ICETI 2025

BOOK OF PROCEEDINGS

October 22-26 2025

9TH INTERNATIONAL CONFERENCE ON ENGINEERING TECHNOLOGY AND INNOVATION

WWW.ICETI.ORG

9TH INTERNATIONAL CONFERENCE ON ENGINEERING TECHNOLOGY AND INNOVATION

ISSN: 2687-2323

PROCEEDINGS OF THE 9TH INTERNATIONAL CONFERENCE ON ENGINEERING TECHNOLOGY AND INNOVATION (ICETI) OCTOBER 22-26, 2025 SARAJEVO, BOSNIA AND HERZEGOVINA

Edited by

Prof. Dr. Özer Çınar

Published, 2025

info@iceti.org
www.iceti.org

info@iceti.org

SCIENTIFIC COMMITTEE

1. Prof. Dr. Adisa Parić - University of Sarajevo - Bosnia and Herzegovina
2. Prof. Dr. Aleksandar Dimitrov - Ss. Cyril and Methodius University - Macedonia
3. Prof. Dr. Anita Grozdanov - Ss. Cyril and Methodius University - Macedonia
4. Prof. Dr. Asif Šabanović – International University of Sarajevo - Bosnia and Herzegovina
5. Prof. Dr. Christos Douligeris - University of Erlangen-Nurnberg - Germany
6. Prof. Dr. Dragutin T. Mihailović - University of Novi Sad - Serbia
7. Prof. Dr. Erkan Şahinkaya – İstanbul Medeniyet University - Turkey
8. Prof. Dr. Falko Dressler - University of Paderborn - Germany
9. Prof. Dr. Houssam Toutanji – Western Michigan University - USA
10. Prof. Dr. Ian F. Akyıldız – Georgia Institute of Technology - USA
11. Prof. Dr. İsmail Usta - Marmara University - Turkey
12. Prof. Dr. Liljana Gavrilovska - Ss Cyril and Methodius University - Macedonia
13. Prof. Dr. Lukman Thalib - Qatar University - Qatar
14. Prof. Dr. M. Asghar Fazel – University of Environment - Iran
15. Prof. Dr. Mehmet Akalin - Marmara University - Turkey
16. Prof. Dr. Mehmet Kitiş – Süleyman Demirel University - Turkey
17. Prof. Dr. Muammer Koç - Hamad bin Khalifa University - Qatar
18. Prof. Dr. Özer Çınar – Yıldız Technical University - Turkey
19. Prof. Dr. Perica Paunovik - Ss. Cyril and Methodius University - Macedonia
20. Prof. Dr. Rifat Škrijelj – University of Sarajevo - Bosnia and Herzegovina
21. Prof. Dr. Samir Đug - University of Sarajevo - Bosnia and Herzegovina
22. Prof. Dr. Tanju Karanfil – Clemson University - USA
23. Prof. Dr. Tibor Biro - National University of Public Service, Budapest - Hungary
24. Prof. Dr. Ümit Alver – Karadeniz Technical University - Turkey
25. Prof. Dr. Wolfgang Gerstacker - University of Erlangen-Nurnberg - Germany
26. Prof. Dr. Yılmaz Yıldırım - Bülent Ecevit University - Turkey
27. Prof. Dr. Yousef Haik - Hamad bin Khalifa University - Qatar
28. Assoc. Prof. Dr. Alaa Al Hawari - Qatar University - Qatar
29. Assoc. Prof. Dr. Izudin Dzafic - International University of Sarajevo - Bosnia and Herzegovina
30. Assoc. Prof. Dr. Muhamed Hadziabdic - International University of Sarajevo - Bosnia and Herzegovina
31. Assoc. Prof. Dr. Nusret Drešković - University of Sarajevo - Bosnia and Herzegovina
32. Assist. Prof. Dr. Faruk Berat Akçeşme - University of Health Sciences - Turkey
33. Assist. Prof. Dr. Fouzi Tabet - German Biomass Research Center - Germany
34. Assist. Prof. Dr. Haris Gavranovic - International University of Sarajevo - Bosnia and Herzegovina
35. Assist. Prof. Dr. Murat Karakaya - Atılım University - Turkey
36. Assist. Prof. Dr. Sasan Rabieh - Shahid Beheshti University - Iran
37. Assist. Prof. Dr. Ševkija Okerić - University of Sarajevo - Bosnia and Herzegovina
38. Assist. Prof. Dr. Ivana Plazonić - University of Zagreb - Croatia
39. Assist. Prof. Dr. J. Amudhavel - VIT Bhopal University - India
40. Assist. Prof. Dr. Hasan Bora Usluer - Galatasaray University - Turkey
41. Dr. Muhammet Uzun - RWTH Aachen University - Germany
42. Dr. Zsolt Hetesi - National University of Public Service, Budapest - Hungary
43. Dr. Zsolt T. Németh - National University of Public Service, Budapest - Hungary

ORGANIZATION COMMITTEE

Chairman of the Conference

Prof. Dr. Özer Çınar – Yıldız Technical University - Turkey

Members of the Committee

Prof. Dr. M. Asghar Fazel – University of Environment - Iran

Prof. Dr. Ümit Alver - Karadeniz Technial University - Turkey

Assoc.Prof.Dr. Agim Mamuti - Mother Teresa University - North Macedonia

Assoc. Prof. Dr. Lukman Thalib - Qatar University - Qatar

Assoc. Prof. Dr. Nusret Drešković - University of Sarajevo - Bosnia and Herzegovina

Assist. Prof. Dr. Sasan Rabieh - Shahid Beheshti University - Iran

Alma Ligata - Zenith Group - Bosnia and Herzegovina

Ismet Uzun - Zenith Group - Bosnia and Herzegovina

Musa Kose - Zenith Group - Bosnia and Herzegovina

WELCOME TO ICETI 2025

On behalf of the organizing committee, we are pleased to announce that the International Conference On Engineering Technology And Innovation is held on October 22-26, 2025 in Sarajevo, Bosnia and Herzegovina. ICETI 2025 provides an ideal academic platform for researchers to present the latest research findings and describe emerging technologies, and directions in Engineering Technology And Innovation. The conference seeks to contribute to presenting novel research results in all aspects of Engineering Technology And Innovation. The conference aims to bring together leading academic scientists, researchers and research scholars to exchange and share their experiences and research results about all aspects of Engineering Technology And Innovation. It also provides the premier interdisciplinary forum for scientists, engineers, and practitioners to present their latest research results, ideas, developments, and applications in al lareas of Engineering Technology And Innovation. The conference will bring together leading academic scientists, researchers and scholars in the domain of interest from around the world. ICETI 2025 is the oncoming event of the successful conference series focusing on Engineering Technology And Innovation. The International Conference on Engineering Technology and Innovation (ICETI 2025) aims to bring together leading academic scientists, researchers and research scholars to exchange and share their experiences and research results about all aspects of Engineering Technology and Innovation. It also provides the premier interdisciplinary forum for scientists, engineers, and practitioners to present their latest research results, ideas, developments, and applications in all areas of Engineering Technology and Innovation.The conference will bring together leading academic scientists, researchers and scholars in the domain of interest from around the world. The conference's goals are to providea scientific forum for all international prestige scholars around the world and enable the interactive exchange of state-of-the-art knowledge. The conference will focus on evidence-based benefits proven in technology and innovation and engineering experiments.

Best regards,

Prof. Dr. Özer ÇINAR

Mechanical Stability of Different Black Inks Printed on Packaging Paper With a Metallised Surface

Maja Rudolf[1], Irena Bates[1], Ivana Plazonić[1], Dino Priselac[1]*

[1]University of Zagreb, Faculty of Graphic Arts

*maja.rudolf@grf.unizg.hr

Abstract:

Today in the graphic industry we are witnessing the rapid growth of the packaging sector, with the consumption of paper and printing inks increasing every year. Therefore, it is important to study the interaction between the printing substrates (paper and cardboard) and the applied inks in order to ensure the quality of the packaging prints and the adequate protection of the wrapped products. The goal of this study is to evaluate the mechanical stability (i.e., rubbing resistance) of printed paper-based packaging materials, which is crucial for the quality of the printed packaging. In the study, metallised paper was used as a printing substrate, while printing was carried out using four different black inks (three conventional and one UV ink) by offset printing technique. All metallized printing papers are coated with a white base ink in one or two layers before being printed with black ink. The mechanical stability of all samples was measured as a colour difference (ΔE2000) before and after the performed rubbing test. The test was done in three iterations of the rubbing cycle under two different weights, simulating the conditions of the packaging during transport and general handling. The colorimetric values of CIE L*a*b* components were further analysed to determine if the colour difference was greater in lightness or in tone components of colour. These findings can help determine which combination of printing substrate and ink should be used, ensuring the optimal quality of product packaging.

Keywords: Black Ink, Colour Difference, Mechanical Stability, Offset Printing

FULL-TEXT ARTICLES

Enhancing Thermal Efficiency of Panel Radiators Through Turbulence-Inducing Modifications in Water Channels

Umut Uçak[1*], Çisil Timuralp[2]

Abstract

This study aims to improve the thermal efficiency of panel radiators, which are widely used in building heating systems. By promoting turbulent flow inside the water channels, heat transfer can be increased without raising energy use. In the modified design, small 90° rectangular blocks (5 mm wide and 10 mm long) are added inside the vertical channels to create turbulence. This design distinguishes itself from other methods by creating turbulence through indentations in the sheet metal that forms the channel, rather than using additional turbulators such as twisted tape inserts or wire coil inserts. This unique strategy not only boosts heat transfer performance but also leads to energy savings, reduced emissions, and more efficient heating in both residential and commercial buildings. With the new design, due to increased friction, the pressure rise shows a 1.36% increase compared to the current design. Meanwhile, the temperature of the water circulating in the panel radiator channels has decreased further compared to the current design, resulting in a 1.21% increase in ΔT. This indicates that the new design transfers more heat from the water to the environment, causing the water's outlet temperature to drop more significantly. Additionally, with the new design, the water contact surface area has decreased by 0.64 m², while the water heat flux has increased by approximately 25% compared to the current design. The higher heat flux achieved with a smaller water contact surface area demonstrates the efficiency of the system compared to the existing design.

Keywords: *Energy efficiency, panel radiator, turbulent flow, heat transfer enhancement, water channel*

1. INTRODUCTION

In the context of modern energy policies, energy efficiency is regarded as a critical parameter for both sustainable economic development and environmental Protection [1]. Enhancing the thermal efficiency of panel radiators, which are widely utilized in building heating systems, directly contributes to reducing energy demand in the residential and commercial sectors. High-efficiency radiator designs can deliver the required thermal output with lower energy input, thus minimizing fossil fuel consumption, decreasing operational costs, and improving national energy security by reducing dependency on external energy sources. Moreover, the reduction in energy consumption results in lower greenhouse gas emissions, thereby contributing to climate change mitigation and global sustainability goals [2]. From a thermal engineering perspective, the enhancement of convective heat transfer can be achieved through various flow and surface modification strategies. Among these, promoting turbulent flow is particularly effective. In contrast to laminar flow characterized by orderly fluid layers turbulent flow involves chaotic motion, eddies, and vortices, which reduce the thermal boundary layer thickness at the fluid–solid interface and significantly increase the convective heat transfer coefficient [3]. To this end, surface optimization through methods such as increasing roughness, implementing extended surfaces (fins), introducing swirl flow, and applying turbulence-enhancing geometrical modifications are frequently employed in heat transfer enhancement studies. The analysis results provide essential insights for radiator design optimization, offering practical contributions toward improving thermal performance, reducing energy consumption, and minimizing environmental impact in heating applications.

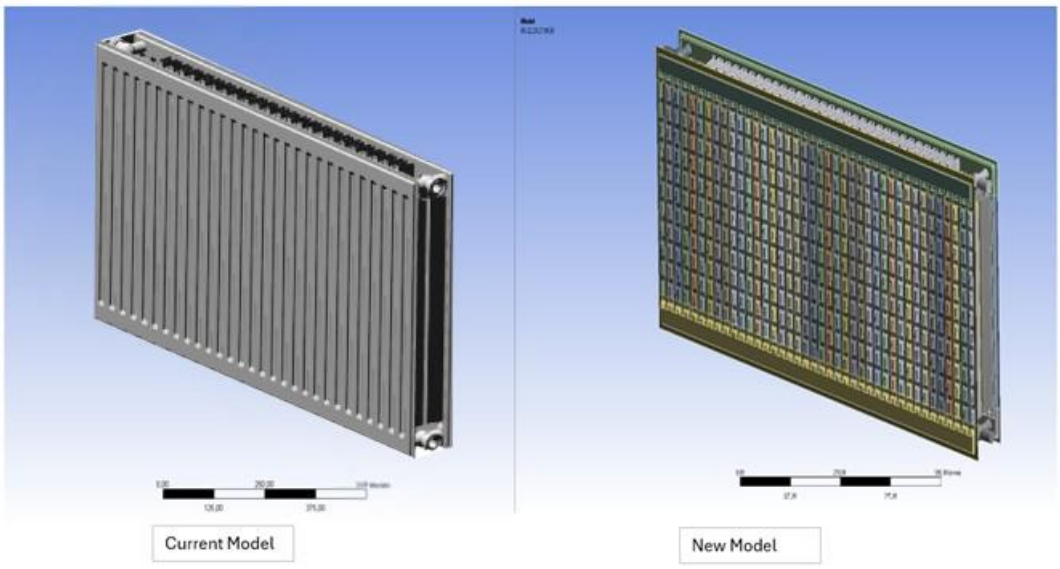
[1*] *Corresponding author: Türk Demirdöküm Factory Inc. R&D department, Bozüyük / Türkiye, umut.ucak@vaillant-group.com*

[2] *Eskisehir Osmangazi University, Faculty of Engineering, Department of Mechanical Engineering, Eskisehir / Turkey. cisil@ogu.edu.tr*

2. MATERIALS AND METHODS

As we look at the literature, we can see that various turbulent flows contribute positively to thermal efficiency. For example, in the study by Pour Razzaghi et al. numerical analyses on turbulent fluid flow and heat transfer in twisted flat tubes are presented. In the research, three different geometries for the cross-section of the flat tube were examined, and water with constant thermophysical properties was used as the working fluid. The Reynolds number ranges from 5000 to 20000. The results show that twisted flat tubes increase efficiency by nearly 70% compared to smooth tubes with the same surface area and hydraulic diameter. Additionally, the alternating helical direction increases efficiency by almost 60%, which is noted as a good choice. Due to the effect of helical pitch on the flow path, the alternating geometry performed better at some Reynolds numbers. These results are explained using flow characteristic contours such as turbulent kinetic energy, tangential velocity, and temperature. The results are presented in numbers and plots for a more accurate comparison [4]. In another study conducted by R. Yang and F. P. Chiang, experiments were conducted using water as the working medium, flowing through a varying-curvature curved pipe, forming a double-pipe heat exchanger. The heat transfer coefficients were determined using the Wilson plot method. The effects of Dean, Prandtl, Reynolds numbers, and curvature ratio on heat transfer and friction factors were examined. The results showed that a higher Dean number leads to a higher heat transfer rate. Compared to a straight pipe, the heat transfer rate can be increased by up to 100%, while the friction coefficient increased by less than 40%. These findings suggest that using S-shaped pipes instead of straight pipes could be advantageous for performance enhancement in heat exchangers, such as solar collectors [5].A separate investigation by A. Lanani and R. Benchabi explores the two-dimensional heat transfer characteristics within a corrugated channel., heat exchangers are vital in energy management across various industries, designed to maximize heat transfer through specialized surfaces. Techniques like corrugated surfaces enhance turbulence and heat transfer. A study using Fluent CFD code explores the impact of Reynolds numbers on flow and heat transfer, revealing that friction and Nusselt numbers are affected by Reynolds numbers and outperform those in smooth channels [6]. In different study conducted by Soorena Azarhazin et al. it was found that utilizing different pipe geometries, specifically flattened pipes, increases turbulence and unsteadiness, making them promising candidates for thermal applications due to enhanced heat transfer coefficients [7].

