelli and A. C. Bovik, "Multiscale structural similarity for image quality assessment", *In The Thrity-Seventh Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1398-1402, 2003.

[15] D. Asatryan, "Image blur estimation using gradient field analysis", *Computer Optics*, vol. 41, no. 6, pp. 957-962, 2017. DOI: 10.18287/2412-6179-2017-41-6-957-962.

Աղավաղված պատկերների որակի վերլուծություն նորմալացված փոխադարձ ինֆորմացիայի միջոցով

Մարիամ Ե. Հարությունյան[1], Դավիթ Գ. Ասատրյան[1,2] և Կարեն Ա. Մաստոյան[1]

[1]ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտ, Երևան, Հայաստան
[2] Հայ-Ռուսական համալսարան, Երևան, Հայաստան
e-mail: armar@sci.am, dasat@iiap.sci.am, kmastoyan@gmail.co

Ամփոփում

Այս հետազոտության մեջ մենք ստուգանմուշի օգտագործմամբ ուսումնասիրում ենք, թե ինչպես են տարբեր տեսակի աղավաղող ալգորիթմները ազդում պատկերի որակի ամբողջական գնահատման վրա, մասնավորապես, երբ ներառված են սուբյեկտիվ որակի գնահատականները: Մենք օգտվում ենք Tid2013 տվյալների բազայից, որը պարունակում է3000 պատկեր, որոնք աղավաղված են 24 տարբեր ալգորիթմներով, կարծիքների միջին միավորների (MOS) հետ համատեղ: Մենք համեմատում ենք նորմալացված փոխադարձ ինֆորմացիայի արդյունքները (NMI) պատկերի որակի գնահատման համար Վեյբուլի բաշխման վրա հիմնված W^2 արդյունքների, հայտնի PSNR նմանության չափի և MOS-ի հետ: Մենք պնդում ենք NMI- ի ինտեգրումը պատկերի որակի գնահատման չափման մեծությունների ցանկում:

Բանալի բառեր՝ Պատկերի որակ, աղավաղման տեսակներ, գնահատման չափումներ, նորմալացված փոխադարձ ինֆորմացիա:

Анализ качества искаженных изображений с помощью нормализованной меры взаимной информации

Мариам Е. Арытюнян[1], Давид Г. Асатрян[1,2] и Карен А. Мастоян[1]

[1]Институт проблем информатики и автоматизации НАН РА, Ереван, Армения
[2]Российско-Армянский университет, Ереван, Армени
e-mail: armar@sci.am, dasat@iiap.sci.am, kmastoyan@gmail.com

Аннотация

В этой статье методом сравнения с эталоном мы исследуем, как различные типы алгоритмов искажения влияют на оценку качества изображения с полной ссылкой, особенно при включении субъективных оценок качества. Мы опираемся на базу данных TID2013, которая содержит 3000 изображений, искаженных 24 различными алгоритмами, в сочетании со средними оценками мнений (MOS) для рейтингов качества. Мы сравниваем результаты нормализованной взаимной информации (NMI) для оценки качества изображения с W^2, на основе распределения Вейбулла, общего показателя сходства PSNR и MOS. Мы выступаем за интеграцию NMI в репертуар показателей оценки качества изображения.

Ключевые слова: Качество изображения, типы искажений, метрики оценки, нормализованная взаимная информация.

Mathematical Problems of Computer Science 61, 15–26, 2024.
doi: 10.51408/1963-0112

UDC 519.24, 539.1, 004.4

Pythia8 MC Tuning Validation Using the Professor2 Package

Hazaravard M. Ghumaryan

A.I. Alikhanyan National Science Laboratory, Yerevan, Armenia
e-mail: hazar@yerphi.am

Abstract

This article presents a method for multi-parameter, simultaneous tuning of the Monte Carlo event generator. It is validated on the Pythia8 Monte Carlo event generator widely used in High Energy Physics (HEP). The obtained results show that the method can be used to constrain the free parameters of phenomenological models while properly taking into account the correlations existing among the parameters.
Keywords: Monte Carlo tuning, hadronization, Pythia8, Belle II experiment, Professor2 package.
Article info: Received 3 March 2024; sent for review 19 March 2024; accepted 21 May 2024.

1. Introduction

Historically, to constrain free parameters in multi-parameter models, a tuning procedure is applied based on the change-check method i.e., parameters are changed "one-by-one" and the impact of a particular parameter on the physics observable of interest is studied. Obviously, for any complex system with too many free parameters, this method becomes non-optimal in a practical sense as it also suffers not taking into account the correlations existing between parameters. The state-of-art method is to parameterize the generator behaviour in a simultaneous change of multi-parameter set by fitting the generator response with a polynomial for each physics observable entering the tuning list:

$$MC_b(p) \approx f^b(p) = \alpha_0^b + \sum_i \beta_i^{(b)} p_i' + \sum_{i \le j} \gamma_{ij}^{(b)} p_i' p_j'. \tag{1}$$

In formula (1), MC_b is the true response of the MC generator in a bin b, $f^b(p)$ is a set of functions that model the true MC response for each observable bin b when changing the parameter vector $\mathbf{p}$ and $\mathbf{p}' = \mathbf{p}$ - $\mathbf{p}_0$ is the parameter vector shifted from its nominal value $\mathbf{p}_0$. To get the optimal, i.e., "tuned" values of the parameters, the response function is minimized concerning the reference sample by performing the χ^2 minimization as shown in

formula (2):

$$\chi^2(p) = \sum_O \omega_O \sum_{b \in O} \frac{(f^b(p) - R_b)^2}{\Delta_b^2}, \tag{2}$$

where R_b is the reference value for bin b, Δ_b is the total uncertainty of the bin b and ω_O is the statistical weight for each observable.

Recent experimental results have inspired more modeling work in the theory community. Simultaneously, efforts to refine the existing models and parameters are ongoing. Although Pythia [1] has been extensively compared to LHC, Belle, and Belle II data, its constraints from e^+e^- colliders haven't been updated since 2009, relying on an undocumented tuning effort with the Professor [2] tool. In this work, we address this issue using the Pythia8 Monte Carlo event generator from the Belle II analysis framework.

2. Toolkits for Tuning

We created a tuning framework to model the continuum data [3] produced by the Pythia8 Monte Carlo event generator. This involves the utilization of a dedicated computing platform at AANL (belle2.yerphi.am) connected to the Worldwide GRID computing system [4]. To stay updated with the latest Belle II software releases, we established access to the CernVM File System (CVMFS) [5] repository on our local server. Being a member of the Belle II international collaboration, we have full access to KEK [6] and DESY [7] computing platforms (PC Farms), enabling the simultaneous distribution of jobs across different systems.

We have designed an automated framework capable of processing different parameter sets as input and deploying corresponding jobs to the GRID. Likewise, samples generated from the Worldwide computing system are collected and organized into folders corresponding to the simulation settings. On the local "belle2.yerphi.am" machine, the Belle II Analysis Software Framework (basf2) [8] environment has been configured for the author of this article. A set of jobs is then sent to the Worldwide GRID computing platform under the periodic control and monitoring of the DIRAC [9] project. Access to this platform is facilitated through the use of the GRID certificate.

3. Tuning Procedure

To address the challenges associated with multi-parameter optimization problems, a specialized package, i.e., Professor2 has been developed as an alternative to manual adjustment methods. The approach used by the Professor2 package is known as parametrization-based tuning. A notable advantage of this package is its capacity to handle correlations among different parameters, enabling multivariate minimization within the parameter space.

The primary objective of this method is to determine the correspondence function or characteristic polynomial between the generated Monte Carlo data and the reference data, ultimately minimizing the multi-dimensional χ^2 function. In our study, we utilize the Pythia8 model for generating physics events. We present a comparison and tuning of selected observables extracted from reference data and the Belle II off-resonance Monte Carlo event generator.

One of the important aspects of the tuning procedure is to choose the free parameters from

the models to which the obsevables of interest are sensitive. This is done by performing sensitivity checks explained in Section 3.2. Having the list of sensitive parameters the tuning procedure is performed by using the Professor2 toolkit. It requires samples to be generated with different sets of sensitive parameters. The Professor2 framework provides four types of parameter sampling: grid, uniform, sobol and latin hypercube sampling. The "sobol" and "latin" hypercube distributions aim at covering the space more evenly for low sample sizes. In this work, we used "uniform" sampling with $N = 1500$ parameter points as shown below:

prof2-sample -o output params.dat -n 1500

which creates 1500 sets of parameters within predefined intervals for each. These sets are used to generate Monte Carlo samples which are later used to extract the observables of interest and the model parameters to be tuned for. The observables of interest used in this work are described in detail in Section 2.1. The generated data sets were saved in ROOT file format, which was later analyzed to extract the distributions of interest in the binned (histogram) format. It is important to mention that a special care is taken to ensure that the histograms did not contain any empty bins, otherwise, the empty bins are encountered by setting the corresponding weights to be equal to zero.

To model the MC response, a characteristic polynomial from the generated Monte Carlo (MC) samples is constructed as shown below:

prof2-ipol runsdir ipolfile=ipol.dat –order=5

To consider the multiple correlations of parameters, the order of the polynomial can be changed accordingly. In this work, the order of polynomial is set to 5. Consequently, the output of the interpolation is the file "ipol.dat" , which contains the interpolation results essential for the tuning procedure, such as the MC response extracted for each observable in terms of bin content variation.

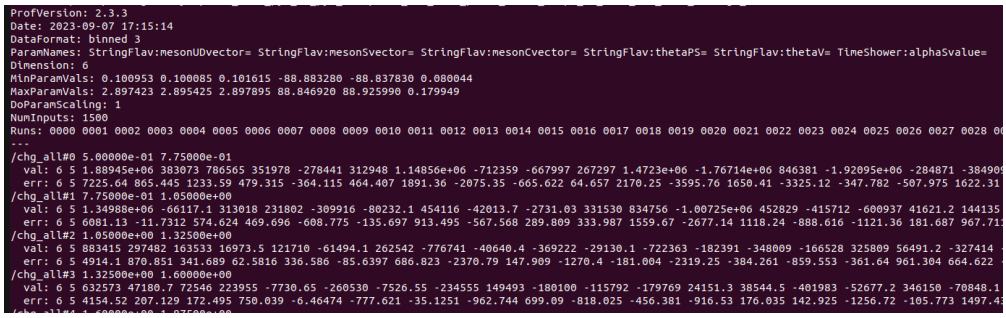
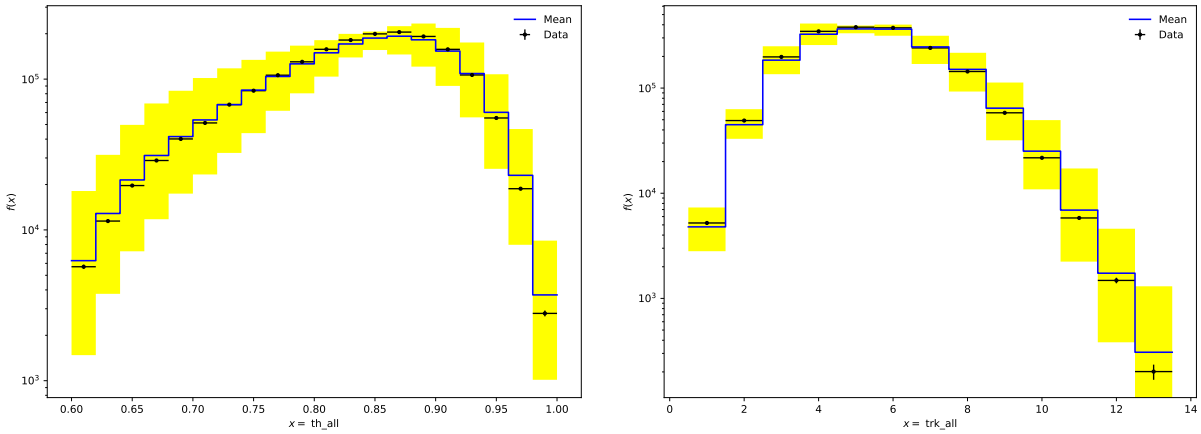


Fig. 1. "ipol.dat" file.

To ensure the full coverage of Monte Carlo samples generated with different parameter sets on the reference data, the "envelope" plots are used that effectively visualize the access of the generated data on the full spectra of the reference data:

prof2-envelopes mc -d refdir

The final step of the tuning procedure is performed using the prof2-tune tool. It minimizes the MC response function given in the form of the characteristic polynomial and accumulated in the "ipol.dat" file by comparing it with the reference data for each distribution obtained



(a) Envelope plot of Thrust

(b) Envelope plot of number of tracks in event

Fig. 2. The yellow band corresponds to the coverage of the generated MC samples by using different Pythia8 parameter settings and black points are from the reference sample.

from various MC runs with different parameter sets. These run combinations can be generated uniquely and randomly at runtime by prof2-tune, or they can be provided through a plain text file.

prof2-tune -d refdir ipolfiles=ipol.dat -r runsdir=refdir/../mc

3.1 Observables Used in Tuning

The variables (observables) used in the tuning procedure are tabulated in Table 1. and Table 2. The selection of a particular variable is motivated by the specifics of the tuning, namely, for what purpose the tuning of the generator is performed, thus, what model(s) entering the generator should be constrained. In this work, two-stage tuning was performed. First, three variables (see Table 1) of interest were selected. In the second stage, one more variable is added to the list (see Table 2).

Table 1. Observables used for the first stage tuning.

Thrust
Inclusive charge particle momentum spectra
Number of tracks in event

It is important to mention that adding more variables to the tuning list increases the possible correlation between the parameters in the tuning list thus making it more difficult to minimize multidimensional χ_2 for the MC response function.This study used three and four variables consequently, achieving a good agreement between the reference sample and MC simulations. In this work, the "Event" and "Event shape" variables are used to study the hadron production process later modeled in Pythia8, as well as to separate the events with different quark origins produced in the Belle II experiment. Specifically, the Thrust and

Table 2. Observables used for the second stage tuning.

Thrust
Inclusive charge particle momentum spectra
Number of tracks in event
visibleEnergyOfEventCMS

FoxWolframR2 [10] variables are crucial in distinguishing continuum events from BantiB events. At the same time, the use of event variables is essential to avoid possible issues related to Particle Identification (PID) inefficiency effects and background interference. The Thrust axis T is determined by the direction in which the sum of the longitudinal momenta of particles reaches its maximum. The thrust T is connected to the Thrust axis by

$$T = \frac{\sum |\mathbf{p_i} \cdot \mathbf{T}|}{\sum |\mathbf{p_i}|}, \tag{3}$$

where $\mathbf{p_i}$ represents the momentum of each particle.
The Fox-Wolfram moments are defined as follows:

$$H_x^l = \sum_{i,j=1}^{N} W_{ij}^x P_l(\cos\Omega_{ij}), \tag{4}$$

where W_{ij}^x is a weight factor, and $P_l(\cos\Omega_{ij})$ is the Legendre polynomial, and the FoxWolframR2 is given by the ratio

$$foxWolframR2 = \frac{H_2}{H_0}. \tag{5}$$

The inclusive charged particle momentum spectrum in high-energy particle physics refers to the distribution of momenta for all charged particles produced in a collision of beams. This observable provides insights into the overall behavior and characteristics of particle production within a given experiment.

Another important observable in the experiment is the "visibleEnergyOfEventCMS", which is defined as a sum of the energies of all particles that leave observable signals in the detector:

$$E_{\text{vis}} = \sum_{\text{visible particles}} E_i. \tag{6}$$

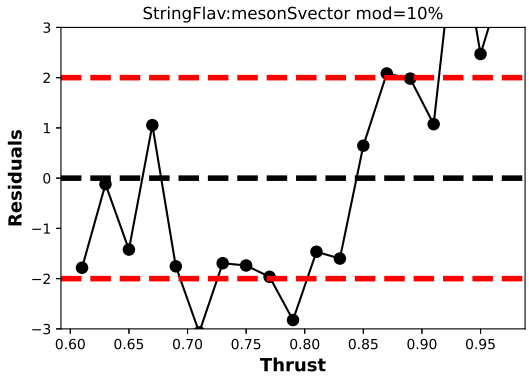
3.2 Sensitivity Checks

To reveal the sensitive parameters for the selected list of observables, the parameter sensitivity checks are perfomed using the normalized residuals extracted from two sample tests, as shown below :

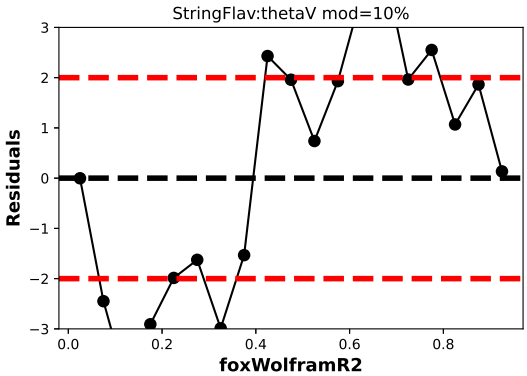
$$r_i = \frac{n_i - N\hat{p}_i}{\sqrt{N\hat{p}_i}\sqrt{(1 - N/(N+M))(1 - (n_i + m_i)/(N+M)}}. \tag{7}$$

Table 3. The list of six different parameters that are selected from the Pythia Monte-Carlo event generator for tuning the "Event" and "Event shape" variables.

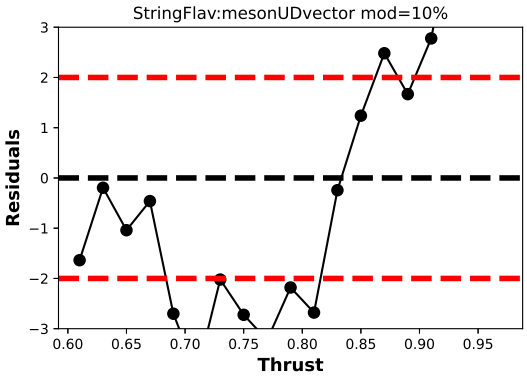
Parameters	Values
StringFlav:mesonUDvector	(default = 0.50; min = 0., max = 3.)
StringFlav:mesonSvector	(default = 0.55; min = 0., max = 3.)
StringFlav:mesonCvector	(default = 0.88; min = 0., max = 3.)
StringFlav:thetaPS	(default = -15.; min = -90., max = 90.)
StringFlav:thetaV	(default = 36.; min = -90., max = 90.)
TimeShower:alphaSvalue	(default = 36.; min = -90., max = 90.)



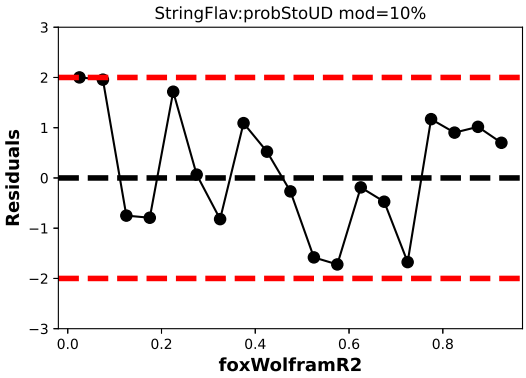
(a) The sensitivity plot of StringFlav:meson Svector Pythia parameter



(b) The sensitivity plot of StringFlav:thetaV Pythia parameter



(c) The sensitivity plot of StringFlav:meson UDvector Pythia parameter



(d) The sensitivity plot of StringFlav:probStoUD Pythia parameter

Fig. 3. The sensitivity check for parameters given in Table 2 based on Thrust and FoxWolfram distributions.

When normalized residuals exceed the 2 sigma level (see Fig. 3) for certain observables, we consider them to be sensitive to a particular parameter. Conversely, for other parameters, there is an absence of sensitivity, as indicated by residuals within the 2 sigma window. The most sensitive parameters concerning the "event" and "event shape" variables are listed in Table 3.

3.3 Validation Scheme

For the validation of the developed scheme, the Monte Carlo (MC) simulations provided by Belle II collaboration were used as reference samples.

4. Results

Our studies show that the validation of the Pythia8 MC tuning procedure using the Professor2 package is highly affected by the statistics and the number of generated samples, as well as the correlations between Pythia8 parameters. Controlling these factors is crucial for obtaining meaningful and reliable results from simulations using Pythia8. Meanwhile, by increasing the statistics and the number of generated MC samples for u,d,s,c quark combinations and using interpolation with a 5th order characteristic polynomial, we reproduce the distributions extracted from the reference sample as it can be seen in Figures 5-6. It is important to mention that due to correlations, the tuned values of parameters can differ from their default values although the spectra of observables of interest from tuned and reference samples are in a very good agreement.

