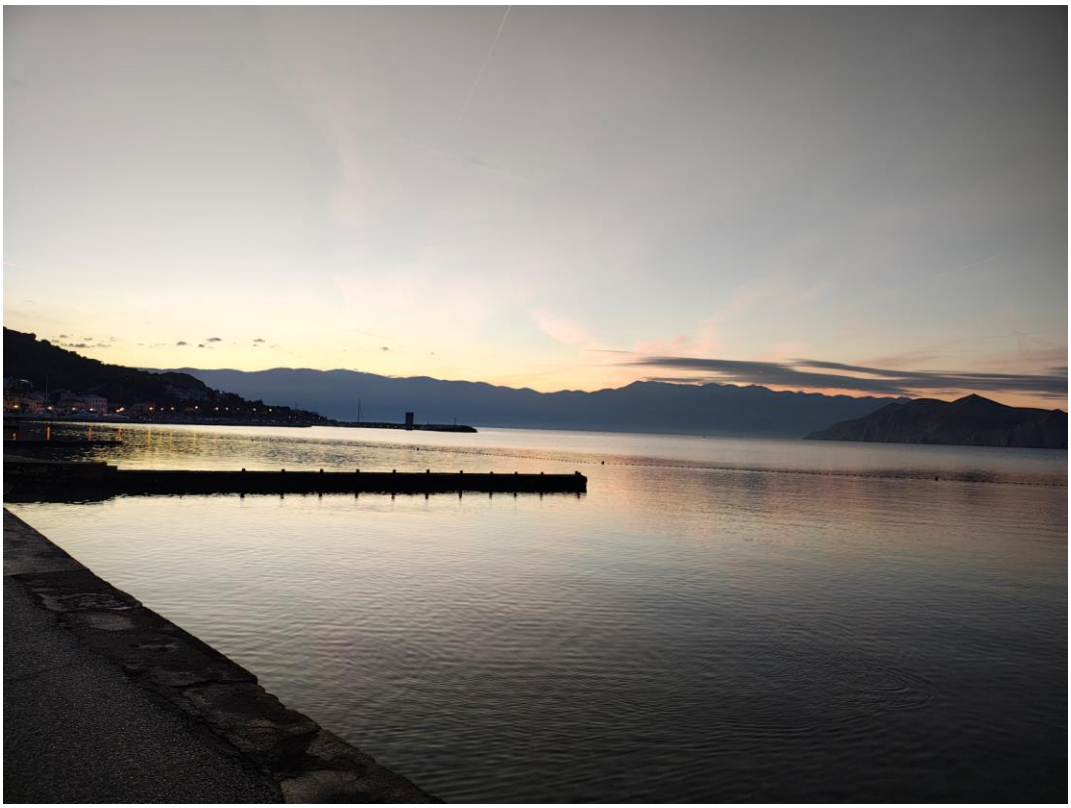
Baška SIF (Spatial Intelligence Forum) Meeting 2025

Baška, Krk Island, Croatia,

21th - 24th September, 2025

From GPS/GNSS, to spatial AI - even the sky is no more the limit!

PROCEEDINGS



Organised by

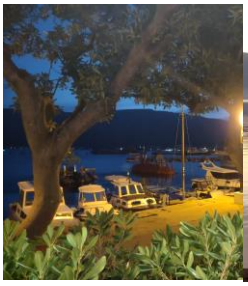
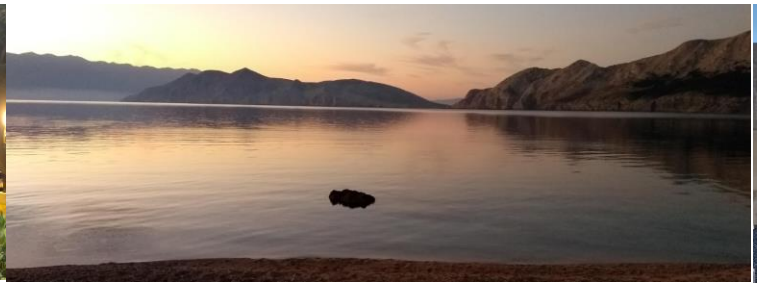
Faculty of Engineering, University of Rijeka, Rijeka, Croatia

Technical co-sponsors:

Hrvatsko Zagorje Krapina University of Applied Sciences, Krapina, Croatia

Virovitica University of Applied Sciences, Virovitica, Croatia

Media coverage: Coordinates


BAŠKA SIF (SPATIAL INTELLIGENCE FORUM) METING 2025

From GPS/GNSS, to spatial AI -
even the sky is no longer the limit!

BAŠKA, KRK ISLAND, CROATIA, 21 – 24 SEPTEMBER, 2025

The Faculty of Engineering, University of Rijeka, Rijeka, Croatia organised The Baška SIF (Spatial Intelligence Forum) Meeting 2025, in person. The conference was held in Corinthia Baška Sunny Hotel by Valamar, Baška, Krk Island, Croatia, 21st - 24th September, 2025.

Baška, Krk Island, Croatia, a popular holiday resort on the magnificent Adriatic Krk Island, host the traditional annual scientific Baška SIF Meeting 2025. Well-known for its natural scenic beauty, rich history and culture, and unforgettable hospitality, Baška has been inspiring and facilitating science, research, international collaboration and understanding since its establishment in times of the Old Roman Empire.

Organised by Faculty of Engineering, University of Rijeka, Rijeka, Croatia, and co-sponsored by Hrvatsko Zagorje Krapina University of Applied Scences, Krapina, Croatia, and Virovitica University of Applied Sciences, Virovitica, Croatia, Baška SIF Meeting 2025 convened scientists and researchers, practitioners, developers, operators, policy-makers, regulators, and users interested in utilisation of spatial data for knowledge, technology, business, and capacity development. The Baška SIF Meeting 2025 provided the framework for exchange of idea, presentation of the latest research results, the knowledge transfer and sharing, as well as to facilitate international multi-disciplinary scientific, research & development, and commercial projects establishment and collaboration.

Published by:
Faculty of Engineering, University of Rijeka, Rijeka, Croatia

For the Publisher:
Prof Jonatan Lerga, PhD,
Faculty of Engineering, University of Rijeka, Rijeka, Croatia

Editors:
Marta Alvir, PhD, Faculty of Engineering, University of Rijeka, Rijeka, Croatia
Prof Dr Renato Filjar, Hrvatsko Zagorje Krapina University of Applied Sciences, Krapina, Croatia, and
Faculty of Engineering, University of Rijeka, Rijeka, Croatia

Front-page photo credits:
Prof Dr Renato Filjar

Text formatting:
Faculty of Engineering, University of Rijeka, Rijeka, Croatia

Electronic publishing:
Faculty of Engineering, University of Rijeka, Rijeka, Croatia

Address:
Faculty of Engineering, University of Rijeka, Rijeka, Croatia
Vukovarska 58
51000 Rijeka
Croatia
Phone: +385 (0) 51 651 444
Fax: +385 (0)51 651 416
URL: https://riteh.uniri.hr/
E-mail: dekanat@riteh.uniri.hr

Published in May, 2026.

ISBN 978-953-8246-08-1

BAŠKA SIF MEETING 2025
INTERNATIONAL PROGRAMME AND ORGANISING COMMITTEE

Chair

Jonatan Lerga, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*

Co-chairs

Jasna Prpić-Oršić, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*
Teodor B Iliev, *University of Ruse, Bulgaria*

Members

Marta Alvir, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*
Nikola Andjelić, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*
Jeffrey Auerbach, *U.S. Department of State, Washington, DC*
Marina Banov, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*
Peter Brida, *University of Žilina, Žilina, Slovakia*
Enes Ciriković, *Virovitica University of Applied Sciences, Virovitica, Croatia*
Margarida Coelho, *University of Aveiro, Aveiro, Portugal*
Robert Crane, *Sam Houston State University, Huntsville, TX*
Christopher Erickson, *US Space Force, Washington, DC*
Iakovos Fantidis, *International Hellenic University, Kavala, Greece*
Sharafat Gadimova, *International Committee on GNSS (ICG), UN OOSA, Vienna, Austria*
Dorota A Grejner-Brzezinska, *Ohio State University, Columbus, OH*
Martin Grzebellus, *NavCert, Braunschweig and Munich, Germany*
Rick Hamilton, *U. S. Coast Guard, Washington, DC*
Haris Haralambous, *Frederick University, Nicosia, Cyprus*
Ivan Hedji, *Virovitica University of Applied Sciences, Virovitica, Croatia*
Manuel Hernandez-Pajares, *Universitat Politechnic de Catalunya, Barcelona, Spain*
Sanna Kaasalainen, *Finnish Geosptial Research Institute (FGI), Helsinki, Finland*
Attila Komjathy, *NASA Jet Propulsion Laboratory, Pasadena, CA*
Panagiotis Kogias, *International Hellenic University, Kavala, Greece*
Danijel Koprivanac, *Virovitica University of Applied Sciences, Virovitica, Croatia*
Andrzej Krankowski, *University of Warmia and Mazury, Olsztyn, Poland*
Kristijan Lenac, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*
Zolzaya Lkhamsuren, *Mongolian Geospatial Association, Ulaanbaatar, Mongolia*
Ivana Lučin, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*
Yenca Migoya-Orué, *The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy*
Lana Miličević, *independent researcher, Karlovac, Croatia*
Jade Morton, *University of Colorado, Boulder, CO*
Aleksandar Nešković, *School of Electrical Engineering, University of Belgrade, Belgrade, Serbia*
Washington Ochieng, *Imperial College, London, UK*
Ashik Paul, *Institute of Radio Physics and Electronics, University of Calcutta, Kolkata, India*
Babatunde Rabiu, *UN African Regional Centre for Space Science, Technology, & Education, Ile-Ife, Nigeria*
Nenad Sikirica, *Hrvatsko Zagorje Krapina University of Applied Sciences, Krapina, Croatia*

Milja Simeunović, *Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia*
Arian Skoki, *Morplo, Faculty of Engineering, University of Rijeka, Rijeka, Croatia*
Katarina Stefanović, *Telecom Serbia, Belgrade, Serbia*
Ivaylo Stoyanov, *University of Ruse, Ruse, Bulgaria*
Darko Špoljar, *Hrvatsko Zagorje Krapina University of Applied Sciences, Krapina, Croatia*
Goranka Štimac Rončević, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*
Mladen Tomić, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*
Karen Van Dyke, *GPS/GNSS Expert, Rockport, MA*
Saša Vlahinić, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*
Ivan Volarić, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*
Lucija Žužić, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*

BAŠKA SIF FORUM 2025 LOCAL ORGANISING COMMITTEE

Chair

Mladen Tomić, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia*

Members

Luca Maras, *Valamar Hotels & Resorts, Poreč, Croatia*
Renato Filjar, *Faculty of Engineering, University of Rijeka, Rijeka, Croatia, and Hrvatsko Zagorje Krapina University of Applied Sciences, Krapina, Croatia*
Magdalena Supan Budimilić, *Valamar Hotels & Resorts, Baška, Croatia*
Filip Šklebar, *Hrvatsko Zagorje Krapina University of Applied Sciences, Krapina, Croatia*
Ivana Topić, *Tourist Board of the Baška Municipality, Baška, Croatia*

LIST OF REVIEWERS

Marta Alvir
Renato Filjar
Marko Mikša
Lana Miličević
Filip Šklebar

CONTENT

Automatic Image Segment Classification Using Contour Features

Lucija Žužić
Department of Computer Engineering
Faculty of Engineering, University of Rijeka
Rijeka, Croatia
and
Center for Artificial Intelligence and Cybersecurity
University of Rijeka
Rijeka, Croatia
e-mail: lucija.zuzic@uniri.hr
ORCID: https://orcid.org/0009-0007-6788-7522

Jonatan Lerga
Department of Computer Engineering
Faculty of Engineering, University of Rijeka
Rijeka, Croatia
and
Center for Artificial Intelligence and Cybersecurity
University of Rijeka
Rijeka, Croatia
e-mail: jonatan.lerga@riteh.uniri.hr
ORCID: https://orcid.org/0000-0002-4058-8449

***Abstract*—In today's world, there is a significant need for automatic image segmentation and object identification. Images can have large dimensions, which can be problematic when processing them in their entirety. This paper aims to perform image segmentation classification using meaningful contour features rather than the full object. The object contour is first converted into a series of features. Then, the correlation coefficients with the class are examined to select the features for further processing. Different families of classifier methods are applied to the prepared input data. Candidate models using various algorithms and the full or reduced predictor set are compared, allowing for the final model to be selected for possible deployment based on the accuracy and execution time. This allows for fully accurate segment classification in the test set at a low temporal cost, using a Naive Bayes (NB) classifier with contour height and extent as inputs. This work introduces a novel, lightweight pipeline suitable for embedded underwater inspection systems. Results demonstrate that a Naive Bayes classifier, utilising only height and extent features, achieves the highest accuracy on the test set with negligible inference time, outperforming complex models in efficiency.**

***Keywords*—image segmentation, contour features, machine learning, Naive Bayes classifier, shape analysis, underwater inspection**

I. Introduction

Image segmentation and classification are fundamental tasks in computer vision, forming the basis for higher-level scene understanding and object recognition. Traditionally, these tasks have relied on manual annotation or supervised learning techniques, demanding extensive classified data and domain-specific knowledge. In this paper, an automatic method for image segment classification is proposed, leveraging contour features to associate semantic labels with image regions, eliminating the need for large-scale human annotation.

Our approach uses a contour-based representation of image segments, enabling the extraction of robust, informative shape-related features. These features are input to a classifier trained on a dataset with three classes [1], [2], allowing for scalable classification across a wide range of images [3], [4]. The proposed method achieves high accuracy while maintaining computational efficiency by integrating contour geometry with spatial and contextual cues.

This work introduces a novel contour feature extraction technique tailored for segment classification. This serves to develop a classification framework that combines local and global shape descriptors. The approach is validated through extensive experiments on a full or reduced feature set, using various families of classifiers, demonstrating its effectiveness across the three classes in the classification dataset. The images contain cracks (missing pieces) [5], [6] and indentations (structural damage) in underwater concrete walls [7]. We also identify ropes that partially cover these walls. This work paves the way for more autonomous visual systems capable of understanding complex scenes with minimal supervision.

The main contributions of this paper are:

- A computationally efficient pipeline for automatic classification of underwater concrete defects using contour geometry.
- A comparative analysis of eight classifier families, validating that simple geometric features (height and extent) are sufficient for this domain.
- Validation of a model suitable for resource-constrained embedded systems, balancing high accuracy with low inference latency.

II. Related Work

Deep learning has revolutionised automated defect detection, with architectures such as U-Net [8], [9] and variants of You Only Look Once (YOLO) [10], [11] becoming standard for segmentation tasks. Specific to infrastructure, models such as MultiScaleCrackNet [12] and dual-encoder networks [13] have demonstrated high precision in identifying fine-grained cracks by leveraging multiscale learning. To address real-time constraints, recent research has shifted toward lightweight networks and pruning techniques to reduce computational

overhead [14]. However, in challenging underwater environments, these data-hungry black-box models often struggle due to data scarcity and high turbidity, sometimes necessitating the generation of synthetic data using a Generative Adversarial Network (GAN) [15]. Unlike these computationally intensive approaches, the proposed method prioritises interpretability and speed. By focusing on explicit geometric contour features rather than learned textural representations, the proposed pipeline achieves high classification accuracy suitable for resource-constrained embedded systems without the need for massive annotated training sets.

III. Methodology

The flowchart summarising the full pipeline in Figure 1 includes input image processing and segmentation, contour extraction, feature calculation, feature selection, application of various classifiers, and final output classification for each contour, described by the provided set of contour features.

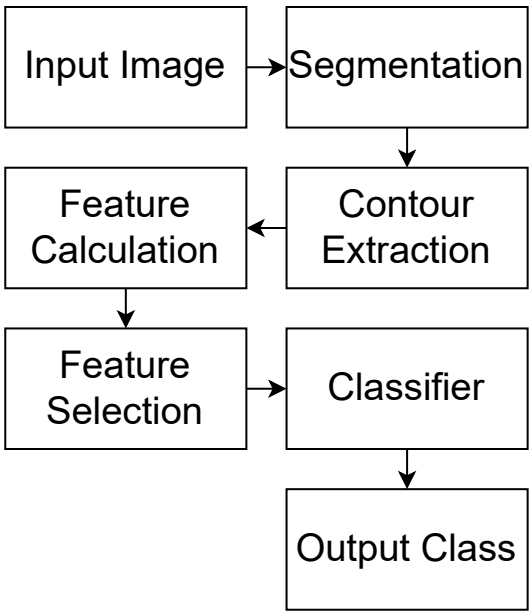


Fig. 1. The flowchart illustrating the full pipeline.

A Support Vector Machine (SVM) with a Polynomial Kernel, the Naive Bayes (NB) method, Flexible Discriminant Analysis (FDA), the Partial Least Squares (PLS) method, a C5.0 Decision Tree (DT), the Random Forest (RF) method, eXtreme Gradient Boosting (XGB), and a shallow Neural Network (NN) were tested based on their ability to classify a set of classified image segments based on various contour features, into one of the predefined object labels.

The methods were selected because they represent larger families of classification methods. SVM models are supervised, maximum-margin models. DT models also apply supervised learning in a tree structure. NB classifiers are probabilistic classifiers that can be parametric or nonparametric; this study uses a nonparametric approach. The FDA uses multiple nonparametric linear regression models to produce a nonparametric probabilistic classification. PLS is a non-parametric linear regression model. RF and XGB use ensemble learning, combining multiple weak classifiers based on simple models such as a DT, reducing errors and increasing robustness. NN models mimic the brain using artificial neurons to produce outputs based on inputs and activation functions. NN models require a predefined structure and hyperparameter tuning. All tested models were applied according to the research by Kuhn for the *R caret* package [16], [17].

Hyperparameters were optimised using a grid search within the *caret* package, with performance evaluated via 10-fold cross-validation repeated 10 times [18].

A. Support Vector Machine

To capture non-linear relationships between geometric features and defect classes, a Support Vector Machine (SVM) with a Polynomial Kernel [19], [20] was used. While linear models [21] are computationally cheaper, the separation between classes, representing ropes and cracks, necessitated a non-linear boundary in the feature space. A kernel probability density function was fitted to ensure compatibility with the probabilistic evaluation framework [22]. The polynomial degree, scale, and cost parameters were optimised via repeated k-fold cross-validation.

B. Naive Bayes

A non-parametric Naive Bayes (NB) classifier [23] was selected for its computational efficiency [24] and robustness on small feature sets [25]. Despite the assumption of feature independence, NB effectively separates classes when geometric distributions (height or extent) differ. We employed a kernel density estimate for the feature distributions, tuning the bandwidth adjustment and Laplace correction to smooth the data and prevent zero-probability issues [25].

C. Flexible Discriminant Analysis

Flexible Discriminant Analysis (FDA) was employed to model non-linear class boundaries [26] by combining Multivariate Adaptive Regression Splines (MARS) with Linear Discriminant Analysis (LDA) [27]. This approach allows for more flexible decision surfaces than standard LDA, accommodating potential irregularities and non-normality in the contour feature distributions [28]. The interaction degree and pruning terms were determined through cross-validation while keeping execution time to a minimum [29]. The FDA relies on the assumption of feature normality and equal group covariances [30], so it may not be the best fit for the complex structures examined in this work.

D. Partial Least Squares

Partial Least Squares (PLS) [31]–[33] modelling was tested to handle potential multicollinearity among the full set of 16 contour features. By projecting predictors into a latent space that maximises covariance with the categorical class labels [32], PLS reduces noise and feature redundancy. The number of latent components was the primary hyperparameter tuned during validation.

E. Decision Tree

A single Decision Tree (DT) [34] was trained using the C5.0 algorithm to establish a baseline for interpretability [35]. Unlike complex ensembles, this model generates explicit, human-readable rules based on feature thresholds, such as specific height cutoffs for ropes [36]. As a single-tree implementation, no hyperparameters need to be tuned, making it a deterministic benchmark for ensemble methods.

F. Random Forest

To mitigate the high variance often associated with single decision trees, a Random Forest (RF) [37] classifier was implemented. By aggregating predictions across multiple uncorrelated trees trained on bootstrap samples [38], [39], this ensemble method reduces overfitting [40], a crucial factor given the distinct but overlapping geometric properties of the defect classes. The number of randomly selected predictors at each split was optimised via cross-validation.

G. Gradient Boosting

The eXtreme Gradient Boosting (XGB) [41] algorithm was applied to sequentially correct classification errors using gradient descent. The tree-based booster was tuned with a comprehensive set of hyperparameters, including shrinkage (η), maximum tree depth, and minimum loss reduction (γ), to balance model complexity and generalisation. This method was chosen to test if boosting could exploit subtle feature interactions that simpler models might miss.

H. Neural Networks

A shallow Neural Network (NN) [42] was trained to capture complex non-linear interactions [43] between features without the high computational cost of deep architectures [44]. The model utilises a feed-forward structure [45], with weights adjusted via backpropagation to minimise empirical risk [46]. To prevent overfitting on the dataset, the number of hidden units and the weight decay parameter were rigorously tuned using 10-fold cross-validation [47].

IV. Dataset

The authors gathered 5413 classified objects from 3 classes across 1209 underwater concrete wall images [1], [2], and the objective was to classify them using machine learning and various contour features. The classes represent missing pieces in the wall (cracks), structural weakness and fractures (indentations), and parts of rope that potentially obstruct the wall from view. This is illustrated in Figure 2, which shows a masked image with labels.

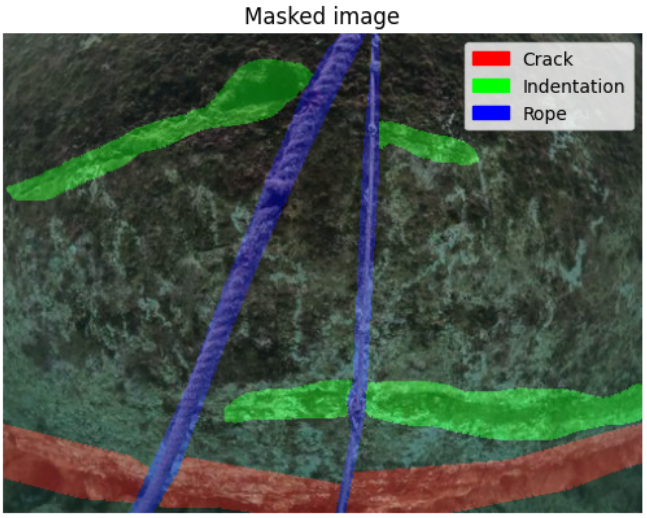


Fig. 2. An example of a masked image containing all three classes, crack (red), indentation (green), and rope (blue).

This could pave the way for automatic classification after unsupervised image segmentation, using other well-known algorithms such as Gaussian Mixture Models (GMMs) [48], or deep learning [49], [50] with architectures like YOLO [11], the Segment Anything Model (SAM) [51], [52], U-Net [8], [9], Residual Networks (ResNet) [53], and MobileNet [54].

A. Contour Features

OpenCV is the largest non-profit open-source computer vision library, encompassing more than 2500 algorithms and running since June 2000. The 16 utilised contour features from the *OpenCV* library are defined in Table I, and listed as follows:

- the width, height, aspect ratio, and contour area
- the bounding box rectangle area, and the convex hull area
- the extent, solidity, equivalent diameter, and orientation
- the top and left rotated bounding rectangle coordinates
- the width and height of the bounding rectangle's midpoints
- the bottom and right bounding rectangle coordinates

The calculation methods for aspect ratio, extent, solidity, and the equivalent diameter are given in the continuation of the test, along with explanations of the convex hull, bounding rectangle, and orientation calculation.

1) Bounding Rectangle: A straight rectangle is used in this paper, and the rotation of the object is not considered. The bounding rectangle's area will be minimised in this case. The contour area, the bounding rectangle area, the width, the height, the right coordinate, and the bottom coordinate are used as features. The midpoint's width and height are also used as model input.

2) Aspect Ratio: The aspect ratio is the contour width divided by the height, made clear by Equation 1.

TABLE I
THE 16 GEOMETRIC CONTOUR FEATURES EXTRACTED FOR CLASSIFICATION.

Feature Name	**Description**	**Formula / Definition**
Width	Width of the straight bounding rectangle	OpenCV `boundingRect`
Height	Height of the straight bounding rectangle	OpenCV `boundingRect`
Aspect Ratio	Ratio of object width to height	$\frac{\text{Width}}{\text{Height}}$
Contour Area	Total number of pixels inside the contour	OpenCV `contourArea`
Bounding Box Area	Area of the straight bounding rectangle	Width $\times$ Height
Convex Hull Area	Area of the smallest convex polygon enclosing the contour	Sklansky's Algorithm ($O(N \log N)$)
Extent	Ratio of contour area to bounding rectangle area	$\frac{\text{Contour Area}}{\text{Bounding Box Area}}$
Solidity	Ratio of contour area to convex hull area	$\frac{\text{Contour Area}}{\text{Convex Hull Area}}$
Equivalent Diameter	Diameter of a circle with the same area as the contour	$\sqrt{\frac{4 \times \text{Contour Area}}{\pi}}$
Orientation	Angle of the ellipse fitted to the contour (0° to 180°)	Fitzgibbon & Fisher Algorithm
Rotated Rect. Top	Top coordinate of the rotated bounding box	OpenCV `minAreaRect`
Rotated Rect. Left	Left coordinate of the rotated bounding box	OpenCV `minAreaRect`
Middle Width	Width at the bounding rectangle's midpoint	Geometric Calculation
Middle Height	Height at the bounding rectangle's midpoint	Geometric Calculation
Bounding Rect. Right	Rightmost coordinate of the bounding rectangle	$x + \text{Width}$
Bounding Rect. Bottom	Bottommost coordinate of the bounding rectangle	$y + \text{Height}$

$$Aspect\ Ratio = \frac{Width}{Height} \tag{1}$$

3) Extent: The extent is the contour area divided by the bounding rectangle area, as defined using Equation 2.

$$Extent = \frac{Contour\ Area}{Bounding\ Box\ Rectagle\ Area} \tag{2}$$

4) Convex Hull: The area of the convex hull is another tested model parameter. The convex hull will look similar to contour approximation, and both may provide the same results in some cases, but not always. In this study, a curve is checked for convexity defects and corrected, yielding a convex curve that encompasses all points within the contour. This paper finds the convex hull of a two-dimensional (2D) point set using Sklansky's algorithm [55] with $O(N \log N)$ complexity in the current implementation. Generally speaking, convex curves are the curves that are always bulged out, or at least flat. A curve bulged on the inside has a convexity defect, which is the local maximum deviation of the hull from the contours.

5) Solidity: The solidity diameter is the contour area divided by the convex hull area, evident in Equation 3.

$$Solidity = \frac{Contour\ Area}{Convex\ Hull\ Area} \tag{3}$$

6) Equivalent Diameter: The equivalent diameter is the diameter of a circle with an equal aggregate sectional area, which is calculated using Equation 4.

$$Equivalent\ Diameter = \sqrt{\frac{4 \times (Contour\ Area)}{\pi}} \tag{4}$$

7) Orientation: Since the bounding box does not account for the object's rotation, the next step is to fit an ellipse. The algorithm by Fitzgibbon and Fisher [56] returns the rotated rectangle inside of which the ellipse is inscribed. The top and left coordinates of the rotated rectangle are used as predictors, as well as the orientation of the fitted ellipse as an angle between 0° and 180°.

B. Training and testing dataset

There were 965 training and 244 testing images, split at a $80/20$ ratio, as shown in Table II. The split was stratified to ensure that all labels and labels in each class follow the same ratio. All labels from the same image belong to the same split. The prevalence of cracks, indentations, and ropes in the combined dataset, training, and testing is 25%, 42%, and 33%, respectively.

TABLE II
THE NUMBER OF IMAGES AND INSTANCES OF EACH CLASS IN TRAINING AND TESTING IMAGES.

Set	Images	Labels	Crack	Indentation	Rope
Train	965	4324	1075	1828	1421
Test	244	1089	270	460	359

V. RESULTS

Candidate models were assessed to determine the optimal method and predictor set for the final model to be applied in real settings. The choice of predictors was based on theoretical knowledge, visual inspection, and variable correlation coefficient values. The reasoning behind this approach was the assumption that reducing the set of predictors would reduce model complexity and computation time. Confusion matrix indicators were evaluated individually for the crack, indentation, and rope classes, with one class designated as positive and the others as negative.

A. Feature Selection

The height has the largest correlation coefficient of 0.53 with the class in Figure 3, a heatmap of variable correlation. Ropes (class 2) are usually elongated vertically rather than horizontally. The strongest negative correlation coefficient, -0.57, was found between the labels and the extent variable. The objects in classes 0 and 1 represent cracks and indentations, respectively. Unlike ropes, cracks and indentations do not have

many gaps or blank areas in their outlines. The height and extent were selected as predictors for the experiment because they provide a clear separation of the labels, particularly for the rope class.

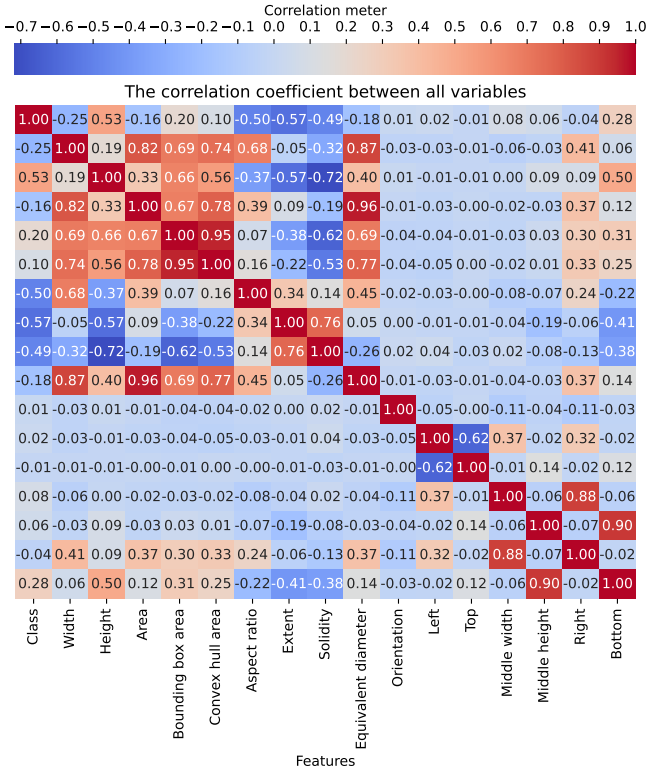


Fig. 3. A heatmap of the correlation between all variables used in this study. Red represents a strong positive correlation, blue represents a strong negative correlation, and white represents a weak correlation. Variables are fully correlated with themselves, so values on the secondary diagonal equal 1. The matrix is symmetric with respect to the secondary diagonal because the same combination of variables results from swapping the row and column.

This is clear in the scatter plot of the objects' heights and extents, labelled in Figure 4. The height has a maximum value of 375 because of the size of the image mask. The mean of each feature was subtracted from the values used in the study, and the resulting values were divided by the standard deviation. Normalising the data before feeding it into the model ensures independence from image dimensions and scale.

B. Candidate Model Comparison

Models were compared using performance indicators from confusion matrices. Balanced Accuracy (BA) is the arithmetic mean of the True Positive Rate (TPR) and the True Negative Rate (TNR), weighted equally across the positive and negative classes.

Another metric that can be used instead of accuracy to consider each class separately is the $F1$ score. The $F1$ score is the harmonic mean of the TPR and Precision, also called the Positive Predictive Value (PPV).

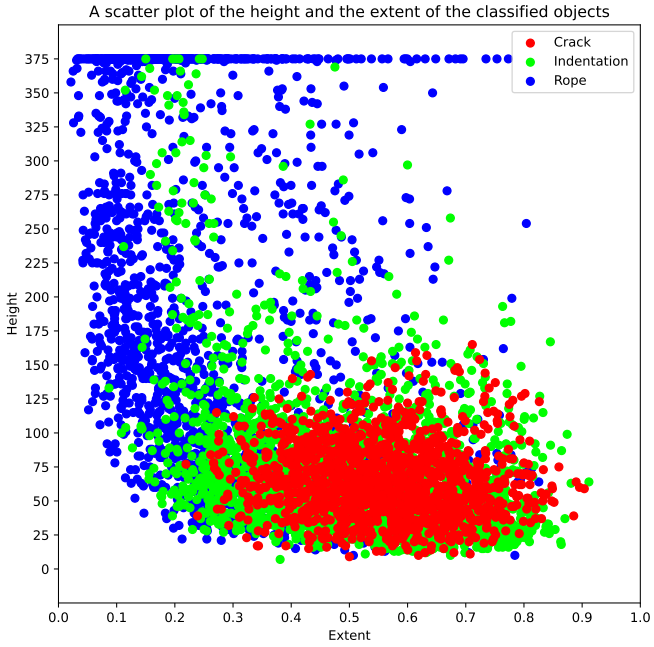


Fig. 4. A scatter plot of the height and the extent of the classified objects containing all three classes, crack (red), indentation (green), and rope (blue).

The PPV describes the proportion of true positives among all positive classifications. Similarly, the Negative Predictive Value (NPV) is the number of true negatives divided by the total number of negative classifications.

When using all available contour features as predictors, the NN method correctly classified all cracks and reported no other objects as cracks, as seen in Table III. While 100% accuracy on the test set often suggests overfitting, in this specific domain, the classes (ropes, cracks, and indentations) exhibit highly distinct geometric properties (as shown in the correlation heatmap). The stratification of the dataset further mitigates bias. However, performance may vary with different water turbidity levels or camera angles.

The NB method was the second-best regarding cracks, with all variables as predictors, achieving TPR, TNR, PPV, NPV, accuracy, BA, and an $F1$ score of 97.04%, 97.56%, 92.91%, 99.01%, 97.43%, 97.3%, and 94.93%, respectively.

When using every input variable for classifying cracks, the PLS method exhibited the worst values of 51.11%, 85.14%, 82.09%, 71.71%, and 58.6% for the TPR, NPV, accuracy, BA, and $F1$ score, respectively. With the same settings, the C5.O tree method yielded the lowest TNR and PPV values of 87.67% and 63.67%, respectively.

When using all available contour features as predictors, the NN method accurately identified all indentations and did not misidentify any objects as indentations, as seen in Table IV. The NB method was the second-best regarding indentations with all variables as predictors, with a 93.26%, 98.09%, 97.28%, 95.22%, 96.05%, 95.68%, and 95.23% TPR, TNR, PPV, NPV, accuracy, BA, and $F1$ score, respectively. The PLS

TABLE III
THE TPR, TNR, PPV, NPV, ACCURACY, BA, AND $F1$ SCORE REGARDING THE CRACK CLASS FOR EACH CANDIDATE MODEL DEVELOPED USING DIFFERENT METHODS AND ALL CONTOUR FEATURES AS PREDICTORS. THE BOLD VALUES FOR EACH COLUMN (PERFORMANCE METRIC) INDICATE THE BEST (LARGEST) VALUE.

Model	TPR	TNR	PPV	NPV	Acc	BA	$F1$
SVM	61.11	90.6	68.18	87.6	83.29	75.85	64.45
NB	97.04	97.56	92.91	99.01	97.43	97.3	94.93
RF	72.59	93.16	77.78	91.16	88.06	82.88	75.1
FDA	55.19	91.33	67.73	86.08	82.37	73.26	60.82
PLS	51.11	92.31	68.66	85.14	82.09	71.71	58.6
C5.0	65.56	87.67	63.67	88.53	82.19	76.61	64.6
NN	**100**	**100**	**100**	**100**	**100**	**100**	**100**
XGB	67.41	93.53	77.45	89.7	87.05	80.47	72.08

method consistently had the lowest value for each performance metric when predicting indentations with all available predictors.

TABLE IV
THE TPR, TNR, PPV, NPV, ACCURACY, BA, AND $F1$ SCORE REGARDING THE INDENTATION CLASS FOR EACH CANDIDATE MODEL DEVELOPED USING DIFFERENT METHODS AND ALL CONTOUR FEATURES AS PREDICTORS. THE BOLD VALUES FOR EACH COLUMN (PERFORMANCE METRIC) INDICATE THE BEST (LARGEST) VALUE.

Model	TPR	TNR	PPV	NPV	Acc	BA	$F1$
SVM	81.09	74.88	70.24	84.41	77.5	77.98	75.28
NB	93.26	98.09	97.28	95.22	96.05	95.68	95.23
RF	83.26	84.9	80.13	87.4	84.21	84.08	81.66
FDA	80.22	75.99	70.96	84.01	77.78	78.11	75.31
PLS	69.57	74.56	66.67	77.01	72.45	72.06	68.09
C5.0	72.17	82.67	75.28	80.25	78.24	77.42	73.7
NN	**100**	**100**	**100**	**100**	**100**	**100**	**100**
XGB	83.04	81.08	76.25	86.73	81.91	82.06	79.5

When using all available contour features as predictors, the NN method correctly classified all ropes and identified no other objects as ropes, as seen in Table V. The NB method was the second-best among the methods when classifying ropes with all variables as predictors, with a TPR, TNR, PPV, NPV, accuracy, BA, and an $F1$ score of 98.89%, 98.49%, 96.99%, 99.45%, 98.62%, 98.69%, and 97.93%, respectively. When using every input variable for classifying ropes, the PLS method exhibited the worst values of 85.34%, 73.77%, 84.85%, 84.59%, and 78.49% for the TNR, PPV, accuracy, BA, and $F1$ score, respectively. With the same settings, the SVM method yielded the lowest TPR and NPV values of 80.22% and 90.82%, respectively.

When using only the height and the extent as predictors, the NN and NB methods correctly classified all cracks and classified no other objects as cracks, as seen in Table VI. When using height and extent for crack classification, the PLS method yielded the worst values of 0% for TPR, PPV, and $F1$ score. The PLS method also had a TNR of 100% because it did not identify any objects as cracks. With the same settings, the RF method presented the lowest values of 83.27% and 71.53% for the TNR and accuracy, respectively. With the same settings, the SVM method presented the lowest values of 75.12% and 49.77% for the NPV and BA, respectively.

TABLE V
THE TPR, TNR, PPV, NPV, ACCURACY, BA, AND $F1$ SCORE REGARDING THE ROPE CLASS FOR EACH CANDIDATE MODEL DEVELOPED USING DIFFERENT METHODS AND ALL CONTOUR FEATURES AS PREDICTORS. THE BOLD VALUES FOR EACH COLUMN (PERFORMANCE METRIC) INDICATE THE BEST (LARGEST) VALUE.

Model	TPR	TNR	PPV	NPV	Acc	BA	$F1$
SVM	80.22	96.16	91.14	90.82	90.91	88.19	85.33
NB	98.89	98.49	96.99	99.45	98.62	98.69	97.93
RF	90.53	95.34	90.53	95.34	93.76	92.94	90.53
FDA	87.47	95.21	89.97	93.92	92.65	91.34	88.7
PLS	83.84	85.34	73.77	91.48	84.85	84.59	78.49
C5.0	87.47	92.33	84.86	93.74	90.73	89.9	86.15
NN	**100**	**100**	**100**	**100**	**100**	**100**	**100**
XGB	86.91	94.38	88.39	93.61	91.92	90.65	87.64

TABLE VI
THE TPR, TNR, PPV, NPV, ACCURACY, BA, AND $F1$ SCORE REGARDING THE CRACK CLASS FOR EACH CANDIDATE MODEL DEVELOPED USING DIFFERENT METHODS AND THE HEIGHT AND THE EXTENT AS PREDICTORS. THE BOLD VALUES FOR EACH COLUMN (PERFORMANCE METRIC) INDICATE THE BEST (LARGEST) VALUE.