Within the scope of these studies, the aim is to improve the thermal performance of panel radiators by promoting turbulent flow through increased friction, achieved via specially designed geometries within the vertical water channels. Figure 1-a displays the reference radiator with straight vertical channels, whereas Figure 1-b illustrates the modified configuration featuring 90° rectangular constrictions (5 mm in width and 10 mm in length) integrated into the channels. Unlike other studies that use additional turbulators such as twisted tape inserts or wire coil inserts, this approach achieves turbulence through indentations made in the sheet metal forming the channel, offering a unique method for enhancing heat transfer efficiency.



(a) (b)

Figure 1. *(a) Radiator with existing water channel ,(b) Radiator with indented water channel [8]*

As shown below in Figure 2, the side view of the existing radiator with a straight water channel, the side view of the new model featuring indented channels, and the geometric details of the indentation form are presented. The cross-sectional view reveals the detailed structure of the indentation: The total width of the water channel remains at 12 mm.The stepped indentations are 5 mm deep on both sides, separated by a 10 mm vertical segment.The alternating indentations create a more complex internal surface compared to the flat wall in the original design.This geometry not only increases the wetted perimeter (and thus the surface area for thermal exchange) but may also improve flow mixing inside the channel, thereby improving the radiator's thermal performance.

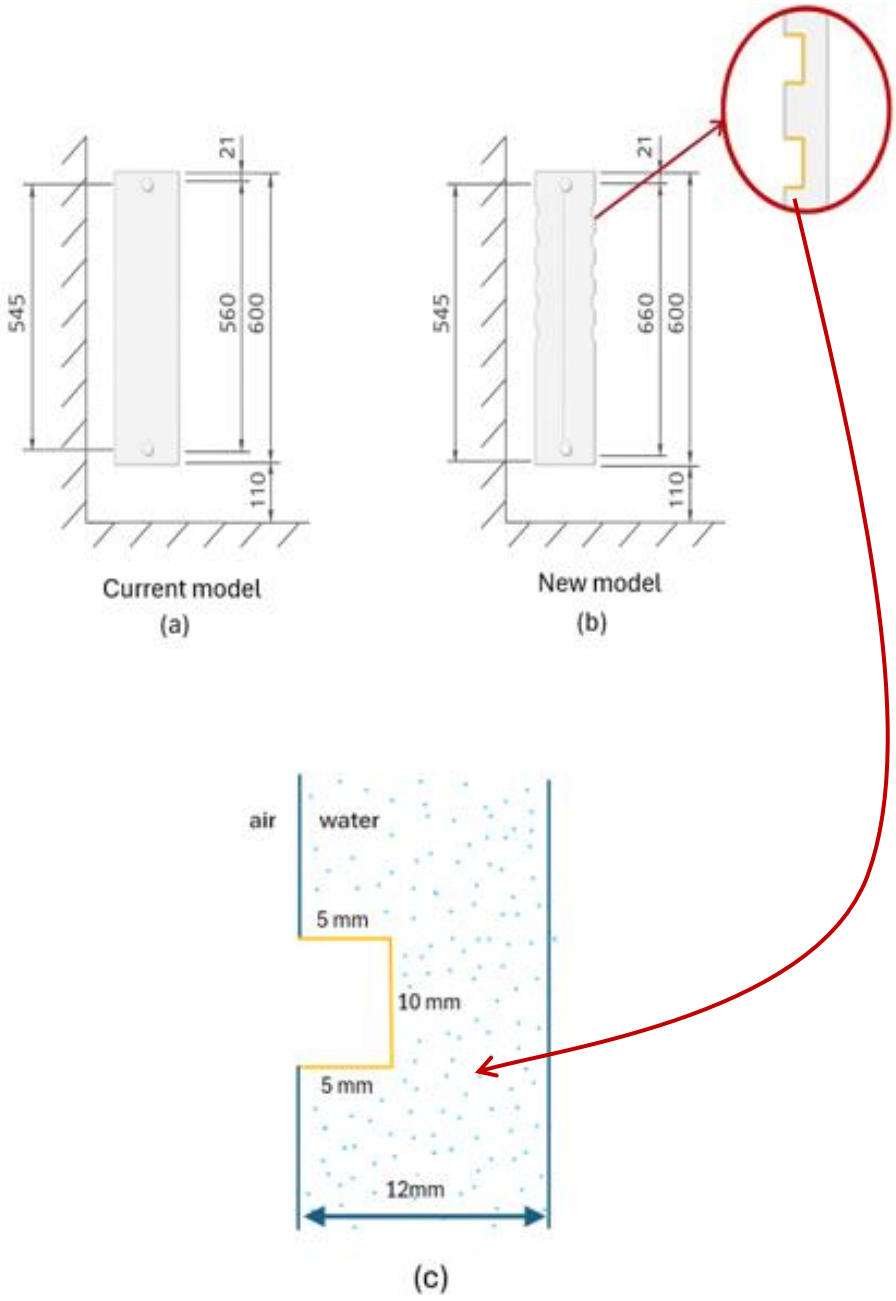


Figure 2. *(a) Side view of radiator with existing water channel, (b) Side view of radiator with indented water channel , (c) dimensions of the indentation form [9]*

Flow analyses were performed using Ansys Fluent 18.02 to assess the heat transfer efficiency of both designs. In the analysis settings, the viscous model for the current design is set to Laminar, while the new design employs the k-Ɛ Realizable model with Non-equilibrium Wall functions, offering a more sophisticated method for capturing turbulent flow characteristics. The k-Ɛ Realizable model was chosen for its ability to provide more accurate predictions of turbulent flow behavior, especially in complex geometries and high Reynolds number flows [10]. Furthermore, the material properties used in the analysis are detailed in Table 1.

***Table 1**. Properties of materials used during the Analysis phase*

Part	Material	Density (kg/m³)	Specific Heat (J/kg·K)	Thermal Conductivity (W/m·K)	Viscosity (kg/m·s)
Fluid	Water	977,78	4189,5	0,6663	0,000404
Panel	DC01	7870	505	5,4	
Convector	DC01	7870	505	5,4	
Elbow	DC01	7870	505	5,4	

In this analysis, the performance and heat transfer of a radiator were examined. The water inlet and outlet connections of the radiator are important factors that affect the direction of water flow and consequently the heat transfer. In this scenario, the water inlet is provided from the top and the outlet from the bottom. The water flow is maintained at a rate of 0.0103 kg/s, and the water temperature is assumed to be 75 ℃. It was chosen this way based on the assumption that the heat transfer efficiency would be higher in cross-flow. This temperature is sufficient to ensure effective heat dissipation by the radiator. As the water circulates within the radiator, it transfers heat to the air across the surface area.The air flow is modeled as natural flow, with an initial temperature set at 20 ℃. Natural air flow refers to the movement of air caused by temperature differences, without the use of any mechanical fan or pump.

This type of flow generally occurs at lower speeds, and therefore, the heat transfer may be more limited compared to mechanical air flow.During the analysis, a mesh with approximately 30 million elements was created for the radiator and the surrounding flow area. This high-resolution mesh allows for more precise and accurate results in fluid dynamics simulations. The average mesh size was determined to be 0.25, indicating a rather fine mesh structure, which is a suitable choice for detailed analysis. This detailed mesh structure enables better modeling of the movement of water within the radiator, air flow, and heat transfer.The quality and size of the mesh are factors that directly affect the accuracy of the simulation. Therefore, the mesh structure used in this analysis is critically important for accurately evaluating the performance and heat transfer of the radiator.

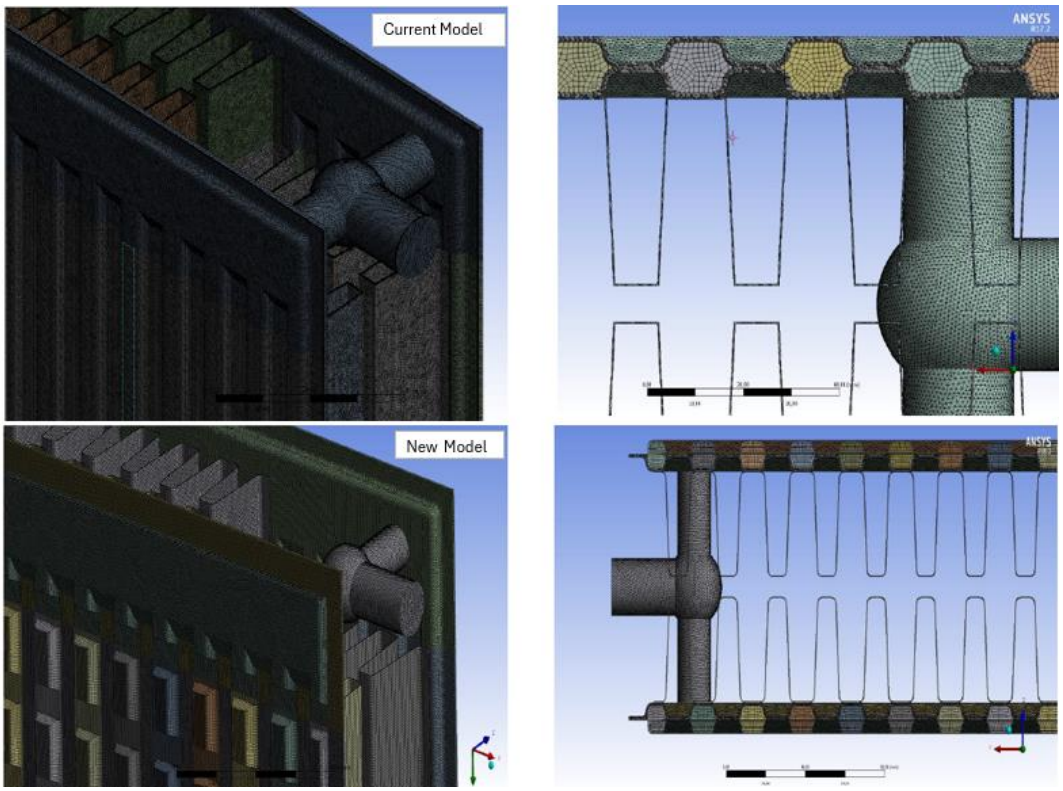


Figure 3. *Mesh structure [8]*

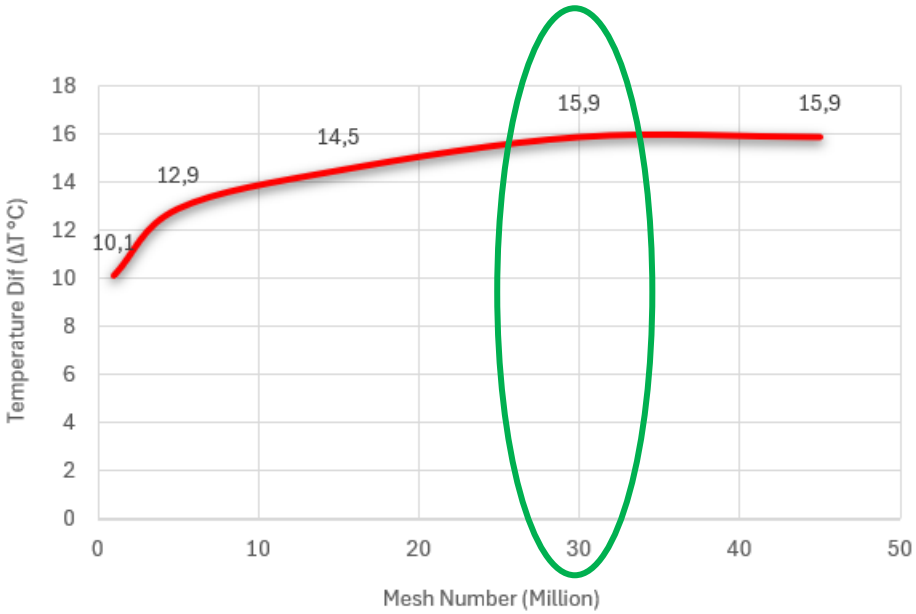
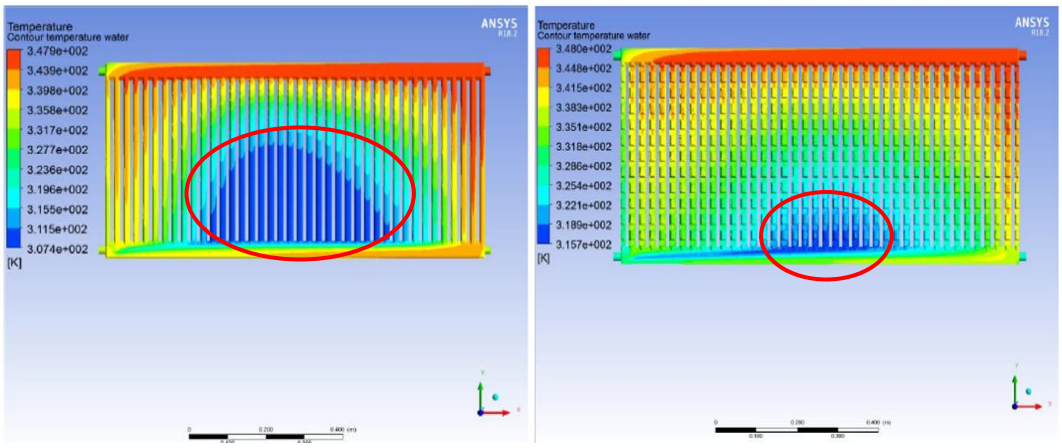


Figure 4. *Results of independence from Mesh*

3. RESULT AND DISCUSSION

From both analyses, sections were taken across three different planes to obtain visuals of velocity, pressure, and temperature. Figure 5 shows the temperature distribution visuals for the analysis of both designs under the same conditions.



(a) (b)

Figure 5. *Temperature distributions (a) Current design, (b) New design [8]*

As can be seen in Figure 5, the amount of heat transfer has been increased with the new design, and the blue area, referred to as the "cold region," has significantly reduced in size. As shown in Table 2, the water contact surface in the new design has decreased compared to the current design. Despite this, it is observed that the increase in water heat flux is at the level of 20%. Therefore, it is evident that the turbulent flow provided in the new design allows for better heat transfer of the water compared to the current design. Additionally, a difference of 4.2 °C in water outlet temperatures between the two designs has been observed. These results demonstrate that the new design will have better thermal performance compared to the current design.

Table 2*. Analysis Outcomes*

Design	Input pressure (Pa)	Output pressure (Pa)	Input temperature (K)	Output temperature (K)	Water contact surface area (m^2)	Water heat flux (W/m^2)
Current	5,9	0,6	348	332,1	1,8171	384,562
New	5,9	0,68	348	327,9	1,8055	480,383

4. CONCLUSIONS

This study demonstrates that incorporating geometrical modifications to induce turbulence in the water channels of panel radiators can significantly enhance thermal performance. The new design achieved higher heat flux and lower outlet temperatures despite a slight reduction in contact surface area. Overall, the findings confirm that the proposed design offers improved heat transfer efficiency compared to the conventional configuration.

5. ACKNOWLEDGMENT

Turk Demirdokum Fabrikaları A.Ş. is acknowledged for their valuable contributions.