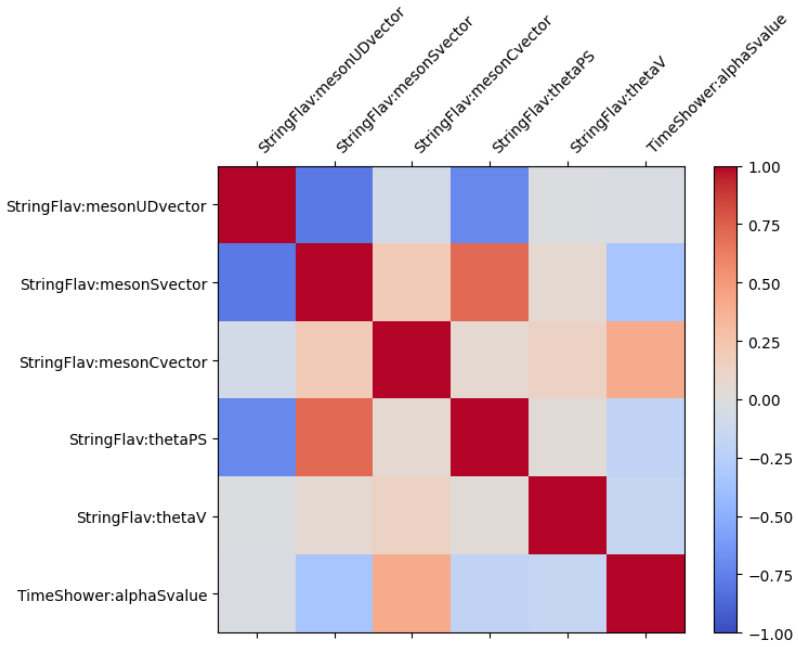


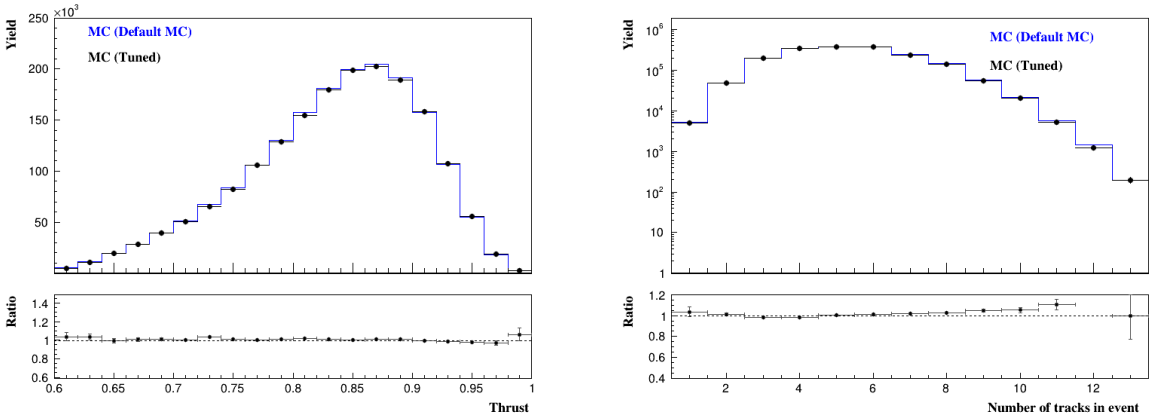
Fig. 4. Correlations values of the parameters for the first stage tune.

For the first tune we have chosen 3 variables, they are:

Table 4. Observables used for the first stage tuning.

Thrust
Inclusive charge particle momentum spectra
Number of tracks in event

Pythia8 parameters	Default	Tuned	Comment
StringFlav:mesonUDvector	=0.5	=0.444457	! Light-flavour vector suppression
StringFlav:mesonSvector	=0.55	=0.434957	! Strange vector suppression
StringFlav:mesonCvector	=2.8	=2.428458	! Charm vector suppression
StringFlav:thetaPS	=-15	=-13.500045	! Mixing angle *thetaPS*
StringFlav:thetaV	=36	=29.243306	! Mixing angle *thetaV*
TimeShower:alphaSvalue	=0.1365	=0.135795	! Effective alphaS(mZ) value



(a) (b)

Fig. 5. Comparing Thrust and Number of tracks in event distributions: Default Monte Carlo (MC) in blue vs. tuned in black.

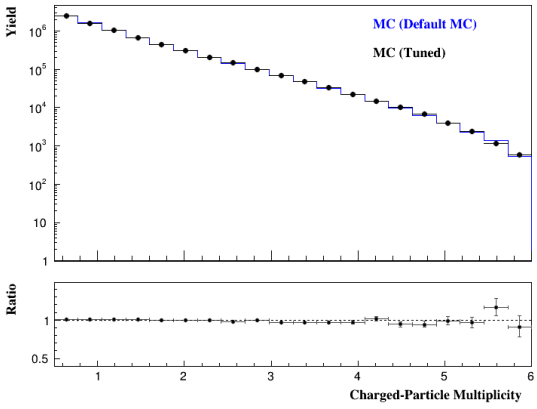


Fig. 6. Comparing Inclusive charge particle momentum distributions: Default Monte Carlo (MC) in blue vs. tuned in black.

The default and tuned values for the Pythia8 parameters.

For three variables, the parameters' default and tune values were quite similar, considering the errors.

Pythia8 parameters	MIGRAD errors
StringFlav:mesonUDvector	=4.685612e-02
StringFlav:mesonSvector	=8.267797e-02
StringFlav:mesonCvector	=1.690820e-01
StringFlav:thetaPS	=7.967455e+00
StringFlav:thetaV	=8.819869e+00
TimeShower:alphaSvalue	=4.741711e-04

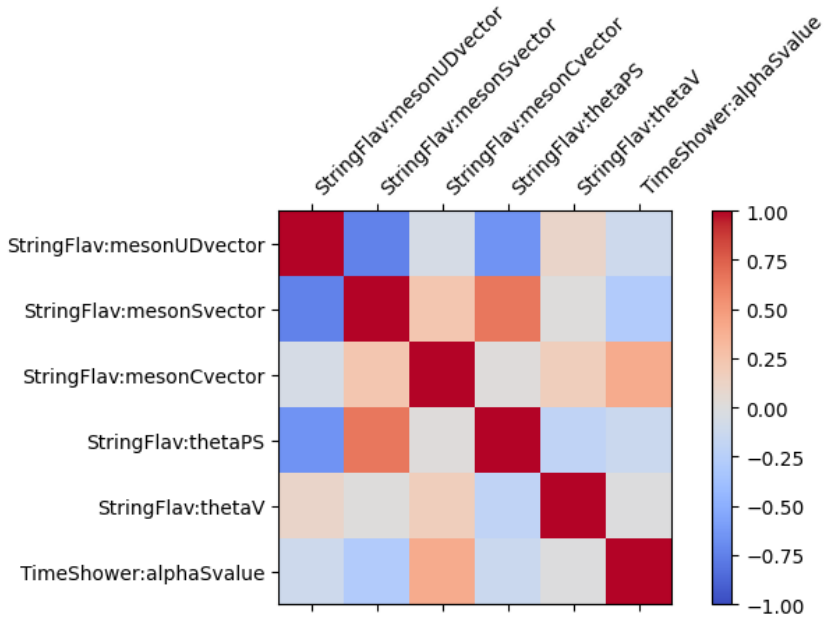


Fig. 7. Parameter correlation values for the second-stage tuning.

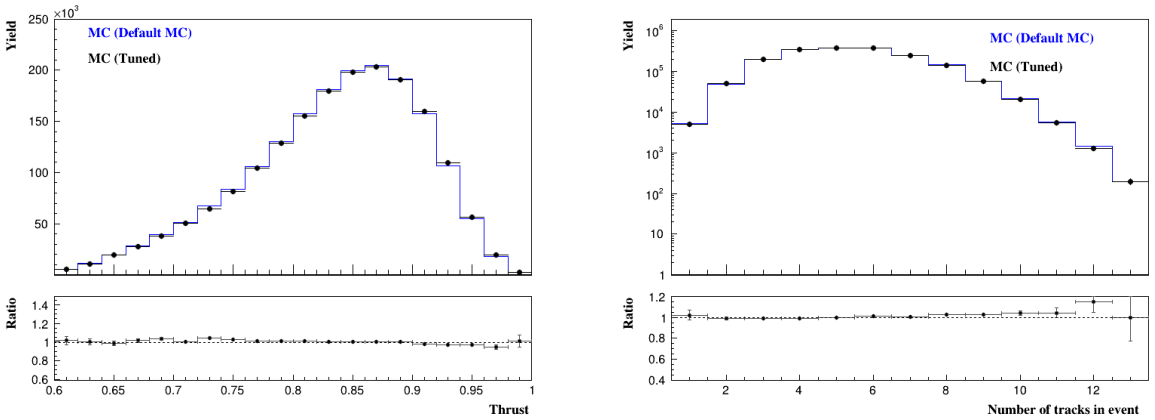
Next, we included the visibleEnergyOfEventCMS variable in our list of observables to study the impact of correlations when tuning with an increased number of variables. This tune was also performed with 5th-order polynomial to ensure the robustness for the minimization results.

As it can be seen from Fig. 8 (b), the comparison between tuned and reference samples for visibleEnergyOfEventsCMS observable is also a very good agreement.

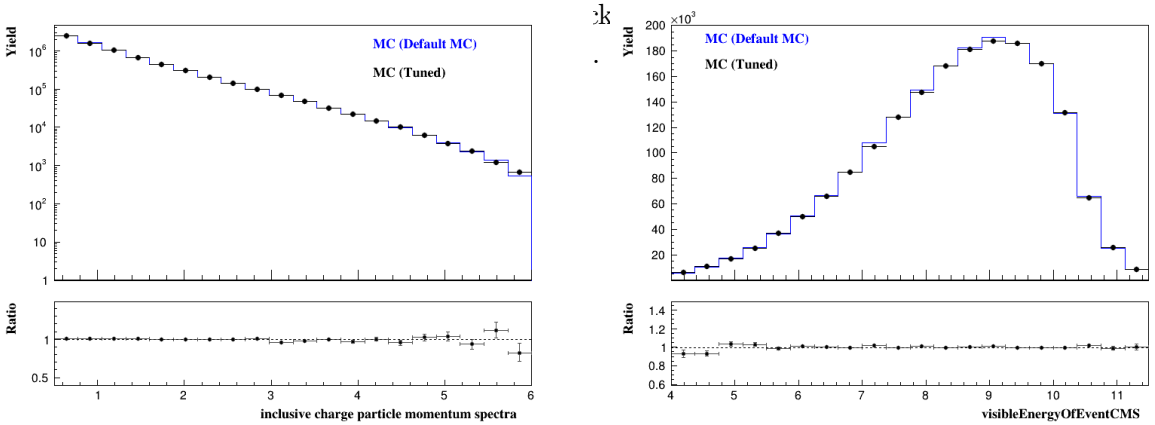
Pythia8 parameters	Default	Tuned	Comment
StringFlav:mesonUDvector	=0.5	=0.454673	! Light-flavour vector suppression
StringFlav:mesonSvector	=0.55	=0.432484	! Strange vector suppression
StringFlav:mesonCvector	=2.8	=2.388490	! Charm vector suppression
StringFlav:thetaPS	=-15	=-14.416715	! Mixing angle *thetaPS*
StringFlav:thetaV	=36	=32.127201	! Mixing angle *thetaV*
TimeShower:alphaSvalue	=0.1365	=0.135604	! Effective alphaS(mZ) value

The Default and tuned values for Pythia8 parameters. For four variables, the parameters default and tune values were quite similar, considering the errors.

Pythia8 parameters	MIGRAD errors
StringFlav:mesonUDvector	=4.270846e-02
StringFlav:mesonSvector	=7.325382e-02
StringFlav:mesonCvector	=1.705115e-01
StringFlav:thetaPS	=6.570061e+00
StringFlav:thetaV	=6.979057e+00
TimeShower:alphaSvalue	=4.257399e-04



(a) (b)



(a) (b)

Fig. 9. Comparing Inclusive charge particle momentum andvisibleEnergyOfEventCMS distributions: Default Monte Carlo (MC) in blue vs. tuned in black.

5. Conclusion

Validation of the tuning procedure developed to tune a certain set of parameters from the Pythia8 Monte Carlo event generator is performed using the Professor2 package. The tuning is done simultaneously for "event" and "event shape" variables using six parameters from Pythia8 MC. The obtained results show that the comparison between the reference and tuned spectra is in a very good agreement. The developed procedure can be extended to any model with free parameters constraining it by using real experimental data.

References

[1] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun.,2008, [Online]. Available: https://pythia.org/.

[2] Andy Buckley, Hendrik Hoeth, Heiko Lacker, Holger Schulz, and Jan Eike von Seggern, *Systematic event generator tuning for the LHC*,arXiv:0907.2973v1,2009, [Online]. Available: https://professor.hepforge.org/.

[3] Dennis Weyland, *Continuum Suppression with Deep Learning techniques for the Belle II Experiment*, 2017.

[4] [Online]. Available: https://wlcg-public.web.cern.ch/

[5] [Online]. Available: https://cernvm.cern.ch/fs/

[6] [Online]. Available: https://kekcc.kek.jp/service/kekcc/support/en/01/

[7] [Online]. Available: https://www.desy.de/

[8] [Online]. Available: https://software.belle2.org/development/sphinx/index.html

[9] [Online]. Available: https://dirac.readthedocs.io

[10] Geoffrey C . FOX and Stephen WOLFRAM,*Event Shapes in e^+e^+ annihilation*, North-Holland Publishing Company, 1978.

Pythia8 Մոնտե Կառլո գեներատորի կարգաբերում՝ Professor2 փաթեթի օգտագործմամբ

Հազարավարդ Մ. Ղումարյան

Ա. Ի. Ալիխանյանի անվան ազգային գիտական լաբորատորիա, Երևան, Հայաստան
e-mail: hazar@yerphi.am

Ամփոփում

Այս հոդվածում ներկայացված է Մոնտե Կառլո դեպքերի գեներատորի բազմապարամետրային, միաժամանակյա Թյունինգի մեթոդը: Այն հաստատվել է Pythia8 Մոնտե Կառլո իրադարձությունների գեներատորի վրա, որը լայնորեն օգտագործվում է բարձր էներգիայի ֆիզիկայում (HEP): Ստացված արդյունքները ցույց են տալիս, որ մեթոդը կարող է օգտագործվել ֆենոմենոլոգիական մոդելների ազատ պարամետրերը կարգաբերելու համար հաշվի առնելով պարամետրերի միջև առկա կոռելացիաները:

Բանալի բառեր՝ Մոնտե Կառլո թյունինգ, հադրոնիզացիա, Pythia8, Belle II գիտափորձ, professor2 փաթեթ:

Проверка настройки Pythia8 MC с использованием Professor2 пакета

Азаравард М. Гумарян

Национальная научная лаборатория имени А.И. Алиханяна, Ереван, Армения
e-mail: hazar@yerphi.am

Аннотация

В данной статье представлен метод для многопараметрической настройки генератора событий по методу Монте-Карло. Метод апробирован на настройке Монте Карло генератора Pythia8, широко используемого для физики высоких энергий (ФВЭ). Полученные результаты показывают, что метод может быть использован для определения свободных параметров феноменологических моделей, одновременно позволяя учесть корреляции существующие между параметрами.

Ключевые слова: Настройка Монте-Карло, адронизация, Pythia8, эксперимент Belle II, пакет профессор2.

Mathematical Problems of Computer Science 61, 27–49, 2024.

doi: 10.51408/1963-0113

UDC 519.237, 519.25

Comparative Analysis of Univariate SARIMA and Multivariate VAR Models for Time Series Forecasting: A Case Study of Climate Variables in Ninahvah City, Iraq

Sameera A. Othman, Hasan H. Jameel and Sadeq T. Abdulazeez

University of Duhok, College of Basic Education, Duhok, Iraq

e-mail: Sameera.Othman@uod.ac, hasan.hazim@uod.ac, sadiq.taha@uod.ac

Abstract

This study involves a comparison between the application of the univariate SARIMA model and the utilization of VAR methods (vector autoregressive models) for multivariate time series analysis. The analysis is conducted using three-time series variables derived from data representing the monthly average of Humidity (H), Rainfall (R), and Temperature (T) in Ninahvah City, Iraq. Both univariate and multivariate time series approaches are employed to model these series. The paper also outlines the implementation of vector autoregressive, structural vector autoregressive, and structural vector error correction models using the 'vars' package. Additionally, it provides functions for diagnostic testing, estimation of constrained models, prediction, causality analysis, impulse response analysis, and forecast error variance decomposition. Furthermore, it introduces three fundamental functions, VAR, SVAR, and SVEC, for estimating these models. The comparison between the methods is based on evaluating the mean error produced by each approach. The findings of the study indicate that univariate linear stationary methods outperform multivariate models. The analysis of the data was carried out using the R software platform. The primary objective is to assess the performance of univariate and multivariate time series models in handling the given data. The research gap lies in the need for a comparative evaluation of SARIMA and VAR methods for time series analysis in the context of monthly environmental variables. These models were chosen due to their effectiveness in capturing temporal dependencies and interactions among multiple variables in time series data, providing a comprehensive analysis of climatic patterns in Ninahvah City, Iraq. The study aims to address the research gap by comparing these models and justifying their selection based on their capabilities to analyze the specified time series data.

Keywords: Modeling, Doers, Energizers, Classifiers, Cognizing, Fundamentals, Dynamicity.

Article info: Received 22 October 2023; sent for review 10 November 2023; received in revised form 1 January 2024; accepted 15 January 2024.

1. Introduction

A multivariate time series (MTS) comprises numerous time-related variables, and it is crucial to understand that each variable's dependence is not solely influenced by its previous values but also by interactions with other variables. Future values are forecasted using this dependency. This dependency is used for forecasting future values. The goals of multivariate time series analysis are to investigate the complex establishing links among variables and enhancing forecast precision. In the early 1980s, the authors in [1] critiqued vector autoregressive models (VARs) led to vector autoregressive models becoming a standard instrument in econometrics. This strategy was immediately improved by the incorporation of non-statistical prior information since statistical tests are commonly utilized to identify connections and intricate associations among variables. In contrast to deterministic repressors, VAR models describe endogenous variables entirely through their own histories. Structured vector autoregressive models (SVARs) facilitate the explicit modeling of contemporaneous interdependencies between the variables on the left. Consequently, these models attempt to address the deficiencies associated with VAR models. Sims posed a challenge to the established multiple structural equation model paradigm initially developed by the Cowles Foundation during the 1940s and 1950s. However, Granger in [2] and later Engle and Granger in [3] introduced a powerful tool to the field of econometrics for simulating and evaluating economic relationships: the concept of co-integration.

In recent times, the study of these fields has witnessed a convergence through the application of vector error correction models (VECM) and structural vector error correction models (SVEC). A comprehensive theoretical exposition of each of these models can be found in the monographs authored by Lutkepohl [4], Hendry [5], Johansen [6], Hamilton [7], and Banerjee et al. [8]. The main aim of this study is to compare the effectiveness of the univariate SARIMA model with the utilization of VAR methods for analyzing multivariate time series data. The motivation behind this research is to understand which approach is more suitable for modeling three specific time series variables related to Humidity, Rainfall, and Temperature in Ninahvah City, Iraq. The study explores various modeling techniques, including vector autoregressive, structural vector autoregressive, and structural vector error correction models, using the 'vars' package in R. It also offers a range of functions for diagnostic testing, model estimation, prediction, causality analysis, impulse response analysis, and forecast error variance decomposition.

2. Methodology

2.1. Stationary

A time series is classified as stationary when its statistical characteristics remain consistent throughout its duration. These characteristics, such as the mean and variance, remain unchanged over time [9]. Conversely, these properties fluctuate significantly, the time series is considered non-stationary. In practical terms, one can assess the stationary of a time series by visualizing it through a plot. A time series is termed "purely stationary" when the joint distribution of $\mathrm{Z}(\mathrm{t}_1), \cdots, \mathrm{Z}(\mathrm{t}_\mathrm{n})$, and $\mathrm{Z}(\mathrm{t}_1 + \tau), \cdots, \mathrm{Z}(\mathrm{t}_\mathrm{n} + \tau)$, where $z(t)$ represents the random variable at time t, remains constant is the same for all $\mathrm{t}_1, \cdots, \mathrm{t}_\mathrm{n}, \tau$. To put it another way, the joint distributions are mostly

unaffected by changing the time origin by a specific sum; instead, they must be determined by the intervals between $t_1, \cdots, t_n$ [10]. The time series Z_t is deemed to exhibit weak stationarity when two conditions are met: (a) The expected value of $Z_t = \mu$, which is a constant vector with $k-$dimensions, and (b) The covariance of $Z_t = E[(Z_t - \mu)(Z_t - \mu)'] = \Sigma_z$, a constant $k \times k$ positive-definite matrix. The random vector Z's expectation and covariance matrices are indicated by the letters E(Z) and Cov(Z), respectively. To establish if the time series is stationary, the collection of autocorrelations for the time series can also be used. The univariate time series stationary is examined using the Unit Root Test, and a multivariate time series is examined using the Co-integration test [11]. Consider the following two situations:

- When each univariate time series within an MTS item exhibits stationarity, the MTS item itself is considered to be stationary.
- If any of the individual time series within a multivariate time series (MTS) exhibit non-stationarity, a cointegration test should be conducted to verify that the MTS as a whole is also non-stationary. We may make the Z_t series stationary by differencing if it has not already been done. $Z_t = Z_t - Z_{t-1} = \nabla Z_t$ denotes the differenced series. Below are the definitions of the Backward Shift Operator B:

$B^m Z_t = Z_{t-m}$. The backward difference operator ∇ is defined by $\nabla = 1 - B$. Another method for determining whether the data is stationary or not, at lag k, the autocorrelation function is defined as:

$$\rho_k = \frac{E[(Z_t - \mu)(Z_{t+k} - \mu)]}{E[(Z_t - \mu)^2 (Z_{t+k} - \mu)^2]},$$

where z_t : stands for observation. μ: Mean of observation, ρ_k: autocorrelation function.

The cross-correlation for lag k given two time series variables X_t and Y_t is given as $r_{xy} = c_{xy}/s_x s_y$ where, $c_{xy} = \frac{1}{n}\sum_{t=1}^{t-k}(x_t - \bar{x})(y_{t+k} - \bar{y})$, $k = 0, 1, 2 \ldots$; $\bar{x}$ and $\bar{y}$ are the sample means of x_t and y_t, s_x and s_y are the sample standard deviations, respectively [12], Process of White Noise. A white noise process with the formula $a_t = (a_{1t}, \cdots, a_{kt})'$ is a continuous random vector that satisfies the conditions $E(a_t) = 0$, $E(a_t a_t') = \Sigma_a$, and $E(a_t a_s') = 0$ for $s \neq t$. Unless otherwise specified, the $\Sigma_a =$ covariance matrix is assumed to be non-singular as pointed in [11].