Model	TPR	TNR	PPV	NPV	Acc	BA	$F1$
SVM	2.59	96.95	21.88	75.12	73.55	49.77	4.64
NB	**100**	**100**	**100**	**100**	**100**	**100**	**100**
RF	35.93	83.27	41.45	79.77	71.53	59.6	38.49
FDA	37.04	87.18	48.78	80.77	74.75	62.11	42.11
PLS	0	**100**	0	75.21	75.21	50	0
C5.0	36.67	84.62	44	80.21	72.73	60.64	40
NN	**100**	**100**	**100**	**100**	**100**	**100**	**100**
XGB	38.89	84.49	45.26	80.75	73.19	61.69	41.83

When using only the height and the extent as predictors, the NN and NB methods correctly classified all indentations, reporting no other objects as indentations, as seen in Table VII. When using the height and the extent for classifying indentations, the RF method exhibited the worst values of 60%, 68.38%, 61.89%, 61.64%, and 57.08% for the TPR, NPV, accuracy, BA, and $F1$ score, respectively. With the same settings, the SVM method yielded the lowest TNR and PPV values of 46.1% and 53.5%, respectively.

TABLE VII
THE TPR, TNR, PPV, NPV, ACCURACY, BA, AND $F1$ SCORE REGARDING THE INDENTATION CLASS FOR EACH CANDIDATE MODEL DEVELOPED USING DIFFERENT METHODS AND THE HEIGHT AND THE EXTENT AS PREDICTORS. THE BOLD VALUES FOR EACH COLUMN (PERFORMANCE METRIC) INDICATE THE BEST (LARGEST) VALUE.

Model	TPR	TNR	PPV	NPV	Acc	BA	$F1$
SVM	84.78	46.1	53.5	80.56	62.44	65.44	65.6
NB	**100**	**100**	**100**	**100**	**100**	**100**	**100**
RF	60	63.28	54.44	68.38	61.89	61.64	57.08
FDA	73.48	58.51	56.43	75.1	64.83	65.99	63.83
PLS	85	46.26	53.64	80.83	62.63	65.63	65.77
C5.0	67.61	61.05	55.94	72.05	63.82	64.33	61.22
NN	**100**	**100**	**100**	**100**	**100**	**100**	**100**
XGB	69.57	59.62	55.75	72.82	63.82	64.59	61.9

When using only the height and the extent as predictors, the NN and NB methods correctly classified all ropes and classified no other objects as ropes, as seen in Table VIII. When using height and extent for rope classification, the RF

method yielded the worst values of 88.36%, 75.57%, 83.38%, 80.81%, and 74.4% for TNR, PPV, accuracy, BA, and $F1$ score, respectively. With the same settings, the XGB method yielded the lowest TPR and NPV values of 69.36% and 86.35%, respectively.

TABLE VIII
THE TPR, TNR, PPV, NPV, ACCURACY, BA, AND $F1$ SCORE REGARDING THE ROPE CLASS FOR EACH CANDIDATE MODEL DEVELOPED USING DIFFERENT METHODS AND THE HEIGHT AND THE EXTENT AS PREDICTORS. THE BOLD VALUES FOR EACH COLUMN (PERFORMANCE METRIC) INDICATE THE BEST (LARGEST) VALUE.

Model	TPR	TNR	PPV	NPV	Acc	BA	$F1$
SVM	77.44	93.15	84.76	89.36	87.97	85.29	80.93
NB	**100**	**100**	**100**	**100**	**100**	**100**	**100**
RF	73.26	88.36	75.57	87.04	83.38	80.81	74.4
FDA	71.31	96.03	89.82	87.19	87.88	83.67	79.5
PLS	79.67	89.86	79.44	89.99	86.5	84.76	79.55
C5.0	75.49	94.93	87.99	88.73	88.52	85.21	81.26
NN	**100**	**100**	**100**	**100**	**100**	**100**	**100**
XGB	69.36	95.34	87.99	86.35	86.78	82.35	77.57

The experiment was run on the 6^{th} Generation Intel(R) Core(TM) i5-63000HQ Central Processing Unit (CPU) of a personal computer, running at a base speed of 2.3 GHz, instead of a Graphics Processing Unit (GPU), to estimate the execution time for mobile devices and built-in systems that have significantly lower computational capabilities but must perform object detection. The model training and the inference time in seconds (s) utilising the *R system.time* function for all candidate models are displayed in Table IX. The PLS and C5.0 methods are simple and require the least training and inference time. The SVM method has high computational demands due to kernel estimation in training, and the XGB method trains multiple trees for boosting.

TABLE IX
THE MODEL TRAINING AND INFERENCE TIME IN SECONDS FOR EACH CANDIDATE MODEL DEVELOPED USING DIFFERENT METHODS AND SETS OF PREDICTORS.

	All predictors			Height and extent		
Model	Train	Test	Total	Train	Test	Total
SVM	6934.05	0.33	6934.91	5795.31	0.42	5796.2
NB	265.75	2.49	268.65	77.44	0.59	78.51
RF	1171.66	0.1	1172.19	181.41	0.14	182.0
FDA	80.71	0.05	81.31	27.69	0.01	28.17
PLS	21.27	0.07	21.77	9.14	0.04	9.66
C5.0	64.94	0.12	65.47	14.94	0.11	15.72
NN	564.05	0.03	564.49	334.31	0.04	334.84
XGB	3509.78	0.02	3510.26	2811.36	0.02	2811.91

While XGBoost and SVM achieved competitive accuracy, their training times (2811.36 s and 5795.31 s, respectively) make them less suitable for rapid on-site retraining than Naive Bayes (77.44 s), which maintained equivalent accuracy.

VI. DISCUSSION

Based on the previously presented data classification results, it can be concluded that the desired automatic shape classification was achieved. However, the results vary considerably across classes.

Regarding the rope class, it is noticeable that these shapes differ from cracks and indentations, influencing the final result. Separating ropes from other objects exclusively improves accuracy. However, separating the indentations and cracks is much more challenging. If results are not compared across methods and predictor sets, but only for the best-performing model, it can be concluded that all objects can be correctly classified. In other words, automatic object classification largely depends on the choice of the classification algorithm and input variables. The predictors can be selected based on correlation, as demonstrated in this paper.

The accuracy increases with a larger set of predictors, as expected. When examining only the two most accurate methods, the differences in classification across predictor sets are insignificant. As expected, using a larger predictor set takes longer. On the other hand, the training time for the SVM, RF, NN, and XGB methods is significantly larger than for the others, regardless of the choice of predictors. Considering that the computational complexity of the NN algorithm is much higher than for the NB method, which is equally as successful when using the height and the extent as predictors, and the second most successful for the full predictor set, it can be concluded that the NB method is the better choice. Although inference time is similar across models, the NB algorithm dominates in terms of training time and accuracy on a reduced predictor set, suggesting it is the optimal choice for the final model.

VII. LIMITATIONS

This study utilised a specific dataset of underwater concrete walls. The selected features (height and extent) are robust for linear objects such as ropes and cracks, but may require augmentation with texture features for more complex biological fouling or non-linear structural damage. While the proposed Naive Bayes model utilising height and extent features achieved maximal accuracy on this test set, several limitations must be acknowledged. First, the method relies heavily on the quality of the upstream segmentation; high levels of environmental noise or turbidity, common in underwater inspection, could result in fragmented contours that misrepresent the object's true geometry. Second, the reliance on simple geometric descriptors assumes a distinct shape separation between classes (linear ropes and irregular cracks). This approach may generalise poorly to objects with complex, overlapping morphological features, such as biological fouling or nonlinear structural fractures, where deep learning models that utilise texture analysis typically excel. Finally, while the dataset is representative of concrete wall defects, deployment in varying underwater conditions would require retraining the classification thresholds to account for different camera angles and object scales.

VIII. CONCLUSION

This paper explores the performance and results of automatic image segment classification using contour features. Classification was attempted not with a single algorithm but

with various algorithms chosen to ensure representation of different classifier families. The *OpenCV* library was used to convert the classified image segments into a series of contour features. The objects were then reorganised into a training and testing dataset to ensure a similar class distribution across the split. Finally, the candidate models were trained using various methods with all predictors or a subset of the two variables most highly correlated with the object class, the contour height and extent, used as input. Results were compared using performance indicators from the confusion matrix. The process successfully identified the best candidate model for final deployment, combining the NB method and the reduced predictor set for optimal accuracy and speed. It is crucial to emphasise that, based on the test set results, complete accuracy in label identification for image segments can be achieved using the selected contour features for the crack, indentation, and rope classes. This enables unsupervised classification of additional objects identified and separated using previously developed image segmentation algorithms.

ACKNOWLEDGMENT

The authors extend their sincere gratitude to the company Vectrino [1] (https://www.vectrino.hr; Bože Milanovića 2b, 51000 Rijeka, Croatia; info@vectrino.eu) for providing the dataset that made this research possible. We also thank Deni Klen for hand-labelling the images [2].

This work was fully supported by the EU Horizon 2020 project INNO2MARE under the grant no. 101087348, Ministry of Science, Education and Youth of Croatia under the bilateral Croatian and Slovenian science project "Predicting Anomalous Trajectories Using Machine Learning", University of Rijeka projects no. uniri-mzi-25-1 and uniri-zip-2103-4-22, and IRI3 ("Povećanje razvoja novih proizvoda i usluga koji proizlaze iz aktivnosti istraživanja i razvoja", grant no. "Autonomna rješenja za nadzor i sigurnost priobalja").

REFERENCES

[1] Vectrino, "Homepage Vectrino," 2025, Accessed: 22.05.2025. [Online]. Available: https://www.vectrino.hr

[2] D. Klen and J. Lerga, "Comprehensive Annotation of Underwater Data for Image Segmentation," in *MY FIRST CONFERENCE 2024 8th Annual PhD Conference on Engineering and Technology*. University of Rijeka, 2024. [Online]. Available: https://mfc.uniri.hr/wp-content/uploads/2024/09/MFC_2024_Book_of_Abstracts.pdf

[3] A. Saberironaghi and J. Ren, "DepthCrackNet: A Deep Learning Model for Automatic Pavement Crack Detection," *Journal of Imaging*, vol. 10, no. 5, p. 100, 2024. [Online]. Available: https://doi.org/10.3390/jimaging10050100

[4] J. Wu and X. Zhang, "Tunnel Crack Detection Method and Crack Image Processing Algorithm Based on Improved Retinex and Deep Learning," *Sensors*, vol. 23, no. 22, p. 9140, 2023. [Online]. Available: https://doi.org/10.3390/s23229140

[5] S. Shim, "Self-training Approach for Crack Detection Using Synthesized Crack Images Based on Conditional Generative Adversarial Network," *Computer-Aided Civil and Infrastructure Engineering*, vol. 39, no. 7, pp. 1019–1041, 2024. [Online]. Available: https://doi.org/10.1111/mice.13119

[6] S. Kulkarni, S. Singh, D. Balakrishnan, S. Sharma, S. Devunuri, and S. Korlapati, "CrackSeg9k: A Collection and Benchmark for Crack Segmentation Datasets and Frameworks," in *Computer Vision – ECCV 2022 Workshops*. Springer Nature Switzerland, 2022, pp. 179–195. [Online]. Available: https://doi.org/10.1007/978-3-031-25082-8_12

[7] U. Orinaitė, V. Karaliūtė, M. Pal, and M. Ragulskis, "Detecting Underwater Concrete Cracks with Machine Learning: A Clear Vision of a Murky Problem," *Applied Sciences*, vol. 13, no. 12, p. 7335, 2023. [Online]. Available: https://doi.org/10.3390/app13127335

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer Nature Switzerland, 2015, pp. 234–241. [Online]. Available: https://doi.org/10.1007/978-3-319-24574-4_28

[9] C. Dechesne, P. Lassalle, and S. Lefèvre, "Bayesian U-Net: Estimating Uncertainty in Semantic Segmentation of Earth Observation Images," *Remote Sensing*, vol. 13, no. 19, p. 3836, 2021. [Online]. Available: https://doi.org/10.3390/rs13193836

[10] R. Khanam and M. Hussain, "YOLOv11: An Overview of the Key Architectural Enhancements," *arXiv preprint arXiv:2410.17725*, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2410.17725

[11] Y. Tian, Q. Ye, and D. Doermann, "YOLOv12: Attention-Centric Real-Time Object Detectors," *arXiv preprint arXiv:2502.12524*, 2025. [Online]. Available: https://doi.org/10.48550/arXiv.2502.12524

[12] N. Russel and A. Selvaraj, "MultiScaleCrackNet: A Parallel Multiscale Deep CNN Architecture for Concrete Crack Classification," *Expert Systems with Applications*, vol. 249, no. Part B, p. 123658, 2024. [Online]. Available: https://doi.org/10.1016/j.eswa.2024.123658

[13] J. Zhang, Z. Zeng, P. Sharma, O. Alfarraj, A. Tolba, and J. Wang, "A Dual Encoder Crack Segmentation Network with Haar Wavelet-Based High–Low Frequency Attention," *Expert Systems with Applications*, vol. 256, p. 124950, 2024. [Online]. Available: https://doi.org/10.1016/j.eswa.2024.124950

[14] J. Xu, S. Wang, R. Han, X. Wu, D. Zhao, X. Zeng, R. Yin, Z. Han, Y. Liu, and S. Shu, "Crack Segmentation and Quantification in Concrete Structures Using a Lightweight YOLO Model Based on Pruning and Knowledge Distillation," *Expert Systems with Applications*, vol. 283, p. 127834, 2025. [Online]. Available: https://doi.org/10.1016/j.eswa.2025.127834

[15] F. Zhang, Y. Gu, L. Yin, J. Song, C. Qiu, Z. Ye, X. Chen, and J. Wu, "Research on the Generation and Evaluation of Bridge Defect Datasets for Underwater Environments Utilizing CycleGAN Networks," *Expert Systems with Applications*, vol. 262, p. 125576, 2025. [Online]. Available: https://doi.org/10.1016/j.eswa.2024.125576

[16] M. Kuhn, "Building Predictive Models in R Using the Caret Package," *Journal of Statistical Software*, vol. 28, no. 5, pp. 1–26, 2008. [Online]. Available: https://doi.org/10.18637/jss.v028.i05

[17] M. Kuhn and K. Johnson, *Applied Predictive Modeling*. Springer New York, 2013. [Online]. Available: https://doi.org/10.1007/978-1-4614-6849-3

[18] M. Kuhn, "Caret: Classification and Regression Training," *CRAN: Contributed Packages*, 2024. [Online]. Available: https://doi.org/10.32614/CRAN.package.caret

[19] B. Boser, I. Guyon, and V. Vapnik, "A Training Algorithm for Optimal Margin Classifier," in *COLT '92: Proceedings of the fifth annual workshop on Computational learning theory*. ACM, 1992, pp. 144–152. [Online]. Available: https://doi.org/10.1145/130385.130401

[20] T. Hastie, S. Rosset, J. Zhu, and H. Zou, "Multi-Class AdaBoost," *Statistics and Its Interface*, vol. 2, no. 3, pp. 349–360, 2009. [Online]. Available: https://doi.org/10.4310/SII.2009.v2.n3.a8

[21] D. Meyer, F. Leisch, and K. Hornik, "The Support Vector Machine Under Test," *Neurocomputing*, vol. 55, no. 1-2, pp. 169–186, 2003. [Online]. Available: https://doi.org/10.1016/S0925-2312(03)00431-4

[22] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 2007, ISBN: 9780521880688.

[23] D. Hand and K. Yu, "Idiot's Bayes: Not So Stupid after All?" *International Statistical Review*, vol. 69, no. 3, pp. 385–398, 2007. [Online]. Available: https://doi.org/10.1111/j.1751-5823.2001.tb00465.x

[24] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Pearson, 2022, ISBN: 9780134610993.

[25] G. John and P. Langley, "Estimating Continuous Distributions in Bayesian Classifiers," *arXiv preprint arXiv:1302.4964*, 2013. [Online]. Available: https://doi.org/10.48550/arXiv.1302.4964

[26] G. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*. Wiley, 1992. [Online]. Available: https://doi.org/10.1002/0471725293

[27] J. Cohen, P. Cohen, S. West, and L. Aiken, *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Routledge, 2013. [Online]. Available: https://doi.org/10.4324/9780203774441

[28] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. Springer New York, 2009. [Online]. Available: https://doi.org/10.1007/978-0-387-84858-7

[29] C. Reynès, R. Sabatier, and N. Molinari, "Choice of B-splines with Free Parameters in the Flexible Discriminant Analysis Context," *Computational Statistics & Data Analysis*, vol. 51, no. 3, pp. 1765–1778, 2006. [Online]. Available: https://doi.org/10.1016/j.csda.2005.11.018

[30] D. Wetcher-Hendricks, *Analyzing Quantitative Data: An Introduction for Social Researchers*. Wiley, 2011, ISBN: 9780470526835.

[31] H. Abdi, "Partial Least Squares Regression and Projection on Latent Structure Regression (PLS Regression)," *WIREs Computational Statistics*, vol. 2, no. 1, pp. 97–106, 2010. [Online]. Available: https://doi.org/10.1002/wics.51

[32] S. Sæbø, T. Almøy, A. Flatberg, A. Aastveit, and H. Martens, "LPLS-Regression: A Method for Prediction and Classification Under the Influence of Background Information on Predictor Variables," *Chemometrics and Intelligent Laboratory Systems*, vol. 91, no. 2, pp. 121–132, 2008. [Online]. Available: https://doi.org/10.1016/j.chemolab.2007.10.006

[33] V. González, R. Giraldo, and V. Leiva, "PLS1-MD: A Partial Least Squares Regression Algorithm for Solving Missing Data Problems," *Chemometrics and Intelligent Laboratory Systems*, vol. 240, p. 104876, 2023. [Online]. Available: https://doi.org/10.1016/j.chemolab.2023.104876

[34] G. Louppe, "Understanding Random Forests: From Theory to Practice," Ph.D. dissertation, University of Liège, Faculty of Applied Sciences, Department of Electrical Engineering & Computer Science, 2014. [Online]. Available: https://doi.org/10.13140/2.1.1570.5928

[35] X. Wu, V. Kumar, R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. McLachlan, A. Ng, B. Liu, P. Yu, Z. Zhou, M. Steinbach, D. Hand, and D. Steinberg, "Top 10 Algorithms in Data Mining," *Knowledge Information Systems*, vol. 14, no. 1, pp. 1–37, 2008. [Online]. Available: https://doi.org/10.1007/s10115-007-0114-2

[36] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014, pp. 212–218. [Online]. Available: https://doi.org/10.1017/CBO9781107298019.019

[37] L. Breiman, A. Cutler, A. Liaw, and M. Wiener, "RandomForest: Breiman and Cutlers Random Forests for Classification and Regression," *CRAN: Contributed Packages*, 2024. [Online]. Available: https://doi.org/10.32614/CRAN.package.randomForest

[38] T. Ho, "Random Decision Forests," in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, vol. 1. IEEE, 1995, pp. 278–282. [Online]. Available: https://doi.org/10.1109/ICDAR.1995.598994

[39] Y. Amit and D. Geman, "Shape Quantization and Recognition with Randomized Trees," *Neural Computation*, vol. 9, no. 7, pp. 1545–1588, 1997. [Online]. Available: https://doi.org/10.1162/neco.1997.9.7.1545

[40] R. Caruana and A. Niculescu-Mizil, "An Empirical Comparison of Supervised Learning Algorithms," in *ICML '06: Proceedings of the 23rd international conference on Machine learning*. ACM, 2006, pp. 161–168. [Online]. Available: https://doi.org/10.1145/1143844.1143865

[41] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 785–794. [Online]. Available: https://doi.org/10.1145/2939672.2939785

[42] A. Brahme, Ed., *Comprehensive Biomedical Physics*. Elsevier, 2014, ISBN: 9780444536334.

[43] J. Olden and D. Jackson, "Illuminating the "Black Box": A Randomization Approach for Understanding Variable Contributions in Artificial Neural Networks," *Ecological Modelling*, vol. 154, no. 1-2, pp. 135–150, 2002. [Online]. Available: https://doi.org/10.1016/S0304-3800(02)00064-9

[44] C. Bishop, *Pattern Recognition and Machine Learning*. Springer New York, 2006, ISBN: 9780387310732.

[45] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer New York, 2013. [Online]. Available: https://doi.org/10.1007/978-1-4757-3264-1

[46] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[47] P. Probst, A. Boulesteix, and B. Bischl, "Tunability: Importance of Hyperparameters of Machine Learning Algorithms," *Journal of Machine Learning Research*, vol. 20, no. 53, pp. 1–32, 2019. [Online]. Available: https://jmlr.org/papers/v20/18-444.html

[48] M. Schwab, A. Mayr, and M. Haltmeier, "Deep Gaussian Mixture Model for Unsupervised Image Segmentation," in *Machine Learning, Optimization, and Data Science: 10th International Conference, LOD 2024*. Springer Nature Switzerland, 2024, pp. 339–352. [Online]. Available: https://doi.org/10.1007/978-3-031-82484-5_25

[49] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning," in *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, vol. 48. PMLR, 2016, pp. 1050–1059. [Online]. Available: https://proceedings.mlr.press/v48/gal16.html

[50] T. Fisher, H. Gibson, Y. Liu, M. Abdar, M. Posa, G. Salimi-Khorshidi, A. Hassaine, Y. Cai, K. Rahimi, and M. Mamouei, "Uncertainty-aware interpretable deep learning for slum mapping and monitoring," *Remote Sensing*, vol. 14, no. 13, p. 3072, 2022. [Online]. Available: https://doi.org/10.3390/rs14133072

[51] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. Berg, W. Lo, P. Dollár, and R. Girshick, "Segment Anything," in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, 2023, pp. 3992–4003. [Online]. Available: https://doi.org/10.1109/ICCV51070.2023.00371

[52] N. Ravi, V. Gabeur, Y. Hu, R. Hu, C. Ryali, T. Ma, H. Khedr, R. Rädle, C. Rolland, L. Gustafson, E. Mintun, J. Pan, K. Alwala, N. Carion, C. Wu, R. Girshick, P. Dollár, and C. Feichtenhofer, "SAM 2: Segment Anything in Images and Videos," *arXiv preprint arXiv:2408.00714*, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2408.00714

[53] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2016, pp. 770–778. [Online]. Available: https://doi.org/10.1109/CVPR.2016.90

[54] A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv preprint arXiv:1704.04861*, 2017. [Online]. Available: https://doi.org/10.48550/arXiv.1704.04861

[55] J. Sklansky, "Finding the Convex Hull of a Simple Polygon," *Pattern Recognition Letters*, vol. 1, no. 2, pp. 79–83, 1982. [Online]. Available: https://doi.org/10.1016/0167-8655(82)90016-2

[56] A. Fitzgibbon and R. Fisher, "A Buyer's Guide to Conic Fitting," in *BMVC '95: Proceedings of the 6th British conference on Machine vision (Vol. 2)*. BMVA Press, 1995, pp. 513–522. [Online]. Available: https://doi.org/10.5244/C.9.51

Statistical Properties of GNSS Positioning Errors

1st Lana Miličević
Faculty of Civil Engineering
University of Zagreb
Zagreb, Croatia
lana.milicevic10@gmail.com

***Abstract*—Using the Global Navigation Satellite System (GNSS) with high accuracy remains challenging in urban environments, where positioning performance is degraded by the multipath effect. This recognised source of error in high-accuracy GNSS applications is caused by buildings, vegetation, and other obstacles in satellite signal propagation. As demand for precise positioning increases in various location-based applications, it has become essential to understand and avoid systematic positioning errors in order to design reliable navigation systems.**

This study examines the statistical properties of GNSS positioning errors in urban environments, with the aim of improving the understanding of their causes and mitigation. GNSS positioning data was collected at two distinct, multipath-prone locations. RTKLIB software was used to generate positioning solutions for satellite elevation mask angles ranging from 0° to 40°, in increments of 5°. A practical upper limit of 40° was chosen to balance signal quality and satellite availability for meaningful analysis, as such high mask angles leave very sparse data to work with, introducing substantial error. The data were analysed using a statistical process developed in the R programming language. The analysis revealed that there are optimal mask angles for minimising horizontal positioning errors, with the optimal angle varying by environment. Systematic directional bias was observed in positioning errors across all mask angle settings. Circular statistical analysis using the Rayleigh test confirmed high directional consistency, with concentration values exceeding 0.94 and p-values indicating statistically significant non-uniform distributions.

The findings of this study show that environmental multipath effects create systemic positioning biases that are largely independent of satellite elevation filtering strategies, highlighting the need for the development of error mitigation strategies that go beyond conventional methods - integrated solutions combining predictive modelling with advanced, environment-specific filtering techniques.

***Index Terms*—Multipath, GNSS positioning, Elevation mask angle, Error directionality**

I. Introduction

Accurate and reliable positioning is vital for applications ranging from autonomous vehicles and surveying to urban navigation and location-based services. The Global Navigation Satellite System (GNSS) has made worldwide positioning possible and widely available, but its performance is still challenged in urban environments, largely due to the multipath effect. The multipath effect is to date among the most prominent factors influencing the accuracy of positioning using Global Navigation Satellite Systems (GNSS), and is the dominant error source in GNSS applications that require high accuracy. [1] This phenomenon arises when signals from satellites are reflected or scattered by buildings, trees, and other obstacles. Signal reflection and diffraction caused by these obstacles results in the signal travelling in multiple paths to the receiver. Due to that, the signal reaches the reciever delayed in comparison to an ideal path, and thereby introduces error to the receiver's measurements for determining its position. Because low-elevation signals are especially susceptible to reflections from vertical surfaces and ground bounce, practical receivers commonly rely on filtering out the signals from sattelites at lower angles (also known as mask angle), as a first line of defence. Studies on mask angle selection indicate that raising the cutoff can reduce low-elevation multipath, but may also degrade accuracy by reducing satellite availability. Optimal mask values depend strongly on the local environment and application. [2] Other work has moved beyond static mask tuning towards site-specific elevation cutoff models derived from fisheye imagery or 3D urban models, as well as integrated multipath mitigation schemes that combine elevation masks with advanced signal processing and environmental modelling. [3]

This study explores the statistical properties of GNSS positioning errors in urban settings, with a focus on assessing mask angle optimisation and statistically characterising positioning error. This approach helps understand how environmental factors and satellite mask angle choices interact to create predictable patterns of error.

GNSS positioning readings were conducted during a 15-minute time frame in environments prone to multipath errors - next to reflecting buildings, and underneath a tree. These measurements are subsequently used in analysing how varying the satellite mask angle affects positioning precision, and to find an optimal mask angle as a trade-off between excluding low-quality signals and preserving adequate satellite coverage.

II. Methodology

A. Data acquisition and processing

GNSS positioning data was collected using the GNSSLogger tool at the two aforementioned environments - next to reflective buildings and under a tree canopy. Each recording lasted approximately 15 minutes produced RINEX observation logs in the `.25o` format.

The data was processed using RTKLIB, an open-source GNSS analysis package. Processing requires the package's `rtkpost` tool, which takes a broadcast ephemeris file (`.brdc`) containing satellite orbital and clock information, together with the observation log file (`.25o`) containing raw

pseudorange and carrier-phase measurements. RTKLIB then outputs a positioning solution file (`.pos`), which provides estimated receiver coordinates under specified processing conditions. Positioning solutions were generated for elevation mask angles between 0° and 40° in 5° increments. The positioning solution was computed using single-mode positioning, Broadcast ionosphere correction, Saastamoinen troposphere correction, and the signals from all available satellite constellations. The corresponding navigation file (`2025/brdc1360.25n`) used in the positioning solutions calculation was obtained from EarthData.

The final dataset consists of positioning solutions for each measuring station and mask angle. The positioning solutions contain time-stamped position estimates and associated quality indicators derived from the pos format - GPS time (`gpst`), latitude (`lat`), longitude (`lon`), solution quality (`q`), number of satellites (`ns`), error standard deviations (`sdn`, `sdne`, `sdeu`, `sdun`), solution age (`age`), ambiguity resolution ratio (`ratio`). The mask angle and station are marked alongside each positioning solution entry. The reference elevations and coordinates of the measurement points were obtained and verified using the OpenStreetMap and Open Elevation services.

B. *Programmatic statistical analysis*

Programmatic analysis in relation to mask angles and the multipath effect was performed using the R programming language and various specialised libraries for data management, geospatial computation and statistics (Tidyverse, Geosphere, Circular, CircStats, and Purrr). The first step of the analysis aims to determine the optimal mask angle that minimises horizontal positioning errors for each scenario, if such exists. To describe positioning accuracy, the R script calculates horizontal distance errors calculated as the great-circle distance between each estimated position and the site reference using the Haversine distance formula.

Should a global optimal mask angle exist such that horizontal error values smoothly decrease with mask angle towards it and increase afterwards, we continue to examine the existence and statistical properties of the directionality of positioning errors in terms of their azimuthal direction. The directionality of the error is analysed by computing the azimuth bearing of the error vector from the true location for each position reading. The distribution of these error azimuths is visualised using polar histograms at each mask angle to determine whether errors tend to have a preferred direction that indicates the nature and position of the environmental obstruction. The azimuth data is statistically analysed using circular statistical techniques such as the circular mean direction and mean resultant length, and applying the Rayleigh test to assess angular uniformity. Summary statistics of these metrics are aggregated for each mask angle and site to support the interpretation of GNSS error nature and structure in different conditions.

III. Results and Discussion

A. *Horizontal positioning error*

The programmatic analysis of the elevation mask angle confirms that an optimal one exists in terms of minimising the horizontal positioning error. This optimal angle is a point at which a balance is reached: increasing the mask angle from zero degrees initially reduces error by excluding low-quality signals affected by environmental multipath interference. However, increasing the angle beyond the optimum point increases error as useful satellite signals are filtered out. The results confirmed this satellite signal propagation behaviour in both environments (Fig. 1). Under the tree canopy, the minimum horizontal error occurred at 20°, equaling 46.3 m. Near the building, the optimum was reached at 10°, with the horizontal positioning error equaling 35.4 m.

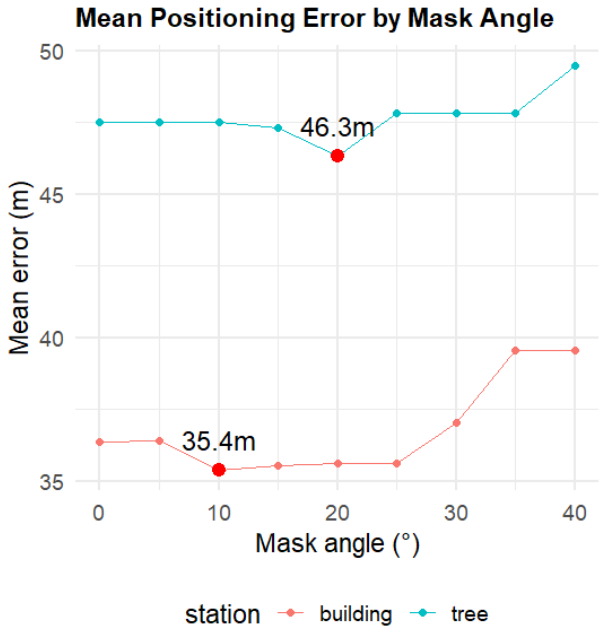


Fig. 1: Mean horizontal error across mask angles varying from 0 to 40 degrees. The plot shows the existence of an mask angle optimising the horizontal positioning error at both measuring stations. Mask angle values lower than the optimal don't filter out multipath noise, whereas mask angle values higher than the optimal begin filtering out useful satellite signals.

The existence of an optimal mask angle is consistent with the physics of multipath propagation, in which amplitude and phase distortions depend strongly on satellite elevation and the geometry between receiver, satellite, and surrounding surfaces. At low elevations, reflected signals arrive with delays and phase shifts similar to the direct signal, amplifying interference, while at higher elevations the likelihood and severity of such reflections diminishes. The elevation mask functions as a spatial filter, discarding low-quality signals dominated by multipath and producing an optimal trade-off between signal quality and availability. [4]

Fig. 2: North and East components of the ENU error per site, hinting at the deterministic nature of horizontal error directions. The errors demonstrate a stable general direction at all mask angles, but varying variance.

B. Error directionality

Having confirmed that an optimal mask angle in areas under the influence of the multipath effect exists, the directionality of the error was examined to check for systematic bias within positioning errors. Plotting the North and East components of the computed ENU error per mask angle and site reveals that the error directions are indeed not random and remain having a stable general direction for an observed station across all mask angle settings (Fig. 2). Near the building, errors predominantly point towards the southwest, while under the tree, they consistently point eastward. Systematic directionality suggests that the persistent bias created by multipath effects and satellite geometry cannot be eliminated solely by adjusting the mask angle, though it may be partially mitigated.

A further distinction comes to light when comparing the geometry and nature of the measurement site environments. Positioning errors were consistently higher in the canopy environment than in the building scenario, suggesting the former environment causes more severe signal degradation than the latter. Generally, multipath effects differ fundamentally between built and vegetated environments. Buildings introduce strong, specular reflections that create systematic and repeatable positioning biases, particularly evident in azimuthal error clustering. By contrast, vegetation tends to attenuate and scatter signals, producing more diffuse multipath and increased noise, leading to less predictable and greater variablity positioning errors. [1]

The directional errors are further quantified in the form of polar histograms (Fig. 3). One polar histogram is plotted by each mask angle to show how the spread of errors changes as the angle varies. The visualisations demonstrate error distributions with pronounced peaks in specific directions. For the building location, the peak error azimuth remained consistently around 202° (SSW) for lower mask angles, shifting slightly westwards to approximately 225° (SW) for higher mask angles (35° and 40°). Under the tree, the peak error direction remained stable at approximately 90° (E) across all mask angles. Furthermore, all graphs show that the error is narrower around the optimum and for lower mask angles, while larger mask angles cause the error to spread more widely.

C. Statistical validation of directional bias

Utilising circular statistical analysis based on the Rayleigh test for uniformity, quantitative evidence of the systematic nature of the observed errors is obtained. Two key metrics reported on in Table I are:

- **Concentration** ($\bar{r}$) - measures how tightly the error directions cluster around the mean direction. A value of 1 indicates that all errors point in the same direction, while 0 corresponds to a completely random (uniform) distribution.
- **p-value** - tests the null hypothesis that the observed azimuths are uniformly distributed (i.e., no preferred direction). Small p-values indicate that the distribution is unlikely to be uniform.

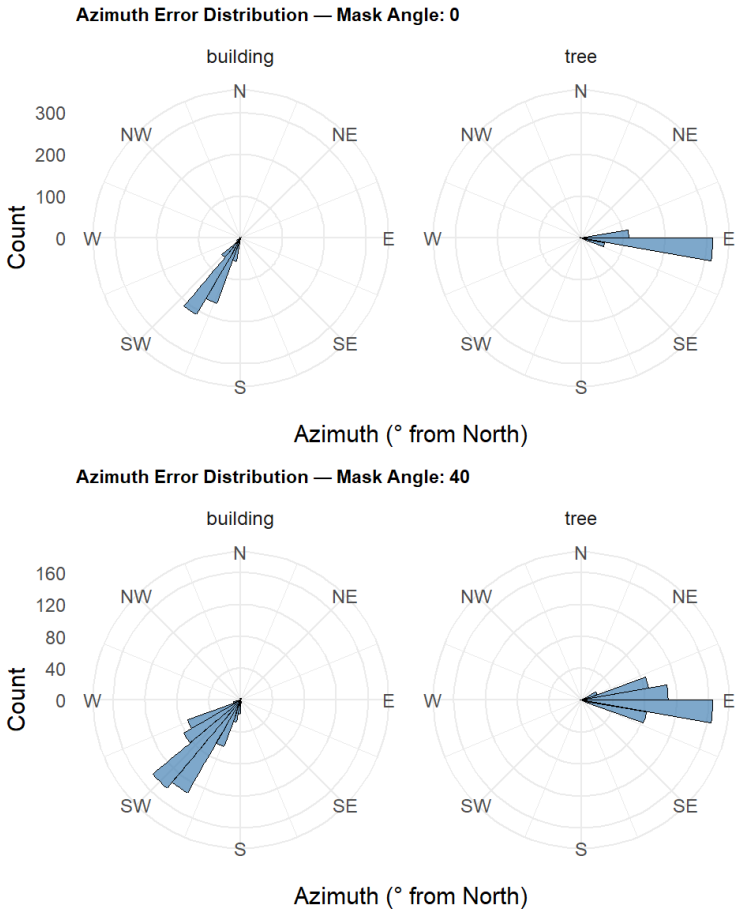


Fig. 3: Chosen histograms of azimuth error at minimum and maximum value mask angles (0° and 40°, respectively). The errors retain azimuthal direction across mask angles at each site. The error spread is wider at higher mask angles, indicating higher error directionality variance.

TABLE I: Summary of azimuthal bias statistics, including sample sizes, mean azimuths, correlation coefficients, and uniformity test p-values, for different mask angles across building and tree environments. Consistent high concentration values and low p-values confirm statistically significant error directionality.

MA (°)	Station	n	Mean Azimuth (°)	$\bar{r}$	p_{uniform}
0	building	511	212	0.972	2.81×10^{-210}
5	building	518	212	0.972	3.59×10^{-213}
10	building	546	210	0.948	1.11×10^{-213}
15	building	560	209	0.948	2.79×10^{-219}
20	building	572	209	0.949	2.01×10^{-224}
25	building	572	209	0.948	7.08×10^{-224}
30	building	587	210	0.983	4.25×10^{-247}
35	building	558	221	0.959	2.07×10^{-223}
40	building	558	221	0.959	1.87×10^{-223}
0	tree	484	94.1	0.996	2.42×10^{-209}
5	tree	484	94.1	0.996	2.38×10^{-209}
10	tree	481	94.4	0.996	4.00×10^{-208}
15	tree	474	95.2	0.995	2.51×10^{-204}
20	tree	498	91.6	0.986	3.78×10^{-211}
25	tree	551	88.9	0.984	1.47×10^{-232}
30	tree	551	88.9	0.984	1.47×10^{-232}
35	tree	551	88.9	0.984	1.47×10^{-232}
40	tree	474	89.4	0.980	1.93×10^{-198}

Under the tree canopy, concentration values ($\bar{r}$) ranged from 0.98 to 0.996, indicating extremely high directional consistency, while p-values were uniformly very low, statistically rejecting the hypothesis of random azimuthal distribution.

Concentration values in the vicinity of the building remained high (0.948-0.983) with consistently low p-values, thereby confirming significant directional preference in error patterns. The statistical measures demonstrate that, irrespective of mask angle selection, error direction remains highly clustered, not random, and strongly dependent on the local environment, thereby supporting the hypothesis that multipath effects are affected by environment geometry. The persistence of these directional patterns across mask angle variations suggests that environmental multipath effects create systematic biases that are largely independent of satellite elevation filtering. This indicates that optimal GNSS accuracy in challenging environments requires site-specific calibration or advanced processing techniques beyond simple mask angle optimisation.

It is important to note that histogram peaks indicate the most frequent error direction, whereas the summary statistics report on the circular mean direction. The two measures are related

but not identical: the mean azimuth reflects the average over all samples, while the histogram peak highlights the mode, i.e. the direction where errors are most concentrated.

IV. Conclusions

Multipath-induced errors are particularly critical for high-precision GNSS applications, and are prone to systematic biases in positional estimates. Simple mitigation tools, such as adjusting the elevation mask angle to exclude low-elevation satellites that are more susceptible to multipath, offer only partial solutions. [5], [6] Optimising the mask angle may reduce average error, but also risks discarding valuable signals and does not eliminate biases that are specific to the local environment's geometry. [4]

The presented investigation revealed that the angle of the satellite elevation mask exerts a non-linear influence on the accuracy of GNSS positioning, with optimal mask angles varying between environmental contexts (in this case, the building and the tree canopy site). Further increases in mask angle diminish accuracy, as excluding beneficial satellites outweighs the benefits of rejecting multipath-affected ones.