REFERENCES

[1]. Chavanne, X. (2013). Energy Efficiency: What It Is, Why It Is Important, and How to Assess It. Energy Science, Engineering and Technology. New York.

[2]. Çağlar, A. (2025). Effects of using BN/water nanofluid on the thermal performance, energy saving, and power consumption of a panel radiator heating system. Case Studies in Thermal Engineering. Retrieved from ScienceDirect.

[3]. A. Tohidi, S. M. Hosseinalipour, M. Shokrpour, and A. S. Mujumdar, "Heat transfer enhancement utilizing chaotic advection in coiled tube heat exchangers," Applied Thermal Engineering, vol. 76, pp. 185-195, 2015. Available: https://doi.org/10.1016/j.applthermaleng.2014.10.073

[4]. M. J. Pour Razzaghi, M. Ghassabian, M. Daemiashkezari, A. N. Abdulfattah, H. Hassanzadeh Afrouzi, and H. Ahmad, "Thermo-hydraulic performance evaluation of turbulent flow and heat transfer in a twisted flat tube: A CFD approach," Case Studies in Thermal Engineering, vol. 35, 2022. doi: https://doi.org/10.1016/j.csite.2022.102107.

[5]. R. Yang and F. P. Chiang, "An experimental heat transfer study for periodically varying-curvature curved-pipe," Int. J. Heat Mass Transf., vol. 45, pp. 3199-3204, 2002.

[6]. A. Lanani and R. Benchabi, "Two-Dimensional Study of Heat Transfer Characteristics Flow in a Corrugated Channel," Mechanics of Fluids and Gases, vol. 22, no. 2, 2016. doi: https://doi.org/10.5755/j01.mech.22.2.13321.

[7]. S. Azarhazin, M. Soleimani, J. Kozinski, and A. Tarokh, "The influence of changing pipe cross section on the turbulent flow and heat transfer," International Communications in Heat and Mass Transfer, vol. 166, 2025, doi: https://doi.org/10.1016/j.icheatmasstransfer.2025.109126.

[8]. Ansys Fluent 18.02

[9]. DemirDöküm, (2014). Technical archive.

[10]. A. J. Jaworski (Ed.). (2018). Selected Problems in Fluid Flow and Heat Transfer. Energies, 11, 1825. Retrieved from www.mdpi.com/journal/energies.

K-Nearest Neighbors as a Transparent Baseline for Automated EEG Sleep Staging

Ahmet Sertol Köksal [1*]

Abstract

Sleep staging plays a key role in evaluating sleep health, but manually annotating polysomnography data is both time-consuming and prone to inconsistencies. This study presents the first comprehensive baseline evaluation of the K-Nearest Neighbors algorithm for EEG-based sleep staging, using a nested subject-wise cross-validation approach. We assessed 24 configurations combining six data scalers and four distance metrics on the ISRUC dataset. Overall, KNN delivered stable performance, with macro-F1 scores between 0.59 and 0.62 and Cohen's κ ranging from 0.55 to 0.57. Among scalers, the Normalizer consistently performed the worst (macro-F1≈0.52), while Power transform, Standard, and Quantile scalers produced more reliable outcomes. The choice of distance metric had a relatively minor impact, but Euclidean distance offered the best trade-off—slightly improving accuracy while delivering runtimes up to five times faster than Cosine. Hyperparameter tuning consistently favored k ≈ 30 with distance weighting, indicating that extensive parameter searches may not be necessary. At the individual class level, Wake (F1≈0.79) and N3 (F1≈0.82) stages were identified with high accuracy, whereas N2 (F1≈0.68) was moderately accurate. REM (F1≈0.56) and particularly N1 (F1≈0.27) remained difficult to classify, though some setups improved performance by up to 8%. In summary, while KNN does not match deep learning in raw accuracy, it provides valuable benefits in terms of transparency, reproducibility, and interpretability. We recommend using Euclidean distance, k≈30 with distance weighting, and avoiding the Normalizer as a practical and interpretable baseline for future EEG-based sleep analysis.

Keywords: Distance metric, EEG, KNN, Sleep staging

1. INTRODUCTION

Sleeping is a biological phenomenon central to all humans, and accurate evaluation is essential in clinical diagnosis and in understanding human health. Although the sleep monitoring gold standard, polysomnography (PSG), involves time-consuming, scorer variability-prone by-hand annotation of 30-second epochs into stages (Wake, N1, N2, N3, REM), these concerns have motivated development over recent years of automated sleep staging methods, from classical machine learning (ML) through recent deep learning (DL) methods. DL models, despite attainable accuracy levels impressive on paper [1-4], typically necessitate considerable computational power, complex designs, and are in large part "black-box" systems, thus limiting reproducibility and interpretability in clinical use. In contrast, simpler ML methods [5-8] are more interpretable, computationally lightweight, and capable, in certain conditions, of producing robust results with modest datasets, at points reaching accuracy levels close to DL systems.

Out of these methodologies, the K-Nearest Neighbors algorithm provides a natural baseline. KNN is intuitive, instance-based, and was successfully implemented in a wide variety of application domains. However, accuracy is well known to be significantly impacted by scaling, distance measure, and hyperparameters such as *k*-neighbors and weighting. Several case studies in a variety of application domains have demonstrated that choice of distance measure can significantly impact classification decisions. Abualfeilat et al. [9], for example, reported that careful selection of metric provided better predictive power in healthcare applications. Ma et al. [10] noted that alternative metrics, namely Minkowski, surpassed the usual Euclidean distance in tasks in the application domain of medical imaging. Mladenova and Valova [11] compared several distance functions in a systematic case study on text classification and confirmed that choice of metric significantly interacted with properties within the features, but noted that preprocessing by normalization impacted the results as well. Accordingly, Zhang et al. [12] suggested hybrid distance functions and demonstrated that combining metrics can remove ambiguities happening in the case of multiple equi-distance neighbors. In conclusion, the above case studies confirm that accuracy in KNN is significantly linked with distance measure choice.

In spite of all this, nothing is well established regarding the effect of distance metrics on KNN in particular in EEG-based sleep staging. To our knowledge, there are barely any works tackling the former, notably the paper by Qureshi et al. [13], comparing Euclidean, Manhattan, and Chebyshev distance in a pilot study with 25 subjects and 10-fold cross-validation. This, however, was a restricted study in several important aspects: it did not do subject-wise data division, it was restricted to

[1] * *Corresponding author: Yozgat Bozok University, Department of Computer Engineering, 66100, Yozgat, Turkey.*
ahmet.koksal@bozok.edu.tr

a single scaling method, and it did not methodically search hyperparameters such as *k* or weighting. Hence, the collective influence of scaling, metric choice, and weight giving to neighbors has never been seriously taken into account, and thus a seminal methodological fallacy in the field goes undetected.

This disparity suggests the need for a systematic and reproducible investigation into KNN settings in the context of EEG-based sleep stage classification. In the paper, we close that gap by comparing six scalers and four distance functions in 24 combinations on the ISRUC dataset (100 subjects, 1-channel EEG). The rigoruous nested subject-wise cross-validation scheme was applied in order to avoid leakage, and both the performance (macro-F1, Cohen's Kappa, per-class F1) and the computational cost were evaluated. Rather than comparing against deep learning systems, we strive at presenting a systematic and reproducible KNN baseline. Through insights on how scaling, choice of metric, and hyperparameters affect results, we provide practical suggestions on how to set up KNN, thereby offering a reference point in itself in the way of methodological improvement in the future as well as a useful tool in applied practice, in situations in which simplicity, interpretability, and efficiency are paramount. This baseline is most useful in low-resource clinical settings and in wearable or home monitoring settings, in which the depth and resource intensity of deep learning systems are infeasible.

2. MATERIALS AND METHODS

2.1. Dataset

We used the ISRUC-Sleep dataset [14], providing polysomnography recordings during a complete night in 100 participants. A single overnight EEG was employed in each person. Segmentation was performed on epochs in 30 seconds, consistent with established sleep stage practice. The initial sample frequency was 200 Hz but was reduced by half, i.e., 100 Hz, in order to reduce the computational burden. Since a low-cost, single-channel setting was aimed at and existing works made comparable, a single EEG channel, i.e., C4–A1, was used. Artifact rejection and band-pass filtering were noted as performed during dataset preparation and therefore no further preprocessing was needed.

2.2. Feature Engineering and Selection

40 candidate features were computed on a 30-second basis from every epoch during statistical, spectral, and nonlinear domains (Supplementary Table S1). Feature selection was performed with a combined correlation, information-theoretic, and model-based method. The overall process is described in Supplementary Section S1. Following selection, 15 features were retained in ISRUC (Supplementary Table S2).

2.3. Modeling and Evaluation

We utilized a KNN classifier in order to compare the effect of preprocessing and distance functions on sleep staging based on EEG recordings. We compared six scalers, namely Standard (STD), Robust (ROB), MinMax (MMX), Quantile (QNT), Power Transform (Yeo Johnson) (PYJ), and Normalizer (NRM), and four distance functions were used in each scaler, namely Euclidean (EUC), Manhattan (MAN), Chebyshev (CHB), and Cosine (COS). This experimental setup resulted in a total of 24 scalar–metric combinations. For all configurations, we tuned the value of the number of neighbors *k* in the range 1--50, and both uniform and distance weighting schemes were used.

The model was evaluated by a nested cross-validation scheme. The outer loop utilized a 20-fold subject-wise cross-validation in order to obtain unbiased estimates of the generalization error. A 3-fold subject-wise cross-validation was utilized in the inner loop in order to optimize the hyperparameters. Both loops used strict subject separation in order to prevent data leakage, thus no epoch coming from the same individual was present simultaneously in both train and test folds. The same care was also taken in preprocessing: all scalers were trained only on the train folds and then applied on the corresponding test ones, in order to prevent distributional information leakage.

Hyperparameter optimization (HPO) was carried out through Optuna [15]. The optimization was configured to run 60 trials per (fold, scaler, metric) combination, and we chose the Tree-structured Parzen Estimator [16] as the sampler along with Successive Halving [17] as a pruning strategy. The first 12 iterations were left as exploration in order to enjoy enough diversity in candidate solutions. The above settings made it efficient in search without redundant computation. Interestingly, the optimization was always converging towards $k \approx 30$ with distance weighting, a fact on which we comment in the Results section.

We assessed mainly through the use of macro-F1 and Cohen's Kappa. We chose macro-F1 as the primary metric as it treats minority classes such as N1 and REM in a balanced fashion, and we included Kappa as it has been adopted as a de facto standard in sleep stage classification. We also provided extra measures, accuracy, precision, recall, and per-class F1 scores, in order to provide a fuller accounting of a joint accuracy and cost profile. We compared computational efficiency in terms

of HPO time, time to fit model, time to make a prediction, and overall runtime per configuration, and thus a joint evaluation of accuracy and cost is possible.

2.4. Implementation Details

All the calculations were performed in Python (3.12.3). The machine learning experiments were implemented in scikit-learn (1.6.1). The data manipulations and numerical computations were performed in Pandas (2.2.3) and NumPy (1.26.4), respectively. All the experiments were conducted on a Windows 11 Pro computer running on a 12th Gen Intel(R) Core(TM) i9-12900 CPU (2.40 GHz) and 128 GB RAM. Training and validation were done exclusively on CPU and without a GPU.

3. RESULTS & DISCUSSION

3.1. Overall Performance

To illustrate the distribution of performance on scaler–metric combinations, we give a sample subset of the results on macro-F1 and Kappa in Table 1, presenting the three highest ranked combinations and MMX and NRM as a reference, and results of all 24 combinations in detail in Supplementary Table S3.

As seen in Table 1, macro-F1 values fell within a narrow band of approximately 0.59–0.62, representing a difference of approximately 5%. The confidence intervals overlap significantly. Although this implies that the choice of scaler and metric has limited statistical significance on performance, the range of results is not entirely uniform. PYJ and QNT were the best performing scalers, followed closely by STD and ROB. MMX lagged slightly behind, and NRM consistently underperformed.

Table 1. Representative subset of scaler–metric combinations with macro-F1 and Kappa results.

Scaler	Metric	Macro-F1 (mean ± CI95)	Kappa (mean ± CI95)
PYJ	EUC	0.620 ± 0.017	0.577 ± 0.021
PYJ	MAN	0.619 ± 0.017	0.576 ± 0.021
QNT	EUC	0.618 ± 0.016	0.576 ± 0.020
MMX	MAN	0.598 ± 0.016	0.548 ± 0.021
NRM	MAN	0.540 ± 0.020	0.487 ± 0.024

Among scalers, strongly similar results were obtained by PYJ, STD, QNT, and ROB (≈0.60–0.62), and MMX was somewhat lagging (≈0.59), yet statistically within the same group as the leaders. The outlier was NRM, producing significantly lower results (macro-F1 ≈0.52–0.53, with non-overlapping CI95).

In distance calculations, EUC, MAN, CHB, and COS all behaved alike (≈0.59–0.60). Statistically significant differences were indicated by Friedman's test ($\chi^2 = 280.6, p < 10^{-45}$), but in light of the small absolute effect sizes, practical significance is restricted.

In addition to F1, we also examined Cohen's Kappa, a popular choice in sleep stage classification. Estimations of Kappa were also strongly and invariably in the range 0.55–0.57 over scaler–metric pairs, with minor variations and overlapping intervals of confidence. Following the Landis & Koch scale [18], this corresponds to moderate agreement, in that, despite KNN providing a workable baseline, it does not reach levels typically desired in clinical-grade systems of "substantial" agreement. Our findings then confirm that Kappa mirrors the F1 findings: differences in performance are small over preprocessing decisions, with NRM continuing to underperform.

Figure 1 heatmap visualizations and Figure 2 grouped bar charts ((a) and (b)) also confirm these findings, demonstrating consistency in performance across scaler–metric combinations. NRM is consistently lagging in all the visualizations, confirming the conclusion that preprocessing and metric choice are not decisive in overall performance.

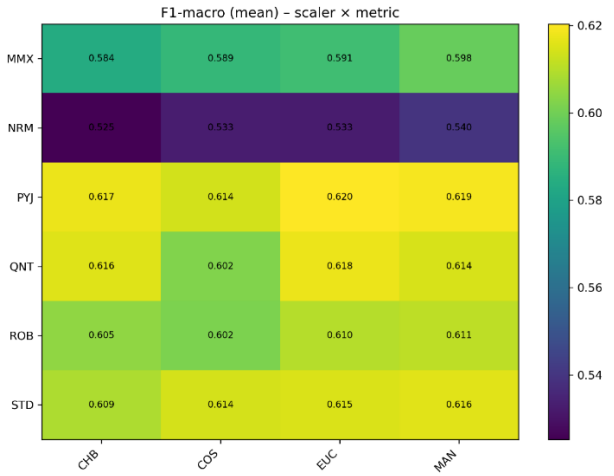


Figure 1. Heatmap of macro-F1 scores across all scaler–metric combinations.

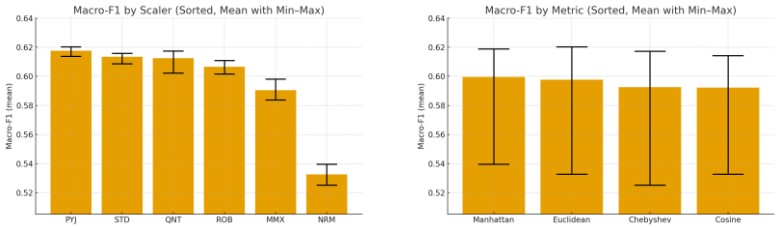


Figure 2. (a) Bar chart of best-performing metric per scaler. (b) Bar chart of best-performing scaler per metric.