2.2. Vector Autoregressive (VAR) Model

One approach to representing the interplay among multiple time-varying variables is through the utilization of the vector autoregressive (VAR) model. This model provides a streamlined representation of dynamic interactions, wherein each internal variable is influenced by its own past values as well as the past values of all other internal variables. The simple p-lag Vector autoregressive VAR (p) method looks like this:

$$Z_t = c + \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_p Z_{t-p} + a_t \,; t = 0, \pm 1, \pm 2, \cdots, \quad (1)$$

where $Z_t = (Z_{1t}, \cdots Z_{kt})'$ is a $(k \times 1)$ vector of time series variable, $\emptyset_i$ are fixed $(k \times k)$ coefficient matrices, $c = (c_1, \cdots, c_k)'$ is a fixed $(k \times 1)$ vector of intercept terms, $a_t = (a_{1t}, \cdots a_{kt})'$ is a white noise procedure with $(k \times 1)$. The procedure can be written clearly in matrix form:

$$\begin{pmatrix} Z_{1t} \\ Z_{2t} \\ \vdots \\ Z_{kt} \end{pmatrix} = \begin{pmatrix} \emptyset_{11}^1 & \emptyset_{12}^1 & \cdot & \cdot & \emptyset_{1k}^1 \\ \emptyset_{21}^1 & \emptyset_{22}^1 & \cdot & \cdot & \emptyset_{2k}^1 \\ \vdots & \vdots & \cdot & \cdot & \vdots \\ \emptyset_{k1}^1 & \emptyset_{k2}^1 & \cdot & \cdot & \emptyset_{kk}^1 \end{pmatrix} \begin{pmatrix} Z_{1t-1} \\ Z_{2t-1} \\ \vdots \\ Z_{kt-1} \end{pmatrix} + \begin{pmatrix} \emptyset_{11}^2 & \emptyset_{12}^2 & \cdot & \cdot & \emptyset_{1k}^2 \\ \emptyset_{21}^2 & \emptyset_{22}^2 & \cdot & \cdot & \emptyset_{2k}^2 \\ \vdots & \vdots & \cdot & \cdot & \vdots \\ \emptyset_{k1}^2 & \emptyset_{k2}^2 & \cdot & \cdot & \emptyset_{kk}^2 \end{pmatrix} \begin{pmatrix} Z_{1t-2} \\ Z_{2t-2} \\ \vdots \\ Z_{kt-2} \end{pmatrix} + \cdots +$$

$$\begin{pmatrix} \emptyset_{11}^p & \emptyset_{12}^p & \cdot & \cdot & \emptyset_{1k}^p \\ \emptyset_{21}^p & \emptyset_{22}^p & \cdot & \cdot & \emptyset_{2k}^p \\ \vdots & \vdots & \cdot & \cdot & \vdots \\ \emptyset_{k1}^p & \emptyset_{k2}^p & & & \emptyset_{kk}^p \end{pmatrix} \begin{pmatrix} Z_{1t-p} \\ Z_{2t-p} \\ \vdots \\ Z_{kt-p} \end{pmatrix} + \begin{pmatrix} a_{1t} \\ a_{2t} \\ \vdots \\ a_{kt} \end{pmatrix}. \quad (2)$$

2.3. Stable VAR (p) Processes [13]

If every root of the matrix lies within the unit circle, and the absolute values of the roots of matrix $\emptyset_i$ are less than 1, then process 1 exhibits stability. That is, if $\det(I_n - \emptyset_1 Z - \cdots - Z^p) \neq 0$ for $|Z| \leq 1$, then a stationary VAR (p) process Z_t ; $t = 0, \pm 1, \pm 2, \cdots$ is stable.

2.4. A Stable VAR(p) Process' Autocovariances

The result of deducting the mean from VAR (p) is

$$Z_t - \mu = \emptyset_1(Z_{t-1} - \mu) + \cdots + \emptyset_p(Z_{t-p} - \mu) + a_t. \quad (3)$$

After dividing both sides by $(Z_{t-1} - \mu)'$ and calculating the expectation, having at $l = 0$ by utilizing:

$$\Gamma_z(i) = \Gamma_z(-i)',$$

$$\Gamma_z(0) = \emptyset_1(Z_{t-l} - \mu) + \cdots + \emptyset_p(Z_{t-p} - \mu) + \Sigma_a,$$

$$= \emptyset_1\Gamma_z(1)' + \cdots + \emptyset_p\Gamma_z(p)' + \Sigma_a. \quad (4)$$

If $\mu > 0$, then

$$\Gamma_z(l) = \emptyset_1\Gamma_z(l-1)' + \cdots + \emptyset_p\Gamma_z(l-p)' + \Sigma_a, \quad (5)$$

If $\emptyset_1, \cdots, \emptyset_p$ and $\Gamma_z(p-1)$ are provided, the auto covariance functions $\Gamma_z(l)$ for $l \geq p$ can be derived from these equations.

2.5. A Stable VAR (p) Process's Autocorrelation

Obtaining the autocorrelations of a stable VAR (p) process is achieved by extracting information from the matrix:

$$R_z(l) = D^{-1}\Gamma_z(l)D^{-1}, \tag{6}$$

hence, D is a diagonal matrix with the Z_t component's standard deviation on the main diagonal. Consequently

$$D^{-1} = \begin{bmatrix} \frac{1}{\sqrt{\gamma_{11}(0)}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sqrt{\gamma_{kk}(0)}} \end{bmatrix}, \tag{7}$$

and $Z_{i,t}$ and $Z_{j,t-1}$ have the following correlation:

$$\rho_{ij}(l) = \frac{\gamma_{ij}(l)}{\sqrt{\gamma_{ii}(0)}\sqrt{\gamma_{jj}(0)}}, \tag{8}$$

which is just the $ij-$ th element of $R_z(l)$. The model's characteristic roots are, once again, the inverses of the solutions. As a result, stationarity necessitates that all characteristic roots have a modulus of less than one. The ACF satisfies the difference equation$\left(1 - \emptyset_1 B - \emptyset_1 B^2 - \cdots - \emptyset_p B^p\right)^p = 0,$ for $p \geq 0$ a stationary AR (p) sequence. The ACF plot of a stationary AR (p) model will display a blend of damped sinusoidal and exponential decay patterns, influenced by the unique source it originates from, leading to varying levels of similarity in the shapes observed.

2.6. Order Selection by VAR

The three selection criteria that will be utilized to evaluate the VAR process order p are as follows:

(i) Employing the Akaike Information Criterion (AIC) method [14], as introduced in [15],

$$AIC(p) = \ln\left|\widetilde{\Sigma}_\varepsilon(p)\right| + \frac{2}{N}\ (\text{number of estimated parameter})$$

$$= \ln\left|\widetilde{\Sigma}_a(p)\right| + \frac{2pk^2}{N}.$$

(ii) Given Hannan and Quinn [16], the Hannan-Quinn Criterion (HQC), follows

$$HQC(p) = \ln\left|\widetilde{\Sigma}_a(p)\right| + \frac{2\ln k}{N} \text{ (Parameters that are freely estimated)}$$
$$= \ln\left|\widetilde{\Sigma}_a(p)\right| + \frac{2\ln(\ln N)}{N} pk^2.$$

(iii) Using the Bayesian Information Criterion (BIC) [17],

$$BIC(p) = \ln\left|\widetilde{\Sigma}_a(p)\right| + \frac{\ln k}{N} \text{ (Parameters that are freely estimated} = \ln\left|\widetilde{\Sigma}_a(p)\right| + \frac{\ln N}{N} pk^2,$$

where the VAR order is p,

The estimated white noise covariance matrix Σ_ε is represented by Σ_a . In a vector time series, there are k different time series components. N is the sample size. Each estimate is selected to minimize the criterion's value in each of the aforementioned parameters.

2.7. Forecasting

If it is determined that the fitted model in 1 is sufficient, forecasts can be made being used. The following estimates are used to create forecasts:

$$\widehat{Z}_t = \hat{c} + \widehat{\emptyset}_1 Z_{t-1} + \widehat{\emptyset}_2 Z_{t-2} + \cdots + \widehat{\emptyset}_p Z_{t-p} + a_t \,; t = 0, \pm 1, \pm 2, \cdots. \quad (9)$$

Given the forecast origin t., the forecasts so produced are those with the smallest mean square error [4].

Using vector moving average models for forecasting (VMA). Considering that the model is recognized and serves as a source for prediction. The VMA forecast (q). Generally, for h-step forward forecast with $h \leq q$, as occurs

$$Z_t(h) = \mu - \sum_{i=1}^{q} \theta_i a_{t+h-i}. \quad (10)$$

Utilizing VARMA models for prediction, we are employing the criterion of minimizing mean-squared error to delve into the future projections of a time series Z_t with a VARMA(p, q) structure, similar to the VAR models of (9). As stated below, for the VARMA (p, q) model

$$Z_t(h) = \emptyset_0 + \sum_{i=1}^{p} \emptyset_i Z_{t+h-i}. \quad (11)$$

3. Applications

Data pertaining to the monthly averages of temperature (T), rainfall (R), and humidity (H) for Ninavah, Iraq, ranging from 1976 to 2001, were examined using the R program. In this example, we'll refer to humidity as (H, Z_{1t}), precipitation as R (R, Z_{2t}), and temperature as (T, Z_{3t}). The multivariate time series can therefore be described using the random vector $Z_t = (Z_{1t}, Z_{2t}, Z_{3t}$. The time series data for these three variables are presented in Figure 1 in a variety of graphical formats. The core scientific challenge outlined in the text revolves around the thorough analysis and modeling of these multivariate time series data, specifically focusing on climate variables—humidity, precipitation, and temperature. The overarching goal is to uncover intricate relationships and discern patterns within the dataset. Additionally, the aim is to construct a robust multivariate model capable of accurate forecasting and in-depth analysis.

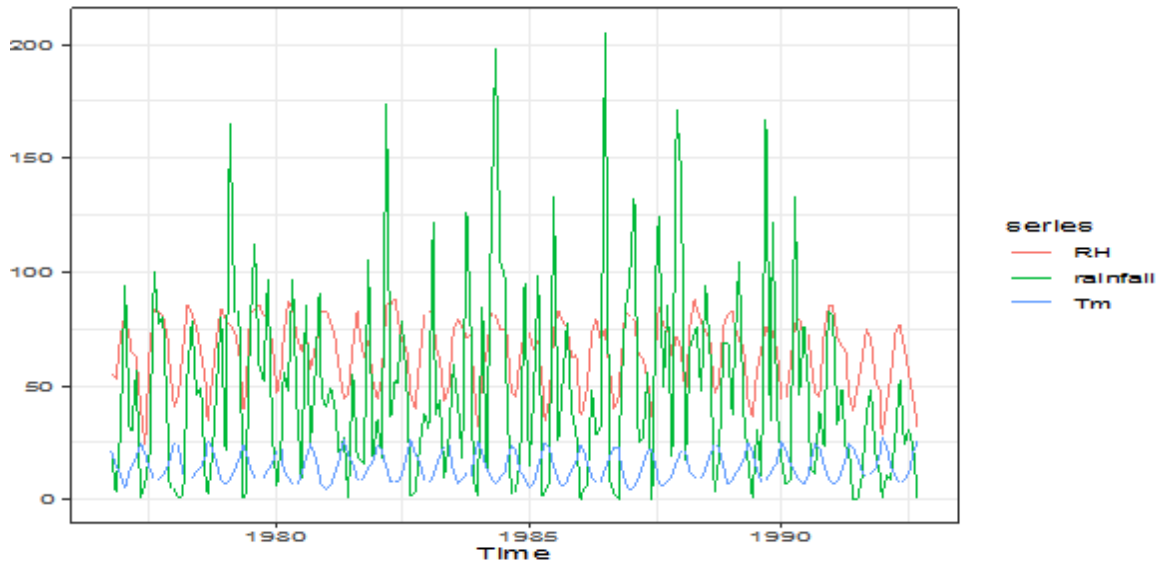


Fig. 1. The three raw series' time series plot (H, R,T).

The Unit Root test is used to determine whether univariate time series datasets are stationary. In contrast, the Co-integration test (original) is used to examine stationarity in multivariate time series datasets. The Augmented Dickey-Fuller (ADF) test can be used to determine whether a series has a unit root. This is predicated on the assumption that a trend-lined series will display a unit root and a significant p-value.

H_0: The data is non-stationary and has a unit root.

H_1: The data are static and have not yet produced the results in Table (1).

Table 1. Original and transformed data stationary testing of Nineveh time series data sets for the period 1976 – 2001

Stationary testing							
Datasets	Responses	Dickey-Fuller	p-value	Phillips-Perron Unit Root Test	p-value	KPSS Level	p-value
R	Z_t	-7.339	0.01	-70.58	0.01	0.25798	0.1
H	Z_t	-5.9331	0.01	-154.35	0.01	0.15987	0.1
T	Z_t	-8.5794	0.01	-67.175	0.01	0.03855	0.1

The alternate proposition becomes relevant when the p-value rejects the null hypothesis and exceeds the 0.05 threshold. In the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, the p-value surpasses the 0.05 threshold, signifying the absence of a unit root and, consequently, stationarity within the series. To ascertain trend stationarity, researchers will assess the null hypothesis; a low p-value suggests the presence of a non-trend stationary signal with a unit root. To test the Stability, all characteristic roots should have a modulus of less than one. The results in Table (2) represent the original data.

Table 2. Roots of the stability characteristic polynomial

The characteristic polynomial's roots							
0.9018	0.9081	0.9081	0.9049	0.9049	0.8934	0.8825	0.8825
0.8803	0.8803	0.871	0.871	0.8214	0.8214	0.8148	0.8148
0.8036	0.8036	0.7866	0.7866	0.7509	0.7509	0.691	0.691
0.4724	0.2125	0.2125					
Log Likelihood -1795.702							

All of the roots k are inside the unit circle. We have no strenuous roots. Our system is generally stable.

3.1. Co-Integration Test

Co-integration testing is a method used to assess the accuracy of long-term linkages between variables because none of them now exhibit stationarity. If the variables exhibit co-integration, it implies that they have an ongoing link, even if they are not stationary at the moment [18]. They also offered the Maximum Eigen Value test and the Trace test as two more methods for counting co-integrated vectors. While the Trace test looks into the potential of r+1 co-integrating vectors, the Maximum Eigen Value test looks into the possibility of a maximum of r co-integrating vectors [19]. They claim that the Maximum Eigen Value test is the best technique for determining the number of co-integrating vectors. After d distinct differentiations, an integrated sequence of order d, designated as I(d), becomes stationary.

H_0: no co-integration of variables H_1: co- integration of variables

The results in Table 3 represent the data.

Table 3. Findings from Johansen's Co-integration Examination for H, R, T

Unrestricted Co-integration Rank Test (Trace)			
Co-integration rank(r)	Eigenvalue	Trace stat.	Critical Value 5%
r = 0*	8.766264e-02	35.55	34.91
r <= 1*	5.444259e-02	18.76	19.96
r <= 2*	4.546822e-02	8.52	9.24
Unrestricted Co-integration Rank test (Maximum Eigenvalue)			
Co-integration rank(r)	Eigenvalue	Trace stat	Critical Value 5%
r = 0*	8.766264e-02	16.79	22.00
r <= 1*	5.444259e-02	10.24	15.67
r <= 2*	4.546822e-02	8.52	9.24

The Trace test reveals the presence of three co-integrating equations with a significance level of 0.05. The asterisk (*) signifies the rejection of the hypothesis at the same 0.05 significance level.

The column of r in Table (3) represents the rank and we know that this is some indication of the number of co-integrating relationships. When r = 0, the test statistic is 35.55 > 34.91. This implies that we do not accept the null hypothesis, which proposes that r > 0. As such, there is some co-integration present. When r < 1, we do not find enough evidence to reject the null hypothesis because 18.76 < 19.9. When r < 2, this again means that we do not find enough evidence to reject the null hypothesis because 8.52 < 9.24. Therefore, we conclude that there is at most 1 co-integrating relationship that presents the Johansson test when we choose the maximal eigenvalue statistic. We are unable to dismiss the null hypothesis. None of the statistical values falls below the 5 percent threshold. It means no co-integrating relationships present the Johansson test.

3.2. The Raw Data Correlation Matrix

The three variables are highly connected, as shown by the correlation matrix ($Z_{t(corr.)}$) below. As a result, the multivariate technique will take into account the interrelation between the variables

$$\text{Corr (H,R,T)} = \begin{matrix} \\ \text{H} \\ \text{R} \\ \text{T} \end{matrix} \begin{matrix} \begin{matrix} \quad 1 \quad & \quad 0.636 \quad & \quad -0.868 \end{matrix} \\ \begin{pmatrix} 1 & 0.636 & -0.868 \\ 0.636 & 1 & -0.461 \\ -0.868 & -0.461 & 1 \end{pmatrix} \end{matrix},$$

$$\text{Cov (H,R,T)} = \begin{matrix} \\ \text{H} \\ \text{R} \\ \text{T} \end{matrix} \begin{matrix} \begin{matrix} \quad 229.1 \quad & \quad 403.2 \quad & \quad -79.9 \end{matrix} \\ \begin{pmatrix} 229.1 & 403.2 & -79.9 \\ 403.2 & 1764 & -117.4 \\ -79.9 & -117.4 & 36.9 \end{pmatrix} \end{matrix}.$$

3.3. The Cross Correlations

Table 4 shows the cross-correlation matrices at various lags (lags 1–12).

High values demonstrate that the variables are interdependent and a multivariate model can be successfully fitted to the data. A simple matrix $s_\ell = [s_{\ell,ij}]$ is constructed for each sample CCM $\hat{\rho}_\ell$ as follows:

$$s_{\ell,ij} = \begin{cases} + & if\ \hat{\rho}_{\ell.ij} \geq 2/\sqrt{T}, \\ - & if\ \hat{\rho}_{\ell.ij} \leq -2/\sqrt{T}, \\ . & if\ |\hat{\rho}_{\ell.ij}| < 2/\sqrt{T}, \end{cases}$$

where: $\hat{\rho}_\ell$ is a consistent estimate of ρ_ℓ, T is a total number.

The results in Table (4) represent the original data.

Table 4. displays example Cross-Correlation Matrices depicting the Monthly Simple Returns of three different Indexes in their raw form (H, R, T).

Lag 1	lag 2	lag3	lag 4	lag 5	lag 6
$\begin{bmatrix} + & + & - \\ + & . & - \\ - & - & + \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$	$\begin{bmatrix} - & - & + \\ - & \cdot & + \\ + & + & - \end{bmatrix}$	$\begin{bmatrix} - & - & + \\ - & \cdot & + \\ + & + & - \end{bmatrix}$	$\begin{bmatrix} - & \cdot & + \\ - & \cdot & + \\ + & + & \cdot \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$
Lag7	**lag 8**	**lag 9**	**lag 10**	**lag11**	**lag 12**
$\begin{bmatrix} + & + & - \\ + & + & - \\ - & - & + \end{bmatrix}$	$\begin{bmatrix} + & + & - \\ + & + & - \\ - & - & + \end{bmatrix}$	$\begin{bmatrix} + & + & - \\ + & \cdot & - \\ - & - & + \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$	$\begin{bmatrix} - & - & + \\ - & - & + \\ + & + & - \end{bmatrix}$	$\begin{bmatrix} - & - & + \\ - & \cdot & + \\ + & + & - \end{bmatrix}$

Table (4) illustrates the simplified CCM for monthly data of (H, R, T). Notable cross-correlations, which are statistically significant at the estimated 5% level, are mainly observable at the lags of 8 and 9.

3.4. Selecting a Model

AIC, BIC, and HQC at various lags are shown in both Table (5), which represents the original data, and Figure (2), which depicts the data. At lag 9, a three- selection process reaches the minimal values (the bolded values). VAR (9) is, therefore, the model of choice in Table (5): Empirical Lag Selection**.**

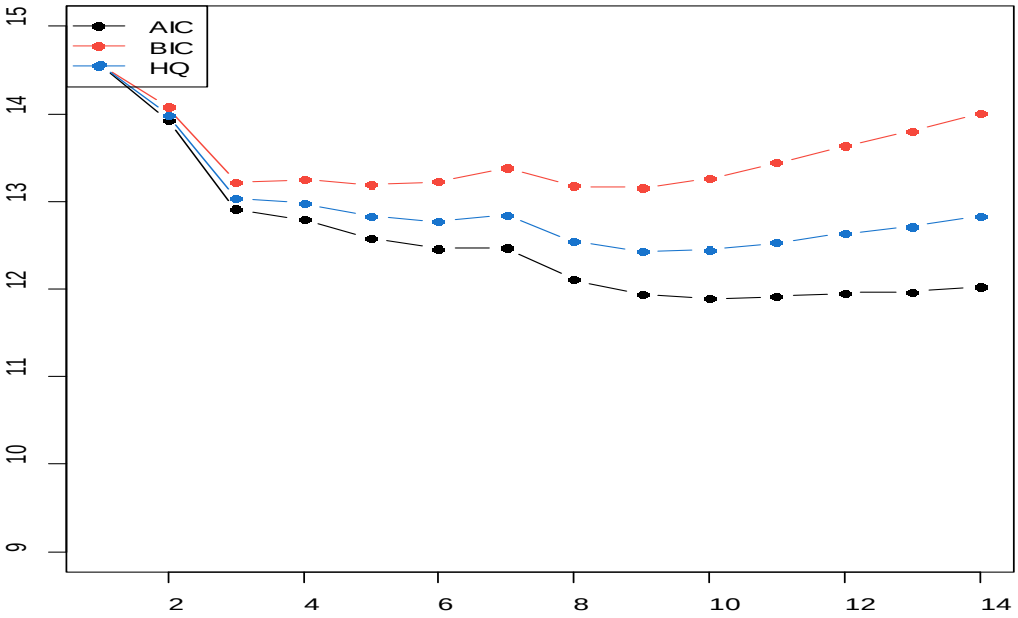


Fig. 2. Illustrates the information criteria applied to (H, R, T) data. The lines, depicted as solid, dashed, and dotted, correspond to AIC, BIC, and HQ, respectively.