An interesting finding is the presence of a strong, systematic directional bias in positioning errors that persists across all mask angle settings. Errors oriented towards the southwest near the building and a consistent eastward bias under the tree canopy suggest that environmental multipath effects generate systematic positioning shifts that cannot be mitigated by elevation masking alone. However, increasing the mask angle does abandon good satellites, resulting in wider scatter. High directional concentration values ($\bar{r} > 0.94$) and statistically significant non-uniform distributions confirm that directional biases are predictable and environment-specific, rather than being caused by random measurement noise.

The magnitude of the observed errors (35 and 46 metres at optimal mask angles), together with the persistence of systematic directional bias, indicate that precision navigation technology must go beyond basic elevation masking to address both the magnitude errors and directional bias that arise systematically in obstructed settings due to the multipath effect presence. Such GNSS positioning errors can be partially reduced by using advanced signal processing techniques, environmental modeling, or multi-sensor fusion approaches. Contemporary strategies include: (i) integrating environmental models (such as LIDAR or 3D city models), (ii) employing statistical multipath mitigation algorithms (e.g. sparse estimation models), (iii) hybrid programmatic solutions requiring hardware support (e.g. antenna array beamforming, coupled amplitude delay lock loop), and other. [6], [7] State-of-the-art development prioritises physics-based multipath models that exploit detailed environmental data to predict and correct systematic biases [7], [8]. At the same time, the use of more robust statistical metrics such as RMSE or quantile-based measures improves the capturing of variability and extreme errors. [9] Validation across temporal conditions and satellite constellations is important to distinguish persistent environmental effects from transient geometric factors [10]. Ultimately, progress in this domain will depend on integrated solutions that combine predictive modelling with advanced filtering and sensor fusion techniques to mitigate the systematic biases inherent to specific environments [5], [11].

References

[1] B. M. Hannah, *Modelling and simulation of GPS multipath propagation*. PhD thesis, Queensland University of Technology, 2001.

[2] G. A. McGraw, P. D. Groves, and B. W. Ashman, "Robust positioning in the presence of multipath and nlos gnss signals," *Position, navigation, and timing technologies in the 21st century: integrated satellite navigation, sensor systems, and civil applications*, vol. 1, pp. 551–589, 2020.

[3] W. W. Wen and L.-T. Hsu, "3d lidar aided gnss nlos mitigation in urban canyons," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 18224–18236, 2022.

[4] P. J. Teunissen, O. Montenbruck, *et al.*, *Springer handbook of global navigation satellite systems*, vol. 10. Springer, 2017.

[5] J. Zhu, H. Zhou, Z. Wang, and S. Yang, "Improved multi-sensor fusion positioning system based on gnss/lidar/vision/imu with semi-tight coupling and graph optimization in gnss challenging environments," *IEEE Access*, vol. 11, pp. 95711–95723, 2023.

[6] P. M. C. Pereira, H. D. M. da Silva, and C. M. G. S. Lima, "Advancements in multipath mitigation for gnss receivers: Review of channel estimation techniques," *Space: Science & Technology*, vol. 5, p. 0278, 2025.

[7] J. Lesouple, T. Robert, M. Sahmoudi, J.-Y. Tourneret, and W. Vigneau, "Multipath mitigation for gnss positioning in an urban environment using sparse estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1316–1328, 2018.

[8] P. Closas, C. Fernandez-Prades, and J. A. Fernandez-Rubio, "A bayesian approach to multipath mitigation in gnss receivers," *IEEE journal of selected topics in signal processing*, vol. 3, no. 4, pp. 695–706, 2009.

[9] R. Sieradzki and J. Paziewski, "Analyzing the stochastic properties of code observation using various low-cost gnss receivers," *GPS Solutions*, vol. 29, no. 4, p. 168, 2025.

[10] R. C. Littell, "Statistical analysis of experiments with repeated measurements," *HortScience*, vol. 24, no. 1, pp. 37–40, 1989.

[11] M. Elsanhoury, J. Koljonen, P. Välisuo, M. Elmusrati, and H. Kuusniemi, "Survey on recent advances in integrated gnsss towards seamless navigation using multi-sensor fusion technology," in *Proceedings of the 34th international technical meeting of the satellite division of the institute of navigation (ION GNSS+ 2021)*, pp. 2754–2765, 2021.

Application of Copernicus Marine Data for Planning Offshore Renewable Energy Projects: A Case Study in the Adriatic Sea

Marta Alvir
Faculty of Engineering, University of Rijeka
Rijeka, Croatia
marta.alvir@uniri.hr

Ivana Lučin
Faculty of Engineering, University of Rijeka
Rijeka, Croatia

Martina Ivić
Faculty of Engineering, University of Rijeka
Rijeka, Croatia

Zvonimir Mrle
Faculty of Engineering, University of Rijeka
Rijeka, Croatia

Matej Mališa
Faculty of Engineering, University of Rijeka
Rijeka, Croatia

Stefan Ivić
Faculty of Engineering, University of Rijeka
Rijeka, Croatia

***Abstract*—Effects of climate change and electricity demand growth, increased the need for renewable energy sources. Hence, there is more focus on planning and construction of offshore renewable energy projects, which have a higher cost, but also higher power production compared to onshore energy. Offshore renewable energy projects have a strong impact on the marine environment and the local community, due to their large size, high price and complex construction. Therefore, adequate planning of their location and technology is necessary. Accordingly, this research presents the application of Copernicus Marine Data as a freely available data source of physical conditions. Results are compared for four pilot locations of potential offshore wind farms in the Adriatic Sea, indicating relatively small changes in most parameters. Furthermore, limitation of this type of dataset is discussed.**

***Index Terms*—Copernicus Marine Data, Offshore Renewable Energy, Adriatic Sea, Physical Conditions, Site Selection**

I. Introduction

Global population growth and civilisation advances result in exponential growth in energy demand [1], which causes an increase in greenhouse gas emissions. For dealing with this problem and combating climate change, renewable energy is becoming crucial. Technological advances and supportive policies worldwide are accelerating the deployment of renewable energy, making it more accessible and cost-effective. Offshore renewable energy sources are becoming increasingly important due to more consistent energy generation and the possibility for larger-scale projects. The European Commission's Offshore Renewable Energy Strategy [2] aims to reach 61 GW in the European Union from offshore renewable energy (wind and ocean energy) by 2030, and 340 GW by 2050.

To achieve these goals, choosing an optimal location and detailed planning are important for maximising energy generation, reducing maintenance costs, and minimising environmental effects. One of the factors that is considered during decision making is a comprehensive dataset of physical conditions, including sea currents, waves, salinity, temperature, sea surface height and others. These datasets can be expensive to obtain, especially for large areas at sea due to the high costs of marine operations and equipment.

One of the options in preliminary location planning is to use databases such as Copernicus Marine Service (CMEMS), which provides free, open, and high-resolution marine data products [3]. In research [4], CMEMS data were implemented for maritime spatial planning and a decision support system by monitoring and assessing ocean conditions, while [5] analysed the implementation of CMEMS data for determining wave energy potential and spectral variability. Previous research [6] showed a high degree of accuracy of the CMEMS data compared to the wind speed and significant wave height measured at buoys, indicating that this CMEMS dataset can be applicable in studies focusing on offshore wind potential.

The objective of this study is to demonstrate CMEMS oceanographic products for offshore renewable-energy planning, with a specific focus on offshore wind farms. This research presents an overview of the possible application of data from CMEMS, including wave parameters, sea surface height, sea currents, salinity and sea temperature for planning offshore renewable energy and the implementation of such data for comparison of possible wind-farm locations in the Adriatic Sea.

II. METHODOLOGY

A. Study Area

Due to the increasing need for renewable sources and the development of technology, new locations for offshore renewable energy projects are being investigated, which have

not been considered. Because of the lower average wind speeds and greater sea depth compared to the sea in the northern part of Europe, higher cost and less expected energy output is expected, therefore offshore wind farms have not been set up in the Adriatic Sea. However, through project BEYOND (Blue Economy sYnergies fOr sustaiNable Development) funded by the Interreg Italy-Croatia 2021-2027 program, potential locations are being investigated. The aim is to prioritise electricity production and the synergistic benefits of blue economy sectors, including tourism, fisheries and marine conservation. Four pilot locations are presented in Fig. 1. Two locations are in Italy (Veneto and Apulia region) and two in Croatia (Istria and Dalmatia region).

Fig. 1. Pilot locations. The purple circle presents the Veneto region, the orange Istria region, the yellow Dalmatia region and the green Apulia region.

B. Data Source

The Copernicus program provides open and free services based on data from airborne, seaborne and ground-based sources, with a combination of data from satellites and numerical models implemented for Earth observation. It is managed by the European Commission, in partnership with the Member States, European Space Agency and the European Organisation for the Exploitation of Meteorological Satellites and other organisations. The program goal is to provide open, free and continuous data with a wide range of applications, including mitigation of climate change effects, sustainable environment management and civil security.

Part of it is CMEMS, which provides data about oceans' dynamics, physical state, and marine ecosystems on a global scale and for European regional seas, including the Mediterranean Sea. CMEMS offers products such as near-real-time data and forecast data, which can be used for planning the construction, installation, and maintenance of offshore renewable energy projects, as well as reanalysis of data from previous periods, to support site and design selection.

For planning offshore wind farms and comparing locations, it is crucial to analyse the physical conditions used for choosing optimal locations and optimising wind turbine layout. The most crucial physical natural conditions for maximisation of energy production and minimisation of cost are wind speed and sea depth. However, additional physical conditions should be considered, including sea temperature, salinity, sea currents, sea surface height and wave conditions, which can be obtained from CMEMS.

Due to the influence on operational efficiency of turbines and maintenance, seawater temperature should be analysed while planning the location of offshore renewable energy projects. Changes in air density due to seawater temperature reduce wind energy production, affect corrosion and material degradation. Moreover, it can cause thermal contraction and expansion of material, which can influence the stability of structures. Salinity can also affect the corrosion of metal components in wind turbines; therefore, proper coating and material selection should be determined based on expected values. Large magnitudes of sea currents exert hydrodynamic forces on wind turbine foundations, threatening stability and creating scour effects. Sea currents are able to increase the risk of damage to subsea power cables. The stability of wind turbines can be impacted by waves, where parameters usually analysed are the significant wave height and the wave mean period. Therefore, wind turbine structures should be designed to withstand wave impact, fatigue stress and cyclic loading. Frequency analysis of extreme wave events is important for the installation and maintenance of offshore renewable energy projects.

C. Data Processing

Case Study analyses of physical conditions for four pilot locations in the Adriatic Sea are based on Mediterranean Sea - High Resolution L4 Sea Surface Temperature Reprocessed [7], Mediterranean Sea Physics Reanalysis [8], Mediterranean Sea Waves Reanalysis [9]. Mediterranean Sea - High Resolution L4 Sea Surface Temperature Reprocessed is obtained by merging multi-sensor (L3S) and optimally interpolated (L4) satellite-based estimates. Mediterranean Sea Physics Reanalysis uses the NEMO (Nucleus for European Modelling of the Ocean) model for ocean modelling with assimilation of satellite altimetry tracks, salinity profiles and in-situ temperature using the OceanVar 3DVar scheme. Mediterranean Sea Waves Reanalysis uses the WAM (Mediterranean Wave Forecast) wave model with an assimilation data scheme of satellite observations for significant wave height. Information about the analysed data is presented in Tab. I. Data about the temporal extent in the table refers to data valid at the time of conducting research. All data are 2D results except sea currents and salinity.

Data is obtained for 10-year periods from January 1, 2013, to December 31, 2024. Variables from the Mediterranean Sea Physics Reanalysis dataset can also be obtained as hourly, monthly or yearly, while Mediterranean Sea Waves Reanalysis can also be obtained as monthly values; however, it affects a

TABLE I
ANALYSED PHYSICAL CONDITIONS FOR PLANNING OFFSHORE RENEWABLE ENERGY PROJECTS

Physical condition	Variables	Unit	Spatial resolution	Temporal resolution	Temporal extent	Dataset
Sea temperature	analysed_sst	K	0.05° × 0.05°	Daily	1982-present	[7]
Salinity	so	psu	0.042° × 0.042°	Daily	1987-present	[8]
Sea currents	uo and vo	m/s	0.042° × 0.042°	Daily	1987-present	[8]
Sea surface height	zos	m	0.042° × 0.042°	Daily	1987-present	[8]
Significant wave height	VHM0	m	0.042° × 0.042°	Hourly	1985-present	[9]
Wave mean period	VTM10	s	0.042° × 0.042°	Hourly	1985-present	[9]

TABLE II
COMPARISON OF SEA SURFACE TEMPERATURE FOR PILOT LOCATIONS.

	Veneto region	Istria region	Dalmatia region	Apulia region
Minimum	8 °C	10.5 °C	12.9 °C	12.6 °C
Maximum	30 °C	28.9 °C	28.3 °C	28.6 °C
Average	17.5 °C	18.4 °C	19.3 °C	18.9 °C

TABLE III
COMPARISON OF SALINITY FOR PILOT LOCATIONS.

	Veneto region	Istria region	Dalmatia region	Apulia region
Minimum	34 ppt	35.4 ppt	36.7 ppt	36.8 ppt
Maximum	38.5 ppt	39.1 ppt	39.2 ppt	39.1 ppt
Average	36.7 ppt	38.15 ppt	38.5 ppt	38.4 ppt

different range of temporal extent. Although there are several ways to download data from CMEMS, including the API Copernicus Marine Toolbox, this research uses the MyOcean Pro viewer [10] to download data in NetCDF file format and analyse it using the Python programming language. Values of temperature are recalculated as Celsius instead of Kelvin. For data processing, several Python modules were used, including numpy, matplotlib, netCDF4 and datetime.

For each location, values are spatially averaged, considering that each analysed area consists of several points of CMEMS data. Spatial resolution of points is around 4-5 kilometres, which is suitable for the initial stage of planning, but for more detailed analysis, other modelling and measurement techniques should be employed. The biggest limitation of this data is the inability to capture small-scale processes, such as local vorticity and river plumes, especially when complex bathymetry and coastal shape are present.

III. RESULTS

A. Sea Surface Temperature

In Tab. II comparison of sea surface temperature is presented. It can be observed that the lowest temperature was in the Veneto region, which is expected due to the influence of cooler water from the river Po. Considering that this location has the smallest sea depth, the highest sea surface temperature can be observed with the additional influence of the direction of sea currents circulation in the Adriatic Sea. The Istria region is also under the influence of the shallower sea and the river Po, but a smaller river influence can be observed than in the Veneto region due to a greater distance. Southern pilot regions in the Adriatic Sea have very similar values. Overall, it can be concluded that sea surface temperature has a small difference between locations, with an average value of around 18.5 °C.

B. Salinity

Tab. III-B represents the minimum, maximum and average values of salinity for each pilot location. The smallest values of salinity are in the Veneto region, also due to the influence of the river Po. It can be observed that other locations have similar values of salinity. Data show slightly lower values in the Istria region due to the proximity of the Po River. Salinity strongly depends on season; therefore, the highest values of salinity are expected during the summer.

C. Sea Currents

Fig. 2 represents a rose diagram of sea currents for the potential location of an offshore wind farm in the Veneto region. As can be observed, sea currents are predominantly south and southwest directions, with maximum magnitudes up to 0.2 m/s and average values of 0.04 m/s.

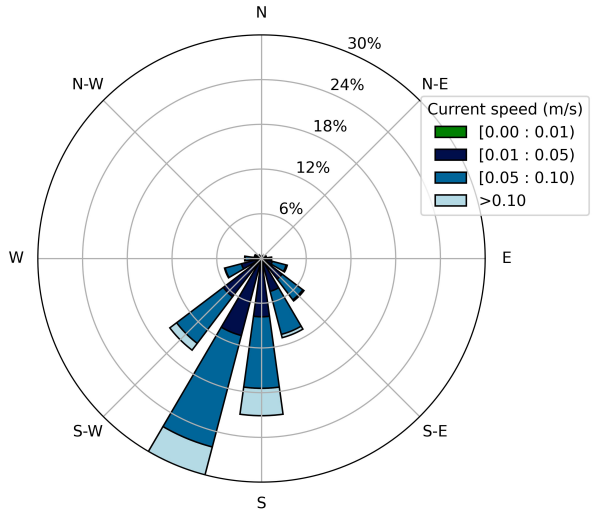


Fig. 2. Sea currents rose diagram for the potential location of an offshore wind farm in the Veneto region, Italy.

From Fig. 3, which represents a rose diagram of sea currents for the Istria region, the south and southeast direction can be observed with an average value of 0.03 m/s and a maximum observed value of 0.5 m/s.

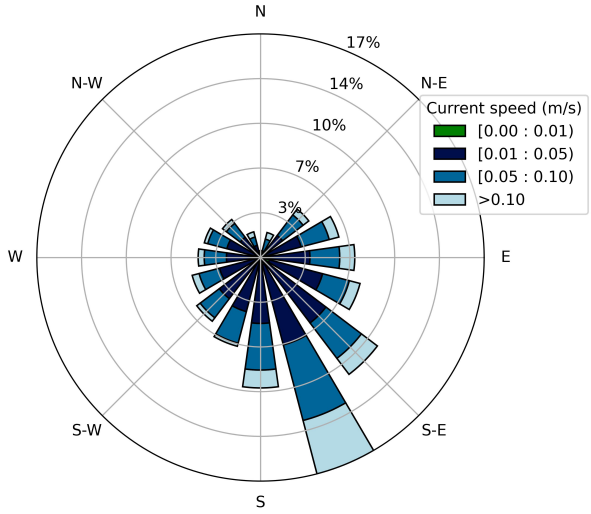


Fig. 3. Sea currents rose diagram for the potential location of an offshore wind farm in the Istria region, Croatia.

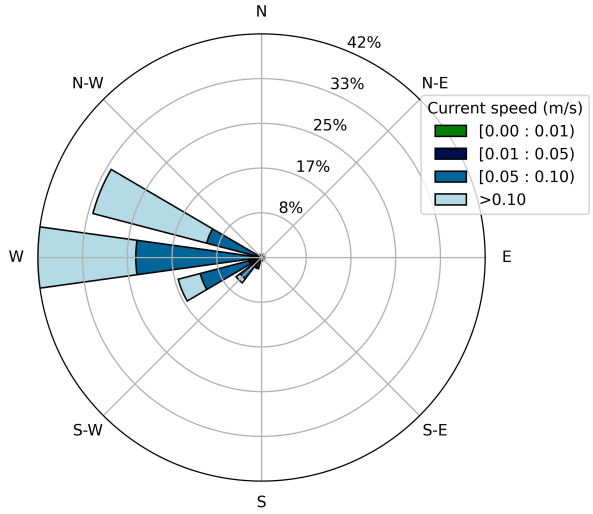


Fig. 5. Sea currents rose diagram for the potential location of an offshore wind farm in the Apulia region, Italy.

Sea currents analysis for the Dalmatia region presented in Fig. 4 showed the highest values compared to other regions. The average magnitude of sea currents was 0.06 m/s, while the maximum value was 0.6 m/s, predominantly in southeast and east directions.

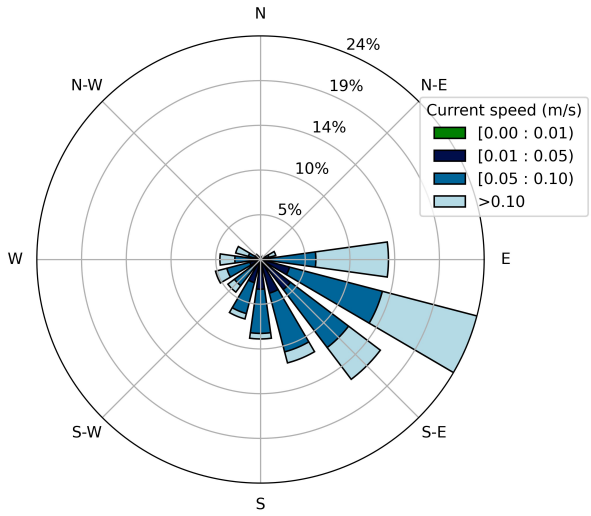


Fig. 4. Sea currents rose diagram for the potential location of an offshore wind farm in the Dalmatia region, Croatia.

Analysis of sea currents for 10 years in the Apulia region, presented in Fig. 5, shows a dominant west direction with maximum magnitudes up to 0.4 m/s and an average value of 0.01 m/s.

D. Waves

Fig. 6 represents wave conditions at the Veneto region based on hourly values for 10 years period. From the analysis of how often values of the same wave significant height (H_s) and wave mean period (T_p) appear, it can be concluded that most of the time, waves with small heights and short periods can be expected, with several extreme conditions. The highest value of wave significant height from the data is 5.3 m, while the average value is 0.25 m.

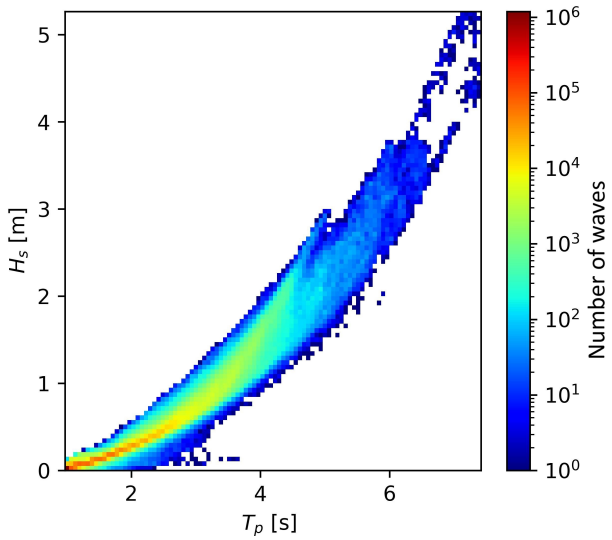


Fig. 6. Wave analysis for Veneto region.

Wave conditions for the Istria region are presented in Fig. 7. Results for the density diagram of wave significant height (H_s) versus wave mean period (T_p) indicate that most of the waves have small heights and short wave mean periods; however, values are expected to be bigger than in the Veneto region. The averaged value of wave significant height was 0.4 m, while the

maximum value was 5.9 m.

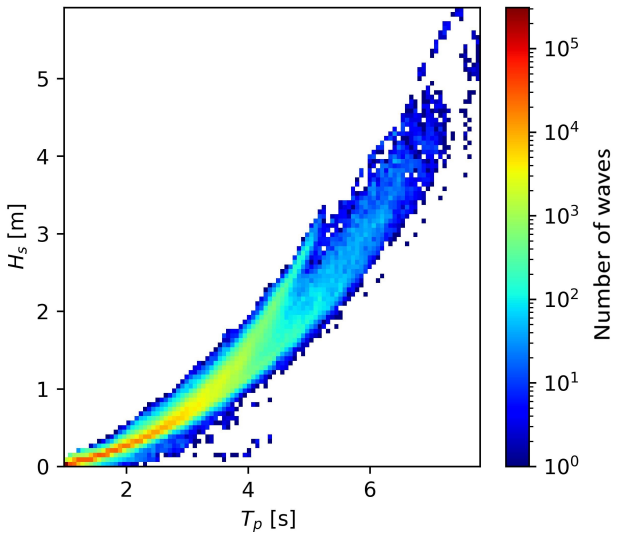


Fig. 7. Wave analysis for Istria region.

Density diagram of wave significant height (H_s) versus wave mean period (T_p) for the Dalmatia region is presented in Fig. 8. The diagram indicates a slightly higher number of extreme events with higher significant height and longer mean period compared to the previous two regions, which is expected due to the southern position in the Adriatic Sea. The highest wave significant height obtained from the CMEMS model was 5.6 m, while the average value was 0.42 m.

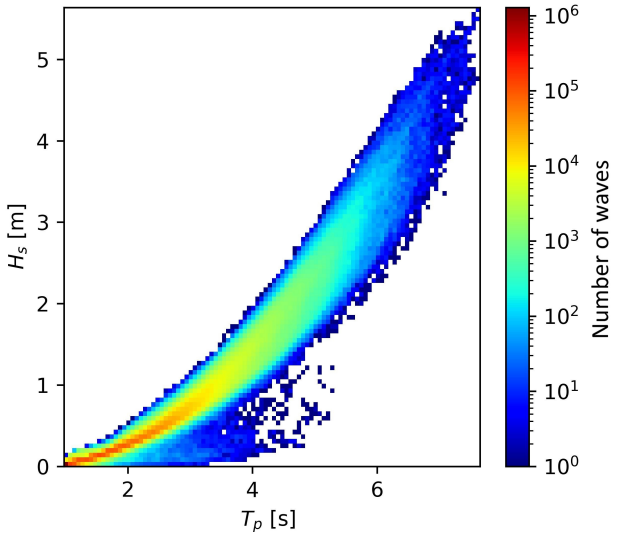


Fig. 8. Wave analysis for Dalmatia region.

Fig. 9 represents the density diagram of wave significant height (H_s) versus wave mean period (T_p) for the Apulia region. Results show the largest number of events with higher height and especially a longer mean period compared to the other three pilot regions. It indicates larger and more frequent loads on offshore foundations with extreme wave conditions and consequently higher costs. The average value of wave significant height for the Apulia region was 0.48 m, while the highest value was 5.35 m.

TABLE IV
COMPARISON OF SEA SURFACE HEIGHT FOR PILOT LOCATIONS.

	Veneto region	Istria region	Dalmatia region	Apulia region
Minimum	-0.9 m	-0.89 m	-0.85 m	-0.85 m
Maximum	0.17 m	0.08 m	0.02 m	0
Average	-0.44 m	-0.45 m	-0.45 m	-0.48 m

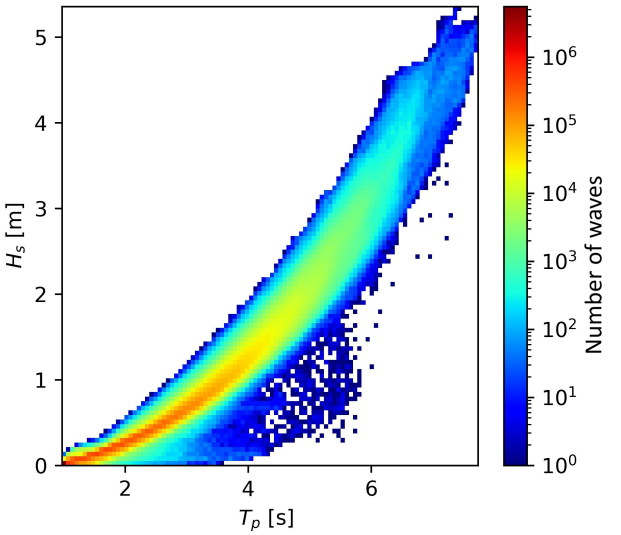


Fig. 9. Wave analysis for Apulia region.

E. Sea Surface Height

Tab. IV represents sea surface height for each pilot location. Results indicate that all locations have similar values, with values less than the reference mean sea level. Slightly higher values are expected in the Veneto region. It can be assumed that during planning, the design of the foundations should be about 0.5 meters lower than the expected sea level.

IV. CONCLUSION

Due to the urgent need to reduce greenhouse gas emissions and combat climate change, offshore renewable energy is becoming increasingly important. To choose adequate technology and location, analysis of the data obtained for the potential location is crucial. In this research, the application of CMEMS data for planning offshore renewable energy is presented. Analysis was done for sea physical condition, including sea temperature, salinity, sea currents, sea surface height, sea surface height, significant wave height and wave

mean period. Results are compared for four pilot locations in the Adriatic Sea. Compared data is consistent with expected values due to the influence of the river Po, indicating successful implementation of such data for planning offshore renewable energy, especially in the initial stages of projects. Such data can also be beneficial in other stages and more detailed analysis, along with planning construction and maintenance activities. However, for more detailed planning in later construction phases, other methods should be considered due to the relatively low resolution of data and the inability to capture local circulations. Considering that CMEMS data is based on numerical models and estimations, there is still uncertainty and biases which should be taken into account when detailed planning. Therefore, in situ measurements and detailed numerical models should be used for later stages.

ACKNOWLEDGMENT

This work was fully supported by the Interreg Italy – Croatia program through the project BEYOND (Blue Economy sYnergies fOr sustaiNable Development). The datasets used in this study were obtained from the Copernicus Marine Environment Monitoring Service (CMEMS).

REFERENCES

[1] A. Olabi and M. A. Abdelkareem, "Renewable energy and climate change," *Renewable and Sustainable Energy Reviews*, vol. 158, p. 112111, 2022.

[2] European Commission, "An EU Strategy to harness the potential of offshore renewable energy for a climate neutral future," European Commission, Tech. Rep., 2020. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2020:741:FIN

[3] P. Y. Le Traon, A. Reppucci, E. Alvarez Fanjul, L. Aouf, A. Behrens, M. Belmonte, A. Bentamy, L. Bertino, V. E. Brando, M. B. Kreiner *et al.*, "From observation to information and users: The copernicus marine service perspective," *Frontiers in marine science*, vol. 6, p. 234, 2019.

[4] A. Abramic, A. G. Mendoza, V. Cordero-Penin, M. Magalhaes, Y. Fernández-Palacios, C. Andrade, H. Calado, S. Kaushik, G. Carreira, N. Nogueira *et al.*, "Site selection within the maritime spatial planning: Insights from use-cases on aquaculture, offshore wind energy and aggregates extraction," *Ocean & Coastal Management*, vol. 251, p. 107051, 2024.

[5] N. Vidjajev, S. Rikka, and V. Alari, "Pre-evaluation of wave energy converter deployment in the baltic sea through site limitations using cmems hindcast, sentinel-1, and wave buoy data," *Energies*, vol. 18, no. 14, p. 3843, 2025.

[6] I. G. F. d. Freitas, H. B. Gomes, M. Peña, P. Mitsopoulos, T. S. V. Nova, K. M. R. d. Silva, and A. J. P. Calheiros, "Evaluation of wind and wave estimates from cmems reanalysis for brazil's offshore energy resource assessment," *Wind*, vol. 2, no. 3, pp. 586–598, 2022.

[7] E.U. Copernicus Marine Service Information (CMEMS) , "Mediterranean sea - high resolution l4 sea surface temperature reprocessed," Marine Data Store (MDS), 2025, accessed: June 11, 2025. [Online]. Available: https://doi.org/10.48670/moi-00173

[8] E.U. Copernicus Marine Service Information (CMEMS), "Mediterranean sea physics reanalysis," Marine Data Store (MDS), 2025, accessed: June 11, 2025. [Online]. Available: https://doi.org/10.25423/CMCC/MEDSEA_MULTIYEAR_PHY_006_004_E3R1

[9] E.U. Copernicus Marine Service Information (CMEMS) , "Mediterranean sea waves reanalysis," Marine Data Store (MDS), 2025, accessed: June 11, 2025. [Online]. Available: https://doi.org/10.25423/cmcc/medsea_multiyear_wav_006_012

[10] E.U. Copernicus Marine Service Information (CMEMS), "Myocean pro viewer," 2025, accessed: June 11, 2025. [Online]. Available: https://data.marine.copernicus.eu/viewer/expert

Analysis of the Concepts of Using Microservice Architecture on Open Operating Systems

Davor Cafuta
Department of Information Technology and Computing
Zagreb, University of Applied Sciences
Zagreb, Croatia
davor.cafuta@tvz.hr

Leonardo Štavalj-Ladišić
Department of Information Technology and Computing
Zagreb, University of Applied Sciences
Zagreb, Croatia
leonardo.stavalj@tvz.hr

Danijela Pongrac
Department of Information Technology and Computing
Zagreb, University of Applied Sciences
Zagreb, Croatia
danijela.pongrac@tvz.hr

Ivica Dodig
Department of Information Technology and Computing
Zagreb, University of Applied Sciences
Zagreb, Croatia
ivica.dodig@tvz.hr

Brigitta Cafuta
Department of Information Technology and Computing
Zagreb, University of Applied Sciences
Zagreb, Croatia
brigitta.cafuta@tvz.hr

Mario Golubić
Department of Information Technology and Computing
Zagreb, 4IT Solutions d.o.o.
Zagreb, Croatia
mario.golub@gmail.com

***Abstract*—This paper explores microservice architecture, its advantages and challenges, and its use in modern web application. Microservices enable the decomposition of an application into smaller, independent services that enable flexibility, resilience and easy scalability. Application is demonstrated using a practical scenario of an e-commerce system, the use of tools such as Docker, Kubernetes, Kafka and Keycloak, as well as approaches to authentication, communication between services and system monitoring. In the paper we emphasizes advantages such as modularity and robustness, but also addresses challenges such as complexity management and distributed transactions.**

***Index Terms*—distributed systems, Docker, Kubernetes, DevOps, resilience, observability**

I. Introduction

Microservices are an architectural style of software design in which an application is divided into a series of small, interconnected but independent services. Each microservice performs a specific functionality and focuses on a business problem, so the application consists of modules that easily communicate via standardized APIs (Application Programming Interface), usually REST (Representational State Transfer) or gRPC (Google Remote Procedure Calls) protocols. In contrast to traditional monolithic architectures, in which the entire application is firmly connected to an overall system, the microservice architecture enables modularity and flexibility, which improves the scalability, manageability and resilience of applications [1].

This research is funded by the project Application of VR technology and neural networks in the field of computer security and digital forensics - NPOO2024-1 - EU

Each microservice can be developed, tested, deployed and scaled independently of other parts of the application. This independence allows teams to work in parallel, speeding up development cycles and making it easier to update and debug individual parts of the system without affecting the entire system. A microservice for the customer account may manage customer data and preferences, while another service processes orders, tracks inventory or enables payment [2].

Due to the modularity and autonomy of each microservice, it is possible to develop each service in different programming languages or frameworks, depending on the specific needs or preferences of the team developing it [2].

For example, a payment microservice could be developed in Java for its robustness and security benefits, while analytics or recommendation systems could use Python for data processing and machine learning support. In this way, microservice architecture facilitates the use of the best tools and technologies for different parts of the application.

This approach is particularly useful in environments with a high need for scalability and fault tolerance, as it allows teams to simply scale a particular microservice when demand increases, rather than scaling the entire application.

The next chapter describes the related work. Then, it explains the main differences between monolithic and microservices architecture. The fourth chapter presents a case study of the architecture of the entire system. The conclusion briefly summarizes the advantages and challenges of microservice architecture.

II. Related work

Martin Fowler and James Lewis first popularized the term microservices through numerous technical articles in which they explained the need to change the architecture of large applications [1].

Netflix has introduced a microservice architecture to better manage the user experience and performance of its service. This allows individual parts to be updated more quickly without affecting the entire platform [3].

Netflix is one of the pioneers of microservice architecture, which it uses to manage a variety of functions such as streaming services, user authentication, recommendations and content personalization. By implementing microservices, Netflix ensures that every aspect of its platform can scale independently, while issues and upgrades can be fixed quickly without impacting other parts of the application. This approach allows Netflix to provide its users with a highly personalized viewing experience and handle large amounts of traffic worldwide [4].

The development of microservices has also been aided by the advent of virtualization and later container technologies such as Docker and Kubernetes. With Docker containers, microservices can be packaged together with all the necessary dependencies, ensuring consistent behavior regardless of the environment in which the service is running [5]. This allowed for easier testing, deployment and management of microservices, and Kubernetes enabled their orchestration in large, distributed systems [6]. The microservice architecture has thus become indispensable for the development of modern cloud-native applications.

Each microservice is isolated from the others and can function as an independent unit. This allows the team to develop, test and deploy a single service without affecting the rest of the application. If one service has a problem, it doesn't necessarily mean that the entire application stops working. The independence of microservices allows teams to work in parallel and independently of each other [1]. For example, the Google Maps API is structured as a microservice and can be used as a standalone service within different applications [7] [8] [9].

In a microservice architecture, the services usually have their own databases, which reduces the dependency between them. This facilitates independent scaling of the service, but also brings challenges in terms of data synchronization and consistency. [10] [11].

Amazon began introducing microservices to give teams more flexibility in developing new features. Amazon's transition to microservices enabled faster release of new features and simplified maintenance of the infrastructure that supports the global e-commerce platform. For example, Amazon's microservices manage various functions such as order fulfillment, inventory tracking, and payment services, allowing teams to innovate independently and respond more quickly to changing market demands [10] [11].

Microservices make it possible to scale certain parts of the application without affecting the rest of the system. Uber, for example, uses Node.js for its real-time requirements, while for analytics it prefers Python due to its strong data processing libraries [12].

Uber uses microservices to support critical functions such as ride management, payment processing and customer support. Microservices allow Uber to be highly scalable and adapt quickly to the market, and the application can easily respond to the growth of users and driver partners. This architecture allows Uber's engineering teams to make changes without affecting the entire platform, which is especially useful in a rapidly changing industry like transportation [12].

The scalability and resilience of microservice systems are the most important advantages of this architecture. Horizontal scaling enables the addition of new microservice instances to reduce load and respond to increased demand. Unlike vertical scaling, where resources such as CPU or memory are added to a single server, horizontal scaling involves the multiplication of microservice instances, allowing for better traffic distribution.

An example of the application of horizontal scaling are systems such as Google Cloud, which uses horizontal scaling to adapt to large loads in real time, such as high demand during global events or campaigns [7] [8] [9]. This strategy enables flexible and cost-effective capacity expansion as needed.

In microservice systems, fault tolerance is achieved by implementing various patterns that allow the system to recover quickly or minimize damage in the event of an error. One of the most important patterns is the circuit breaker, which temporarily interrupts requests to an unresponsive service to prevent further overload. This allows other services to continue working even if a particular service is temporarily unavailable [13].

Netflix has pioneered the use of the Circuit Breaker pattern and also uses mechanisms such as Retry and Timeout to retry or terminate the request if a response does not arrive within the expected time. This combination allows systems to adapt to temporary failures and maintain stability [3].

Monitoring and logging are key to maintaining the stability and performance of microservice architecture as they provide insight into system health and early detection of problems. Prometheus is often used to collect performance data, while the ELK stack (Elasticsearch, Logstash, Kibana) enables detailed log monitoring and system event analysis. Real-time monitoring enables proactive management of problems, while logging makes it easier to identify the cause of the problem and enables rapid diagnosis. Amazon, for example, uses the ELK stack for centralized monitoring of its applications and rapid troubleshooting, ensuring high availability and a quick response to potential challenges [10] [11].

III. Monolit vs. Microservice Architecture

Monolithic applications are easier to develop initially because all parts of the application use a common code base and database. They are easier to implement and test at the beginning of development because everything is integrated into a single application. However, as the system grows, this simplicity becomes a limitation [1].

A monolithic architecture means that the entire application is integrated into a single system where all parts are interconnected and use a common set of resources, such as a single database and a set of APIs. One of the main advantages of a monolithic architecture is the simplicity of development and implementation. Since the entire application is in a single code package, it is easier to understand, develop and implement. Another significant advantage of a monolithic architecture is increased performance efficiency. Since all parts of the application are in the same operating environment, the different services do not need to communicate over the network, resulting in lower latency. In a microservices architecture, communication between microservices can sometimes lead to delays and bottlenecks, whereas in a monolithic system, interaction between components is much faster and more direct. This internal connectivity also means that a monolith is easier to test and integrate, as all parts are tested together in a single environment.

A monolithic architecture also has the advantage of simpler transaction management and data consistency. Since all functions share a single database, it is easier to ensure data consistency and transaction coordination. This is particularly useful for applications that require complex transactions or need to ensure a high level of data accuracy and security need to ensure consistency. In a microservices architecture, where each service has its own database, ensuring consistency can be more challenging and complex.

