APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

August 2025
Vol. 40 No. 8
ISSN 1054-4887

The ACES Journal is abstracted in INSPEC, in Engineering Index, DTIC, Science Citation Index Expanded, the Research Alert, and to Current Contents/Engineering, Computing & Technology.

The illustrations on the front cover have been obtained from the ARC research group at the Department of Electrical Engineering, Colorado School of Mines

Published, sold and distributed by: River Publishers, Alsbjergvej 10, 9260 Gistrup, Denmark

AUGUST 2025 REVIEWERS

Sabry A. H.
Dinesh Kumar Reddy Basani
Biswajit Dwivedy
Venkata Surya Teja Gollapalli
Srilakshmi Inuganti
Bashra Kadhim
Lida Kouhalvandi
Matteo Bruno Lodi
Vinoth M.
Enrique Melgoza
Durga Prasad Mishra
Andrew Peterson
Kannadhasan S.
Ehsan Akbari Sekehravani
Fusun Oyman Serteller
Somchat Sonasang
Jun Wang
Wei-Chung Weng
Mingwen Zhang
Xuemei Zheng

THE APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

Vol. 40 No. 8 August 2025

TABLE OF CONTENTS

Deep Learning-Based Hybrid Multivariate Improved ResNet and U-Net Scheme for Satellite Image Classification to Detect Targeted Region

P. Pabitha Muthu[1], S. Siva Ranjani[1], and M. Manikandakumar[2]

[1]Sethu Institute of Technology
Kariapatti, Virudhunagar District-626115, Tamil Nadu, India
pabithapraveen@gmail.com, Sivaranjani222@gmail.com

[2]Department of AIML and Data Science
Christ University, Kengeri Campus, Karnataka-560074, India
manikandakumar.m@christuniversity.in

***Abstract* –** In the field of remote sensing, the process of segmentation and classification of satellite images is a challenging task attributable to different types of target detection. There are problems in recognizing a target and clutter region. Then, there is a necessity to consider these problems regarding the classification of satellite image using an effectual approach. In this approach, a deep learning dependent automated segmentation, detection, and classification of satellite images is carried with artificial intelligence methods. Initially, the input image is preprocessed, segmented using Edge-ROI and YOLO v3 based segmentation in which the parameter is tuned by means of multi-heuristic tuna swarm optimization (MH-TSO) approach and is classified using hybrid multivariate improved residual network (ResNet) and U-Net classifier approach. The stage of Edge-ROI segmentation and YOLO v3 based segmentation is employed to extract regions. The preprocessing is carried using median average filtering along with adaptive histogram equalization. A scheme of deep learning-based multivariate improved residual neural network for classification of satellite images is proposed effectively. The proposed technique performance is estimated for three kinds of dataset, namely Salinas, Pavia University, and Indian Pines satellite image datasets, and the results obtained are shown, which proves the efficiency of the suggested mechanism.

***Index Terms* –** Hybrid multivariate improved residual neural network and U-Net, multi-heuristic tuna swarm optimization, remote sensing, satellite image, YOLO v3 segmentation.

I. INTRODUCTION

At present, scholars have focused on image classification which is thus regraded as a pivotal technology for the recognition of pattern and a computer vision scheme. Usually, the satellite images are in a digital form [1]. The technique of image processing could be employed for the retrieval of precise set of data from the stored images [2][3]. Therefore, this technique will be obliging in improving the visual perception and to modify or repair an image which depends on blurring of image, image deformation, or image deterioration [4]. Few methods are still available to analyze the image which relies on some specified issue requirements [5-7]. Image segmentation and classification algorithms are utilized in various regions of image in a thematic group. The kinds of image segmentation are categorized as four types from varied range of angles like feature space clustering, region-based, model-driven, and threshold-based [8,9]. Presently, the feature space clustering approach is regarded as the most significant and popular method for image segmentation due to its easy calculation, simple principle, and good segmentation. Furthermore, the images are normally segmented using a single clustering technique, and it is probable to get stuck in a local optimum since the boundaries are blurred and this displays unclear boundaries with deprived visual effects.

The primary problems are not only lies in capturing images, it also lies in processing the captured information in a fast manner and disseminating them at any instant of stated target classification and detection which is estimated with their accuracy range [10]. The key technique for such issue which is associated to the target recognition was thus to separate several mechanisms of target by recognizing the one which is fascinating from a residual one. The method of classification is still a complex task and it is a valid concern for the detection of target satellite systems [11,12]. There are problems in recognizing target and clutter regions. Thus, it is a necessity to consider these problems regarding the classification of satellite image using an effectual approach.

This paper is systematized as follows. Section II is the portrayal of several traditional techniques that are prevailing. Section III offers a brief explanation

Submitted On: June 13, 2024
Accepted On: July 25, 2025

https://doi.org/10.13052/2025.ACES.J.400801

regarding the overall work of the proposed methodology and novel mechanism. Section IV is the performance and comparative estimation of the proposed technique with traditional ones. Section V concludes the entire work objective.

II. RELATED WORKS

This section is a study of several techniques currently available. The author in [13] proposed Deep Convolutional Neural Network (DCNN) for the purpose of classifying the targeted regions from the input images with the utilization of convolution layers and without fully connected layers. In [14], the author presented an architecture that has limited training data that were termed a deep CNN highway unit intended for target classification in remote sensing images. The technique presented was found to be better in attaining a targeted region with a better range of specificity.

A target recognition model was presented in [15] that has some additional features that were dependent on CNN architecture with features for the remote sensing images. The suggested CNN technique consists of two steps. First is the extraction of features using max-pool and average pooling functions. Then training was performed with fully-connected layers. The suggested technique attains high recognition accuracy of about 94.38% for about 10 classes in the military target's detection in the MSTAR dataset.

An automatic target detection of satellite images was projected in [16] with the use of two kinds of machine learning schemes which offers better classification accuracy. Primarily, the support vector machine (SVM) was utilized along with a proper optimization approach for developing a set of local features. Then, the novel structure of CNN was implemented and both methods were analyzed with practical aspects.

A new technique of target recognition was suggested in [17] for the purpose of target recognition using CNN. Initially, from MSTAR datasets the targeted images are classified with the use of CNN. Then, the features were extracted from CNN output from convolution layer from which the position of targets were located in the image with the use of clustering and sampling approach. The suggested technique not only comprehends detection of target efficiently but also offers contented accuracy similarly in classifier function.

A multi feature-based convolution neural network (MFCNN) was presented in [18] for the recognition of targets from input image. First, from the satellite images the features were extracted. Then it was aggregated by the complementary relations as a single column vector. Finally, the target was identified with the use of fully connected networks. The suggested method has the advantage that it has the capability to recognize targets of MSTAR dataset images accurately on comparing other existing methods and was capable of offering information regarding the pose.

A speckle-noise-invariant CNN was projected in [19] along with regularization for decreasing the noise present in image and so as to recognize targets from images. Image despeckling was performed using Lee sigma filter to extract features. Then, the targets were recognized from the despeckled images using CNN training. The suggested CNN along with approach of regularization improves the recognition accuracy on relating traditional methods.

A recurrent neural network (RNN) [20] was proposed for automatic target recognition (ATR). Once electromagnetic waves from a radar incident on the targets, scattering of the incident energy might rise and this scattered signal was recognized as a target radar signature. The anticipated RNN based technique offers an accuracy of 93% in terms of classification.

Remote sensing image classification was a typical application [21]. In an attempt to augment the remote sensing image classifier outcomes, multiple classifiers combinations were utilized for classifying the Landsat-8 Operational Land Imager (Landsat-8 OLI) images. Few classifier techniques and algorithms combinations were scrutinized. The combined classifier encompassing of five classifiers adherent was made. The consequences of member in both classifiers were evaluated. The voting strategy was examined to take part the classification consequences of classifier member. The results exemplifies that the classifiers have unrelated performances and the grouping of several classifiers brings enhanced outcome on comparing the single classifier, thereby this approach conquers overall higher accuracy of classification. An approach was supported in [22] regrading the flexible multiple-features hashing learning context which is meant for LSRSIR which took several features that were balancing as an input which acquires the function of feature mapping hybrid approach [23].

A pre-trained convolution neural network (CNN) was presented in [24] for ATR. The main shortcoming of ATR was that there was an availability of only limited datasets for training purposes and for this reason a deep learning model was employed. The projected scheme was dependent on employing a pre-trained CNN specifically AlexNet which was employed for feature extraction and after which the output features were trained using multiclass SVM classifier. The pre-trained network proposed offers an accuracy of about 99.4% for three targeted modules or classes.

An approach in [25] aims at tackling problems associated with ATR. The suggested scheme depends on the approach of transfer learning at which three varied CNN that were pretrained are employed as feature extractors

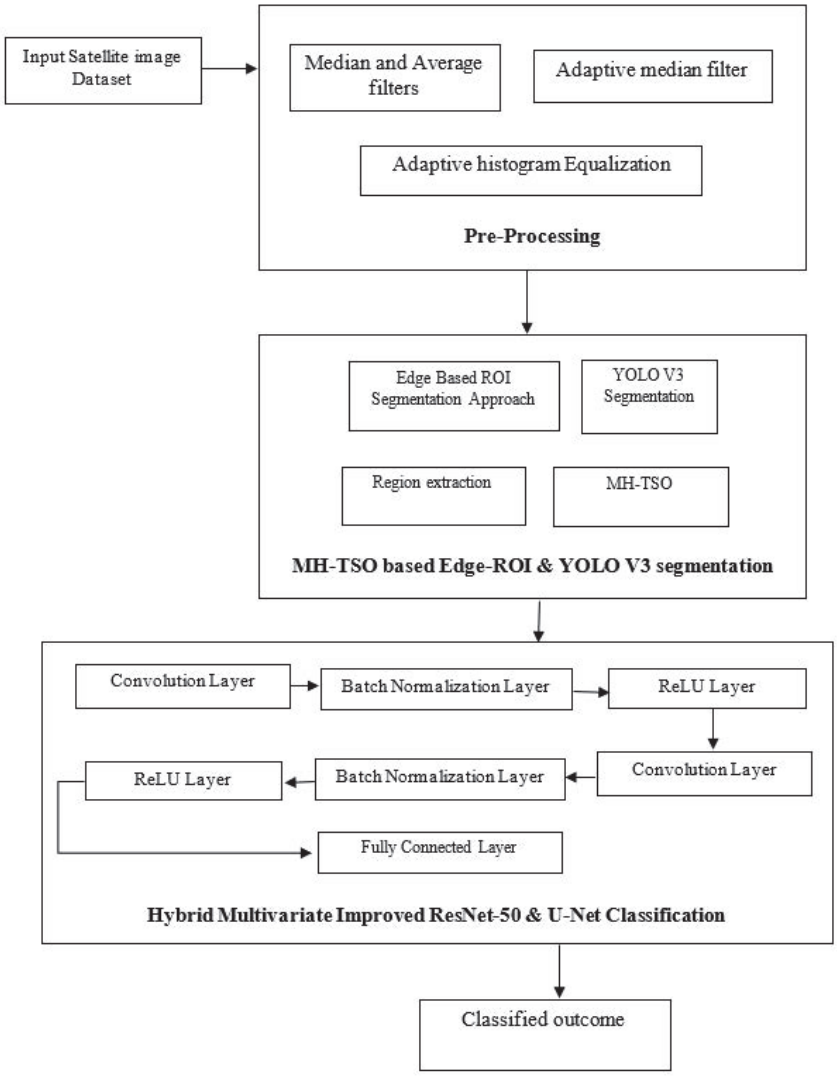


Fig. 1. Overall flow depiction of suggested scheme.

along with the integration of SVM. In this approach, the suggested classifier models are AlexNet, Google Net, and VGG16. On analyzing the performance, those three CNNs were capable of competing the issues of ATR at which AlexNet offers accuracy of about 99.27%.

A comparison was made in [26] among pre-trained CNN and SVM for the recognition of target so as to extract features and it was found that CNN offers better accuracy than others. Since there were several techniques employed for the classification and recognition of target in satellite images, there is a need to enhance the accuracy level so as to attain better detection of targets.

III. PROPOSED WORK

Our proposed methodology is explained briefly in this section. The overall flow depiction of the suggested method is given in Fig. 1. A deep learning-based model is implemented in this work for the better recognition and classification of targets from the satellite images. Initially, the input image is preprocessed, segmented using Edge-ROI and YOLO v3 based segmentation at which the parameter is tuned by means of multi-heuristic tuna swarm optimization (MH-TSO) approach and was classified using hybrid multivariate improved residual network (ResNet) and U-Net classifier approach. The stage of Edge-ROI segmentation and YOLO v3 based segmentation is employed to extract regions. The preprocessing is supported using median, average filtering along with adaptive histogram equalization. A scheme of deep learning-based multivariate improved residual neural network for classification of satellite images is proposed.

A. Input image pre-processing

Once the input image is collected, preprocessing is the initial and significant step for classifying satellite images. The process of preprocessing is employed so as to confirm the accessibility and reliability of the input images. In this, every step is significant since it aims at reducing the workload of further functions. This preprocessing step is carried out with the use of filters, histogram equalization techniques for detecting unintentional errors which could impair the capability of artifacts to prevent noise in the image. The abnormality areas are extracted from the satellite images. A nonlinear filtering system is termed as adaptive median filter that is most widely employed in eradicating noise present in image. The major benefit of this noise reduction process is that it is an archetypal step to yield better performance. Figure 2 shows the input image and noisy input image.

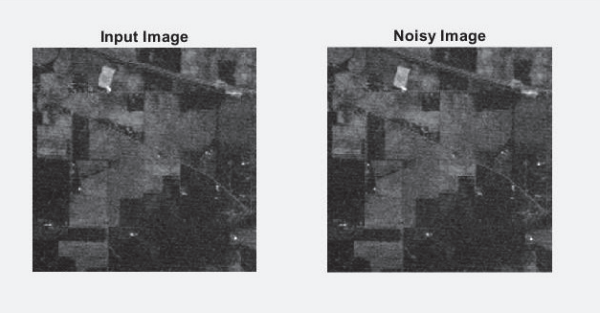


Fig. 2. Input image and noisy image.

This is the preliminary step of the suggested approach which eliminates background noises. Salt and pepper along with Gaussian noise will be present in satellite images which should be removed to attain a high-quality image. The preprocessing step is accomplished to delete these redundant noises in the image. Unnecessary errors are identified by employing filtering and histogram techniques. In this phase, the extracted noise is filtered through an adaptive filtering technique which is employed for the purpose of isolating noise from the image and it acts as a non-linear optical filterer. A typical approach for noise eradication is carried for improving the efficiency of the system and thus to augment the image contrast. Gaussian noise will be rectified by means of adaptive Wiener filter (AWF). The filtering process aids in offering bright and clear images. The use of this AWF aids in offering smooth and less variant images.

A non-linear spatial filter that is most widely employed is the median filter which in turn replaces the pixel values by the gray level medians in the desired neighborhood of that pixel. The adaptive median filter is a kind of statistic order since their response depends on pixel ranking of the image area covered by means of a filter. It is widely employed since this offers reduction of noise by blurring fewer edges. This noise reducing effect depends on two major factors: number of noise pixels involved in the computation of median and the spatial extent of their neighborhood.

Normally, for augmenting the consistency of the system, histograms are equalized. Histogram equalization is regarded as the computerized process of enhancing the image contrast. The most significant values of sensitivity are increased by this. The spectrum strength of the image is widened. After this process the average image contrast is augmented.

Let us assume q signifies the normalized histograms for intensities that are all possible. This can be expressed as:

$$q^{x} = (\text{number of pixels with x intensity}/\text{total number of pixels}), \quad (1)$$

where x = 0, 1…, x-1.

Thus, the histogram equalized image is signified as:

$$H_{i,j} = \text{base}(\,(x-1)\sum_{x=0}^{b_{i,j}} p^{x}. \quad (2)$$

Here, base signifies the adjacent integer and is equivalent in transforming the intensity of pixel. The probability distribution function PDF of this is expressed as:

$$\frac{\partial N}{\partial y}\left(\int_{0}^{N} pN(y)\,dz\right) = \partial N(N)(y^{-1})(N))d/dN. \quad (3)$$

The probability distributed standardization function can be signified as $\frac{\partial N}{\partial y}$.

At the time of preprocessing, the grayscale channel is blocked. Usually, the poor contrast image appears with satellite objects. The process of preprocessing is employed for augmenting the greyscale channel's contrast. Histogram equalization is a technique employed for escalating the image contrast and is followed by growing usual values of intensity effectively. This is expressed as:

$$\text{histogram equalization}\,(q) = \text{intensity sum of the pixel}/\text{pixel average number}. \quad (4)$$

This approach not only enhances image quality but also augments the performance of the system.

B. Edge-ROI and YOLO-V3 based segmentation approach for extraction of regions

Detection of boundaries in the image border is carried out in this step. Understanding the characteristics of image edge detection is the most crucial step since edges comprise significant and necessary information and features. An image size is reduced considerably after filtering process so as to retain indispensable functional properties of an image. Typically, edge detection is employed

for extracting boundaries or edges of image and is extensively used in image segmentation process. In the initial derivatives of the image, edges are sensed by gradient approach at which the edges are minimal. For this reason, suitable thresholding is employed to attain sharp edge isolation. Region of Interest (RoI) is evaluated that is divided to compute dimensions at which target edges will be extracted from the desired region. Figures 3 and 4 show the segmented image at frame 0 and frame 30.

The approach of segmentation determines whether the neighbor of the region should be included or not. A regional mechanism is employed for combining the sub regions or pixels as bordered areas. A pixel aggregation is the simpler one which begins with collection of seed point from grown regions once the neighbor pixel has same features which will be attached to every seed points. This is expressed as:

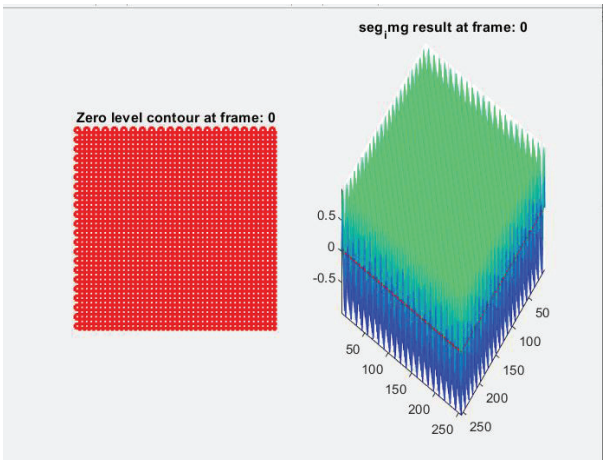


Fig. 3. Segmented image at frame 0.

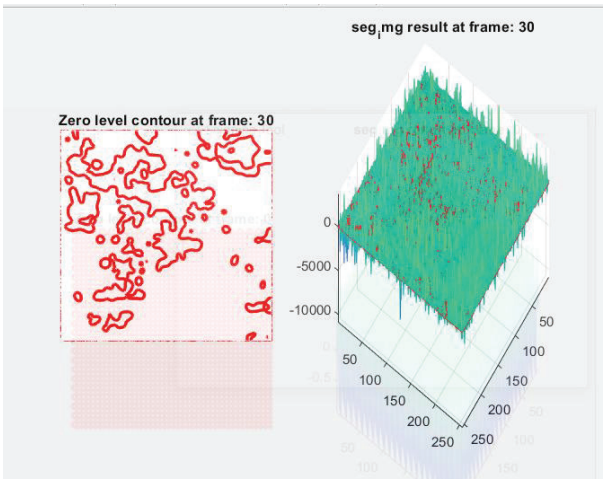


Fig. 4. Segmented image at frame 30.

$$\mathrm{ROI}^{\mathrm{Segmentation}} = \sum_{\{i,j\}\in Q_2}^{T} V_2\left(L_i, L_j\right).M.\log_{c_i} + \gamma \int c_i dx, \tag{5}$$

where $\mathrm{ROI}^{\mathrm{Segmentation}}$ refers to region of interest segmentation, velocity gradient is represented by V_2, L_i, L_j denotes the lower and higher pixel values of image, M signifies pixel numbers block of the image, $\log_{c_i}$ denotes the image's spatial size, γ signifies the image's frequency coefficient, c_i denotes the pixel distance.

The image frequency co-efficient is thus tuned by means of MH-TSO technique. The optimal features selection thus enhances the segmentation outcome. The optimal range of features are thus selected using MH-TSO algorithm.

Initial step is the spiral foraging model at which tuna will swim once it is nursed thus forming spiral ones for derivation of shallow water feed which is simply attacked. Next step is the parabolic foraging at which corresponding tuna thus swims subsequently making a parabola shape for enclosure of its prey. Using these techniques, tuna forage successfully. Mathematical modeling of this optimization technique is:

$$y_i^{in} = rand.(ub - lb) + lb, \qquad i = 1, 2, \ldots, NP, \tag{5}$$

where y_i^{in} refers to the ith initial individual, ub and lb are the upper and lower boundaries of search space, NP refers to the number of tuna population and *rand* signifies the random vector which is uniformly distributed and ranges from 0 to 1.

The anticipated mathematical modelling of spiral foraging is:

$$y_i^{t+1} = \begin{cases} \alpha_1.\left(y_{rand}^t + \beta.\left|y_{rand}^t - y_i^t\right|\right) + \alpha_2.y_i^t, & i = 1 \\ \alpha_1.\left(y_{rand}^t + \beta.\left|y_{rand}^t - y_i^t\right|\right) + \alpha_2.y_{i-1}^t, & i = 2, 3, \ldots, Np, \end{cases} \tag{6}$$

where y_{rand}^t signifies the reference point which is randomly generated at search space. Specifically, typical algorithms that are metaheuristic might transmit wide global exploration at a previous stage and then transfer to exact local exploitation slowly.

The anticipated mathematical modelling of parabolic foraging is:

$$y_i^{t+1} = \begin{cases} y_{best}^t + rand.\left(y_{best}^t - y_i^t\right) + TF.p^2.\left(y_{best}^t - y_i^t\right), & if\ rand < 0.5 \\ TF.p^2.y_i^t, & if\ rand \geq 0.5 \end{cases} \tag{7}$$

$$p = \left(1 - \frac{t}{t_{max}}\right)^{(t/t_{\max})}, \tag{8}$$

where TF refers to the random number having the value 1 or -1. The tuna hunt cooperates over two approaches of foraging and identifies the prey.