3.2. Stability Across Folds

At the fold level, it was observed that variability was low, and the standard deviations lay within 0.036–0.041 and CV% lay within 5–7% among all strong scalers. Again, NRM did not only underperform but also exhibited the highest variability (CV% ≈9), a fact that both indicates instability and inaccuracy.

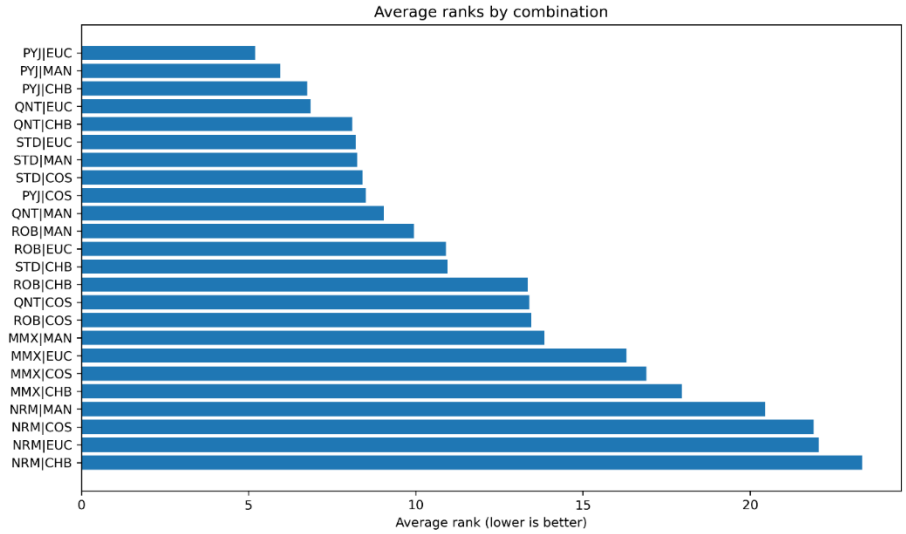


Figure 3. Average ranks of scaler–metric combinations across 20 folds.

Average rank analysis (Figure 3) supported the above results also: PYJ and QNT were first overall, STD and ROB were in a close position, MMX trailed a pace, and NRM was consistently the poorest.

The main effect tests confirmed these trends. As Figures 2 (a) and (b) show, scalers are highly precise in their segregation (PYJ, STD, QNT, ROB all close together, MMX fairly weak, and NRM well down), and metrics generate roughly equal means (MAN, EUC, CHB, COS ~0.59–0.60).

3.3. Class-wise Performance

Class-level analysis offered a less sensationalistic description. The total macro-F1 changes were minuscule, yet in all the classes the optimal combination was evident.

Table 2. Best scaler–metric combination per class based on macro-F1.

Class	Best combo	F1
N3	PYJ+COS	0.815
W	STD+EUC	0.792
N2	QNT+CHB	0.679
REM	PYJ+MAN	0.556
N1	STD+EUC	0.276

As seen in Table 2, the model had little trouble identifying N3 (F1 = 0.815) and Wake (F1 = 0.792), both of which were detected with high confidence. N2 was somewhat less accurate (F1 = 0.679), but still reasonable. In contrast, REM (F1 = 0.556) and especially N1 (F1 = 0.276) proved much harder for the model to classify correctly.

For N1, there was low recall (~0.21), but higher precision (~0.40), as if the model rarely made N1 predictions but was quite precise in so doing. The 0.25 vs 0.27 F1 discrepancy appears minuscule in absolute form, but represents ≈8% relative improvement, a potentially large impact in clinical or application settings. This shows that no scaling or choice of metric on its own can adequately address the problem of classifying N1/REM. However, the modest gains observed with certain combinations (e.g., STD+EUC being somewhat effective on N1) potentially could be a significant point of departure if supplemented with certain preprocessing or augmentation strategies particular to each class.

In order to achieve full transparency, the class-wise results on all combinations of scalers and metrics are presented in the Supplementary Table S4. This table shows, e.g., that NRM delivers significantly worse performance for N1 (≈0.18), while all other scalers are closely clustered around 0.25–0.27. This verifies the above conclusion that NRM is consistently inferior, while the remaining scalers and metrics are different from one another insignificantly.

3.4. Timing and Computational Cost

The timing analysis indicated that unless accuracy relies on metrics, they contribute substantially to computation time. Figure 4 provides a summary in grouped bar plots, in which a group represents a scaler and four distance measures are shown as adjacent bars. The three panels are displayed in a row and are associated with HPO time, fit time, and runtime in total.

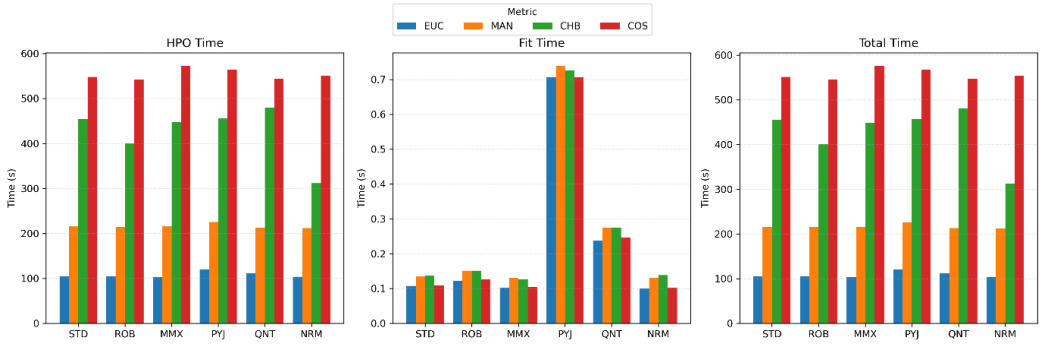


Figure 4. Grouped bar plots of HPO, fit, and total runtime across scaler–metric combinations.

EUC was by far the fastest, with total runtime from ~105–122 s. MAN took approximately 210–224 s (~2× slower), CHB 420–515 s (~4× slower), and COS was the slowest at ~550–590 s (~5–6× slower). The predictor time was also comparable in trend (EUC≈0.15 s, COS≈3.4 s). Among scalers, PYJ took a somewhat larger fitting overhead (~0.7 s compared to 0.1–0.3 s), but total runtime was nevertheless predominantly determined by the choice of metric rather than scaler.

The bar charts in Figure 5 clearly illustrate these differences: COS is the costliest option, while EUC has the same prediction accuracy at much less cost. Because it offers a balanced solution between performance and computational efficiency, we recommend that EUC be the default choice in resource-constrained environments.

3.5. Hyperparameter Optimization

The HPO results were remarkably consistent: (i) The optimal k-value was almost always 31; interquartile ranges across folds were narrow ($IQR \leq 2.5$). (ii) "weights = distance" was mostly superior; uniform weighting was rarely chosen but never among the best. (iii) For NRM, k ranged from 15–21 but always performed poorly. These observations demonstrate that KNN can be used as a consistent and computationally efficient baseline for EEG sleep staging.

3.6. Limitations

This study has several limitations. First, the analysis was conducted only on the ISRUC dataset; therefore, the generalizability of the findings to other datasets may be limited. Second, only a single EEG channel was used, which limits spatial information. Third, only 40 features were extracted from the EEG data, which may prevent the model from reflecting potential insights from a broader feature space. Finally, no data augmentation strategy was implemented, which limits classification performance from reaching optimal levels, particularly for minority classes such as N1 and REM.

3.7. Future Work

In light of the noted limitations, it will be important for future studies to validate on different datasets and modalities. Class-specific strategies should be investigated to improve the underperformance, particularly in the N1 and REM stages. For this purpose, advanced feature engineering, cost-sensitive learning, and data augmentation methods appear promising. Furthermore, the performance of different distance metrics such as Mahalanobis and Braycurtis can be evaluated. Tests on large-scale databases such as Sleep-EDF and MASS will also demonstrate the robustness of the proposed approach. Finally, the methods we use here—specifically, subject-wise cross-validation and data leakage prevention—may also provide a useful framework for algorithms other than KNN.

4. CONCLUSIONS

This study presented a baseline analysis of KNN in EEG-based sleep staging for 24 combinations of 6 scalers and 4 distance metrics. Data leakage was prevented in the analyses using extensive nested cross-validation. For most combinations, performance differences were small. The Normalizer scaler consistently performed poorly, while the other scalers were similar. The Quantile scaler provided more consistent results, while the Standard scaler produced a more balanced picture across classes. While there were minor differences between scalers, the overall performance was primarily determined by the algorithm's structure.

The effect of distance metrics on classification performance was also limited. The Euclidean distance was both more accurate and significantly faster than alternatives such as Cosine. Therefore, it offered the best balance between accuracy and speed. Runtime analysis showed that the Power transform scaler had the worst fit-time performance. There were no significant differences for the other scalers. The choice of distance metric was crucial in terms HPO time. Metrics other than Euclidean were more costly. For HPO, the weighting setting should definitely be "distance"; the number of neighbors should be approximately 30. It can be argued that further parameterization is often unnecessary.

Class-based results showed that Wake and N3 are more easily distinguished, while N1 and REM are difficult to predict. While the scaler and metric combinations offered small absolute gains, especially for N1, the relative increase of 8% can be meaningful in practice.

In conclusion, this study establishes a transparent and reproducible KNN baseline for EEG sleep staging. With distance-based weighting, approximately 30 neighbors and Euclidean distance are strongly recommended as a default choice. Furthermore, Power transform and Quantile scalers can be used to improve overall performance, while the Standard scaler is preferred in resource-constrained situations. The Normalizer scaler is not recommended. These findings provide a reliable reference for future comparative studies in the field of KNN EEG sleep staging.

5. ACKNOWLEDGEMENT

This work has been supported by Yozgat Bozok University Scientific Research Projects Coordination Unit under grant number FED-2025-1548.

REFERENCES

[11]. H. Phan, K. P. Lorenzen, E. Heremans, O. Y. Chén, M. C. Tran, P. Koch, and M. De Vos, "L-SeqSleepNet: Whole-cycle long sequence modeling for automatic sleep staging," *IEEE Journal of Biomedical and Health Informatics,* vol. 27, no. 10, pp. 4748–4757, 2023.

[12]. M. Perslev, S. Darkner, L. Kempfner, M. Nikolic, P. J. Jennum, and C. Igel, "U-Sleep: resilient high-frequency sleep staging," *NPJ Digital Medicine,* vol. 4, no. 1, p. 72, 2021.

[13]. A. Supratak and Y. Guo, "TinySleepNet: An efficient deep learning model for sleep stage scoring based on raw single-channel EEG," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC),* Jul. 2020, pp. 641–644.

[14]. H. Phan, K. Mikkelsen, O. Y. Chén, P. Koch, A. Mertins, and M. De Vos, "Sleeptransformer: Automatic sleep staging with interpretability and uncertainty quantification," *IEEE Trans. Biomed. Eng.,* vol. 69, no. 8, pp. 2456–2467, 2022.

[15]. C. H. Tai, T. Y. Liao, S. P. Chen, and M. H. Chung, "Sleep stage classification using light gradient boost machine: exploring feature impact in depressive and healthy participants," *Biomed. Signal Process. Control,* vol. 88, p. 105647, 2024.

[16]. S. K. Satapathy, D. Loganathan, and P. Narayanan, "Automated sleep staging analysis using sleep EEG signal: A machine learning based model," in *Proc. Int. Conf. Advance Comput. Innovative Technol. Eng. (ICACITE),* Mar. 2021, pp. 87–96.

[17]. E. Moris and I. Larrabide, "Evaluating sleep-stage classification: how age and early-late sleep affects classification performance," *Med. Biol. Eng. Comput.,* vol. 62, no. 2, pp. 343–355, 2024.

[18]. S. K. Satapathy and D. Loganathan, "Multimodal multiclass machine learning model for automated sleep staging based on time series data," *SN Comput. Sci.,* vol. 3, no. 4, p. 276, 2022.

[19]. H. A. Abu Alfeilat, A. B. Hassanat, O. Lasassmeh, A. S. Tarawneh, M. B. Alhasanat, H. S. E. Salman, and V. S. Prasath, "Effects of distance measure choice on k-nearest neighbor classifier performance: a review," *Big Data,* vol. 7, no. 4, pp. 221–248, 2019.

[20]. X. Ma, X. Han, and L. Zhang, "An improved k-nearest neighbor algorithm for recognition and classification of thyroid nodules," *J. Ultrasound Med.,* vol. 43, no. 6, pp. 1025–1036, 2024.

[21]. T. Mladenova and I. Valova, "Analysis of the KNN classifier distance metrics for Bulgarian fake news detection," in *Proc. 3rd Int. Congr. Human-Computer Interaction, Optimization Robotic Appl. (HORA),* Jun. 2021, pp. 1–4.

[22]. C. Zhang, P. Zhong, M. Liu, Q. Song, Z. Liang, and X. Wang, "Hybrid metric K-nearest neighbor algorithm and applications," *Math. Probl. Eng.,* vol. 2022, no. 1, p. 8212546, 2022.

[23]. S. Qureshi, S. Karrila, and S. Vanichayobon, "Human sleep scoring based on K-nearest neighbors," *Turk. J. Electr. Eng. Comput. Sci.,* vol. 26, no. 6, pp. 2802–2818, 2018.

[24]. S. Khalighi, T. Sousa, J. M. Santos, and U. Nunes, "ISRUC-Sleep: A comprehensive public dataset for sleep researchers," *Comput. Methods Programs Biomed.,* vol. 124, pp. 180–192, 2016.

[25]. T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining,* Jul. 2019, pp. 2623–2631.

[26]. J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Proc. 24th Int. Conf. Neural Inf. Process. Syst. (NeurIPS),* Granada, Spain, Dec. 2011, pp. 2546–2554.

[27]. K. Jamieson and A. Talwalkar, "Non-stochastic best arm identification and hyperparameter optimization," in *Proc. 19th Int. Conf. Artif. Intell. Statist. (AISTATS),* vol. 51, A. Gretton and C. C. Robert, Eds. PMLR, 2016, pp. 240–248.

[28]. J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics,* vol. 33, no. 1, pp. 159–174, 1977.

BIOGRAPHY

Ahmet Sertol KÖKSAL is an Assistant Professor in the Department of Computer Engineering at Yozgat Bozok University. He received his Ph.D. in Electronics and Computer Education from Sakarya University in 2014. His primary research interests include artificial intelligence, cybersecurity, and chaos-based applications in engineering.

Biodegradability of gift-wrapping papers: A comparison of different materials in the context of environmental sustainability

Ivana Plazonić[1], Ivana Ćubela[1], Iva Kvesić[1]*

Abstract

The biodegradation of eight different types of commercially available gift-wrapping papers was analysed to assess their ecological sustainability. They differ visually and tactilely from each other and were produced from different raw materials: natural, pigmented, unprinted, printed, and metallized. Filter paper made exclusively from high-quality cotton cellulose fibres was used as the reference material. All papers were tested in laboratory conditions for thickness, grammage, bulk density, moisture content, colorimetric values and surface smoothness according to Bekk. Cut into 2 cm x2 cm dimensions, the paper samples were initially weighed and subjected to biodegradation in soil in laboratory beakers and covered with transparent foil for the purpose of maintaining humidity and temperature. The biodegradation process lasted a total of 70 days, and every 14 days, 3 samples of the same paper type were taken from the soil, dried in a drying oven at 40 °C, manually cleaned of soil with a brush, and analysed. Changes in the mass of the samples were determined by weighing on an analytical balance, while microscopic analysis and colorimetric measurements provided additional confirmation of the level of material degradation. The results of this research contribute to a better understanding of the speed and potential for degradability of different types of paper, so that consumers can be informed about the importance of environmentally friendly choices when purchasing decorative paper in the context of sustainable consumption and environmental protection.