In addition, a monolithic architecture may be a more cost-effective choice for smaller applications or those with a relatively stable feature set. Developing and maintaining a monolithic application often requires fewer resources than a microservices architecture, which requires specialized solutions for managing, orchestrating and monitoring a large number of independent services. A monolithic architecture can be the right choice for projects that do not require high scalability or for organizations looking for a simple, cost-effective solution for their applications.

Although the monolithic architecture has certain limitations, its advantages in terms of simplicity, consistency and lower costs make it a suitable choice for less complex or limited applications and for companies that want a cost-effective and easy-to-maintain infrastructure for their applications.

advantages of microservices compared to monoliths are particularly evident in complex and large applications that require high scalability, flexibility and development speed. The microservices architecture, which is based on the division of an application into smaller, self-contained modules, the microservices, brings with it a number of advantages that monolithic architectures cannot always achieve. Thanks to their modularity, microservices enable more agile resource management, easier updates and faster adaptation to market changes, which is crucial in the development of modern web applications.

To illustrate the key differences between the two architectural approaches, Table I compares the characteristics of monolithic and microservice architectures:

TABLE I
THE DIFFERENCE BETWEEN MONOLITHIC AND MICROSERVICE ARCHITECTURE

Characteristics	**Monolithic architecture**	**Microservice architecture**
Development	The entire system is developed as a single code pool	Each service is developed independently, often by separate teams
Deployment	The entire application must be redeployed each time a change is made.	It is possible to deploy each service independently
Scalability	More difficult, the entire system is scaled together	Easier, only the service that requires greater scalability is scaled
Technological freedom	Limited, the entire system must use the same programming language and framework	Each service can use different technologies, polyglot development
Resilience	A failure of one component can jeopardize the entire system	The failure is isolated, other services continue to function.
Complexity management	Simpler at the beginning, but more difficult to maintain over time	More complex at the beginning, but more modular and sustainable in the long term
Team organization	One large team works on everything	Several independent teams work on individual services
CI/CD	One pipeline for the entire application	Multiple CI/CD flows for each service individually
Monitoring diagnostics	Simple for a smaller number of components	Requires an advanced observability stack for monitoring services

IV. CASE STUDY

In modern software development, microservice architecture is increasingly becoming a standard approach for creating scalable, distributed and easily maintainable applications. Its main advantage lies in the ability to divide an application into smaller, independent components, services, which can be developed, tested, scaled and distributed separately. This approach significantly increases team agility, facilitates continuous delivery and enables the effective application of DevOps practices [2].

However, the transition from theoretical models to the actual implementation of microservice architecture brings with it a number of challenges:

- how to ensure configuration consistency across different environments (development, test, production)?
- how to reproduce production errors locally during development,
- how to achieve robust and reliable communication between the services,
- how to monitor the performance of distributed components in real time using appropriate monitoring tools.

These challenges were the direct motivation for the creation of this practical example of a microservices-based e-commerce system, which pursues the following goals:

- Enable realistic replication of the production environment on-premises using tools such as Kubernetes in Docker (kind), Docker containers and standard cloud-native tools for development, configuration and monitoring
- Support rapid iteration and debugging in an environment that uses the same service declarations (e.g. manifests and configurations) as cloud systems
- Show the entire data flow through the system, from the user request in the React application, authentication and authorization through the Keycloak server, traffic management and resilience policies in the Spring Cloud Gateway, to synchronous and asynchronous communication between microservices, data storage and sending notifications
- Implementation of an observability approach through OpenTelemetry instrumentation, centralized collection of metrics with Prometheus, trace tracking with Grafana Tempo and log management with Grafana Loki, all displayed in a unique Grafana interface

This project thus serves as a bridge between academic theory and concrete practices in an enterprise environment. It aims to provide a controlled but sufficiently complex environment to experiment with microservices before implementing them in real business systems. This approach enables a better understanding not only of architectural principles, but also of the day-to-day challenges of developing, orchestrating and monitoring distributed applications.

The developed e-commerce system based on the microservice architecture is structured in a main repository called tvz-microservices. Each logical component of the system is organized in its own subdirectory containing source code, configuration files, tests and related resources. Such a modular organization promotes independence of development, testing and scaling of individual parts of the application. Figure 1 is a visual representation of the overall system architecture, illustrating the relationships and data flows between the user interface, authentication server, API gateway, microservices, databases and observation layer.

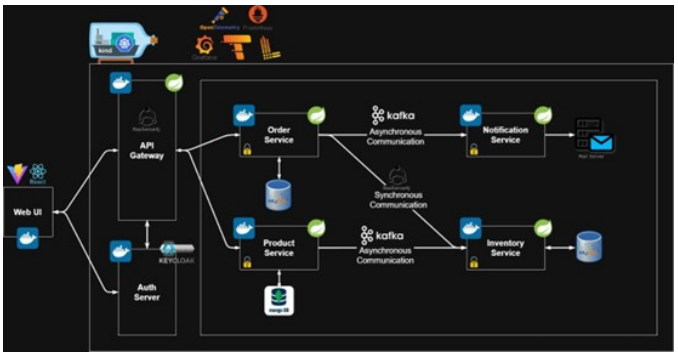


Fig. 1. Visual representation of the overall system architecture

The architecture described above is implemented in a local Kubernetes cluster started with the kind tool, which enables a realistic simulation of a production environment. The main components of the system are described below:

A. API gateway

This service is located in the api-gateway directory and uses the Spring Cloud Gateway as the entry point for all user requests. Key functions include:

- Validation of JWT tokens generated via the Keycloak server
- Managing security rules with classes such as SecurityConfig.java and CorsConfig.java
- Definition of routing rules in the Routes.java file, which determines the routing of requests to individual microservices
- Application of Resilience4j patterns for resilient behavior in the event of failures, including interruption and retry mechanisms

B. Microservices

The system consists of four independent Spring Trunk microservices, each of which operates in its own Docker container and has a clearly defined responsibility:

- Product Service: manages products, uses MongoDB, publishes events to Kafka topics
- Order Service: manages orders, uses MySQL, sends events and uses synchronous communication with the Inventory Service
- Inventory Service: manages inventory, uses MySQL, responds synchronously to requests and updates inventory status asynchronously
- Notification service: without its own database, consumes events from the Kafka topic and sends e-mail notifications via an external mail server

Each service has its own REST API, uses JPA or MongoTemplate and provides support for testing with JUnit5 and the Testcontainers libraries for integration tests in containers.

The microservices described communicate with each other using the event-driven architecture model with the use of Kafka topic as a mechanism for distributing events. In cases where immediate validation or confirmation is required (e.g. inventory check), synchronous HTTP communication is used. This hybrid combination enables for both high flexibility and the required precision, depending on the context of the request.

C. Front-end application

The microservices-web directory contains a frontend application developed in React and deployed with the Vita tools. Application:

- Creates a static bundle that is provided via the Nginx web server
- Communicates with the backend system via HTTP calls to the API gateway
- Authentication management is based on the OpenID Connect protocol (configuration in oidcConfig.ts) and the tokens are stored and managed by authStore.ts

D. Docker configurations

The docker/ directory contains all configuration files for infrastructure services and supporting systems:

- Observability tools: prometheus.yml, tempo.yml, dashboard.json for Grafana
- Database initialization: init.sql for MySQL, mongo-init.js for MongoDB
- Keycloak configuration: realm-export.json with user, client and role definitions

E. Kubernetes manifests

The k8s/ directory contains all scripts and manifests needed to start and manage a local cluster:

- create-kind-cluster.sh, delete-kind-cluster.sh and kind-config.yaml for managing child clusters.
- manifests/applications/ contains Kubernetes deployment and service objects for each microservice, API gateway and frontend.
- manifests/infrastructure/ contains manifest files for the infrastructure such as Kafka, Zookeeper, Prometheus, Tempo, Loki and the databases PostgreSQL and MongoDB

F. Infrastructure services

Each infrastructure element runs as a Docker pod within a Kubernetes cluster and includes:

- Apache Kafka and Zookeeper - distributed communication between the services
- Schema Registry - validation and versioning of Kafka messages
- Prometheus, Grafana, Tempo, Loki - observation layer for metrics, traces and logs
- ConfigMaps and Secrets - for managing configurations, TLS certificates and sensitive data

For a microservice system to be sustainable and reliable, observability is an indispensable part of the architecture. As part of the e-commerce system, a comprehensive monitoring system was developed that combines metrics, traces and logs. It is used not only to monitor status, but also to quickly detect problems, analyze performance and achieve a high level of availability.

The integration with OpenTelemetry and Micrometer was used in all microservices. OpenTelemetry enables the collection of traces and metrics from the application, while Micrometer provides a unified layer for exporting data to systems such as Prometheus.

The collected metrics are sent to Prometheus, which regularly retrieves data from the endpoints of each service (scrapes). The configuration in prometheus.yml defines jobs for each microservice and infrastructure component.

Grafana Tempo is a distributed system for storing and retrieving traces that is compatible with the OpenTelemetry Protocol (OTLP). It is used to monitor the flow of requests across multiple services. Each service generates traces that represent steps within a single request. These traces are visualized in Grafana, where you can see which service consumed the most time, where an abort or an error occurred.

Tempo enables end-to-end tracking. For example, a user request passes through the API gateway, then the product Service, the inventory Service and finally ends up in the database. Each of these steps is measured and timestamped, and all are linked by a unique traceid.

The logs of all services are sent to Grafana Loki, which only indexes metadata such as the name of the service, the message level and the traceid, while the rest of the content remains unindexed. This enables a fast and efficient search.

The log entries are structured and contain timestamp, log level, thread name and attributes such as orderNumber, skuCode, etc. Loki therefore makes it possible to search logs according to the same criteria as traces, allowing errors to be correlated across all layers of the application.

The architecture as a whole offers numerous advantages. First of all, there is a clear division of tasks, with each service managing only its own area, which simplifies development, testing and documentation. System resilience is achieved using the use of circuit breaker patterns and asynchronous processing, avoiding chain crashes in the event of a single service failure. Furthermore, each service can be independently scaled horizontally, enabling real-time load adjustment. If a single module fails, the rest of the system continues to operate, which contributes significantly to high availability.

G. Communication between services

Communication between microservices is a fundamental functionality of any distributed architecture and has a direct impact on data consistency, system robustness and overall performance. In this e-commerce system, a hybrid strategy has been implemented that combines synchronous and asynchronous communication, depending on the domain requirements and the required level of consistency.

Synchronous HTTP calls are used for operations that require immediate confirmation, such as checking the availability of items or reducing stock in real time. For example, when initializing an order, the order Service synchronously calls the inventory Service to check the stock status (GET /api/inventory). This approach enables clear process control and direct exception handling within a single transaction. The use of @Transactional in Spring enables atomic execution of operations, reducing the risk of partially executed transactions.

However, synchronous calls also come with challenges, such as latency due to network communication, as well as the possibility of cascading failure if a service becomes unavailable. To mitigate these risks, resilience patterns such as circuit breakers and retry mechanisms implemented via the Resilience4j library have been integrated into the system. This configuration allows the system to terminate calls to faulty services and provide fallback behavior, thus protecting other components from overload.

On the other hand, asynchronous communication was performed using Apache Kafka as a distributed message broker.

This form is used for operations that do not need to be executed immediately and should not block the main flow of the application, such as initializing the inventory or sending email confirmations. For example, after saving a new product in the database, the Product Service publishes the InventoryCreatedEvent event in the Kafka topic inventory-created, which the Inventory Service retrieves asynchronously. Similarly, the Order Service publishes an OrderPlacedEvent, which the Notification Service picks up and uses to send personalized email notifications to users.

However, asynchronous communication comes with its own challenges. Firstly, the data between the services can only be consistent to a limited extent, as changes are passed on with a time delay. In addition, the sequence of events, the processing of repeated messages (idempotency) and the possibility of errors occurring only during message processing must be taken into account. In this project, protection is implemented through dead-letter topics, where unprocessed messages are forwarded to a special channel for later analysis. In addition, the use of schemas and the schema registry ensures compatibility between event versions, which facilitates the maintenance and further development of the system.

This combination of communication patterns enables a balance between speed, consistency and reliability. Critical operations such as inventory confirmation rely on synchronous validation, while operations that can be processed independently are processed asynchronously without affecting user-friendliness. The system therefore utilizes the advantages of both approaches: synchronous for precise flow control and asynchronous for scalability and reliability.

V. Conclusion

The microservices architecture represents a paradigm shift in the way modern distributed applications are designed, developed and maintained. Its main strength lies in its modularity, i.e. the division of a complex system into independent, specialized components that can be developed, tested, implemented and scaled independently of each other. This approach enables greater flexibility, speeds up the provision of new functionalities and reduces dependencies between teams.

Based on the analysis of an e-commerce system, a concrete implementation of microservices with modern technologies such as Spring Trunk, Docker, Kubernetes, Apache Kafka and Keycloak is shown. The combination of synchronous and asynchronous communication patterns has proven to be optimal for achieving a balance between the speed of response and the system's resilience. The use of DevOps tools and CI/CD pipelines enabled fast iterations and automated testing, while the observability layer (Prometheus, Grafana, Tempo, Loki) provided visibility and control over the performance and stability of the system.

Despite its many advantages, microservices architecture is not without its challenges. Complexity management, distributed security, message versioning and eventual consistency are just some of the issues that require additional infrastructure and operational maturity. Implementing the service mesh layer, orchestrating distributed transactions through the Sagas pattern and harmonizing configurations through GitOps practices are key elements that help overcome these challenges.

Examples from leading global technology companies such as Netflix, Amazon and Uber clearly demonstrate the power of microservices in large, scalable environments. At the same time, for smaller projects with limited resources, a monolithic architecture may still be the better choice. Microservices are not a universal solution, but a sophisticated tool that must be carefully tailored to the technical and organizational requirements of the system in question.

Ultimately, microservice architecture is not just a technical choice, but a strategic approach to developing a scalable and resilient software ecosystem in which technological flexibility becomes a critical factor for innovation and competitive advantage.

References

[1] C. Richardson, Microservices Patterns With examples in Java, Manning Publications Co., 2018

[2] S. Newman, Building Microservices, O'Reilly, 2015

[3] The Netflix tech blog, Medium, https://netflixtechblog.com/ (accessed July 12, 2025).

[4] K.Vardarska, "Understanding design of microservices architecture at Netflix",Techaheadcorp, https://www.techaheadcorp.com/blog/design-of-microservices-architecture-at-netflix (accessed July 19, 2025).

[5] Get Docker, Dockerdocs, https://docs.docker.com/ (accessed July 12, 2025).

[6] Production-Grade Container Orchestration, Kubernetes, https://kubernetes.io/ (accessed July 12, 2025).

[7] Google Cloud Documentation, Google Cloud, https://cloud.google.com/docs (accessed July 11, 2025).

[8] Secure and grow your apps with Google Identity, Google Identity, https://developers.google.com/identity (accessed July 11, 2025).

[9] Google Maps Platform Documentation, Google Maps Platform, https://developers.google.com/maps/documentation (accessed July 11, 2025).

[10] AWS Lambda Documentation, AWS, https://docs.aws.amazon.com/lambda/ (accessed July 10, 2025).

[11] Implementing Microservices on AWS, AWS, https://docs.aws.amazon.com/whitepapers/latest/microservices-on-aws/microservices-on-aws.html (accessed July 10, 2025).

[12] Introducing Domain-Oriented Microservice Architecture, Uber Blog, https://www.uber.com/en-HR/blog/microservice-architecture/ (accessed July 12, 2025).

[13] Circuit Breaker pattern, Microsoft Ignite, https://learn.microsoft.com/en-us/azure/architecture/patterns/circuit-breaker (accessed July 12, 2025).

Double Autorization System for Grade Entrance System in Higher Education Institution

Brigitta Cafuta
Department of Information Technology and Computing
Zagreb, University of Applied Sciences
Zagreb, Croatia
brigitta.cafuta@tvz.hr

Danijela Pongrac
Department of Information Technology and Computing
Zagreb, University of Applied Sciences
Zagreb, Croatia
danijela.pongrac@tvz.hr

Davor Cafuta
Department of Information Technology and Computing
Zagreb, University of Applied Sciences
Zagreb, Croatia
davor.cafuta@tvz.hr

Ivica Dodig
Department of Information Technology and Computing
Zagreb, University of Applied Sciences
Zagreb, Croatia
ivica.dodig@tvz.hr

Mario Golubić
Department of Information Technology and Computing
Zagreb, 4IT solutions d.o.o.
Zagreb, Croatia
mario.golub@gmail.com

***Abstract*—Computer security has recently played an important role in all digitized processes. When digitizing processes, we often strive to simplify data entry as much as possible. Very often we grant authorizations and open systems with little supervision to make work easier for users with little IT knowledge. The concept of a single shared password particularly contributes to this situation. In this paper, we present a dual-permission grade entry system upgrade designed to prevent unauthorized users from entering grades if they intercept the teacher's password.The architecture of such a system is presented, as well as a proposed implementation method using an embedded system. The goal in building the system is simplicity and cost efficiency.**

***Index Terms*—Embedded system, IOT, Computing security**

I. Introduction

To simplify work, especially after the Covid years, digitization of processes is increasing. With this digitization of processes, paper offline systems are increasingly being switched off. Digital information is trusted.

In higher education, all grade records were kept manually and, when grades were given, they were manually entered into the index and the application form on paper. With the introduction of the information system, the data was also kept digitally. This meant that at any one time there were three different pieces of information about the grade in the system itself. The biggest problem in such a system is that these three pieces of information can differ from each other. In that case, the user must intervene to determine the accuracy of the information.

This research is funded by the project Application of VR technology and neural networks in the field of computer security and digital forensics - NPOO2024-1 - EU

Over time, the paper application forms were partially digitized and printed through the information system, so we can assume that the grade and the data on the forms were identical to the entries in the information system. The grade was therefore stored via the paper index and the computer system. The application form and the examination list printed and signed from the information system formed a backup copy of the information system on paper.

The index was considered less accurate than the information system, as the index was in the student's possession at all times and, therefore, susceptible to falsification. Secretaries or student services workers have entered grades into the information system based on the teacher's examination record. The paper record was retained because it prevented manipulation during entry. With such a system, one could fully rely on the grade in the information system.

Over time, teachers began to enter grades into the system, initially only from a few certified computers. Once entered, they would sign the printed examination lists. The security was based on a limited number of computers and addresses from which the system could be accessed. Over time, the system became more open to access, primarily to accommodate teaching staff. Today, most higher education institutions use the Croatian national ISVU Teacher Portal for grade entry, which is available as a website with no restriction on network addresses.

The AAIEdu username and password, which are defined as level 2 of 4 existing levels in the national classification, are used to access the portal. The same username is used to access e-mail, the institution's portal, and grade entry. As a result,

most teachers do not set a complex password for the system. The password used to enter grades is the same one used in front of students when they connect to the lab computer or to the LMS system of the institution. This makes the grade entry system insecure, and there is a high risk of a third party stealing the password and entering grades for themselves or even selling grades. If grades are entered by the third person later during colloquium time, when there may be more than 200 students in a course, it is unlikely that the teacher will notice the difference.

Most teachers reduced the amount of their own record keeping when the digital information system was introduced. They are lulled by the security of the information system and sign the exam list without additional checks. For these reasons, it is time to put a system in place that aims to secure access to grade entry, in the sense that the access to the interface for accessing the computer and the portal is separated from the security-risk grade entry system.

The next chapter describes the related work. The system architecture is presented in the following chapter. The paper ends with a conclusion.

II. Related work

Identity theft is a crime in which someone illegally obtains your personal information or uses it to impersonate another person, usually for financial gain or to commit other fraud. In the real world, this is equivalent to stealing your home keys and gaining access to another person's home and all the information it contains. There is a risk of accessing financial accounts, credit cards and even health data. For this reason, authentication is introduced, the main task of which is to confirm correct identification.

The simplest form of authentication is the password. Before any sophisticated protection, a password is what prevents an unauthorized person from accessing unauthorized resources. In most cases, passwords alone are no longer sufficient in today's threat environment, but they are still the first factor in multi-factor authentication systems.

In practice, it is recommended to use long and complex passwords with a combination of upper and lower case letters, numbers and special characters. Unique passwords for each account and regular, but not too frequent, password changes for the most sensitive accounts to avoid encouraging the use of simple passwords [1].

Passwords are an indispensable part of digital life. Even though more and more advanced authentication methods are being developed, good password management remains the basis for the security of your personal and business computers [2] [3] [4].

Despite their importance, passwords are often the weak link in the security chain due to human error. People are less likely to use simple, easy-to-remember passwords that are vulnerable to dictionary or brute force attacks. Reusing the same passwords for multiple services is also a major risk [5] [6].

In addition to passwords, there are two other similar methods that are based on information that the user knows. These are shorter numerical codes or PINs, which are often used with bank cards or mobile devices. Shorter numeric codes are basically simplified passwords and do not increase security compared to passwords [7]. For this reason, we will not look at them in detail in this paper. Alternative passwords are security questions that only the user knows the answer to (e.g. "What is your mother's maiden name?"). When security questions are simple, they become vulnerable to social engineering. Another disadvantage of security questions is the input of the answers, as the answers usually consist of free characters where a single space, capital letter or punctuation mark can make a significant difference to authentication. With such systems, it makes more sense to include a human factor that can judge whether the user has given a sufficiently similar answer [8].

There is also a group of authorization mechanisms that are based on user characteristics. Their common name is biometric authentication. Examples are fingerprint, recognition of unique facial features, iris scanning or voice recognition. All of these methods are extremely secure, but for various reasons are not suitable for most systems, including ours [9].

Systems based on network authentication (access via a browser from the network) must in this case exchange extremely sensitive data (fingerprint or voice sample). For reasons of data protection, such options can not be implemented without the user's consent and a number of other measures to ensure the security of the computer system against external manipulation.

Browsers as end-user tools very often do not have a device to retrieve the user's biometric characteristics, as the computer is not equipped with such a device or it is not accessible to the browser to retrieve data from it due to the security of the device.

The other authentication methods are based on the principle of a device (application program or real device) that generates passwords. These devices are called tokens. The principle of operation is that a temporary PIN is created or a confirmation is sent by the application, which ultimately creates the same PIN [10].

Originally, these were small devices with a simple display and a built-in unique algorithm or random definition (initial password) that generated a new unique code each time it was pressed or generated a more complex code based on the built-in algorithm by entering the initial code. Today, these devices are often replaced by an application on a cell phone that takes over the hardware function of the device. For even more security, there are also smart cards or USB keys that retain the logic of hardware generation but have no interface, as the computer to which the device is connected serves as the interface [11].

Originally, the token was used as the only means of authorization, which required its distribution to all users. In addition to the complicated distribution process, redistribution in the event of token loss or the risk of it being compromised poses an even greater problem.

For this reason, computer security has evolved from single-factor authorization to dual or multiple factors. With the introduction of multiple factors, authorization is reduced to several steps, all of which must be fulfilled in order for a person to be authenticated.

For example, single-factor authentication, which is based on a password, is upgraded with double authorization, where after entering the correct password, the user pin must also be entered so that the person is authenticated. In the case of multiple authentication, we could additionally request confirmation of the data we have sent to the e-mail address after the user pin has been entered correctly. We could also require the presence of a specific hardware device or a digital key in the form of a certificate on the device itself [12].

III. System Autorization

For example, single-factor authentication, which is based on a password, is upgraded with double authorization, where after entering the correct password, the user pin must also be entered so that the person is authenticated. In the case of multiple authentication, we could additionally request confirmation of the data we have sent to the e-mail address after the user pin has been entered correctly. We could also require the presence of a specific hardware device or a digital key in the form of a certificate on the device itself.

Given the problematic distribution of tokens to employees and external teachers, of whom there are over 400 in the institution, the simplest authorization method is reduced to double authentication via a known password and a user PIN.

The one-step authorization is not sufficient, because the teaching staff often use very simple passwords, as the College Computer Center in Zagreb implements a single sign-on system via the AAIEdu identity. The AAIEdu identity does not support double authorization and is used for all systems in higher education, e.g. access to the institution's website, email and grading system. Its security has been clearly proven by the NIAS system of the Republic of Croatia, where it has been assigned level 2 out of 4 available levels.

Teachers and external collaborators very often connect to the institution's computers in front of the students, and very often they use browsers with a screen visible on the projector, which briefly display the password letter instead of the usual asterisk. With all the necessary secure preparations when entering the password, the student can obtain the teacher's password, for example, by reading over their shoulder or recording them while entering the password. In this paper, we do not address the potential vulnerability of the computer network itself and the operating system of the client computer because communication interception tools can be placed there or on the network itself to intercept the password.

Biometric systems were immediately discarded because of legal requirements, but also because of the logistics of collecting data from teachers and external collaborators. We also discarded the possibility of security questions, as the database for such questions would have to be very large, which is not in line with the actual ideas of the authorization aforementioned above. In addition, given the freedom of response, it is desirable for the user to see what they have entered, which can also be seen by a larger group of potential threats in the user's environment if the user is sitting at a shared computer, for example.

For this reason, a randomly generated user pin offers the best security, because if third parties manage to obtain the teacher's password, they will not have access to the algorithm for generating the user pin. Since the created user pin is different every time, it is not possible to use it by interception. In addition, a simple script can warn the user to change the password if the computer system detects a significant increase in incorrectly entered user pins for a single user.

The only weakness of the created user pin is its length, as a brute force attack is possible, where combinations are tried until the PIN is guessed. Entering the password protects the system from guessing the pins immediately. There are also restrictions on the number of attempts before the password is re-entered or possibly the captcha system is activated, which prevents automatic brute force scripts.

There are several options for submitting the PIN. A string of characters can be sent by email or SMS. The string can also be created using an authenticator in an application on a cell phone. The simplest application is a mobile authenticator application. During distribution, the user receives a QR code which they load into the application. The application assigns temporary passwords based on the initial key according to a time or sequence logic. The disadvantage of this system is that it can be installed on multiple devices, as the QR code can be used multiple times for installation. Even with secure issuance, where the cell phone is set up under the supervision of a secretary without the user a code, most authenticators support copying the code (algorithm) to another device. For this reason, this level of security is not sufficient.

The goal is to set up the system so that the user can authorize the entry of grades from only one device, in a way that makes the distribution of such a concept as easy as possible. All teachers and external collaborators must provide an email address and cell phone number to contact the institution where they work. Email is very insecure as a third party can access it if they know the user's password. A cell phone, on the other hand, solves the problem as it is a unique device and sending an SMS message ensures that only the user has access to the SMS message. It is not expected that the user will make it available to a third party and their screen will be shared on a projector. To prevent possible screen sharing, the user's PIN appears after two lines in the message itself so that it is not directly visible on the main screen in the overview view of the incoming SMS message.

IV. System Architecture

The system architecture is structured as follows. The basis of the system is moj.tvz.hr as the main internal website of the college, which is available to all teaching staff. The website uses its own token, which it constantly transmits via a GET call and stores in the database for security reasons.

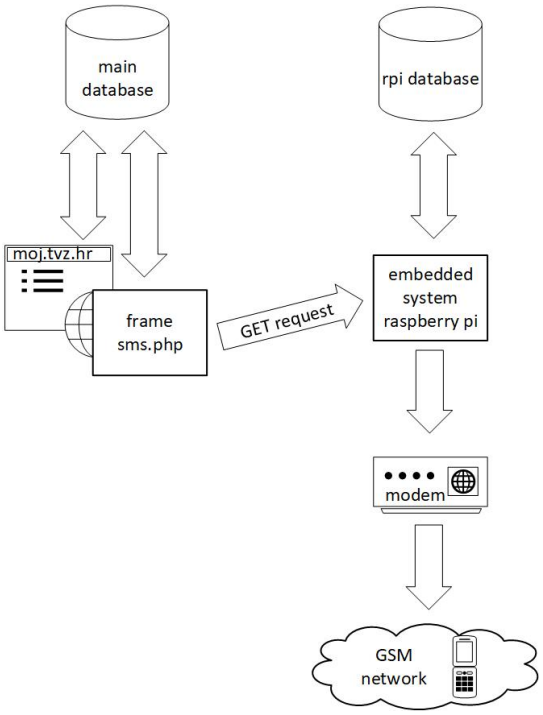


Fig. 1. System Architecture

The portal is accessed through the Office 365 service over Microsoft 2FA authorization with an authorization key from Microsoft. We believe that the aforementioned authorization model is not sufficient for entering grades, as Microsoft's authorization mechanism can be installed on multiple cell phones and computers. In addition, users can use the option to remember the session settings in the browser while working on a public computer, making it accessible to a third person. In such a case, there would be a risk that grades could be entered by a third party on a public computer.

Within the website, teachers who have appropriate authorization have access to the course for which they wish to enter grades. If they attempt to log in, the system will block access until they complete two-factor authentication via SMS. Through the web interface, they can request a temporary PIN, which is sent to their registered mobile phone. Once the temporary PIN is entered, the teacher has 120 minutes to submit grades.

To prevent issues that may arise from leaving an authorized session active on a public computer, the system monitors the timing of session requests. If there is a significant period of user inactivity, the system automatically ends the session and prompts the user to generate and enter a new PIN in order to continue working. The system is implemented using a database. An additional 2FA authorization table has been added, which stores the phone number to which the message was sent, the PIN code that was sent, the session identifier of the website from which the request originated, the date and time the message was sent, and the user who requested it.

Access to the message-sending function is managed through a separate PHP script, which is integrated into the existing website using an iframe. This page is hosted on the same server and under the same HTTPS certificate to ensure proper display and compatibility with browser security settings. The page features a form with a send button, and both the session and user ID are passed to it. Data transmitted between the server script and the authorization script during a GET request is protected with a symmetric encryption key.

When the send button is clicked, a new record is added to the table, and a GET request is simultaneously sent to the embedded system responsible for sending the SMS. Communication between the embedded system and the server script is secured using a newly established symmetric key. The embedded system receives the request on its network server and places the data to be sent into a queue. Confirmation of the request's receipt is then sent. The queue is managed through a database table. The script on the embedded system starts automatically, monitors the queue, and sends SMS messages via the modem in the order they were received. If a request waits longer than one minute or if the same user submits duplicate requests, the system automatically rejects them. This helps prevent unwanted system blockages, as most users tend to make multiple requests in a short period when they do not receive an SMS promptly. Since all requests are recorded in the database, confirmation of just one is sufficient. Once the modem confirms the transmission, the system updates the database to indicate that the message has been sent. The modem's average capacity is approximately one message every five seconds.

When the user receives the SMS message, they enter the code into the iframe interface to complete the two-factor authentication. The system verifies the code against the database, and if it matches, the grade entry system is activated by updating the authorization timestamp to allow access for the next 120 minutes for the specified session.

The embedded system was built using a Raspberry Pi 4 equipped with a GSM SIM7600 modem. The Raspberry Pi is installed in the server room of the higher education institution. It is secured by being located within a separate network, which prevents direct access to its services without explicit firewall rules allowing such access.

An Apache web server with PHP support is used to facilitate communication with the server. The PHP environment enables encryption functions to secure communication and database access. A MariaDB database is installed on the Raspberry Pi to manage the message queue. A PHP script provides a web page through which SMS send requests can be made; these requests are automatically recorded in the database table that serves as the queue.

A service is set up within the embedded Raspberry Pi system that runs a microservice written in Python. This microservice reads the database table, retrieves send requests from the queue, and communicates with the modem via the serial interface using AT commands. The advantage of using the Python language lies in the blocks for catching errors, which enable problems to be sent automatically to the servers

for monitoring by central IT support. For example, if the GSM network fails, a notification is automatically sent to the central computer service server. Besides monitoring the modem, the script also oversees system parameters of the Raspberry Pi, such as hard disk status and the power quality.

In the event of an unexpected crash of the microservice, which was particularly important in the initial phase of testing the computer system, a Python script for monitoring was set in a time-controlled trigger. The time-controlled script (placed in CRON) checks at short intervals whether the microservice is operational and ready for serve.

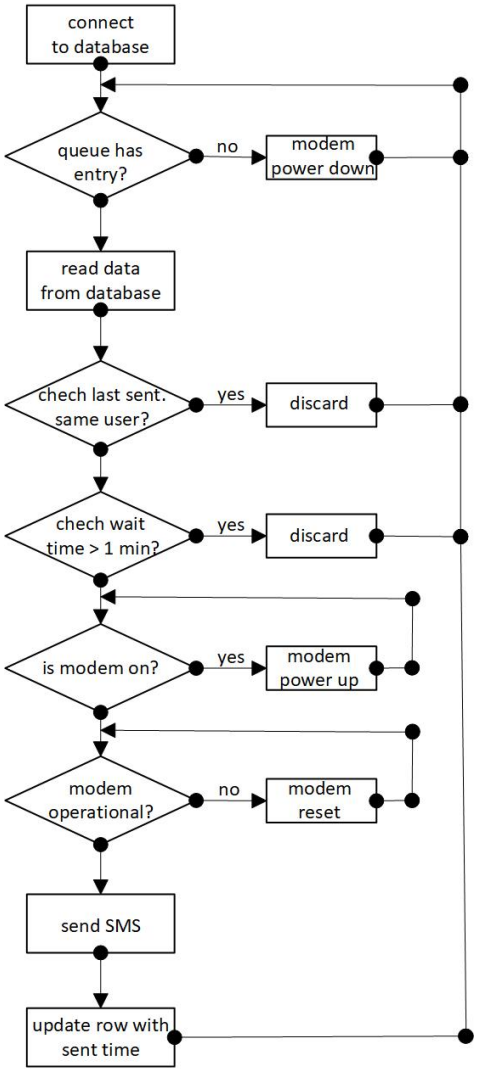


Fig. 2. Algorithm of Raspberry PI Modem Python Microservice

V. Conclusion

In today's digitally connected world, where information is a valuable resource, understanding and applying the principles of computer security have become essential for individuals, organizations and governments to ensure the confidentiality, integrity and availability of data. Without adequate safeguards, systems are vulnerable to various threats such as malware, hacking, identity theft and social engineering, which can lead to severe financial loss, reputational damage and loss of trust.

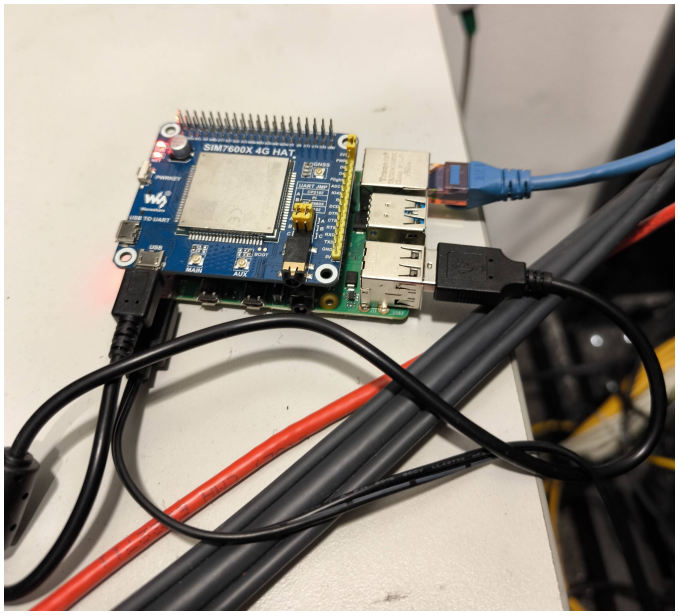


Fig. 3. Fully functional Embedded System

When grades were manually entered into the system, the above threats did not exist, but neither did the benefits of a digital system. Digitization has led to automation of processes, faster access to information, lower resource consumption and lower operating costs.

Originally, digitization was carried out in local environments and was not exposed to the public network and therefore secure. It was precisely to simplify the process for teachers, especially those with little IT knowledge, that a major threat to the digital process itself arose.

This paper examines the possibilities of implementing adequate security protection of the system as simply as possible, given that the existing computer system does not support the concept of protection and is not in the focus in the near future to be developed. The aim is to propose a solution that is easy to implement even for teachers with little IT knowledge and that can be easily set up and kept in operation to be suppoered for a large number of teachers and external collaborators.

The contribution of the article is to propose a protection concept and to describe the architecture and implementation of a protection system with double authentication. The aforementioned system is a sufficient solution until the aforementioned problem is solved at the level of applications developed by the University Computing Center in Zagreb through the NIAS concept, which is supported as a comprehensive authorization concept in the Republic of Croatia.

References

[1] R. Wash and E. Rader, 'Prioritizing security over usability: Strategies for how people choose passwords', JOURNAL OF CYBERSECURITY, vol. 7, no. 1, pp. 1–17, 2021.

[2] L. Tam, M. Glassman, and M. Vandenwauver, 'The psychology of password management: a tradeoff between security and convenience', BEHAVIOUR & INFORMATION TECHNOLOGY, vol. 29, no. 3, pp. 233–244, 2010.
[3] E. Stobert and R. Biddle, 'The Password Life Cycle', ACM TRANSACTIONS ON PRIVACY AND SECURITY, vol. 21, no. 3, Jun. 2018.
[4] J. Li, L. Stecker, E. Zeigler, T. Holland, and D. Liang, 'Scramble the Password Before You Type It', in TRENDS AND ADVANCES IN INFORMATION SYSTEMS AND TECHNOLOGIES, VOL 2, 2018, vol. 746, pp. 1097–1107.
[5] I. A. M. Abass, L. F. Hussein, T. Kallel, and A. Ben Aissa, 'New Textual Authentication Method to Resistant Shoulder-Surfing Attack', INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS, vol. 13, no. 1, pp. 490–496, Jan. 2022.
[6] B. Ur, 'SIGCHI Outstanding Dissertation Award - Supporting Password Decisions with Data', in CHI 2018: EXTENDED ABSTRACTS OF THE 2018 CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2018.
[7] M. Mehrubeoglu and V. Nguyen, 'Real-time Eye Tracking for Password Authentication', in 2018 IEEE INTERNATIONAL CONFERENCE ON CONSUMER ELECTRONICS (ICCE), 2018.
[8] A. Braeken and A. Touhafi, 'Efficient Mobile User Authentication Service with Privacy Preservation and User Untraceability', in PROCEEDINGS OF 2020 5TH INTERNATIONAL CONFERENCE ON CLOUD COMPUTING AND ARTIFICIAL INTELLIGENCE: TECHNOLOGIES AND APPLICATIONS (CLOUDTECH'20), 2020, pp. 39–46.
[9] E. D. Dimaunahan, A. H. Ballado Jr, F. R. G. Cruz, and J. C. Dela Cruz, 'MFCC and VQ Voice Recognition Based ATM Security for the Visually Disabled', in 2017 IEEE 9TH INTERNATIONAL CONFERENCE ON HUMANOID, NANOTECHNOLOGY, INFORMATION TECHNOLOGY, COMMUNICATION AND CONTROL, ENVIRONMENT AND MANAGEMENT (IEEE HNICEM), 2017.
[10] Z. Xu, J. Xu and L. -D. Kuang, "A Token-based Authentication and Key Agreement Protocol for Cloud Computing," 2021 IEEE 6th International Conference on Smart Cloud (SmartCloud), Newark, NJ, USA, 2021, pp. 38-43, doi: 10.1109/SmartCloud52277.2021.00014.
[11] Y. Wang, 'Password Protected Smart Card and Memory Stick Authentication against Off-Line Dictionary Attacks', in INFORMATION SECURITY AND PRIVACY RESEARCH, 2012, vol. 376, pp. 489–500.
[12] F. Aloul, S. Zahidi and W. El-Hajj, "Two factor authentication using mobile phones," 2009 IEEE/ACS International Conference on Computer Systems and Applications, Rabat, Morocco, 2009, pp. 641-644, doi: 10.1109/AICCSA.2009.5069395.