Deep learning-based YOLO v3 is employed for the segmentation process to get accurate segmented regions.

Typical versions of YOLO do not have localization error resolution. The newer version of this was not capable of identifying the smaller objects. YOLO v3 is designed for handling the existing limitations and thus offers an enhanced output during testing phase. Version 3 offers enhanced outcome for small size objects. It does not offer precise output for huge objects. The YOLO v3 architecture depends on Darknet-53 strategy that utilizes 3×3 and 1×1 convolutional filters having scarce shortcut connections with 53 convolution layers which is twice as fast as ResNet-152. YOLO v3 is stimulated through Feature Pyramid Network (FPN) that covers heuristics like up-sampling, residual blocks and skip connections. It provides Darknet53 as a ground network, that integrates 53 more layers to detect objects more easily. Like FPN, YOLO v3 utilizes 1×1 convolution to detect objects. Feature maps are produced at three different levels. It down samples the input by a factor of 32, 16 and 8 at three varied levels. On employing a stride of 32, in the 82nd layer the result tensor is 13×13 feature map which is used to detect with the help of 1×1 convolution. After applying a stride of 16, the detection is done after the 94^{th} layer. Adding some convolutions on the 79^{th} layer is done and then combined with 61^{st} layer on 2x upsampling to get a 26x26 feature map. After applying a stride of 8, detection is made in the 106^{th} layer with the help of 52×52 feature map. Some convolutions are combined to 91^{st} layer and then integrating with the 36^{th} layer with the help of 1x1 kernel, the down-sampled feature maps are integrated with the up-sampled maps at various locations to get fine grained features to detect the smaller objects of different sizes. The varied feature maps 52×52, 26×26, and 13×13 are employed for detecting huge, medium, and smaller objects correspondingly.

C. Classification using hybrid multivariate improved ResNet and U-Net classification

A ResNet and U-Net classifier is employed in this section after the segmentation process. The satellite image is recognized and is classified so as to recognize areas of vegetative lands from the input satellite images. This is the method employed for pre-trained convolution. CNN permits the valuation of inconsistencies among one or two objective variable quantity and one or both of them. The hybrid multivariate improved residual network and U-Net classifier offers a likelihood and the statistical distributions are usable. First, CNN reads the input thereby performing a resizing operation by measuring class likelihoods of validation phase. CNN in scrips a substantial development in the analysis and identification of images.

A hybrid multivariate improved ResNet and U-Net classifier is well-organized in the following layer forms: Convolution layers, Pooling layers, ReLU layers, Fully connected layers. This has predetermined preprocessing procedures, taking into consideration other images for classification purposes. This hybrid multivariate improved ResNet and U-Net classifier is employed for several determinations in many fields.

Hyperparameter selection is done manually or automatically. In this model, manual selection is carried out due to the highly complex nature and computational cost of the automatic models. The dataset is evaluated by the suggested ResNet classifier model by employing an optimization approach. The hyperparameters that are needed for estimating the model parameters are hidden layers, activation functions, data augmentation, kernel size, padding, and stride. The hyperparameters that are employed for controlling the training algorithm behavior are learning rate, batch size, momentum, and number of epochs. In this model, batch normalization is employed after the convolution layer and before the activation function. This technique is employed with suitable optimizer like stochastic gradient descent (SGD) optimizers at which learning rate and weight decay of ResNet system provides an improved stage of final trainings and validation accuracy that overcomes issue of overfitting thereby reducing the computational complexity of training process. Hyperparameter tuning parameters for the presented ResNet system are shown in Table 1.

Table 1: Hyperparameter tuning parameters for the presented ResNet system

Hyperparameter	**Value**
Batch size	128
Epochs	200
Rate of learning	0.04
Momentum	0.9
Weight decay	$5e^{-4}$
Drop out fully connected	50%
Optimizer	Stochastic gradient descent
Activation function	ReLU

This classifier makes it possible to estimate the inconsistency among one or more of the variables. The multipath improved ResNet classifier estimates the probabilities. In this method, the classifier will initially calculate, redistribute, and then interpret the class probability of the image:

$$\text{Fea} = \det[\text{p}] - \text{k}\,(\text{classify}\ (\text{N}))^2\,, \tag{9}$$

where Fea signifies the feature, p denotes the pointed feature. The following expressions are made:

$$\det[\text{p}] = \beta_1\beta_2, \tag{10}$$

$$\text{classify}\,(\text{p}) = \beta_1\beta_2, \tag{11}$$

where $\beta_1\beta_2$ represents the features classified. The CNN classification was concluded as:

$$\text{Fea} = \beta_1\beta_2 - \text{L}(\beta_1 + \beta_2)^2, \quad (12)$$

where L is the empirical constant.

Algorithm 1: Multivariate improved-residual network classification

---

Input: Feature selected and extracted-processed image δ_{im}
Output: Classified image ψ_c

Convolution layer

Sample's data feature vector is

$F_i = [F_1, F_2, \ldots, F_n]$ where n is the number of input features.

The output results are computed

$$conv_u^K = Fun_A(B_u + \sum_{m=1}^{T} wt_u^m F_u^{m-1})$$

Where $conv_u^K$ is convolution feature map of the K^{th} layer and u^{th} feature map.

$Fun_A(.)$ is activation function for the convolution layer.

B_u is bias term of the u^{th} feature map.

wt_u^m is weight for the u^{th} feature map and m^{th} filter index.

T is the kernel / filter index.

K is the layer index.

Max-pooling layer:

The pooling of a feature map in a layer

$$P_u^K = \max_{w \in W_o}(C_{u*P_s+w}^K)$$

where W_o is the pooling window size and

P_s is the pooling stride

Residual block:

The activation of this gate is calculated by

$$Fun_A = S_l(wt^{vec}[F_n, blk_{n-1}, mem_{n-1}] + B_u)$$

where F_n is the input feature vector.

blk_{n-1} is the previous block output.

mem_{n-1} is the previous block memory.

wt^{vec}
denotes separate weight vectors of each feature vector.

S_l is logistic sigmoid function.

B_0 is bias vector for the feature map.

The output gate is the section at which the current block output is being generated.

$$mem_n = F_{act} * mem_{n-1} tanh([I_n, blk_{n-1}, mem_{n-1}] + B_c)$$

where B_c is the bias results for convolutional layer.

$$S_l = \sigma(wt^{vec}[F_n, blk_{n-1}, mem_n]) + B_0$$

$$blk_n = S_l . \tanh(mem_n)$$

where blk_n is the current block output and

B_0 is bias vector.

U-Net architecture comprises a constricting growing lower side and upper side. The constricting side consist of two 3×3 convolution layers intended for the iterative model. These two convolutional layers are the modified U-Net layer, and the down sampling layer has a function of 2×2 extreme pooling with two steps. The count of channel attributes becomes twice in every down sampling step. At the growing path, every step comprises of upsampling attributes that are tracked through a 2×2 convolution layer which in turn halves the counts of attributes. Finally, the block of 1×1 convolution is thus employed for mapping the sets of 64 attributes vectors for the class number. Moreover, normalization is added to the input layer at the actual U-Net structure. It was assured that input data dispersion at each layer is stable at the time of input normalization. The entire data is utilized to learn the full convolution through the computation of feed forward and the back propagation process.

Algorithm 2: Deep U-Net algorithm for prediction

Optimize threshold value using gradient descent that

increases Jaccard Index
Step 1 : For each sample data, optimization takes place:
Input sample data = 1800
Convert input sample data to pre-processed data=Pre_data
D_orig = D channel of Pre_datav
Normalize D_orig values between 0 and 1
Employ equalization to D channel of Pre_data
Train_data [channel 0,1,2] = convert pre − processed data into sample data
Train_data [channel 3] = D_orig
Add Train_data to sample set
Step 2: Train U − Net with10000 repetitions
Chosen from the 7200 sample data
Step 3 : Set threshold value as 0.5
Step 4 : For each 1800 sample data
Prediction of targeted region
Employ bin fill
threshold limit is used to
obtain the relevant product
Step 6 : Employ bin fill

Classification is performed using hybrid multivariate improved ResNet and U-Net classifier. As a result of the classification process, vegetative land covers are

identified effectively. The performance investigation of the anticipated system is revealed in the performance analysis section.

IV. PERFORMANCE ANALYSIS

This section is the presentation of comprehensive elucidation of the performance analysis of the projected system. The proposed methodology performance is enumerated by calculating several metrics of performance including accuracy, specificity, precision, recall, F-measure and sensitivity.

A. Accuracy

Accuracy A_i depends on the number of targets that are classified correctly and is evaluated by:

$$A_i = \frac{TP+TN}{TP+TN+FP+FN}, \tag{13}$$

where *TP* is True Positive, *TN* is True Negative, *FP* is False Positive, *FN* is False Negative.

B. Recall sensitivity

Sensitivity measures how many positives are correctly identified as positives and is defined as

$$S_n = \frac{TP}{TP+FN}. \tag{14}$$

C. Specificity

Specificity measures how many non-targets are correctly identified as non-targets and is defined as

$$S_p = \frac{TN}{TN+FP}. \tag{15}$$

D. Precision

Precision is defined as the ratio of the number targets that are classified to the number of targets present in an image

$$Precision = \frac{TP}{TP+FP}. \tag{16}$$

E. F-Measure

F-Measure is obtained by combining precision and recall

$$F-measure = \frac{2*precision*recall}{precision+recall}. \tag{17}$$

The datasets employed in this work are listed below.

F. Dataset description

1. Indian Pines

This is the segmentation of hyperspectral image dataset. The input source data encompasses of hyperspectral bands over a sole scenery in Indiana, USA (Indian Pines) dataset that has pixels of about 145×145. In support of each pixel, data set encompasses of bands with spectral reflectance of around 220 which indicates several portions of electromagnetic spectrum in the range of wavelength $0.4-2.5\cdot10^{-6}$.

2. Salinas

Denoted as the dataset of hyperspectral image which is gathered through 224-band AVIRIS sensor over the Salinas Valley, California, and is characterized over high spatial resolution (about 3.7-meter pixels). The enclosed area encompasses of 512 lines having the sample of around 217, and the water absorption bands of about 20 were being discarded. This image is being accessible as the data of senor radiance and it includes bare soils, vegetables, and vineyards fields. The ground truth Salinas encompasses of 16 classes.

3. Pavia University

The dataset Pavia University is a hyperspectral image type which is accumulated by a labelled sensor as a system of reflective optics imaging spectrometer (ROSIS-3) over Pavia City, Italy. It encompasses of pixels 610×340 by 115 spectral bands. The image is thus divided as nine classes having total labelled samples of about 42,776, with trees, meadows, gravel, bitumen, asphalt, bare soil, brick, metal sheet, and shadow.

G. Performance analysis of suggested system

Fig. 5. Denoised image.

Performance assessment of the anticipated result is revealed in Fig. 5. This shows the denoising process effect on the input image. Likewise, Fig. 6 indicates the segmentation outcome obtained as a result of the effective segmentation process by varying the level of iterations. The range of iteration set are 100, 200, and 500.

The segmented outcome with various iteration values (100, 200 and 500) is shown. Figure 7 is the segmented outcome attained for 100 iterations. Figure 8 is the segmented outcome attained for 200 iterations. Figure 9 is the segmented outcome for 500 iterations.

Classified outcomes are illustrated in Fig. 10 which is assessed for the dataset Salinas. The identified and classified regions from the satellite image Salinas dataset is revealed. Each color signifies one parameter or area

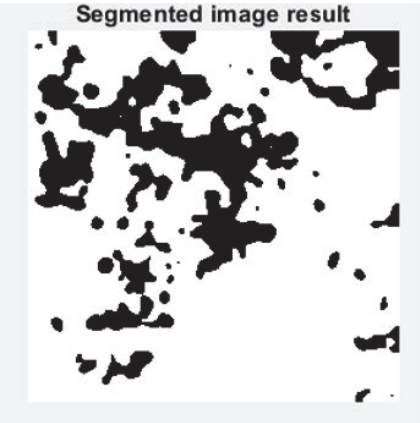


Fig. 6. Segmented outcome.

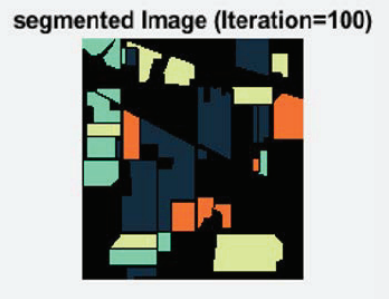


Fig. 7. Segmented image at iteration 100.

Fig. 8. Segmented image at iteration 200.

of vegetative land categorization. The categorization is labelled in Fig. 10 for clarity.

The portrayal of classified consequence for Pavia University dataset is projected in Fig. 11. The recognized and classified regions from the satellite image Pavia University dataset are exposed. Each color indicates one parameter or area of different vegetative land category. The categorization is labelled in Fig. 11 for clarity.

The representation of classified result for Indian Pines dataset is given in Fig. 12. The classified and recognized regions from the satellite image Indian Pines

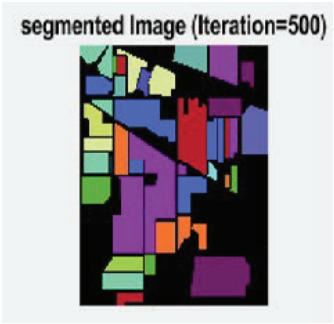


Fig. 9. Segmented image at iteration 500.

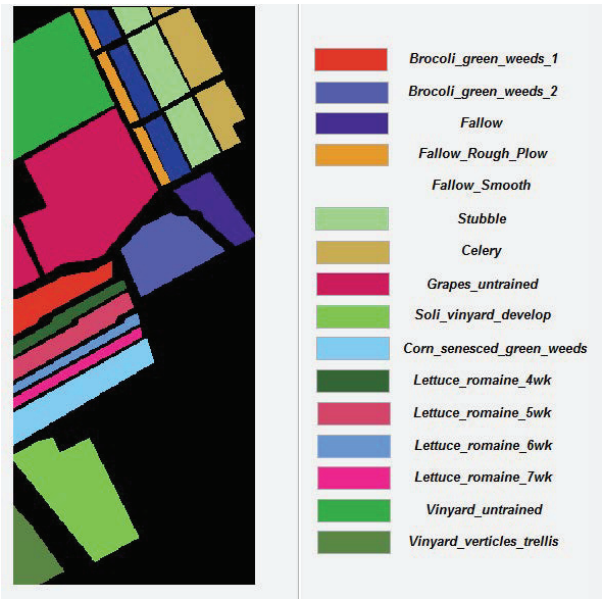


Fig. 10. Classified outcomes for Salinas dataset.

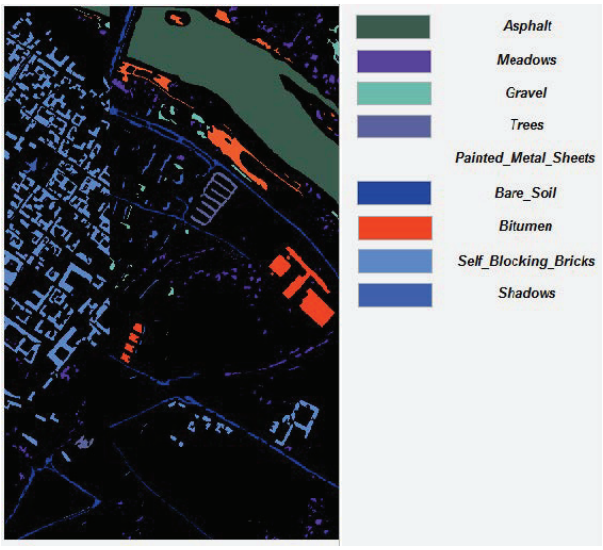


Fig. 11. Classified outcomes for Pavia University dataset.

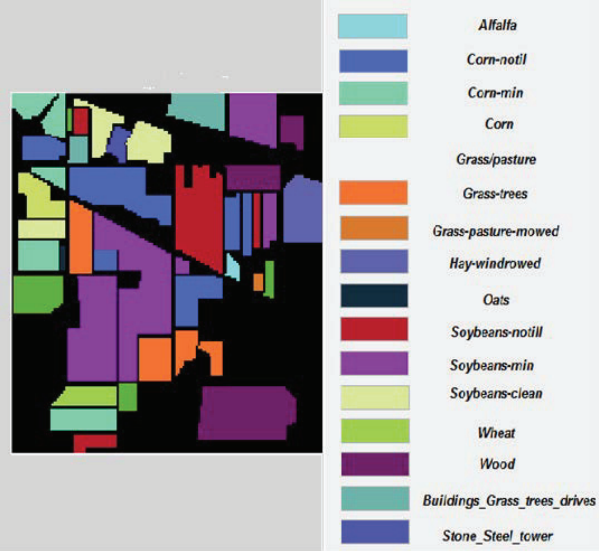


Fig. 12. Classified outcomes for Indian Pines dataset.

dataset are given. Each color implies one parameter or area of vegetative land categorization. The categorization is labelled in Fig. 12 for clarity.

Table 2 illustrates performance investigation of anticipated methodology in terms of PSNR, SSIM, MSE, NCC and IEF.

H. Comparative analysis of proposed system

The classifier function is estimated and the outcomes for each dataset taken are given and the outcome is compared with existing algorithms for exposing the efficiency of proposed system. Table 3 shows the Salinas dataset performance analysis. The epochs are varied and the activation function is increased to estimate the outcomes of proposed system classifier.

Table 4 shows the Pavia University dataset performance analysis. The epochs are varied and the activation function is increased to estimate the outcomes of proposed system classifier.

Table 5 shows the Indian Pines dataset performance analysis. The epochs are varied and the activation function is increased to estimate the outcomes of proposed system classifier.

Table 2: Performance analysis of proposed system error measure

Measure	Proposed Method
Peak signal to noise ratio (PSNR)	46.4019
Mean Square Error (MSE)	1.2202
Structure Similarity Index (SSIM)	0.999
Normalized Cross Correlation (NCC)	1.003
Image Enhancement Factor (IEF)	0.9717

Table 3: Salinas dataset performance analysis

Salinas Dataset				
Epochs	50	150	250	350
Activation Function				
Tanh	98.75	98.16	99.32	99.78
Leaky ReLU	97.12	97.28	97.34	98.91
ReLU	96.27	96.39	96.51	96.36

Table 4: Pavia University dataset performance analysis

Pavia University Dataset				
Epochs	50	150	250	350
Activation Function				
Tanh	96.24	96.75	97.36	97.28
Leaky ReLU	95.87	95.48	95.37	96.71
ReLU	94.25	94.58	94.68	94.12

Table 5: Indian Pines dataset performance analysis

Indian Pines Dataset				
Epochs	50	150	250	350
Activation Function				
Tanh	91.72	91.57	92.38	92.66
Leaky ReLU	90.28	91.24	92.34	93.51
ReLU	91.45	92.35	92.68	93.11

Table 6 shows the comparative analysis of proposed system classifier result. The performance of projected algorithm is related with prevailing performances like Deep CNN, VGG-16, and ResNet-50 to demonstrate the effectiveness of anticipated mechanism hybrid multivariate improved ResNet-50 and U-Net classifier. The classifier's elapsed time is 54.064668 sec, OA (total accuracy) is 72.91%, AA (average accuracy is 89.83%), and kappa value is 0.684.

Table 6: Comparative analysis of proposed classifier outcome

Measures	Deep CNN	VGG-16	ResNet-50	Hybrid Multivariate Improved ResNet-50 and U-Net
Accuracy	84.6579	87.9083	91.1588	96.0345
Sensitivity	5.6000	7.0000	9.3333	18.6667
Specificity	100	100	100	100
Precision	100	100	100	100
Recall	5.6000	7.0000	9.3333	18.6667
F-Measure	0.1061	0.1308	0.1707	0.3146

The performance of projected algorithm is compared with existing techniques like OF (original spectral feature), PCA (principle component analysis), and SF (sparse feature) [27] in Fig. 13 to demonstrate the effectiveness of anticipated mechanism for varied classes in terms of accuracy for Pavia University dataset.

The performance of the projected algorithm is compared with existing techniques like OF, PCA, and SF in Fig. 14 to demonstrate the effectiveness of the anticipated mechanism for varied classes in terms of OA, kappa, and AA for Pavia University dataset.

The performance of the projected algorithm is compared with existing techniques like OF (original spectral feature), PCA (principle component analysis), and SF (sparse feature) in Fig. 15 to demonstrate the effectiveness of the anticipated mechanism for varied classes in terms of accuracy for the Salinas dataset.

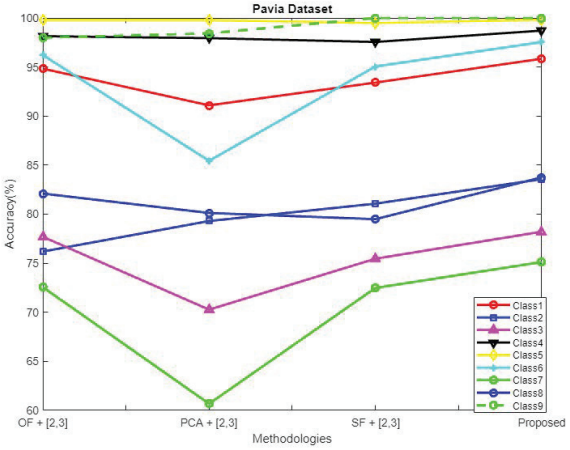


Fig. 13. Comparative analysis of accuracy for Pavia University dataset.

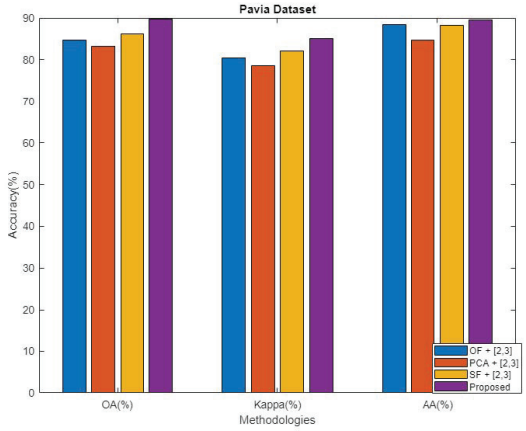


Fig. 14. Comparative analysis for Pavia University dataset in terms of OA, kappa, and AA.