Table 5. Empirical Lag Selection

	AIC(n)	HQ(n)	SC(n)	FPE(n)
selection	9	8	8	9
1	14.28293	14.34716	14.44137	15.95881
2	13.20192	13.33038	13.51880	541441.48394
3	13.04705	13.23973	13.52237	463836.94951
4	12.84215	13.09906	13.47591	378027.24396
5	12.73620	13.05735	13.52840	340211.74568
6	12.63510	13.02047	13.58574	307749.95214
7	12.24344	12.69305	13.35252	208259.68283
8	12.04285	12.55669	13.31037	170670.10872
9	11.97888	12.55694	13.40484	160410.83869
10	12.00919	12.65148	13.59359	165760.36272

3.5. Model Presentation

The VAR (9) model with significant parameters is represented in matrix form as seen in Table (6), which represents the original data, utilizing equation (2) in the approach. The optimal lag value is p = 9 according to AIC and FPE, p = 8 based on the HQ criterion, and p = 7 according to the SC criterion. They performed diagnostic analyses on the residuals after calculating a VAR with both a constant and a trend as deterministic predictors for each of the nine different lag orders.

$$\begin{pmatrix} Z_{1t} \\ Z_{2t} \\ Z_{3t} \end{pmatrix} = \begin{pmatrix} 0.341 & 0.019 & -0.187 \\ 0.567 & -0.056 & 0.393 \\ -0.011 & -0.009 & 0.466 \end{pmatrix} \begin{pmatrix} Z_{1t-1} \\ Z_{2t-1} \\ Z_{3t-1} \end{pmatrix} + \begin{pmatrix} -0.126 & 0.023 & 0.672 \\ -0.887 & 0.106 & -0.964 \\ 0.027 & -0.004 & -0.149 \end{pmatrix} \begin{pmatrix} Z_{1t-2} \\ Z_{2t-2} \\ Z_{3t-2} \end{pmatrix} +$$

$$\begin{pmatrix} 0.084 & 0.036 & 0.236 \\ -0.267 & 0.123 & -0.724 \\ -0.039 & 0.002 & -0.008 \end{pmatrix} \begin{pmatrix} Z_{1t-3} \\ Z_{2t-3} \\ Z_{3t-3} \end{pmatrix} + \begin{pmatrix} 0.007 & 0.0139 & 0.315 \\ -0.059 & 0.089 & 1.083 \\ 0.016 & 0.005 & -0.082 \end{pmatrix} \begin{pmatrix} Z_{1t-4} \\ Z_{2t-4} \\ Z_{3t-4} \end{pmatrix} +$$

$$\begin{pmatrix} 0.101 & 0.035 & 0.121 \\ -0.483 & 0.147 & -1.02 \\ -0.008 & -0.004 & -0.131 \end{pmatrix} \begin{pmatrix} Z_{1t-5} \\ Z_{2t-5} \\ Z_{3t-5} \end{pmatrix} + \begin{pmatrix} -0.084 & 0.0001 & 0.632 \\ -0.523 & 0.153 & 2.891 \\ 0.068 & -0.003 & 0.148 \end{pmatrix} \begin{pmatrix} Z_{1t-6} \\ Z_{2t-6} \\ Z_{3t-6} \end{pmatrix} +$$

$$\begin{pmatrix} 0.202 & -0.001 & -0.232 \\ 0.381 & 0.208 & -2.207 \\ -0.034 & -0.000 & 0.177 \end{pmatrix} \begin{pmatrix} Z_{1t-7} \\ Z_{2t-7} \\ Z_{3t-7} \end{pmatrix} + \begin{pmatrix} 0.277 & -0.039 & -0.497 \\ 1.002 & -0.118 & 2.337 \\ -0.069 & 0.011 & 0.291 \end{pmatrix} \begin{pmatrix} Z_{1t-8} \\ Z_{2t-8} \\ Z_{3t-8} \end{pmatrix} +$$

$$\begin{pmatrix} -0.083 & 0.006 & -0.109 \\ 0.196 & -0.075 & -0.087 \\ 0.112 & -0.010 & 0.053 \end{pmatrix} \begin{pmatrix} Z_{1t-9} \\ Z_{2t-9} \\ Z_{3t-9} \end{pmatrix} + \begin{pmatrix} a_{1t} \\ a_{2t} \\ a_{3t} \end{pmatrix}. \qquad (12)$$

The information is presented in Table (6) along with the summarized results and the graphical representation of equation (12).

Table 6. Results for the Endogenous variables: H, R, T

Statistic	H	R	T
Multiple R-Squared	0.991	0.6873	0.9906
F-statistic	635	12.7	608.7
Adjusted R-squared	0.9894	0.6332	0.989
Residual standard error	6.893	38.87	1.643
p-value	< 2.2e-16	< 2.2e-16	< 2.2e-16

$$\text{Cov of residuals} = \begin{matrix} \\ H \\ R \\ T \end{matrix} \begin{matrix} \begin{matrix} H & R & T \end{matrix} \\ \begin{pmatrix} 47.514 & 169.310 & -2.705 \\ 169.310 & 1510.815 & -3.646 \\ -2.705 & -3.646 & 2.699 \end{pmatrix} \end{matrix},$$

$$\text{Corr of residuals} = \begin{matrix} \\ H \\ R \\ T \end{matrix} \begin{matrix} \begin{matrix} H & R & T \end{matrix} \\ \begin{pmatrix} 1 & 0.6319 & -0.2388 \\ 0.6319 & 1 & -0.0571 \\ -0.2388 & -0.0571 & 1 \end{pmatrix} \end{matrix}.$$

4. Diagnostic Testing

Once the multivariate model 12 has been acquired, the next step is to verify the correctness of the model fit. The following diagnostic techniques are used to this end.

4.1. Residual Autocorrelation Function

The following hypothesis is used, as described in Section (1.1.3) of the methodology**:**

$$H_0: \rho_{uv}\, i = 0 \text{ versus } H_1: \rho_{uv}\, i \neq 0$$

We had a total of n= 192 series.

As a result, the residual autocorrelation function's boundary state has the form $\frac{2}{\sqrt{192}} = 0.144$, and H_0 is rejected if $|r_{uv}, i| > \frac{2}{\sqrt{N}} = 0.144$.

When examining the values of autocorrelations in the residual correlation matrices at various lags (lags 12), it was found that none of the residual autocorrelations exceeds 0.144. Figure (1) represents the original data with$|r_{uv}, i|$. This suggests that the residuals conform to a pattern consistent with white noise. To put it another way, the fitted model is sufficient.

1) Test auto correlation for serial correlation (PT) [20]

The graphs, one for each equation, demonstrate the ACF and PACF of the discrepancies, along with a discrepancy plot and a practical distribution chart. Additional justifications are provided by the plot approach for changing its design. Figures (3-5) represent the original data.

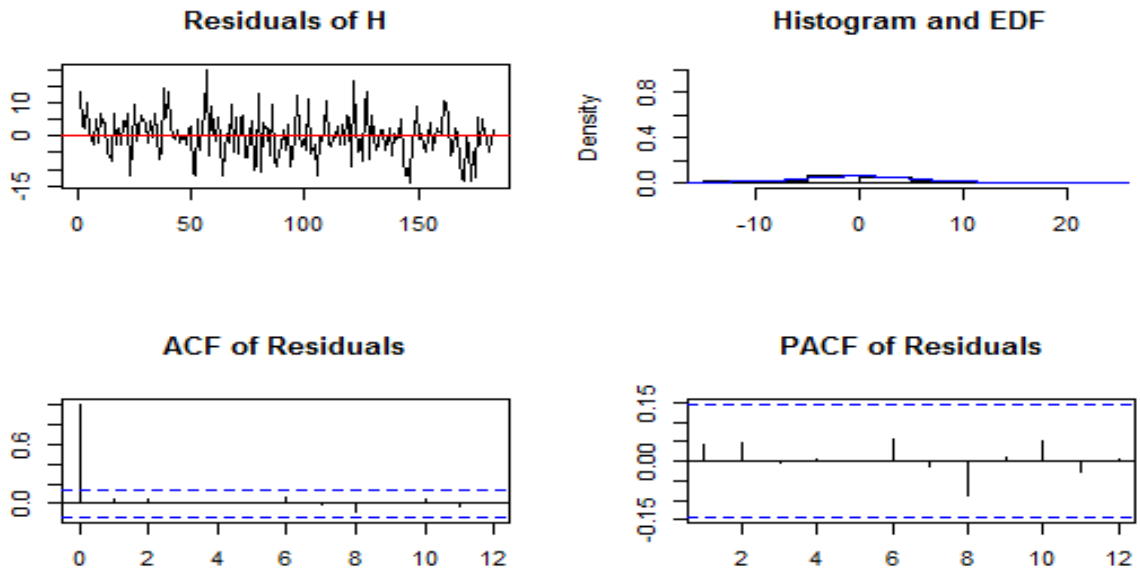


Fig. 3. Explains the Time Series Plots of Residuals (a_{1t}) for H.

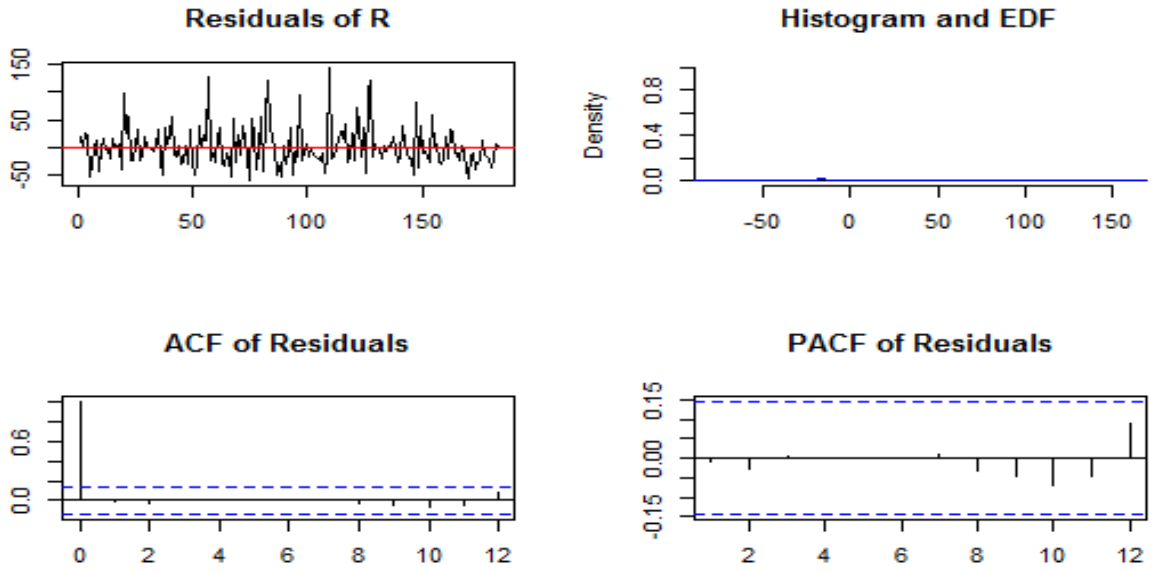


Fig. 4. Explains the Time Series Plots of Residuals (a_{2t}) for R.

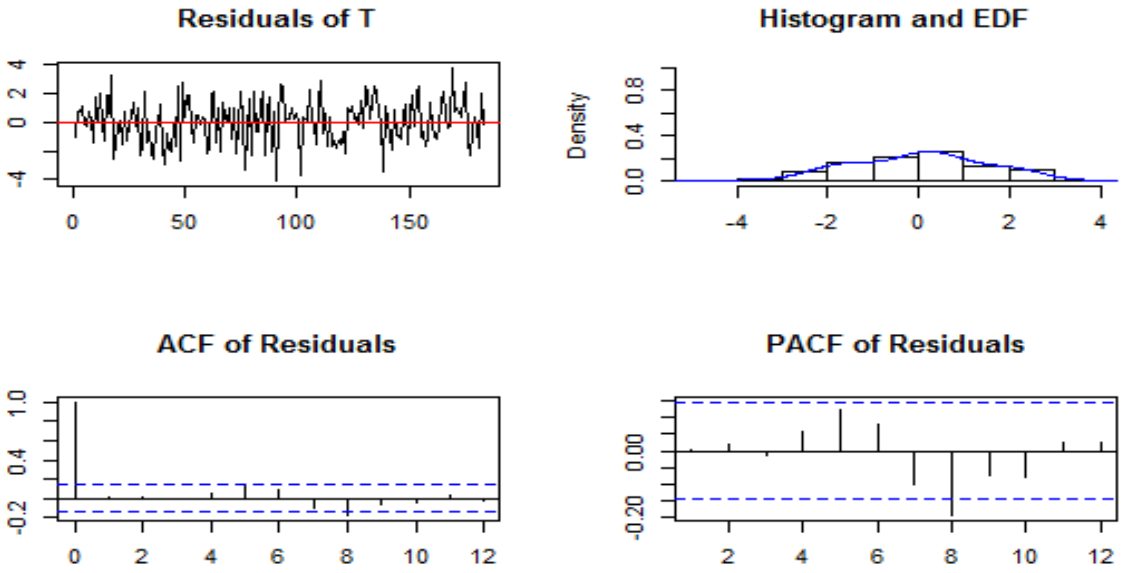


Fig. 5. Explains the Time Series Plots of Residuals (a_{3t}) for T.

Explain the Time Series Plots of Residuals (a_{3t}) for T. Heteroscedasticity: ARCH test ([21], [22]

A statistical model called autoregressive conditional heteroscedasticity (ARCH) is used to assess and forecast volatility in time series. The following regression is the foundation for the multivariate ARCH-LM test. (The univariate test is considered a specific case of the exhibit below and will be omitted):

$$vech(\hat{u}_t\,\hat{u}_t^T) = \beta_0 + B_1 vech(\hat{u}_{t-1}\hat{u}_{t-1}^T) + \ldots + B_q vech(\hat{u}_{t-q}\hat{u}_{t-q}^T) + v_t,$$

$E(u_t) = 0$ and positive time invariant unambiguous covariance matrix $E(u_t\, u_t^T) = \Sigma_u$ (white noise) define u_t as a K-dimensional process [23]. In this context, v_t represents a spherical error process, and the operator 'vech' is used to stack the columns of symmetric matrices, starting from the main diagonal and moving downward. The dimension of β_0 is $\frac{1}{2}K(K+1)$, and for the coefficient matrices B_i where $i = 1, \ldots, q$, $\frac{1}{2}K(K+1) \times \frac{1}{2}K(K+1)$. The null hypothesis is: $H_0 := B_1 = B_2 = \ldots = B_q = 0$ and the alternative is: $H_1 : B_1 \neq 0 \cap B_2 \neq 0 \cap \ldots \cap B_q \neq 0$. The test statistic is explained as: $VARCH_{\mathrm{LM}}(q) = \frac{1}{2}TK(K+1)R_m^2$ with $R_m^2 = 1 - \frac{2}{K(K+1)}\, tr(\widehat{\Omega}\, \widehat{\Omega}_0^{-1})$, and $\widehat{\Omega}$ assigns the above-mentioned regression model's covariance matrix. $\chi^2(qK^2(K+1)^2/4)$ is the distribution of this test statistic.

3) Normality: Jarque & Bera (JB), Skewness, Kurtosis

The Jarque-Bera tests for univariate and multivariate series, as well as separate tests for multivariate skewness and kurtosis (p), are performed on the VAR residuals. By performing the Jarque-Bera test on the residuals following standardization via the Choleski decomposition of the variance-covariance matrix for the centered residuals, one can create a multivariate version of this test. For the multivariate scenario, the test statistics are as follows:

$$JB = \frac{\hat{S}^2(r)}{6\,/T} + \frac{(\widehat{K}\,(r) - 3)^2}{24/T},$$

where T is the sample size, $\hat{S}^2(r)$, $\widehat{K}\,(r)$ are skewness and kurtosis determined the from sample data, and $\widehat{K}\,(r) - 3$ is the excess kurtosis. More specifically, if $\{r_1\,, \ldots\,, r_T\}$ is a variable with T observations. Beloware the definitions for sample skewness and kurtosis.

$$\hat{S}(r\,) = \frac{1}{(T-1)\hat{\sigma}_r^3} \sum_{t=1}^{T} (r_t - \bar{r})^3\,, and\ \widehat{K}(r) = \frac{1}{(T-1)\hat{\sigma}_r^4} \sum_{t=1}^{T} (r_t - \bar{r})^4,$$

when $\hat{\sigma}_r^2$ considering the statistics related to sample variance, $\bar{r}$ is the sample mean of $\hat{S}(r\,)$, it is important to note that both $\hat{S}(r\,)$ and $\widehat{K}(r)$ follow a normal distribution with zero mean and variances of $6/T$ and $24/T$, respectively. This is based on the assumption of normality in the data. As a result of this assumption, the JB statistic conforms to a Chi-square distribution with two degrees of freedom in the asymptotic sense.

To evaluate whether the data conforms to a normal distribution, we can use the JB statistic. If JB exceeds the critical value $JB > \chi^2_{2,1-\alpha}$, where α represents the significance level, then we have grounds to reject the null hypothesis (H_0), which posits that the data follows a normal distribution. These findings are in line with the research conducted in [24], as presented in Table 7, which showcases the original data results.

Table 7. Diagnostic tests of VAR (9) for H, R, T

Null Hypothesis Test		Statistic	p-value
no autocorrelation	PT	92.059	0.00991
no suffer from heteroscedasticity	ARCH	143.25	0.9799
not normality	JB	69.979	4.13e-13
	Kurtosis	37.419	3.751e-08
	Skewness	32.56	3.988e-07

The p-value of 0.00991 is less than the significance level of 0.05, disproving the null hypothesis that there is no autocorrelation. On the other hand, the p-value of the heteroscedasticity (ARCH) test is greater than the 0.05 level of significance, which encourages us to keep the null hypothesis in place. Practically speaking, this means that as the fitted values of the response variable increase, the variance of the residuals should not increase as well. Regarding the Portmanteau Test (PT), the p-value of the normalcy test is below the 0.05 significance level, which allows us to reject the null hypothesis.

4) Structural Stability (SVC) [25]

The stability test is used to determine if there are any structural breaks. If we are unable to test for structural breaks and one occurs, the entire estimate may be thrown off. To avoid this, we use a simple inspection technique that involves plotting the cumulative total of subsequent residuals. A structural change has occurred at that particular junction if the total sum of the data points on the chart exceeds certain essential criteria. Fig. 6, which shows the unedited dataset, serves as an illustration of this occurrence.

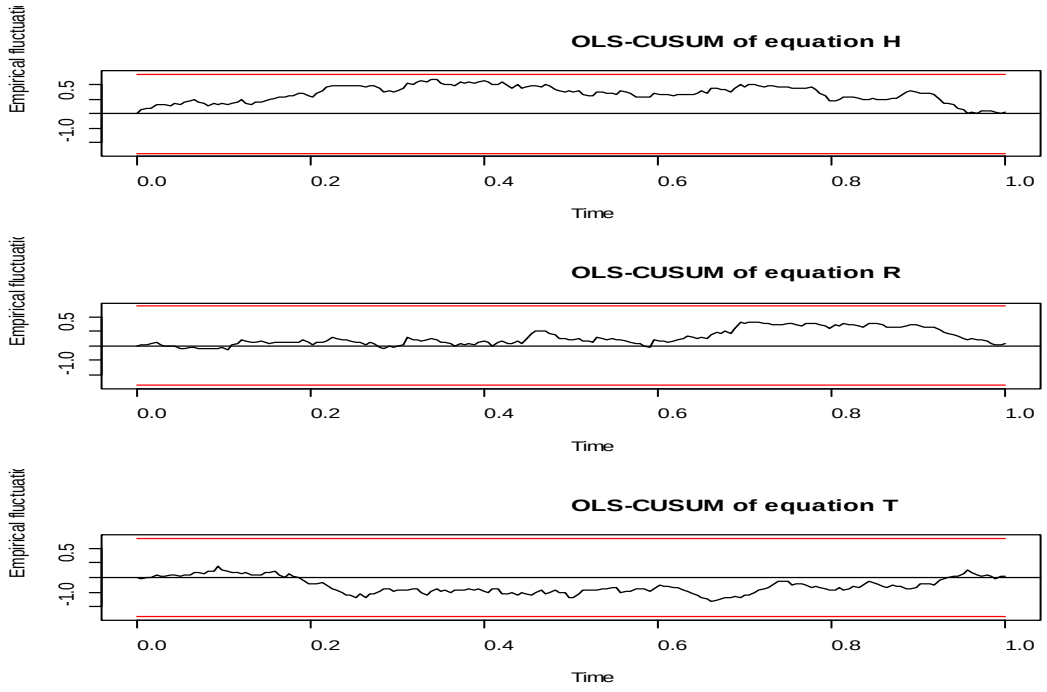


Fig. 6. CUSUM Test for H, R,T

There are no points on this graph beyond the two red lines, so the system is stable.

4.2. Granger Causality

The interdependence structure of the underlying systems of multi-variate time series was investigated. Utilizing Granger causality analysis, we can rephrase the content related to the outcomes presented in Table (8), which encapsulates the unaltered dataset.

Table 8. Causality tests for H, R, T

Null Hypothesis	Statistic(F-test)	p-value
H does not Granger-cause R ,T	3.3014	6.448e-06
R does not Granger-cause H ,T	1.9389	0.01179
No instantaneous causality between: H and R, T	55.968	7.028e-13
T does not Granger-cause H ,R	3.6086	1.0 24e-06
No instantaneous causality between: R and H ,T	53.091	2.962e-12
No instantaneous causality between: T and H,R	12.242	0.002196

We reject the null hypothesis (H_0) due to the p-value being below the significance level of 0.05.

4.3. Forecasting

The built model can be used to generate forecasts since it meets the basic assumption of the model adequacy. The MSE values produced using the program R, are shown in Table (9), which represents the data along with the multivariate model's forecasts for the period (Oct.2000 - May.2001). Table (9) represents the optimal parameters of the multivariate, univariate and MSE for the fitted ARIMA model.