Applying the Technology Acceptance Model to Predict the Adoption of IoT Technologies in Smart Homes

Marko Mikša
University of Applied Sciences
Hrvatsko Zagorje Krapina,
Krapina, Croatia
marko.miksa@vhzk.hr

Filip Šklebar
University of Applied Sciences
Hrvatsko Zagorje Krapina
Krapina, Croatia
filip.sklebar@vhzk.hr

Nenad Sikirica
University of Applied Sciences
Hrvatsko Zagorje Krapina
Krapina, Croatia
nenad.sikirica@vhzk.hr

***Abstract*—The Internet of Things (IoT) is the foundation of smart home development, as it offers greater comfort, security, and energy efficiency. The acceptance of these solutions depends on users' attitudes and perceptions, which can be assessed prior to implementation using the Technology Acceptance Model (TAM). TAM assumes that perceived ease of use (PEOU) influences perceived usefulness (PU), and both constructs shape the behavioral intention to use (BI), whitch refers to the user's intention or willingness to use the technology. The proposed framework is based on a survey with a Likert scale, measurement of key constructs, and regression analysis. This paper outlines guidelines for conducting such an assessment, without applying them to a specific case. It is expected that simpler and more useful systems will significantly increase the intention to adopt IoT.**

Keywords—technology adoption, internet of things, IoT, smart homes, technology acceptance model, TAM

I. INTRODUCTION

The technology known as the Internet of Things (IoT) encompasses multiple interconnected electronic devices that communicate with each other without human assistance, often using smart scenarios, all to make users' lives easier [1]. Smart homes, based on IoT devices such as thermostats, sensors, security cameras, lighting, smart sockets, and similar devices, can provide greater comfort, energy efficiency, and security [1]. Despite their technical availability, the adoption of these solutions in practice depends on the attitudes and perceptions of end users. The Technology Acceptance Model (TAM) is one of the most commonly used frameworks for assessing technology acceptance before implementation. The core constructs of TAM are [2] [3]:

1. Perceived usefulness (PU) – the extent to which a person believes that using a technology enhances their performance or quality of life;
2. Perceived ease of use (PEOU) – the degree to which using the technology is free of effort;
3. Attitude toward use (ATT) – the individual's overall evaluative judgment regarding the technology;
4. Behavioral intention to use (BI) – Intention of the individual to use the technology strongly predicts actual usage.

II. METHOD AND MATERIAL

The examination of use and acceptance of the IoT, particularly in the context of smart homes, can be designed as a survey-based study targeting potential users of smart home technologies. The primary measurement instrument would be a structured questionnaire developed based on validated items from the TAM, carefully adapted to the IoT domain.

When the TAM is applied to the smart home context, its constructs can measure the following dimensions:

1. PU – users expect energy savings, increased security, and greater comfort;
2. PEOU – simplicity of the application, ease of connecting devices;
3. ATT – positive attitude toward using the smart home;
4. BI – intention for broader usage or purchasing new devices.

Each of these constructs would be operationalized through multiple survey items, typically measured on a Likert scale ranging from 1 = strongly disagree to 5 = strongly agree. This scale balances sensitivity with ease of response. Demographic information (e.g., age, gender, educational background, and prior experience with IoT devices) may also be included to control for user differences.

Example survey items for each construct:

1. Perceived Usefulness (PU)
    a) PU1: Using smart home devices would improve my daily life efficiency.
    b) PU2: Smart home systems would make my home more comfortable and secure.
    c) PU3: IoT-based smart homes would help me save energy and money.
2. Perceived Ease of Use (PEOU)
    a) PEOU1: Learning how to operate smart home devices would be easy for me.
    b) PEOU2: I find it easy to control smart home devices through a mobile application.
    c) PEOU3: Interacting with smart home systems would be clear and understandable.
3. Attitude Toward Use (ATT)
    a) ATT1: Using smart home technology would be a good idea.
    b) ATT2: I have a positive feeling about living in a smart home environment.
4. Behavioral Intention (BI)

a) BI1: I intend to use smart home devices in the near future.

b) BI2: I will recommend smart home technologies to others.

c) BI3: I am likely to purchase more IoT-based devices for my home.

Data analysis would include descriptive statistics to profile the sample and reliability testing (e.g., Cronbach's alpha) to assess internal consistency of constructs. Hypothesized relationships between constructs (e.g., PEOU → PU; PU and PEOU → BI) could be evaluated using regression analysis or structural equation modeling (SEM). The recommended tool for data processing is the R programming language, as it offers a wide range of statistical and analytical packages, enables easy data visualization, and is particularly suitable for survey data analysis and modeling. This methodological approach provides a systematic way to anticipate the constructs and their components that impact the adoption of IoT solutions in smart homes.

III. Expected Research Results

Based on previous research, PEOU is expected to significantly enhance PU, while both factors will serve as significant predictors of BI [2]. Therefore, users are more likely to adopt a smart home if the IoT system is simple to use, i.e., has a user-friendly interface. Adoption is further encouraged when the practical benefits of IoT, such as energy savings, security, and remote control, are clearly presented.

Several empirical studies have explored the applicability of the TAM and its extensions in predicting the adoption of IoT-based smart home systems. In line with these assumptions, de Boer, van Deursen, and van Rompay (2019) found that IoT skills strongly influence both PU and PEOU, and that perceived usefulness directly predicts IoT use, indicating that pragmatic considerations rather than attitudes drive adoption [4]. These findings are further confirmed by Nikou (2019), who demonstrated that both perceived usefulness and perceived ease of use significantly influence the intention to adopt smart home technology, while compatibility emerged as the strongest direct predictor [5]. Moreover, perceived cost was found to affect adoption negatively, underlining the role of pragmatic considerations in shaping user acceptance [5]. Using the UTAUT model as an extension of TAM, Salah Hashim and Al-Sulami (2020) showed that performance expectancy (PE) and effort expectancy (EE), which correspond to TAM's perceived usefulness and ease of use, significantly and positively influence users' behavioral intention to adopt IoT services [6].

This comprehensive approach underscores the necessity of considering both perceived ease of use and perceived usefulness, alongside other influencing factors such as user skills and facilitating conditions, for a holistic understanding of IoT adoption [4]. In addition, Zheng et al. (2018) highlight that adoption decisions are not only shaped by these established TAM variables but also by privacy concerns specific to the physical nature of IoT devices [7]. Their interviews with smart home owners revealed that users often trade privacy for convenience, rely heavily on trust in manufacturers, and remain unaware of risks stemming from inference algorithms applied to non-audio/visual data, making privacy perceptions a critical determinant of adoption [7].

In line with these expected outcomes, the proposed methodological framework illustrates how such results would be generated in an empirical application. Survey data would first be analysed to assess construct reliability and validity, followed by statistical testing of hypothesised relationships between perceived usefulness, perceived ease of use, and behavioural intention. The interpretation of results would be conducted only after the presentation of these statistical findings, focusing on the strength and significance of relationships rather than descriptive trends alone. Spatial variables, such as geographic location or type of settlement, could further enable comparative analysis of adoption patterns across different contexts. This illustrative deployment demonstrates the capacity of the proposed methodology to support meaningful inference in future empirical studies.

IV. Discussion

The application of the TAM in the context of smart homes provides researchers and industry with a clearer understanding of the factors influencing the acceptance of new technologies. The practical implications are evident: manufacturers should invest in intuitive and user-friendly interfaces, communicate benefits such as security and energy efficiency more transparently, and actively address concerns related to privacy and data protection [1].

Although this study is conceptual, its value lies in illustrating how empirical research could be designed using TAM as a theoretical foundation [2] [3]. Future studies could combine TAM with extended models such as UTAUT, which incorporates additional factors like social influence and facilitating conditions, thereby providing a more comprehensive understanding of smart home adoption across diverse contexts [8]. Prior empirical studies support this perspective by showing that user skills, perceived usefulness, and ease of use remain central determinants of IoT adoption [4][5]. Extended models such as UTAUT further emphasize the role of performance and effort expectancy [6]. At the same time, IoT-specific privacy concerns highlight the importance of trust in manufacturers and user awareness of data risks [7].

Several practical and contextual challenges should be acknowledged. Measurement validity may vary depending on how TAM constructs such as perceived usefulness, perceived ease of use, and behavioral intention are operationalized, while cultural differences among respondents may further influence perceptions of technology, privacy, and trust, thereby reducing the comparability of results across contexts. In addition, the use of self-reported data may introduce response bias, particularly with regard to attitudes toward privacy and data security. Beyond these methodological considerations, the application of TAM in smart home environments reveals inherent contextual limitations, as such technologies operate within highly personal and domestic settings where acceptance decisions may be shaped by shared household use, trust in manufacturers, and perceived data sensitivity. Consequently, standard TAM relationships may be moderated by privacy-related and contextual factors that are not fully captured by perceived usefulness and perceived ease of use alone. Therefore, models of smart home adoption should not only account for usefulness and ease of use but also integrate broader digital skills, contextual factors, and perceptions of privacy. This comprehensive approach is essential for both researchers and practitioners in developing strategies that foster the responsible and sustainable adoption of IoT technologies.

V. Conclusion

This study contributes by offering a conceptually grounded framework for applying TAM to the investigation of IoT adoption in smart home environments. Although no

empirical validation has been conducted, the framework highlights key contextual and privacy-related challenges that future research should consider when operationalizing TAM in domestic IoT settings.

An analysis of existing smart homes with implemented IoT technologies would indicate the continued relevance of TAM as a theoretical framework in this context. Perceived usefulness and perceived ease of use remain key factors influencing users' attitudes toward the adoption of IoT technologies, while future research is encouraged to extend the model by incorporating additional constructs such as trust, privacy concerns, spatial aspects, and digital literacy.

A deeper understanding of these aspects will enable both researchers and industry professionals to design smarter, safer, and more technologies tailored to the user. In this way, TAM continues to serve not only as a main model but also as a flexible framework capable of evolving with emerging technologies.

REFERENCES

[1] Matasović, I., Galović, M., Kokanović, M. i Crnac, Z. (2025). Sigurnosni izazovi IOT uređaja. Festung, 1 (1), 73-80. Available at: https://hrcak.srce.hr/334235

[2] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319 –340. Available at: https://doi.org/10.2307/249008

[3] Marangunić, N. i Granić, A. (2012). TAM - četvrt stoljeća istraživanja. Suvremena psihologija, 15 (2), 205-223. Available at: https://hrcak.srce.hr/111605

[4] de Boer, P. S., van Deursen, A. J. A. M., & van Rompay, T. J. L. (2019). Accepting the Internet-of-Things in our homes: The role of user skills. Telematics and Informatics, 36, 147–156. https://doi.org/10.1016/j.tele.2018.12.004

[5] Nikou, S. (2019). Factors driving the adoption of smart home technology: An empirical assessment. Telematics and Informatics, 45, 101283. https://doi.org/10.1016/j.tele.2019.101283

[6] Salah Hashim, H., & Al-Sulami, Z. A. (2020). A model of factors influencing users' adoption of Internet of Things services: A case study of Iraqi educational institutions. IOP Conference Series: Materials Science and Engineering, 769(1), 012006. https://doi.org/10.1088/1757-899X/769/1/012006

[7] Zheng, S., Apthorpe, N., Chetty, M., & Feamster, N. (2018). User Perceptions of Smart Home IoT Privacy. Proceedings of the ACM on Human-Computer Interaction, 2, 1. https://doi.org/10.1145/3274469

[8] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425–478. https://doi.org/10.2307/30036540

Latency Wars Beyond the Last Hop: A Comparative Study of Starlink and 5G/LTE Routing Using RIPE Atlas and Khipu

Stefanović Katarina
Radio Access Network Optimization
Engineer and RIPE Atlas Ambassador
Telecom Serbia
Belgrade, Serbia
0009-0005-9847-8620
katarinastef@telekom.rs

***Abstract*— Reliable low-latency connectivity is a critical requirement for modern Internet services, especially for remote, underserved, and long-haul communication scenarios where terrestrial infrastructure may be limited or inefficient. The rapid deployment of Low Earth Orbit (LEO) satellite constellations has introduced a new paradigm in global Internet connectivity, challenging traditional terrestrial mobile networks. This paper compares latency and routing performance between Starlink satellite broadband and LTE/5G mobile networks using RIPE Atlas measurements. ICMP-based ping and traceroute data toward distant targets (e.g., Singapore) are analyzed using Khipu to visualize hop count, routing structure, and failure patterns. Results show that Starlink achieves lower RTTs (~177 ms) over 15–17 hops, while LTE/5G paths are longer (up to 31 hops) and more variable (~260–350 ms). Starlink routes remain largely within AS14593 and show fewer intercontinental detours. In contrast, mobile traces often suffer from CGNAT delays and core-related latency. These findings suggest that modern LEO satellite systems can match or outperform terrestrial mobile networks on long-haul routes. Future work will explore TLS latency, 4G vs. 5G differences, IPv4 vs. IPv6 paths, and Starlink's inter-satellite routing effects.**

Keywords—LEO, RIPE atlas, measurements, LTE, NR, NTN

I. INTRODUCTION

The modern world continues to face a pronounced digital divide in both internet access and usage. While mobile networks today cover the vast majority of the global population, a significant portion of people still lacks access to reliable broadband connectivity. According to the GSMA report [1], approximately 4% of the global population (~330 million people) remains outside the footprint of mobile broadband networks (coverage gap), while an additional 39% (~3.1 billion people) live within coverage but do not use mobile internet due to affordability, digital literacy, or device limitations (usage gap). In total, over 3.4 billion people—about 37% of the world's population—remain offline and excluded from the digital economy.

Beyond the population divide, a spatial divide also exists: terrestrial mobile networks cover only around 15% of the Earth's surface, leaving vast areas such as oceans, mountains, deserts, and remote rural zones disconnected. This coverage gap not only affects everyday users but also professionals and outdoor populations. For instance, a recent study by the Bullitt Group [2] reports that over 57 million hikers, 60 million anglers, 9 million skiers, and more than 1 million emergency responders often operate in areas completely lacking connectivity. Similarly, workers in agriculture, forestry, freight, and remote construction regularly encounter communication "dead zones" .

To address these limitations, satellite-based networks—particularly Low Earth Orbit (LEO) constellations—have emerged as a promising solution. Unlike traditional Geostationary Earth Orbit (GEO) satellites at ~36,000 km, LEO satellites operate at altitudes between 500 and 2,000 km, enabling lower latency and higher throughput. Starlink, for example, already operates over 6,000 active LEO satellites, offering global broadband with latency and speed comparable to terrestrial ISPs [3], [4].

As shown in Fig. 1, the number of global subscribers to satellite-based services—particularly in the areas of IoT, broadband, and Non-Terrestrial Networks (NTN) mobile—is projected to increase significantly through 2030. This growth underscores both the rising reliance on NTN solutions and the urgency of evaluating their performance and integration with terrestrial systems.

To systematically assess the performance of LEO-based access compared to mobile terrestrial networks, this study leverages the RIPE Atlas measurement platform [5]. Through active measurements such as ping and traceroute, we analyze latency, jitter, and routing characteristics between Starlink LEO and LTE/5G networks across Europe and intercontinental links. The visual analysis tool Khipu is used to trace route stability and detect anomalies.

This paper is structured as follows. Section II provides an overview of NTN architecture and its integration into 3GPP standards. Section III describes the measurement methodology using RIPE Atlas, probe selection, and target distribution. Section IV presents latency and routing results with comparative visualizations. Section V concludes with key findings and outlines future work, including deeper protocol-level measurements, IPv4 vs. IPv6 comparison, and analysis of Starlink's inter-satellite link (ISL) impact. Section VI provides the references and data sources used for validation and reproduc

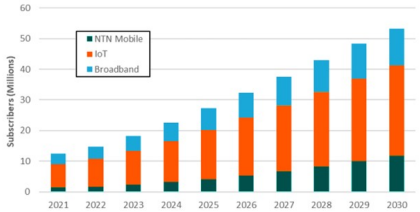


Fig. 1. Projected global growth of NTN subscribers by service category (IoT, broadband, NTN Mobile), 2021–2030. Source: ABI Research.

This paper is structured as follows. Section II provides an overview of NTN architecture and its integration into 3GPP standards. Section III describes the measurement methodology using RIPE Atlas, probe selection, and target distribution. Section IV presents latency and routing results with comparative visualizations. Section V concludes with key findings and outlines future work, including deeper protocol-level measurements, IPv4 vs. IPv6 comparison, and analysis of Starlink's inter-satellite link (ISL) impact. Section VI provides the references and data sources used for validation and reproducibility.

II. Satellite Communication Architecture

A. LEO Satellite Architectures and 3GPP NTN Integration

LEO satellites, enable significantly lower propagation delays and improved link performance compared to geostationary systems [6], [7]. The deployment of large-scale LEO constellations such as Starlink, OneWeb, and Project Kuiper has led the 3rd Generation Partnership Project (3GPP) to incorporate support for NTN into the 5G standard, starting with Release 17 (2022) [8], [9]. These standards primarily adopt a "transparent payload" model, where the satellite relays user traffic while all processing remains in the terrestrial infrastructure [10].

3GPP defines four primary NTN architectures (Fig. 2), each suited for different deployment scenarios [11]–[13]. The Transparent Satellite model (Rel-17) relays user traffic to a ground-based gNB, offering simple implementation but requiring dense ground coverage. The Regenerative Satellite model (planned for Rel-19) integrates gNB functions onboard the satellite, enabling direct service in remote areas. The CU/DU Split model (Rel-18) places the Distributed Unit (DU) on the satellite and retains the Central Unit (CU) on the ground, optimizing bandwidth usage on the feeder link. Finally, the Satellite Backhaul model allows terrestrial gNBs to connect to the 5G Core via LEO satellites, useful for isolated or emergency scenarios.

Each architecture introduces specific trade-offs between complexity, coverage, latency, and resource flexibility. Accurate user location and timing, typically provided via GNSS, are critical for synchronization and beam tracking [14]. Moreover, LEO satellite motion introduces Doppler effects and frequent handovers, necessitating specialized mobility and scheduling mechanisms in the radio access network (RAN) [15-16].

Commercial NTN systems reflect these designs: Starlink implements regenerative payloads with optical inter-satellite links (ISLs), supporting global routing independent of ground relays [17]. OneWeb and Kuiper emphasize transparent and hybrid payloads, targeting mobile and enterprise backhaul [18], [19]. Additionally, systems like AST SpaceMobile and Lynk Global aim to enable direct-to-device (D2D) connectivity from unmodified smartphones [20], [21].

NTN use cases span across rural IoT, maritime and aviation connectivity, emergency response, and resilient or primary backhaul for under-connected regions [22], [23]. As part of the 5G multi-connectivity framework, NTNs enhance global coverage, enabling seamless switching between terrestrial and satellite links [24].

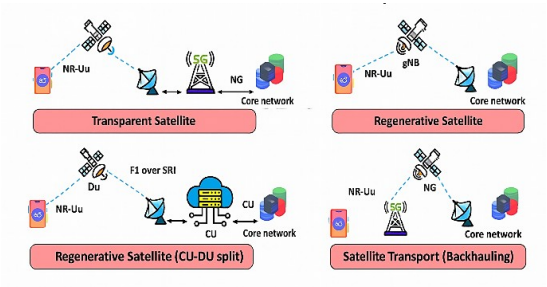


Fig. 2. Overview of the four 3GPP-defined NTN architectures: (a) Transparent Satellite (Rel-17), (b) Regenerative Satellite (planned Rel-19), (c) CU/DU Split Satellite (Rel-18), and (d) Satellite Backhaul (Rel-18). Each path shows the location of gNB functions, user-plane routes, and interaction with the 5G Core Network.

III. Methodology

This study aims to evaluate and compare the end-to-end latency and routing characteristics of Starlink's LEO satellite broadband and terrestrial LTE/5G mobile networks across Europe. The RIPE Atlas platform was employed for conducting systematic active measurements under unified conditions. This measurement methodology follows similar approaches in previous studies that utilized RIPE Atlas to evaluate latency and routing behavior across mobile and satellite networks [24]–[26]. However, measurement-based comparison between LEO satellite broadband and LTE/5G mobile networks across long-haul intercontinental paths remains limited. The overarching motivation lies in quantifying whether modern LEO systems can achieve latency performance comparable to or better than terrestrial mobile infrastructure, particularly in long-haul and latency-sensitive contexts.

Starlink was selected as the satellite benchmark due to its innovative LEO architecture, which advertises round-trip times (RTTs) as low as 20–50 ms—significantly outperforming legacy GEO systems (>600 ms). LTE/5G mobile probes served as terrestrial baselines, reflecting widespread real-world deployments across the European market. Despite inherent complexity in mobile radio access and core processing, LTE/5G remain dominant for rural and mobile broadband access.

The working hypothesis is that Starlink can achieve comparable or superior RTTs to LTE/5G under stable network conditions. To test this, a measurement campaign was designed involving geographically distributed and technologically diverse probes.

A. Measurement Targets and Scheduling

A curated set of probes was deployed to ensure diversity in terms of geography, network operators, and access technologies. Table I presents a sample configuration of probes included in the experiment.

Each probe conducted ICMP echo request (ping) measurements toward five target IPs representing a range of topological distances:

- Local Anycast DNS nodes: Google (8.8.8.8), Cloudflare (1.1.1.1)
- Continental anchors: RIPE Atlas nodes in Frankfurt (217.243.20.18) and Amsterdam (193.0.0.164)
- Long-haul anchor: Singapore node (139.180.132.31, AS20473)

Measurements were performed for ten consecutive days, collecting three pings per target per cycle.

TABLE I. SAMPLE PROBES USED IN MEASUREMNET

PROBES		
Technology/Operator	***Country/ASN***	***Notes***
Satellite/Starlink (SpaceX)	Germany/14593	Very stable 90d
Satellite/Starlink (SpaceX)	Germany/14593	Stable 30d
Satellite/Starlink (SpaceX)	France/14593	Stable 30d
Satellite/Starlink (SpaceX)	Netherlands/14593	Reliable, stable 90d
Satellite/Starlink (SpaceX)	Bulgaria/14593	Stable 30 d
LTE/Vodafon.DE	Germany/3209	Very stable 90d
LTE/T-Mobile	Germany/3320	Stable 30d
LTE/Free Mobile	France/51207	Stable 30d
LTE/Vodafon.NL	Netherlands/13308	Stable 30d
5G/Headoff.com	Bulgaria/12716	Stable 30d

Each probe conducted ICMP echo request (ping) measurements toward five target IPs representing a range of topological distances:

- Local Anycast DNS nodes: Google (8.8.8.8), Cloudflare (1.1.1.1)
- Continental anchors: RIPE Atlas nodes in Frankfurt (217.243.20.18) and Amsterdam (193.0.0.164)
- Long-haul anchor: Singapore node (139.180.132.31, AS20473)

Measurements were performed for ten consecutive days, collecting three pings per target per cycle.

B. Traceroute Configuration

To enrich latency measurements with routing path data, UDP-based Paris traceroutes were configured as follows:

- Protocol: User Datagram Protocol (UDP) over port 80
- Maximum Time To Live (TTL): 32 hops
- Probes per flow: 3 packets
- Paris flows: 16 (to capture route diversity)
- Timeout per response: 4000 ms
- Duplicate timeout: 10 ms
- Packet size: 48 bytes
- Don't Fragment: disabled

This configuration supports detection of Carrier-Grade Network Address Translation (CGNAT), mobile core encapsulation, and dynamic handovers in both mobile and satellite environments.

C. Metrics and Analysis Strategy

Four key performance indicators (KPIs) were selected:

- RTT Distribution: Minimum, average, and percentile RTT values across targets
- Jitter: Measured as interquartile range (IQR) of RTTs
- Packet Loss: Percentage of failed ping responses
- Path Depth and Variability: Number of traceroute hops and route changes

These metrics enable a comparative evaluation of latency efficiency and stability. Starlink is hypothesized to provide lower jitter and more consistent paths due to vertically integrated infrastructure and localized ground egress. In contrast, LTE/5G is expected to show higher RTT variation due to scheduling delays, CGNAT traversal, and core network processing.

IV. RESULTS

A. RTT Behavior and Platform-Specific Trends

The latency measurements were conducted continuously over 11 consecutive days, encompassing multiple test cases for both the Starlink LEO satellite and terrestrial LTE/5G mobile networks. The results presented in Fig. 3 and 4 illustrate the variation of RTT for each test case, highlighting both minimum and average values per measurement instance/per case/per day.

For the LTE/5G mobile network, the minimum RTT varied between approximately 4 ms and 30 ms, depending on factors such as server proximity, path optimization, and routing efficiency. In Case 2, an anycast route to a nearby Domain Name System (DNS) node resulted in a notably low RTT of below 5 ms, demonstrating that under ideal network conditions, mobile technologies can achieve near edge-level responsiveness.

However, the average RTT values were significantly higher—typically within the 75–85 ms range—reflecting the impact of longer end-to-end routes and suboptimal backbone paths. In several scenarios, traffic routed through distant anchor points, particularly towards Asian regions, elevated RTTs to beyond 250 ms. Occasional improvements to approximately 187 ms were observed when optimized transit paths were used.

Minor packet loss events occurred sporadically, which is consistent with ICMP rate-limiting or filtering mechanisms employed in mobile core networks.

In contrast, the Starlink LEO satellite network exhibited remarkably stable and low-latency performance. The minimum RTT consistently ranged between 15–18 ms, corresponding well with theoretical propagation delays for satellites orbiting at ~550 km altitude. The average RTT remained steady between 55–60 ms, with modest increases observed for transcontinental paths—most notably to East Asia (~180 ms).

Throughout all test days, no packet loss was recorded, and latency stability remained exceptional, underscoring the robustness of Starlink's routing architecture and congestion control mechanisms.

Overall, the data confirms that while terrestrial LTE/5G mobile networks can occasionally achieve lower instantaneous latency through localized routing, Starlink's LEO infrastructure provides superior stability, predictability, and global performance uniformity across measurement intervals.

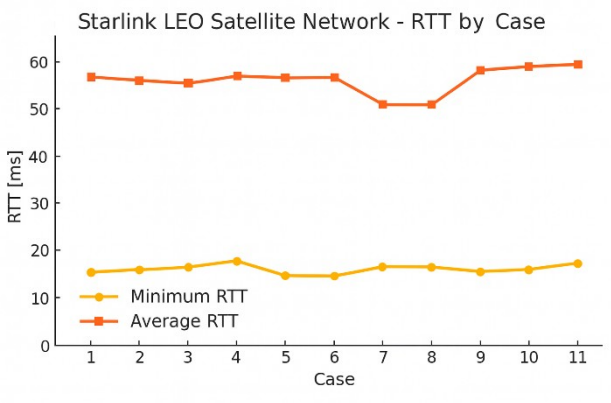


Fig. 3. Round-Trip Time (RTT) per case over 11 days of measurement Starlink LEO satellite network

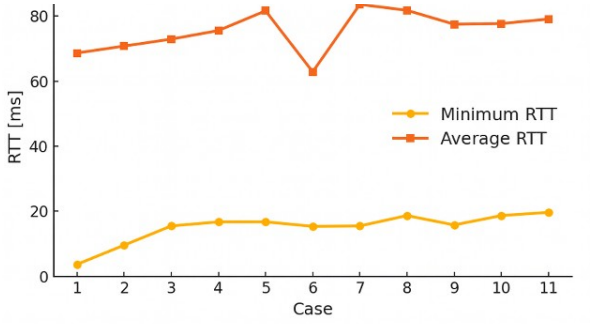


Fig. 4. Round-Trip Time (RTT) per case over 11 days of measurement for LTE/5G mobile networks.

These results suggest that Starlink's RTT is not only within terrestrial broadband ranges but exhibits superior consistency across geography and time. The implications are particularly relevant for hybrid terrestrial-satellite 5G networks.

B. *Jitter Characteristics and Temporal Stability*

The analysis of jitter, expressed as IQR of RTT variation, reveals clear performance differences between the mobile (LTE/5G) and Starlink LEO networks (see Fig. 5). Across all 11 measurement cases, the mobile network exhibited substantially higher jitter, typically within the 55–75 ms range, with notable peaks in Cases 1 and 2, where route switching and transient Radio Resource Control (RRC) state transitions likely increased delay variability. Beyond these initial cases, jitter stabilized slightly but remained elevated compared to satellite performance.

Conversely, the Starlink network demonstrated consistently low jitter values, remaining within a narrow IQR band of 37–41 ms, indicative of stable queuing and routing behavior throughout the 11-day observation period. The minimal dispersion in Starlink's jitter distribution highlights the deterministic latency behavior characteristic of LEO satellite systems, which are less influenced by terrestrial congestion and handover events.

Overall, these results confirm that while both technologies achieve comparable median latency values, Starlink provides far superior latency stability—a key factor for time-sensitive applications such as real-time streaming, Voice over Internet Protocol (VoIP), and low-latency gaming.

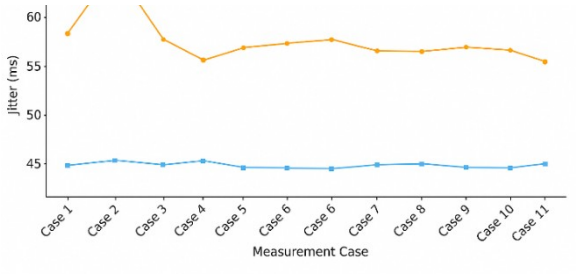


Fig. 5. Jitter per case over 11 days of measurement for LTE/5G mobile networks and Starlink

C. *Khipu-Based Visualization and Long-Haul Path Diagnostics*

To investigate routing dynamics and long-haul performance across satellite and terrestrial networks, this study employs Khipu — an interactive traceroute visualization platform developed by RIPE NCC [26]. Additionally, this analysis enables the comparison of jitter characteristics between terrestrial and satellite paths, highlighting the typically higher jitter observed in mobile networks due to dynamic routing and core processing, in contrast to the more stable latency patterns of Starlink's LEO infrastructure. Khipu provides two core views of traceroute data collected via RIPE Atlas: an IP-level view, where each hop is displayed individually, and an ASN-level aggregation, where nodes are grouped by their respective Autonomous Systems. This structure enables intuitive interpretation of route stability, divergence points, and per-hop performance. In addition, Khipu allows filtering paths based on AS number, country, latency thresholds, or reachability, facilitating efficient comparison between network types, as illustrated in Fig. 6.

To illustrate the contrast in routing performance and structure between satellite and mobile networks, we selected two representative traceroute cases for in-depth analysis using Khipu. These two cases capture both failure and success scenarios, making them ideal for demonstrating Khipu's capabilities in visualizing performance-critical routing patterns across different network technologies.

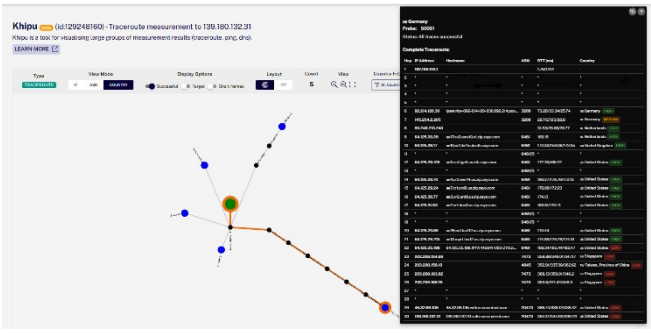


Fig. 6. Traceroute visualization using the Khipu platform.

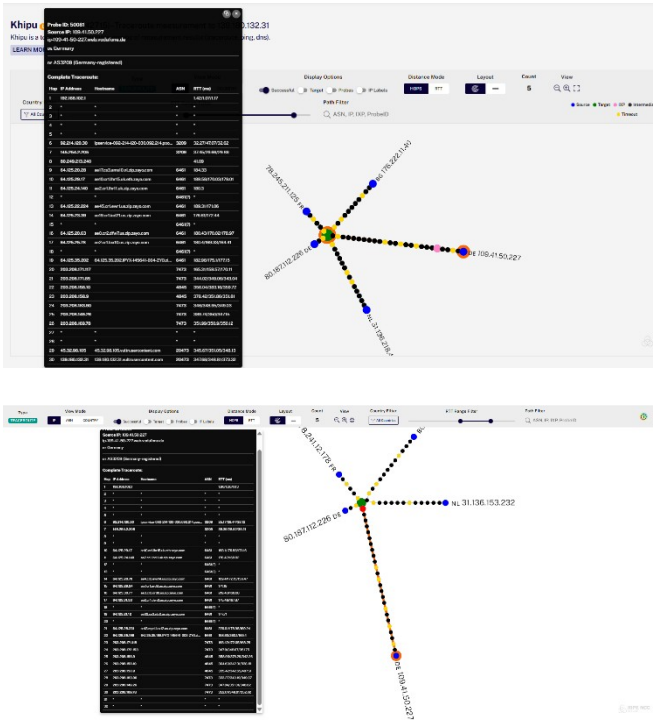


Fig. 7. Failed mobile traceroute and successful mobile traceroute to the same target

Figure 7 contrasts two traceroute attempts from the same LTE probe in Germany to a target in Singapore. The first trace failed after 32 hops, likely due to ICMP filtering, CGNAT, or upstream peering issues. This is visualized in Khipu as a red terminal node with truncated path segments. The second trace succeeded, reaching the target in 31 hops, but exhibited higher latency (~355 ms) and a more complex intercontinental path than typical Starlink traces

Such incomplete traceroutes are common in mobile networks, where intermediate nodes often block ICMP "TTL expired" replies. As explained in NetBrain's traceroute guide [28], routers may deprioritize or drop these replies due to processing overhead. Additionally, mobile paths traverse the core network (e.g., SGW/PGW or 5G UPF), introducing latency through NAT gateways and multiple processing layers.

These contrasting results—failure versus success from the same origin probe—highlight the inherently variable nature of mobile network routing over long-haul paths. In contrast, Starlink's paths remain shorter (15–17 hops), more predictable, and operate primarily within a single ASN (AS14593), resulting in significantly lower and more stable RTTs. This is attributed to Starlink's direct ground station routing and "bent-pipe" satellite relaying. Overall, the contrasting outcomes—failure versus success from the same origin—underscore the dynamic and less predictable nature of long-haul mobile routing compared to Starlink's more stable, latency-optimized design. This difference is further quantified in Fig. 8, where RTT values from Germany to Singapore are consistently lower and more stable over Starlink, compared to LTE/5G.

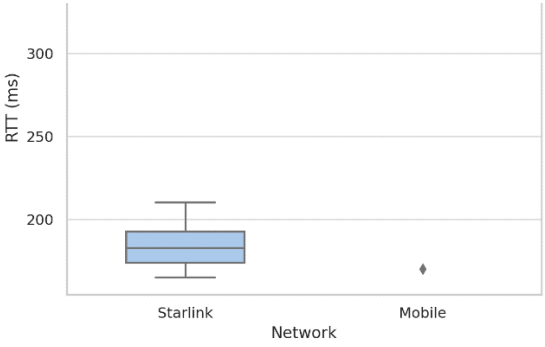


Fig. 8. RTT Distribution from Germany Probes to Singapore Target: Starlink vs. Mobile Networks

Conclusion and Future Work

This paper presented a comprehensive latency and routing analysis comparing Starlink's LEO) satellite network with terrestrial LTE/5G mobile infrastructures, focusing on both intra-European and intercontinental communication. Results show that for intra-European destinations, Starlink and mobile probes exhibit comparable RTTs, typically in the 15–35 ms range. In several cases, Starlink delivered lower latency and jitter, benefiting from direct peering and reduced core processing. Conversely, mobile probes exhibited higher RTT variability due to EPC overhead, CGNAT traversal, and radio scheduling delays/congestion.

For long-haul destinations such as Singapore, Starlink consistently outperformed LTE/5G mobile probes, averaging ~177 ms over 15–17 hops. Mobile networks frequently exceeded 270–350 ms over more than 30 hops and occasionally failed to complete traceroutes. Khipu-based visualizations confirmed that Starlink routes are typically shorter, more direct, and remain within AS14593 until egress, whereas mobile traces often traverse multiple Tier-1 ASes with greater variability and detours. These findings suggest that modern LEO satellite systems can match—or in some cases outperform—terrestrial mobile networks in terms of global reach and latency consistency.

Future work will focus on the following directions to deepen the analysis of satellite and mobile performance:

- Application-Layer Latency Evaluation: Measure Transport Layer Security (TLS) handshake latency and complete application setup times, extending beyond ICMP to reflect end-user experience more accurately.
- Expanded Starlink Coverage and Geographic Scope: Incorporate all available Starlink-enabled RIPE Atlas probes and expand measurements across multiple continents to enable broader comparisons (e.g., intra-Asia, transatlantic, and African regions).
- 4G vs. 5G Core Latency Comparison: Include mobile probes on both 4G and 5G networks to assess whether core network enhancements yield significant end-to-end latency improvements relative to Starlink.
- Internet Protocol version 4 (IPv4) vs. Internet Protocol version 6 (IPv6) Routing Analysis: Use dual-stack probes to compare protocol-specific path characteristics, investigating how different peering policies affect satellite and mobile performance.

- Detection of Routing Anomalies and Policy-Based Divergence: Leverage Khipu's traceroute visualizations to detect asymmetries, routing loops, and policy-driven detours in backbone infrastructures.
- Impact of Starlink's ISL: Assess how the activation of ISLs affects long-haul routing efficiency and whether space-based forwarding further reduces latency compared to terrestrial backbones.

These directions aim to construct a richer and more nuanced understanding of latency, routing behavior, and architectural trade-offs in next-generation non-terrestrial and terrestrial broadband systems.

REFERENCES

[1] GSMA Intelligence, "The State of Mobile Internet Connectivity 2024," GSMA, 2024.

[2] Bullitt Group, "Off-grid communication report," 2022.

[3] Rohde & Schwarz, "Satellite connectivity and off-grid communication," Taipei Technical Report, 2022.

[4] BCSatellite.net, "Low Earth Orbit Constellations: Architecture and Challenges," 2023.

[5] RIPE NCC, "RIPE Atlas: A global Internet measurement network," 2024.

[6] M. Giordani et al., "Non-Terrestrial Networks in 5G & Beyond: A Survey," IEEE Communications Surveys & Tutorials, vol. 23, no. 4, pp. 2801–2841, 2021.

[7] H. Wang et al., "LEO Satellite Communication for 5G and Beyond," IEEE Network, vol. 35, no. 5, pp. 80–86, 2021.

[8] 3GPP TR 38.811, "Study on New Radio (NR) to support non-terrestrial networks," V15.4.0, Sept. 2020.

[9] 3GPP TR 38.821, "Solutions for NR to support non-terrestrial networks (NTN)," V16.0.0, Dec. 2020.

[10] 3GPP Release 17 Overview, https://www.3gpp.org/release-17

[11] A. Guidotti et al., "Architectures and Key Technical Challenges for 5G Systems Incorporating Satellites," IEEE Transactions on Vehicular Technology, vol. 68, no. 3, pp. 2624–2639, 2019.

[12] M. Rinaldi et al., "On the Integration of Satellite and Terrestrial Networks for Future 5G," in Proc. EuCNC, 2018.

[13] 3GPP TR 38.863, "Study on CU/DU split over satellite for NR," V0.2.0, Sept. 2023.

[14] S. Andrenacci et al., "Satellite Backhauling in Future 5G Networks," in Proc. IEEE Globecom, 2022.

[15] K. Liolis et al., "Challenges and Enablers for Satellite Integration into 5G," IEEE Network, vol. 32, no. 5, pp. 54–61, 2018.

[16] J. Sachs et al., "5G Radio Access for Ultra-Reliable and Low-Latency Communications," in Proc. IEEE VTC-Spring, 2018.