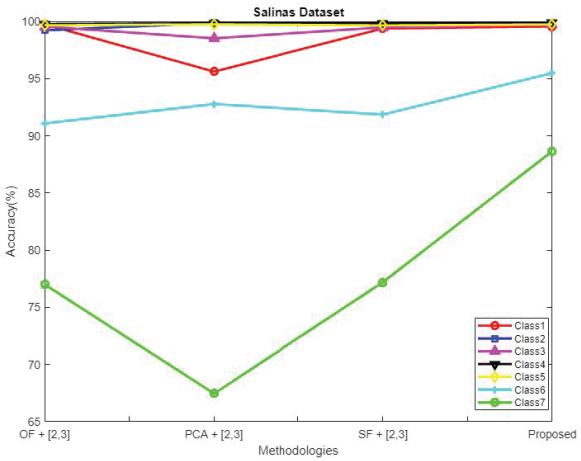


Fig. 15. Comparative analysis of accuracy for Salinas dataset.

The performance of the projected algorithm is compared with existing techniques like OF, PCA, and SF in Fig. 16 to demonstrate the effectiveness of the anticipated mechanism in terms of OA, kappa, and AA for Salinas dataset.

The performance of the projected algorithm is compared with existing techniques [28] like TSVM, LGC, LPA, LPACC, SS-CNN, consistency regularization and pseudo-labelling in Fig. 17 to demonstrate the effectiveness of the anticipated mechanism in terms of OA, kappa, and AA for Indian Pines dataset.

The performance of the projected algorithm is compared with existing techniques like TSVM, LGC, LPA, LPACC, SS-CNN, consistency regularization, and pseudo-labelling in Fig. 18 to demonstrate the effec-

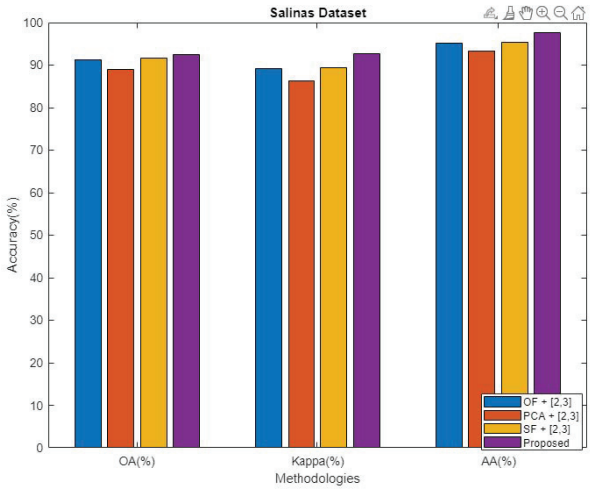


Fig. 16. Comparative analysis for Salinas dataset in terms of OA, kappa, and AA.

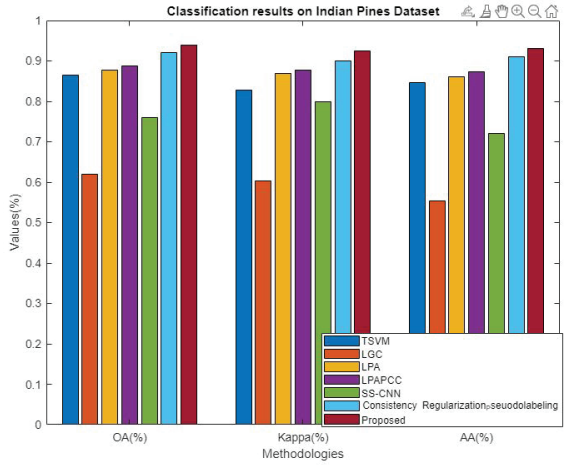


Fig. 17. Comparative analysis for Indian Pines dataset in terms of OA, kappa, and AA.

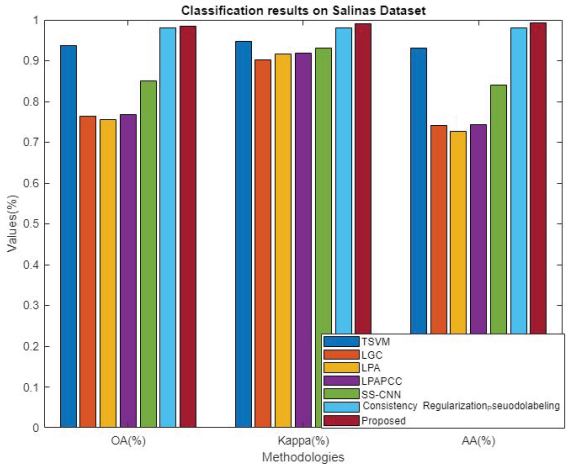


Fig. 18. Comparative analysis for Salinas dataset in terms of OA, kappa, and AA.

tiveness of the anticipated mechanism in terms of OA, kappa, and AA for Salinas dataset.

The performance of the projected algorithm is compared with existing techniques like TSVM, LGC, LPA, LPACC, SS-CNN, consistency regularization, and pseudo-labelling in Fig. 19 to demonstrate the effectiveness of the anticipated mechanism in terms of OA, kappa, and AA for Pavia University dataset.

The accuracy of the existing model is compared with the proposed accuracy in Table 7, which shows enhancement over other compared models [29,30]. The accuracy of the proposed model is attained as 99.8050%.

Table 8 shows the performance comparison of various models on the Indian Pines dataset. The metrics analyzed are OA, AA, and kappa for proposed and existing schemes MS-RPNet and MFFCG-2 [29,30]. The proposed model shows the following results: OA=99.38%, AA=99.06%, and kappa=99.78%, which is higher than existing values.

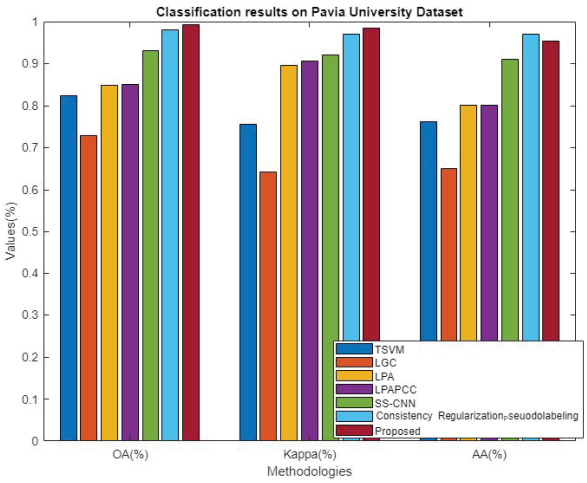


Fig. 19. Comparative analysis for Pavia University dataset in terms of OA, kappa, and AA.

Table 7: Accuracy comparison with existing models

Model	Accuracy
Deep CNN	88.412
VGG-16	90.9150
ResNet-50	92.5727
MS-RPNet	99.6500
MFFCG	98.9400
Proposed	99.8050

Table 9 shows the performance comparison of various models on the Pavia University dataset. The metrics analyzed are OA, AA, and kappa for proposed and existing schemes MS-RPNet and MFFCG-2. The proposed model shows the following results: OA=99.87%, AA=99.86%, and kappa=99.73%, which is higher than existing values.

Table 10 shows the performance comparison of various models on the Salinas dataset. The metrics analyzed are OA, AA, and kappa for proposed and existing scheme MFFCG-2. The proposed model shows the following

Table 8: Performance comparison of different models on Indian Pines dataset

Metrics	MS-RPNet	MFFCG-2	Proposed
OA (%)	98.8700	94.2500	99.3800
AA (%)	NA	93.9900	99.0600
kappa (%)	98.0500	93.2900	99.7800

Table 9: Performance comparison of different models on Pavia University dataset

Metrics	MS-RPNet	MFFCG-2	Proposed
OA (%)	99.7600	90.3400	99.8700
AA (%)	NA	94.1400	98.8600
kappa (%)	99.6800	93.5200	99.7300

Table 10: Performance comparison of different models on Salinas dataset

Metric	MFFCG-2	Proposed
OA (%)	87.1700	98.6700
AA (%)	86.4900	99.0500
kappa (%)	85.1600	99.7100

results: OA=98.67%, AA=99.05%, and kappa=99.71%, which is higher than existing values. From analysis, it was obvious that the proposed model shows enhancement over other existing models.

V. CONCLUSION

In this approach, a deep learning-based automated technique for the detection, segmentation, and classification of satellite images was carried out. Initially, the input image was preprocessed, segmented using Edge-ROI and YOLO v3 model at which parameters are tuned by means of multi-heuristic tuna swarm optimization (MH-TSO), and was classified using hybrid multivariate improved residual network and U-Net classifier approach. The analysis was compared with traditional methods to validate the effectiveness of the proposed scheme. The proposed technique performance is estimated for three kinds of dataset: Salinas, Pavia University, and Indian Pines satellite image datasets. The analysis shows that the proposed system offers better rate of recall, precision, and F-measure than existing techniques. The outcomes show that the designed method overcomes the issues of existing satellite image classification methods through classifying the satellite image in an effective manner. In future, this work can be extended by integrating attention mechanisms or transformer based models to further improve feature representation, and also by applying this framework to other domains like urban planning, agricultural monitoring, and disaster assessment for evaluating their broad applicability.

REFERENCES

[1] X. Zhang and J. Xiang, "Moving object detection in video satellite image based on deep learning," in *Proc. SPIE LIDAR Imaging Detection and Target Recognition*, Xi'an, China, vol. 10605, p. 106054H, 2017.

[2] F. A. Mianji, Y. Gu, Y. Zhang, and J. Zhang, "Enhanced self-training superresolution mapping technique for hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 4, pp. 671-675, 2011.

[3] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. J. I. A. Abbas, "Modern trends in hyperspectral image analysis: A review," *IEEE Access,* vol. 6, pp. 14118-14129, 2018.

[4] J. Chen and A. Zipf, "Deep learning with satellite images and volunteered geographic information," in *Geospatial Data Science Techniques and Applications*, J. Chen and A. Zipf, Eds. Boca Raton, FL: CRC Press, p. 63, 2017.

[5] C. Avolio, M. Á. M. Armenta, A. J. Lucena, M. J. F. Suarez, P. V. Martin-Mateo, and F. L. González, "Automatic recognition of targets on very high resolution SAR images," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 2271-2274, 2017.

[6] R. J. Soldin, "SAR target recognition with deep learning," in *2018 IEEE Applied Imagery Pattern Recognition Workshop (AIPR),* Washington, DC, USA, pp. 1-8, 2018.

[7] Y. E. García-Vera, A. Polochè-Arango, C. A. Mendivelso-Fajardo, and F. J. Gutiérrez-Bernal, "Hyperspectral image analysis and machine learning techniques for crop disease detection and identification: A review," *Sustainability*, vol. 16, no. 14, p. 6064, 2024.

[8] P. Ilamathi and S. Chidambaram, "Integration of hyperspectral imaging and deep learning for sustainable mangrove management and sustainable development goals assessment," *Wetlands*, vol. 45, no. 9, 2025.

[9] N. Dahiya, G. Singh, S. Singh, and V. Sood, "Crop land assessment with deep neural network using hyperspectral satellite dataset," in *Hyperautomation in Precision Agriculture*. Cambridge, MA, USA: Academic Press, pp. 159-167, 2025.

[10] P. Sevugan, V. Rudhrakoti, T. H. Kim, M. Gunasekaran, S. Purushotham, and R. Chinthaginjala, "Class-aware feature attention-based semantic segmentation on hyperspectral images," *PLoS One*, vol. 20, no. 2, p. e0309997, 2025.

[11] R. Luo, P. Lu, P. Chen, H. Wang, X. Zhang, and S. Yang, "Hyperspectral classification of ancient cultural remains using machine learning," *Remote Sensing Applications: Society and Environment*, vol. 37, p. 101457, 2025.

[12] Q. Zhe, W. Gao, C. Zhang, G. Du, Y. Li, and D. Chen, "A hyperspectral classification method based on deep learning and dimension reduction for

ground environmental monitoring," *IEEE Access*, 2025.

[13] S. Chen, H. Wang, F. Xu, and Y.-Q. Jin, "Target classification using the deep convolutional networks for SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4806-4817, 2016.

[14] Z. Lin, K. Ji, M. Kang, X. Leng, and H. Zou, "Deep convolutional highway unit network for SAR target classification with limited labeled training data," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 7, pp. 1091-1095, 2017.

[15] J. H. Cho and C. G. Park, "Multiple feature aggregation using convolutional neural networks for SAR image-based automatic target recognition," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 12, pp. 1882-1886, 2018.

[16] I. Fedorchenko, A. Oliinyk, A. Stepanenko, T. Zaiko, S. Shylo, and A. Svyrydenko, "Development of the modified methods to train a neural network to solve the task on recognition of road users," *Eastern-European Journal of Enterprise Technologies*, vol. 2, no. 9, pp. 46-55, 2019.

[17] H. He, S. Wang, D. Yang, and S. Wang, "SAR target recognition and unsupervised detection based on convolutional neural network," in *Proc. 2017 Chinese Automation Congress (CAC)*, pp. 435-438, 2017.

[18] I. M. Gorovyi and D. S. Sharapov, "Comparative analysis of convolutional neural networks and support vector machines for automatic target recognition," in *Proc. 2017 IEEE Microwaves, Radar and Remote Sensing Symposium (MRRS)*, pp. 63-66, 2017.

[19] Y. Kwak, W.-J. Song, and S.-E. Kim, "Speckle-noise-invariant convolutional neural network for SAR target recognition," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 4, pp. 549-553, 2018.

[20] B. Sehgal, H. S. Shekhawat, and S. K. Jana, "Automatic target recognition using recurrent neural networks," in *Proc. 2019 Int. Conf. on Range Technology (ICORT)*, Chandipur, India, pp. 1-5, 2019.

[21] H. Shen, Y. Lin, Q. Tian, K. Xu, and J. Jiao, "A comparison of multiple classifier combinations using different voting-weights for remote sensing image classification," *International Journal of Remote Sensing*, vol. 39, no. 11, pp. 3705-3722, 2018.

[22] D. Ye, Y. Li, C. Tao, X. Xie, and X. Wang, "Multiple feature hashing learning for large-scale remote sensing image retrieval," *ISPRS International Journal of Geo-Information*, vol. 6, no. 11, p. 364, 2017.

[23] M. Häberle, M. Werner, and X. X. Zhu, "Geospatial text-mining from Twitter: A feature space analysis with a view toward building classification in urban regions," *European Journal of Remote Sensing*, vol. 52, no. 2, pp. 2-11, 2019.

[24] M. Al Mufti, E. Al Hadhrami, B. Taha, and N. Werghi, "SAR automatic target recognition using transfer learning approach," in *Proc. 2018 Int. Conf. on Intelligent Autonomous Systems (ICoIAS)*, Singapore, pp. 1-4, 2018.

[25] M. Al Mufti, E. Al Hadhrami, B. Taha, and N. Werghi, "Using transfer learning technique for SAR automatic target recognition," in *Proc. SPIE Future Sensing Technologies*, Tokyo, Japan, vol. 11197, p. 111970A, 2019.

[26] A. Khan, A. Sohail, U. Zahoora, and A. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial Intelligence Review*, vol. 53, pp. 5455-5516, 2020.

[27] W. NanLan and Z. Xiaoyong, "Hyperspectral data classification algorithm considering spatial texture features," *Mathematical Problems in Engineering*, vol. 2022, p. 901234, 2022.

[28] X. Hu, T. Zhou, and Y. Peng, "Semisupervised deep learning using consistency regularization and pseudolabels for hyperspectral image classification," *Journal of Applied Remote Sensing*, vol. 16, no. 2, p. 026513, 2022.

[29] H. Chen, T. Wang, T. Chen, and W. Deng, "Hyperspectral image classification based on fusing S3-PCA, 2D-SSA and random patch network," *Remote Sensing*, vol. 15, no. 13, p. 3402, 2023.

[30] U. A. Bhatti, M. Huang, H. Neira-Molina, S. Marjan, M. Baryalai, and H. Tang, "MFFCG: Multi feature fusion for hyperspectral image classification using graph attention network," *Expert Systems with Applications*, vol. 229, p. 120496, 2023.

P. Pabitha Muthu is currently serving as an Assistant Professor in the Department of Computer Science and Engineering at Sethu Institute of Technology, located in Kariapatti, Virudhunagar District 626115, Tamil Nadu, India. She earned her undergraduate degree in Computer Science and Engineering from Pandian Saraswathi Yadav Engineering College. She completed her postgraduate studies in the same field at the same institution in 2014. Her research is focused on satellite image processing and wireless sensor networks.

S. Siva Ranjani is a Professor at Sethu Institute of Technology, Pulloor, Kariapatti Taluk, Virudhunagar District, India. She completed her post-graduation in Computer Science and Engineering from Kalasalingam University in 2007. She was awarded a Ph.D. in the same discipline in 2015, with her research focusing on Wireless Sensor Networks. Her areas of interest include data mining, wireless sensor networks, remote sensing, and image processing.

M. Manikandakumar is currently an Assistant Professor in the department of AIML and Data Science, Christ University, Kengeri Campus, Bangalore, India. He completed his post-graduation in Computer Science and Engineering in 2013 from Anna University, Chennai, where he completed the Ph.D in Information and Communication Engineering in 2023. He is currently an Assistant Professor in the department of AIML and Data Science, Christ University, Kengeri Campus, Bangalore. His research interests include data science, wireless networking, and machine learning, image processing and computer vision.

Enhanced Deep Learning Approach for Multi-parameter Hollow Shaped Cylindrical Dielectric Resonator Antenna Design

Fidan Gamze Kizilcay[1,2] and Muhammet Hilmi Nisanci[1]

[1]Department of Electrical Electronics Engineering
Sakarya University, Sakarya, Türkiye
fidan.kizilcay@ogr.sakarya.edu.tr, nisanci@sakarya.edu.tr

[2]Department of Information Technology
Zonguldak Bülent Ecevit University, Zonguldak, Türkiye
gamze.kizilcay@beun.edu.tr

Abstract – The design of antennas for specific purposes often results in significant time costs due to the lengthy simulation processes required. Adopting deep learning-based approaches in antenna design can offer more efficient solutions. In this study, deep learning methods were applied to accurately and efficiently predict the resonant frequency value of the hollow shaped cylindrical dielectric antenna. For this purpose, a total of 1000 simulations were performed for the considered antenna, and corresponding operational frequencies in 6-12 GHz frequency band were obtained. The data was diversified to search for an optimal solution. A total of 800 simulation results were employed for training, and a series of operations were performed to develop the training model. As a result of these improvements the mean squared error (MSE) was observed to decrease to 0.128. In order to evaluate the performance of the model, the output was obtained by using randomly assigned input parameters. This revealed a difference of 0.49% between the actual result and the model output, which indicates improved prediction accuracy and reliability of the model.

Index Terms – Antenna design, deep learning, dielectric resonator antenna, hollow shaped antenna.

I. INTRODUCTION

Dielectric resonator antennas (DRAs) are predominantly designed for operation at microwave and millimeter-wave frequencies, particularly in applications such as 5G, WiMax, WLAN, and radar [1–3]. The performance of antennas utilized in these technologies is of paramount importance, especially concerning signal accuracy and efficiency in high-frequency bands. The frequency range of 6-12 GHz is a notable bandwidth for applications that demand high resolution and low latency. DRAs operating within this frequency range are distinguished from conventional metal antennas by their low loss, wide bandwidth, high efficiency, and compact structures [4, 5].

Achieving the desired antenna performance during the design phase is challenging. Many software tools provide features for optimizing antenna parameters. Deep learning algorithms offer the capability to perform multi-parameter optimization across a wide range, which not only enhances antenna performance but also accelerates physical prototyping processes. The integration of deep learning and machine learning techniques into antenna design has thus opened up a innovative field of research focused on improving antenna performance.

In [6], Ranjan et al. highlight the advantages offered by machine learning-assisted algorithms in optimizing hybrid DRA designs. This research encompasses improvements in antenna performance achieved through optimizing combinations of dielectric materials and geometric structures. Kushwaha et al. [7] focus on the optimization of cylindrical DRAs, addressing the contributions of machine learning to specific DRA geometries. This approach enables the optimization of critical performance indicators for DRA in the X-band, such as gain, radiation pattern, and bandwidth. The study by Pachori et al. [8] emphasizes the contributions of machine learning-based modeling in predicting the performance of multiple-input multiple-output (MIMO) antennas. This research demonstrates the significant potential of machine learning algorithms in accelerating the optimization process by enabling rapid and accurate predictions of antenna performance. In [9], Fu and Leung examine how evolutionary algorithms, when integrated with machine learning, can enhance efficiency in antenna design. These hybrid methods not only enable high-accuracy optimization of antenna parameters but also reduce computational costs throughout the process. This study illustrates how machine learning can work synergistically with evolutionary algorithms to meet the

Submitted On: December 13, 2024
Accepted On: August 22, 2025

https://doi.org/10.13052/2025.ACES.J.400802

complex electromagnetic requirements of DRAs in the X-band, showcasing the multifaceted benefits these two fields bring to antenna design.

Numerous studies apply deep learning techniques to datasets of antenna and filter designs to produce fast and reliable results [10–17]. Depending on the processing speed of the computer and the size of the data set, initial calculations may take time; however, once the desired data is obtained, this method remains both relevant and efficient.

Although the main focus of the study is to investigate the role of deep learning in antenna design, another objective is to design an antenna with high gain and wide impedance bandwidth operating in the 6-12 GHz frequency region, which distinguishes it from conventional cylindrical DRAs. Various methods have been applied to improve impedance bandwidth in DRAs. These methods generally enhance impedance bandwidth by altering the permittivity distribution within the dielectric resonator [18–21]. In a related study aimed at improving impedance bandwidth [22], cavities of different diameters were introduced into a cylindrical DRA structure to modify the effective dielectric constant, achieving an improved impedance.

Overall, in contrast to traditional antenna designs, the physical parameters of the dielectric used in DRAs are typically analyzed through machine learning techniques [6–8, 23]. What distinguishes this study from similar ones is the involvement of more parameters in the application of machine learning methods in DRAs, as well as the incorporation of deep learning techniques, thereby introducing a novel approach.