Keywords: gift-wrapping papers, biodegradation, sustainability, environmental protection

1. INTRODUCTİON

In nature, different materials biodegrade at different rates. Biodegradability can be defined as the capacity for biological degradation of organic materials by living organisms down to the base substances such as water, carbon dioxide, methane, basic elements and biomass [1]. The speed and degree of biodegradation vary widely depending on the material's structure and composition but also on the environmental conditions. Temperature and the presence and activity of microorganisms play a significant role in biodegradation [2]. Namely, most microorganisms that assist the biodegradation process need light, water and oxygen and they tend to reproduce faster in warmer conditions [3]. Therefor higher temperatures and the presence of suitable microorganisms accelerate the biodegradation process. Paper, which is made from plant-based materials, primarily cellulose, is generally biodegradable material as it can be broken down by microorganisms in the environment [4]. However, the biodegradability of specific paper products can vary based on factors like coatings, additives and used inks for specific printed designs. Today, there is a variety of papers on the market depending on their function and they all have different biodegradability levels as they maintain different processing styles [5]. A special category of paper is wrapping paper, the function of which is to wrap gifts in a way that attracts attention and creates positive excitement and delight in the mind of the gift recipient. There are several different ways to make wrapping more attractive, and shiny, glittery, ostentatious printed papers have always attracted more attention, especially if the gifts are intended for children or for some more festive occasions like weddings, anniversaries, birthdays and holidays.

Wrapping paper can be produced from a variety of raw materials (virgin wood and non-wood fibres, recycled fibres or synthetic materials), but also with different finishes and surface textures - including smooth, texture, matte, glossy, metallic or printed. If wrapping paper contains plastic coatings, laminates, or other synthetic materials that do not break down easily it is not biodegradable. Polypropylene, polyethylene, laminates and other synthetic materials are non-biodegradable [6]. On the other hand, the use of metallic or toxic ink and shiny coatings which makes our wrapped gift pleasing to the eye - shiny and glittery, is the big problem after the present has been unwrapped and the paper thrown away. This kind of paper cannot be recycled or composted, so it ends up in landfill and causing huge detriment to the planet's health [7]. Total wrapping

[1] *Corresponding author: University of Zagreb Faculty of Graphic Arts, Department of Fundamental and general knowledge, 10 000 Zagreb, Croatia. ivana.plazonic@grf.unizg.hr*

paper production reached 18,159,919 kt in 2020 in the World. The highest ranked country, from 59 countries, in wrapping paper production with 7,050,000 kt was China. China itself accounted for 38.8% of wrapping paper production in the world, while together with USA and Sweden hold a 53.2 % share [8].

Gift giving is a wonderful tradition that brings joy to both the giver and the recipient [9]. However, it also creates excess waste due to traditional gift-wrapping practices, which, depending on the type of wrapping paper, can be harmful to the environment. Given that people find it difficult to give up previously acquired habits, it is unreasonable to expect that people will stop wrapping gifts, but it is possible to enlighten them ecologically and show them the advantages and disadvantages of certain types of wrapping paper that will ultimately help them in their decision when buying them.

2. MATERIALS AND METHODS

2.1. Paper samples

Eight types of gift-wrapping paper and filter paper made exclusively from high-quality cotton cellulose fibres as the reference material were used in this research (Table 1). Gift-wrapping papers were chosen to cover diverse offers on the market based on its visual and tactile experience created by their composition. They differ in raw materials and surface treatment resulting in natural, pigmented, unprinted, printed, and metallized gift-wrapping papers.

Table 1. Sample table

Sample No.	Composition	Colour	Print
1	100% cotton fibres (*filter paper*)	white	No
2	25% grass fibre + 75% recycled fibre	natural	No
3	wood pulp (*kraft paper*)	natural	Yes (design)
4	wood pulp (*kraft paper*)	natural	Yes (monochrome blue)
5	100% recycled fibre	white	Yes (design)
6	wood-free (*crepe paper*)	red (pigment)	No
7	100% cellulose (*silk tissue paper*)	red (pigment)	No
8	100% recycled fibre	natural	lamination with green metallized foil
9	synthetic	red (pigment)	Yes (design)

All papers were tested in laboratory conditions for thickness (TAPPI T 411), grammage (TAPPI T410), bulk density, moisture content (TAPPI T412), surface smoothness according to Bekk (TAPPI T 479), and colorimetric values CIE L*a*b*.

2.2. Biodegradation of samples in laboratory conditions

Prior biodegradation analysis all paper samples were cut into 2 cm x 2 cm dimensions. Each sample was initially weighed on analytical balance, KERN ABT 220-4NM. Then, the samples in laboratory beakers were subjected to biodegradation in soil (pH = 7.5) in such a way that 3 samples of the same paper were arranged between 90 g layers of soil and covered with transparent foil to maintain humidity and temperature. In one laboratory beaker 550 g of soil was used for biodegradation test of 15 samples. The biodegradation process lasted a total of 70 days, and every 14 days, 3 samples from the same level were taken from the soil, dried in a drying oven at 40 °C, manually cleaned of soil with a brush, and analysed.

Figure 1. Biodegradation in laboratory conditions

2.3. Analysis

Changes in the mass of the samples due to biodegradation of the material were determined by weighing samples every 14 days on an analytical balance KERN ABT 220-4NM, and the mass loss (Δm) was calculated according to Formula 1 and the results for each gift-wrapping paper were presented as an average value of three performed test.

$$\triangle m = \left(\frac{m_1 - m_2}{m_1}\right) \times 100 \quad (1)$$

where: m_1 is an initial mass of sample; m_2 is a mass of sample after biodegradation in soil in defined period of time.

Spectrophotometric measurements of colorimetric values L*a*b* were made on then points before sample were subjected to biodegradation in soil using spectrophotometer SpectroEye X-rite under conditions of standard illumination D50 and observation angle 2°. After removing the samples from the soil, if the sample retained a shape suitable for measurement, the colorimetric values of the samples were measured again and compared with the initial colorimetric values. The difference in colour (ΔE_{00}^*) is calculated according to Formula 2:

$$\triangle E_{00}{}^* = \sqrt{\left(\frac{\triangle L'}{k_L S_L}\right)^2 + \left(\frac{\triangle C'}{k_C S_C}\right)^2 + \left(\frac{\triangle H'}{k_H S_H}\right)^2 + R_T \frac{\triangle C'}{k_C S_C} \frac{\triangle H'}{k_H S_H}} \quad (2)$$

where: ΔE_{00}^* is the total colour difference; $\Delta L'$ is the lightness difference between gift-wrapping paper before and after performed biodegradability test in soil; $\Delta C'$ is the chroma difference between gift-wrapping paper before and after performed biodegradability test in soil; $\Delta H'$ is the hue difference between gift-wrapping paper before and after performed biodegradability test in soil; R_T is the rotation function; K_L, K_C, K_H are the parametric factors for variation in the experimental conditions; S_L, S_C, S_H are the weighting functions.

Microscopic images of samples were captured by digital microscope Bresser (magnification 50x, image resolution 640x480), before and after removing them from the soil through all defined biodegradation periods, as additional visual confirmation of the material degradation.

3. RESULTS AND DISCUSSION

3.1. Properties of the gift-wrapping papers

Some basic and optical properties of the analysed papers are summarized in Tables 1 and 2. From results presented in Tables 1 and 2 it is visible that chosen gift-wrapping papers differ significantly in grammage, thickness, bulk density, smoothness of the surface and colorimetric values.

Table 2. Basic properties of the papers

Sample No.	Thickness, d (mm)	Grammage, x (g/m^2)	Bulk density, y (g/cm^3)	Moisture content, $w_{moisture}$ (%)	Bekk smoothness (sek)	
					face side (FS)	underside (US)
1	0.185 ± 0.02	84.68 ± 2.72	463.51 ± 37.62	1.08 ± 0.23	1.12 ± 0.08	1.02 ± 0.11
2	0.123± 0.01	69.41 ± 1.26	548.14 ± 33.07	1.58 ± 1.06	7.22 ± 1.23	2.68 ± 0.51
3	0.099 ± 0.01	59.58 ± 0.68	611.75 ± 32.10	1.01 ± 0.09	41.24 ± 1.90	2.14 ± 0.40
4	0.081 ± 0.00	61 .10 ± 0.67	761.31 ± 36.03	1.64 ± 0.70	60.58 ± 4.94	17.66 ± 0.73
5	0.060 ± 0.00	65.28 ± 0.97	1074.00 ± 73.38	0.86 ± 0.12	233.20 ± 7.60	183.40 ± 29.55
6	0.131 ± 0.01	23.85 ± 0.65	189.73 ± 26.96	0.87 ± 0.30	-	-
7	0.035 ± 0.00	19.31 ± 0.24	531.02 ± 17.97	1.23 ± 0.63	87.84 ± 2.47	10.78 ± 0.15
8	0.096 ± 0.00	72.22 ± 0.88	741.62 ± 10.55	0.59 ± 0.12	-	3.82 ± 0.29
9	0.045 ± 0.00	33.58 ± 1.07	733.79 ± 34.70	1.02 ± 0.36	-	-

** Note: - means that by Bekk method measuring paper smoothness for some samples was not appropriate as this method relies on measuring the time it takes for a vacuum to be established between the paper and a glass plate, and excessive permeability (No. 6) or very smooth surface (No. 8 and No. 9) can interfere with the vacuum formation, leading to inaccurate results or a very slow drop in air pressure during the test can make this measuring impossible.*

Table 2. Colorimetric values of the papers

Sample No.	L*	a*	b*
1	95.91 ± 0.17	0.05 ± 0.07	7.26 ± 0.22
2	83.42 ± 0.65	2.01 ± 0.12	11.35 ± 0.38
3	66.6 ± 3.07	6.96 ± 1.51	23.31 ± 1.63
4	28.19 ± 0.77	2.45 ± 0.31	-14.08 ± 0.53
5	47.79 ± 0.34	62.16 ± 0.66	39.78 ± 1.05
6	53.31 ± 0.68	66.48 ± 0.29	30.45 ± 0.45
7	45.87 ± 0.26	58.39 ± 0.49	27.81 ± 0.51
8	32.87 ± 0.39	-18.87 ± 0.27	-1.83 ± 0.14
9	41.92 ± 0.48	56.71 ± 1.10	24.43 ± 0.76

3.2. Biodegradation analysis

The biodegradation of gift-wrapping paper was monitored through the mass loss of the samples (Δm), which is shown in Figure 2 and total colour difference of the samples (ΔE_{00}*), which is shown in Figure 3.

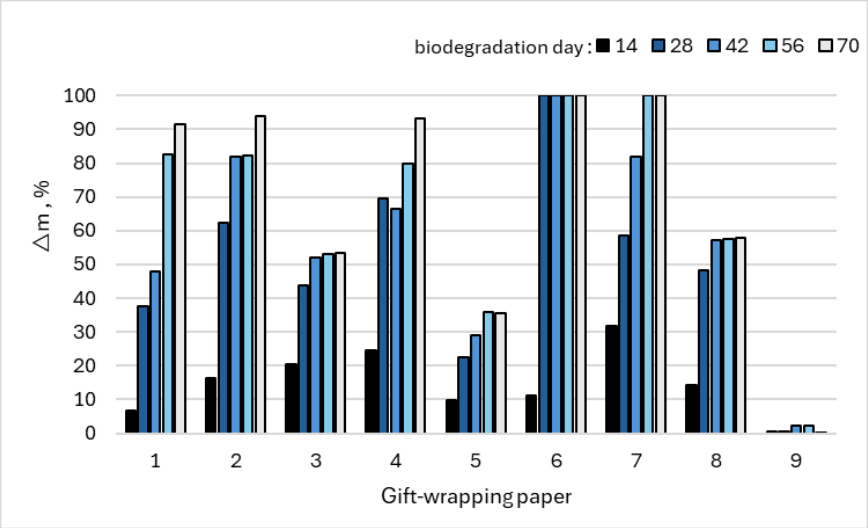


Figure 2. Mass loss of samples due biodegradation test

Gift-wrapping paper made from plant-based materials, cellulose fibres, decomposes much faster, as is evident from the weight loss over 70 days of biodegradation (Figure 2), than synthetic paper, despite its low thickness and grammage (Table 1). Synthetic paper (No.9) decomposes the slowest. It has also been observed that the fastest degrading fibre-based papers are those that do not have any surface treatment in the form of printing or lamination (No.1, No.2, No.6, No.7). These are precisely the papers that are almost completely decomposed in 70 days ($\Delta m_{\text{after 70 d}} > 90\%$). On the other hand, papers that are printed or laminated (No.3, No.4, No.5, No.8) decompose much more slowly.

When it comes to papers based on cellulose fibres, the rate of biodegradation can be linked to the smoothness of its surface. If the paper is unprinted and not laminated, its surface is rougher and therefore more accessible for soil and moisture to penetrate the material. Papers that are printed or laminated have a much smoother surface and are therefore more protected from the effects of soil and moisture. As can be seen from Figure 2, sample No.5 lost the least mass of all fibre-based gift-wrapping papers, and it has the highest smoothness according to Bekk (FS = 233.20 ± 7.60 sek; US =183.40 ± 29.55 sek). It is also visible that fibre-based gift-wrapping papers of very low grammage (**$x \approx 20$ g/m²**), low thickness and high smoothness have completely decomposed.

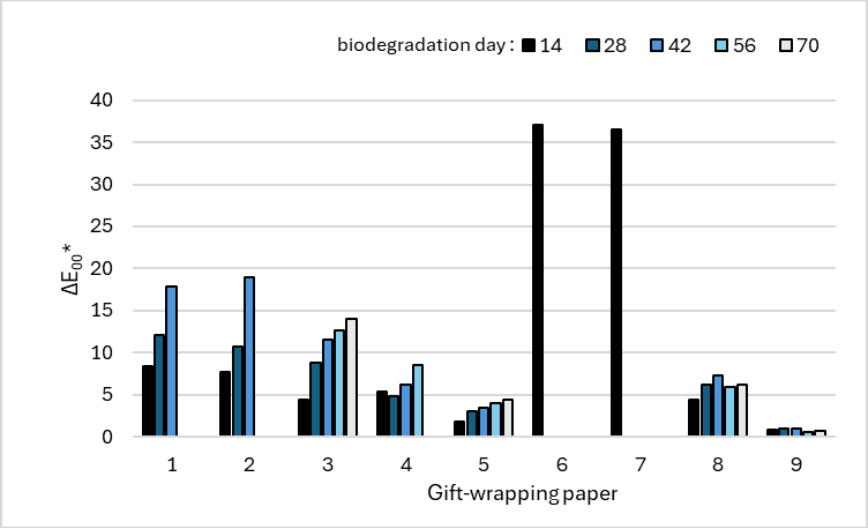


Figure 3. Total colour difference of samples due biodegradation test

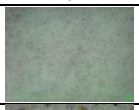
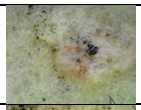
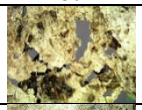
With the extension of the biodegradation time, it is visible that the difference in the surface colour of the gift-wrapping papers increases, except for synthetic paper (No.9), which is consistent with changes in the structure of the material itself due to mass loss. The less the paper decays, the less the change in colour of the paper surface. The greatest change in colour was observed in red pigmented gift-wrapping papers (No.6 and No.7) after only 14 days, the papers that completely degraded the fastest under defined conditions. It can be concluded that red pigment used in crepe wrapping paper and silk wrapping production is organic water soluble.