Table 9. Multivariate VAR (9) model's and univariate and MSE for the fitted ARIMA (H, R,T) time series

Time Series	VAR(9)	SARIMA $(1,0,0)(1,1,1)_8$
	MSE	MSE
H	49.2073	15.6396
R	497.190	366.388
T	3.2624	2.2405

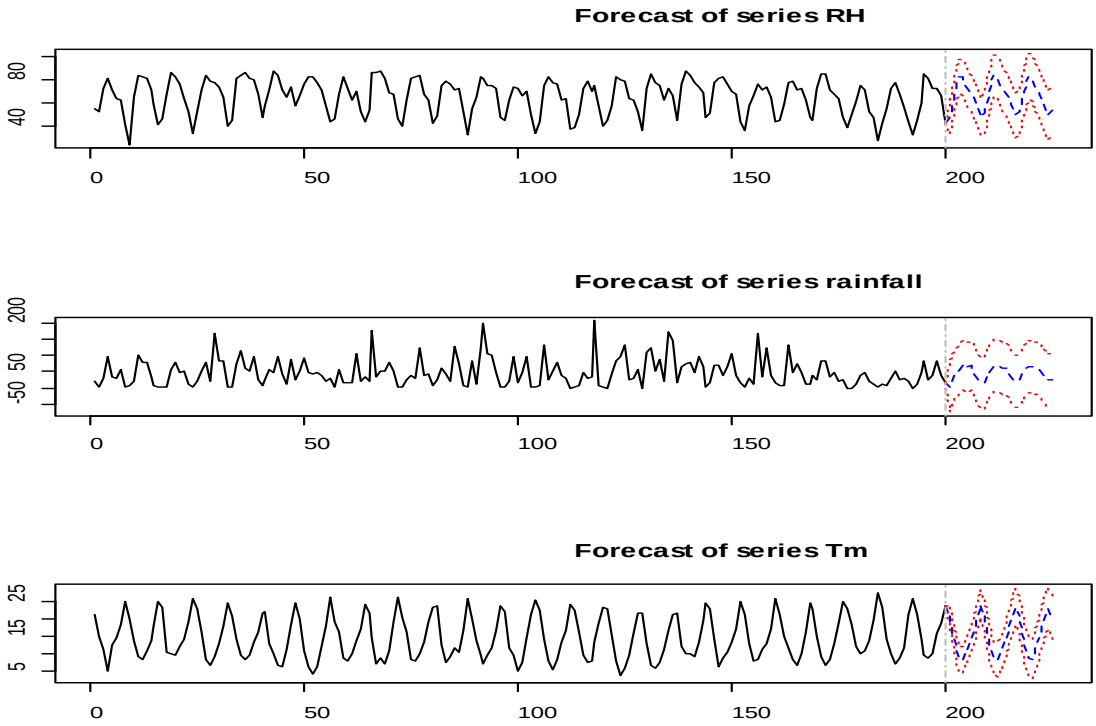


Fig. 7. Forecasts of the Multivariate Model VAR (9)

4.4. Forecast Error Variance Decomposition (FEVD)

A Forecast Error Variance Decomposition assesses the mutual influence of variables through the utilization of the VAR model. To determine the FEVD, we analyze the forecast errors from each equation within the fitted VAR model. Subsequently, the prepared VAR model quantifies the proportion of each error manifestation attributed to unanticipated fluctuations in the counterpart variable (forecast errors). The variance decomposition method aids in the interpretation of the VAR model. The amount of variance in the dependent variable described by each independent variable can be determined. FEVD describes how a potential shock in a one-time series affects the future uncertainty in the other time series of the system. Since this process progresses over time, a shock to a time series can be insignificant in the short run but critical in the long run. When a vector autoregression (VAR) model is used, FEVD, a crucial technique in econometrics and many multivariate time series analytic contexts, helps to comprehend its consequences. The degree to which one variable in the autoregression influences the others is revealed by this decomposition of variance. It evaluates the percentage of forecast error variation for each variable that may be attributable to external shocks affecting the other variables in the context of the data shown in Fig.7.

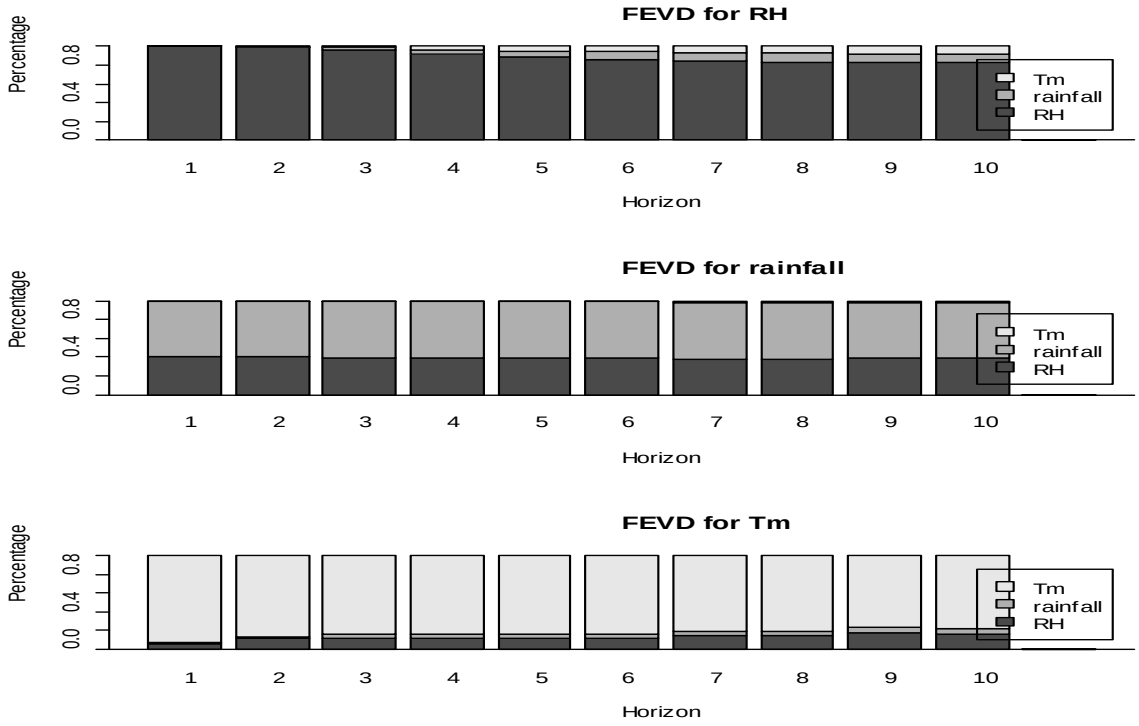


Fig. 8. Forecast Error Variance Decomposition from VAR (9) model fit.

These Fig. 8 graphs show percentages of the shook. The first plot depicts the FEVD for RH starts. It appears that although we were borderline on whether to conclude that Granger causes RH starts, the FEVD reveals that the magnitude of the causality is tiny anyway, while that of RH is greater on rainfall and Tm starts. The second plot shows the FEVD for rainfall. It appears that although we were borderline on whether to conclude that RH starts rate Granger cause rainfall and Tm.

5. Conclusion

The evolution of numerous vars package functions and strategies is described in this article. These improvements give researchers an easy-to-use setting for conducting VAR, SVAR, and SVEC analyses. This is primarily accomplished by putting impulse response function implementation approaches into practice, breaking down forecast error variance, making forecasts, and offering diagnostic testing tools. It also provides tools for determining the model's ideal lag duration, evaluating stability and causation, and performing further diagnostic tests. The article also covers how to determine the co-integrating rank using VECM, which can easily be changed into its level-VAR equivalent. The data was not stationary, as we observed. However, an effective method of transforming a non-stationary series is stationary. To ascertain the model's order, compute the differences and build a correlogram. SARIMA (1,0,0)(1,1,1)8 was chosen for univariate, and the VAR model was then used. Then, using MSE, we assess the forecasting precision. After examining each forecasting accuracy, we concluded that SARIMA would produce better results than VAR in the presence of low-correlated variables and the absence of numerous co-integrations among variables because of its higher forecasting accuracy. They should be aware that there is a

correlation. When variables exhibit a strong correlation, the VAR model can be utilized to yield highly favorable outcomes. Limitations of this study include the focus on monthly environmental variables in Ninahvah City, Iraq. Future research could explore other regions, incorporate additional variables, and assess model performance under diverse climatic conditions.

Reference

[1] C.A. Sims, "Macroeconomics and reality", Econometrica: *Journal of the Econometric Society*, pp. 1-48, 1980.

[2] C.W. Granger, "Some properties of time series data and their use in econometric model specification", *Journal of Econometrics*, vol. 16, no. 1, pp. 121-130, 1981.

[3] R.F. Engle, & C.W.J. Granger, "Co-integration and error correction: representation, estimation, and testing", Econometrica: *Journal of the Econometric Society*, pp. 251-276, 1987.

[4] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*, Springer Science & Business Media, 2005.

[5] D. F. Hendry, *Dynamic econometrics*, Oxford University Press, 1995.

[6] S. Johansen, "Likelihood-Based Inference in Cointegrated Vector Autoregressive Models", *Oxford University Press*, New York, 1995.

[7] J. D. Hamilton, *Time series analysis*, Princeton University Press, Princeton, N.J., 1994.

[8] A. Banerjee, J.J. Dolado, J.W. Galbraith, & D. Hendry, *Co-integration, error correction, and the econometric analysis of non-stationary data*, Oxford University Press, 1993.

[9] J.M. Liu, R. Chen, L.M. Liu, & J.L. Harris, "A semi-parametric time series approach in modeling hourly electricity loads", *Journal of Forecasting*, vol. 25, no. 8, pp. 537-559, 2006.

[10] G.S. Maddala and I.M. Kim, *Unit roots, cointegration, and structural change*, 1998.

[11] W.W. Wei, *Multivariate time series analysis and applications*, John Wiley & Sons, 2018.

[12] R.S. Tsay, *Multivariate time series analysis: with R and financial applications*, John Wiley & Sons, 2013.

[13] A.M. Sathe, & N.S. Upadhye, "Estimation of the Parameters of Vector Autoregressive (VAR) Time Series Model with Symmetric Stable Noise", *arXiv preprint* arXiv:2104.07262, 2021.

[14] K. Yamaoka, T. Nakagawa, & T. Uno, "Application of Akaike's information criterion (AIC) in the evaluation of linear pharmacokinetic equations", *Journal of Pharmacokinetics and Biopharmaceutics*, vol. 6, no. 2, pp. 165-175, 1978.

[15] H. Akaikei, "Information theory and an extension of maximum likelihood principle", *In Proc. 2nd Int. Symp. on Information Theory*, pp. 267-281, 1973.

[16] E.J. Hannan, & B.G. Quinn, "The determination of the order of an autoregression", *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 41, no. 2, pp. 190-195, 1979.

[17] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461-464, 1978.

[18] S. Johansen, "Statistical analysis of cointegration vectors", *Journal of Economic Dynamics and Control*, vol. 12, no. 2-3, pp. 231-254, 1988.

[19] G.S. Maddala and I.M. Kim, *Unit roots, cointegration, and structural change*, 1998.

[20] V. Patilea, & H. Raïssi, "Corrected portmanteau tests for VAR models with time-varying variance", *Journal of Multivariate Analysis*, vol. 116, pp. 190-207, 2013.
[21] H. Bouakez, & M. Normandin, "Fluctuations in the foreign exchange market: How important are monetary policy shocks?", *Journal of International Economics*, vol. 81, no. 1, pp. 139-153, 2010.
[22] O.C. Onifade, & S.O. Olanrewaju, "Investigating performances of some statistical tests for heteroscedasticity assumption in generalized linear model: A Monte Carlo simulations study", *Open Journal of Statistics*, vol. 10, no. 3, pp. 453-493, 2020.
[23] H. Lütkepohl, "Structural vector autoregressive analysis for cointegrated variables", *Allgemeines Statistisches Archiv*, vol. 90, pp. 75-88, 2006.
[24] C.M. Jarque, & A.K. Bera, "A test for normality of observations and regression residuals", *International Statistical Review/Revue Internationale de Statistique*, pp. 163-172, 1987.
[25] H. Herwartz, "Hodges–Lehmann detection of structural shocks–an analysis of macroeconomic dynamics in the euro area", *Oxford Bulletin of Economics and Statistics*, vol. 80, no. 4, pp. 736-754, 2018.

Միաչափ SARIMA և բազմաչափ VAR մոդելների համեմատական վերլուծություն ժամանակային շարքերի կանխատեսման համար. կլիմայի փոփոխականների դեպքի ուսումնասիրություն Իրաքի Նինահվա քաղաքում

Սամիրա Ա. Օթմա, Հասան Հ. Ջամիլ և Սադեք Թ. Աբդուլազիզ

Դուհոկի համալսարան, Հիմնական կրթության քոլեջ, Դուհոկ, Իրաք
e-mail: Sameera.Othman@uod.ac, hasan.hazim@uod.ac, sadiq.taha@uod.ac

Ամփոփում

Այս ուսումնասիրությունը ներառում է համեմատություն SARIMA միաչափ մոդելի կիրառման և VAR մեթոդների (վեկտորային ավտոռեգրեսիվ մոդելներ) օգտագործման միջև բազմաչափ ժամանակային շարքերի վերլուծության համար: Վերլուծությունն իրականացվում է եռաժամանակյա շարքի փոփոխականների միջոցով, որոնք ստացվել են Իրաքի Նինահվա քաղաքում խոնավության (H), տեղումների (R) և ջերմաստիճանի (T) ամսական միջինը ներկայացնող տվյալներից: Այս շարքերը մոդելավորելու համար օգտագործվում են և՛ միաչափ, և՛ բազմաչափ ժամանակային շարքերի մոտեցումները: Հոդվածը ուրվագծում է նաև վեկտորային ավտոռեգրեսիայի, կառուցվածքային վեկտորի ավտոռեգրեսիայի և կառուցվածքային վեկտորի սխալի ուղղման մոդելների իրականացումը vars փաթեթի միջոցով: Բացի այդ, այն ապահովում է ախտորոշիչ

թեստավորման, սահմանափակ մոդելների գնահատման, կանխատեսման, պատճառահետևանքային վերլուծության, իմպուլսային արձագանքի վերլուծության և կանխագուշակման սխալի շեղումների տարրալուծման գործառույթներ: Բացի այդ, այս մոդելները գնահատելու համար ներդրվում են երեք հիմնարար գործառույթներ՝ VAR, SVAR և SVEC: Մեթոդների համեմատությունը հիմնված է յուրաքանչյուր մոտեցման արդյունքում առաջացած միջին սխալի գնահատման վրա: Հետազոտության արդյունքները ցույց են տալիս, որ միաչափ գծային ստացիոնար մեթոդները գերազանցում են բազմաչափ մոդելներին: Տվյալների վերլուծությունը կատարվել է R ծրագրային հարթակի միջոցով: Հիմնական նպատակը տվյալների մշակման մեջ միաչափ և բազմաչափ ժամանակային շարքերի մոդելների կատարողականի գնահատումն է: Հետազոտության բացը կայանում է ամսական բնապահպանական փոփոխականների համատեքստում ժամանակային շարքերի վերլուծության SARIMA և VAR մեթոդների համեմատական գնահատման անհրաժեշտության մեջ: Այս մոդելներն ընտրվել են ժամանակային սերիաների տվյալների մեջ բազմաթիվ փոփոխականների միջև ժամանակային կախվածությունների և փոխազդեցությունների գրանցման արդյունավետության շնորհիվ՝ ապահովելով Իրաքի Նինահվա քաղաքի կլիմայական օրինաչափությունների համապարփակ վերլուծություն: Ուսումնասիրության նպատակն է լրացնել հետազոտության բացը՝ համեմատելով այս մոդելները և հիմնավորելով դրանց ընտրությունը՝ հիմնվելով նշված ժամանակային շարքի տվյալները վերլուծելու նրանց կարողությունների վրա:

Բանալի բառեր՝ Միաչափ ժամանակային շարքեր, Բազմաչափ գործընթաց, Խաչաձև հարաբերակցություն և VAR, Կանխատեսում, ARCH-LM թեստ, Կառուցվածքային կայունություն (SVC)

Сравнительный анализ одномерных SARIMA и многомерных VAR моделей для прогнозирования временных рядов: тематическое исследование климатических переменных в городе Нинахва, Ирак

Самира А. Отман, Асан А. Джамил и Садек Т. Абдулазиз

Университет Духока, Колледж базового образования, Духок, Ирак
e-mail: Sameera.Othman@uod.ac, hasan.hazim@uod.ac, sadiq.taha@uod.ac

Аннотация

Данное исследование включает сравнение применения одномерной модели SARIMA и использования методов VAR (векторных авторегрессионных моделей) для многомерного анализа временных рядов. Анализ проводится с использованием переменных трех временных рядов, полученных на основе данных, представляющих среднемесячные

значения влажности (H), количества осадков (R) и температуры (T) в городе Нинахва, Ирак. Для моделирования этих рядов используются как одномерные, так и многомерные подходы к временным рядам. В статье также описывается реализация моделей векторной авторегрессии, структурной векторной авторегрессии и структурной векторной коррекции ошибок с использованием пакета vars. Кроме того, он предоставляет функции для диагностического тестирования, оценки моделей с ограничениями, прогнозирования, анализа причинно-следственных связей, анализа импульсных характеристик и разложения дисперсии ошибок прогноза. Кроме того, для оценки этих моделей вводятся три фундаментальные функции: VAR, SVAR и SVEC. Сравнение методов основано на оценке средней ошибки, создаваемой каждым подходом. Результаты исследования показывают, что одномерные линейные стационарные методы превосходят многомерные модели. Анализ данных проводился с использованием программной платформы R. Основная цель — оценить эффективность одномерных и многомерных моделей временных рядов при обработке данных. Пробел в исследованиях заключается в необходимости сравнительной оценки методов SARIMA и VAR для анализа временных рядов в контексте ежемесячных переменных окружающей среды. Эти модели были выбраны из-за их эффективности в определении временных зависимостей и взаимодействий между множеством переменных в данных временных рядов, обеспечивая всесторонний анализ климатических моделей в городе Нинахва, Ирак. Исследование направлено на устранение пробелов в исследованиях путем сравнения этих моделей и обоснования их выбора на основе их возможностей анализировать указанные данные временных рядов.

Ключевые слова: Одномерный временной ряд, Многомерный процесс, Взаимная корреляция и VAR, Прогноз, Тест ARCH-LM, Структурная устойчивость (SVC)

Mathematical Problems of Computer Science 61, 50–61, 2024.

doi: 10.51408/1963-0114

UDC 510.5

P-*m*-Mitotic Sets and Arithmetical Hierarchy

Arsen H. Mokatsian[1] and Khachatur A. Barseghyan[2]

Institute for Informatics and Automation Problems of NAS RA, Yerevan, Armenia
Siemens Industry Software, Yerevan, Armenia
e-mail: arsenmokatsian@gmail.com, khachatur.barseghyan@outlook.com

Abstract

Let $\{0,1\}^*$ be the set of all finite strings of elements from $\{0,1\}$, and let **P** be the class of problems recognized by deterministic Turing machines, which run in polynomial time (a problem is simply a subset of $\{0,1\}^*$). This article defines the class $\widehat{\mathbf{P}}$ and shows that $\widehat{\mathbf{P}}$ is isomorphic to the class **P**.

Based on the notions of *T*-mitoticity and *T*-autoreducibility, K.Ambos-Spies introduced the notions of *P*-*m*-mitoticity and *P*-*m*-autoreducibility. The notions of $\widehat{\mathrm{P}}$-*m*-mitoticity and $\widehat{\mathrm{P}}$-*m*-autoreducibility are introduced by analogy with the mentioned notions.

The article proves that the index sets $\{z \mid W_z$ is $\widehat{\mathrm{P}}$-*m*-mitotic$\}$ and $\{z \mid W_z$ is $\widehat{\mathrm{P}}$-*m*-autoreducible$\}$ are Σ_3-complete.

Keywords: Arithmetical hierarchy, *P*-*m*-mitotic set, *P*-*m*-autoreducible set, index set.
Article info: Received 29 February 2024; sent for review 19 March 2024; accepted 24 May 2024.

1. Introduction

Information about the basic concepts of computability theory used in this article, in particular the Turing machine (TM), the numbering of computably enumerable sets $\{W_i\}_{i\in\omega}$ and the arithmetical hierarchy, can be found in Rogers [1], Soare [2].

The two definitions of polynomial time reducibility given by Karp [3] and Cook [4] are just time-bounded versions of many-one reducibility $(\leq_m)$ and Turing reducibility $(\leq_T)$.

Among other works devoted to the research of time-bounded computations and used in this article, we note the works of Ladner [5], Ambos-Spies [6], Hopcroft, Ullman [7](1979), Sipser [8], Arora, Barak [9], Terwijn [10].

Notation. We fix the alphabet $\Lambda = \{0,1\}$.

Given a set Y, the set of all finite strings of elements from Y is denoted by Y^*.

A Turing machine T (deterministic or nondeterministic) runs in polynomial time if there is a polynomial function q such that for every input of length n, any computation sequence of T halts in $q(n)$ or fewer moves.

It is an intuitively appealing notion that **P** is the class of problems that can be solved efficiently.

In this article, we consider the class $\widehat{\mathbf{P}}$ (see Definition 2). Proposition 1 (below Definition 3) shows that the classes **P** and $\widehat{\mathbf{P}}$ are isomorphic (i.e., there is an isomorphic mapping from **P** to $\widehat{\mathbf{P}}$ and vice versa, there is an isomorphic mapping from $\widehat{\mathbf{P}}$ to **P**; with respect to the relations in question) (see the definition of isomorphic mapping in Definition 1).

An oracle Turing machine runs in polynomial time if there exists a polynomial function q such that for *every* input of length n and *any* oracle set X, the machine halts within $q(n)$ steps (see Ladner [5], p.156).