Deep learning, as a subset of machine learning, employs multi-layered artificial neural networks (ANN) that automatically learn complex patterns and extract meaningful features from large and unstructured datasets. Unlike traditional machine learning, which relies on manual feature selection and works best with structured data, deep learning can identify complex structures directly from raw data, making it especially effective for difficult and high-dimensional problems [24, 25]. Building on these advantages, the deep learning models used in this study were further improved, resulting in enhanced predictive performance.

In the current study, the Latin hypercube sampling (LHS) algorithm was applied to accurately and efficiently predict the resonant frequency value of a hollow shaped cylindrical dielectric antenna. In order to increase the impedance bandwidth, two gaps were created within the cylindrical dielectric. These gaps were designed in such a way that they do not disrupt symmetry. The LHS algorithm was used to determine the design parameter values for the antenna to operate at the desired frequency.

The main purpose of LHS is to provide sampling points that represent a wide range of the design space. This comprehensive sampling offers better diversity, allowing for statistically reliable results to be obtained [26, 27]. To enhance prediction accuracy, a multi-layer ANN was developed and optimized using the adaptive moment estimation (Adam) optimization algorithm, widely recognized for its effectiveness in training deep learning models [28, 29]. In addition, advanced techniques such as early stopping and adaptive learning rate adjustments were employed to refine the deep learning model.

This paper is organized as follows. Section 2 explains the proposed cylindrical DRA design specifications as well as the dataset used in this study. Section 2 also describes the methodology for predicting the resonant frequency value of the DRA, the multi-layer ANN and deep learning models, and the optimization algorithm. In section 3, the obtained results are discussed. Conclusions and perspectives for future works are presented in section 4.

II. MATERIALS AND METHOD

A series of systematic steps are followed to accurately and efficiently predict the resonant frequency value of the hollow shaped cylindrical DRA. This process is illustrated in Fig. 1 which shows the five main steps the proposed design procedure includes: antenna design, LHS, algorithms and data, model improvement, and predicted results.

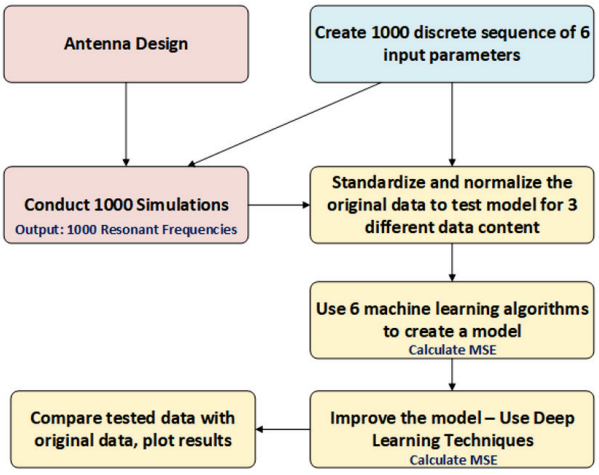


Fig. 1. Design procedure flowchart.

A. Antenna design

Figure 2 shows the proposed hollow shaped cylindrical DRA components and the design parameters. As shown in Fig. 2, the proposed antenna has eight physical and two electrical design parameters: r_1, r_2, r_3, r_4, r_5, h_1,

h_2, h_3, ε_r and loss tg. In order to numerically obtain the data set, a commercial full-wave electromagnetic simulation program Computer Simulation Technology (CST) was used [30] in the simulations. The designed antenna was fed using a 50Ω waveguide port and a microstrip-slot coupling mechanism. The parameters utilized in creating the dataset, moving outward from the center, are: r_1, r_2, and the height from the ground of the solid object between these two values, h_1; r_3, r_4, and the height from the ground of the solid object between these two values, h_2. The other parameters of the antenna were kept constant. The dielectric substrate height is 1.52 mm, while the outer radius and height of the dielectric cylindrical antenna are $r_5 = 6$ mm and $h_3 = 4$ mm, respectively. Since changes in slot size have minimal impact on the results, they were not considered when generating the dataset.

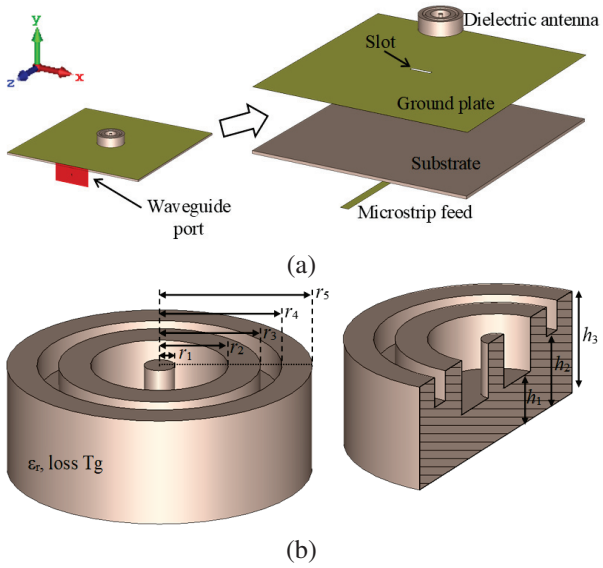


Fig. 2. (a) Antenna design and (b) antenna scheme.

B. Latin hypercube sampling

LHS is a statistical technique used for sampling a multi-variable space. Unlike random sampling methods, LHS generates samples where each variable is evenly distributed across its value range. LHS divides the sample space into a specified number of cells and takes only one sample from each cell. This ensures that different values are selected for each variable, allowing the entire value range to be represented simultaneously for all variables.

A dataset was prepared for the six parameters in the design (r_1, r_2, r_3, r_4, h_1, and h_2), and simulations were performed. As shown in Table 1, the boundaries for each parameter were set as ($r_1 \in [0, 1.50]$, $r_2 \in [1.50, 3.00]$, $r_3 \in [3.50, 4.25]$, $r_4 \in [4.25, 5.00]$, $h_1 \in [0, 3.00]$, $h_2 \in [0, 3.00]$). Subsequently, random values were assigned to each parameter within its limits, generating 1000 different inputs.

Table 1: LHS dataset

#	r_1	r_2	r_3	r_4	h_1	h_2
1	0.366	2.055	3.826	4.331	2.411	2.814
2	1.110	2.855	3.991	4.764	2.350	2.159
3	1.320	2.280	3.825	4.582	0.875	0.587
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1000	0.027	1.550	4.016	4.318	0.929	2.234

C. Algorithms and data

In this study, the resonant frequency value was predicted using different machine learning methods with the values of r_1, r_2, r_3, r_4, h_1, and h_2. First, to assess the predictive success of traditional machine learning methods, the following machine learning algorithms were tested:

- RandomForestRegressor
- GradientBoostingRegressor
- HistGradientBoostingRegressor
- ExtraTreesRegressor
- AdaBoostRegressor
- BaggingRegressor

For each model, hyperparameter optimization was conducted using the Random Search method to obtain the best-performing model for the respective machine learning algorithm. Random search efficiently tunes hyperparameters by randomly sampling the search space, often outperforming similar strategies in many machine learning models [31, 32]. RandomizedSearchCV, an implementation of the random search strategy combined with cross-validation, was employed to optimize the hyperparameters of our machine learning models. Three different scaling methods were applied to each model: the original data, as well as two preprocessed versions (normalized and standardized), to test whether data preprocessing affected model performance.

The performance of each model was obtained using mean squared error (MSE) on the test data. In equation (1), n is the number of tested data, y_i is the predicted resonant frequency by the model and y_i' is the resonant frequency of simulation for the same sequence of six input parameters:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i'\right)^2. \quad (1)$$

MSE, which calculates the average of squared errors, serves as a metric to evaluate the predictive accuracy of the model. The lower the MSE, the better the model's performance is considered to be.

D. Model improvement

To achieve better prediction performance a multi-layer ANN model was created and tested. For this purpose, a model that utilizes the dense layer concept was preferred. A dense layer is a layer in a neural network where every neuron is connected to all the neurons in the previous layer.

Based on the model's output performance, the number of layers and neurons in each layer were optimized. As a result, the proposed ANN consists of nine hidden dense layers, in addition to an input layer and output layer. The number of neurons in the dense layers is set to 64, 128, 128, 64, 64, 32, 32, 16, and 8, respectively, from the first dense layer to the last. Considering factors such as the bimodal, positive, and linear nature of the data, the activation function preferred in this case is the Rectified Linear Unit (ReLU) for all layers. The ANN structure is illustrated in Fig. 3.

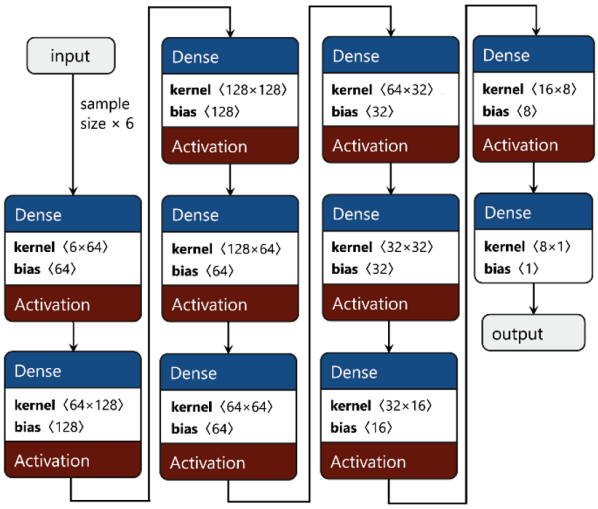


Fig. 3. ANN structure of the model.

Additionally, to prevent overfitting during the training phase, features such as early stopping, automatic learning rate adjustment (ReduceLROnPlateau), and saving the best-performing model (ModelCheckpoint) were used in an effort to obtain the model with the best performance.

In the deep learning model, the MSE, which was used in traditional machine learning methods during testing, was preferred. The optimization algorithm chosen is the Adam optimization algorithm.

III. RESULTS AND DISCUSSION

The bar graph in Fig. 4 illustrates the operating frequency values obtained from simulations conducted within the 6-12 GHz frequency range.

Since the designed antenna is expected to operate actively within the 6-12 GHz range, it is observed that

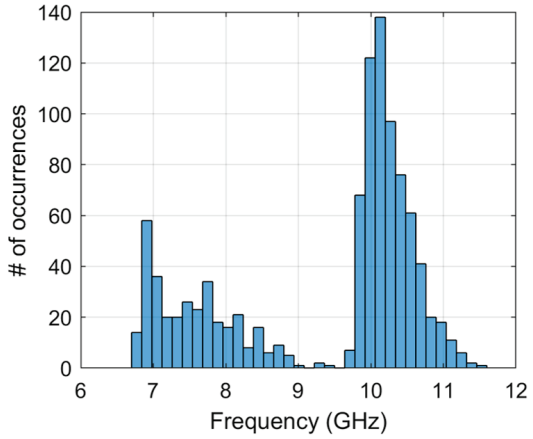


Fig. 4. Obtained operating frequency values.

the results concentrate in the 7-8.5 GHz and 9.5-11 GHz region. However, a small dense region between these is also shown in the graph, which does not align with the general data distribution. This minor density makes accurate predicting challenging. Subsequently, 800 of the data were used for training and 200 for testing.

The original dataset was expanded through the processes of normalization and standardization. This dataset was used distinctly as input for both machine learning and deep learning models. As the output, a single operating frequency value was obtained, allowing for a performance comparison between the machine learning and deep learning models.

Accordingly, initially, all the data was fed into machine learning methods, and MSE was calculated (Table 2).

Table 2: Performance of traditional machine learning methods

#	Model	MSE	Data
1	ExtraTreesRegressor	0.272	Normalized
2	GradientBoostingRegressor	0.276	Normalized
3	RandomForestRegressor	0.299	Normalized
4	HistGradientBoostingRegressor	0.328	Normalized
5	ExtraTreesRegressor	0.333	StandardScale
6	BaggingRegressor	0.340	Normalized
7	ExtraTreesRegressor	0.350	Normal
8	HistGradientBoostingRegressor	0.403	Normal
9	HistGradientBoostingRegressor	0.403	StandardScale
10	RandomForestRegressor	0.408	Normal
11	GradientBoostingRegressor	0.412	StandardScale
12	GradientBoostingRegressor	0.416	Normal
13	RandomForestRegressor	0.422	StandardScale
14	BaggingRegressor	0.429	Normal
15	BaggingRegressor	0.438	StandardScale
16	AdaBoostRegressor	0.617	Normalized
17	AdaBoostRegressor	0.710	StandardScale
18	AdaBoostRegressor	0.714	Normal

Table 2 indicates that ExtraTreesRegressor achieved the best prediction performance, particularly with normalized data, which resulted in an MSE of 0.272, followed closely by GradientBoostingRegressor. In general, normalization improved model performance, while standardization and raw data led to higher errors. AdaBoostRegressor performed the worst, with MSE values significantly higher across all preprocessing methods. These findings suggest that tree-based ensemble methods, especially ExtraTreesRegressor, are well-suited for this dataset, with normalization playing a key role in optimizing predictive accuracy.

In the next step, multi-layered ANNs were used, and the number of neurons in these layers was optimized. Then, early stopping and saving the best-performing methods were applied. Finally, the improving process was completed using the Adam optimization. With the deep learning model, a value of 0.128 was achieved for MSE. Thus, a performance improvement of 53% was obtained compared to the lowest MSE value obtained with traditional methods, which was 0.272. Figure 5 presents a comparison of actual simulation results with predicted values.

Additionally, the MSE distribution of the entire test set, which contains 200 data points, was calculated. As seen in Fig. 6, the best predictions fall particularly within the 8-8.5 GHz and 10-11 GHz ranges, where sufficient training examples were available, allowing the system to learn effectively. Poor predictions were obtained below 7 GHz, which is the band edge. The sharp peak at 7 GHz, along with the limited number of training examples below 7 GHz, have led to uncertainty, causing the model to generate incorrect interpolations. In contrast, better predictions were observed in the 9-10 GHz range and at the upper band edge above 11 GHz. This can be explained by the fact that most of the test data is concentrated around 10 GHz, providing the model with more examples to learn from. As a result, the model was able to capture the underlying patterns more effectively in this region, leading to improved predictions.

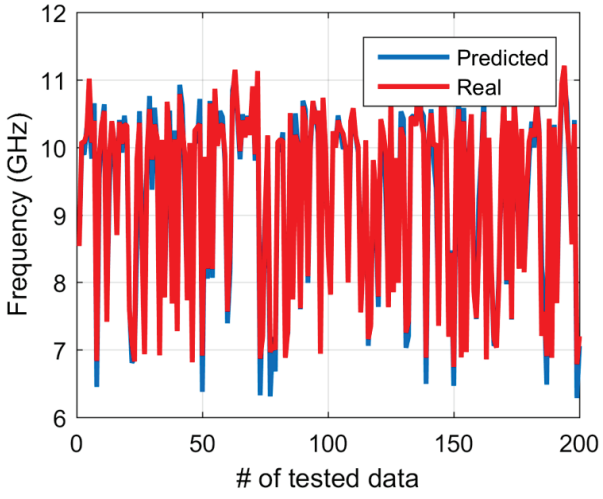


Fig. 5. Predicted and real values.

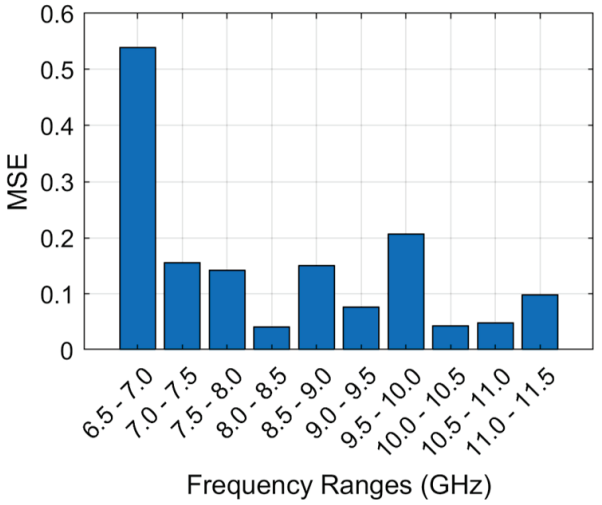


Fig. 6. MSE of the test set.

In order to show the prediction performance of the resonant frequency value of the proposed deep learning methods for the hollow shaped cylindrical dielectric antenna, a numerical analysis was performed with random input parameters by using CST simulation software.

As shown in Fig. 7, DRA operates at 10.25 GHz with a bandwidth of 1.14 GHz and return loss of 37.96 dB. When the same input parameters were fed into the

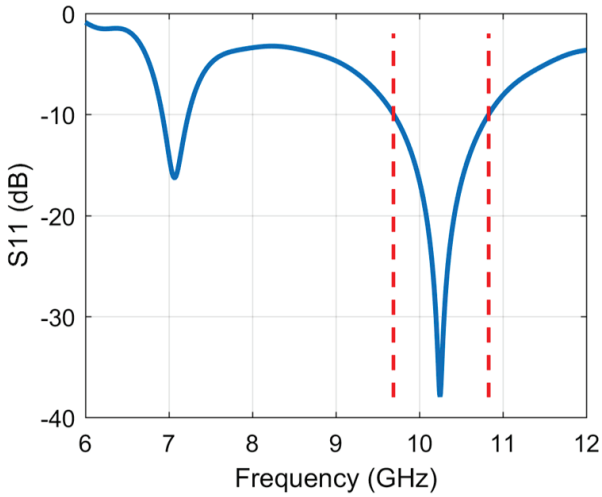


Fig. 7. Simulation results for design parameter values r_1 = 1.5 mm, r_2 = 2.5 mm, r_3 = 3.5 mm, r_4 = 4.75 mm, r_5 = 6 mm, h_1 = 2.5 mm, h_2 = 2.5 mm, and h_3 = 4 mm.

proposed deep learning model, the obtained operating frequency was 10.30 GHz. It was observed that the two results closely match.

On a computer with 256 GB of RAM and an Intel Xeon Gold 6426Y processor and Nvidia RTX A5000 graphics card, 1000 simulation processes for this design were completed in approximately 16 hours, with each simulation taking around 57 seconds. The entire training and testing of the machine learning was completed within 150 seconds. Running deep learning processes took an additional 15 seconds. This process resulted in an improved MSE value when compared to the machine learning approach. To achieve better predictions, the deep learning process was rerun, which took 2 hours to complete. Once the training model was established, the prediction results were generated instantaneously. While traditional machine learning methods demonstrated solid performance, the deep learning model achieved a remarkable improvement, showcasing its superior predictive capabilities. Deep learning offers a promising approach for achieving even greater predictive performance.

IV. CONCLUSION

In this study, a novel approach was implemented to predict the resonant frequency value of the hollow shaped cylindrical dielectric antenna operating in the 6-12 GHz frequency band using a combination of traditional machine learning and deep learning methods. Through the use of Latin hypercube sampling (LHS), a diverse dataset of 1000 simulations were generated. Various preprocessing techniques, including normalization and standardization, were applied to enhance data representation, and multiple machine learning models were evaluated for performance. Among traditional methods, the ExtraTressRegressor demonstrated the best performance, achieving an MSE of 0.272.

To further improve prediction accuracy, a multilayer ANN was developed, optimized using the Adam optimization algorithm. Advanced techniques such as early stopping and adaptive learning rate adjustments were employed to refine the deep learning model. This approach achieved an MSE of 0.128, representing a 53% improvement over the best-performing traditional method.

In order to show the performance of the improved model, a simulation was performed using random input parameters, and the simulation result was compared with the model output. The operating frequency obtained from the simulation was 10.25 GHz, while the model output was 10.30 GHz. Thus, the difference between the two results remained at a level of 0.49%. The consistency between the predicted and simulation results underscores the efficacy of the deep learning model.

The findings highlight the potential of integrating machine learning and deep learning techniques to enhance predictive accuracy in antenna design. This methodology not only accelerates the design process but also improves reliability, paving the way for further advancements in the field of electromagnetic simulation and optimization. Future work could focus on expanding the dataset and exploring alternative optimization techniques to further enhance the prediction accuracy.

ACKNOWLEDGMENT

The authors thank Abdullah Oguz Kizilcay from Zonguldak Bulent Ecevit University for his significant contributions to the field of deep learning, which have provided valuable insights for this study.

REFERENCES

[1] B. Mukherjee, P. Patel, and J. Mukherjee, "A review of the recent advances in dielectric resonator antennas," *Journal of Electromagnetic Waves and Applications*, vol. 34, no. 9, pp. 1095-1158, June 2020.

[2] N. K. Mishra, J. Acharjee, V. Sharma, C. Tamrakar, and L. Dewangan, "Mutual coupling reduction between the cylindrical dielectric resonator antenna using split ring resonator-based structure," *AEU - Int. J. Electron. Commun.*, vol. 154, p. 154305, Sep. 2022.

[3] A. Vahora and K. Pandya, "A miniaturized cylindrical dielectric resonator antenna array development for GPS / Wi-Fi / wireless LAN applications," *Electron. Energy*, vol. 2, p. 100044, 2022.

[4] A. Petosa, *Dielectric Resonator Antenna Handbook*. Norwood, MA: Artech House, 2007.

[5] S. Zheng, X. Chen, N. Yang, B. Qian, L. Zhao, and A. A. Kishk, "Broadband dual-polarized cup dielectric resonator antenna with stable omnidirectional radiation patterns," *Appl. Comput. Electromagn. Soc.*, vol. 39, no. 10, pp. 908-915, 2024.

[6] P. Ranjan, H. Gupta, S. Yadav, and A. Sharma, "Machine learning assisted optimization and its application to hybrid dielectric resonator antenna design," *Facta Univ. - Ser. Electron. Energ.*, vol. 36, no. 1, pp. 31-42, 2023.

[7] A. K. Kushwaha, V. Rai, G. Kumar, V. Kumar, A. Pandey, and R. K. Barik, "Cylindrical dielectric resonator antenna optimization: A machine learning perspective," in *2023 Int. Conf. Comput. Electron. Electr. Eng. their Appl. IC2E3 2023*, pp. 1-6, 2023.

[8] K. Pachori, "Performance prediction of dielectric resonator-based MIMO antenna for sub-6.0 GHz using machine learning algorithms," *Electromagnetics*, vol. 43, no. 8, pp. 539-550, 2023.