[17] SpaceX, "Starlink Mission Page," [Online]. Available: https://www.spacex.com/launches/

[18] OneWeb, "Connecting the World," [Online]. Available: https://www.oneweb.world/

[19] Amazon Kuiper, "Project Kuiper Overview," [Online]. Available: https://www.aboutamazon.com/

[20] AST SpaceMobile, "Space-Based Cellular Broadband," [Online]. Available: https://ast-science.com/

[21] Lynk Global, "Direct-to-Phone Connectivity," [Online]. Available: https://lynk.world/

[22] Keysight Technologies, "NTN Test Solutions for 5G," Whitepaper, 2023.

[23] 5G Americas, "The Role of Non-Terrestrial Networks in 5G," 2023.

[24] R. Richter, V. Ververis, and V. Bajpai, "Breaking Through the Clouds: Performance Insights into Starlink's Latency and Packet Loss," [Online]. Available: https://arxiv.org/abs/2301.05587

[25] N. Mohan, A. E. Ferguson, H. Cech, R. Bose, P. R. Renatin, M. K. Marina, and J. Ott, "A Multifaceted Look at Starlink Performance", ACM Internet Measurement Conf. (IMC) 2022, pp. 374–387.

[26] M. Handley, "Delay is Not an Option: Low Latency Routing in Space", Proc. ACM HotNets, 2018, pp. 85–91.

[27] RIPE NCC, "Khipu: Visualising Traceroutes with RIPE Atlas," [Online]. Available: https://atlas.ripe.net/visualisations/khipu/

[28] NetBrain Technologies, "Understanding traceroute and its limitations," NetBrain, Accessed: Oct. 26, 2025. [Online]. Available: https://www.netbraintech.com

From Radial to Rectangular: A New Method for UNET-Based Path Loss Prediction

Jelena Mladenović
Department of Telecommunications
School of Electrical Engineering,
University of Belgrade
Belgrade, Serbia
jelenam@etf.rs

Aleksandar Nešković
Department of Telecommunications
School of Electrical Engineering,
University of Belgrade
Belgrade, Serbia
an@etf.rs

Nataša Nešković
Department of Telecommunications
School of Electrical Engineering,
University of Belgrade
Belgrade, Serbia
natasha@etf.rs

***Abstract*—This paper introduces a novel methodology for large-scale path loss prediction using building maps. In order to enable path loss prediction across an entire area of interest (i.e., a radio map) in a single inference step, we propose a new approach that transforms the inherently radial nature of radio signal propagation into a rectangular representation. This transformation addresses the limitations of convolutional neural networks in processing images with radial symmetry. Our method employs supervised learning and is trained on a limited set of simulated radio maps corresponding to building maps. Once trained, it can predict path loss across the entire map in under half a second, achieving an RMSE of 2.63 dB on a test set of 640 images.**

Keywords—path loss prediction, deep learning, image processing, radio communications

I. Introduction

Accurate prediction of path loss (PL) is one of the first and main steps in the development of radio systems because it represents a basis for good base station site selection and appropriate frequency planning [1]. Without it, performing coverage analysis and determining key metrics such as received signal strength, link budget, and signal-to-noise ratio becomes impossible [1], [2].

Fast and accurate path loss prediction has long been a challenging task, further intensified by the demands of emerging technologies such as 5G and beyond. Radio propagation prediction was traditionally addressed with two main approaches, which led us to statistical and deterministic radio propagation models. Statistical models are derived from measurement data and are inherently biased toward the specific environments in which those measurements were taken. They do not consider detailed environmental characteristics and instead rely on broad environment classifications (e.g., urban, suburban, rural environment), which limit their prediction accuracy. However, due to their low computational complexity and fast execution, statistical models remain a practical choice in scenarios where coarse estimation is acceptable and computational speed is a priority. On the other hand, deterministic propagation models are site-specific and highly dependent on detailed environmental characteristics. They are based on the physical laws of electromagnetic wave propagation and offer high accuracy but require detailed information about the environment, which is often difficult or even impossible to obtain. These models are computationally expensive, making them practical only in use cases where high-precision path loss estimation justifies the cost. Thus, traditional radio propagation models require a tradeoff between precision and computational efficiency, making the search for optimal solutions an ongoing area of research.

In recent years, the integration of artificial intelligence (AI) has demonstrated significant potential in addressing the challenges of radio propagation modeling. One key advantage of AI-based approaches is their ability to predict signal characteristics accurately by leveraging diverse data sources, including satellite imagery and OpenStreetMap (OSM) data of the target areas. Early AI-driven models primarily utilized multilayer perceptron (MLP) architectures and hand-crafted features, focusing on point-to-point prediction tasks (e.g., [1], [2], [3], [4]). With the advancement of deep learning, more sophisticated architectures such as convolutional neural networks (CNNs) have been introduced (e.g., [5], [6], [7]), allowing the use of raw spatial data such as images. However, relatively few studies have explored approaches in which models predict signal propagation for an entire area in a single inference step (e.g., [8], [9], [10]), rather than on a per-point basis.

The objective of this paper is to develop a novel image-to-image deep-learning approach for path loss prediction by implementing a radial-to-rectangular transformation that enables efficient UNet-based inference over entire radio maps. This transformation facilitates convolutional neural networks to more effectively capture the propagation mechanisms of radio signals coming from the base station. Our model is based on image-to-image regression for predicting path loss across an entire area in a single inference step, achieving an RMSE of 2.63 dB on a test set of 640 images. To the best of the authors' knowledge, the proposed methodology offers a new perspective for addressing the challenge of radial signal propagation prediction. The remainder of the paper is organized as follows: Section II describes the dataset and outlines the preprocessing steps used to transform radial signal patterns into a rectangular representation. Section III details the network architecture, as well as the training and evaluation procedures. Section IV presents experimental results and performance analysis. Finally, Section V concludes the paper and discusses potential directions for future work.

II. Dataset Description and Preprocessing

To generate a labeled dataset that simulates real-world network scenarios, we created 3D building maps that simulate urban environments. These maps are artificially generated using the Matlab software tool to simulate real OpenStreetMap (OSM) maps, where rectangles or combinations of rectangles represent buildings. However, unlike real OSM-based building maps, the generated maps do not follow any urbanistic planning or geometric structure. As a result, the buildings are randomly positioned within the image, and the total number of buildings per map is lower

This work was financially supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia under contract number: 451-03-137/2025-03/200103.

compared to real-world data. These limitations may lead to weaker prediction performance because the neural network encounters a different spatial organization of buildings in every sample and therefore cannot consistently learn meaningful structural patterns that could improve the prediction. Simulated building heights range from 10 to 45 meters, and the observed area corresponds to a microcell scenario. Input images have dimensions of 256×256 pixels, where each pixel represents an area that corresponds to 1 by 1 meter in a real-world scenario. Radio maps were obtained using theoretical PL calculations considering both knife-edge diffraction and free-space PL. In all cases, the transmitter is positioned at the center of the observed area, operating at a frequency of 700 MHz with a 2.15 dBi omnidirectional antenna. In accordance with typical microcell deployment specifications, the transmitter antenna height is set at 12 meters, while the receiving antenna height is set at 1.5 meters, see Table I.

TABLE I. SIMULATION PARAMETERS

Radio map area	256 × 256 m^2
Operating frequency	700 MHz
Transmit antenna gain	2.15 dBi
Receive antenna gain	2.15 dBi
Transmit antenna height	12 m
Receive antenna height	1.5 m

Positive values in decibels are used in our radio maps to represent path loss information. Fig. 1 shows one example of a building map and its corresponding radio map. Knowing the level of the transmitted signal, it is straightforward to calculate the level of the received signal as:

$$P_{RX}[dBm] = P_{TX}[dBm] - PL[dB] \quad (1)$$

Although we acknowledge that real-world measurements or ray tracing simulations are necessary for developing a highly reliable model, we used the proposed dataset because it was easy to generate. It still captured the key characteristics of signal propagation, which were sufficient for evaluating the proposed prediction methodology. However, we consider this a valid first step toward the development of more advanced models based on the proposed method.

A. Grayscale conversion

In order to create PNG images from the raw numeric dataset, radio maps and building maps are preprocessed utilizing gray conversion. Radio maps are converted into grayscales ranging from 0 to 255 using min-max normalization, with the minimum and maximum values that correspond to the lowest PL and highest PL values in the whole dataset, including both the training and testing datasets. After being saved as PNG files, these pictures are fed into the model for training. An identical preprocessing procedure is applied to building maps using minimum and maximum values extracted from building maps. In this normalization process, we take into account both training and testing datasets to enable straightforward inversion and model evaluation in dB units.

B. Radial to rectangular transformation

One particular challenge in using CNNs for radio signal prediction is the radial nature of radio signal propagation. Given that CNNs are primarily made to detect local features in images, the challenge lies in the fact that the model must learn to relate local details to the larger context. This includes understanding how buildings and distance from the transmitter together influence path loss. To address this, a new approach is proposed in which both input and target images are transformed from radial into rectangular format. This transformation is performed so that radial directions from the center of the original image become columns in the transformed image. The first radial direction is defined by the line connecting the central pixel and the pixel at position (256, 128). The process is then repeated for a predefined number of radial directions in the counterclockwise direction. All radial directions have equal length, defined as the distance from the transmitter (i.e., center of the image) to its nearest edge, which is equivalent to half of the area's length dimension. Based on a predefined number of radial directions and steps per direction, rows and columns in the new image are defined. The formal mathematical definition of the proposed transformation, *T*, is given below.

$$T: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{k \times l} \quad (2)$$

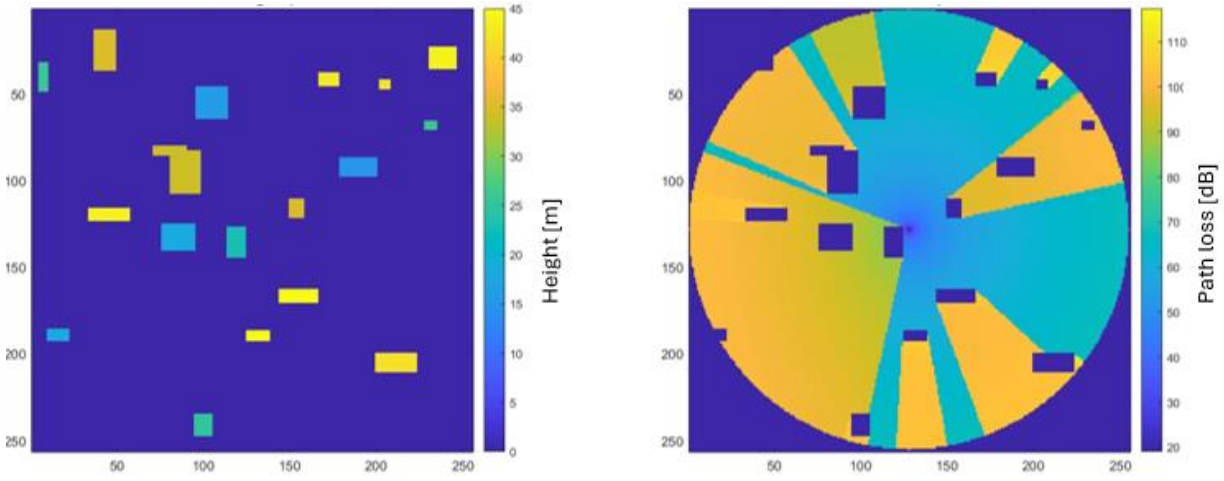
Here, m and n denote the initial number of rows and columns of the image, while k and l denote the number of rows and columns in transformed image. The number of radial directions and steps per direction are set to 896 each, resulting in an image size of 896×896 pixels. These values were selected based on simulations that demonstrated that this configuration decreases transformation errors to zero while keeping a reasonable image size. Fig. 2 illustrates an example of an original radio map and its corresponding transformed radio map. The transformed images, both inputs and targets, are used for neural network training. After prediction, the results are converted back into the original radial format.

III. NETWORK ARCHITECTURE AND TRAINING PROCEDURE

Our goal was to develop a predictive model capable of accurately generating a radio map from a given building map with the assumption that the transmitter is always positioned at the center of the area. To achieve this, we adopt a supervised learning approach based on the UNet network architecture [11].

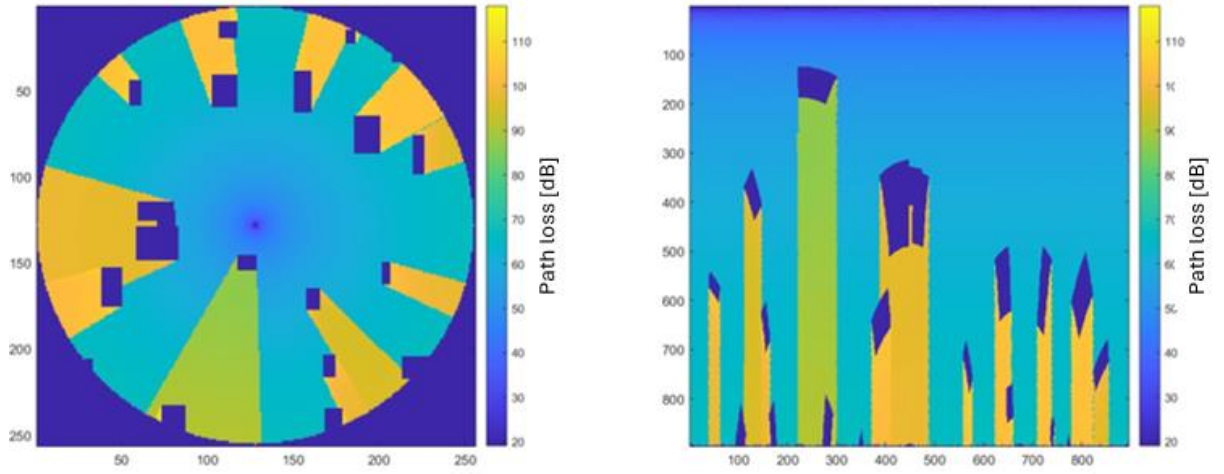
Our model is a 7-level deep UNet designed for image-to-image regression tasks. The architecture follows the typical UNet structure, consisting of a contracting path (encoder) and a symmetric expanding path (decoder), connected through skip connections. These skip connections enable high-resolution features from the encoder to be combined with the upsampled features in the decoder, thus preserving fine spatial details.

Each of the seven convolutional blocks that form the encoder includes two convolutional layers that are followed by ReLU activations. The receptive field is expanded in the first two blocks by applying dilated convolutions with a dilation rate of 2. The number of feature channels doubles with each block in the encoder (from 16 to 1024). After each convolution block there is a max-pooling layer, which is used to gradually decrease the spatial resolution. A 2048-channel bottleneck layer connects the encoder and decoder.



(a) Building map (b) Radio map

Fig. 1 Building map and its corresponding radio map with the transmitter located at the center of the area



(a) Radio map (b) Transformed radio map

Fig. 2 Original radio map and its corresponding transformed radio map

The decoder reconstructs spatial resolution through transposed convolutions while also incorporating the corresponding encoder features via skip connections. Each upsampling step is followed by a convolutional block. Finally, a 1×1 convolution layer produces a single-channel grayscale output image of the same spatial dimensions as the input. Using this architecture, our model performs image-to-image translation, effectively transforming a building map into the corresponding radio map.

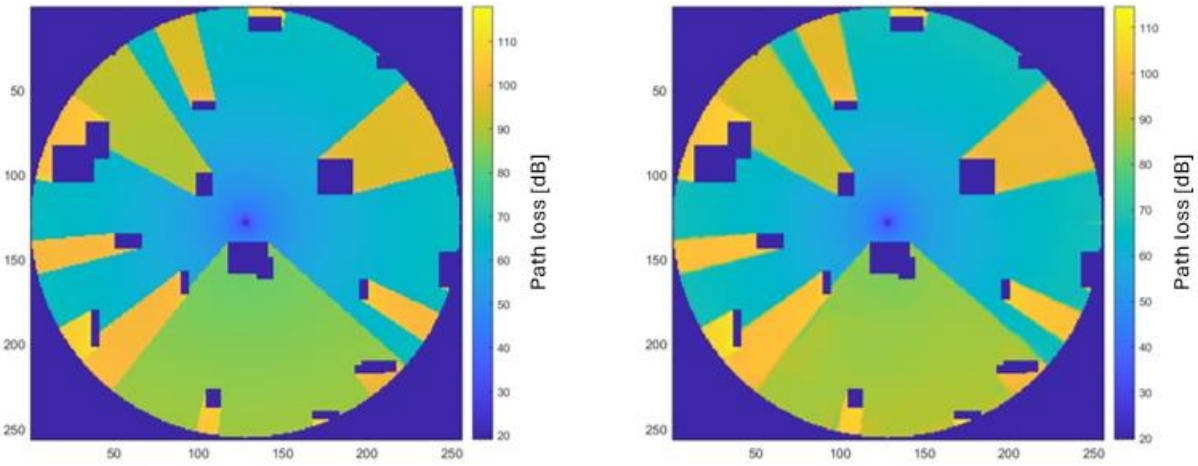
The model was implemented using the PyTorch framework, leveraging resources of the National AI Platform. The neural network was trained on a dataset consisting of 2200 input-target image pairs. Before training, each image from the dataset is preprocessed as previously described. Training was conducted over 83 epochs with a learning rate of 10^{-4} using the Adam optimizer, with a batch size of 5. The training process stopped using early stopping with a patience parameter of 6 epochs.

IV. Results and Discussion

An additional dataset of 640 images was used to evaluate the accuracy of the developed model. The test dataset was preprocessed using radial-to-rectangular transformation and grayscale conversion, just like the training dataset. The statistical metrics used to assess the accuracy of the created model were root-mean-squared-error (RMSE), standard deviation (STD), and mean. The values of these parameters were calculated by comparing the target radio maps with the radio maps that are result of model prediction. Only pixels within the radial range, which do not belong to buildings, were considered for statistical evaluation. The values of statistical metrics presented in Table 2 were obtained by averaging the statistical parameters for each of the images obtained by prediction based on the test dataset. Fig. 3 shows one example of a target radio map and a corresponding predicted radio map. Prediction on the whole test dataset took approximately 279.9 seconds using a computer equipped with an NVIDIA Quadro M1200 graphics card.

TABLE II. STATISTICAL METRICS VALUES

Statistical metrics	**Values [dB]**
RMSE	2.64
STD	2.63
MEAN	0.23



(a) Original radio map (b) Predicted radio map

Fig. 3 Original radio map and its corresponding predicted radio map

Considering the dynamic nature of radio signal propagation, these results show a high degree of accuracy. Although this study is limited by the use of synthetic data and the lack of multipath modeling beyond free-space and diffraction, the suggested model can successfully capture the complex propagation effects and obtain comparatively low error metrics. The achieved performance is comparable to the results reported by state-of-the-art methodologies cited in the literature. However, a detailed direct comparison would require identical datasets, which are not available in the present study. Considering the limited GPU hardware, computational efficiency indicates that the method can be scaled and modified for more complex and larger datasets. These results suggest that the model has lots of promise for real-world use in the planning and optimization of radio networks, while the identified limitations will be addressed in future work.

V. Conclusion

This work introduces new methodology for path loss prediction, based on image transformation and the application of convolutional neural networks. The obtained results indicate that the proposed approach is both accurate and computationally efficient, even when executed on less powerful hardware. It enables path loss prediction for a microcell area in a single inference step in less than half a second, with an average RMSE of 2.64 dB.

The presented model represents the first step in the author's research on how the application of the new methodology can contribute to the development of accurate and efficient propagation models with many opportunities for further improvement. Improving the model does not only involve optimizing the architecture and parameters used in the training process but also considering the use of additional databases. Given the fact that the accuracy of the created model is determined by the accuracy of the target data, the goal of further work, among other things, is to create and evaluate the model using the results of field measurements or ray-tracing simulations that would be used as target values. The achieved results have demonstrated the model's capacity to efficiently learn the connection between path loss and the environment.

Acknowledgment

The authors gratefully acknowledge the Palace of Science for providing the facilities and resources that made this research possible.

References

[1] N. Nešković, A. Nešković, “Microcell electric field strength prediction model based upon artificial neural networks,” *AEU-International Journal of Electronics and Communications*, vol. 64, no. 8, pp. 733-738, Aug. 2010, doi: 10.1016/j.aeue.2009.05.005

[2] H.-S. Jo, C. Park, E. Lee, H. K. Choi, and J. Park, “Path loss prediction based on machine learning techniques: Principal component analysis, artificial neural network, and gaussian process,” *Sensors*, vol. 20, no. 7, p. 1927, Mar. 2020, doi: 10.3390/s20071927

[3] S. Sotiroudis, K. Siakavara, J.N. Sahalos, “A neural network approach to the prediction of the propagation path-loss for mobile communications systems in urban environments, ” *Piers Online*. 3. pp. 1175-1179, 2007, doi: 10.2529/PIERS070220023434

[4] L. Wu, D. He, B. Ai, J. Wang, H. Qi, K. Guan, Z. Zhong, "Artificial neural network based path loss prediction for wireless communication network," in *IEEE Access*, vol. 8, pp. 199523-199538, 2020., doi: 10.1109/ACCESS.2020.3035209.

[5] J. Thrane, et al., "Deep learning-based signal *strength prediction using geographical images and expert knowledge," GLOBECOM 2020-2020 IEEE Global Communications Conference*, 2020, pp. 1-6, doi: 10.1109/GLOBECOM42002.2020.9322089

[6] Sotirios P. Sotiroudis, Sotirios K. Goudos, Katherine Siakavara, “Deep learning for radio propagation: Using image-driven regression to estimate path loss in urban areas,” *ICT Express*, vol. 6, Issue 3, 2020, pp. 160-165, doi: 10.1016/j.icte.2020.04.008

[7] M. Ribero et al., "Deep learning propagation models over irregular terrain," *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 4519-4523, doi: 0.1109/ICASSP.2019.8682491.

[8] R. Levie, Ç. Yapar, G. Kutyniok and G. Caire, "RadioUNet: Fast Radio Map Estimation With Convolutional Neural Networks," in *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 4001-4015, June 2021, doi: 10.1109/TWC.2021.3054977

[9] J. -H. Lee and A. F. Molisch, "A Scalable and Generalizable Pathloss Map Prediction," in *IEEE Transactions on Wireless Communications*, vol. 23, no. 11, pp. 17793-17806, Nov. 2024, doi: 10.1109/TWC.2024.3457431

[10] V. V. Ratnam *et al*., "FadeNet: Deep Learning-Based mm-Wave Large-Scale Channel Fading Prediction and its Applications," in *IEEE Access*, vol. 9, pp. 3278-3290, 2021, doi: 10.1109/ACCESS.2020.3048583

[11] O. Ronneberger, P. Fischer, and T. Brox, “UNet: Convolutional networks for biomedical image segmentation,” *Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015* , doi: 10.48550/arXiv.1505.0459

Exploring Students' Attitudes Toward Learning Business Communication With AI Assistance

Tamara Tolnauer-Ackermann, Lucija Bačić and Marija Krstinić
Zagreb University of Applied Sciences, Zagreb, Croatia
tamara.tolnauer.ackermann@tvz.hr, lucija.bacic@tvz.hr, marija.krstinic@tvz.hr

***Abstract* —The growing integration of artificial intelligence (AI) tools in higher education presents new opportunities for enhancing students' professional language competencies. This study investigates students' perceptions of AI tools in the context of learning business communication in English. Based on survey responses from 148 students enrolled at Zagreb University of Applied Sciences, the paper examines usage patterns, confidence levels, and challenges encountered when applying AI in communication-related tasks. Results indicate that students predominantly use AI for grammar checking, vocabulary acquisition, and report writing, and that such use is associated with increased confidence in written English communication. Furthermore, findings reveal that the majority recognize the importance of business communication for their future careers, with 95% agreeing it should be part of all study programs. The majority of participants reported using AI tools with some regularity, most commonly on a weekly basis. A smaller proportion engages with these tools daily. Additionally, a considerable number of participants reported feeling more confident writing and speaking in English with the help of AI tools. However, concerns regarding the accuracy and contextual understanding of AI-generated content persist. The study highlights the importance of integrating AI tools into business communication curricula in a pedagogically sound and ethically guided manner.**

Key words: artificial intelligence, education, business communication, English for academic purposes

I. INTRODUCTION

In today's professional environment, shaped by globalization and rapid technological advancement, business communication represents a key skill in developing both professional and educational competencies. In addition to acquiring essential theoretical and practical expertise, students increasingly need to develop and refine communication skills necessary for the effective exchange of information in business contexts, as well as for successful collaboration in international and digitalized workplaces.

At the same time, with the unstoppable advancement of technology, artificial intelligence (AI) has, in recent years, found broader applications and entered educational processes more strongly. As a tool for fostering learning and the development of diverse competencies, AI enhances and offers new opportunities for personalized learning, interactive content, and automated feedback. AI-based tools, such as chatbots, language-learning platforms, and intelligent evaluation systems are significantly transforming the ways in which students acquire and improve business communication skills. This technology enables faster access to information, individualized learning, and the development of communication confidence, particularly valuable in English-language communication. However, despite the growing application of AI in education and its numerous advantages, the scientific literature still records relatively few studies focusing exclusively on students' perceptions and attitudes toward its role in learning business communication and developing essential skills. Students' views on the use of AI in this context remain underexplored.

This study aims to address this gap by examining how students perceive and adopt this innovative approach to learning. The paper seeks to fill this research gap by analyzing students' attitudes toward the use of artificial intelligence in the process of learning business communication. The focus is placed on perceived benefits, drawbacks, and challenges, as well as on possible implications for future positive educational practices. The findings can contribute to a deeper understanding of the role of AI in education, while also offering practical guidance for teachers and higher education institutions in improving existing models and designing new, more effective approaches to teaching.

Furthermore, this study emphasizes the need for the continuous adaptation of curricula to effectively integrate emerging AI technologies. A proactive pedagogical approach that balances technological support with critical human reasoning and oversight is essential in reducing the risks of over-reliance on AI. This includes structured education on digital literacy and the ethical use of AI to better prepare students for the real challenges of business communication.

Research Problem, Subject, and Objectives

Despite the increasing presence of artificial intelligence (AI) in education, there is still a limited

number of studies addressing students' attitudes toward the use of AI in learning business communication. While global literature highlights the potential of AI tools to enhance linguistic and communication competencies, in the Croatian educational context it remains insufficiently explored how students perceive this technology; whether they view it as a supportive resource or as an obstacle to developing their own skills.

The research problem can be defined through the following question: *To what extent do students perceive AI as a useful tool in learning business communication, and are there differences in attitudes depending on their personal experiences and characteristics?*

The subject of the research is students' attitudes toward learning business communication with the support of AI, with a particular emphasis on the perceived advantages, limitations, opportunities, and challenges in applying AI tools in the educational process.

The main objective of the study is to examine and analyze students' perceptions of the role of AI in learning business communication and to determine the extent to which this technology contributes to their motivation, self-confidence, and learning effectiveness. The results of the study are expected to provide deeper insights into the potential and limitations of AI use in education, while also serving as practical guidance for teachers and educational institutions in designing more effective teaching models.

Review of Previous Research

Business communication has held an important place in education for decades, particularly in the fields of economics, management, and technical sciences, where students are expected to develop linguistic and communication competencies adapted to a global environment. Numerous studies emphasize that communication skills are among the most sought-after competencies in the labor market, with increasing focus on business communication in English as the lingua franca of the modern business world (Indeed, 2025; Savvas, 2024). The development of digital technologies has enabled innovative approaches to teaching and learning communication skills. E-learning, online platforms, and interactive applications have been shown to increase students' motivation and provide more flexible educational environments (ScienceDirect, 2022). In recent years, artificial intelligence (AI) has attracted particular interest, introducing a new dimension to the educational process. AI is increasingly being applied for automated evaluation of written and oral tasks, language content generation, speech recognition, and the adaptation of teaching materials to individual learners' needs (U.S. Department of Education, 2023; ScienceDirect, 2023). Previous studies indicate that students perceive AI tools such as intelligent tutors and chatbots positively, as these provide instant feedback and allow them to practice different communication scenarios in a safe environment, without the fear of making mistakes (Nelson, Santamaría, Javens, & Ricaurte, 2025). On the other hand, some literature points out potential drawbacks—overreliance on technology, reduced critical thinking, the risk of superficial knowledge acquisition, as well as challenges related to ethical behavior. At the global level, there is an abundance of research addressing the role of AI in education more broadly, yet far fewer studies focus specifically on business communication. In the Croatian educational context, such studies are even more scarce, which creates additional space for exploring students' perceptions of the effectiveness, risks, and challenges of using AI in learning business communication. For this reason, this paper seeks to connect theoretical insights on the application of modern technologies in education, particularly in the acquisition of business communication knowledge and skills, with an empirical investigation of students' attitudes, thereby contributing to a broader understanding of the role of AI in modern education.

The Significance of Communication Skills and the Contribution of AI Tools to Their Acquisition

In today's fast and dynamic world, efficient communication is key to success in any business environment. For students studying at technical universities, especially those studying engineering, mastering business communication cannot be overstated. Although engineering expertise is essential in creating new, innovative solutions and advancing technological progress, the ability to transfer ideas clearly and establish successful collaboration with different stakeholders as well as participate in efficient negotiations and meetings is of equal importance. As the business world becomes more complex and more connected, learning how to communicate within a business environment is becoming one of the key skills for students aspiring to enter the global workforce.

In recent years, AI has started to play a more influential role in the learning process, including business communication education. The use of AI has been deeply integrated into modern society, and this influence is increasingly reflected in educational settings (Vieriu & Petrea, 2025). Tools, platforms, and virtual assistants based on artificial intelligence are now available to assist students in improving their business communication skills in various ways. Real-time feedback on written and spoken communication, help with improving message clarity, and business environment simulations are just a few areas in which AI can be of great help. While technical skills are often considered the most important for employability, many employers highly value 'soft skills' such as critical thinking, communication, and teamwork, sometimes even preferring them over purely technical expertise (Mwita, Mwilongo, & Mwamboma, 2024). It is especially so for engineering students whose focus is often more on acquiring technical expertise and less on developing their 'soft skills' such as communication. AI's ability to analyse vast amounts of

data and provide personalized feedback to each student enables students to gain insight into their communication styles and areas which perhaps need improvement. “The intelligent system can automatically adjust the teaching content and difficulty according to the students’ learning progress and performance to meet the learning needs of different students” (Cui & Yue, 2024). Moreover, AI technologies such as natural language processing (NLP) can assist students’ writing by detecting grammar, tone, and clarity issues while speech analysis tools can help improve verbal communication and presentation skills. Furthermore, it can automatically correct students’ homework and provide detailed feedback (Cui & Yue, 2024).

Although the integration of AI into business communication education brings numerous advantages, it also represents challenges and potential drawbacks which should be considered. Tailor-made approaches to learning are definitely beneficial for each student. Immediate feedback enables students to learn and progress faster and with greater self-confidence. Simulation of realistic business scenarios additionally helps students improve their presentation and negotiation skills. However, such an approach can lead to over-reliance on AI and the weakening of human interaction and critical and analytical thinking skills (Zhai, Wibowo, & Li, 2024). Another serious concern which should not be overlooked is the ethical implications of using AI in business communication (Nicholas, 2025). As AI systems collect vast amounts of data which include students’ personal communication styles, privacy and security issues may arise. Also, there is a problem of potential biases in AI algorithms which could lead to unfair preferences during employment procedures (Pandey, Sharma, Tiwari, & Tyagi, 2024). Depending on how the AI tools were designed and tested they could perpetuate certain stereotypes and biases in communication. In an increasingly digital and globalized professional environment, effective business communication in English has become a vital skill (Bhatia & Bremner, 2012). As AI becomes more embedded in the academic context, understanding students’ perceptions of its usefulness, reliability, and impact on their learning is critical for developing effective pedagogical strategies (Godwin-Jones, 2023).

II. EMPIRICAL RESEARCH AND DISCUSSION

Research questions

- RQ1: How do students perceive the impact of AI tools on their ability to learn and practice business communication in English?
- RQ2: For which business communication tasks in English do students most frequently use AI tools, and what challenges do they encounter?

Methodology

The study involved a total of 148 undergraduate students enrolled at Zagreb University of Applied Sciences. A structured online questionnaire was developed, containing multiple-choice and Likert-scale items related to students' use of AI tools, confidence in English communication, and experiences with business communication tasks. The survey was administered in the summer semester of 2025. Participation was voluntary and anonymous. Descriptive statistics were used to analyze patterns in responses, with a particular focus on AI-related behavior and communication skills development.

III. RESULTS

This study explores Zagreb University of Applied Sciences students’ attitudes toward acquiring business communication skills, with particular focus on the perceived role of artificial intelligence (AI) tools in enhancing their learning experience (Zhang & Zou, 2022). Based on a survey of 148 respondents, key findings reveal that the majority recognize the importance of business communication for their future careers, with 95% agreeing it should be part of all study programs. Furthermore, the majority of participants reported using AI tools with some regularity, most commonly on a weekly basis. A smaller proportion engages with these tools daily. These findings suggest a growing familiarity and integration of AI technologies into everyday academic or professional activities (Godwin-Jones, 2023). Additionally, a considerable number of participants reported feeling more confident writing and speaking in English with the help of AI tools, indicating that AI may play a supportive role in developing language confidence (Zhang & Zou, 2022). Eighty percent believe AI tools (e.g., ChatGPT) positively contribute to their learning of business communication. Supporting these results, Picek (2022) highlights that AI-driven chatbots in higher education serve as interactive platforms that simulate real-world business communication scenarios, providing immediate feedback and fostering practical skills such as responsiveness and clarity.

This experiential learning approach bridges the gap between theory and practice, enhancing student engagement and preparing them for professional communication challenges. Survey data show that 39% of students agreed or strongly agreed with the statement that they feel more confident writing in English when supported by AI tools. An additional 43% remained neutral, while 18% expressed disagreement. Furthermore, over half of the participants reported that AI helps them understand technical vocabulary, which is particularly relevant in business contexts. These findings indicate that AI use correlates with a measurable increase in communication confidence among students. As shown in **Fig. 1**. students reported varying levels of confidence when writing in English with the support of AI tools.

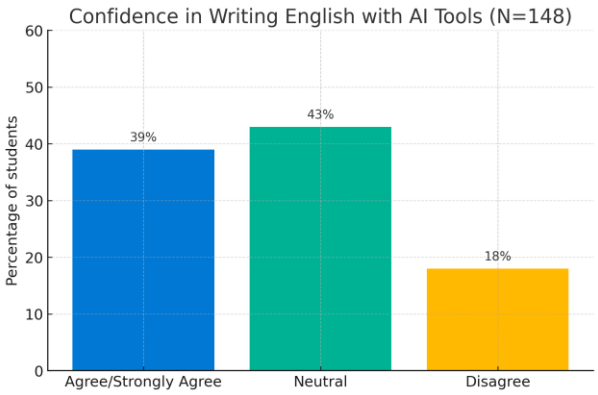


Fig 1. Display of confidence in writing English with AI tools

The most frequently reported AI-supported tasks were related to written business communication. Specifically, 74 students indicated using AI tools for grammar and spell-checking, while 37 students reported relying on AI for acquiring specialized vocabulary. Additionally, 29 students utilized AI to assist with writing reports and essays, and 30 students employed AI tools for translation tasks from Croatian to English. These findings suggest that students predominantly engage with AI in ways that support linguistic accuracy, lexical development and the structural demands of professional writing. The distribution of students' use of AI tools for specific business communication tasks is illustrated in **Fig. 2.**

Fig. 2. Display of AI Use for Business Communication Tasks in English

Despite the perceived benefits of AI tools, students also reported several notable concerns regarding their use. Among the most commonly cited issues were the generation of inaccurate or misleading outputs, difficulties in understanding contextual nuances, and the presence of occasional grammatical errors within AI-generated content. These limitations highlight the importance of fostering critical digital literacy among students and suggest the need for complementary instructional support to guide effective and responsible use of AI in communication-related tasks.

IV. DISCUSSION

Perceived Benefits of AI in Business Communication Learning

The findings confirm that students view AI tools as valuable aids in acquiring business communication skills in English, particularly in written contexts. Respondents reported frequent use of AI for grammar correction, vocabulary acquisition, and writing tasks, indicating a strong integration of these technologies into their academic routines. Notably, a significant number of students associated AI use with increased confidence in written expression, suggesting that these tools offer both linguistic and psychological support. This supports previous claims that AI can serve as a scaffold in language learning environments, helping students manage complex language tasks more effectively (Godwin-Jones, 2023).

In addition, the growing number of hybrid and remote positions in the global business environment requires new communication competencies, particularly skills in using digital communication platforms and managing intercultural interactions, competencies that artificial intelligence tools can help simulate and strengthen. The inclusion of AI-driven scenario-based learning modules could further enhance students' preparedness for contemporary business challenges.

Limitations and Challenges

Despite the reported benefits, students also expressed concerns about the accuracy, reliability, and contextual appropriateness of AI-generated content. Instances of incorrect outputs, misunderstandings of nuance, and even grammatical errors highlight the limitations of current AI systems. These concerns point to the importance of fostering critical digital literacy among students to ensure they can evaluate and refine AI-assisted outputs effectively. Curricula should integrate comprehensive AI literacy programs that address not only operational knowledge of AI tools but also broader socio-ethical implications. Such a dual approach is essential for preparing students to use AI responsibly while maintaining the necessary human judgment. Over-reliance on AI, especially without adequate instruction on its limitations, may hinder the development of independent communicative competence, which is essential in professional contexts.

To mitigate the risk of over-reliance, educational institutions should consider blended learning models that combine AI tools with traditional pedagogical methods, emphasizing critical thinking, problem-solving, and interpersonal communication. Encouraging metacognitive strategies for reflecting on the reliability of AI-generated outcomes can further contribute to the development of authentic and independent learners.

Expanding the Role of AI Beyond Writing

While most students reported using AI primarily for writing-related tasks, the study found limited engagement with AI tools for developing spoken business communication skills. This is a notable gap, considering that business communication often involves verbal interaction, such as meetings, negotiations, and presentations, that require fluency, persuasion, and cultural awareness. AI-driven applications like speech recognition software, pronunciation coaches, or conversation simulators represent promising but underused resources in this area. Integrating these tools into curricula could enhance students' oral communication proficiency and better prepare them for real-world professional interactions. The integration of AI tools for oral communication with interactive, immersive scenarios that simulate business meetings, negotiations, and presentations can provide engaging environments in which students practice and acquire essential knowledge and skills. These tools can be designed to adapt to individual learner needs, offering personalized feedback on fluency, pronunciation, pragmatics and interpersonal dynamics.

V. CONCLUSION

This study demonstrates that AI tools play a meaningful and increasingly prominent role in supporting students' development of business communication skills in English. The majority of students reported using AI primarily for written tasks, such as grammar correction, vocabulary development, and academic or professional writing, activities that align closely with the structural demands of business correspondence. Moreover, students associated AI usage with increased confidence, particularly in written expression. However, concerns about accuracy, context sensitivity, and occasional grammatical errors suggest that AI tools are not yet fully trusted for autonomous language generation and still require human oversight and critical evaluation. To fully harness the potential of AI in business communication education, there is a need to expand its application to spoken communication as well. Currently underutilized, AI-powered tools such as voice assistants, speech-to-text applications, real-time pronunciation coaches, and AI-driven conversation simulators offer promising avenues for developing oral proficiency and communicative confidence in professional settings. These tools can replicate business scenarios such as meetings, negotiations, and presentations, contexts that students often find particularly challenging.