From Table 3 it is evident that microscopic images prove the results of biodegradability rate monitored through mass loss (Δm) and total colour difference (ΔE_{00}^{*}) of samples.

Namely, reference paper (No.1) made only from cotton fibres takes longer to start decomposition compared to papers made from wood pulp (No.3, No.4), recycled fibres pulp (No.5, No.8) or recycled fibres with the addition of grass fibres (No.2), which is logical because cotton fibres are far longer and stronger than recycled fibres. Although the degradation of the reference sample occurs more slowly during the first few weeks, after 70 days these papers reach approximately the same degree of biodegradation.

Gift-paper No.6 and No.7 after only 14 days most drastically change the colour by losing red pigment. These two papers were completely degraded in soil, No.6 after 28 days and No.7 after 56 days. It is evident that for both cellulose-based papers of the same grammage, the permeability and smoothness of the surface have a major impact on the rate of biodegradation. The more permeable and rougher the paper, the faster it decomposes. It is also evident that papers with extremely high smoothness achieved by printing, lamination or synthetic composition (No.5, No.8, No.9) are extremely stable in the soil.

Table 3. Microscopic images of papers during biodegradation process

Sample No.	Day of biodegradation					
	0	14	28	42	56	70
1						
2						
3						
4						
5						
6						
7						
8						
9						

4. CONCLUSION

To ensure biodegradability, consumers should choose cellulose-based gift-wrapping paper labelled as recyclable and biodegradable, and should avoid those with plastic, metallic, or glittery coatings. Choosing wrapping paper from natural materials, consumers contribute to reducing waste and protecting the environment. Unprinted, uncoated cellulose-based paper without metallic or plastic is the best choice for gift wrapping considering its biodegradability. While glossy or printed papers may take slightly longer to break down than paper without additional surface treatment. Laminated papers are far less susceptible to biodegradation, while synthetic papers are not biodegradable.

REFERENCES

[29]. N. T. Joutey, W. Bahafid, H. Sayel and N. El Ghachtouli, *Biodegradation - Life of Science*, Chapter: Biodegradation: Involved Microorganisms and Genetically Engineered Microorganisms,R. Chamy and F. Rosenkranz, Ed. Publ. InTech Open Science Europe, Jun. 2013.

[30]. J.Poluszyńska, T. Ciesielczuk, M. Biernacki and M. Paciorkowski, "The effect of temperature on the biodegradation of different types of packaging materials under test conditions", Archives of Environmental Protection, vol. 47, pp. 74–83, 2021.

[31]. V. Zambare and M. F. Md. Din, *Bioremediation for Global Environmental Conservation*, Chapter: Introductory Chapter: Biodegradation – New Insights, N.i Shiomi, V. Zambare and M. F. Md. Din, Ed. Publ. InTech Open Science Europe, Dec. 2023

[32]. M. A. Hubbe, J. S. Daystar, R. A. Venditti, J. J. Pawlak, M. C. Zambrano, M. Barlaz, M. Ankeny, and S. Pires, "Biodegradability of cellulose fibers, films, and particles: A Review," BioResources, vol. 20, pp. 2391–2458, 2025.

[33]. S. Ahmed, A. M. Hall and S. F. Ahmed, "Comparative Biodegradability Assessment of Different Types of paper", Journal of Natural Sciences Research, vol. 8, pp. 9–20, 2018.
[34]. S. Shaikh, M. Yaqoob and P. Aggarwal, "An overview of biodegradable packaging in food industry," *Current Research in Food Science*, vol. 4, pp. 503–520, 2021.
[35]. A. Puri and A. H. Gupta, "A review on recent trend in sustainable gift wrapping," *International Journal of Creative Research Thoughts (IJCRT)*, vol. 12, pp. f522–f534, Jan. 2024
[36]. (2025) Helgi Library website. [Online]. Available: https://www.helgilibrary.com/charts/which-country-produces-the-most-wrapping-paper/
[37]. D. J. Howard, "Gift-Wrapping Effects on Product Attitudes: A Mood-Biasing Explanation," *Journal of Consumer Psychology*, vol. 1, pp. 197–223, 1992.

Analysis of the Quality of Packaging Prints Obtained with Four Different Black Inks on Metallised Substrates

Irena Bates[1], Ivana Plazonić[1], Maja Rudolf[1], Dino Priselac[1]*

Abstract

Packaging has become an integral part of our everyday lives. It is used for a wide range of products. Printing on packaging provides information about the products themselves, the manufacturer and other important information required by regulations. Although its main purpose is to protect goods from damage, contamination and external influences during storage and transport, packaging also serves as an effective means of communication between the product and the consumer. The appearance of the packaging has a strong influence on consumer perception and purchasing decisions. Packaging is therefore often a decisive factor when choosing between products of similar value. This paper analyses the quality of packaging prints obtained with four different black (UV and three conventional) inks on metallised substrates after mechanical impact. In order to assess the quality of the packaging prints, an analysis of the prints surface was performed, defining the glossiness of the print and the stability of the prints after mechanical impact. Based on this analysis, results were obtained showing that the print with UV ink (marked B) and the print with conventional ink (marked A1) showed the greatest changes in spectral reflectance and ink coverage of the print surface after mechanical impact, while the gloss values changed the most after mechanical impact on the print marked A2 with conventional ink, and the print B with UV ink showed stable gloss values.

Keywords: packaging print, quality, black ink, metallised substrates

1. INTRODUCTİON

The earliest packaging was made from natural materials such as leaves. Later, materials such as woven containers and ceramics were used. Glass and wood have been used for 5,000 years [1]. Paper packaging has been used since the invention of paper by Tsai Lun around 105 AD, who produced paper from bamboo fibres or cambric grass [2].

Today, packaging has become an integral part of our everyday lives. It is used for a wide range of products. Although its main purpose is to protect the goods from damage, contamination and external influences during storage and transport, and at the same time to keep items together in one box, bag, or a container. However, packaging is much more than just protection, it serves as an effective means of communication between the product and the consumer [3].

Printing on packaging provides information about the products themselves, the manufacturer and other important information required by regulations. At the same time, the appearance and design of the packaging have a strong influence on consumer perception and purchasing decisions. For this reason, packaging is often a decisive factor when choosing between products of similar value [4]. In addition to its protective function, packaging also plays a key role in identifying and presenting goods. Its design must remain intact and unchanged by mechanical damage, as it is essential for the product's recognition and overall presentation. In essence, packaging combines practicality, protection, and promotion, making it an indispensable part of modern commerce [5].

In the printing industry, traditional printing techniques are still used for mass production of printed products, each offering distinct advantages and limitations. Offset printing remains the dominant printing technique for paper and cardboard packaging. This printing technique achieves the highest quality packaging with economic efficiency, making it the preferred choice for luxury packaging, speciality beverages, chocolates, perfumes, and cosmetics [2].

In conventional offset inks, drying occurs through penetration and oxidation/polymerisation in two stages. First, the printed ink film penetrates between the fibres of the printed substrate, resulting in rapid curing, with the resin in the ink playing an important role. Finally, drying is completed by oxidation and polymerisation of the ink on the printing substrate. During oxidation, oxygen initiates the polymerisation of free radicals from unsaturated oils, completing the drying of the ink on the printing substrate. UV-curable inks are specially formulated inks that contain a photoinitiator, which, upon exposure to UV

[1*] *Corresponding author: University of Zagreb Faculty of Graphic Arts, Printing Department, 10000, Zagreb, Croatia.*
irena.bates@grf.unizg.hr

energy from 200 nm to 400 nm, creates free radicals that react with the vehicle of the ink and initiate polymerisation. This method produces a completely dry layer of ink on the printing surface very quickly [6].

While conventional water-based inks can be used for all packaging, UV-curable inks may only be used for food packaging if the conditions of EuPIA's Good Manufacturing Practice are met. Whereby printers should conduct migration testing or migration modelling to cover each relevant category and application structure of each food contact material.

This paper aim was to analyse the overall quality of packaging prints obtained with four different offset black inks (one UV-curable and three conventional) on metallised substrates after exposure to mechanical impact. To assess the overall difference in packaging print quality, the print surface was analysed and the print gloss determined. In addition, spectral changes in reflectance on the opposite side of the print was measured before and after mechanical impact.

2. MATERİALS AND METHODS

The metallised paper substrates used in this research are intended for packaging to enhance its functionality, barrier and aesthetic properties. They are produced by laminating a thin film of biaxially oriented metallised polyester, 12.00 ± 0.24 mm thick and 16.80 ± 0.34 g/m², onto a paper substrate with a grammage of 250 g/m². The metallised polyester has a surface tension of 3.8 N/m^2 (ASTM D2578) [7] and is first printed with a single layer of white ink. A KBA Rapida 105 printing press was used to print three conventional black inks from different manufacturers (marked as A1, A2, A3 prints) at a speed of 8,000 sheets per hour, while a KBA RA 106 X-7+L RS printing press was used to print UV-curable black ink (marked as B print) at a speed of 9,000 sheets per hour. The descriptions and marks of all prints are presented in Table 1. During the printing process, the densitometric values of the black printed fields were measured and defined using a Techkon densitometer (M1 conditions, D50/2°).

Table 3. Mark and description of prints based on used inks

Mark of prints	Description of black ink used
A1	Conventional ink from 1. manufacturer
A2	Conventional ink from 2. manufacturer
A3	Conventional ink from 3. manufacturer
B	UV-curable ink

After printing, the samples were tested for mechanical impact resistance using a Hanatek RT4 tribometer. The test was conducted by placing the printed sample with the printed side on the unprinted substrate between discs, with the pressure regulated by placing a weight (m = 0.4535 kg) on the upper disc (standard BS 3110) [8]. After 60 revolutions, the device was stopped, and the analysis was performed.

To define the change in quality of the prints after mechanical impact, three analyses were performed. First, an image analysis of the ink surface coverage on the prints before and after treatment was carried out, along with an assessment of the change in print gloss and the change in the reflection spectrum of the opposite unprinted substrate before and after treatment.

Image analysis of the prints before and after mechanical impact was performed using a Dino-Lite Edge digital micrometer at 50× magnification and ImageJ image analysis software.

The gloss of the prints before and after mechanical impact was measured using an Elcometer 480 (ISO 2813:2014, DIN 67530) at a 60° angle, with an observation area of 8 mm × 16 mm and a repeatability of ± 0.2 GU [9,10].

The spectral reflectance values of the unprinted opposing substrate before and after mechanical impact were determined using a Techkon spectrophotometer under M1 conditions, with standard illumination D50 and a viewing angle of 2° (sensor pixel pitch < 3 nm), according to ISO 13655 [11].

All values were calculated and presented as mean values and standard deviations of 20 measurements.

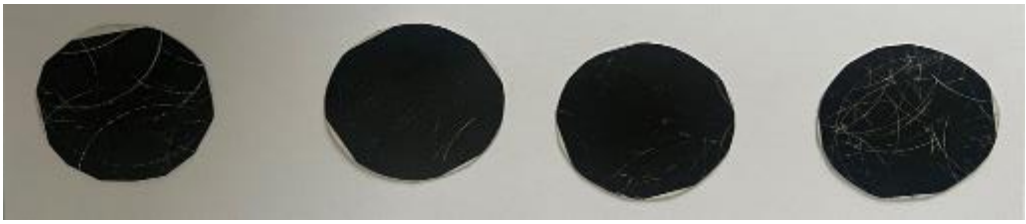
3. RESULTS AND DISCUSSION

During the printing of the entire print run, the densitometric values defined by colour management were monitored to ensure consistent and accurate colour reproduction. Table 2 presents the densitometric values of black inks printed on metallised paper, where the prints show a uniform result produced from a single digital file.

Table 2. Densitometric values of analysed prints obtained with the black ink

Mark of prints	Densitometric values
A1	1.6
A2	1.4
A3	1.1
B	0.8

After testing the samples for mechanical impact resistance using the Hanatek RT4 tribometer, a visually apparent degradation in print quality, in the form of circular arcs, was observed. The degradation in print quality of the samples is shown in Figure 1.



A1	A2	A3	B

Figure 1. Appearance of black prints after mechanical impact resistance testing

From the photograph in Figure 1, it is evident that the greatest deterioration in print quality after the mechanical impact resistance test occurred with print B, followed by print A1. Although very small, but present, deterioration in print quality was obtained with samples A2 and A3.

3.1. Image analysis of the ink surface coverage on prints before and after the mechanical impact resistance test

Image analysis of prints was performed using a digital microscope at 50x magnification. To analyse the print at all locations, 20 images of each print were captured and processed using an image analysis software. Figure 2 shows the threshold applied to the images.

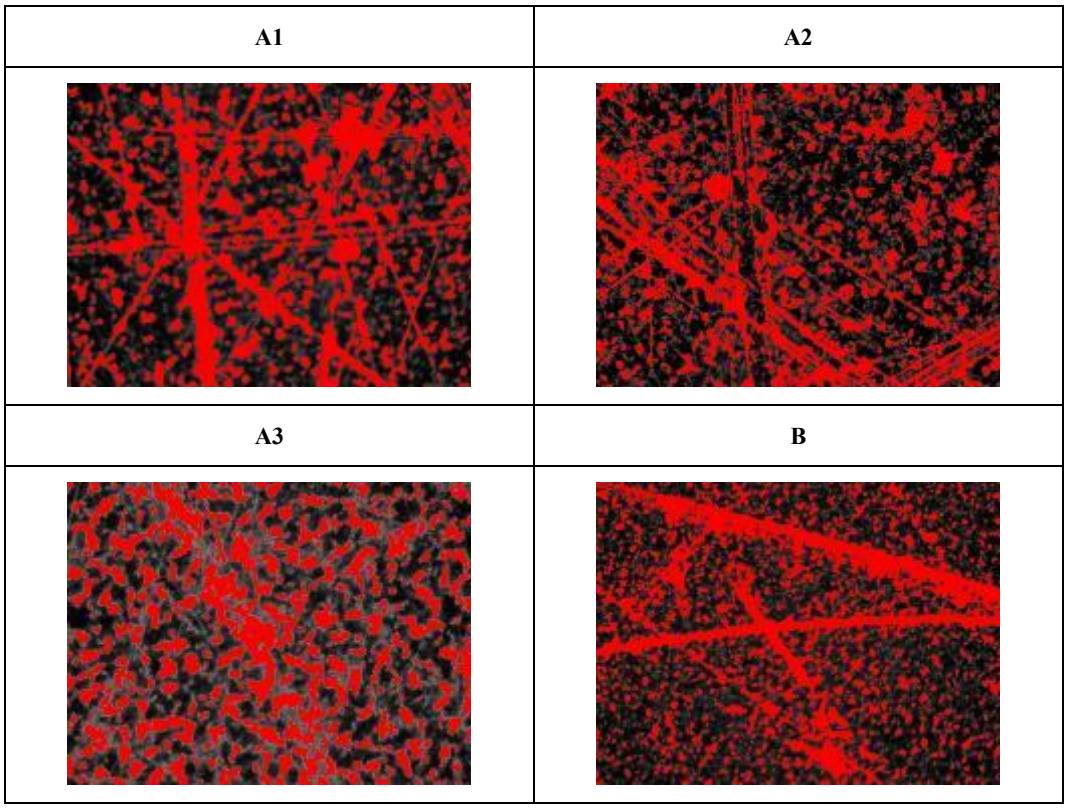


Figure 2. Image analysis of black prints with the threshold

By summarising all measured values, the mean values of ink surface coverage on the prints and their standard deviations were calculated, which are presented in Figure 3.