[9] K. Fu and K. W. Leung, "A machine learning enhanced evolutionary algorithm for antenna

design," in *IEEE Conf. Antenna Meas. Appl. CAMA*, pp. 50-53, 2023.

[10] B. Şenel and F. A. Şenel, "Bandpass filter design using deep neural network and differential evolution algorithm," *Arab. J. Sci. Eng.*, vol. 47, pp. 14343-14354, 2022.

[11] N. Calik, M. A. Belen, and P. Mahouti, "Deep learning base modified MLP model for precise scattering parameter prediction of capacitive feed antenna," *Int. J. Numer. Model. Electron. Networks, Devices Fields*, vol. 33, no. 2, Mar. 2020.

[12] Y. Kim, "Application of machine learning to antenna design and radar signal processing: A review," in *2018 International Symposium on Antennas and Propagation (ISAP)*, Busan, Korea (South), pp. 1-2, 2018.

[13] N. Wang, Z. Kong, X. Ren, S. Chen, G. Dai, and K. Han, "A wideband butterfly antenna based on deep learning parameter optimization algorithm," in *2020 Cross Strait Radio Science & Wireless Technology Conference (CSRSWTC)*, Fuzhou, China, pp. 1-3, Dec. 2020.

[14] M. M. Khan, S. Hossain, P. Mozumdar, S. Akter, and R. H. Ashique, "A review on machine learning and deep learning for various antenna design applications," *Heliyon*, vol. 8, no. 4, p. e09317, Apr. 2022.

[15] A. M. Montaser and K. R. Mahmoud, "Deep learning based antenna design and beam-steering capabilities for millimeter-wave applications," *IEEE Access*, vol. 9, pp. 145583-145591, 2021.

[16] F. Mir, L. Kouhalvandi, L. Matekovits, and E. O. Gunes, "Automated optimization for broadband flat-gain antenna designs with artificial neural network," *IET Microwaves, Antennas Propag.*, vol. 15, no. 12, pp. 1537-1544, Oct. 2021.

[17] T. Sallam and A. M. Attiya, "Convolutional neural network for coupling matrix extraction of microwave filters," *Appl. Comput. Electromagn. Soc.*, vol. 37, no. 7, pp. 805-810, 2022.

[18] R. Chair, S. L. S. Yang, A. A. Kishk, K. F. Lee, and K. M. Luk, "Aperture fed wideband circularly polarized rectangular stair shaped dielectric resonator antenna," *IEEE Trans. Antennas Propag.*, vol. 54, no. 4, pp. 1350-1352, 2006.

[19] W. Huang and A. A. Kishk, "Compact dielectric resonator antenna for microwave breast cancer detection," *IET Microwaves, Antennas Propag.*, vol. 3, no. 4, pp. 638-644, 2009.

[20] P. Gupta, D. Guha, and C. Kumar, "Dielectric resonator working as feed as well as antenna: new concept for dual-mode dual-band improved design," *IEEE Trans. Antennas Propag.*, vol. 64, no. 4, pp. 1497-1502, 2016.

[21] S. Keyrouz and D. Caratelli, "Dielectric resonator antennas: Basic concepts, design guidelines, and recent developments at millimeter-wave frequencies," *International Journal of Antennas and Propagation*, pp. 1-20, 2016.

[22] R. Chair, A. A. Kishk, and K. F. Lee, "Experimental investigation for wideband perforated dielectric resonator antenna," *Electron. Lett.*, vol. 42, no. 3, pp. 15-16, 2006.

[23] J. K. Rai, P. Ranjan, and R. Chowdhury, "Frequency reconfigurable wideband rectangular dielectric resonator antenna for sub-6 GHz applications with machine learning optimization," *AEU - Int. J. Electron. Commun.*, vol. 171, p. 154872, July 2023.

[24] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Elektron Mark.*, vol. 31, pp. 685-695, 2021.

[25] I. H. Sarker, "Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions," *SN Comput. Sci.*, vol. 2, p. 420, 2021.

[26] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, pp. 239-245, 1979.

[27] S. Shakya, C. Schmüdderich, J. Machacek, L. F. Prada-Sarmiento, and T. Wichtmann, "Influence of sampling methods on the accuracy of machine learning predictions used for strain-dependent slope stability," *Geosciences*, vol. 14, no. 2, p. 44, 2024.

[28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *23rd International Conference on Learning Representations (ICLR)*. 2015.

[29] H. Wang, D. Li, Y. Li, G. Zhu, and R. Lin, "Method for remaining useful life prediction of turbofan engines combining Adam optimization-based self-attention mechanism with temporal convolutional networks," *Appl. Sci.*, vol. 14, p. 7723, 2024.

[30] *Computer Simulation Technology*, CST Studio Suite, 2023. http://www.cst.com

[31] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, pp. 281-305, 2012.

[32] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing*, vol. 415, pp. 295-316, 2020.

Fidan Gamze Kizilcay received her B.Sc. and M.Sc. degrees in Electronics and Communication Engineering from Suleyman Demirel University, Isparta, Türkiye. She is pursuing her Ph.D. degree in Electrical Electronics Engineering at Sakarya University. She is currently working as an Instructor at Zonguldak Bülent Ecevit University. Her research interests include electromagnetics and microwave techniques, biomedical engineering and metamaterials.

Muhammet Hilmi Nisanci was born in Istanbul, Turkey, in 1983. He received the B.S. and M.S. degrees from Suleyman Demirel University, Isparta, Turkey, in 2006 and 2009, respectively, both in electronic and telecommunication engineering, and the Ph.D. degree in electrical engineering from the University of L'Aquila, L'Aquila, Italy, in 2013. Since November 2021, he has been working as an Associate Professor with the Department of Electrical and Electronics Engineering, Sakarya University, Sakarya, Turkey. He was involved in the research activities with the UAq EMC Laboratory, L'Aquila, Italy, from February 2007 to March 2009. He is extending his research to include grooved covers and/or cavities. His research interests include the numerical analysis of general electromagnetic problems, reverberation/anechoic chambers, interaction of electromagnetic field with dielectrics and composite media, and their modeling and application for EMC.

An Electromagnetic Imaging Algorithm Based on Generative Adversarial Network for Limited Observation Angle

Chun Xia Yang[1], Xirui Yang[1, 2], Jian Zhang[3], Chi Zhou[1], and Mei Song Tong[4]

[1]Shanghai Engineering Research Center of Intelligent Education and Bigdata
Shanghai Normal University, Shanghai 200234, China
chunxiay@shnu.edu.cn, 1000448084@smail.shnu.edu.cn, 1000528179@smail.shnu.edu.cn

[2]School of Electronic and Information Engineering
Guangzhou City University of Technology, Guangzhou 510800, China

[3]Xpeedic Technology, Inc.
Shanghai 201210, China
jian.zhang@xpeedic.com

[4]Department of Electronic Science and Technology
Tongji University, Shanghai 201804, China
mstong@tongji.edu.cn

***Abstract* –** In the context of long-distance detection and obstacle occlusion, the limited observation angle of electromagnetic imaging poses significant challenges for accurate reconstruction. To address this issue, we propose a hybrid electromagnetic reconstruction algorithm based on a generative adversarial network (GAN). This algorithm utilizes the diffraction tomography (DT) method to generate an initial image, which serves as input for the GAN. Through adversarial training between the generator and the discriminator, the algorithm produces a reconstructed image with enhanced accuracy. Firstly, unlike complete learning-based reconstruction methods that rely solely on scattering field data, our approach effectively integrates both scattering characteristics and *a priori* information from the DT image model, thus improving the accuracy and generalizability of the neural network. Secondly, compared to other linear approximation algorithms, the DT algorithm incorporates fast Fourier transform (FFT) to enhance computational efficiency. Thirdly, this study employs a Fourier spatial data extrapolation technique to mitigate the limitations of insufficient data and improve imaging fidelity. Numerical simulations demonstrate that even at a narrow observation angle of 90°, the proposed algorithm exhibits excellent reconstruction performance and notable generalization ability.

***Index Terms* –** diffraction tomography (DT), generative adversarial network (GAN), inverse scattering, limited angle.

I. INTRODUCTION

In recent years, with the rapid development of artificial intelligence, deep learning technology has made great progress in the electromagnetic field [1, 2], which is expected to bring new ideas for solving inverse scattering problems. Using deep learning to solve electromagnetic inverse problems, the direct approach is to take the measured scattering data as input of the neural network and output the detected target image, which is called complete learning or direct learning reconstruction [3]. In the complete learning reconstruction, neural networks basically operate as a black box, with the user's workload being minimal while the neural network undertakes a heavy task. Due to the unnecessary computational cost spent on learning known physical models or laws, only very simple scatterers can usually be reconstructed.

Recent advancements indicate that the integration of physics-based and data-driven methodologies can substantially improve the analysis and prediction of complex systems. A prominent strategy involves addressing the inverse scattering problem by incorporating deep learning as a complementary computational tool within the framework of traditional physics-based methods. For instance, Guo et al. [4] employed the supervised descent method to solve a two-dimensional microwave imaging challenge. They utilized a neural network to learn the average descent direction between the initial and target models, thereby accelerating model updates and avoiding computationally intensive partial derivative calculations. This approach not only streamlines the

Submitted On: January 15, 2025
Accepted On: July 31, 2025

https://doi.org/10.13052/2025.ACES.J.400803

process but also substantially reduces overall computational costs compared to conventional non-learning algorithms. Another effective strategy is to leverage deep learning techniques to generate an initial estimate for the traditional electromagnetic inverse scattering model. Chen et al. [5] trained a convolutional neural network (CNN) to learn a complex mapping function from magnetic resonance T1 images to dielectric property images. The dielectric image generated by this CNN serves as the initial estimate for microwave imaging, which is then refined through a physics-based image processing step. Similarly, Sanghvi et al. [6] used the signal subspace component of the total contrast source as input to train a CNN, obtaining an initial estimate of the total contrast source. This estimate is subsequently refined using traditional iterative optimization techniques.

Additionally, another strategy takes data-driven methods as the main imaging approach, in which the input data, labels, loss functions, and neural network structures are designed based on physical laws [7]. Some scholars propose a two-step inversion strategy that combines traditional electromagnetic imaging algorithms with neural networks [8, 9]. From a computational efficiency perspective, linear approximation inversion algorithms are commonly employed, where the preliminary imaging results serve as inputs for the neural network instead of scattering data [10, 11]. This approach significantly reduces the learning complexity of the network [12, 13].

While deep learning holds significant potential for electromagnetic inverse scattering, its practical application encounters several critical challenges: (1) Data scarcity: Collecting adequate high-precision scattering field data in real-world scenarios is both expensive and resource-intensive, and the limited availability of observation angles further compounds this issue; (2) Model generalizability: Purely data-driven approaches are prone to overfitting to the training data distribution, which limits their adaptability to complex targets or non-ideal conditions; (3) Physical consistency: Black-box models may produce solutions that contradict fundamental electromagnetic principles, resulting in physically implausible outcomes.

Traditional physics-based methods, such as diffraction tomography (DT) [14, 15], establish explicit mappings between scattering fields and target parameters through Fourier transforms and linear approximations. Despite being constrained by limited observation angles and approximation inaccuracies, these methods offer physically meaningful constraints and computationally efficient initial estimates. In our hybrid physics-data-driven framework, the physical priors of DT are integrated into GAN [16, 17] training to effectively tackle the aforementioned challenges. Specifically, DT's low-complexity initial estimates alleviate the learning burden on the network, while Fourier-based data extrapolation [18] augmented with prior knowledge mitigates data insufficiency. Furthermore, adversarial training improves both generalizability and physical consistency.

This paper is organized as follows: Section II introduces the physical model and formulation of the inverse scattering problem. In Section III, we present a hybrid reconstruction algorithm based on GAN. Section IV provides numerical and experimental results along with their analyses. Finally, Section V summarizes the conclusions.

II. THE ELECTROMAGNETIC INVERSE SCATTERING PROBLEM

Consider a two-dimensional electromagnetic inverse scattering problem. As illustrated in Fig. 1, DOI represents an imaging region in free space containing an unknown scatterer that needs to be detected. The transmitting antenna, located in the observation region S, emits electromagnetic waves to illuminate the imaging region DOI. This induces currents on the scatterer, generating scattered fields. The receiving antennas measure the scattered fields. Then the electromagnetic parameter distribution in imaging region is calculated by inversion algorithms, and the target's spatial position, contour, and internal structure are reconstructed. Assuming that the permeability of the unknown scatterer is the same as the background medium, the objective of reconstruction is the distribution of the relative permittivity in the imaging region.

The scattering integral equation is shown as follows:

$$E^{\mathrm{tot}}(\rho) = E^{\mathrm{inc}}(\rho) + k_b{}^2 \int_{\mathrm{DOI}} g(\rho,\rho') \cdot O(\rho') E^{\mathrm{tot}}(\rho') \mathrm{d}\rho', \rho \in \mathrm{S} \quad (1)$$

where ρ represents the position of the receiving antenna, ρ' represents the position of any point in the imaging region. $g(\rho,\rho') = \frac{i}{4} H_0^{(1)}(k_b|\rho-\rho'|)$ is the two-dimensional Green's function in free space where $H_0^{(1)}$ is the zeroth order Hankel's function of the first kind and k_b is the wavenumber of the background medium. $O(\rho') = \varepsilon_r(\rho') - 1$ represents the contrast distribution, $E^{\mathrm{tot}}(\rho)$ represents the total field, $E^{\mathrm{inc}}(\rho)$ represents the incident field, and the integral term corresponds to the scattered field $E^{\mathrm{sca}}(\rho)$.

III. THE HYBRID RECONSTRUCTION ALGORITHM BASED ON GAN

A. Diffraction Tomography

In Eq.(1), the total field equals the superposition of the incident field and the scattered field. When the scattered field is much weaker relative to the incident field, the incident field can be used to approximate the total field according to the first-order Born approximation. Furthermore, assuming that both the transmitting

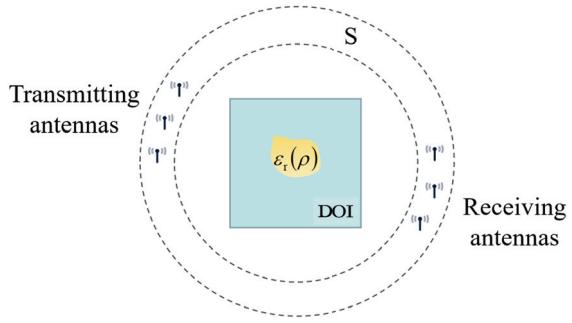


Fig. 1. Electromagnetic inverse scattering model.

antenna and the receiving antenna are located in the far field region of DOI, after a series of derivations, the scattered field can be expressed as [19]:

$$E^{\mathrm{sca}}(\rho_T,\rho_R) \simeq \frac{ik_b}{8\pi\sqrt{|\rho_T||\rho_R|}} e^{ik_b(|\rho_T|+|\rho_R|)} \int_{\mathrm{DOI}} e^{-i(k_b(\hat{\rho}_T+\hat{\rho}_R))\rho'} \cdot O(\rho')d\rho', \quad (2)$$

where $E^{\mathrm{sca}}(\rho_T,\rho_R)$ denotes the scattered field detected by the receiver located at position vector ρ_R, which is generated by the transmitter situated at position vector ρ_T. $\hat{\rho}_T$ and $\hat{\rho}_R$ represent unit vectors in the ρ_T and ρ_R directions, respectively.

Eq. (2) is the formula of the DT algorithm, where the term on the right-hand side conforms to the form of a Fourier transform. Therefore, after measuring the scattering field data at the observation points, the contrast distribution can be obtained by applying simple coefficient multiplication and inverse Fourier transform. Discretizing the imaging region into N pixels, the computational cost of DT is only $O(NlogN)$. The computational complexity of other linear inversion approaches such as the backpropagation method (BP) is $O(N_{inc}N_fNlogN)$ [20, 21], where N_{inc} is the number of incidences, and N_f is the number of iterations for solving the forward problem.

B. Data Extrapolation Based on Fourier Transform

The practical realization of an optimal antenna configuration for achieving comprehensive coverage around the imaging area poses substantial engineering challenges. The restricted viewing angle range reduces diversity in scattering data, posing a challenge for precise reconstruction. To address these challenges, this study integrates a Fourier transform-based extrapolation method, incorporating *a priori* knowledge, to infer missing data in addition to the conventional DT algorithm, thereby enhancing the accuracy and reliability of the results. The following is a concise overview of the steps for implementation, with comprehensive details provided in reference [18].

For convenience, the equation (2) is abbreviated as

$$E(\rho) \simeq \mathscr{F}(O(\rho')), \quad \rho \in \mathrm{S}, \quad \rho' \in \mathrm{DOI}. \quad (3)$$

The limitation of the observation angle results in only a restricted segment of $E(\rho)$ being measurable, which is denoted as

$$E_1(\rho) = E(\rho)g(\rho), \quad \text{where } g(\rho) = \begin{cases} 1, & \rho \in \text{known area}, \\ 0, & \text{otherwise}. \end{cases} \quad (4)$$

Step 1: The contrast distribution in the imaging space is obtained by inverse Fourier transform of the scattering data.

$$O_n(\rho') = \mathscr{F}^{-1}(E_n(\rho)), \quad (5)$$

where $\mathscr{F}^{-1}$ is the inverse operator of $\mathscr{F}$.

Step 2: The target distribution is updated by adjusting the contrast values obtained in the previous step based on *a priori* information.

$$\widetilde{O}_n(\rho') = \begin{cases} O_n(\rho'), & \text{if } O_n(\rho') > 0 \\ 0, & \text{otherwise} \end{cases}. \quad (6)$$

Step 3: The updated contrast data is transformed by the operator $\mathscr{F}$ to obtain the corresponding scattering data.

$$\widetilde{E}_n(\rho) = \mathscr{F}(\widetilde{O}_n(\rho')). \quad (7)$$

Step 4: The initial measurement data is employed for local recovery of the previously updated scattering data.

$$E_{n+1}(\rho) = \widetilde{E}_n(\rho) + \left[E_1(\rho) - \widetilde{E}_n(\rho)\right] g(\rho) = \begin{cases} E_1(\rho), & \rho \in \text{known area} \\ \widetilde{E}_n(\rho), & \text{otherwise} \end{cases}. \quad (8)$$

Step 5: When the termination condition is met, the iteration ends; otherwise, return to Step 1 and increase n by 1. By iteratively following these steps, it becomes feasible to preserve the original measurement data while gradually inferring missing data, thereby significantly enhancing the quality of target reconstruction. The effective contrast range for numerical simulation is defined as $[0,\infty)$ in this study aligning with the typical parameters commonly employed in inverse scattering problems. For practical engineering scenarios, more sophisticated prior conditions can be specified.

The Fourier extrapolation enhances reconstruction accuracy through two key mechanisms. First, Fourier-domain completion. Under limited observation angles, the measurable scattering data corresponds to a truncated segment of the target's Fourier spectrum. By iteratively applying prior constraints (e.g., non-negativity of contrast, spatial sparsity) during extrapolation, we infer missing high-frequency components (as described in Eq. (8)). This process mitigates artifacts and refines edges in the DT image, providing higher-quality inputs for the GAN. Second, physical consistency enforcement. The iterative updates outlined in Eqs. (6)–(8) ensure

that the extrapolated data are consistent with both the measured scattering fields and prior knowledge (e.g. expected permittivity ranges). This approach suppresses non-physical oscillations in the reconstructed image, which are commonly encountered in purely data-driven methods under sparse measurement conditions.

C. Generative Adversarial Network

In this paper, we utilize the GAN as the network model for the data-driven component of the hybrid inversion algorithm. The GAN framework comprises two key components: a generator and a discriminator [22]. The unique adversarial training mechanism allows the model to more effectively capture complex data features and generate high-precision images, making it especially suitable for addressing the inverse scattering problem discussed in this paper. However, traditional purely data-driven GANs face challenges in handling missing data under limited observation angles and may produce non-physical solutions. Our algorithm incorporates physical constraints through DT initialization and Fourier-space extrapolation. Firstly, DT's linear approximations efficiently generate initial target estimates, avoiding direct learning of complex scattering mappings. Secondly, extrapolation utilizes prior knowledge (e.g., target contrast range, spatial sparsity) to constrain solution spaces and suppress non-physical outputs. Finally, the U-Net generator retains the low-frequency features of DT via skip connections, while the Convolutional Block Attention Module (CBAM) [23] module enhances edge reconstruction, thereby ensuring synergistic optimization of physical principles and data-driven features.

1) Generator. The structures of the generator and discriminator networks are illustrated in Fig. 2. In this work, the generator model adopts the structure of U-Net [24]. U-Net is a convolutional neural network commonly used for image segmentation. It has a U-shaped structure consisting of a symmetric encoder and decoder. The encoding part encodes the input rough image through a sequence of convolutional layers to obtain higher-dimensional feature representations. The decoding part decodes the latent space vector through a series of upsampling to reconstruct an image resembling the training data eventually. The feature maps of corresponding dimensions in the downsampling and upsampling processes are fused by employing the skip connection structure. In this paper, during each downsampling process, the CBAM is incorporated to enhance the representation capability of the generator network. This is achieved by selectively emphasizing informative features while suppressing irrelevant ones. The reconstruction accuracy of the relative permittivity and position of the target scatterer is further improved.