Therefore, higher education institutions should not only integrate AI tools into writing-focused curricula but also explore and adopt interactive, speech-based AI applications to support holistic communication competence. Therefore, a balanced educational approach is essential; one that combines AI-assisted learning with explicit instruction in digital literacy and ethical awareness. Future research should investigate the effectiveness of such tools in improving fluency, pragmatic accuracy and interpersonal communication strategies in business contexts. Ultimately, the question of how students perceive and accept AI as a useful tool and innovative approach for developing their communication competences and whether challenges and obstacles exist in its application, will continue to grow in importance. It is likely to remain a subject of future scientific inquiry both in the global and Croatian educational context.

REFERENCES

[1] V. K. Bhatia and S. Bremner, "English for business communication," *Language Teaching*, vol. 45, no. 4, pp. 410–445, Oct. 2012. [Online]. Available: https://doi.org/10.1017/S0261444812000171

[2] Y. Cui and S. Yue, "Business English teaching with intelligent assistance—Research on case-driven and artificial intelligence application improvement," Journal of Educational Research and Policies, vol. 6, no. 10, pp. 152–160, 2024. [Online]. Available: https://doi.org/10.53469/jerp.2024.06(10).18

[3] R. Godwin-Jones, "AI in language learning: Friend, assistant, or competitor?," Language Learning & Technology, vol. 27, no. 1, pp. 1–7, 2023. Available: https://www.lltjournal.org/item-detail/1053/

[4] K. Hyland, Teaching and Researching Writing, 3rd ed. London, U.K.: Routledge, 2016. [Online]. Available: https://doi.org/10.4324/9781315717203

[5] Indeed, "The importance of business communication: 6 reasons why," Indeed Career Guide, July 25, 2025. [Online]. Available: https://www.indeed.com/career-advice/career-development/importance-of-business-communication

[6] K. Mwita, N. Mwilongo, and I. Mwamboma, "The role of soft skills, technical skills and academic performance on graduate employability," International Journal of Research in Business and Social Science, vol. 13, no. 5, pp. 767–776, 2024. [Online]. Available: https://doi.org/10.20525/ijrbs.v13i5.3457

[7] A. S. Nelson, P. V. Santamaría, J. S. Javens, and M. Ricaurte, "Students' perceptions of generative artificial intelligence (GenAI) use in academic writing in English as a foreign language," Education Sciences, vol. 15, no. 5, art. 611, 2025. [Online]. Available: https://doi.org/10.3390/educsci15050611

[8] A. J. Nicholas, "AI and business education: Ethical challenges and practical applications," in Proc. World Conf. Education and Teaching, vol. 4, no. 1, pp. 89–99, 2025. [Online]. Available: https://doi.org/10.33422/etconf.v4i1.1082

[9] A. S. Pandey, Y. Sharma, A. Tiwari, and S. Tyagi, "Ethical implications of AI-powered communication tools," in Proc. IC3SE Conf., pp. 1–6, 2024. IEEE. [Online]. Available: https://doi.org/10.1109/IC3SE62002.2024.10593350

[10] R. Picek, "Application of chatbots in higher education environment," in Proc. Int. Conf. Information and Communication Technology. Zagreb, Croatia: Univ. of Zagreb, 2022. [Online]. Available: https://www.croris.hr/crosbi/publikacija/prilog-skup/696764

[11] Savvas, "The role of communication skills in business success," Savvas Learning Company Blog, 2024. [Online]. Available: https://www.savvas.com/resource-center/blogs-and-podcasts/college-and-career-readiness/career-paths/role-of-communication-skills-in-business-success

Statistical analysis of single-frequency GNSS positioning error under a suspected spoofing attack

Renato Filjar
Hrvatsko Zagorje Krapina University of Applied Sciences
Krapina, Croatia, and
Faculty of Engineering
University of Rijeka
Rijeka, Croatia
ORCID: https://orcid.org/0000-0002-7040-9931

Lana Miličević
Independent researcher
Karlovac, Croatia
ORCID: https://orcid.org/0009-0000-4537-9286

Jurica Kaponja
Laboratory for Spatial Intelligence
Hrvatsko Zagorje Krapina University of Applied Sciences
Krapina, Croatia
email: jurica.kaponja@vhzk.hr

Darko Špoljar
PhD student
Faculty of Engineering
University of Rijeka
Rijeka, Croatia
ORCID: https://orcid.org/0000-0002-7440-7087

***Abstract*—This electronic document is a "live" template and already defines the components of your paper [title, text, heads, etc.] in its style sheet. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.** (*Abstract*)

Keywords—component, formatting, style, styling, insert (*key words****)***

I. INTRODUCTION

The Global Navigation Satellite System (GNSS) has become one of the cornerstones of modern civilisation, a public goods, and the enabler of a growing number of technology and socio-economic applications, systems and services [1, 2]. Its inherited shortcomings and vulnerabilities transform into GNSS-based applications risks, thus undermining the quality and sustainability of GNSS applications [1, 3, 4]. The GNSS shortcomings and vulnerabilities had been initially caused by natural sources of the Positioning, Navigation, and Timing (PNT) performance degradation and disruptions [1, 3, 4, 6]. Maintaining the GNSS PNT performance problem has recently been complicated further with adversarial artificial cyberattacks on GNSS PNT. The intentional forceful provision of false GNSS signals and navigation messages/data aimed at convincing a GNSS receiver to estimate predefined wrong position, a process known as GNSS spoofing, has become the most dangerous form of a GNSS cyberattack [7-9]. GNSS spoofing attack may cause a severe GNSS PNT performance degradation, and, consequently, a considerable damage and loss of lives [1, 3, 10]. The GNSS resilience development address mitigation of both the natural and adversarial sources of the PNT performance degradations and disruptions [10-14].

Here we contribute to the understanding of the nature of GNSS spoofing with a statistical analysis of individual GNSS positioning performance in the case of suspected GNSS spoofing operation in the Eastern Mediterranean. Based on the event reconstruction derived from experimental raw GNSS pseudoranges collected at the International GNSS Service (IGS) reference station in Nicosia, Cyprus, the research aims to identify evidence of the GNSS spoofing conducted by provision of the false navigation messages, and to assess the level of exposure to suspected GNSS spoofing attack of different GNSS systems: the US-operated GPS, the China-operated Beidou (BDS), and the EU-operated Galileo.

This Section introduces a reader to the problem, and outlines the motivation, aim, objectives and hypothesis of the research. Section II. presents comprehensively the methodology of research, and the data/material used. Section III. presents the research results obtained. Section IV. discusses research results in terms of the aim, objectives, and hypothesis of the research, identifies new knowledge, and concludes the manuscript with inference, list of contributions, and suggestions for the future research.

II. METHOD AND MATERIAL

Section II. outlines the methodology and presents a description of material used in the presented research, in relation of the aims and objectives set.

A. Method

The presented research aims at: (i) identification of the suspected GNSS spoofing attack in terms of the forceful provision of a false navigation message to a single-frequency commercial-grade GNSS receiver, and (ii) estimation of the level of exposure to/consequences of the GNSS spoofing attack on various GNSS positioning performance error. The Nicosia, Cyprus GNSS reference station site is selected for a number of records of the temporal exposure to GNSS spoofing attacks. Two cases used in research are selected, one with the evidence of significantly deteriorated GNSS PNT performance, and the other one with usual GNSS PNT performance serving as a control/benchmark in the study. The case selection is conducted based on the reports of GPS performance degradation derived from the ADS-B navigation message analysis conducted by a trusted third-party. The presented research derives GNSS positioning estimates from the provided raw GNSS pseudorange observations, as described in Section II. B. Cases are selected in periods of quiet, undisturbed ionospheric conditions, and the GNSS ionospheric delay correction using broadcast Klobuchar model parameters are conducted to remove the ionospheric effects from the positioning error estimates. Two sets of time series are derived for every case of every GNSS system deployed, using (i) broadcast navigation message, and (ii) received navigation message. The procedure emphasises

potential difference in errors caused by suspected false navigation message-based GNSS spoofing attack. Using precisely known ground-truth position of the IGS reference station, the procedure yields time series of positioning error components, per GNSS system (GPS, GLONASS, BDS, and Galileo). Exploratory statistical analysis is performed on single GNSS data, following the performance comparison between the systems. The methodology flow-chart is depicted in Figure 1.

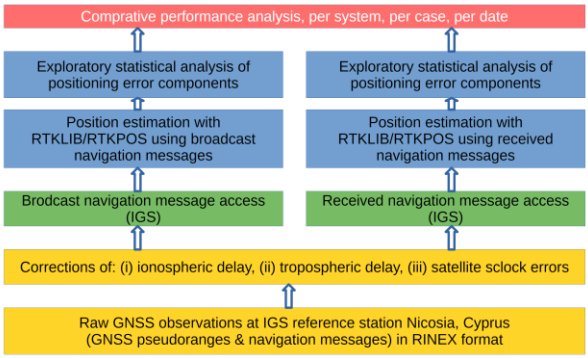


Fig. 1. Flow chart of the research methodology.

GNSS position estimation based on GNSS pseudorange observations, and broadcast and received navigation messages are performed using the RTKLIB/RTKPOST Software-Defined Radio (SDR) GNSS receiver [15], an open-source research tool commonly used in studies of GNSS PNT performance [4, 5]. The RTKLIB/RTKPOST is utilised in the presented research configured in the post-processing mode as a single-frequency commercial-grade GNSS receiver, as depicted in Figure 2.

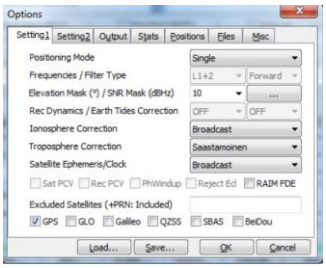


Fig. 2. RTKPOST configuration.

Corrections of ionospheric and tropospheric delays, as well as that of satellite atomic clocks are used as broadcast by related satellites. The RTKLIB/RTKPOST GNSS SDR receiver is being configured to utilise a single GNSS system (either GPS, GLONASS, BDS, or Galileo), as required by the specific case of the research.

Exploratory statistical analysis is performed on position estimates obtained from the RTKLIB/RTKPOST GNSS SDR receiver using a software develop by authors in the open-source R environment for statistical computing [16].

B. Material

The IGS reference station in Nicosia, Cyprus is identified as a source of observations for its records of operation in the area frequently exposed to GNSS spoofing attacks. The IGS reference station collects the raw 30 s- sampling rate GNSS pseudorange and navigation message observations in continuous manner, 24 hours a day, through the year [17]. It utilises a GNSS receiver capable of collection all four GNSS systems' pseudoranges in the site vicinity. The observation records contain measurements and data as received, thus remaining contaminated with ionospheric and GNSS spoofing effects.

Position of the IGS reference station in Nicosia, Cyprus is depicted in Figure 3.

Fig. 3. Position of the IGS reference station Nicosia, Cyprus, maps created in the R environment for statistical computing with the software utilising the *leaflet* library and OpenStreetMap data

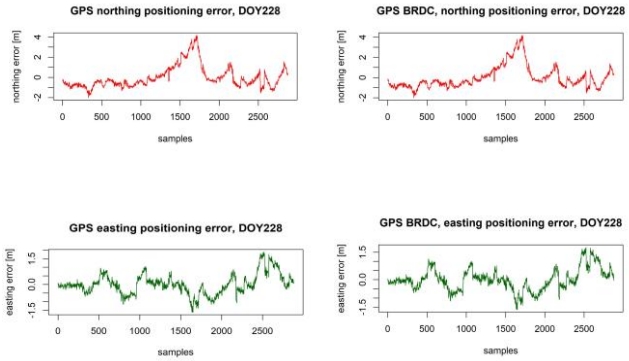
Selection of research cases: (i) DOY288 in 2025 (16th August), as a day with a significant GNSS PNT degradation potentially due spoofing, and (ii) DOY240 in 2025 (28th August) as a day with the usual GNSS PNT performance, is performed based on the third-party GPSjam.org historic records of anomalous GPS PNT performance [18], as derived from the aviation ADS-B reports. Methodology of estimation is described by the GPSjam.org [18].

Observations are presented and accessed in the computer-readable formats, and analysed using software developed by authors for the purpose of the presented research in the R environment for statistical computing [16].

III. Research Results

This Section contains the results of exploratory statistical analysis of derived single-frequency commercial-grade GNSS positioning error components observations at the Nicosia, Cyprus IGS reference station on DOY228 (suspected GNSS spoofing attack) and DOY240 (control, no considerable GPS PNT degradation identified from ADS-B messages), both in 2025. Statistical analysis comprises statistical properties of datasets, as well as the correlation analysis, of positioning error estimates resulting from utilisation of the received navigation messages, as compared with the GPS PNT performance obtained using the broadcast GPS message on the DOY240 of 2025.

Time series of the GPS positioning error on DOY228 in 2025, as an example of initial exploratory statistical analysis performed on all observations are shown in Figure 4.



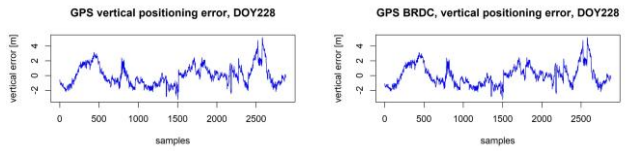


Fig. 4. Time series of GPS positioning errors during the suspected spoofing case.

The GPS positioning errors are derived from GPS pseudorange measurements based on broadcast and received navigation message, respectively. An initial assessment of diagrams does not show a visible difference. However, a formal statistical analysis is to be deployed to justify the conclusion objectively.

The range of observed positioning error per components, as well as their variance are depicted with box-and-whisker plots (boxplots) in Figure 5.

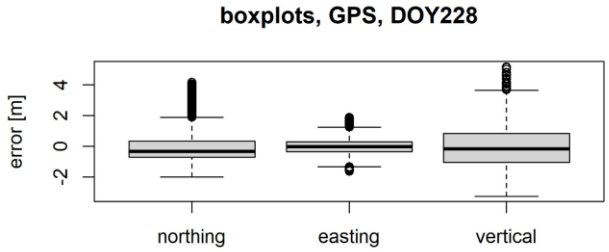


Fig. 5. Box-and-whisker plots of GPS positioning error components, derived from GPS positioning estimates based on received navigation message

Boxplots reveal ordinary common situation of a daily GNSS PNT performance. Similar analysis is performed on all GNSS positioning error sets, per identified cases.

Boxplots are examined for potential mutual association, using the comparative boxplot analysis, supported in the R environment for statistical computing, as shown in Figures 6, 7, and 8, for northing, easting, and vertical errors, respectively.

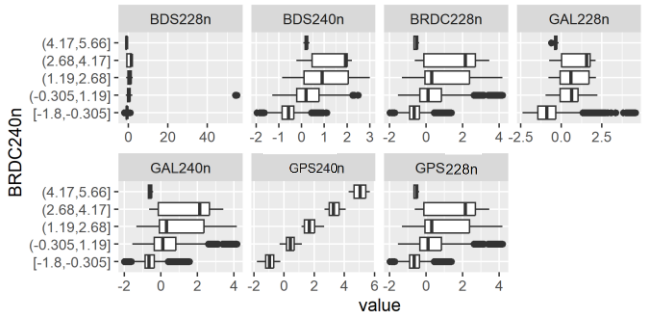


Fig. 6. The comparative boxplot analysis, with broadcast message northing GPS error as the benchmark for comparison (BRDC … GPS positioning error component obtained with the GPS broadcast navigation messages, GPS, BDS, and GAL ... positioning error component obtained with the related received navigation messages of GPS, BDS, and Galileo, respectively)

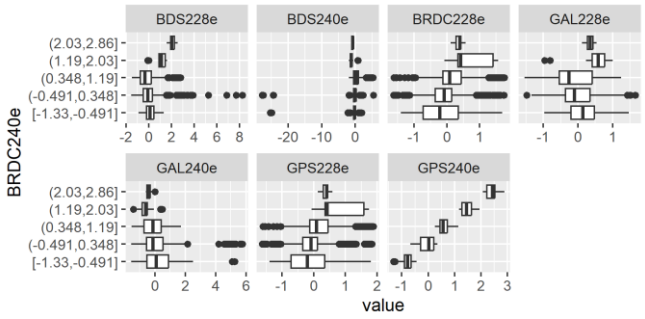


Fig. 7. The comparative boxplot comparative analysis, with broadcast message easting GPS error as the benchmark for comparison(BRDC … GPS positioning error component obtained with the GPS broadcast navigation messages, GPS, BDS, and GAL ... positioning error component obtained with the related received navigation messages of GPS, BDS, and Galileo, respectively)

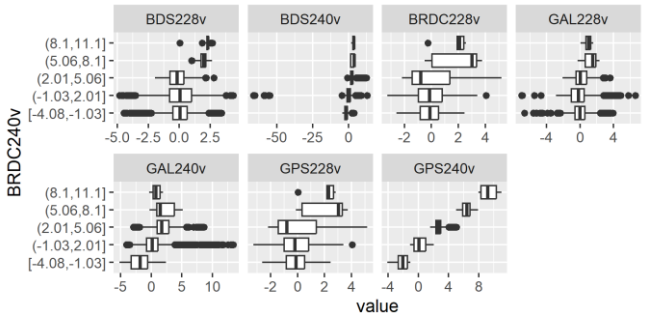


Fig. 8. The comparative boxplot comparative analysis, with broadcast message vertical GPS error as the benchmark for comparison (BRDC … GPS positioning error component obtained with the GPS broadcast navigation messages, GPS, BDS, and GAL ... positioning error component obtained with the related received navigation messages of GPS, BDS, and Galileo, respectively)

The analysis points to a high association with targeted GPS positioning error component determined using broadcast and received GPS navigation message, respectively. The other component pairs extend far weaker association, as it may be expected.

Potential association between pairs of GNSS positioning error components of different cases are assessed independently using the correlation analysis, with the results for northing, easting, and vertical GNSS positioning errors depicted in Figures 9, 10, and 11, respectively.

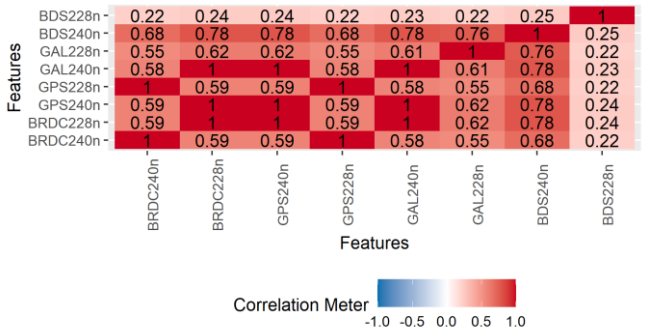


Fig. 9. Correlation analysis of northing positioning errors (BRDC … GPS positioning error component obtained with the GPS broadcast navigation messages, GPS, BDS, and GAL ... positioning error component obtained with the related received navigation messages of GPS, BDS, and Galileo, respectively)

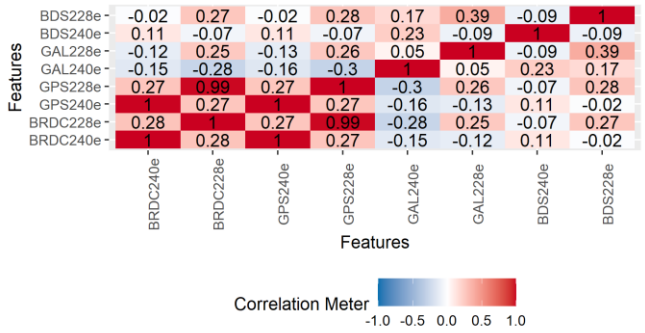


Fig. 10. Correlation analysis of easting positioning errors (BRDC … GPS positioning error component obtained with the GPS broadcast navigation messages, GPS, BDS, and GAL ... positioning error component obtained with the related received navigation messages of GPS, BDS, and Galileo, respectively)

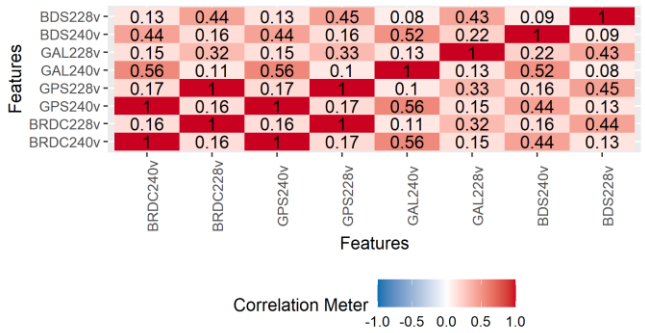


Fig. 11. Correlation analysis of easting positioning errors (BRDC … GPS positioning error component obtained with the GPS broadcast navigation messages, GPS, BDS, and GAL ... positioning error component obtained with the related received navigation messages of GPS, BDS, and Galileo, respectively)

Statistical tests are performed that proved the statistical similarity in bias (t-test), and statistically significant difference in variance in all pairs of components.

IV. Discussion and Conclusion

Presented research aims at detection of false navigation message-class of GNSS spoofing attack in selected cases of the Eastern Mediterranean GNSS positioning performance degradation. Data are statistically analysed in the case of suspected GNSS spoofing, and in the control case without evidence of the GNSS performance degradation. Positioning performance of the single-frequency commercial-grade GPS, Beidou, and Galileo PNT are analysed, and mutually compared, based on broadcast and received navigation messages. The GLONASS PNT performance is not examined due to insufficient number of observations.

Exploratory analysis of individual GNSS positioning error components, for both types of navigation messages used, returned occasional outcomes that showed no proven association with the suspected times of spoofing in effect. The GPS positioning performance was in both cases stable, but with notable variance, especially in the case of suspected spoofing. The Beidou positioning performance was generally accurate, but with a number of distinctive outliers, in the case of spoofing. The Galileo positioning performance extends a large variance, and a number of distinctive outliers, especially in the case of suspected spoofing.

Advanced statistical analysis shows that positioning components follows different statistical distributions, with the t-tests showing statistical similarity in bias, and F-tests differences in variance of all pairs of GNSS positioning error components examined.

As the result, the research finds no evidence of the navigation message-related spoofing. Since the natural sources (ionosphere, troposphere, multipath) of potential interference and impact were successfully mitigated, the research presents the evidence adversarial GNSS positioning performance degradation, of different levels of exposure for different GNSS systems.

The research contributes to the subject addressed with: (1) methodology for false navigation message-based GNSS spoofing detection, (2) case scenarios selection and assessment, and (3) a tailored software research framework developed in the R environment for statistical computing.

The research will continue with development of new and advanced methods for GNSS spoofing detection, identification, and mitigation in GNSS pseudorange/observations domain, based on utilisation of the applied statistics, statistical learning, and information theory.

References

[1] R. Filjar, M. C. Damas, T. B. Iliev, “Resilient Satellite Navigation Empowers Modern Science, Economy, and Society,” in Proc CIEES 2020, Borovets, Bulgaria, IOP Conf. Ser: Mater Sci Eng 1032, 012001 (10 pages), 2020. [Online] doi:10.1088/1757-899X/1032/1/012001

[2] EUSPA, Report on Consumer Solutions User Needs and Requirements. Prague, Czechia: EUSPA, 2024. [Online] Available at: https://www.gsc-europa.eu/sites/default/files/sites/all/files/report_on_consumer_solutions_user_needs_and_requirements.pdf

[3] London Economics, The economic impact on the UK of a disruption to GNSS. London, UK: London Economics, 2023. [Online] Available at: https://tinyurl.com/5ybfej3v

[4] R. Filjar, “An application-centred resilient GNSS position estimation algorithm based on positioning environment conditions awareness,” Proc ION ITM 2022. Long Beach, CA: ION, 2022, pp. 1123 – 1136. [Online] doi: https://doi.org/10.33012/2022.18247

[5] M. Alvir, D. Koprivanac, E. Ciriković, E. Malić, I. Heđi, R. Filjar, “A case study of smartphone-grade GNSS positioning performance analysis during a massive geomagnetic storm,” Proc TELFOR 2024. Belgrade, Serbia: TELFOR, 2024, pp. 1-4. doi: https://doi.org/10.1109/TELFOR63250.2024.10819088

[6] J. Sanz Subirana, J. M. Juan Zornoza, and M. Hernandez-Pajares, GNSS Data Processing – Volume I: Fundamentals and Algorithms. Nordwijk, The Netherlands: ESA, 2013. ISBN 978-92-9221-886-7. [Online] Available at: http://bit.ly/1tDzJIQ

[7] N. O. Tippenhauer, C. Poepper, K. B. Rasmussen, S. Čapkun, “On the Requirements for Successful GPS Spoofing Attacks,” Proc of the 18th ACM conference on Computer and communications security. Chicago, IL, 2011, pp. 75-86. doi:10.1145/2046707.2046719

[8] M. Filić, “Foundations of GNSS Spoofing Detection and Mitigation with Distributed GNSS SDR Receiver,” TransNav, 12(4), 2018, pp. 649 – 656, 2018. doi:10.12716/1001.12.04.01

[9] L. Meng, L. Yang, W. Jang, L. Zhang, “A Survey of GNSS Spoofing and Anti-Spoofing Technology,” Remote Sensing, 2022, 14, 4826. [Online] doi: https://doi.org/10.3390/rs14194826

[10] A. J. Kerns, D. P. Shepard, J. A. Bhatti, T. E. Humphreys, “Unmanned Aircraft Capture and Control via GPS Spoofing,” J of Field Robotics, 31(4), 2014, pp. 617-636. [Online] doi: https://doi.org/10.1002/rob.21513

[11] A. Jafarnija-Jahromi, et al, “GPS Vulnerability to Spoofing Threats and a Review of Antispoofing Techniques,” Int J of Nav and Obs, Article ID 127072, 16 pages, 2012. doi:10.1155/2012/127072

[12] D. P. Kingma, M. Welling, “GNSS Spoofing Detection Based on Opportunistic Position Information,” 2025, arXiv:2506.12580v1 [cs.CR] 14 Jun 2025. [Online]. Available: https://arxiv.org/html/2506.12580v1

[13] N. Spens, D.-K. Lee, F. Nedelkov, D. Akos, “Detecting GNSS Jamming and Spoofing on Android Devices,” NAVIGATION: Journal of the Institute of Navigation, 69(3), 2022, navi.537; DOI: https://doi.org/10.33012/navi.537

[14] R. Filjar, S. Kos, D. Brčić, “Jamming-Spoofing-Meaconing-Resilient GNSS Performance at the Open Sea,” Proc of 8th Annual Baška GNSS Conference. Baška, Krk Island, Croatia: Faculty of Maritime Studies, 2014, pp. 25 - 33. Available at:

https://www.pfri.uniri.hr/web/hr/dokumenti/zbornici-gnss/2014-GNSS-8.pdf

[15] rtklibexplorer, “RTKLIB: An Open Source Program Package for GNSS Positioning, ver. RTKLIB-EX 2.5.0,” 2025, gitHub repository. [Online]. Available at: https://github.com/rtklibexplorer/RTKLIB/releases

[16] R Core Team, “R: A Language and Environment for StatisticalComputing,” open access repository, 2025. Vienna, Austria; R Foundation for Statistical Computing. [Online]. Available at: https://www.r-project.org/

[17] NASA, ”NASA EarthData CDDIS NASA’s Archive of Space Geodesy Data,” open access upon the free registration,” 2024, open-access repository [Online]. Available at: https://cddis.nasa.gov/archive/gnss/data/daily/

[18] J. Wiseman, ”GPSjam web-site and historic records of spatial estimation of GPS interference based on ADS-B messages,” 2025, open-access repository [Online]. Available at: https://gpsjam.org/?lat=47.77619&lon=20.81488&z=3.0

Artificial intelligence in tourism – optimization of work costs and competitiveness

Zdenko Bolfek
University of Applied Sciences Hrvatsko Zagorje Krapina
Krapina, Croatia
zdenko.bolfek@gmail.com

Morana Bolfek
Independent researcher
Zagreb, Croatia
morana.bolfek@gmail.com

***Abstract* – Artificial intelligence (AI) is increasingly influencing the tourism industry by supporting personalized services, automating routine operations, and improving the quality of business decisions. This paper examines how AI technologies are reshaping tourism, taking into consideration its application in digital customer support, predictive analytics, dynamic pricing, robotic assistants, and individualized offers for guests. The analysis highlights the economic effects of AI adoption, including gains in operational efficiency, cost savings, and changes in the labour market, as well as challenges connected to data protection, maintaining human interaction, and ensuring technological access for smaller enterprises.**

A distinctive feature of tourism businesses compared to many other sectors lies in their relatively high share of employee-related costs, which directly affects profitability. Due to the specific nature of tourism activities, technological progress in the past had only a limited impact on cost reduction and business performance. Recent developments in artificial intelligence, however, are creating new opportunities to overcome these barriers. The paper therefore discusses potential economic implications for tourism enterprises that may arise from the continued advancement of AI-based solutions.

Keywords – artificial intelligence, economic impacts, employee costs, tourism

I. INTRODUCTION

Tourism is among the most dynamic and influential economic sectors worldwide and plays a particularly important role in the economies of Mediterranean countries such as Croatia. In Croatia, tourism accounts for almost 20% of the gross domestic product (GDP), making it a major engine of growth, employment, and service exports. At the same time, the global tourism sector faces a number of persistent challenges — shifts in consumer preferences, growing guest expectations, labour shortages, and the constant pressure to achieve greater efficiency in daily operations.

For most businesses, especially those that are labour-intensive, employee-related expenses represent one of the largest cost components in theirs profit and loss accounts. These costs are largely inevitable and have a direct influence on profitability. However, cost management should not be confused with indiscriminate cost reduction. Business experience shows that investing in employees — through fair pay, professional development, and various motivation programs — can lead to higher productivity, lower employee turnover, and improved service quality over time.

On the other hand, excessive cost-cutting in human resources often produces the opposite effect: dissatisfied and unmotivated employees tend to perform worse, which ultimately undermines customer satisfaction and competitiveness. Effective management of labour costs therefore requires a careful balance between financial sustainability and the creation of a positive, stimulating work environment that supports both individual growth and business stability.

Within this context, artificial intelligence (AI) is emerging as a key driver of innovation in tourism. AI tools are increasingly used to enhance service quality, streamline operations, and tailor the customer experience. They enable predictive demand analysis, automate communication through chatbots, support dynamic pricing of accommodation and transport, and facilitate the use of robotic assistants in hotels, restaurants, and airports. Together, these applications help raise the competitiveness of tourism enterprises as well as destinations as a whole.

The purpose of this paper is to examine how artificial intelligence is reshaping the tourism industry — its main areas of application, the economic impacts it produces, and the challenges and ethical questions that accompany its growing use.

II. THEORETICAL FRAMEWORK – ARTIFICIAL INTELLIGENCE IN SERVICE INDUSTRIES

Artificial intelligence encompasses a range of technologies that allow computer systems to perform tasks that would normally require human intelligence, such as recognizing speech or images, understanding natural language, making decisions, and generating predictions based on data analysis. Within service-oriented industries, AI holds particular promise because it can automate repetitive and administrative work, support the delivery of personalized services, and enhance the overall customer experience.

In tourism, AI technologies are increasingly applied to improve guest interaction, streamline internal operations, and assist managers in making data-informed decisions. The most common applications of AI in tourism include:

- **Chatbots and virtual assistants** — providing 24/7 support to guests, handling routine queries,

managing bookings and responding automatically to frequently asked questions.

- **Predictive analytics** — used to anticipate demand patterns, analyse seasonality, adjust pricing strategies, and optimize the use of accommodation and transport capacity.
- **Personalized offers and recommendations** — by analysing guest preferences and behaviour, AI systems can tailor suggestions for hotels, restaurants, activities, and local attractions.
- **Robotic assistants and smart systems** — ranging from robots that help with check-in, room deliveries, or information service, to smart rooms that enable guests to adjust conditions via voice commands or mobile applications.
- **AI in digital marketing** — optimizing advertising, analysing user reviews and sentiment on social media, and targeting specific tourist segments.

The application of AI technologies in tourism builds upon the broader concept of servitization of the economy, where competitiveness increasingly depends on the quality and personalization of services. AI enables the creation of new data-driven business models and enhances the industry's ability to respond quickly to shifts in tourist behaviour and changing market conditions.

In addition to improving service quality and increasing operational efficiency, AI supports strategic decision-making based on big data collected through digital channels. In this way, artificial intelligence becomes a key tool for maintaining the competitiveness of tourist destinations and enterprises in the context of growing global competition.

III. APPLICATION OF ARTIFICIAL INTELLIGENCE IN TOURISM

Employee costs in tourism enterprises are generally higher than in most manufacturing sectors. Tourism businesses, especially those in accommodation and food service activities, are characterized by a high share of labour due to the nature of the work, which requires intensive human effort, while at the same time material costs—which typically represent the largest share in an average business entity—are relatively lower.

Tourism also has a very high share of jobs classified as elementary occupations according to the National Classification of Occupations [1], such as cleaners, washers, laundry workers and pressers, and kitchen helpers. According to the Institute of Economics, Zagreb [2], as many as 67.5% of individuals working in elementary occupations are employed in accommodation services, which explains why tourism industry is associated with lower average wages.

The average share of employee costs in all Croatian business entities subject to corporate income tax, according to 2023 data, amounts to 12.5% of average business revenues [3]. Within this, employee costs in manufacturing account for 13.3%, while in tourism they stand at 24.2% (hotels and similar accommodation), 22.3% (restaurants and other food service activities), and a below-average 9.8% in travel agencies. Overall, within the accommodation and food service sector, the share of employee costs in business revenues amounts to 22.9%. Based on the above, it is evident that the share of employee costs in tourism enterprises (excluding travel agencies) is more than 80% higher than in the average Croatian company.

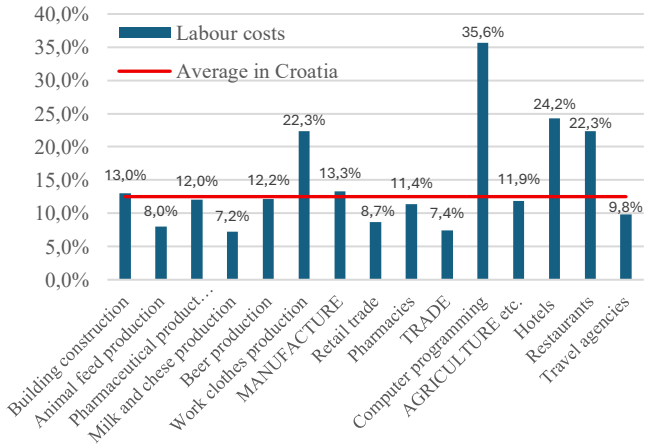


Fig. 1. Share of employee costs in business revenues of Croatian business entities in 2023 by selected sectors

Source: Author's elaboration based on [3]

In 2024, the average share of employee costs in business revenues of Croatian enterprises increased to 14.0%, and in the accommodation and food service sector to 24.6%. The total employee cost in this segment amounts to EUR 1,732 million, which represents 7.4% of the total employee costs of entities subject to corporate income tax, or 2% of Croatia's GDP.

In the early stages of the manufacturing industry, employee costs were also very high due to the strong reliance on human labour. With the introduction of the conveyor belt at Ford, the early phase of the automation process began, resulting in higher productivity and, consequently, profitability. Although technology was intensively introduced in the following periods, the share of labour costs in manufacturing remained stable, as new tasks and sectors generated additional demand for labour — in other words, the substitution of people by machines was balanced by the effect of new job creation. From the second half of the 1980s, the situation changed. The substitution effect of people by machines became stronger, while the job creation effect weakened, which resulted in a decrease in the share of employee costs in the manufacturing industry. Between 1987 and 2017, in the United States, the demand for labour in manufacturing declined by approximately 30% [4].

Tourism is a service activity which, due to its nature, requires a large amount of labour, but the work involved is mostly of lower added value. Therefore, average labour costs per hour in tourism are lower than in most other sectors, while due to the larger number of employees required, the total employee cost is higher. Because of the specific characteristics of service activities, the substitution of people by machines has not been possible as in

manufacturing, resulting in a recently significantly higher share of employee costs in tourism compared to manufacturing and even other industries.

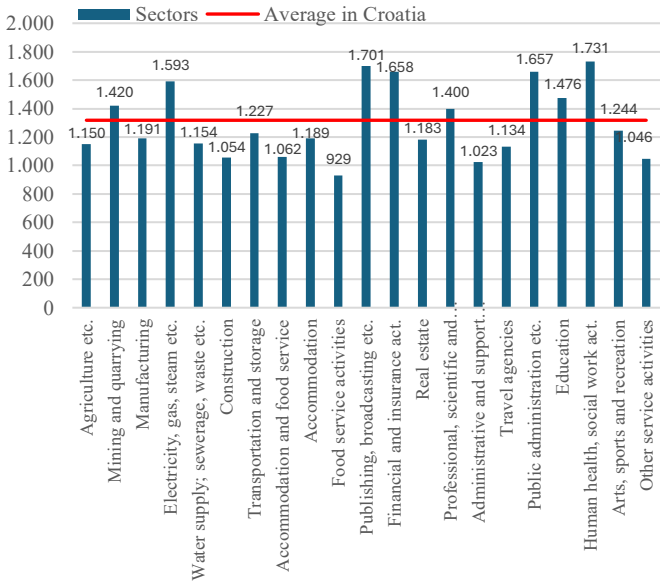


Fig. 1. Average monthly net salary per employee in legal entities in Croatia in 2024 by sectors

Source: Author's elaboration based on [5]

It was only with the emergence of artificial intelligence that more significant optimization of business processes and employee costs in tourism became possible. Due to the currently much higher share of employee costs in this sector, the potential savings effects are greater compared to other industries.

The application of artificial intelligence in tourism encompasses a wide range of technologies and solutions that enhance all stages of the customer experience — from planning and booking a trip, through the stay at the destination, to analysing guest satisfaction and loyalty after their return. AI technologies assist tourism enterprises in improving business processes, personalizing services, and optimizing marketing and sales activities.

A. Digital Assistants and Chatbots

Digital assistants and chatbots represent one of the most widespread and impactful applications of artificial intelligence in modern tourism. Systems of this kind are now widely used on websites, mobile applications, and social media platforms. They are based on Natural Language Processing (NLP) technology, which enables them to understand and interpret user queries. This technology is not only used to answer frequently asked questions but also to provide support in more complex situations, such as making reservations or selecting specific services.

The use of digital assistants brings numerous benefits to both tourism companies and their clients. Above all, it ensures constant availability of service, 24 hours a day, which increases customer satisfaction and loyalty. Continuous access to accurate and relevant information is particularly important in today's tourism environment, characterized by high expectations and the need for fast, high-quality service. Moreover, chatbots help reduce the workload of employees, especially in reception, reservation, and customer support departments, leading to better organization and cost optimization. Another important advantage of artificial intelligence technologies is the ability to automate various services, such as accommodation booking, ordering food, spa treatments, transportation, and other amenities. By integrating chatbots with hotel and tourism management systems, guest needs can be met faster and more efficiently, which further enhances the customer experience and increases revenue. Modern solutions also often include multilingual support, making communication with guests from different parts of the world easier thereby enabling market expansion.

B. Personalization of Offers

Personalization of offers is based on in-depth analysis of customer data. Using advanced data processing algorithms such as machine learning and deep learning, AI analyses search history, bookings, reviews, and online behaviour to generate personalized recommendations. These recommendations include different services such as accommodation, restaurants, tourist activities, and other attractions customized for the individual user. Such an approach not only increases customer satisfaction but also stimulates higher spending. For example, platforms like Booking.com use AI to suggest accommodations based on a user's previous reservations, preferred destinations, and budget, thereby increasing the probability of realized sales [6].

Personalization through AI relies on the system's ability to process large volumes of data in real time, considering variables such as demographics, seasonal trends, and user habits. This technology enables predictive analytics, allowing the system to anticipate future customer needs based on past behaviour. For instance, Airbnb uses personalization algorithms to suggest experiences such as cooking classes or adventure tours, based on user interests gathered from previous interactions with the platform [7]. Such recommendations enhance the customer experience and increase customer loyalty, because users feel recognized and valued.