Note that the definitions of R. Ladner [11] and other authors are based on the concept of a multitape Turing machine.

Based on the notions of T-mitoticity and T-autoreducibility, Ambos-Spies [6] introduced the notions of P-m-mitoticity and P-m-autoreducibility. By analogy with the mentioned notions we introduce the notions of $\widehat{\mathrm{P}}$-m-mitoticity and $\widehat{\mathrm{P}}$-m-autoreducibility (see Definitions 15,16) and also give the definitions of index sets $M(P\text{-}m)=\{z \mid W_z$ is $\widehat{\mathrm{P}}$-m-mitotic$\}$ and $A(P\text{-}m) = \{\mathrm{z} \mid W_z$ is $\widehat{\mathrm{P}}$-T-autoreducible$\}$.

This article studies the location of index sets $\{z \mid W_z$ is $\widehat{\mathrm{P}}$-m-mitotic$\}$ and $\{\mathrm{z} \mid W_z$ is $\widehat{\mathrm{P}}$-m-autoreducible$\}$ in the arithmetical hierarchy.

2. Preliminaries

Notation. Let ω be the set of all nonnegative integers.

We will denote the Λ^* elements by lowercase Greek letters $\sigma, \tau, \ldots$

Let us denote that $\sigma\hat{}\tau$ denote the *concatenation* of string σ followed by τ.

Let $<$ be the natural order on Λ^* ($\lambda < 0 < 1 < 00 < 01 < \cdots$), where λ represents the empty string.

We will denote the subsets ofΛ^* by uppercase Greek letters $\Xi, \Theta, \cdots$, as well as by the Latin letter P with subscripts (P_i).

If $\sigma \in \Lambda^*$, then $|\sigma|$ denote the length of σ.

If $\Xi \subseteq \Lambda^*$, then

$$\Xi(\sigma) = \begin{cases} 1, & \text{if } \sigma \in \Xi \\ 0, & \text{if } \sigma \notin \Xi. \end{cases}$$

If $A \subseteq \omega$, then $A(x) = \chi_A(x)$(where χ_A is a characteristic function of a set A).

Define the mappings h_0, h_1 as follows:

Let h_0 be a 1-1 mapping from ω onto Λ^*, $h_0(0) = \lambda$, $h_0(n+1) = (n+2)$-nd string according to the order of strings on Λ^*.

Let h_1 be a 1-1 mapping from Λ^*onto ω.

$h_1(\lambda) = 0$;

$h_1(n+1$ string according to the order of strings on $\Lambda^*)= n$ (In fact, $h_1 = h_0^{-1}$).

Definition 1. (*i*) Let two sets $\mathfrak{M}$ and $\widetilde{\mathfrak{M}}$ be given. Let there be defined any sort of relations between the elements of each of these sets.

If it is possible to place the two sets into one-to-one correspondence so that the mapping preserves the relations; that is, if with every element a of $\mathfrak{M}$ there can be associated an element b of $\widetilde{\mathfrak{M}}$ in a biunique manner so that the relations existing between any elements $a, b, \cdots$ of $\mathfrak{M}$ also exist between the associated elements $\bar{a}, \bar{b}, \cdots$ and vice versa, then the two sets are called *isomorphic* (with respect to the relations in question), and we write $\mathfrak{M} \cong \widetilde{\mathfrak{M}}$. The mapping itself is called an *isomorphism* (see Waerden [11], pp. 25-26).

(*ii*) If in two sets $\mathfrak{M}$ and $\mathfrak{N}$ certain relations are defined (such as $a < b$ or $ab = c$) and if to each element a of $\mathfrak{M}$ an image $\bar{a} = \varphi a$ is assigned in such a manner that all relations between the elements of $\mathfrak{M}$ also hold for the images (so that, for example, $a < b$ implies $\bar{a} < \bar{b}$ in the cases of the relation <), then φ is called a *homomorphic mapping* or *homomorphism* from $\mathfrak{M}$ to $\mathfrak{N}$ (see Waerden [11], p. 28).

Remark. It can be proved that the mapping $h_1: \Lambda^* \to \omega$ is an isomorphism.

It is known that there exist effective enumerations of the sets $P_0, P_1, \ldots$ and oracle Turing machines $\boldsymbol{M}_0, \boldsymbol{M}_1, \ldots$, where P_i denotes the set recognized by the Turing machine (also denoted by P_i), which runs in polynomial time, and $\boldsymbol{M}_i$ denotes the oracle Turing machine, which runs in polynomial time. $\boldsymbol{M}_i(A)$ denotes the set recognized by $\boldsymbol{M}_i$ with oracle A (see Ladner [5], p.157).

Notation. For a given function f, $f\restriction x$ denotes the restriction of f to arguments y<x, and $A\restriction x$ denotes $\chi_A\restriction x$.

(Note that any string $\sigma \in \Lambda^*$can be considered as a partial function from ω into Λ.)

Let $h_0(A) = \{\tau|(\exists x)\,[h_0(x) = \tau \;\&\; x \in A]\}$,$h_1(\Xi) = \{x|(\exists\tau)[h_1(\tau) = x \;\&\; \tau \in \Xi]\}$.

Let $\sigma \in \Lambda^*$. By σ' we denote a string γ such that $h_0(\gamma) = h_0(\sigma) + 1$.

Let $\hat{h}$ be a computable function from ω onto ω^2.

Let Q_e be the Turing program with code number e (also called *index* e) in the standard listing (of programs), and let φ_e be the partial function computed by Q_e (see Soare [2], p.14).

We write $\varphi_{e,s}(x) = y$ if $x, y, e < s$ and y is the output of $\varphi_e(x)$ in $< s$ steps of the Turing program Q_e. If such a y exists, we say $\varphi_{e,s}(x)$ *converges*, which we write as $\varphi_{e,s}(x)\downarrow$, and $\varphi_{e,s}(x)\uparrow$ otherwise. Similarly, we write $\varphi_e(x)\downarrow$ if $\varphi_{e,s}(x)\downarrow$ for some s, and we write $\varphi_e(x)\downarrow= y$ if $\varphi_e(x)\downarrow= y$ and $\varphi_e(x) = y$ and similarly for $\varphi_{e,s}(x)\downarrow= y$ (see Soare [2], pp.16-17).

$W_e = \operatorname{dom}\varphi_e$.

Based on the available numbering of computably enumerable (c.e.) sets $\{W_i\}_{i\in\omega}$, the available numbering of computable operators $\{\boldsymbol{\Phi}_i\}_{i\in\omega}$, and the available enumeration of polynomials $\{q_i\}_{i\in\omega}$, we define for an arbitrary i (proceeding from the fact that $\hat{h}(i) = (i_0, i_1)$)

1) the *set* $\widehat{P}_i$ as follows: $(\forall x)(\forall s \geq q_{i_1}(x))\left[\hat{P}_{i,s}(x) = W_{i_0,q_{i_1}(x)}(x)\right]$,
it is obvious that $(\forall x)(\forall s \geq q_{i_1}(x))\left[\hat{P}_{i,q_{i_1}(x)}(x) = \hat{P}_{i,s}(x) =_{dfn} \widehat{P}_i(x)\right]$;

2) the *oracle Turing machine* $\widehat{\boldsymbol{M}}_i$ as follows:
$(\forall x)\left(\forall s \geq q_{i_1}(x)\right)(\forall \sigma)\left[\widehat{\boldsymbol{M}}_{i,s}(\sigma)(x) = \boldsymbol{\Phi}_{i_0,q_{i_1}(x)}(\sigma)(x)\right]$,
it is obvious that $(\forall x)\left(\forall s \geq q_{i_1}(x)\right)(\forall \sigma)\left[\widehat{\boldsymbol{M}}_{i,q_{i_1}(x)}(\sigma)(x) = \widehat{\boldsymbol{M}}_{i,s}(\sigma)(x) =_{dfn} \widehat{\boldsymbol{M}}_i(\sigma)(x)\right]$.

Definition 2. $\widehat{\mathbf{P}} = \{\hat{P}_i \mid i \in \omega\}$ (note, that $\mathbf{P} = \{P_i \mid i \in \omega\}$).

Based on the above and similar statements , which are also presented, for example, by Hopcroft, Ullman [7], Sipser [8], Arora, Barak [9],Terwijn [10], the following conclusion is presented in [9]:

All low-level choices (number of tapes, alphabet size, etc..) in the definition of Turing machines are immaterial, as they will not change the definition of **P** (see Arora, Barak [9], p. 30).

Thus, since neither the number of tapes nor the way the inputs and outputs are presented (binary coding or natural numbers) significantly affect, we can assert that

$$(\forall i)(\exists j)(\forall x)\left[\hat{P}_i(x) = P_j\left(h_0(x)\right)\right] \,\&\, (\forall j)(\exists i)(\forall \sigma)\left[P_j(\sigma) = \hat{P}_i\left(h_1(\sigma)\right)\right]$$

and

$$(\forall i)(\exists j)(\forall x)(\forall A)\left[\widehat{\boldsymbol{M}}_i(A)(x) = \boldsymbol{M}_j(h_0(A))(h_0(x))\right] \,\&\, (\forall j)(\exists i)(\forall \sigma)(\forall \Xi)\left[\boldsymbol{M}_j(\Xi)(\sigma) = \widehat{\boldsymbol{M}}_i(h_1(\Xi))(h_1(\sigma))\right].$$

In [12], the existence of a homomorphic mapping from $\widehat{\mathbf{P}}$ to **P** and, vice versa, the existence of a homomorphic mapping from $\widehat{\mathbf{P}}$ to **P** (with respect to the relations in question) were proved.

Now we will prove that **P** and $\widehat{\mathbf{P}}$ are isomorphic (with respect to the relations in question).

Define the relations in **P** and $\widehat{\mathbf{P}}$.

Definition 3. (*i*) Let $P_i, P_j \in \mathbf{P}$. P_i is to the *left* of $P_j\left(P_i <_L P_j\right)$ if $(\exists\, \gamma \in \Lambda^*)(\forall\, \tau \leq \gamma)$
$[P_i(\tau) = P_j(\tau)\ \&\ P_i(\gamma') < P_j(\gamma')]$ (i. e. $P_i(\gamma') = 0\ \&\ P_j(\gamma') = 1$);

(*ii*) Let $\hat{P}_i, \hat{P}_j \in \widehat{\mathbf{P}}$. $\hat{P}_i$ is to the *left* of $\hat{P}_j\left(\hat{P}_i <_L \hat{P}_j\right)$ if $(\exists\, x)(\forall\, y < x)$
$\left[\hat{P}_i(y) = \hat{P}_j(y)\ \&.\ \hat{P}_i(x+1) < \hat{P}_j(x+1)\right]$(i. e., $\hat{P}_i(x+1) = 0\ \&\ \hat{P}_j(x+1) = 1$).

It is shown in [12] that there is a homomorphic mapping from **P** to $\widehat{\mathbf{P}}$ and vice versa, there is a homomorphic mapping from $\widehat{\mathbf{P}}$ to **P** (with respect to the relations in question).

Proposition 1. The classes **P** and $\widehat{\mathbf{P}}$ are isomorphic.

Let's define the mapping $\varrho\colon \omega \to \omega$.

Let j_0 be such that $(\forall \sigma)\left[P_0(\sigma) = \hat{P}_{j_0}\left(h_1(\sigma)\right)\right]$.

(As noted above, for P_0 there exists such $\hat{P}_{j_0}$)

1) Define $\varrho(0) = j_0$.

n+1) Suppose that $(\forall k_0 \leq n)(\forall k_1)\left(\varrho(k_0 \neq k_1)\right)$.

Let m_0 be such that $(\forall \sigma)\left[P_n(\sigma) = \hat{P}_{m_0}\left(h_1(\sigma)\right)\right]$.

If $m_0 \notin \{\varrho(0), \varrho(1), \cdots \varrho(n-1)\}$, then define $\varrho(n) = m_0$.

If $m_0 \in \{\varrho(0), \varrho(1), \cdots \varrho(n-1)$, then let m_1 be such that $(\forall \sigma)\left[P_n(\sigma) = \hat{P}_{m_1}\left(h_1(\sigma)\right)\right]$ & $m_1 \notin \{\varrho(0), \varrho(1), \cdots \varrho(n-1)\}$. Such m exists because according to the Padding Lemma (see

Soare [2], p.15, Rogers [1], p. 22), $(\forall v_0)(\exists v \geq v_0)$ [v is the index of c.e. set (i.e. the domain of the p.c. function φ_v) such that $(\forall x)(W_v(x) = \hat{P}_m(x))$ and for all x $W_v(x)$ is computed in the same time as $\hat{P}_m(x)$].

Then define $\varrho(n) = m_1$. (Thus the definition of mapping ϱ is completed.)

Let P_i,P_j are such that P_i<P_j .

Then $(\exists\, \gamma \in \Lambda^*)(\forall\, \tau \leq \gamma)[P_i(\tau) = P_j(\tau)$ & $P_i(\tau) = \hat{P}_{\varrho(i)}\big(h_1(\tau)\big)$ & $P_i(\gamma') = 0$ & $P_j(\gamma') = 1\,]$. Since $P_i(\gamma') = \hat{P}_{\varrho(i)}\big(h_1(\gamma')\big) = 0$ and $P_j(\gamma') = \hat{P}_{\varrho(j)}\big(h_1(\gamma')\big) = 1$ then $\hat{P}_{\varrho(i)}\big(h_1(\gamma')\big) < \hat{P}_{\varrho(j)}\big(h_1(\gamma')\big)$.

As $(\forall\, \tau \leq \gamma)\big[\hat{P}_{\varrho(i)}\big(h_1(\tau)\big) = \hat{P}_{\varrho(j)}\big(h_1(\tau)\big)\big]$ then $\hat{P}_{\varrho(i)} < \hat{P}_{\varrho(j)}$ (according to Definition 3).

So, if $P_i < P_j$, then $\hat{P}_{\varrho(i)} < \hat{P}_{\varrho(j)}$.

Thus, there is a mapping $\mathbf{P}\rightarrow\widehat{\mathbf{P}}$ such that it preserves the order, i.e., there is an isomorphic mapping from $\mathbf{P}$ to $\widehat{\mathbf{P}}$.

Similarly, one can prove the existence of an isomorphic mapping from $\widehat{\mathbf{P}}$ to $\mathbf{P}$. So, we can say that the classes $\mathbf{P}$ and $\widehat{\mathbf{P}}$ are isomorphic (with respect to the relations in question).

2.1. Premliminaries about P-T-mitoticity

Definition 4. Define $\Theta \leq_T^P \Xi$, if there exists an i such that $B = \boldsymbol{M}_i(A)$ (see Ladner [5], Ambos-Spies [6]).

Definition 5. Define $B \leq_T^{\hat{P}} A$ if there is an i such that $B = \widehat{\boldsymbol{M}}_i(A)$.

Definition 6. A *splitting of* A is a pair A_1, A_2 of c.e. sets such that $A_1 \cap A_2 = \emptyset$. We sometimes will write A=$A_1 \sqcup A_2$ if A_1, A_2 is a splitting of A (see Downey, Stob [13], p. 4).

Definition 7. A c.e. set$A$ is *T-mitotic* if there is a splittingA_1, A_2 of A such that $A_1 \equiv_T A_2 \equiv_T A$ (see Downey, Stob [13], p. 83, Lachlan [14], pp. 9-10).

Let us recall some information about *T*-autoreducibility.

Definition 8. We say that a partial recursive functional $\boldsymbol{\Psi}$ is an *autoreduction* if, for all X and n, the computation of $\boldsymbol{\Psi}(X, n)$ includes no question of the form "$n \in X$?". A set A is *T-autoreducible* if there exists an autoreduction $\boldsymbol{\Psi}$ such that $A = \boldsymbol{\Psi}(\mathrm{A})$ (see Trakhtenbrot [15], Ladner [16], p. 199).

From the definition of *T*-autoreducibility it follows:

A is T-autoreducible $\Leftrightarrow (\exists e)(\forall x)(\boldsymbol{\Phi}_e(A \cup \{x\})(x)) = A(x)) \Leftrightarrow$

$(\exists e)(\forall x)(\boldsymbol{\Phi}_e(A - \{x\})(x)) = A(x))$.

Ambos-Spies introduced the following notions:

a) A computable set Ξ is *P-T-mitotic* if there is a set $\Theta \in \mathbf{P}$ such that $\Xi \equiv_T^P \Xi \cap \Theta \equiv_T^P \Xi \cap \overline{\Theta}$. Otherwise, Ξ is *non-P-T-mitotic* (see Ambos-Spies [6], p. 4).

b) A computable set Ξ is *P-T-autoreducible* if for some $n \in \omega$ and every $\sigma \in \Lambda^*$, $\Xi\,(\sigma) = \boldsymbol{M}_n(\Xi\, - \{\sigma\})(\sigma)$ (see Ambos-Spies [6], p.19).

(Ambos-Spice prefers the expression "$\Xi\,(\sigma) = \boldsymbol{M}_n(\Xi\, - \{\sigma\})(\sigma)$" instead of the equivalent expression "$\Xi\,(\sigma) = \boldsymbol{M}_n(\Xi\, \cup \{\sigma\})(\sigma)$". For the sets of nonnegative numbers, the expression "$A(x) = \boldsymbol{M}_n(A \cup \{x\})(x)$" is used in the definition of *T*-autoreducibility, for example, in Downey, Slaman [17], p. 121.)

Ambos-Spies has proved that

(*i*) if Ξ is *P-T*-mitotic, then Ξ is *P-T*-autoreducible(see Ambos-Spies [6], p.19),

(*ii*) there is a computable set Ξ, which is *P-T*-autoreducible, but not *P-T*-mitotic (see Ambos-Spies [6], p. 21).

We represent the definitions of $\hat{P}$-*T*-mitoticity and $\hat{P}$-*T*-autoreducibility according to Ambos-Spies with slight modifications (see Ambos-Spies [6]).

Definition 9. A computable set A is $\hat{P}$-*T-autoreducible* if for some $n \in \omega$ and every $x \in \omega, A(x) = \widehat{\boldsymbol{M}}_n(A \cup \{x\})(x)$.

Definition 10. A computable set A is $\hat{P}$-*T-mitotic* if there is a set $B \in \widehat{\mathbf{P}}$ such that $A \equiv_T^{\hat{P}} A \cap B \equiv_T^{\hat{P}} A \cap \bar{B}$. Otherwise, A is *non-*$\hat{P}$*-T-mitotic*.

Let us give the definitions of index sets $T(\hat{P})M,\ AT(\hat{P})$.

Definition 11. $T(\hat{P})M=\{z \mid W_z$ is $\hat{P}$-*T*-mitotic$\}$,

$AT(\hat{P}) = \{z \mid W_z$ is $\hat{P}$-*T*- autoreducible$\}=\{z \mid (\exists i)(\forall x)[M_i(W_z \cup \{x\})(x) = W_z(x)]$ & (W_z is computable)$\}$.

2. 2. Premliminaries about P-m- mitoticity

Definition 12. (Computing a function and running time)

Let $f: \Lambda^* \to \Lambda^*$ and let $T: \omega \to \omega$ be some functions, and let M be a Turing machine (TM). We say that *M computes f in* $T(n)$*-time* (we write $T(n)$-time instead of T-time, for emphasis that T is applied to the input length), if for every $\sigma \in \Lambda^*$, if M is initialized to the start configuration on input σ, then after at most $T(|\sigma|)$ steps it halts with $f(\sigma)$ written on its output tape.

We say that *M computes f* if it computes f in $T(n)$ time for some function $f: \omega \to \omega$. (see Arora, Barak [9], p. 17)

Definition 13. $\{f_n: n \epsilon \omega\}$ is the effective enumeration of ***PF*** (the class of deterministically polynomial time computable functions from Λ^* to Λ^*).

Ξ is *polynomial time many-one* (*P-m*) *reducible* to Θ ($\Xi \leq_m^P \Theta$), if for some n, $(\forall \sigma \in \Lambda^*)$ $\left(\Xi(\sigma) = \Theta\left(f_n(\sigma)\right)\right)$ (see Ambos-Spies [6], p.2).

By analogy, for arbitrary n we will define the function $\hat{f}_n: \omega \to \omega$.

Let $\{\varphi_i\}_{i \epsilon \omega}$ be the enumeration of the partial computable (p.c.) functions of one variable and T_j be the Turing machine which computes the p.c. function φ_j (see Soare [2], p.12, Rogers [1], p. 12). Remind, that $\hat{h}$ is a computable function from ω onto ω^2. Then we define (proceeding from the fact that $\hat{h}(i) = (i_0, i_1)$) the function $\hat{f}_i$ (for all i) as follows:

Definition 14. *a*) For arbitrary i let T_{i_0} be initialized to the start configuration on input x. Then define the function $\hat{f}_i$ as follows:

$$\hat{f}_i(x) = \begin{cases} \text{the total number of 1's, appearing anywhere on the} \\ \text{tape, after u-th step,} \quad \text{if } \left(\exists u \leq q_{i_1}\right)\left(T_{i_0} \text{ stops at } u\text{-th step}\right); \\ \text{the total number of 1's, appearing anywhere on the} \\ \text{tape, just after } q_{i_1}\text{-th step,} \quad \text{otherwise .} \end{cases}$$

b) A is $\hat{P}$-*m-reducible* to B $\left(A \leq_m^{\hat{P}} B\right)$, if $(\exists i)(\forall x)(\forall s_1 \geq q_{i_1}(x))(\exists s_2 \geq s_1)$.

Definition 15. A computable set Ξ is *P-m-mitotic* if Ξ is finite or cofinite if there is a set $\Theta \epsilon \boldsymbol{P}$ such that $\Xi \equiv_m^P \Xi \cap \Theta \equiv_m^P \Xi \cap \overline{\Theta}$ (see Ambos-Spies [6], p. 4).

Definition 16. A computable set Ξ is *P-m-autoreducible* if Ξ is finite, or cofinite, or if for some $f \epsilon \boldsymbol{PF}$, $\Xi \leq_m \Xi$ via f and $(\forall \sigma \in \Lambda^*)\,(f(\sigma) \neq \sigma)$ (see Ambos-Spies [6], p.19).