2) Discriminator. Compared with other network architectures, the GAN stands out due to its incorporation of a discriminator. The discriminator network contains several convolutional layers and finally passes through an average pooling layer followed by a sigmoid activation operation. In this study, the discriminator receives two distinct types of input data: one consists of real images from the training set, which represent the true distribution of scatterers; the other comprises synthetic images generated by the generator G. The discriminator's role is to estimate the probability that an input

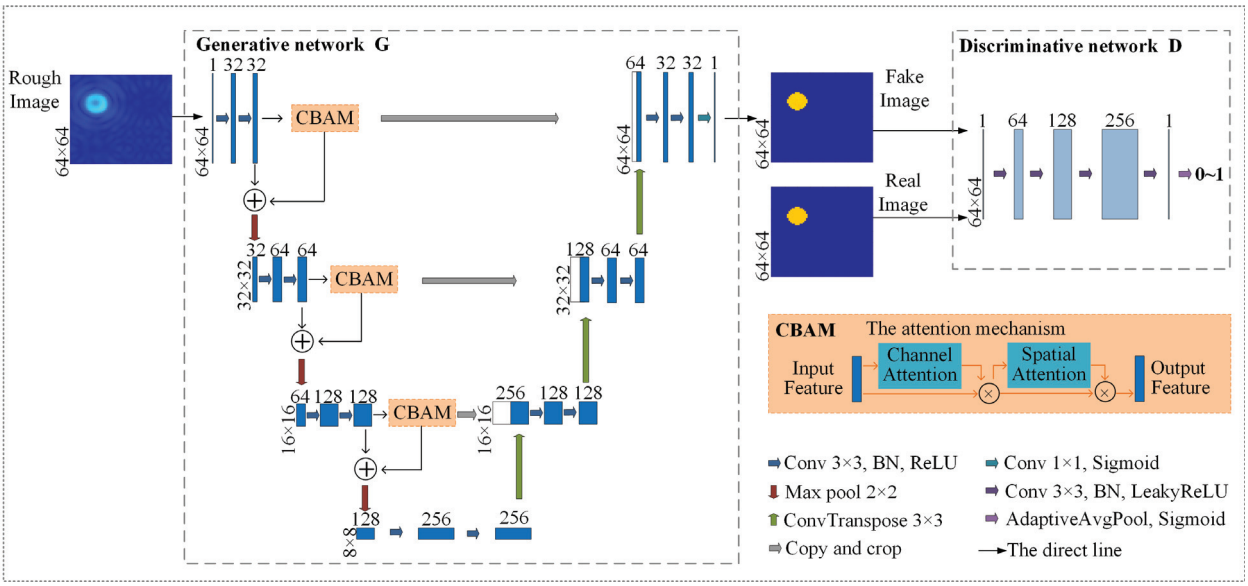


Fig. 2. The structure of the GAN.

sample originates from the training data rather than from the generator, thereby assessing the authenticity of the generated samples and providing feedback to the generator. This feedback mechanism prompts the generator to iteratively refine its parameters to produce increasingly realistic samples.

3) Training process. The generator and discriminator are trained in an adversarial manner until a Nash equilibrium is reached. This can be mathematically described as:

$$\min_{\theta_{\mathrm{G}}}\max_{\theta_{\mathrm{D}}} V(\theta_{\mathrm{G}},\theta_{\mathrm{D}}) = \mathbb{E}_{I^{real}\sim P_{real}(I^{real})}[\log \mathrm{D}_{\theta_{\mathrm{D}}}(I^{real})] + \mathbb{E}_{I^{rough}\sim P_{rough}(I^{rough})}[\log(1-\mathrm{D}_{\theta_{\mathrm{D}}}(\mathrm{G}_{\theta_{\mathrm{G}}}(I^{rough})))]. \tag{9}$$

Here, $\mathbb{E}$ denotes the mathematical expectation of the data distribution, I^{real} represents the real image, I^{rough} denotes the rough image input to the generator, while θ_{G} and θ_{D} refer to parameters to be trained in the generator and discriminator, respectively.

The loss function for the generator is defined as:

$$L_{\mathrm{G}} = \alpha L_{\mathrm{GAN}}^{\mathrm{G}} + L_{\mathrm{MSE}}. \tag{10}$$

The first term on the right-hand side of the equation represents the adversarial loss, where α serves as a weighting parameter, and L_{MSE} denotes the pixel-wise mean square error between the generated image and the real image distribution. The specific formula is as follows:

$$L_{\mathrm{MSE}} = \frac{1}{N}\sum_{i=1}^{N}\left(I_{\rho_i}^{real} - \mathrm{G}_{\theta_{\mathrm{G}}}(I^{rough})_{\rho_i}\right)^2. \tag{11}$$

In this paper, the generator is pre-trained using L_{MSE} prior to adversarial training in order to accelerate the learning process and prevent convergence to local optima. The adversarial loss is defined as:

$$L_{\mathrm{GAN}}^{\mathrm{G}} = -\log\left(\mathrm{D}_{\theta_{\mathrm{D}}}\left(\mathrm{G}_{\theta_{\mathrm{G}}}(I^{\mathrm{Rough}})\right)\right). \tag{12}$$

This equation indicates that the control parameters θ_{D} of the discriminator remains constant, while the parameters θ_{G} of the generator is optimized to enable the output image to deceive the discriminator to the greatest extent possible.

Similarly, when the generator parameters θ_{G} are held constant, training the discriminator parameters θ_{D} enhances its ability to accurately differentiate between real and generated images. Throughout this adversarial training process, the discriminator's ability to discern between real and synthetic samples also progressively improves. Ultimately, the model converges to an equilibrium where the generator produces samples so realistic that the discriminator can no longer reliably distinguish between real and generated data.

IV. NUMERICAL ANALYSIS

A. Test One: Circular Scatterers

The simulation setup is defined as follows. The imaging area is a square with each side measuring 0.8 m, divided into a grid of 64×64 pixels. The center of the imaging area serves as the coordinate origin. The observation area is equipped with 17 transmitting antennas and 17 receiving antennas, uniformly distributed along an arc spanning from 0° to 90°, with a radius of 2 m from the origin. Each transmitting antenna successively emits incident waves containing three frequencies (that is, 0.5 GHz, 1.0 GHz and 1.5 GHz) to illuminate the imaging area. Meanwhile, all receiving antennas synchronously capture the scattered fields generated by each individual incident wave.

The target is a circular homogeneous dielectric with a relative permittivity ranging from 1.2 to 3 (contrast of 0.2 to 2 with a step size of 0.1) and a radius range between 0.03 m and 0.08 m (with a step size of 0.01 m). The center coordinates are distributed within the range $(-0.2$ to $0.2, -0.2$ to $0.2)$ m. 3000 sets of training data and 300 sets of testing data were generated via simulation using the Method of Moments (MoM). The GAN adopts the Adam optimizer with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The learning rates for the generator and discriminator are both set to 1×10^{-5}. The batch size is set to 8, and the training process has been conducted for 200 epochs. The adversarial loss parameter α is set to 1×10^{-3}. The setting of these hyperparameters was guided by relevant literature [22][24], and their optimal values were identified through systematic numerical experiments.

By comparing the simulation results with the GAN-based direct inversion algorithm (referred to as "GAN") and the hybrid algorithm combining BP and GAN (referred to as "BP+GAN"), we validate the performance of the proposed hybrid electromagnetic reconstruction algorithm that combines extrapolated enhanced DT and GAN (referred to as "DT+GAN"). Specifically, the GAN algorithm utilizes scattered field data directly, whereas BP+GAN uses images reconstructed by the BP algorithm as input for its neural network. The results are presented in Fig. 3, where the first column displays the test sample with a radius of 0.04 m and a contrast value of 1.1. Meanwhile, the second column exhibits the sample with a radius of 0.07 m and a contrast value of 0.8. The hybrid inversion methods exhibit a superior capability in differentiating scatterers from the background, significantly outperforming the direct learning reconstruction method. In particular, the DT+GAN method achieves a more precise reconstruction of contrast values compared to the BP+GAN approach.

To quantify the reconstruction accuracy, the imaging error is calculated using the following formula[25]:

$$error = \frac{100\%}{N}\frac{\sum_{n=1}^{N}\left|\widetilde{O}(\rho) - O(\rho)\right|^2}{\sum_{n=1}^{N}|O(\rho)|^2}. \tag{13}$$

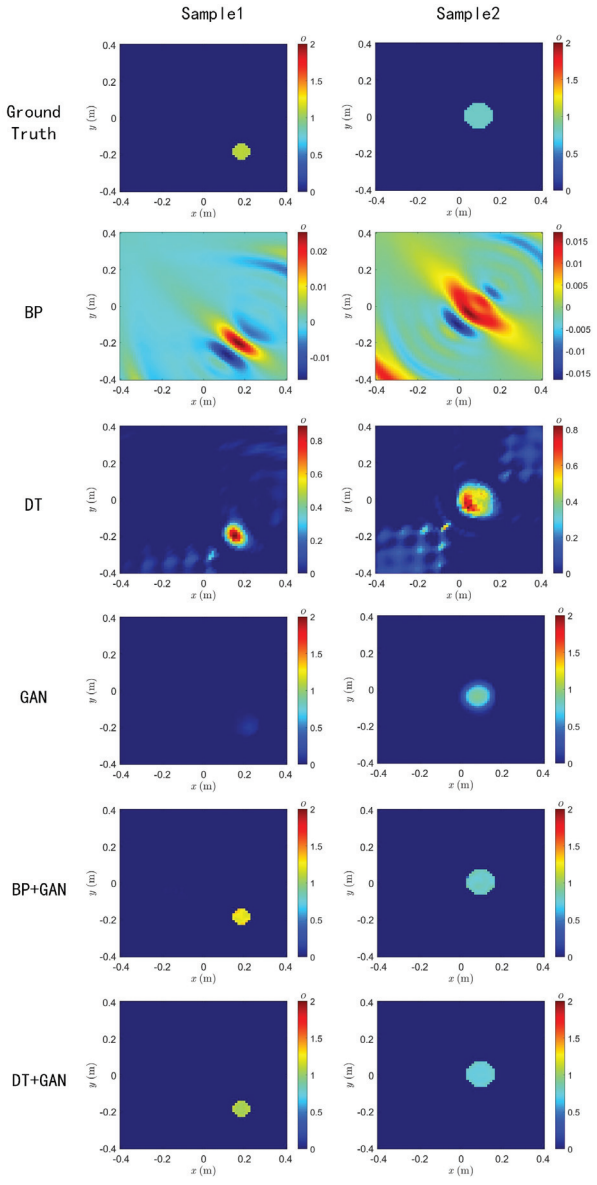


Fig. 3. The reconstruction results of a single circle under $0^\circ - 90^\circ$ observation. The first row illustrates the ground truth distribution, while the second, third, fourth, fifth, and sixth rows depict the reconstruction results of BP, DT, direct GAN, BP+GAN, and DT+GAN respectively.

The contrast of a cell in the imaging region predicted by the neural network is denoted as $\tilde{O}(\rho)$, while $O(\rho)$ represents the actual contrast of the corresponding position. Here, N refers to the number of cells in the imaging region. The cumulative distribution functions (CDFs) of the inversion error of BP+GAN and DT+GAN in the test set are shown in Fig. 4. The average reconstruction error for BP+GAN is $4.63 \times 10^{-6}\%$, while for DT+GAN it is $2.54 \times 10^{-6}\%$.

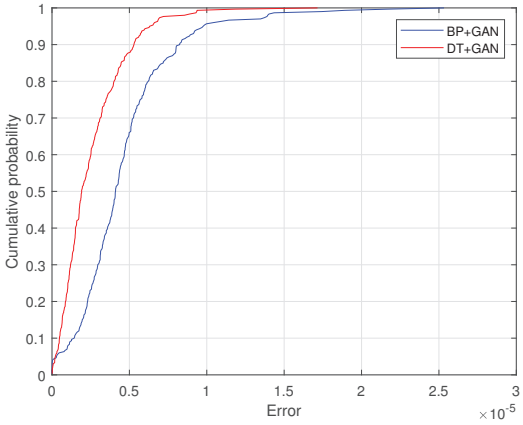


Fig. 4. CDFs of the reconstruction error for a single circular object.

The average execution times for the components BP, DT and GAN in hybrid algorithms based on 300 test samples have been analyzed. The BP algorithm recorded an average execution time of 0.0749 s, while the DT algorithm was significantly faster at 0.0050 s. The GAN component, when integrated with both BP and DT in the hybrid algorithms (BP+GAN and DT+GAN), demonstrated an average processing time of 0.2357 s, with negligible variation across different configurations. The execution time is measured on a computer equipped with an Intel i5-12400F CPU and an NVIDIA GeForce RTX 4060 GPU. These results highlight the superior computational efficiency of the DT algorithm compared to BP, even when an extrapolation step is included. The difference in solution efficiency can be attributed to the fact that the computation time of the BP algorithm increases significantly with the increase in the number of excitation waves, while this factor has a relatively minor impact on the DT algorithm.

To assess the generalization capability of the inversion algorithms, two dielectric circles were employed as reconstruction targets to evaluate the above neural networks. By testing 300 randomly generated samples, the BP+GAN model achieves an average error of $5.90 \times 10^{-3}\%$, whereas the DT+GAN model achieves an average error of $5.20 \times 10^{-3}\%$. Fig. 5 illustrates the reconstruction results for two samples, while Fig. 6 plots the CDFs of the reconstruction errors. In Sample 1, the contrast values of the two circles are set to 1.6 and 0.4, with corresponding radii of 0.03 m and 0.05 m, respectively. In Sample 2, the contrast values are adjusted to 0.2 and 0.7, with corresponding radii of 0.03 m and 0.06 m, respectively. The simulation results demonstrate that both direct GAN and BP+GAN tend to lose information regarding the target with low contrast during the reconstruction of the two-circle target,

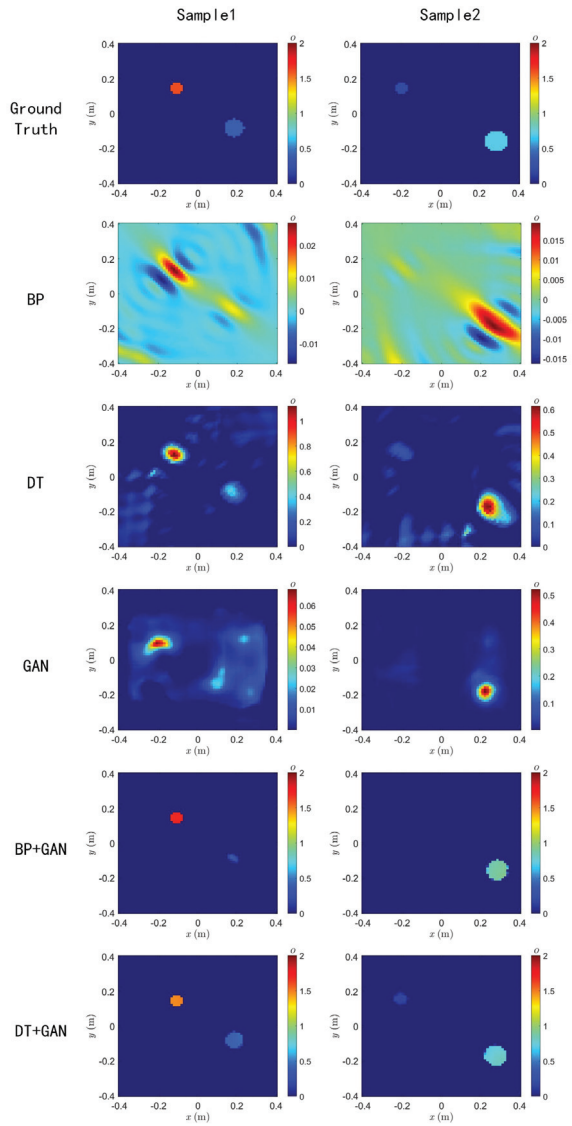


Fig. 5. Reconstruction results of two circular structures in the test of generalization ability.

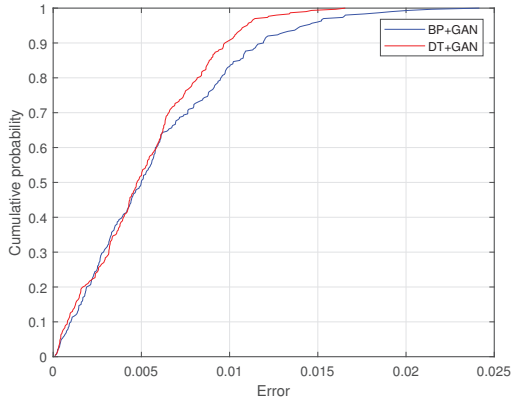


Fig. 6. CDFs of two circle reconstruction errors used to evaluate generalization ability.

whereas DT+GAN does not exhibit this issue. Although there is a slight deviation in capturing scatterer structure details, DT+GAN successfully achieves accurate inversion of both the target location and contrast value.

B. Test Two: Handwriting Digits

In this section, the reconstruction target is derived from the MNIST dataset [26], a widely utilized database of handwritten digits in the field of machine learning. The scatterers are modeled as dielectric materials with randomly assigned relative permittivities ranging between 1.2 and 3.0. Each image is upscaled from its original size of 28×28 pixels to 64×64 pixels to align the spatial resolution with the imaging setup described in Section IV-A. This adjustment standardizes the problem scale across tests and increases the number of unknowns, thereby intensifying the complexity of the inverse problem. Such a practice enables robust comparisons of reconstruction fidelity under heightened computational demands. The imaging model and neural network parameters were set identical to those used in test one. A simulation employing the MOM generated a training set consisting of $10,000$ samples, along with a validation set of $1,000$ samples and a testing set of 300 samples. Fig. 7 presents reconstructed profiles from two representative examples, while Fig. 8 displays the CDFs of the reconstruction errors. Since both BP and DT are linear approximate inversion algorithms, and given the large training dataset used in this experiment, the difference in reconstruction quality between the BP+GAN model and the DT+GAN model is not pronounced.

To evaluate the generalization performance of the algorithms, a randomly distributed circular scatterer is introduced into the imaging region, following the settings from test one, which may cause partial occlusion of the handwritten digits. Fig. 9 displays images of two reconstructed samples, whereas Fig. 10 illustrates the CDFs of the reconstruction errors. The results indicate that, compared to BP+GAN, DT+GAN demonstrates superior generalization capability, allowing for better differentiation of distinct targets within the imaging region and achieving a reconstructed contrast value that is more closely aligned with the ground truth.

C. Test Three: Experimental Data

In this section, we validate the proposed reconstruction method using experimental data provided by the Fresnel Institute. The measurement system is detailed in reference [27]. The scatterer is an off-centered circular cylinder with a significant vertical dimension, presenting an almost two-dimensional structure. Slightly deviating from the previous simulation model, the transmitter remains fixed while the scatterer rotates about the central vertical axis from 0° to 350° in increments of 10°. To assess the reconstruction performance

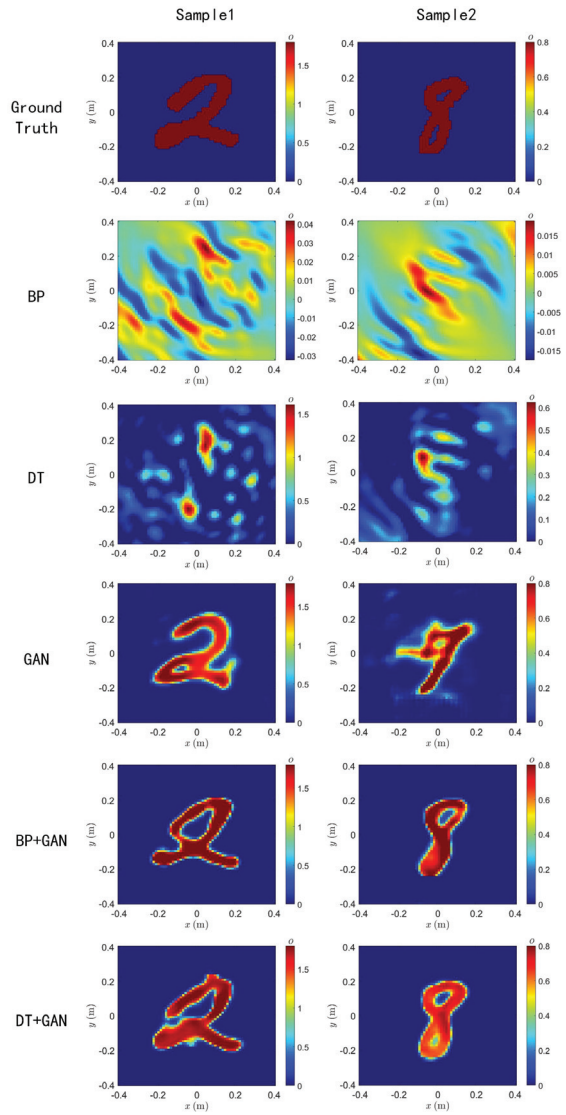


Fig. 7. The reconstructed results of handwritten digits from MNIST database.

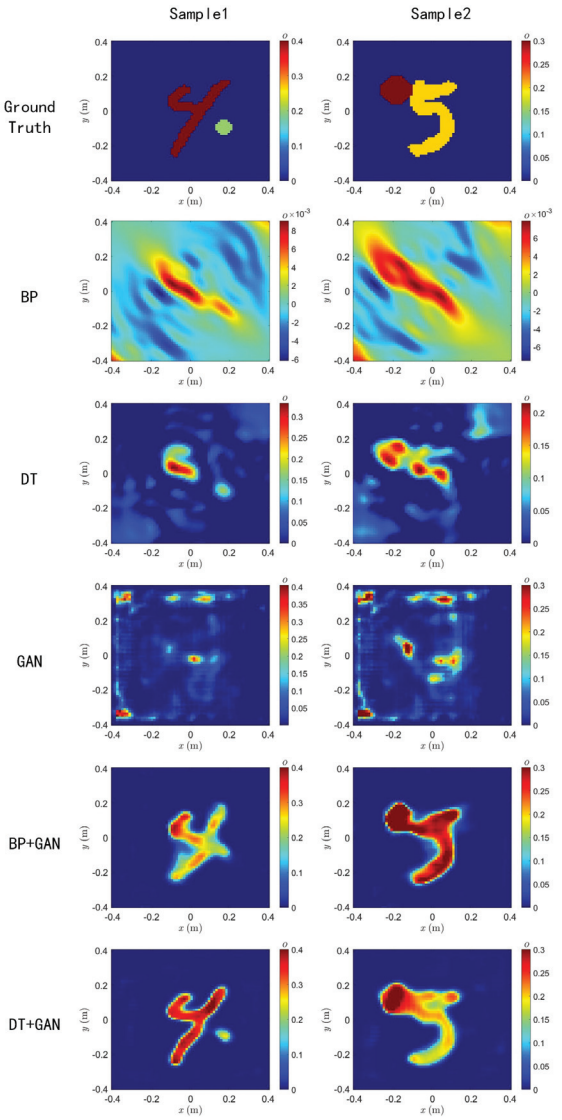


Fig. 9. The reconstructed results of handwritten digits with an extra circle.