Except improving customer satisfaction, AI-driven personalization has a significant economic impact. According to McKinsey & Company [8], companies using personalized marketing can reduce customer acquisition costs by 50%, increase revenue by 5% to 15%, and boost marketing return on investments (ROI) by 10% to 30%. In practice, this means that hotels and travel agencies can optimize their offerings, reduce advertising costs, and increase booking rates. For example, the Marriott hotel chain uses AI to send personalized email offers to guests, including discounts for destinations they previously visited or add-on services such as spa treatments, resulting in higher engagement and repeat bookings [9].

Despite its many advantages, applying AI in personalization requires careful data management to ensure user privacy and compliance with regulatory frameworks, such as the General Data Protection Regulation (GDPR) in European Union. Ethical challenges, such as potential algorithmic bias or excessive manipulation of customer decisions, are also significant. Nevertheless, when properly implemented, AI personalization not only improves the customer experience

but also provides a competitive advantage for companies in the dynamic tourism sector.

C. Dynamic Pricing

Dynamic pricing represents a modern revenue optimization strategy that allows the adjustment of prices for offered services in real time. AI systems use predictive analytics and machine learning to continuously monitor factors such as market conditions, seasonality, competitor pricing, consumer behaviour etc. This enables price adjustments in goal to achieve an optimal balance between revenue maximization and market competitiveness.

For instance, many hotels use advanced Revenue Management Systems (RMS) that automatically adjust room rates based on current demand, major local events, weather conditions, competitor activity and numerous other factors. If a popular event such as a concert or conference is taking place at the destination, prices may increase significantly, while during periods of low demand, prices are reduced to encourage bookings and increase occupancy. This dynamic approach enables hoteliers to offer "the right price at the right time." For example, Uber also apply dynamic pricing, known as surge pricing, where ride costs fluctuate depending on demand, traffic conditions, and driver availability. Amazon uses sophisticated AI algorithms to collect data on competition and inventory, adjusting product prices multiple times per day to remain competitive in the market. Airbnb uses AI to determine optimal accommodation prices, adjusting them based on local events, seasons, and competition, thereby contributing to maximizing revenue for its hosts [10].

In the airline industry, dynamic pricing has been used for decades, with prices varying based on number of factors such as overall customer demand, real-time supply, competitor pricing etc. Although Delta Airlines states that they are testing artificial intelligence to eliminate manual processes, accelerate analysis and adjustments, according to Singer [11] they are moving to personalized pricing in which the price depends on the personal circumstances of the customer. Some emphasised that such practice is unfair, bad for consumers or, in some cases, even illegal. However, Delta airlines plan that by the end of this year, 20% of prices on domestic flights within the U.S. will be set with the help of artificial intelligence. All aforementioned uses show how AI-based dynamic pricing has become an indispensable tool in modern tourism and transportation, improving business efficiency and customer experience.

D. Robotic Assistants and Smart Rooms

Robotic assistants and smart rooms are becoming increasingly common in luxury hotels, modern airports and other transport hubs. Robotic assistants take on various roles, such as receptionists, waiters, housekeepers, and guides, facilitating everyday tasks and improving the speed and quality of service. For example, major hotel chains such as Hilton and Marriott use AI platforms and chatbots to automate communication with guests and handle simple inquiries, such as information on opening hours or service availability. This frees up staff to focus on more complex tasks and personalized guest attention [12].

Smart rooms in these hotels are equipped with AI technology that allows guests to control lighting, climate, and multimedia through voice commands or mobile applications. The environment in rooms can be adjusted to each guest's preferences in order to increase comfort. For instance, a guest can change the room temperature, play a movie, or adjust lighting simply by voice, eliminating the need for physical controls and contributing to a modern, luxurious experience.

In airports, railway stations and bus terminals robotic assistants serve as guides for passengers, helping them with information about terminals, security procedures, and even luggage handling. These robots can use voice and facial recognition technologies, enabling multilingual interaction. Such automation reduces congestion at information desks and shortens waiting times, making travel more pleasant and less stressful.

Based on preferences and feedback, these systems can adapt services to be as relevant and convenient as possible. This not only increases guest satisfaction but also improves operational efficiency and reduces costs in hospitality, airports and other transport hubs. The advancement of robotic assistants and smart rooms reflects the broader trend of digital transformation in the tourism and hospitality industry, where these technologies are increasingly recognized as essential for sustainable and innovative business practices.

E. AI in Marketing and Customer Experience Analysis

AI in marketing and customer experience analysis plays a crucial role in enabling a more detailed and efficient approach to understanding customer needs and behaviours. Using techniques such as sentiment analysis, AI can automatically process large volumes of comments, reviews, and feedback from social media, review platforms, or direct communication channels. This analysis helps recognize positive, negative, or neutral customer attitudes, uncover key issues, and identify opportunities for improvement in offerings. Furthermore, AI enables market trend detection by monitoring shifts in preferences, emerging customer demands, or potential issues. This allows marketing teams to adjust their strategies, products, or communications in a timely manner when needed. For example, if AI detects rising interest in a specific product or service, marketing campaigns can focus on that segment, with the goal to increase revenues and profit at the end.

Advertising automation further transforms the way customers are reached. Through machine learning algorithms, AI can optimize budget allocation across different channels and ad formats, adjust messages in real time based on user reactions and personalize content for specific demographic and behavioural profiles. This achieves a high level of precision in targeted promotion, with lower costs and higher returns on investment in marketing. Through continuous monitoring and analysis of feedback, companies can react more quickly to market changes, increase customer satisfaction and build long-term loyalty and trust. AI allow the transformation of traditional marketing into digital, agile, and data-driven business.

IV. ECONOMIC EFFECTS OF ARTIFICIAL INTELLIGENCE APPLICATION IN TOURISM

In tourism the application of artificial intelligence brings numerous economic benefits, but also challenges, to

entrepreneurs, destinations, and the economy as a whole. By increasing operational efficiency, improving the customer experience, and adapting to market changes, AI has the potential to significantly influence the competitiveness of the tourism sector, employment, and economic growth. On the other hand, automation may lead to a reduction in jobs in certain segments, raising the question of labour market adaptation.

The most important economic effects of applying AI in tourism are:

- **Increased operational efficiency and cost reduction**

AI enables the automation of routine and administrative tasks such as bookings, invoicing, inventory management, and customer support. This reduces operating costs and increases the speed of delivered service. For example, fewer employees are needed at reception desks and customer service thanks to chatbots, bookings are processed more quickly, prices are dynamically set, etc. The result is reflected in the reduction of employee costs in the profit and loss accounts of entrepreneurs.

- **Revenue growth through personalization and predictive analytics**

Personalized offers based on AI analytics encourage higher guest spending and increase revenues. AI predictive models help businesses better plan capacity utilization, set optimal prices, and respond promptly to changes in demand. For instance, dynamic pricing enables higher occupancy rates at optimal prices, while greater average revenue per guest is achieved through personalized recommendations and additional services. The impact can be seen through an increase in revenue in the profit and loss accounts.

- **Creation of new jobs and specialized occupations**

Although AI reduces the need for certain types of jobs (e.g., administrative staff), it simultaneously creates demand for new professions such as:

- AI system management specialists
- Data analysts in tourism
- Digital marketing experts specialized in AI tools
- IT support for robotic assistants and smart systems

This opens opportunities for the development of a digital workforce in tourism and encourages lifelong learning. Moreover, tourism, traditionally a sector with below-average wages, has the potential to raise its average income levels. For example, the lower-paid role of a receptionist can be replaced by specialized AI system management experts.

- **Risk of reduced employment in routine jobs**

The greatest challenge of AI application in tourism relates to the potential reduction of jobs susceptible to automation, such as receptionists, reservation operators, and tour guides. Although new jobs will be created in digital tourism, some workers risk being excluded from the labour market if reskilling measures are not implemented. For example, according to Palrão et al. [13], by 2030 robots will account for around 25% of the "workforce" in the hotel industry.

- **Increased competitiveness and destination differentiation**

Destinations that integrate AI technologies into their operations in a timely manner can achieve a competitive advantage through:

- Improved customer experience
- More efficient destination management
- More precise targeting of tourists through AI marketing

This is particularly important for tourism-oriented countries such as Croatia, where seasonality and strong competition are ongoing challenges.

Ultimately, the total employee costs in the accommodation and food service sector in Croatia in 2024 amounted to EUR 1,732 million, or 24.6% of business revenues. If, through the utilization of artificial intelligence and other advanced technologies, employee costs were reduced to the average share of employee costs in business revenues of Croatian companies, savings of EUR 746 million would be achieved, representing 3.2% of all employee costs in Croatia in 2024, or 0.87% of that year's GDP.

Such a change would lead to a significant improvement in the performance indicators of tourism companies. Assuming *ceteris paribus*, based on the results of the accommodation and food and beverage service sector for 2024, a reduction in costs would increase the EBITDA margin to 30.2%, the net profit margin to 16.2%, the return on assets to 7.6%, and the return on equity to 17.3%. From a sector with below-average returns on assets and equity, tourism would become a sector with above-average returns — noticeably higher than the overall average in Croatia.

TABLE 1. OVERVIEW OF SELECTED FINANCIAL INDICATORS OF THE ACCOMMODATION AND FOOD SERVICE SECTOR FOR 2024

	Total accommodation and food		**All industries in Croatia**
	Actual (2024)	***Estimated (under cost reduction assumption)***	
EBITDA margin	19.6%	30.2%	11.8%
Net profit margin	7.3%	16.2%	5.9%
ROA	3.4%	7.6%	4.5%
ROE	7.8%	17.3%	10.8%

Source: Authors calculation based on [14]

V. CHALLENGES AND ETHICAL ISSUES IN THE APPLICATION OF ARTIFICIAL INTELLIGENCE IN TOURISM

Although the application of artificial intelligence in tourism brings numerous benefits, it also opens up a series of challenges [15], particularly regarding data privacy, ethical standards in the service sector, and impacts on the labour market. To ensure the positive effects of AI

technologies, it is necessary to address potential negative consequences of their use in a timely manner.

The most important challenges and ethical issues include:

- **Privacy and security of customer data**

AI systems in tourism rely on the collection and processing of large amounts of personal guest data — from booking and payment information to preferences and behaviour during their stay. Privacy concerns become particularly sensitive because:

- Tourism companies often use data for personalized offers without providing clear information to users.
- The transfer and storage of data, including outside the country, may compromise the security of personal information.
- AI systems may inadvertently generate bias and discrimination based on processed data.

- **Loss of human contact in service activities**

One of the key challenges is the potential erosion of the authentic customer experience, which in tourism has traditionally been based on personal interaction. Automation and robotic assistants can reduce the warmth and emotional component of service, which some guests may perceive negatively. In fact, the majority of tourists — particularly high-spending visitors — prefer personal contact over automated systems. According to a study [16], the segment of high luxury travellers (tourists spending at least USD 750 per night) most strongly expects digital detoxing and opportunities to meet new people while traveling.

- **Unequal access to technology**

The application of AI requires significant financial investment, placing small and medium-sized tourism enterprises at a disadvantage compared to large chains and international platforms. This raises the risk of further market consolidation, with smaller service providers losing market share. High market concentration can lead to an increase in service prices.

- **Responsibility for incorrect AI system decisions**

AI systems can occasionally deliver inaccurate recommendations, misjudge demand, reject reservations, or make other errors without justification. In such cases, this raises questions about who is responsible — the developer, the tourism company, or the system itself. Uncertainty about who is to blame inevitably leads to negative customer experiences, which adversely impact business performance.

VI. CONCLUSION

Artificial intelligence is becoming a key tool in the modernization and digital transformation of the tourism industry. Through the application of AI systems, it is possible to improve lots of aspects of tourism operations — from service personalization, dynamic pricing, and customer support automation to smart resource management and market analysis.

The economic effects of AI implementation in tourism are evident in increased operational efficiency, cost reduction, revenue growth, and enhanced competitiveness of tourist destinations. At the same time, AI creates new jobs in the digital sector but also carries the risk of reducing employment in traditional service occupations. There is great potential in reducing the significant labour costs and bringing them down to the average level of other industries. On Croatian example, reducing the share of personnel costs in companies in the accommodation and food and beverage service sectors to the level of the average Croatian company would result in savings equivalent to 0.87% of Croatia's GDP in 2024. Assuming *ceteris paribus*, tourism sector companies would transition from a sector with below-average return on assets and capital to one with above-average returns, thereby increasing interest in investing in Croatian tourism.

Key challenges are related to data privacy, preserving human contact in service activities, the accessibility of technology for small enterprises, and ethical issues regarding accountability for AI system decisions. Addressing these challenges requires coordinated action from legislation, the tourism industry, and the academic community.

In conclusion, AI constitutes an opportunity to redefine the tourism experience and promote the sustainable development of the industry and primarily by enhancing cost efficiency through the reduction of labour expenses.

REFERENCES

[1] Croatian Bureau of Statistics, "National classification of occupations (Nacionalna klasifikacija zanimanja)," Narodne novine, 147/2010-3736, 2010.

[2] The Institute of Economics, Zagreb, "Tourism and hospitality sector profile (Profil sektora turizam i ugostiteljstvo)," 2015. [Online]. Available: https://mint.gov.hr/UserDocsImages/AA_2018_c-dokumenti/180529_cekom_Profil-sektora.pdf,

[3] Fina, "Annual financial report of an entrepreneur (Godišnji financijski izvještaj poduzetnika), All companies, Year 2023", 2024.

[4] D. Acemoglu D. and P. Restrepo, "Automation and new tasks: How technology displaces and reinstates labor," Cambridge: National Bureau of Economic Research, 2019.

[5] Croatian Bureau of Statistics, "Average monthly paid off net earnings per person in paid employment in legal entities, according to NKD 2007, Months, Results from JOPPD forms," 2025. [Online]. Available: https://podaci.dzs.hr/hr/statistika-u-nizu/

[6] Booking.com, "Booking.com enhances travel planning with new AI-powered features for easier, smarter decisions," 2024. [Online]. Available: ews.booking.com/bookingcom-enhances-travel-planning-with-new-ai-powered-features--for-easier-smarter-decisions/

[7] M. Grbović, E. Wu, P. Liu, C. H. Tan, L. Wu, B. Yu, A. Tian, "Machine learning – powered search ranking of Airbnb experiences," Medium, The Airbnb tech blog, 2019. [Online]. Available: https://medium.com/airbnb-engineering/machine-learning-powered-search-ranking-of-airbnb-experiences-110b4b1a0789

[8] McKinsey & Company, "What is personalization?," 2023. [Online]. Available: https://www. mckinsey.com/featured-insights/mckinsey-explainers/what-is-personalization

[9] M. Ramirez, "Marriott International: AI use cases 2024," Pitchgrade, 2024. [Online]. Available: https://pitchgrade.com/companies/marriott-international-ai-use-cases

[10] AIAcademy, "What is dynamic pricing using AI? (Što je dinamičko određivanje cijena uz pomoć AI?)," 2025. [Online]. Available:

https://aiacademy.ba/blogs/wiki/sta-je-dinamicko-odredivanje-cijena-uz-pomocai#:~:text=dinami%C4%8Dko%20odre%C4%91ivanje%20cijena%20na%20osnovu%20lokalnih%20doga%C4%91aja

[11] H. Singer, "Delta's AI-based personalized pricing is bad for flyers. It should probably be banned," The Sling, 2025. [Online]. Available: https://www.thesling.org/deltas-ai-based-personalized-pricing-is-bad-for-flyers-it-should-probably-be-banned/

[12] Tragento, "Artificial intelligence in tourist industry: The newest implementations of known travel brands (Umjetna inteligencija u turističkoj industriji: Najnovije implementacije poznatih travel brendova)," 2025. [Online]. Available: 2025. [Online]. Available: https://tragento.com/umjetna-inteligencija-u-turistickoj-industriji-najnovije-implementacije-poznatih-travel-brendova-2025/#:~:text=AI%20automatizira%20odgovore%20na%20jednostavne%20upite%20gostiju

[13] T. Palrão, R. I. Rodrigues, A. Madeira, A. S. Mendes and S. Lopes, "Robots in Tourism and Hospitality: The Perception of Future Professionals, *Human Behavior and Emerging Technologies*," Volume 2023, Issue 1, [Online]. Available: https://doi.org/10.1155/2023/7172152

[14] Fina, "Annual financial report of an entrepreneur (Godišnji financijski izvještaj poduzetnika), All companies, Year 2024", 2025.

[15] Z. Bolfek, "Challenges of artificial intelligence (Izazovi umjetne inteligencije)", in *Proceedings of the 6th Symposium of the University of Applied Sciences Hrvatsko zagorje Krapina – Encounters 2024: Artificial Intelligence in Business*, M. Hercigonja-Szekeres and N. Sikirica, Eds. Krapina, Croatia: University of Applied Sciences Hrvatsko zagorje Krapina; Medicinska naknada d.o.o., May 23–24, 2024, pp. 26–31.

[16] McKinsey & Company "Updating perceptions about today's luxury traveler," 2024. [Online]. Available: https://www.mckinsey.com/industries/travel/our-insights/updating-perceptions-about-todays-luxury-traveler

WLAN Positioning Using SVM Models and Monte Carlo Simulations of Spatial Partitioning

Marko Burdžić
School of Electrical Engineering,
University of Belgrade,
Belgrade, Serbia
marko.burdzic@etf.rs

Aleksandar Nešković
School of Electrical Engineering,
University of Belgrade,
Belgrade, Serbia
neshko@etf.rs

Nataša Nešković
School of Electrical Engineering,
University of Belgrade,
Belgrade, Serbia
natasha@etf.rs

Abstract*— The increasing demand for precise indoor localization, driven by the limitations of GNSS in indoor environments and the wide availability of WLAN networks, has intensified research on WLAN-based positioning methods. This paper proposes a cascaded SVC-SVR model combined with spatial partitioning generated through Monte Carlo simulations and a Breadth-First Search (BFS) algorithm. The proposed method uses a Monte Carlo approach to find quasi-optimal subspaces, resulting in enhanced positioning accuracy compared to deterministic partitioning. Experimental results show that the best configuration achieved a mean absolute positioning error of 6.84 m, representing an 11% improvement over the best deterministic approach on the same dataset.

Keywords — Support Vector Machines, positioning, WLAN, spatial partitioning.

I. Introduction

The accurate determination of a user's location has become highly significant mainly due to the various applications of Location-Based Services (LBS) in modern-day telecommunications. The rapid growth of smartphones and other wireless devices in the last couple of years has resulted in a wide range of services including indoor localization. This field has been extensively researched over the past decades, particularly in industrial settings, wireless sensor networks, and robotics. User and device localization have wide-scale applications in health sector, industry, surveillance and other sectors [1]. Moreover, it plays an essential role in the development of modern systems such as the Internet of Things (IoT), smart cities, smart buildings, smart grids, and Machine Type Communication (MTC). Positioning systems are expected to provide location information in both outdoor and indoor environments. For outdoor localization, users often use global navigation satellite systems (GNSSs), such as the Global Positioning System (GPS), Galileo, GLONASS etc. GNSSs provide good positioning performance in outdoor environments. However, in indoor environments, where direct line-of-sight is unavailable, GNSSs perform poorly and have very limited usage based on severe indoor channel conditions, including shadowing and multipath fading [2]. Indoor localization is a highly complex and challenging task. Indoor spaces, in contrast to outdoor environments, have more pronounced effects of attenuation, reflection, diffraction, and scattering of signals. These effects significantly impact radio wave propagation, which diminishes the accuracy of determining positioning in indoor settings [3]. The field of indoor positioning has been advancing in different ways for some years and with different technologies: WLAN, Bluetooth, Zigbee, Ultra-wideband (UWB), Radio-identification (RFID), Ultrasound etc. [4].

For indoor positioning, WLAN has proven to be a practical solution because of its existing infrastructure, eliminating the need for additional hardware. The most used technique in WLAN positioning is fingerprinting, which includes two phases: offline and online. During the offline phase, Received Signal Strength Indicator (RSSI) data is collected at reference points (RP) to construct a radio map. In the online phase, the user's current location is estimated based on the similarity between the measured RSSI values and the reference values from the radio map [2]. If the data collection in the offline phase is carried out with sufficient precision, this method can achieve high accuracy.

In recent years, machine learning algorithms have been increasingly applied in numerous research studies to improve positioning accuracy [5]. In this study, Support Vector Machines (SVM) and spatial partitioning were used for the purpose of positioning in WLAN networks. The machine learning model is a cascaded architecture, which consists of Support Vector Classification (SVC) and Support Vector Regression (SVR) layers. In the first stage, the SVC classifier determines the subspace to which the current RSSI sample most likely belongs, while in the second stage, the corresponding SVR model estimates the precise (x, y) coordinates within that subspace. Various spatial partitions were evaluated against the performance of the model to improve accuracy. Previous studies have shown that the positioning problem is more efficiently solved by partitioning the space into subspaces [6], [7]. However, subspaces were defined either geometrically or based on expert intuition. This paper proposes a procedure for quasi-optimal subspace determination using a Monte Carlo approach, resulting in enhanced positioning accuracy.

The paper is structured as follows: Chapter II is the literature review. Chapter III describes the dataset used to train the models. Chapter IV discusses in detail the SVC-SVR model. Chapter V presents the training procedure. Chapter VI discusses the results from the work. The last Chapter VII presents the conclusions.

II. Literature Review

WLAN-based positioning research has grown at an exponential rate over the past two decades as it has proven to be very applicable to many types of systems and services. The first WLAN-based localization framework was RADAR [8] and used a comparison between measured RSSI data from a radio map and the use of a nearest neighbor method in signal space to determine a user's location. Later, the Horus system [9] developed a localization method based on clustering and significantly decreased the computational complexity of the localization process.

Probabilistic methods constitute important advancement over classical deterministic localization methods. Whereas deterministic methods rely only on comparing fingerprints and RSSI of the unknown point, probabilistic methods model the distribution of signals and utilize Bayes' rule to compute the most likely location. One of the first systems of its type was developed by Roos et al. [10] where it was shown

experimentally that using such a framework can produce significant reductions in positioning error in realistic environments. The tests were conducted in an office space of size 16 m × 40 m containing 10 access points and 155 reference points, with a total of 6200 training and 2400 test measurements. In the comparison of methods, the classical Nearest Neighbor (NN) produced an average error of 3.71 m, while the probabilistic versions produced substantially more favorable results: 2.57 m average error with the Kernel method and only 1.56 m average error with the Histogram method.

The KNN algorithm has been widely adopted for indoor positioning, as it is easy to implement and provides reliable results. The focus of the research article [11] is to create and demonstrate a new neighbor-weighting approach, called Improved Weighted K-Nearest Neighbor (IWKNN). In this approach, weight can be a factor not only of distance in the RSSI space, but also of a confidence factor related to the reliability of the signals. This method dramatically decreases the influence of unstable fingerprints, and as a result the system is much more robust. The authors perform a series of positioning experiments in indoor corridors 50 m × 16 m, containing 7 WLAN access points. The database was created from 48 RPs, with each RP being measured 50 times and testing 14 locations in the online positioning phase with an additional 700 samples. The proposed IWKNN approach was able to provide a much lower average positioning error of 1.24 m than the classical WKNN and KNN which had averages of 1.51 m and 1.76 m, respectively. The IWKNN method also had less error variance, providing a more stable localization.

In paper [12], the authors propose a new architecture that operates by combining two deep neural networks (DNN) that complement each other. The first network is shallower with three 128 neuron hidden layers that model global patterns of RSSI distributions, whereas the second deeper network has five 256 neuron layers that model fine-grained variations due to noise and multipath. The outputs are combined using a dynamic fusion mechanism, in which the weights are adapted to each individual input. As an example, this dynamic separation could allow the shallower network, with fewer computational resources, and a better generalization capacity, to be assigned a greater weight coefficient if it is dealing with a RSSI signal with low noise. In contrast, when the input is complex multipath, the deeper network would be assigned a higher weight coefficient, resulting in captures of the fine-grained variations more precisely. Consequently, both networks contribute to each iteration, allowing a trade-off between stability and accuracy, improving the overall localization error. The experimental evaluations were carried out in a lab of 40 m × 20 m, using 100 reference points and 8 AP. In the offline phase, 200 measurements were taken at each RP, and in the online phase, 50 test points with 200 measurements. In total, a combined 30000 samples were measured. The results show that the proposed Fusion-DNN with Bayesian hyperparameter optimization achieved an average positioning error of 0.87 m, compared to individual DNN models which had an average positioning error of 1.21 m, and classical algorithms that had average positioning errors of 1.53 m, 1.67 m, 2.02 m, for WKNN, KNN, SVM algorithms respectively.

Work presented in [13] suggested an ensemble model that combined a fuzzy classifier and Multilayer Perceptron (MLP) for WLAN indoor localization based on RSSI measurements. The experiment was conducted in the underground shopping mall car park in Chongqing (China), with coverage of 2000 m², with a total of 150 RPs. The experiment made use of the existing infrastructure and consisted of 16 access points in total. Part of the experiment involved collecting 400 samples from the RP in the offline phase, and then, the online phase involved collecting 40 test samples from each RP. The proposed methodology and conducted experiments used hierarchical clustering and employed a fuzzy classifier to ascertain the area and then employed an MLP to estimate position. The authors compared a range of algorithms: KNN, SVM, Decision Tree, Gaussian Naive Bayes (NB), Random Forest **(**RF), AdaBoost and Bagging. The results showed that the ensemble learning approach combining KNN and MLP achieved the best performance. The ensemble framework yielded an average positioning error of 3.015 m, representing a significant improvement compared to the standalone MLP (3.603 m) and RF (11.136 m). The main contribution of the paper is the potential of a fuzzy-MLP ensemble model which countered noise and interference from the signal to produce a valid positioning solution in complex spaces.

In studies [14], the authors used advanced neural networks, specifically Recurrent Neural Networks (RNN), in conjunction with the wireless location-fingerprinting method. The Long Short-Term Memory (LSTM) architecture was particularly researched. As in earlier work, there was an offline and online phase in the modeling. During the offline phase, 365 reference points were defined and 365000 measurements were collected. In the online phase, testing was conducted at 175 locations. Testing occurred over a 21 m × 16 m space with six access points. This study showed that RNN models outperform the traditional methods of positioning using KNN-RADAR algorithm [8], with the LTSM model having the best accuracy, generating an average positioning error of only 0.75 m.

The use of SVM algorithms in the context of indoor positioning was analyzed in [15], primarily focused on how the training set size, quality of the RSSI data, and quantity of access points affected localization performance. Research was completed in an indoor space of 221 m², with 130 RPs located at 1.5 m intervals. In the offline phase, RSSI values were filtered to remove unusable values, while in the online phase, the RSSI values from the last eight consecutive measurements taken from an unknown location were implemented to stabilize the signal. As such, the effect of disturbances which inhibit precise localization were significantly minimized. Within this experiment the SVR model was compared to a probabilistic approach and a neural network, displaying superiority with an average positioning error of 0.68 m. Furthermore, a version of the model was created to operate with resource-constrained devices, which only required 81 KB of memory.

III. DATASET

The dataset used in this research was collected at the School of Electrical Engineering, University of Belgrade. The WLAN system consists of 8 Access Points (AP) arranged to provide optimal radio coverage for wireless internet access. The database includes 1612 measurements collected

across 403 reference points (RP), with four measurements taken at each RP, each with RSSI values from AP1–AP8 and the corresponding (x, y) position. The measurement methodology, equipment used, and other details related to data collection are described in [6]. The entire space, with dimensions 147.1 m × 66.1 m is divided into 62 smaller space units.

IV. METODOLOGY

The aim of this research is to minimize the mean positioning error using machine learning techniques and spatial partitioning. In each iteration of the simulation, a cascaded SVC-SVR machine learning model is trained, consisting of two levels: an SVC classifier and an SVR regressor (Fig. 1). All rooms in the observed space are grouped into a predefined number N of subspaces, with each group consisting of mutually adjacent rooms. In order to efficiently group rooms, a graph is formed where nodes represent rooms, and edges connect adjacent ones. Spatial partitioning into N subspaces is performed by applying Breadth-First Search (BFS) algorithm [17]. For each node, the BFS identifies the node's neighbors. First, N initial nodes are randomly chosen and assigned to N different groups - one for each subspace. Afterwards, the BFS is started independently from each node, expanding toward neighboring nodes. The algorithm runs in cycles to avoid favoring a particular group. Each iteration will take one node from each group's queue. Every neighbor is explored, and those not yet assigned to any group, are added to the current node's group. This guarantees that any node can only belong to one group. This iterative process continues until all rooms are assigned to one of the subspaces, forming a spatial partitioning map. This process is executed within a Monte Carlo simulation, where many different spatial partitions into N subspaces are randomly generated. The positioning prediction process is defined as follows: RSSI data is input into the SVC classifier, which determines the corresponding subspace. Based on the output, one of the N SVR models is selected to predict the (x, y) coordinates.

V. TRAINING PROCESS

All models in this research were developed in the Python programming language, using the open-source *scikit-learn* library that offers a complete suite of machine learning algorithms and utilities. The training process commences with data preprocessing. The collected RSSI data were normalized using min-max normalization to fit the values into a [0,1] range. Using this preprocessing step leads to improved numerical stability, and computational efficiency during model training. The data were randomly divided into training and test datasets, with a 90:10 split. For each Monte Carlo iteration, a new spatial partition of the environment is created into N mutually exclusive subspaces, with each subspace containing rooms that are spatially adjacent to each other. In each partition of N subspaces, a new cascaded SVC-SVR model is trained. The first level of the model is the SVC classifier that predicts the class of subspace (1 to N) using as inputs RSSI vectors. The output class of the SVC serves as a conditional input to the second level, where corresponding SVR model predicts the current position (x,y). The SVC model is trained on the complete training dataset, while each SVR model is trained on a subset of training dataset belonging to specific spatial subspace. To enhance model

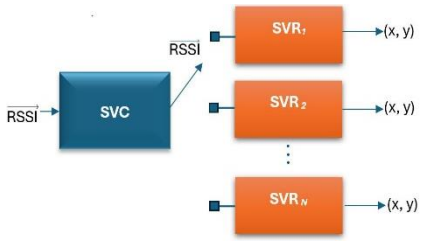


Figure 1 Cascade architecture of SVC and SVR models

generalization ability and further tune the hyperparameters during the training procedure, the *GridSearchCV* procedure with cross-validation was employed. By systematically training the model across all predefined hyperparameter combinations, the procedure enables the identification of the best-performing configuration. The following hyperparameters were optimized: *C*, *gamma* and *epsilon*. Besides hyperparameters, the selection of the kernel function is also quite significant. The kernel function enables the solution for nonlinear problems by mapping data into a higher-dimensional space where the problem becomes linearly separable. While solutions in higher-dimensional space may still use additional resources, the kernel trick enables the kernel function to be used in a higher-dimensional space without mapping or transformation of the data into that space [16]. Different kernel functions, including polynomial, sigmoid, and Radial Basis Function (RBF), were evaluated. The RBF kernel produced the best results for modeling the non-linear structure of indoor signal propagation. The C regularization parameter governs the trade-off between achieving low training error and good generalization. Higher values of C lead to narrow soft margins with smaller training error and an increased risk of overfitting, whereas low values of C have less tolerance to error that may lead to improved performance on unseen data. The values of parameter C were tested over the range [0.1, 1, 10, 50, 75, 100, 200, 300, 400, 500, 1000]. The best C performances were typically observed in the range of C=100. The *gamma* parameter is an important factor in the decision boundary with the RBF kernel. Small values of *gamma* lead to smoother boundaries and more generalization ability. The considered values for gamma were *scale* and *auto*. The *scale* value was selected since it automatically calculates the *gamma* value based on the number of features and the variance of the data. *Epsilon* is a parameter specific to SVR to define the margin in which errors are not penalized. The smaller the *epsilon*, the more sensitive the model is to changes in the training data and the greater the risk of overfitting. The considered values for epsilon were from 0 to 1 in increments of 0.1. The *categorical cross-entropy* function was used as the optimization function for the SVC model, while the *Root Mean Square Error* (RMSE) optimization function was used for the SVR model.

VI. RESULTS

After the training phase, the models were evaluated on the test dataset using spatial partitions with N = 4, 6, 8, 10, 12 and 14 subspaces. For each fixed number of subspaces, Monte Carlo simulations were then performed 100, 500, 1000, 5000, 10 000, 20 000, and 50 000 iterations. Each iteration represents different random spatial partitions of rooms within the same number of subspaces, allowing for a detailed analysis of how the number of subspaces and the

diversity of stochastic spatial configurations influence positioning accuracy. An analysis of various spatial partitions using the Monte Carlo method showed that the lowest positioning error was obtained when the space was divided into 12 subspaces, with 50 000 iterations and mean error of 6.84 m (Table 1). The results show the combination of Support Vector Machines (SVC and SVR) algorithm and the use of Monte Carlo Simulations with the BFS algorithm to spatially segment the space improved positioning accuracy than previously established [6], [7]. More precisely, the value achieved in this study represents an improvement of approximately 11% compared to the previous study on the same dataset with a mean error of 7.69m when using deterministic, geometric spatial segmentation [7]. The advantage of the Monte Carlo approach compared to fixed geometric partitioning of space lies in its ability to explore numerous spatial partitions, assessing the model's stability and error variability across configurations. In this way, the probability of identifying a spatial partition with minimal positioning error is increased. Apart from mean error, for each of the best configurations per number of subspaces, the standard deviation, and 67th and 95th percentiles, were also analyzed (Table 1), indicating overall stability and reduced variability across different configurations. In Fig. 3 the cumulative distribution functions (CDFs) are presented for the cases N = 4, N = 8, and N = 12. It can be observed that all three curves exhibit a similar increasing trend. The 67th and 95th percentiles also confirm the advantage of N = 12 (Table 1), demonstrating that it is better overall as it provides a more balanced trade-off between accuracy and stability of results.

VII. Conclusion

This paper presents a WLAN positioning method based on a combination of Support Vector Machine algorithms (SVC and SVR) and spatial partitioning via Monte Carlo simulations with the BFS algorithm. Unlike previous methods that were based on deterministic space partitioning, the proposed approach stochastically generates multiple spatial segmentations, increasing the likelihood of identifying configurations with the lowest positioning error. The experimental results demonstrate the method provides the best results with N = 12 subspaces and 50 000 iterations for a mean error of 6.84 m, which indicates 11% improvement over the best results of prior work on the same dataset. Furthermore, this study shows that the proposed method reduces the mean positioning error and ensures more stable and consistent results. The findings are clear that using Monte Carlo simulations with BFS based partitioning of space is a viable method for improving accuracy of indoor positioning systems in WLAN systems. Future work may extend the current method to include using unsupervised learning methods as an independent validation and to potentially reduce positioning error.

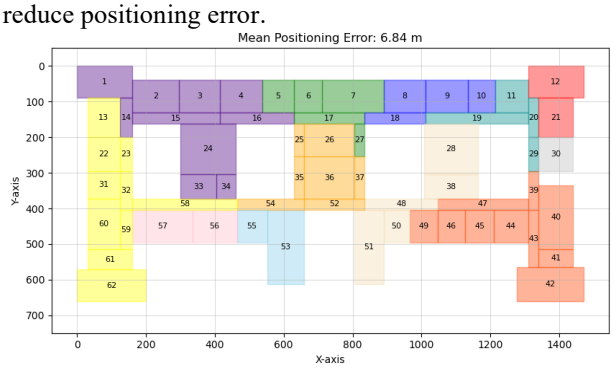


Figure 2 Spatial partitioning into N = 12 subspaces

Table 1 Model performance for different spatial partitions

Number of subspaces	4	6	8	10	12	14
Number of iterations	50 000	50 000	50 000	50 000	50 000	50 000
Mean positioning error [m]	7.34	7.17	7.18	7.13	6.84	7.08
Mean positioning error of SVC&SVR model [m] [7]	8.06	7.69	8.08	/	/	/
Mean positioning error of the ANN model [m] [6]	9.00	8.97	8.91	/	8.54	/
Standard deviation	6.13	6.95	6.86	7.08	7.02	7.11
Median error [m]	5.32	4.65	5.06	4.93	4.62	4.79
Maximum error [m]	34.45	39.31	39.40	40.07	40.61	38.93
Minimum error [m]	0.28	0.18	0.54	0.25	0.43	0.37
67th percentile	7.66	7.03	7.13	7.34	6.62	7.14
95th percentile	19.61	21.46	21.53	23.03	22.62	21.84

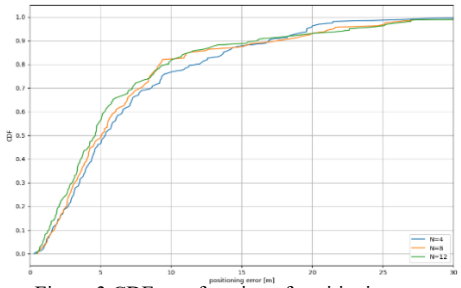


Figure 3 CDF as a function of positioning error

References

[1] F. Zafari, et al. "A survey of indoor localization systems and technologies." *IEEE Communications Surveys & Tutorials* 21.3 (2019)

[2] N. Singh, S. Choe, and R. Punmiya. "Machine learning based indoor localization using Wi-Fi RSSI fingerprints: An overview." *IEEE access* 9 (2021): 127150-127174.

[3] A. Nessa, et al. "A survey of machine learning for indoor positioning.", *IEEE access* 8 (2020): 214945-214965.

[4] T. Yang, A. Cabani, and H. Chafouk. "A survey of recent indoor localization scenarios and methodologies." *Sensors* 21.23 (2021)

[5] P. Roy and C. Chowdhury. "A survey of machine learning techniques for indoor localization and navigation systems." Journal of Intelligent & Robotic Systems 101.3 (2021)

[6] M. Borenović and A. Nešković, "Positioning in WLAN environment by use of artificial neural networks and space partitioning.", Annals of telecommunications-annales des télécommunications 64 (2009)

[7] M. Burdžić, A. Nešković, and N. Nešković, "Positioning in WLAN networks based on SVM algorithms and space segmentation," TELFOR 2024

[8] P. Bahl, V. N. Padmanabhan, "RADAR:An in-building RF-based user location and tracking system." Proceedings IEEE INFOCOM 2000

[9] Youssef, Moustafa, and Ashok Agrawala. "The Horus WLAN location determination system." *Proceedings of the 3rd international conference on Mobile systems, applications, and services*. 2005.

[10] Roos, Teemu, et al. "A probabilistic approach to WLAN user location estimation." *International Journal of Wireless Information Networks* 9.3 (2002)

[11] C. Li, Z. Qiu, and C. Liu. "An improved weighted k-nearest neighbor algorithm for indoor positioning." *Wireless Personal Communications* 96.2 (2017)

[12] F. Mahdavi, H. Zayyani, and R.Rajabi. "RSS localization using an optimized fusion of two deep neural networks." *IEEE Sensors Letters* 5.12 (2021)

[13] Q. Zhu, et al. "Accurate WiFi-based indoor localization by using fuzzy classifier and mlps ensemble in complex environment." *Journal of the Franklin Institute* 357.3 (2020)

[14] M. T. Hoang, et al, "Recurrent neural networks for accurate RSSI indoor localization." *IEEE Internet of Things Journal* 6.6 (2019)

[15] K. Shi, et al. "Support vector regression based indoor location in IEEE 802.11 environments." *Mobile Information Systems* 2015

[16] Tharwat, "Parameter investigation of support vector machine classifier with kernel functions." Knowledge and Information Systems 61 (2019): 1269-1302

[17] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed., MIT Press, 2009