Definition 17. A computable set A is $\hat{P}$-*m-mitotic* if A is finite or cofinite if there is a set $B \epsilon \widehat{\boldsymbol{P}}$ such that $A \equiv_m^{\hat{P}} A \cap B \equiv_m^{\hat{P}} A \cap \bar{B}$ (see Ambos-Spies [6], p. 4).

Definition 18. A computable set A is $\hat{P}$-*m-autoreducible* if A is finite, or cofinite, or if $(\exists i)[A \leq_m A$ via $\hat{\mathrm{f}}_{\mathrm{i}}$, and $(\forall x)\,(\hat{f}_i(x) \neq x))$.

Definition 19. For any given class $\mathcal{E}$ of computably enumerable sets, let $IND_{\mathcal{E}} = \{z | W_z \in \mathcal{E}\}$. If $A = IND_{\mathcal{E}}$ for some $\mathcal{E}, A$ is called an *index set* (see Rogers [1], p. 324).

Let us give the definitions of index sets $M(\hat{P}\text{-}m)$, $A(\hat{P}\text{-}m)$.

Definition 20. $M\left(\hat{P}\text{-}m\right) = \{z \mid W_z$ is $\hat{P}$-m-mitotic$\}$,

$A(\hat{P}\text{-}m) = \{\mathrm{z} \mid W_z$ is $\hat{P}$-m- autoreducible$\}$.

3. Results

To formulate the main results, we remind the following definitions:

Definition 21. A set A is Σ_n-complete (Π_n-complete) if $A \in \Sigma_n(\Pi_n)$ and $B \leq_1 A$ for every $B \in \Sigma_n(\Pi_n)$ (it makes no difference whether we use "$B \leq_m A$" or "$B \leq_1 A$" in the definition of Σ_n-complete and Π_n-complete) (see Soare [2], p. 64).

Definition 22. $Rec = \{z | W_z$ is computable (recursive)$\}$, $Fin = \{z | W_z$ is finite$\}$, $Cof = \{z | \overline{W}_z$ is finite$\}$ (see Soare [2], p. 21).

It is known that Fin is Σ_2-complete, Cof and Rec are Σ_3-complete (see Soare [2], pp. 65-67, Rogers [1], pp. 327-328).

One of the approaches to the problem of lower bounds (called a *reducibility approach* in [1]) is to take certain distinguished sets as standard "*reference points*" and to obtain bounds on the level (and degree) of any other given set by establishing reducibility relationships between it and one or more of the reference sets. In most cases, we shall use sets complete in Σ_n or Π_n $(n > 0)$ as *reference sets*, and we shall use *m*-reducibility. The reducibility approach is particularly useful for getting lower bounds on level (and degree). In conjunction with the Tarski-Kuratowski algorithm (and the strong hierarchy theorem), it sometimes enables us to identify not only the level but, indeed, the recursive-isomorphism type of a given set (see Rogers [1], p. 325).

Lemma 1. *Let* $\mathcal{E}$ *be the class of computably enumerable sets, such that* $IND_{\mathcal{E}} \supseteq Cof$, $\overline{IND}_{\mathcal{E}} \supseteq \overline{Rec}$ *and* $IND_{\mathcal{E}} \in \Sigma_3$. *Then* $IND_{\mathcal{E}}$ *is* Σ_3*-complete* (*note that* $\overline{Rec} = \{z \mid W_z$ *is non-computable*$\}$).

Proof. To prove Lemma 1, we use Rogers' proof of index set *Rec's* Σ_3-completeness (see Rogers [1], pp. 327-328). To do this, the Σ_3-complete *reference set B* is used (where $B = \{x \mid (\exists y)[y \in W_x \,\&\, W_x$ is infinite$]\}$) and it is proved that $B \leq_m Cof$ (namely, such a general computable function g is constructed that $[z \in B \iff \mathrm{g(z)} \in Rec]$. Moreover, the construction is such that eventually $[z \in B \implies \mathrm{g(z)} \in Cof]$ and $[z \notin B \implies \mathrm{g(z)} \in \overline{Rec}]$.

Thus, if the class $\mathcal{E}$ satisfies the requirements of Lemma 1, the abovementioned function g will m-reduce B to $IND_{\mathcal{E}}$ (i.e., $z \in B \iff \mathrm{g(z)} \in IND_{\mathcal{E}}$). And since $IND_{\mathcal{E}} \in \Sigma_3$, then $IND_{\mathcal{E}}$ is Σ_3-complete. □

In the article [12] it is proved that $AT(\hat{P})$ and $T(\hat{P})M$ are Σ_3-complete.

Theorem 1. $A(\hat{P}\text{-}m)$ *is* Σ_3*-complete.*

Proof. Let's first prove that $A(\hat{P}\text{-}m) \in \Sigma_3$.

$z \in A(\hat{P}\text{-}m) \iff [W_z$ is computable$]\,\&\,[[(\exists i)[W_z \leq_m^{\hat{P}} W_z$ via $\hat{f}_i \,\&\, (x) \neq x] \vee (W_z$is finite$) \vee$
$(W_z$ is cofinite$)] \iff (\exists z_1)(\forall n)(\forall u_0)\,(\exists u_1 \geq u_0)[(n \in W_{z,u_1} \,\&\, n \notin W_{z_1,u_1}) \vee$
$(n \notin W_{z,u_1} \,\&\, n \in W_{z_1,u_1}) \,\&\, [[(\exists i)(\forall x)(\forall s_1 \geq q_{i_1}(x))(\exists s_2 \geq s_1)$
$[W_{z,s_2}(x) = W_{z,s_2}\left(\hat{f}_{i,s_2}(x)\right) \,\&\, \hat{f}_{i,s_2}(x) \neq x] \vee (\exists t_0)(\forall t_1)[t_1 \leq t_0 \vee$
$W_{z,t_1} = W_{z,t_0} \vee (\exists v_0))(\forall v_1)(\exists t_2)[v_1 \leq v_0) \vee v_1 \in W_{z,t_2}]] \iff$
$(\exists z_1)(\exists i)(\exists \mathrm{t}_0)(\exists v_0)$
$(\forall n)(\forall u_0)(\forall x)(\forall s_1 \geq q_{i_1}(x))(\forall t_1)(\forall v_1)$
$(\exists u_1 \geq u_0)(\exists s_2 \geq s_1)(\exists t_2)$
$[[(n \in W_{z,u_1} \,\&\, n \notin W_{z_1,u_1}) \vee (n \notin W_{z,u_1} \,\&\, n \in W_{z_1,u_1}) \,\&$
$[[\, W_{z,s_2}(x) = W_{z,s_2}\left(\hat{f}_{i,s_2}(x)\right) \,\&\, \hat{f}_{i,s_2}(x) \neq x] \vee$
$[t_1 \leq t_0 \vee W_{z,t_1} = W_{z,t_0}] \vee [v_1 \leq v_0 \vee v_1 \in W_{z,t_2}]]]$.

Thus, $T(\hat{P})M \in \Sigma_3$.

Since $A(\hat{P}\text{-}m) \supseteq Cof$, $\overline{A(\hat{P}\text{-}m)} \supseteq \overline{Rec}$ and $A(\hat{P}\text{-}m) \in \Sigma_3$ it follows from Lemma 1 that $A(\hat{P}\text{-}m)$ is Σ_3-complete. □

Theorem 2. $M(\hat{P}\text{-}m)$ *is* Σ_3*-complete.*

Proof. Let's first prove that $M(\hat{P}\text{-}m) \in \Sigma_3$.

$z \in M(\hat{P}\text{-}m) \iff [W_z$ is computable$]\,\&\,[[(\exists i_0)[W_z \equiv_m^{\hat{P}} (W_z \cap \hat{P}_{i_0}) \equiv_m^{\hat{P}} (W_z \cap \bar{\hat{P}}_{i_0})\,]$
$\vee (W_z$ is finite$) \vee (W_z$ is cofinite$)] \iff$
$(\exists z_1)(\forall n)(\forall u_0)\,(\exists u_1 \geq u_0)[(n \in W_{z,u_1} \,\&\, n \notin W_{z_1,u_1}) \vee (n \notin W_{z,u_1} \,\&\, n \in W_{z_1,u_1}) \,\&$
$[[(\exists i)(\exists j)(\forall x_1)(\forall s_1 \geq max(q_{i_1}(x_1), q_{j_1}(x_1)))(\exists s_2 \geq s_1)$
$[W_{z,s_2}(x_1) = (W_{z,s_2} \cap \hat{P}_{i,s_2})\left(\hat{f}_{j,s_2}(x_1)\right)] \,\&$
$(\exists k)(\forall x_2)(\forall s_3 \geq max(q_{i_1}(x_2), q_{k_1}(x_2)))(\exists s_4 \geq s_3)$

$[(W_{z,s_4} \cap \hat{P}_{i,s_4})(x_2) = W_{z,s_4}\left(\hat{f}_{k,s_4}(x_2)\right)]$ &
$(\exists l)(\forall x_3)(\forall s_5 \geq max(q_{i_1}(x_3), q_{l_1}(x_3)))(\exists s_6 \geq s_5)$
$[W_{z,s_6}(x_3) = (W_{z,s_6} \cap \bar{\hat{P}}_{i,s_6})\left(\hat{f}_{l,s_6}(x_3)\right)]$ &
$(\exists m)(\forall x_4)(\forall s_7 \geq max(q_{i_1}(x_4), q_{m_1}(x_4)))(\exists s_8 \geq s_7)$
$[(W_{z,s_8} \cap \bar{\hat{P}}_{i,s_8})(x_4) = W_{z,s_8}\left(\hat{f}_{m,s_8}(x_4)\right)]$ $\vee$
$(\exists t_0)(\forall t_1)[t_1 \leq t_0 \vee W_{z,t_1} = W_{z,t_0}] \vee (\exists v_0)(\forall v_1)(\exists t_2)[v_1 \leq v_0 \vee v_1 \in W_{z,t_2}]] \Leftrightarrow$
$(\exists z_1)(\exists i)(\exists j)(\exists k)(\exists l)(\exists m)(\exists \mathrm{t}_0)(\exists v_0)$
$(\forall n)(\forall u_0)(\forall x_1)(\forall s_1 \geq max(q_{i_1}(x_1), q_{j_1}(x_1)))(\forall x_2)(\forall s_3 \geq max(q_{i_1}(x_2), q_{k_1}(x_2)))$
$(\forall x_3)(\forall s_5 \geq max(q_{i_1}(x_3), q_{l_1}(x_3)))(\forall x_4)(\forall s_7 \geq max(q_{i_1}(x_4), q_{m_1}(x_4)))(\forall t_1)(\forall v_1)$
$(\exists u_1 \geq u_0)(\exists s_2 \geq s_1)(\exists s_4 \geq s_3)(\exists s_6 \geq s_5)(\exists s_8 \geq s_7)(\exists t_2)$
$[\left(n \in W_{z,u_1} \,\&\, n \notin W_{z_1,u_1}\right) \vee \left(n \notin W_{z,u_1} \,\&\, n \in W_{z_1,u_1}\right)$ &
$[W_{z,s_2}(x_1) = (W_{z,s_2} \cap \hat{P}_{i,s_2})\left(\hat{f}_{j,s_2}(x_1)\right)]$ &.
$[(W_{z,s_4} \cap \hat{P}_{i,s_4})(x_2) = W_{z,s_4}\left(\hat{f}_{k,s_4}(x_2)\right)]$ &
$[W_{z,s_6}(x_3) = (W_{z,s_6} \cap \bar{\hat{P}}_{i,s_6})\left(\hat{f}_{l,s_6}(x_3)\right)]$ &
$[(W_{z,s_8} \cap \bar{\hat{P}}_{i,s_8})(x_4) = W_{z,s_8}\left(\hat{f}_{m,s_8}(x_4)\right)]$ $\vee$
$[t_1 \leq t_0 \vee W_{z,t_1} = W_{z,t_0}] \vee [v_1 \leq v_0 \vee v_1 \in W_{z,t_2}]]$.

Thus, $M(\hat{P}\text{-}m) \in \Sigma_3$.

Since $M(\hat{P}\text{-}m) \supseteq Cof$, $\overline{M(\hat{P}\text{-}m)} \supseteq \overline{Rec}$ and $M(\hat{P}\text{-}m) \in \Sigma_3$, it follows from Lemma 1 that $M(\hat{P}\text{-}m)$ is Σ_3-complete.□

4. Conclusion

It is known that an effective enumeration of the sets of the class **P** (namely, $P_0, P_1, \cdots, P_i, \cdots$) exists and, thus, $\mathbf{P} = \{P_i | i \in \omega\}$. Based on the available numbering of computably enumerable sets $\{W_i\}_{i\in\omega}$, a sequence of sets of non-negative numbers $\hat{P}_i$ is constructed such that their effective enumeration exists and $\hat{\mathbf{P}} = \{\hat{P}_i \mid i \in \omega\}$ by definition.

It is shown that the class $\hat{\mathbf{P}}$ is isomorphic to the class **P**. Using traditional methods, it is shown that the index sets $A(\hat{P}\text{-}m)$ and $M(\hat{P}\text{-}m)$ are Σ_3-sets. Applying the method used by H.Rogers in proving the Σ_3-completeness of the index set $\{z \mid W_z \text{ is computable}\}$, it is proved that the index sets $A(\hat{P}\text{-}m) = \{z \mid W_z \text{ is } \hat{\mathrm{P}}\text{-}m\text{-autoreducible}\}$ and $M(\hat{P}\text{-}m) = \{z | W_z \text{ is } \hat{\mathrm{P}}\text{-}m\text{-mitotic}\}$ are Σ_3-complete.

References

[1] H. Rogers, *Theory of Recursive Functions and Effective Computability*, McGraw-Hill, 1967.

[2] R.I. Soare, *Recursively Enumerable Sets and Degree: A study of computable functions and computably generated sets*, Perspectives in Mathematical Logic, Springer-Verlag, 1987.

[3] R.M. Karp, "Reducibility among combinatorial problems" , *in Complexity of Computer Computations, R.E. Miller and J.M. Thatcher, Eds, Plenum*, New York, pp. 85-103, 1972

[4] S. A. Cook, "The complexity of theorem proving procedures," *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, pp. 151-158, 1971.

[5] R.E. Ladner, "On the Structure of Polynomial Time Reducibility," *Journal of the Association for Computing Machinery*, vol. 22, no. 1, pp. 155-171, 1975.

[6] K. Ambos-Spies, *Part of the book series: Lecture Notes in Computer Science: P-mitotic sets*, Logic and Machines: Decision Problems and Complexity Proceedings of the Symposium on Recursive Combinatorics, vol. 171**,** pp. 1-23, 1983.

[7] J. E. Hopcroft, J. D. Ullman, *Introduction to Automata theory, Languages and Computation*, Addison-Wesley Publishing Company, 1979.

[8] M. Sipser, *Introduction to the Theory of Computation*, PWS, Boston, MA, 1996.

[9] S. Arora and B. Barak, *Computational Complexity*, A Modern Approach, Cambridge University Press, 2009.

[10] S. A. Terwijn, *Complexity Theory,* Nijmegen, the Netherlands, 2010.

[11] B.L. van der Waerden, *Algebra*, Springer, vol. 1, 2003 (vol. 1 is translated from the German Algebra I, seventh edition, Springer-Verlag Berlin, 1966).

[12] A.H. Mokatsian, "Polynomial Time Turing Mitoticity and Arithmetical Hierarchy", *Pattern Recognition and Image Analysis*, Pleiades Publishing , vol. 34, no. 1, pp. 9–19. 2024.

[13] R.G. Downey and M. Stob, "Splitting Theorems In Recursion Theory," *Ann. Pure Appl. Logic*, vol. 65, pp. 1-106, 1993.

[14] A.H. Lachlan, "The priority method. I," *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, vol. 13, pp. 1–10, 1967.

[15] B. Trakhtenbrot, "On autoreducibility," *Dokl. Akad. Nauk SSSR*, vol. 192, pp. 1224–1227, 1970 (*in Russian*).

[16] R.E. Ladner, "Mitotic recursively enumerable sets," *J. Symb. Log*., vol. 38, pp. 199–211, 1973.

[17] R.G. Downey and T.A. Slaman, "Completely mitotic r.e. degrees," *Ann. Pure Appl. Logic*, vol. 41, no.2, pp. 119–152, 1989.

P-m-միթոտիկ բազմություններ և թվաբանական աստիճանակարգ

Արսեն Հ. Մոկացյան[1] և Խաչատուր Ա. Բարսեղյան[2]

[1] ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտ, Երևան, Հայաստան
[2] Սիմենս Ինդաստրի Սոֆթվեր, Երևան, Հայաստան

e-mail: arsenmokatsian@gmail.com, khachatur.barseghyan@outlook.com

Ամփոփում

Դիցուք $\{0,1\}^*$-ը $\{0,1\}$ բազմության տարրերից կազմված բոլոր վերջավոր շղթաների բազմություն է և **P**-ն այնպիսի *հիմնախնդիրների* դաս է, որոնք ճանաչելի են դետերմինիստական Թյուրինգյան մեքենաների միջոցով, որոնց աշխատանքի ժամանակը բազանդամորեն է կախված մուտքային տվյալների չափից (*հիմնախնդիրը* փաստորեն $\{0,1\}^*$ բազմության ենթաբազմություն է)։

Սույն հոդվածում սահմանված է $\widehat{\mathbf{P}}$ դասը և ցույց է տրված, որ $\widehat{\mathbf{P}}$-ն իզոմորֆ է **P** դասին։

Ելնելով *T*-միթոտիկություն և *T*-ինքնահանգեցում հասկացություններից Կ.Ամբոս-Սպիսը ներմուծել է *P-m*-միթոտիկություն և *P-m*-ինքնահանգեցում հասկացությունները։

Համանմանորեն ներմուծվել են $\widehat{\mathrm{P}}$-*m*-միթոտիկություն և $\widehat{P}$-*m*-ինքնահանգեցում հասկացությունները։

Տվյալ հոդվածում ապացուցված է, որ $\{z \mid W_z$-ն $\widehat{\mathrm{P}}$-*m*-միթոտիկ է$\}$ և $\{z \mid W_z$-ն $\widehat{\mathrm{P}}$-*m*-ինքնահանգեցվող է$\}$ ինդեքսային բազմությունները Σ_3-լրիվ են.

Բանալի բառեր՝ Թվաբանական աստիճանակարգ, *P-m*-միթոտիկ բազմություն, *P-m*- ինքնահանգեցվող բազմություն, ինդեքսային բազմություն։

P-m-митотические множества и арифметическая иерархия

Арсен А. Мокацян[1] и Хачатур А. Барсегян[2]

[1] Институт проблем информатики и автоматизации НАН РА, Ереван, Армения
[2] Сименс Индастри Софтвер, Ереван, Армения

e-mail: arsenmokatsian@gmail.com, khachatur.barseghyan@outlook.com

Аннотация

Пустъ $\{0,1\}^*$ является множеством всех конечных цепочек, составленных из элементов множества $\{0,1\}$ и **P** является классом *проблем*, распознаваемых детерминированными машинами Тьюринга, время работы которых полиномиально зависит от размера входных данных (*проблема* фактически является подмножеством множества $\{0,1\}^*$).

В данной статье определен класс $\widehat{\mathbf{P}}$ и показано, что $\widehat{\mathbf{P}}$ изоморфен классу **P**.

Исходя из понятий T- митотичности и T-автосводимости К. Амбос-Спис ввел понятия P-m- митотичности и P-m- автосводимости.

По аналогии с упомянутыми понятиями введены понятия $\widehat{\mathrm{P}}$-m-митотичности и $\widehat{\mathrm{P}}$-m-автосводимости.

В данной статье доказано, что индексные множества, $\{z \mid W_z -$ $\widehat{\mathrm{P}}$-m-митотично$\}$ и $\{\mathrm{z} \mid W_z - \widehat{\mathrm{P}}$-$m$-автосводимо$\}$ являются Σ_3-полными.

Ключевые слова` Арифметическая иерархия, P-m-митотическое множество, P-m-автосводимое множество, индексное множество.

Mathematical Problems of Computer Science 61, 62–69, 2024.

doi: 10.51408/1963-0115

UDC 004.4

Performance of Linear Algebra Factorization in Multi-Accelerator Architectures

Edita E. Gichunts

Institute for Informatics and Automation Problems of NAS RA, Yerevan, Armenia
e-mail: editagich@iiap.sci.am

Abstract

Hardware and software are required to effectively solve problems in many domains. The idea of creating a hybrid architecture based on graphics processors arose to meet the increasing demands of modern scientific problems. Most of these problems are reduced to solving linear algebra problems. A set of efficient linear solutions has been successfully used to solve important scientific problems for many years. Factorizations play a crucial role in solving linear algebra problems.

This work presents implementations of LU, QR and Cholesky factorizations on two graphics processors using the MAGMA 2.6.0 library. Their performances are given for matrices with real and complex numbers in single and double precision.

Keywords: MAGMA, multiple GPU, Linear Algebra, Factorizations.

Article info: Received 29 February 2024; sent for review 19 March 2024; accepted 16 May 2024.

1. Introduction

Many of the most important scientific programs rely on high-performance algorithms and linear algebra technologies, highlighting their importance and widespread impact from national security to medical breakthroughs. In the current high-performance computing (HPC) environment, parallelization is crucial. With the increasing utilization of video cards worldwide, the development of GPU parallel computing is expected to greatly affect the field of high-performance computing. These possibilities have already generated a great deal of interest in scientific as well as non-scientific circles. After all, the acceleration potential of good parallelization of algorithms is not always tens of times faster. The trend in multi-computing is clearly moving towards parallel algorithms, with most new solutions and initiatives focused in this direction. The current generation of GPUs has a fairly flexible architecture, which, along with high-level programming

languages and hardware-software architectures, reveals these capabilities and makes them more accessible. In the field of high-performance computing, the hottest topic is GPU-based hybrid systems. Hybrid architecture combines the advantages of shared and distributed memory systems. In such architectures, the GPU device is used as a coprocessor or accelerator to handle multi-computation applications. Calculations on GPU are known for being developed and processed very quickly. One of the leading video chip manufacturers, Nvidia, has introduced the CUDA [1] (Compute Unified Device Architecture) platform. CUDA is both a software and hardware technology available to every developer. It is an extension of the C programming language. The only requirement is the use of different programming paradigms typical for parallel computing.