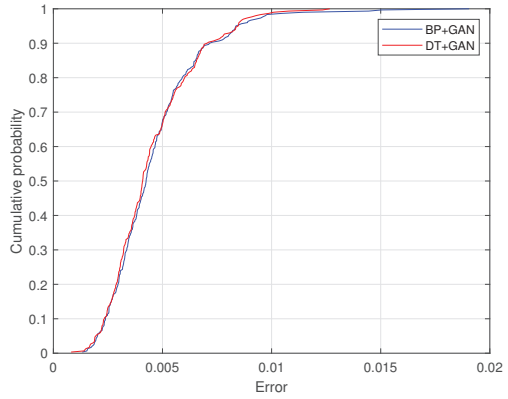


Fig. 8. CDFs of the reconstruction error for handwritten digits.

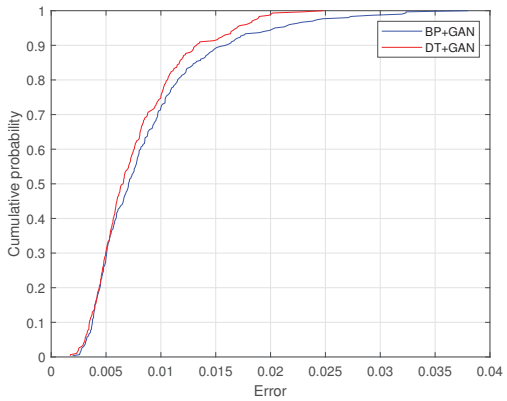


Fig. 10. CDFs of reconstruction error for digits with circles.

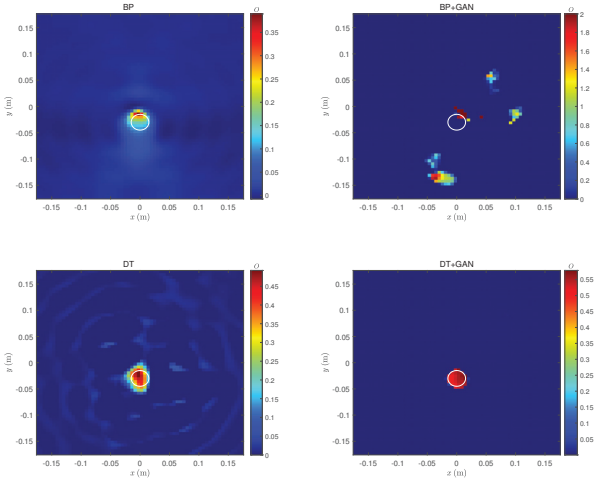


Fig. 11. Reconstruction results using the experimental data (from data set "dielTM_dec8f.exp" at 4 GHz [27]).

under a limited viewing angle, we restrict the rotation angle to the range of 40° to 140° using measurements from the "dielTM_dec8f.exp" dataset at a frequency of 4 GHz.

The neural networks trained in the first test were utilized for verification. Due to alterations in the imaging model and changes in the dimensions of scattering data, the direct application of the GAN method becomes impractical. Fig. 11 illustrates the reconstruction results obtained using BP, DT, BP+GAN, and DT+GAN methods. A white ring on the diagram marks the exact position of the scatterer. Its circular cross-section possesses a radius of 15 mm, with its center positioned at coordinates $(0, -30)$ mm. The estimation of the real part of the relative permittivity yields $\varepsilon_r = 3 \pm 0.3$. The limited observation angle introduces a discrepancy between the BP image and the actual target. Additionally, the imaging model in this scenario significantly deviates from that of the training data, especially concerning the antenna positioning. Specifically, while the excitation wave angle in the training data spans from 0° to 90°, it ranges from 40° to 140° in this case. This substantial difference leads to considerable deviations in the reconstruction outcomes of the BP+GAN technique. However, by employing data extrapolation technology, the DT algorithm can still effectively capture the orientation and size information of scatterers under challenging imaging conditions. Moreover, integrating GAN into DT-based reconstructions (DT+GAN) not only improves the reconstruction accuracy of scatterers but also reduces surrounding artifacts.

Experiments demonstrate that integrating physical priors significantly alleviates the core challenges in deep learning. Specifically, DT's rapid initial estimates reduce reliance on large-scale training data; data extrapolation and prior constraints improve generalizability under narrow observation angles; and adversarial training, guided by physics-informed interactions between the generator and discriminator, effectively suppresses non-physical solutions.

V. CONCLUSION

This paper proposes a hybrid electromagnetic reconstruction algorithm that combines DT and GAN to address the limited viewing angle problem in electromagnetic imaging. This algorithm achieves greater accuracy in reconstructing the scatterer's position, contour, and contrast than the complete learning reconstruction method. This improvement can be attributed to the incorporation of physical mechanisms that effectively reduce the learning complexity of neural networks. Furthermore, compared to other two-step methods like the BP+GAN approach, this algorithm achieves higher solving efficiency by utilizing the fast Fourier transform. Additionally, the introduction of data extrapolation technology effectively addresses reconstruction challenges arising from limited viewing angles. Future research will continue to leverage physical mechanisms to address the inherent nonlinearity and challenges in the mapping process, while extending validation to a wider variety of scatterer geometries, such as multilayered structures, and a more diverse range of observation angles, thereby further evaluating the generalization capabilities.

REFERENCES

[1] Z. Wei and X. Chen, "Solving Full-Wave Nonlinear Inverse Scattering Problems by Deep Learning Schemes," in *2019 IEEE International Conference on Computational Electromagnetics (ICCEM)*, pp. 1-2, IEEE, 2019.

[2] L. Li, L. G. Wang, F. L. Teixeira, C. Liu, A. Nehorai, and T. J. Cui, "DeepNIS: Deep neural network for nonlinear electromagnetic inverse scattering," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 3, pp. 1819-1825, 2018.

[3] X. Chen, Z. Wei, L. Maokun, P. Rocca, et al., "A review of deep learning approaches for inverse scattering problems (invited review)," *Electromagnetic Waves*, vol. 167, pp. 67-81, 2020.

[4] R. Guo, X. Song, M. Li, F. Yang, S. Xu, and A. Abubakar, "Supervised Descent Learning Technique for 2-D Microwave Imaging," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 5, pp. 3550-3554, 2019.

[5] G. Chen, P. Shah, J. Stang, and M. Moghaddam, "Learning-Assisted Multimodality Dielectric Imaging," *IEEE Transactions on Antennas*

and Propagation, vol. 68, no. 3, pp. 2356-2369, 2020.

[6] Y. Sanghvi, Y. Kalepu, and U. K. Khankhoje, "Embedding Deep Learning in Inverse Scattering Problems," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 46-56, 2020.

[7] R. Guo, T. Huang, M. Li, H. Zhang, and Y. C. Eldar, "Physics-Embedded Machine Learning for Electromagnetic Data Imaging: Examining three types of data-driven imaging methods," *IEEE Signal Processing Magazine*, vol. 40, no. 2, pp. 18-31, 2023.

[8] K. Xu, Z. Qian, Y. Zhong, J. Su, H. Gao, and W. Li, "Learning-Assisted Inversion for Solving Nonlinear Inverse Scattering Problem," *IEEE Transactions on Microwave Theory and Techniques*, vol. 71, no. 6, pp. 2384-2395, 2023.

[9] C. X. Yang, J. J. Meng, S. Wei, and M. S. Tong, "A Dual-input Electromagnetic Inverse Scattering Algorithm Based on Improved U-net," *Applied Computational Electromagnetics Society Journal (ACES)*, pp. 961-969, 2024.

[10] X. Ye, N. Du, D. Yang, X. Yuan, R. Song, S. Sun, and D. Fang, "Application of Generative Adversarial Network-Based Inversion Algorithm in Imaging 2-D Lossy Biaxial Anisotropic Scatterer," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 9, pp. 8262-8275, 2022.

[11] R. Song, Y. Huang, K. Xu, X. Ye, C. Li, and X. Chen, "Electromagnetic Inverse Scattering With Perceptual Generative Adversarial Networks," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 689-699, 2021.

[12] M. Li, R. Guo, K. Zhang, Z. Lin, F. Yang, S. Xu, X. Chen, A. Massa, and A. Abubakar, "Machine learning in electromagnetics with applications to biomedical imaging: A review," *IEEE Antennas and Propagation Magazine*, vol. 63, no. 3, pp. 39-51, 2021.

[13] Z. Ma, K. Xu, R. Song, C.-F. Wang, and X. Chen, "Learning-Based Fast Electromagnetic Scattering Solver Through Generative Adversarial Network," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 4, pp. 2194-2208, 2021.

[14] T. Negishi, V. Picco, L. L. Monte, and D. Erricolo, "Dyadic contrast function for the forward model of diffraction tomography of thin cylindrical objects," *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 991-994, 2016.

[15] C. Yang, J. Zhang, and M. S. Tong, "An FFT-accelerated particle swarm optimization method for solving far-field inverse scattering problems," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 2, pp. 1078-1093, 2020.

[16] A. Aggarwal, M. Mittal, and G. Battineni, "Generative adversarial network: An overview of theory and applications," *International Journal of Information Management Data Insights*, vol. 1, no. 1, p. 100004, 2021.

[17] L. Guo, G. Song, and H. Wu, "Complex-valued Pix2pix—Deep neural network for nonlinear electromagnetic inverse scattering," *Electronics*, vol. 10, no. 6, p. 752, 2021.

[18] C. Yang, J. Zhang, and M. Tong, "A novel diffraction tomographic algorithm with narrow observation angle," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 32, no. 5, p. e2583, 2019.

[19] W. C. Chew, *Waves and Fields in Inhomogenous Media*, vol. 16, John Wiley & Sons, 1999.

[20] K. Belkebir, P. C. Chaumet, and A. Sentenac, "Superresolution in total internal reflection tomography," *J. Opt. Soc. Am. A*, vol. 22, no. 9, pp. 1889-1897, Sep. 2005.

[21] Z. Wei and X. Chen, "Deep-learning schemes for full-wave nonlinear inverse scattering problems," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 4, pp. 1849-1860, 2018.

[22] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139-144, 2020.

[23] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," in *Proceedings of the European Conference on Computer Vision (ECCV)*, Sep. 2018.

[24] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pp. 234-241, Springer International Publishing, Cham, 2015.

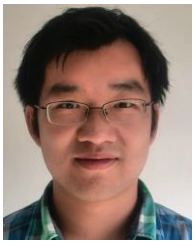
[25] C. X. Yang, J. Zhang, and M. S. Tong, "A hybrid inversion method based on the bat algorithm for microwave imaging of two-dimensional dielectric scatterers," *Progress in Electromagnetics Research M*, vol. 102, pp. 91-104, 2021.

[26] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.

[27] K. Belkebir and M. Saillard, "Testing inversion algorithms against experimental data," *Inverse problems*, vol. 17, no. 6, p. 1565, 2001.

Chun Xia Yang received the Ph.D. degree in electronic science and technology from Tongji University, Shanghai, China, in 2017. During her doctoral studies, she also conducted research at the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Champaign, IL, USA as a visiting student between 2014 and 2016. She is currently an associate professor at the Department of Communication Engineering, Shanghai Normal University, Shanghai, China. Her ongoing research interests primarily revolve around electromagnetic inverse scattering for imaging and computational electromagnetics.

Xirui Yang received her B.S. degree in Communication Engineering and M.S. degree in Electronic Information from Shanghai Normal University, Shanghai, China, in 2021 and 2024, respectively. She is currently affiliated with the School of Electronic and Information Engineering, Guangzhou City University of Technology, Guangzhou, China. Her research focuses primarily on electromagnetic inverse scattering.



Jian Zhang received his B.S. degree in Electronics Science and Technology and his Ph.D. degree in Control Science and Engineering from Tongji University, Shanghai, China, in 2014 and 2020, respectively. He is currently working as an Engineer at Xpeedic Technology, Inc., Shanghai, China. His research mainly focuses on multi-physics numerical simulation techniques.

Chi Zhou received the M.S. degree in Electronic Information from Shanghai Normal University, Shanghai, China, in 2025. He is currently affiliated with Wenyin Cloud Computing Co., Ltd, Shanghai, China. His research mainly focuses on electromagnetic inverse scattering.

Mei Song Tong received the B.S. and M.S. Degrees from Huazhong University of Science and Technology, Wuhan, China, respectively, and Ph.D. degree from Arizona State University, Tempe, Arizona, USA, all in electrical engineering. He is currently the Distinguished Professor and Head of Department of Electronic Science and Technology, and Vice Dean of College of Microelectronics, Tongji University, Shanghai, China. He has also held an adjunct professorship at the University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, and an honorary professorship at the University of Hong Kong, China. He has published more than 700 papers in refereed journals and conference proceedings and coauthored six books or book chapters. His research interests include electromagnetic field theory, antenna theory and design, simulation and design of RF/microwave circuits and devices, interconnect and packaging analysis, inverse electromagnetic scattering for imaging, and computational electromagnetics.

Prof. Tong is a Fellow of the Electromagnetics Academy, Fellow of the Japan Society for the Promotion of Science (JSPS), and Senior Member (Commission B) of the USNC/URSI. He has been the chair of Shanghai Chapter since 2014 and the chair of SIGHT committee in 2018, respectively, in IEEE Antennas and Propagation Society. He has served as an associate editor or guest editor for several well-known international journals, including IEEE Antennas and Propagation Magazine, IEEE Transactions on Antennas and Propagation, IEEE Transactions on Components, Packaging and Manufacturing Technology, International Journal of Numerical Modeling: Electronic Networks, Devices and Fields, Progress in Electromagnetics Research, and Journal of Electromagnetic Waves and Applications, etc. He also frequently served as a session organizer/chair, technical program committee member/chair, and general chair for some prestigious international conferences. He was the recipient of a Visiting Professorship Award from Kyoto University, Japan, in 2012, and from University of Hong Kong, China, 2013. He advised and coauthored 15 papers that received the Best Student Paper Award from different international conferences. He was the recipient of the Travel Fellowship Award of USNC/URSI for the 31th General Assembly and Scientific Symposium (GASS) in 2014, Advance Award of Science and Technology of Shanghai Municipal Government in 2015, Fellowship Award of JSPS in 2016, Innovation Award of Universities' Achievements of Ministry of

Education of China in 2017, Innovation Achievement Award of Industry-Academia-Research Collaboration of China in 2019, "Jinqiao" Award of Technology Market Association of China in 2020, Baosteel Education Award of China in 2021, Carl Friedrich von Siemens Research Award of the Alexander von Humboldt Foundation of Germany in 2023, and Technical Achievement Award of Applied Computational Electromagnetic Society (ACES) of USA in 2024. In 2018, he was selected as the Distinguished Lecturer (DL) of IEEE Antennas and Propagation Society for 2019-2022.

Element Failure Correction for Conformal Antenna Array Using Pre-tuned Non-Dominated Particle Swarm Optimization

Hina Munsif[1], Irfan Ullah[1], Shahid Khattak[1], and Shafqat Ullah Khan[2]

[1]Department of Electrical Engineering
COMSATS University Islamabad, Abbottabad Campus KPK 22060, Pakistan
Hinaa3387@gmail.com, eengr@cuiatd.edu.pk, skhattak@cuiatd.edu.pk

[2]Department of Electronics
University of Buner, Buner, KPK 19290, Pakistan
shafqatphy@yahoo.com

***Abstract* –** This paper presents an enhanced pre-tuned particle swarm optimization (PT-PSO) algorithm for fault compensation in conformal array antennas, addressing both complete element failures and faulty phase shifters. Unlike conventional PSO, which initializes particles randomly across the entire search space (often requiring more iterations and risking local minima), PT-PSO introduces a pre-tuning mechanism that arranges initial amplitudes and phases to guide convergence. Combined with non-dominated sorting, this approach improves multi-objective optimization efficiency by reducing the search space and minimizing local minima, enabling rapid convergence to near-optimal excitation weights. To validate its effectiveness, a 1×8 X-band cylindrical conformal patch array was designed and simulated in HFSS. Results show that PT-PSO successfully restores first sidelobe levels (FSLL) and peak gain under both complete and partial failure scenarios, ensuring accurate pattern recovery. Compared to non-dominated PSO and convex optimization, PT-PSO achieves similar pattern quality with significantly lower computational complexity. The proposed method is particularly applicable to radar and wireless communication systems, where maintaining beam integrity under hardware faults is critical.

***Index Terms* –** Beamforming, conformal array, PSO, sidelobe level.

I. INTRODUCTION

Maintaining a low sidelobe level (SLL) is essential for improving spatial resolution, interference rejection, and overall beamforming performance across a wide range of applications. In radar systems, low SLL reduces clutter and false alarm rates [1-2]. In communication systems, it suppresses co-channel interference and enhances signal-to-noise ratio (SNR). High-resolution imaging and precise target detection (critical in domains such as automotive radar and microwave medical imaging) also rely on tightly controlled sidelobe behavior to avoid ghost targets, imaging artifacts, and performance degradation in safety-critical environments.

In beamforming phased array systems, SLL suppression and beam shaping are typically achieved through amplitude-only weighting or complex weighting (amplitude and phase). Traditional methods such as Taylor tapering and Dolph-Chebyshev distributions, as well as metaheuristic algorithms like Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Differential Evolution (DE), are commonly used to design optimal excitation weights. Conformal beamforming arrays have been adopted in diverse real-world platforms, including UAV and space vehicle communications [3], adaptive wireless communication systems [4], airborne satellite communication [5], millimeter-wave synthetic aperture radar (SAR) [6], and wearable medical imaging technologies [7].

The beamforming process in the above-mentioned systems is obtained using a transmitter/receiver module (TRM). Improper management of switching or damage to TRM (such as due to failure of CMOS transistors) can lead to faulty array antenna elements, severely affecting the performance of the communication system. When an element in the array becomes faulty, either due to complete amplitude failure (zero amplitude) or malfunctioning of phase shifter (generating random phases), it can deteriorate the entire beamforming radiation pattern. Most of the existing literature [8-13] has focused on addressing edge-element failures, where various optimization techniques are used to redistribute complex weights among working elements to recover gain, SLL, and nulls. In beamforming array systems,

Submitted On: April 21, 2025
Accepted On: July 29, 2025

https://doi.org/10.13052/2025.ACES.J.400804

central elements play a more significant role in determining the gain, shaping the main beam, and controlling the first sidelobe level (FSLL). Unlike the edge elements, the center elements act as the core components of the array, carry maximum amplitudes, and significantly influence the overall performance of the radiation pattern. The random faults in central phase shifters introduce destructive interference among array elements and distort the entire array pattern. Studies [14-15] have achieved SLL and gain for the broadside and scanned patterns by calibrating phase perturbations in both edge and center elements. However, these calibration techniques require array output power measurement and depend on calibration hardware setups. The recent software techniques for pattern correction of faulty linear arrays have been reported in [16-21] and for planar arrays in [22-24]. Non-dominated PSO has been studied for conformal array in [25]. A comprehensive comparison of state-of-the-art techniques for center element failure correction, as discussed in [8-25], is presented in Table 1.

The proposed research employs a sub-optimal novel pre-tuned particle swarm optimization (PT-PSO) algorithm, which is shown to be more efficient than optimal

Table 1: Literature review

Ref	Technique(s) Used	Recovery Type	Array Type	Findings	Limitations
[8-13]	GA/PSO/ IFFT/PSO-CS	Amplitude-only/ Complex weights	Linear/ Planar/ Conformal	SLL, nulls, and pattern shape recovered under random edge elements complete failure	No central complete/partial failure
[14-15]	Phase perturbation-based calibration	Phase-only	Arbitrary array shapes	Compensates SLL, gain, HPBW in corrected pattern due to phase perturbation	Output power measurement Dependent on calibration hardware setup Ineffective for complete TRM failure
[16]	Recursive Intelligent Optimizer (RIO)	Phase-only	Linear	Element failure detection Correction of SLL from -12.5 dB to -18.3 dB with 5 edge and center phase shifter failures	Limited recovery of 4-5 dB in SLL No partial failure of central phase shifters Only linear arrays
[17]	NU-CNLMS, GA	Amplitude-only	Linear	SLL of -10 dB and null recovery under 8 edge and center failures	Limited recovery of 2-3 dB in SLL No partial failure of central phase shifters Only linear arrays
[18]	Improved Bat Algorithm (IBA)	Complex weights	Linear	SLL of -35 dB recovery under 10 edge elements complete failure SLL of -20 dB recovered under complete failures at 3rd, 6th, 13th, 19th, 30th	Central complete failure recovery limited to -20 dB No partial failure of central phase shifters Only linear arrays
[19]	PSO	Complex weights	Linear	Full recovery for edge element failure Beam shape recovery for central failure with SLL up to -15.5 dB	Failed to meet -20 dB target for central complete failures No partial failure of central phase shifters Only linear arrays
[20]	PSO, Taguchi	Complex weights	Linear	Effective edge complete failure recovery With central complete failure SLL to -16 dB recovered	Central complete failure recovery limited to -16 dB Main beam distortion, HPBW nearly doubled under center failure No partial failure of central phase shifters Only linear arrays

Table 1: *Continued.*

Ref	Technique(s) Used	Recovery Type	Array Type	Findings	Limitations
[21]	CUSUM-based method	Amplitude-only and complex weights	Linear/ Planar	Single-bit failure in edge and central attenuators and phase shifters SLL of -25 dB (broadside), and -20 dB (scanned) recovery HPBW recovery in planar under complete failure of 10 elements	No complete failure of central elements Defective elements/bits position not stated No conformal array recovery
[22]	GA, PSO, SA	Complex weights	Planar	Improved SNR under edge and center failures	Did not evaluate SLL, gain, HPBW, or nulls No partial failure of central phase shifters Only planar arrays
[23-24]	GA	Phase-only	Planar	Center faulty phase shifters investigated successfully	No complete central failure
[25]	NS-MOPSO	Phase-only	Conformal	Center faulty phase shifters investigated successfully	No complete central failure
This Work	PT-PSO (Pre-Tuned PSO with Non-Dominated Sorting)	Complex weights	Conformal	SLL of -30 dB recovery under central complete failure while maintaining gain, HPBW Central partial failed phase shifters reported Faster convergence and more efficient than non-dominated PSO in [25] and CVX	Increased sidelobes outside mask in complete failure

convex (CVX) [26] and non-dominated PSO optimization algorithm [25] in terms of lower complexity and faster convergence while maintaining competitive beamforming performance. Unlike prior studies that primarily focus on edge-element failures, linear/planar geometries, or require hardware-based calibration, this work addresses a key research gap: accurate and efficient restoration of beamforming performance under central element failures in conformal arrays using purely software-based excitation optimization. The presented approach focuses solely on reconfiguring the excitations to recover the desired pattern and applicable on a wide variety of phased array systems beyond the specific antenna structure. The chosen cylindrical conformal array introduces geometric non-linearity while maintaining computational tractability, and the method is extensible to more complex array configurations. Two major issues related to the center faulty element(s) in a 1×8 beamforming cylindrical array have been investigated using PT-PSO: the first challenge is to recover the SLL and peak gain in the presence of a complete failure of the center elements, while the second is to restore the original pattern in the case of center faulty phase shifters. Then its performance is compared with CVX and standard PSO algorithms.