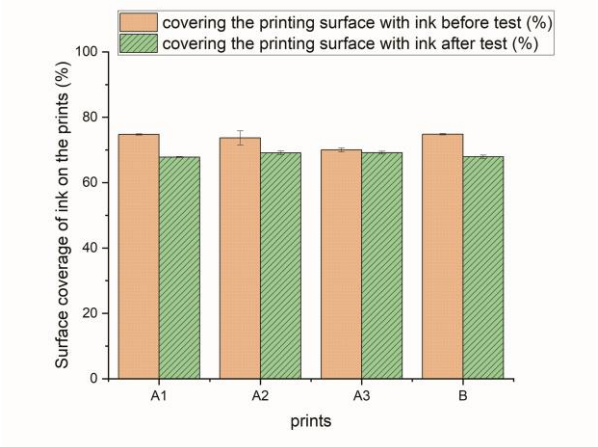


Figure 3. Surface coverage of ink on the prints before and after treatment

The results of the image analysis shown in Figure 3 indicate that the prints have an initial ink coverage between 73.69% and 74.80%, except for the A3 print, which has an ink coverage of 69.99%. It can be seen that none of the prints on the metallised paper achieved complete or nearly complete ink coverage of the substrate. The A3 print shows the least ink coverage initially.

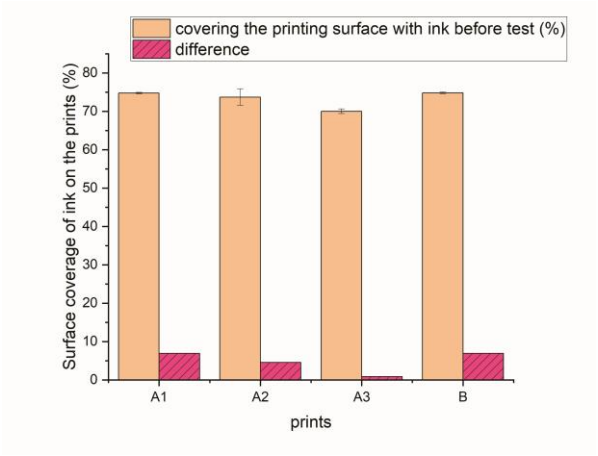


Figure 4. Surface coverage of ink on the prints before and difference after treatment

As some samples exhibit very small differences and changes, an additional Figure 4 was created to better visualise these changes, showing the differences in ink surface coverage before and after treatment. The figure 4 indicates that the A3 print has the smallest change in ink coverage, while the A1 and B prints have the largest change of 6.90% and 6.91%.

3.2. Assessment of the change in print gloss

The gloss measurements of the prints before and after the mechanical impact resistance test were performed on a white standard background, and the mean values and standard deviations of the 20 measured values before and after the test are shown in Figure 5.

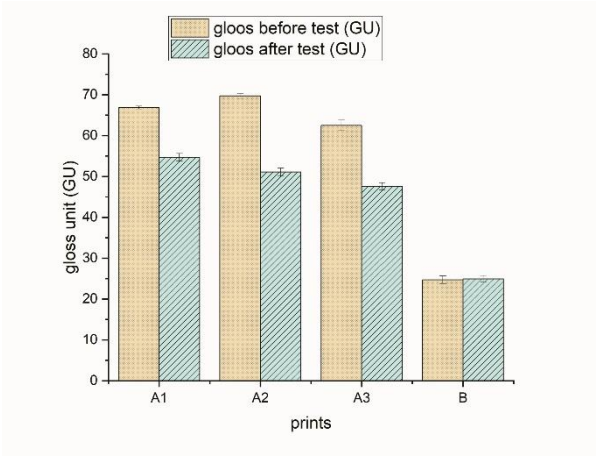


Figure 5. Gloss values of prints before and after the mechanical impact resistance test

From Figure 5 is evident that all analysed prints can be classified as semi-gloss prints as they mean gloss values is between 10 and 70 GU [12]. These results indicate that metallised substrates printed with UV-curable ink (B print), which cures rapidly, have the same gloss before and after the test. The greatest change in gloss values before and after the test was observed in A2 print made with conventional black ink, which also had the highest initial value of 69.71 GU. Among the conventional prints, the A1 print shows the smallest change in gloss values after the mechanical impact resistance test.

3.3. Change in the reflection spectrum of the opposite unprinted substrate before and after mechanical impact resistance test

When testing the mechanical impact on black prints, an opposite unprinted substrate is placed sheet-to-sheet under the pressure of a weight, and the opposite unprinted substrate is rotated for 60 cycles. On the opposite unprinted substrates, after the treatment is complete, ink transfer from the black prints is visually noticeable (Figure 6).

Figure 6. Appearance of opposite unprinted substrate after mechanical impact resistance test

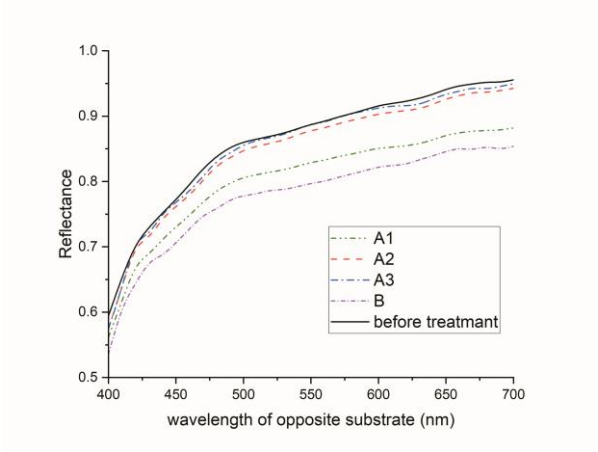


Figure 7. Appearance of opposite unprinted substrate after mechanical impact resistance test

Figure 7 shows that the reflection of the opposite unprinted substrate in the whole visible region of the spectrum decreased the most when the test was performed with print B, followed by print A1.

4. CONCLUSİONS

Based on the results of the analysis of the quality of packaging prints obtained with four different black inks on metallised substrates, using assessments of the ink surface coverage on the prints, changes in the print gloss, and changes in the reflection spectrum of the opposite unprinted substrate before and after testing, the following conclusions were drawn:

The ink surface coverage on all prints is very high, except for print A3, where it is slightly lower than in the other samples. The smallest deterioration in ink surface coverage after testing was observed in sample A3, while the largest deterioration were found in prints B and A1. Analysis of the opposite unprinted substrate confirms that the most stable print on mechanical impact is A3, while print B is the least stable one.

All analysed printed samples belong to semi-gloss prints. On the other hand, from the aspect of the stability of the surface gloss after the mechanical impact resistance test, analysis of changes in print gloss showed that the gloss of A3 print is most impaired, while the gloss values of the B print are the most stable after mechanical impact.

Finally, it can be concluded that it is necessary to perform several analyses to determine the final print quality.

ACKNOWLEDGMENT

This work has been funded by European Union funding under the project "Strateško partnerstvo za istraživanje i razvoj pametnih proizvoda i usluga za povećanje učinkovitosti proizvodnje, rada i korištenja resursa - GrafIT" (IP.1.1.03.0065) and by the University of Zagreb Faculty of Graphic Arts.

REFERENCES

[38]. E. M. S. Selke, D. J. Culter, J. R. Hernandez, *Plastics packaging*, 2nd ed., Hanser publishers, Munich, Germany, 2016.
[39]. H. Kipphan, *Handbook of Print Media*, Berlin, Germany, Springer Publishers, 2001.
[40]. B. Rundh, "Packaging design: Creating competitive advantage with product packaging", *British Food Journal*, vol.111 (9), pp. 988-1002, 2009.
[41]. G. L. Robertson, *Food Packaging: Principles and Practic*e, 3th ed., CRC Press, Boca Raton, USA, 2016.
[42]. C. Liu, M. R. Samsudin, Y. Zou, "The Impact of Visual Elements of Packaging Design on Purchase Intention: Brand Experience as a Mediator in the Tea Bag Product Category", *Behavioral Sciences*, vol. 15(2), pp. 181, 2025. https://doi.org/10.3390/bs15020181
[43]. P. Leach, *The printing ink manual*, 5th ed. Dordrecht, Netherlands, Springer, 2008.
[44]. ASTM D2578 Standard Test Method for Wetting Tension of Polyethylene and Polypropylene Films
[45]. BS 3110:1959 Methods for measuring the rub resistance of print
[46]. ISO 2813:2014 Paints and varnishes — Determination of gloss value at 20°, 60° and 85°
[47]. DIN 67530 Reflectometer as a means for gloss assessment of plane surfaces of paint coatings and plastics
[48]. ISO 13655 Graphic technology — Spectral measurement and colorimetric computation for graphic arts images
[49]. (2025) The Elcometer website. [Online]. Available: https://www.elcometer.com/en/how-to-measure-gloss-using-a-gloss-meter

BIOGRAPHY

Associate professor Irena Bates was born on the 8th December 1978 in Zagreb. After graduating from a Natural Science and Mathematics High School in Zagreb in 1997, she continued her education at the Faculty of Graphic Arts in Zagreb, majoring in technical processes and technology. She graduated from the Faculty of Graphic Arts in 2003 with a thesis entitled "Screen Printing Inks". She was first employed at Sun Chemical" company, a worldwide producer of printing inks and pigments. In 2004, she enrolled in postgraduate studies at the Faculty of Graphic Arts in Zagreb, and in November 2006, she was employed in the Printing Department at the Faculty of Graphic Arts. She participated in research training at several foreign institutions. She earned her PhD in 2013, defending her dissertation entitled "Study of specific parameters of flexographic printing reproduction". Since January 2018, she has led the scientific research project titled "Printability, Quality, and Utilization of Substrates with Non-wood Fibres" (UIP-2017-05-2573), funded by the Croatian Science Foundation. This project lasted until 7 June 2023. She has published 126 scientific papers and 10 professional papers.

Estimating the Number of Contributors in DNA Mixtures Using Machine Learning: A Review

Simona Nikolovska[1*], Marija Nikolovska[2], Ana Cholakoska Cilakova[1], Danijela Efnusheva[1], Bojana Velichkovska[1]

Abstract

The interpretation of DNA profiles usually starts with evaluating the number of contributors (NoC). This affects how DNA evidence is analyzed and compared to known profiles. Traditional statistical models have provided foundational tools for this purpose, based on scientifically validated methods. However, determining the number of contributors in complex DNA profiles with absolute certainty is regarded impossible, due to factors like allele sharing and allelic drop-out. Therefore, the primary intent in research is to bring the mischaracterization to a lower level. With the advancement of machine learning, the task turns to optimizing the use of the available profile information. In this study, we focus on examining the challenges associated with handling genetic data, as well as legal considerations surrounding its collection and sharing for research purposes. Moreover, we provide a comprehensive review of the efforts (methods and data both) for estimating the NoC in DNA mixtures utilizing machine learning technologies.

Keywords: DNA mixtures, DNA profiling, genomics, machine learning, research ethics

1. INTRODUCTION

Genetic data is one of the most sensitive forms of personal data, unique to each individual, containing health and personal characteristics of the respondent. Their uniqueness is what makes studying them difficult, as it involves discovering evolutionary relations between a set of proteins to obtain information about their connection. Therefore, it is essential to have access to comprehensive and quality data, as incomplete or inconsistent information can give a completely different direction to the research and can lead to misdiagnoses, ineffective treatments, and false conclusions [1]. This further highlights the need of a global genomic database or genetic datasets for specific purposes available for study and research. Considering the sensitivity of the data, it is important to create different levels of accessibility to the depth of data that needs to be accessed [2]. There is no doubt that a DNA database can potentially violate human rights, if confidentiality is abused and the sensitive information gets revealed [3].

DNA databases are also a powerful tool for crime investigation [4]. This is validated with a survey compiled by Interpol in 2019 [5], where most of the countries cooperating with Interpol stated that they use DNA profiling in criminal investigations of one form or another. Moreover, police in member countries can submit a DNA profile from offenders, crime scenes, missing persons and unidentified bodies to Interpol's automated DNA database [5]. This database follows the guidance of medical ethics and law by not keeping any nominal data linking a DNA profile to any individual and not containing information about a person's physical or psychological characteristics, diseases or predisposition for diseases. A DNA profile is simply a list of numbers based on the pattern of an individual's DNA, producing a numerical code which is used to differentiate individuals [5]. Having access to this type of data significantly contributes in identifying suspects, linking crime scenes, or proving suspect's innocence.

The interpretation of the DNA profiles usually starts with evaluating the number of contributors (NoC) [6]. The term NoC refers to the estimate number of individuals whose DNA is present in one DNA sample. DNA samples with more than one contributor are referred as mixtures, as they contain genetic material from multiple sources. Minimizing the NoC estimation error is essential to improve forensic analysis and ensure that all potential contributors are identified and accounted for.

1.1. Ethical Considerations in Medical Research

By its nature, genetic data contains very sensitive and personal information about a person, that has a potential to cause harm to the individuals. With its whole set of data, the genome of a person reveals information about other individuals, including parents, or other family members [7]. If the genetic data is not subject to strict privacy and confidentially protections, it can be used against the research participant, by the insurance company, employer, or in court

[1*] *Corresponding author: Faculty of Electrical Engineering and Information Technologies, Ss. Cyril and Methodius University in Skopje, North Macedonia. nikolovskasimona6@gmail.com*

[2] *Faculty of Natural Sciences and Mathematics, Ss. Cyril and Methodius University in Skopje, North Macedonia*

proceedings. In science, this is called genetic discrimination and it occurs when individuals receive different treatment from their employer or the insurance company, because of a genetic mutation that causes or increases the risk of a specific disorder [8]. As a result, the public would normally fear and refuse any genetic testing for genetic research. This refusal could jeopardize the maintenance of individual and public health, if potentially treatable conditions are not detected in a timely manner. Therefore, a number of countries in Europe, America, Asia, and Australia have introduced laws to prevent such discrimination [9].

The other set of challenges, include the process of data collection and preparation. Here, the entire process of preparing, evaluating, revising, and checking the protocols needs to be completed in order to confirm that all steps are in accordance with standard research practices.

1.2. The Usage of Medical Data for Research

Research in medicine requires detailed analysis and processing of medical data, in order to contribute to new scientific knowledge, medicines, treatments, etc. However, considering the limitations imposed by medical ethics, they should be conducted in an appropriate manner.

Informed consent is required when research involves participation of a human, when the research uses human genetic material or biological samples, and when research involves collection of personal information [10]. Informed consent has 3 fundamental requirements [7]:

(1) Research participants must be provided with full information regarding the research,
(2) Research participants must be competent to understand the information and all the implications of the research, and
(3) Research participants must volunteer without any influence

A potential participant must be informed about the research procedures and their purposes, risks, expected benefits and any other alternative procedures, with the opportunity to ask questions and withdraw from the research at any time [7].

Additionally, referring to the concept of confidentiality, clinical data needs to be de-identified to be exported to research databases [11]. The process of de-identification is often done by relabeling participants with unique identifiers, different from the identifiers initially assigned in a clinical trial. With that, the data belonging to an individual in clinical environment can still be associated with the same individual in a de-identified research context [11][12].

2. FORENSIC DNA MIXTURE ANALYSIS

Initial step in interpreting DNA profiles is often determining the NoC. This affects how DNA evidence is analyzed and compared to known profiles.