Linear algebra faces a significant challenge in terms of computational efficiency, which has led to the development of software libraries following advancements in computer architecture. In the mid-1960s, IBM released the Scientific Subroutine Package [2], a collection of FORTRAN subroutines optimized for the IBM System/360 machine. In 1974, Harwood released EISPACK [3], a package of FORTRAN routines that compute the eigenvalues and eigenvectors of a matrix. BLAS's basic linear algebra routines were the first product of a joint project with ACMSIGNUM during 1973–1977 [4], which was based on a proposal made in 1973 [5].The LINPACK library was introduced in 1979 as a collection of routines for solving linear equations and linear least squares problems on supercomputers of the 1970s and 1980s, mostly based on vector processors [6].

LINPACK used the partial rotation engine of LU analysis to solve 100-dimensional problems, allowing the user to evaluate the performance of their memory and processors. The first version of BLAS (BLAS Level 1) implemented scalar-vector and vector-vector operations. BLAS2 (BLAS Level 2) was developed in 1988 as an extension of BLAS1 to exploit the capabilities of vector processors [7, 8]. BLAS2 offers the ability to perform matrix-vector operations. LAPACK [9], released in 1992, replaces LINPACK and EISPACK and provides better performance. LAPACK specializes in solving systems of linear equations, linear least squares problems, eigenvalue problems, and singularity problems. To perform these operations, related calculations are also performed, such as matrix analysis: LU, QR (Q-matrix is unitary or Hermitian, and R is upper trapezoidal), Cholesky, etc.

For GPUs, NVIDIA offers CuBLAS [10], an implementation of BLAS in the NVIDIA CUDA and EM Photonics environments, as well as their CULA solutions [11] as implementations of LAPACK CUDA.

MAGMA [12] is an extension of LAPACK in a hybrid framework. It includes an amazing variety of subroutines for solving linear algebra problems.

MAGMA's research is based on the idea that optimal software solutions for solving complex problems in a hybrid environment should be self-hybridizing, combining the strengths of various algorithms within a single framework. Based on this idea, efforts are being made to develop algorithms for hybrid multi-core and graphics systems. Designed with LAPACK functionality, data storage and interface capabilities, the MAGMA library makes it easy for scientists to port their software components from LAPACK to MAGMA and take advantage of the new hybrid architecture.

LU, QR and Cholesky factorizations play an important role in linear algebra. LU factorization is applied to the problem of finding solutions to a system of linear equations. The first step is to perform the LU factorization of the matrix, and then solutions can be obtained. It is worth mentioning that the paper referenced as [13] presents solutions to a system of linear equations using the types of LU factorization and random butterfly transformation implemented with the MAGMA library on a single graphics processor.

QR factorization is often used to solve the linear least squares (LLS) problem. It is also the basis of the QR algorithm [14,15] for finding the eigenvalue problem.

Cholesky factorization is useful for efficient numerical solutions such as Monte Carlo simulations.
This paper presents implementations of LU, QR, and Cholesky factorizations of widely used linear algebra problems on multiple graphics processor architectures, using the MAGMA 2.6.0 library. Performances of the specified factorizations for matrices with real and complex numbers in both single and double precision are presented.

2. Stages of Implementing Factorizations with Multiple Accelerators

It was observed that when dealing with multiple accelerators, LU, QR, and Cholesky factorizations were implemented by using the MAGMA 2.6.0 library. We have provided descriptions of these factorization subroutines. It should be noted that instead of using types, the letter x was used, which in the case of real numbers is s, and it is d for single and double precision, respectively, while in the case of complex numbers, c and z are used.
*magma_xgetrf_mgpu(ngpu, M, N, d_A, ldda, ipiv, &info)*computes an LU factorization of a general M-by-N matrix A using partial pivoting with row interchanges.
The factorization has the form

A = P * L * U,

where P is a permutation matrix, L is a lower triangular with unit diagonal elements (lower trapezoidal if *M>N*), and U is an upper triangular (upper trapezoidal if *M<N*).

*magma_xgeqrf2_mgpu(ngpu, M, N, d_A, ldda, tau, &info)*computes a QR factorization of an M-by-N matrix A.
The factorization has the form

A = Q * R.

*magma_xpotrf_mgpu(ngpu, uplo, N, d_A, ldda, &info)*computes the Cholesky factorization of a real symmetric and complex Hermitian positive definite matrix A.
The factorization has the form

A = U**H * U, if UPLO = MagmaUpper, or
A = L * L**H, if UPLO = MagmaLower,

where U is an upper triangular matrix and L is a lower triangular.
uplo= MagmaUpper: Upper triangle of A is stored,
uplo= MagmaLower: Lower triangle of A is stored.

Here are the stages of implementing Cholesky factorization when using multiple accelerators:

1. First, the MAGMA library is initialized using the magma_init() function.
2. Memory is allocated for the matrix on the CPU using the function magma_xmalloc_cpu(&h_A, lda*N). Memory is also allocated for the matrix copy on the CPU using magma_xmalloc_pinned(&h_R, lda*N).
3. To allocate memory for the matrix on GPUs, we cycle from GPU to GPU, and in each of them, the function magma_setdevice(dev) is first called, then the memory is allocated using the function magma_xmalloc(&d_A[dev], max_size), where max_size = (1 +N/(nb*ngpu))*nb * magma_roundup(N, nb) and nb= magma_get_dpotrf_nb(N).
4. The matrix is generated using the function magma_generate_matrix(opts, N, N, h_A, lda).

5. We copy the matrix using the lapackf77_xlacpy(MagmaFullStr, &N, &N, h_A, &lda, h_R, &lda) function, which will be sent to the GPU memory.
6. The function magma_xsetmatrix_1D_col_bcycle(N, N, h_R, lda, d_A, ldda, ngpu, nb) transfers the matrix to the GPU memory.
7. We fix the time using the function gpu_time = magma_wtime().
8. The function magma_xpotrf_mgpu(ngpu, uplo, N, d_A, ldda, &info) is called, which performs Cholesky factorization in parallel on GPUs.
9. Using the difference gpu_time=magma_wtime()-gpu_time, we obtain the calculation execution time.
10. After the calculations are completed, the function magma_xgetmatrix_1D_col_bcycl (N, N, d_A, ldda, h_R, lda, ngpu, nb) transfers the results from the GPUs to the CPU memory.
11. We clear the allocated memories on the CPU using the functions magma_free_cpu(h_A) and magma_free_pinned(h_R), and clear the allocated memories on the GPUs by performing a cycle transfer from GPU to GPU, first calling magma_setdevice(dev) and then magma_free(d_A[dev]) functions.
12. At the end of the program, we use magma_finalize() to terminate MAGMA.

3. Experimental Results

Tests were conducted on two NVIDIA Tesla V100-PCIE graphics processors. The cuda-10.2 platform was utilized for parallel computing. To install the MAGMA 2.6.0 library, the BLAS, LaPack, cLaPack and ATLAS libraries were installed. To install the MAGMA library, the gcc, g++, nvcc, and gfortran compilers were used. To compile MAGMA, the following static (.a) and dynamic (.so) libraries are also required: libgfortran.a, libf77blas.a, libcblas.a, libf2c.a, libm.a, libstdc++.a, libpthread.a, libdl. .a, libcublas.so, libcudart.so, libcusparse.so, libcudadevrt.a.

Let us present the results of experiments in the form of graphs.

Figures 1 and 2 display graphs of LU factorization for matrices with real and complex numbers in single and double precision, respectively.

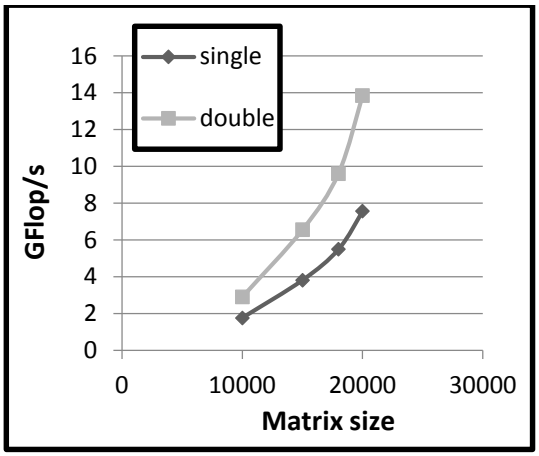


Fig.1. LU real

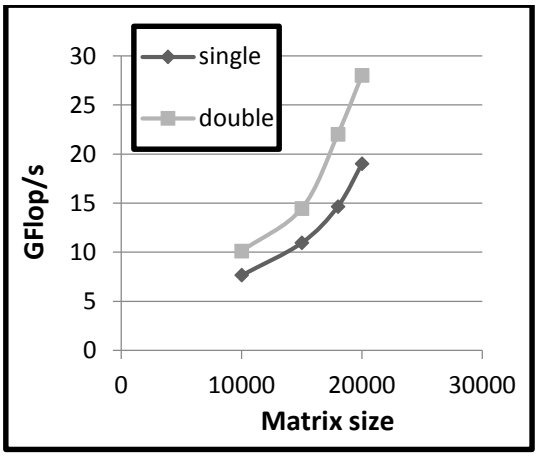


Fig. 2. LU complex

Figures 3 and 4 display graphs of QR factorization for matrices with real and complex numbers in single and double precision, respectively.

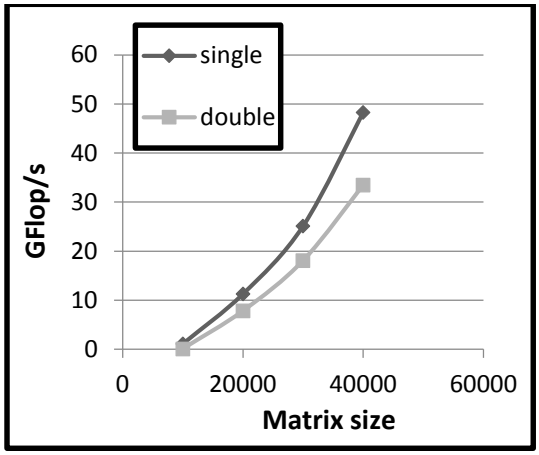


Fig. 3. QR real

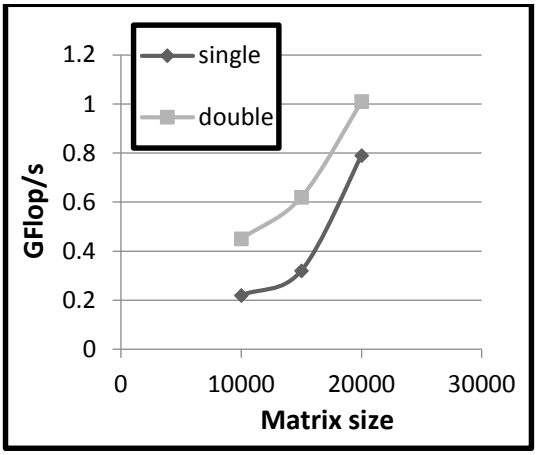


Fig. 4. QR complex

Figures 5 and 6 display Cholesky factorization graphs for matrices with real and complex numbers in single and double precision, respectively.

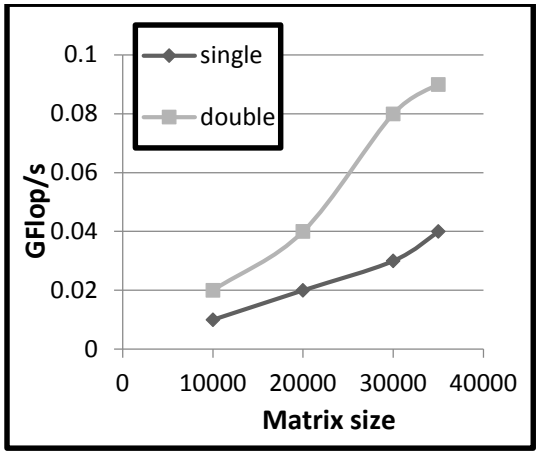


Fig. 5. Cholesky real

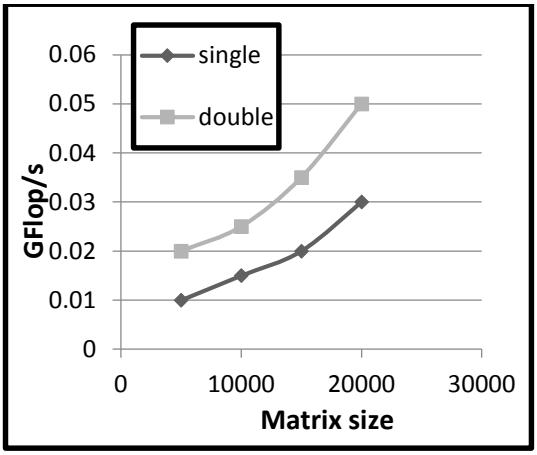


Fig. 6. Cholesky complex

4. Conclusion

We have reached the following conclusions based on the results of our experiments:

For matrices with real numbers, single precision performance in the case of LU factorization is 2 times lower than double precision. For matrices with complex numbers, single precision performance is 1.5 times lower than double precision.

In the case of QR factorization, single-precision performance for matrices with real numbers is 1.5 times higher than binary precision, and for complex numbers, single precision performance is 2 times lower than double precision.

In the case of Cholesky factorization and for matrices with real and complex numbers, single precision performance is 2 times lower than double precision performance.

References

[1] NVIDIA, “NVIDIA CUDA Parallel Computing Platform”. http://www.nvidia.com/object/cuda_home_new.html, NVIDIA, 2013.

[2] International Business Machines Corporation. System/360 Scientific Subroutine Package (360A-CM-03X) Version II, Programmer’s Manual. IBM Technical Publications Department, White Plains, NY, 1967.

[3] B. S. Garbow. EISPACK-a package of matrix eigensystem routines. Computer Physics Communications, 7(4):179–184, 1974.

[4] C. L. Lawson, R. J. Hanson, D. R. Kincaid, and F. T. Krogh. Basic Linear Algebra Subprograms for Fortran Usage. ACM Trans. Math. Softw., 5(3):308–323, September 1979.

[5] R. J. Hanson, F. T. Krogh, and C. L. Lawson. A proposal for standard linear algebra subprograms. ACM Signum Newsletter, 1973.

[6] J. Dongarra, C. B. Moler, J. R. Bunch, and G. W. Stewart. LINPACK Users’ Guide, volume 8. SIAM, 1979.

[7] J. Dongarra, J. Du Croz, S. Hammarling, and R. J. Hanson. Algorithm 656: an extended set of basic linear algebra subprograms: model implementation and test programs. ACM Transactions on Mathematical Software (TOMS), 14(1):18–32, 1988.

[8] Dongarra, J. Du Croz, S. Hammarling, and R. J. Hanson. An Extended Set of FORTRAN Basic Linear Algebra Subprograms. ACM Trans. Math. Softw., 14(1):1–17, March 1988.

[9] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. LAPACK Users’ Guide. Society for Industrial and Applied Mathematics, Philadelphia, PA, third edition, 1999.

[10] CUDA Nvidia. Cublas library. NVIDIA Corporation, Santa Clara, California, 15, 2008.

[11] J. R. Humphrey, D. K. Price, K. E. Spagnoli, A. L. Paolini, and E. J. Kelmelis. CULA: hybrid GPU accelerated linear algebra routines. In SPIE Defense, Security, and Sensing, pages 770502–770502. International Society for Optics and Photonics, 2010.

[12] “MAGMA Matrix Algebra on GPU and Multicore Architectures”, http://icl.cs.utk.edu/magma/, 2014.

[13] H. V. Astsatryan, E. E. Gichunts, “Performances of Methods for Solving a Linear System of Equations in the Architecture of GPU Accelerator”, Transactions of IIAP NAS RA, Mathematical Problems of Computer Science, vol. 45, pp. 44—52, 2016.

[14] B. N. Parlett, *The Symmetric Eigenvalue Problem. Englewood Cliffs*, NJ: Prentice-Hall, 1980.

[15] G. H.Golub, C. F. V.Loan, *Matrix Computations*, 3rd ed. Baltimore: The Johns Hopkins University Press, 1996.

Գծային հանրահաշվի ֆակտորիզացիաների արտադրողականությունները բազմաքանակ արագացուցիչների ճարտարապետությունում

Էդիտա Ե. Գիչունց
ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտ, Երևան, Հայաստան
e-mail: editagich@iiap.sci.am

Ամփոփում

Բազմաթիվ բնագավառների խնդիրների արդյունավետ լուծումների համար պահանջվում է ապարատային և ծրագրային ապահովում: Հիբրիդային ճարտարապետության ստեղծման գաղափարը, որը հիմնված է գրաֆիկական պրոցեսորների վրա, ծագել է ժամանակակից գիտական հիմնախնդիրների աճող պահանջների բավարարման պատճառով: Այդ խնդիրների մեծ մասը բերվում է գծային հանրահաշվի խնդիրների լուծումներին: Արդյունավետ գծային լուծումների հավաքածուն շատ երկար տարիներ կարողանում է լուծել կարևորագույն գիտական խնդիրներ: Գծային հանրահաշվի խնդիրների լուծումներում կարևորագույն դերակատարում ունեն ֆակտորիզացիաները:
Այս աշխատանքում ներկայացված են LU, QR և Խոլեցկու (Cholesky) ֆակտորիզացիաների իրականացումները երկու գրաֆիկական պրոցեսորների վրա՝MAGMA 2.6.0 գրադարանի կիրառմամբ: Տրված են նրանց արտադրողականությունները իրական և կոմպլեքս թվերով մատրիցների համար՝մեկական և երկուական ճշգրտություններում:

Բանալի բառեր՝ MAGMA, բազմակի GPU, գծային հանրահաշիվ, ֆակտորիզացիա:

Производительность факторизации линейной алгебры в мультиускорительных архитектурах

Эдита Е. Гичунц

Институт проблем информатики и автоматизации НАН РА, Ереван, Армения
[2] Сименс Индастри Софтвер, Ереван, Армения
e-mail: editagich@iiap.sci.am

Аннотация

Для эффективного решения проблем во многих областях требуется аппаратное и программное обеспечение. Идея создания гибридной архитектуры на базе графических процессоров возникла для удовлетворения растущих требований современных научных задач. Большинство этих задач сводятся к решению задач линейной алгебры. Набор эффективных линейных решений уже много лет успешно используется для решения

важных научных задач. Факторизации играют решающую роль в решении задач линейной алгебры.

В этой работе представлены реализации факторизации LU, QR и Холецкого на двух графических процессорах с использованием библиотеки MAGMA 2.6.0. Приведены их производительности для матриц с действительными и комплексными числами одинарной и двоичной точности.

Ключевые слова` МАГМА, GPU, линейная алгебра, факторизация.

Կանոններ հեղինակների համար

ՀՀ ԳԱԱ ԻԱՊԻ “Կոմպյուտերային գիտության մաթեմատիկական խնդիրներ” պարբերականը տպագրվում է 1963 թվականից: Պարբերականում հրատարակվում են նշված ոլորտին առնչվող գիտական հոդվածներ, որոնք պարունակում են նոր՝ չհրատարակված արդյունքներ:

Հոդվածները ներկայացվում են անգլերեն՝ ձևավորված համապատասխան “ոճով” (style): Հոդվածի ձևավորման պահանջներին ավելի մանրամասն կարելի է ծանոթանալ պարբերականի կայքէջում՝ http://mpcs.sci.am/:

Rules for authors

The periodical “Mathematical Problems of Computer Science” of IIAP NAS RA has been published since 1963. Scientific articles related to the noted fields with novel and previously unpublished results are published in the periodical.

Papers should be submitted in English and prepared in the appropriate style. For more information, please visit the periodical's website at http://mpcs.sci.am/.

Правила для авторов

Журнал «Математические проблемы компьютерных наук» ИПИА НАН РА издается с 1963 года. В журнале публикуются научные статьи в указанной области, содержащие новые и ранее не опубликованные результаты.

Статьи представляются на английском языке и оформляются в соответствующем стиле. Дополнительную информацию можно получить на веб-сайте журнала: http://mpcs.sci.am/.

The electronic version of the periodical “Mathematical Problems of Computer Science” and rules for authors are available at

http://mpcs.sci.am/

Phone: (+37460) 62-35-51
Fax: (+37410) 28-20-50
E-mail: mpcs@sci.am
Website: http://mpcs.sci.am/

Ստորագրված է տպագրության՝ 27.05.2024
Թուղթը՝ օֆսեթ:
Հրատարակված է ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտի կողմից
Ծավալը՝ 71 էջ: Տպաքանակը՝ 100
ՀՀ ԳԱԱ ԻԱՊԻ Համակարգչային պոլիգրաֆիայի լաբորատորիա
Երևան, Պ. Սևակի 1
Հեռ. +(374 60) 623553
Գինը՝ անվճար

Подписано в печать 27.05.2024
Офсетная бумага.
Опубликовано Институтом проблем информатики и автоматизации НАН РА
Объём: 71 страниц. Тираж: 100
Лаборатория компьютерной полиграфии ИПИА НАН РА.
Ереван, П. Севака 1
Тел.: +(374 60) 623553
Цена: бесплатно

Signed in print 27.05.2024
Offset paper
Published by the Institute for Informatics and Automation Problems of NAS RA
Volume: 71 pages
Circulation: 100
Computer Printing Lab of IIAP NAS RA
Yerevan, 1, P. Sevak str.
Phone: +(374 60) 623553
Free of charge