2.1. Challenges

The assessment of the NoC in forensic mixtures is a complex and often very challenging task. To ensure objective assessment, all of the alternatives must be considered. Several biological and technical factors can complicate this process [13][14]:

- Mixture ratio (unequal contributor proportions by different individuals)
- Allelic drop-in or drop-out (unpredictable appearance of an extra allele or missing allele)
- Allele sharing between contributors
- Stutter peaks and split peaks (predictable extra alleles introduced by mistakes in the DNA replication process)
- Allelic pull-up (software error, where artificial peaks are produced in another dye channel at the same position below a major allele peak)
- Shoulder peaks (low-level distortions at the edge of a primary peak, caused by incomplete DNA replication or software-related artefacts)
- Peak imbalance (alleles does not appear in the expected proportion, due to technical or sample-related factors)
- Imbalanced profiles generated as a result of a genetic phenomena
- Baseline noise
- Overall DNA quality and quantity

2.2. Methods

Considering the complexity of this task, different methods for estimating the NoC in DNA mixtures have been developed. Counting the maximum number of alleles at any locus is the simplest one. The process divides the maximum allele count (MAC) by two and rounds up, because a single individual should only have a maximum of two alleles at a locus [15][16]. Similar approach is the total allele count (TAC) method, which assigns the NoC based on the total allele count across all loci [17]. However, many studies have shown that the allele counting methods are unreliable, especially for mixtures with more contributors due to allele sharing between contributors [15][18][19]. Another approach is applying probabilistic and statistical models to determine the most likely NoC. For instance, the authors of [18] explore the maximum likelihood estimation of NoC based on allele frequency within the profile. Similarly, in [16], the authors provide a computational method that calculates the posterior probability of the NoC based on peaks detected in the DNA profile (their height and frequency) and achieves accuracy of 83% on their laboratory generated DNA profiles. Another method adapted for this purpose is the Markov Chain Monte Carlo (MCMC) method calculating Bayes factors, which then compares the existing hypothesis for NoC using probabilities [20].

Providing more accurate estimation based on scientifically validated statistical methods that quantify uncertainty and improve the accuracy of the analysis, these methods have credibility and are widely accepted in court for interpreting DNA mixtures [21]. Despite that, they can be computationally expensive and difficult to interpret, since they generate probability distributions rather than simple NoC estimate. However, the persisting uncertainty about the composition of the DNA mixtures remains, as a result of the various inherent biological and technical factors.

In recent years, research in this field is increasingly focused on application of machine learning (ML) methods. Unlike previous methods, ML models can analyze complex patterns in forensic DNA profiles and potentially provide more precise contribution estimates.

3. MACHINE-LEARNING BASED WORKS

This section explores ML-based works that have contributed to estimating the NoC in DNA mixtures, both to optimize forensic methodologies and increase reliability of forensic profiling. We analyze current research from two perspectives, data resources and their development and application.

3.1. Data Resources

Firstly, in Table 1 we examine the data developed to support research and facilitate the development of reliable models for estimating the NoC in forensic DNA mixtures. The first dataset (generated by Marciano et al. [22]) contains 1405 profiles from 20 individuals and considered up to four contributors in a mixture. As a comparison to that, the second and third dataset contain profiles with one up to five contributors. Important note here is that the differing number of contributors/classes in the first dataset limits the direct comparability with the further models that use the other datasets.

Moving forward, major difference that the second dataset (generated by Benschop et al. [23]) introduces in this context, is a variety of 1174 unique donor participants for its 590 profiles, unlike the first dataset and the PROVEDIt dataset which use the same combinations of donors for the profiles, multiple times [6][22].

Table 1 Datasets used for estimating NoC in DNA mixtures

Dataset	Profiles	NoC/Classes
Generated by Marciano et al. [22]	1405	1, 2, 3, 4
Generated by Benschop et al. [23]	590	1, 2, 3, 4, 5
PROVEDIt [24][25]	over 25 000	1, 2, 3, 4, 5

Subsequently, the largest and the most widely used open-source forensic DNA mixture dataset to date is PROVEDIt (Project Research Openness for Validation with Empirical Data Initiative) [24]. It contains over 25 000 STR profiles from 1-5 contributors, collected in a four-year period, under 144 different laboratory conditions [25]. The wide time frame ensures comprehensive data collection across diverse forensic scenarios, minimizing the potential gaps that could occur, if the data was collected over a short period of time. Additionally, the diversity of the laboratory conditions allows validation across different protocols, equipment and environmental factors.

3.2. ML Models

Table 2 shows a comparison between related works for estimating the NoC in DNA mixtures. The first column lists the research paper, the second links dataset on which the research was conducted, column three and four show the ML methods and number of features used, column five states the number of profiles in the data, and the last column gives the reported accuracy of the respective works.

Each of the analyzed works, utilized different methods for achieving the best accuracy for NoC estimate. As input features, the models typically used statistical summaries from the profiles, related to the number of alleles in the profile, peak values, peak height ratios, and related metrics.

Table 2 Comparison between machine-learning-based works for estimating NoC in DNA mixture

Research	Dataset	Method	Features	Profiles	Accuracy
PACE (2017) [22]	Custom by the authors (Marciano et al.)	Support Vector Machine (non-linear)	-	1405	0.97
LDA40 (2019) [23]	Custom by the authors (Benschop et al.)	Linear Discriminant Analysis	40	590	0.83
RFC19 (2019) [23]		Random Forest Classification	19		0.83
Kruijver (2021) [6]	Subset of the PROVEDIt dataset	Decision Tree	13	766	0.77-0.85
Alotaibi (2021) [26]		Logistic Regression	-	780	0.95
TAWSEEM (2022) [28]		MLP with 15 layers (each with 64 neurons)	75	6000	0.97
deepNoC (2024) [27]		Hierarchical Multi-task Neural Network	-	743	0.93

In [22], the authors analyzed an approach using classification and regression trees (CART), with an accuracy of 97.5%. Due to overfitting, they instead proposed a nonlinear support vector machine (SVM) approach, with both a testing accuracy and an F1-score of 97%. Similarly, the LDA40 model presented in [23] achieved an accuracy of 83% using linear discriminant analysis with 40 features. Working with the same dataset, the random forest classifier (RFC19 [23]) achieved identical accuracy as LDA40.

The remaining works from Table 2 use the PROVEDIt dataset. All reviewed studies are based on a subset of the dataset that differentiates in size. Similar number of profiles are used by Kruijver et al. [6], Alotaibi et al. [26] and Taylor et al. [27]. The data used by Kruijver et al. and Taylor et al. is almost the same, i.e., both use mixtures amplified by the GlobalFiler multiplex with a 25 second injection protocol. On the other hand, Alotaibi et al. chose to include data with different injection times and cycle numbers. Examining the preprocessing part, Kruijver et al. test different filtering methods (3SD, Oracle and Tree filtering) on the dataset and then extract statistical features from the filtered data, among which are MAC, TAC, and particular statistical summaries about specific markers. On the other hand, Alotaibi et al. do not provide specific information about their preprocessing, only mentioning the challenges they faced including dealing with empty cells, OL values, and deleting the unwanted markers. Taylor et al. build their model on a complex architecture, assuring detailed analysis on each peak, providing 89 pieces of information collected about each peak and ability to capture up to 50 peaks at each locus.

A comparison of the classical ML algorithms applied on the PROVEDIt dataset, shows that the logistic regression model achieved higher accuracy of 95% [26], compared to the performance of the decision tree approach with 77-85% accuracy [6].

TAWSEEM [28] and deepNoC [27] use deep learning to analyze the NoC in mixtures. The authors of [27] achieved 93% accuracy with deepNoC, implementing hierarchical multi-task neural network. On the other hand, TAWSEEM achieved 97% accuracy [28], using multi-layer perceptron neural network. However, Taylor et al. in [27] report that the results of TAWSEEM may be skewed by experimental design.

4. CONCLUSION

DNA profiling, or more specifically, estimating the NoC in complex DNA mixtures remains a challenging task in forensic genetics. Traditional statistical models have provided foundational tools for this purpose, yet they often struggle when working with highly complex profiles. This study offers a review of the recent work in this field, demonstrating the promising elements of ML models, due to their ability to detect complex patterns in data that traditional statistical methods may overlook.

While the integration of ML into forensic DNA analysis holds, several limitations and open questions remain. The data and models need to be generic enough to support different laboratory protocols and populations. Subsequently, the whole process mandates transparency to ensure acceptance within the legal system. This highlights the need for rigorous validation of every step, starting from the data collection, processing, to the interpretation of the results. Another important aspect is that all reviewed studies to date, work with laboratory data that includes up to five contributors. While

this provides a valuable foundation for model development, it may not fully capture the challenges imposed by real-world forensic samples.

In conclusion, the application of ML approaches to NoC estimation represents a valuable direction for advancing forensic DNA interpretation. However, to fully realize their potential and ensure their acceptance within the legal system, further work is required to address ethical, technical and methodological challenges, as well as to ensure access to a comprehensive and good quality data for study and research.

REFERENCES

[1]. H. Kang, "The prevention and handling of the missing data", May 2013, DOI: 10.4097/kjae.2013.64.5.402
[2]. L. Bean and M. R. Hegde, "Gene Variant Databases and Sharing: Creating a Global Genomic Variant Database for Personalized Medicine", March 2016, DOI: 10.1002/humu.22982
[3]. Australian Institute of Criminology, "The Forensic Use of DNA Profiling", November 1990, Available: https://www.aic.gov.au/sites/default/files/2020- 05/tandi026.pdf
[4]. W. Goodwin, A. Linacre and S. Hadi, "An Introduction to Forensic Genetics", 2nd Edition
[5]. The International Criminal Police Organization – Interpol. Available: https://www.interpol.int/en
[6]. M. Kruijver, H. Kelly, K. Cheng, M. Lin, J. Morawitz, L. Russell, J. Buckleton, J. Bright, "Estimating the number of contributors to a DNA profile using decision trees", January 2021, DOI: 10.1016/j.fsigen.2020.102407
[7]. B. E. Ellerin, R. J. Schneider, A. Stern, P. G. Toniolo, S. C. Formenti, "Ethical, Legal and Social Issues Related to Genomics and Cancer Research: The Impending Crisis", November 2005, DOI: 10.1016/j.jacr.2005.03.016
[8]. L. M. Slaughter, "The Genetic Information Nondiscrimination Act: Why Your Personal Genetics are Still Vulnerable to Discrimination", August 2008, Available: https://doi.org/10.1016/j.suc.2008.04.004
[9]. Y. Joly, C. Dupras, M. Pinkesz, S. A. Tovino, M. A. Rothstein, "Looking Beyond GINA: Policy Approaches to Address Genetic Discrimination", January 2020, Available: https://doi.org/10.1146/annurev-genom-111119-011436
[10]. S. Manti and A. Licari, "How to obtain informed consent for research", June 2018, DOI: 10.1183/20734735.001918
[11]. R. Noumeir, A. Lemay, J. Lina, "Pseudonymization of Radiology Data for Research Purposes", December 2006, DOI: 10.1007/s10278-006-1051-4
[12]. M. D. Sorani, J. K. Yue, S. Sharma, G. T. Manley, A. R. Ferguson, The TRACK TBI Investigators, S. R. Cooper, K. Dams-O'Connor, W. A. Gordon, H. F. Lingsma, A. I. R. Maas, D. K. Menon, D. J. Morabito, P. Mukherjee, D. O. Okonkwo, A. M. Puccio, A. B. Valadka, E. L. Yuh, "Genetic Data Sharing and Privacy", October 2014, DOI: 10.1007/s12021-014-9248-z
[13]. The Royal Society, "Forensic DNA analysis A Primer for Courts", November 2017, Available at: https://royalsociety.org/~/media/about-us/programmes/science-and-law/royal-society-forensic-dna-analysis-primer-for-courts.pdf
[14]. P. Gill, B. Sparkes, T. M. Clayton, J. Whittaker, A. Urquhart, J. S. Buckleton, "Interpretation of Mixtures Based on Peak Area - Identification of Genetic Anomalies, Stutters and Other Artefacts", 1998, Available at: https://at.promega.com/~/media/files/resources/conference%20proceedings/ishi%2009/oral%20presentations/03.pdf
[15]. G. M. Dembinski, C. Sobieralski, C. J. Picard, "Estimation of the number of contributors of theoretical mixture profiles based on allele counting: Does increasing the number of loci increase success rate of estimates?", March 2018, Available: https://doi.org/10.1016/j.fsigen.2017.11.007
[16]. H. Swaminathan, C. M. Grgicak, M. Medard, D. S. Lun, "NOCIt: A computational method to infer the number of contributors to DNA samples analyzed by STR genotyping", May 2015, Available: https://doi.org/10.1016/j.fsigen.2014.11.010
[17]. J. Thomson, D. Moore, T. Clayton, "Introducing eNoC – A simple, excel-based tool for improved assignment of the number of contributors (NoC) to a mixture", December 2022, Available: https://doi.org/10.1016/j.fsigss.2022.09.016
[18]. H. Haned, L. Pene, J. R. Lobry, A. B. Dufour, D. Pontier, "Estimating the Number of Contributors to Forensic DNA Mixtures: Does Maximum Likelihood Perform Better Than Maximum Allele Count?", September 2010, Available: https://doi.org/10.1111/j.1556-4029.2010.01550.x
[19]. S. Norsworthy, D. S. Lun, C. M. Grgicak, "Determining the number of contributors to DNA mixtures in the low-template regime: Exploring the impacts of sampling and detection effects", May 2018, Available: https://doi.org/10.1016/j.legalmed.2018.02.001
[20]. D. Taylor, J. Bright, J. Buckleton, "Interpreting forensic DNA profiling evidence without specifying the number of contributors", November 2014, DOI: 10.1016/j.fsigen.2014.08.014
[21]. M. W. Perlin, "Explaining the Likelihood Ratio in DNA Mixture Interpretation", December 2010, Available: https://www.promega.es/~/media/files/resources/conference%20proceedings/ishi%2021/oral%20presentations/perlin.pdf
[22]. M. A. Marciano and J. D. Adelman, "PACE: Probabilistic Assessment for Contributor Estimation- a machine learning-based assessment of the number of contributors in DNA mixtures", March 2017, Available: https://doi.org/10.1016/j.fsigen.2016.11.006
[23]. C. C.G. Benschop, J. van der Linden, J. Hoogenboom, R. Ypma, H. Haned, "Automated estimation of the number of contributors in autosomal short tandem repeat profiles using a machine learning approach", November 2019, Available: https://doi.org/10.1016/j.fsigen.2019.102150
[24]. L.E. Alfonse, A. D. Garrett, H. Swaminathan et al., "The PROVEDIt Initiative: The Development and Release of a Collection of Computational Tools and a Large-Scale Empirical Data Set for Forensic Research and Validation", 2017, Available: https://www.aafs.org/sites/default/files/media/documents/AAFS-2017-B214.pdf
[25]. L. E. Alfonse, A. D. Garrett, D. S. Lun, K. R. Duffy, C. M. Grgicak, "A large-scale dataset of single and mixed-source short tandem repeat profiles to inform human identification strategies: PROVEDIt", January 2018, Available: https://doi.org/10.1016/j.fsigen.2017.10.006
[26]. H. Alotaibi, F. Alsolami, R. Mehmood, "DNA Profiling: An Investigation of Six Machine Learning Algorithms for Estimating the Number of Contributors in DNA Mixtures", December 2021, DOI: 10.14569/IJACSA.2021.0121115
[27]. D. Taylor and M. A. Humphries, "deepNoC: A deep learning system to assign the number of contributors to a short tandem repeat DNA profile", December 2024, Available: https://doi.org/10.48550/arXiv.2412.09803
[28]. H. Alotaibi, F. Alsolami, E. Abozinadah, R. Mehmood, "TAWSEEM: A Deep-Learning-Based Tool for Estimating the Number of Unknown Contributors in DNA Profiling", February 2022, Available at: https://doi.org/10.3390/electronics11040548

ICETI

9TH INTERNATIONAL CONFERENCE ON ENGINEERING TECHNOLOGY AND INNOVATION

Western Michigan University

University of Cincinnati

CNR Group