II. ELEMENT FAILURE CORRECTION USING PRE-TUNED PSO

The eight-element patch antenna array on a cylindrical conformal surface is shown in Fig. 1. The individual antenna elements are fed with attenuators (to control amplitudes) and phase shifters (to control phases) to obtain the desired beamforming array pattern. In case of any faulty attenuator/phase shifter (shown as red in Fig. 1), the desired beamforming pattern is distorted increasing the SLL and reducing the peak gain. These array parameters are severely affected in case of complete/partial failure of center elements particularly. Unlike traditional linear or planar arrays, the cylindrical conformal array used in this work introduces geometric non-linearity due to its curved surface and tilted element orientation [27]. These characteristics make beamforming control more challenging, thereby providing a more rigorous test environment for evaluating the robustness of the proposed excitation-based optimization method. The flow chart of the proposed PT-PSO for element failure correction on a conformal beamforming array is shown in Fig. 2.

The proposed PT-PSO incorporates two key enhancements: pre-tuning of the initial population and non-dominated sorting [25]. The complex excitation

weight vector is expressed as:

$$\boldsymbol{w} = \begin{bmatrix} |A_1^1|\angle\sigma_1^1 & |A_2^1|\angle\sigma_2^1 & \dots & |A_N^1|\angle\sigma_N^1 \\ |A_1^2|\angle\sigma_1^2 & |A_2^2|\angle\sigma_2^2 & \dots & |A_N^2|\angle\sigma_N^2 \\ \vdots & \vdots & \vdots & \vdots \\ |A_1^P|\angle\sigma_1^P & |A_2^P|\angle\sigma_2^P & \dots & |A_N^P|\angle\sigma_N^P \end{bmatrix}, \quad (1)$$

where A_n and σ_n represents the amplitude and phase on nth antenna respectively. The proposed PT-PSO begins by initializing the parameters N (array size), P (particles), and $MaxIter$ (maximum iterations). Corresponding to each p^{th} particle and n^{th} antenna element, random velocity v_n^p and weight w_n^p vectors are generated. These vectors can be real (for partial element failure) or complex (for complete element failure) as indicated in Fig. 2. In the next step, for the complete failure case, 10% of the population is pre tuned using standard Dolph-Chebyshev for amplitudes and compensated phase technique [28,29] for conformal arrays. After pre-tuning, the PT-PSO evaluates objective functions, updates weight vectors, and personal best and global best weight vectors following the non-dominated PSO method outlined in [25]. The objective functions reported in [25] are incorporated in the proposed PT-PSO as:

$$\mathrm{F} = \overbrace{\left(|AF(\theta_S)_{\text{ref}}| - \left|AF(\theta_S)_{\text{PT-PSO}}\right|\right)^2}^{\text{FSLL recovery}} + \frac{1}{2\times \text{FF}} \sum_{-\text{FF}}^{\text{FF}} [|AF(\theta)_{\text{ref}}| - |AF(\theta)_{\text{PT-PSO}}|]^2, \quad (2)$$

where $AF(\theta_s)_{\text{ref}}$ represents the original reference pattern (without element failure) at the desired scan angle θ_s and $AF(\theta_s)_{\text{PT-PSO}}$ is the corrected array pattern using complex weights (1) computed with PT-PSO in case of complete/partial element failure.

The PT-PSO offers significant improvements over the non-dominated PSO method by incorporating pre-tuning, which minimizes the initial population's fitness values and reduces the computational complexity. Unlike [25], where particles are randomly distributed across the entire search space (leading to frequent trapping in local minima), PT-PSO focuses the search near optimal minima. This refinement accelerates convergence and decreases the number of iterations required.

The integration of amplitude and phase tapering techniques, along with tier-based updates for personal and global bests, allows PT-PSO to achieve faster convergence and more precise solutions for multi-objective problems. In PT-PSO, inertia weight decreases gradually from 0.6 to 0.3 over iterations to maintain precision. Unlike standard PSO, where inertia weight starts at 0.9 (allowing large exploration steps initially), PT-PSO employs pre-tuned populations with a reduced search space, requiring smaller steps to avoid leaving this space and to facilitate convergence toward the optimal solution. The cognitive parameter $c1$ and social parameter $c2$ are set to $c1 = 3$ and $c2 = 0.6$.

III. RESULTS AND DISCUSSION

To evaluate the correcting pattern capability of the proposed PT-PSO, two cases of failure are investigated: (A) complete failure → attenuator(s) connected to the array elements fail and therefore those amplitude(s) are considered as zero, and (B) partial failure → phase shifter(s) connected to the array elements become faulty, generating random phase(s). The -30 dB FSLL and peak gain are the goals to be recovered after failures.

A. Complete failure

In this work, central attenuator no. 4 in Fig. 1 is assumed to be failed and its amplitude is set to zero. This choice is taken to evaluate the performance of the PT-PSO for a more critically power contributing central antenna element in generating the array pattern as compared to the edge element, where the power contribution is minimal. As shown in Fig. 3, the FSLL increases and the entire pattern is distorted due to complete failure of the central element. The mask region (-45° to +45°) is applied, within which the restoration of array pattern is required for most of the real-world practical applications [29]. To compensate for the defected pattern, the PT-PSO are implemented on the faulty array to redistribute the complex weights on the remaining working elements to generate the corrected pattern that meets the criteria of the mask. The original (without failure) and PT-PSO (with failure) recalculated amplitudes/phases are reported in Table A1. As shown in Fig. 3, the recalculated weights with PT-PSO have restored the desired FSLL and peak gain within the desired mask region.

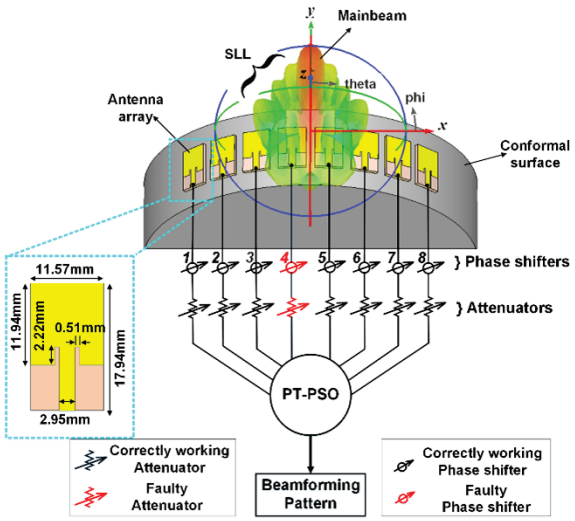


Fig. 1. A 1×8 conformal patch array excited with attenuators and phase shifters.

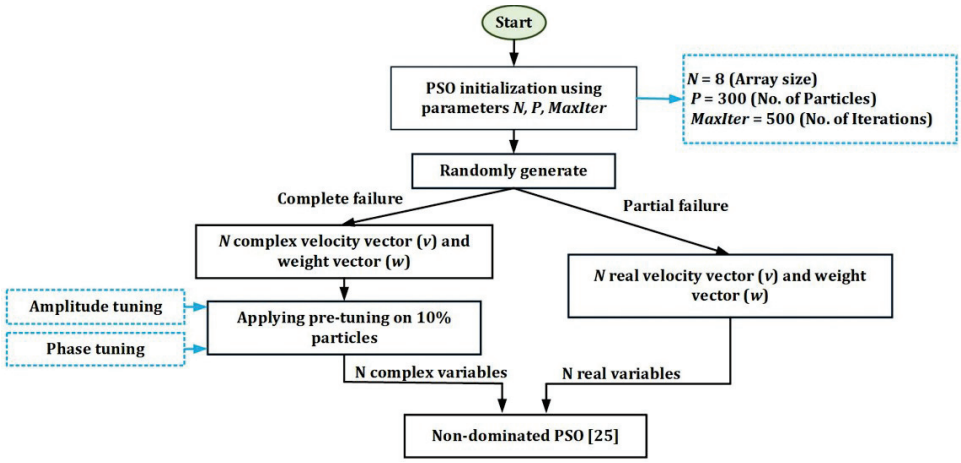


Fig. 2. Flow chart for the proposed PT-PSO.

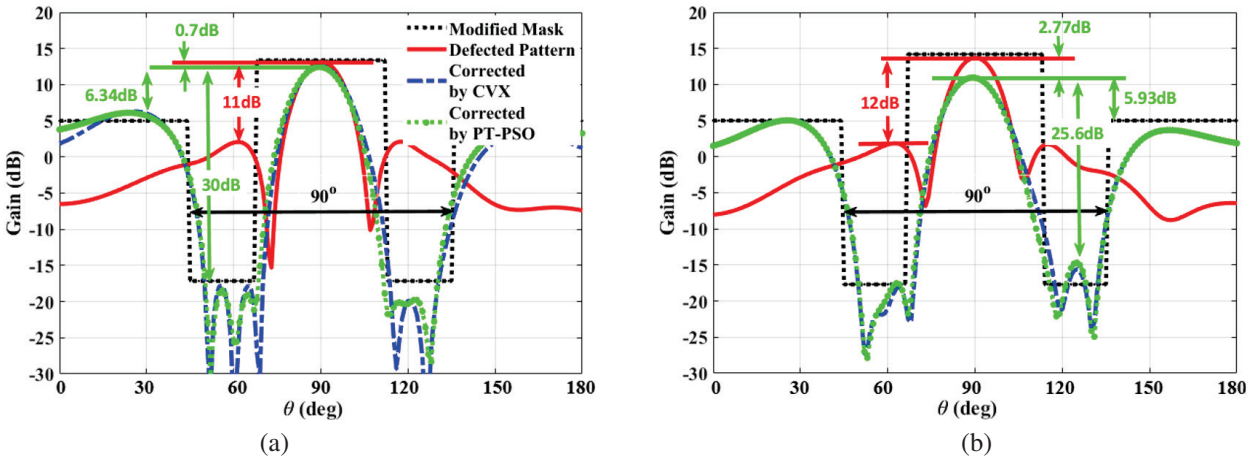


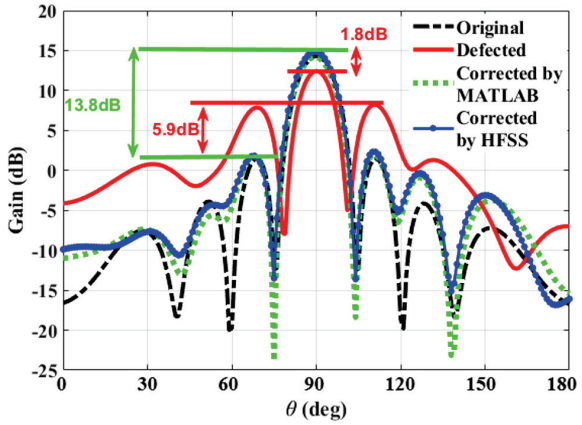
Fig. 3. Simulated pattern recovery of a 1×8 cylindrical array with center (4^{th}) faulty attenuator (a) MATLAB and (b) HFSS.

As expected, the sidelobes outside the desired mask are increased. There is also a 0.7 dB difference in peak gains of the original and corrected array patterns, which cannot be recovered. For comparison, the results obtained with optimal CVX [26] are also shown in Fig. 3. Although the pattern recovering capability of optimal CVX is almost the same as the sub-optimal PT-PSO, its complexity is higher than PT-PSO. This is discussed in section IV. The original and corrected weights for complete failure are reported in Table A1.

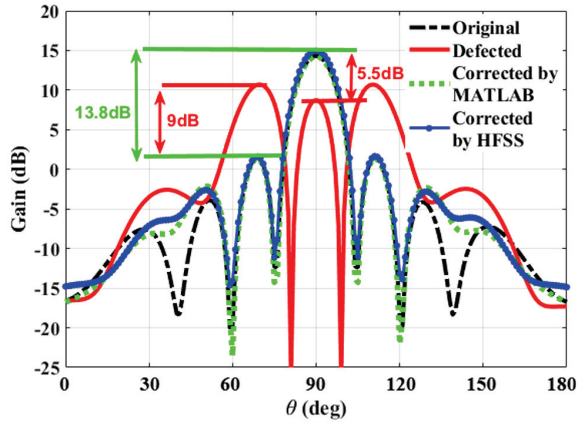
B. Partial failure

For partial failures, single and adjacent central phase shifters are considered. It is assumed that faulty phase shifter(s) can give any random value 0-2π. These random perturbations in original phases causes destructive interference between the antenna element's patterns, which deteriorates the entire array pattern. In this work, central phase shifter 4^{th} and adjacent phase shifters 4^{th} and 5^{th} are considered as faulty, and a phase of π is added to the original phases to demonstrate maximum deformation in the array pattern and the resulting correction capability of PT-PSO. The simulated original, defected and corrected patterns are shown in Fig. 4. As shown in Fig. 4 (a), the gain loss is 1.8 dB and increase in FSLL is 5.9 dB as compared to the original pattern due to single center element phase fault. The PT-PSO has successfully restored the main beam and FSLL to the desired pattern. The corrected phases are computed on the basis of keeping the same phase difference between antenna elements as the phase difference in original phases. This way the corrected patterns in both scenarios have exactly the same gain and FSLL as the original pattern. The case of adjacent phase shifter's fault in Fig. 4 (b) is more severe in terms of gain drop by 5.5 dB and 9 dB increase in FSLL as compared to single phase shifter fault.

The phases recalculated with PT-PSO are able to correct the deformed pattern close to the original pattern.



(a)



(b)

Fig. 4. Simulated pattern recovery of a 1×8 cylindrical array with (a) single center (4th) faulty phase shifter and (b) double adjacent (4th and 5th) faulty phase shifters.

In double faulty phase shifters, when the fault preserves the original progressive phase, the PT-PSO-derived corrected phases successfully restores both gain and FSLL to match the original pattern. When the progressive phase is altered by the faulty phase, PT-PSO adapts by optimizing the phases of the remaining functional phase shifters, thereby reducing FSLL while maintaining the original gain, ensuring robust performance despite hardware limitations. Original and corrected weights are reported in Table A1.

To demonstrate the generality and scalability of the proposed PT-PSO algorithm, we have extended the analysis to include central element failure correction for larger conformal arrays of sizes 1×16, 1×32, 1×64, and 1×128 as shown in Fig. 5. These results clearly illustrate that the proposed approach is effective regardless of array size. In all configurations, central elements were deliberately selected for failure to evaluate the algorithm under the most severe fault scenarios. Specifically, failures were introduced in the following elements:

1×8: 4th element

1×16: 8th and 10th elements

1×32: 16th, 17th, and 19th elements

1×64: 30th, 32nd, 33rd, and 35th elements

1×128: 59th, 62nd, 64th, 65th, 67th, and 70th elements

In each case, the PT-PSO algorithm successfully restored the radiation pattern, achieving SLL≤-30 dB within the SLL region (-45° to +45°). These results support the method's robustness and applicability across a wide range of conformal array sizes.

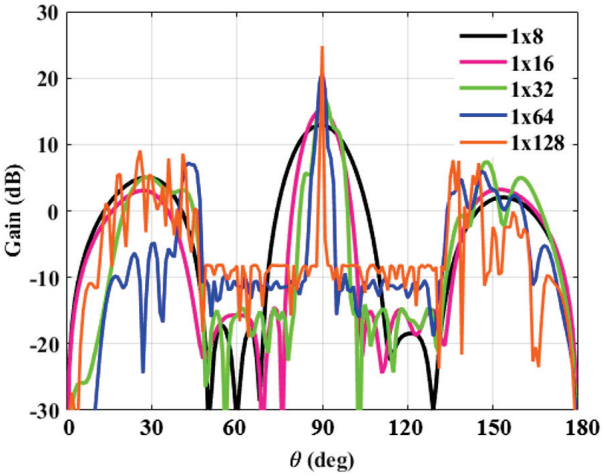


Fig. 5. Center element(s) failure recovery in 1×N conformal arrays.

IV. COMPUTATIONAL COMPLEXITY

In non-dominated PSO [25], the randomly generated particles are distributed across the entire search space, which contains multiple local minima. As a result, particles can become trapped in these local minima, requiring more iterations to achieve a sub-optimal solution in cases of complete/partial failure. In contrast, PT-PSO pre-tunes the randomly initialized particles by recalculating amplitude distribution similar to Dolph-Chebyshev to achieve the desired SLL and compensating the phases to achieve the main beam for failed array. This tuning process reduces the search space, thereby minimizing the number of local minima and decreasing the iterations required to reach the sub-optimal solution. The comparison between non-dominated PSO and PT-PSO is presented in Fig. 6, where both methods were initialized with the same population.

For non-dominated PSO, the convergence began with an initial objective function value (F) of 2.96 dB using (2) and reached a sub-optimal value of -30.75 dB

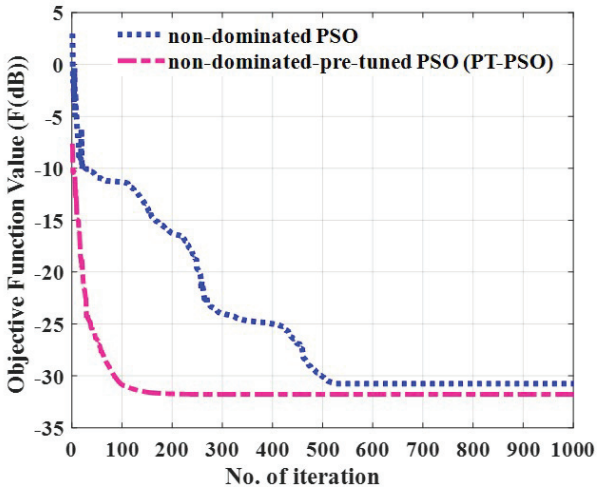


Fig. 6. Objective function value (F) comparison with no. of iteration for non-dominated PSO [25] and proposed PT-PSO.

after 529 iterations. In contrast, PT-PSO, after tuning the weight vectors, started with an initial F value of -5 dB and achieved convergence much faster, requiring only 150 iterations to reach sub-optimal F value of -31.57 dB. This demonstrates that PT-PSO not only converges more quickly but also reduces the chance of particles becoming stuck in local minima, significantly lowering the computational time by decreasing the number of iterations required.

Figure 7 compares the computational complexity of three algorithms: CVX, non-dominated PSO, and PT-PSO. The CVX algorithm employs Newton's method to

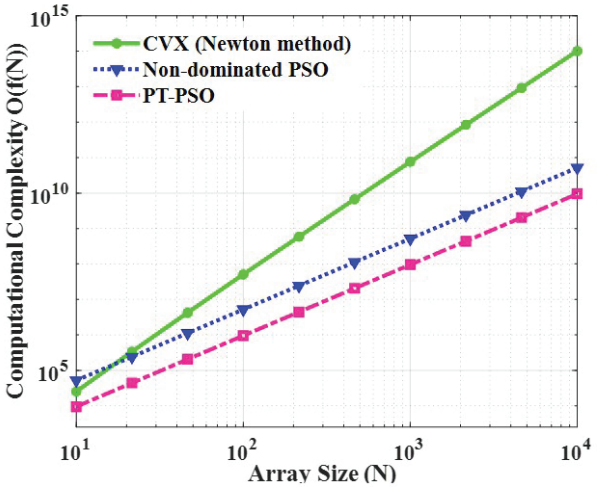


Fig. 7. Computational complexity ($O(f(N))$ comparison of CVX (Newton method), non-dominated PSO [25], and PT-PSO with respect to N.

iteratively solve the Karush-Kuhn-Tucker (KKT) conditions derived from the constraints and objective functions. Newton's method optimizes search directions and step sizes based on penalty functions, making it efficient for problems with well-structured Hessians. The computational complexity of CVX depends on the array size (N) and the handling of the Hessian matrix, resulting in a complexity of $O(iterations \cdot N^3 \cdot log(1/N))$, primarily driven by the matrix calculations in each iteration. In contrast, non-dominated PSO and PT-PSO rely on sorting particles into Pareto-optimal tiers, with a computational complexity of $O(N^2)$ for the sorting process. Consequently, their overall complexity is $O(N^2 \cdot iterations)$, where the number of iterations significantly impacts their performance. This process continues until the stopping criteria is met.

As N grows, the complexity of all three algorithms increases. However, the results shown in Fig. 7 reveal that PT-PSO has the lowest complexity compared to non-dominated PSO and CVX. Particularly, CVX required 11 iterations, non-dominated PSO required 529 iterations, and PT-PSO converged in just 151 iterations. PT-PSO's efficiently optimizes the weight vector while avoiding the computationally steep matrix operations inherent in CVX. Furthermore, PT-PSO achieves faster convergence compared to non-dominated PSO, further reducing its overall complexity. For high-dimensional problems, PT-PSO demonstrates a significant improvement in computational complexity among the three.

To benchmark the performance of the proposed PT-PSO method, a comparison is provided with widely adopted metaheuristic techniques reported in the literature, including GA [17,22-24], IBA [18], Simulated Annealing (SA) [22], and standard PSO [19-20,22,25]. These methods have been primarily applied to linear and planar arrays under various failure scenarios. As summarized in Table 1, PT-PSO achieves a SLL improvement of 21 dB, restoring the corrected SLL to -30 dB, along with enhanced main beam recovery, reduced gain loss, and faster convergence (particularly in cases involving both complete and partial failures of central elements) where traditional algorithms often degrade in performance. Additionally, the pre-tuned initialization significantly reduces computational cost, making PT-PSO highly suitable for real-world conformal array applications involving complex geometries.

V. CONCLUSION

This study presented a pre-tuned particle swarm optimization (PT-PSO) algorithm for fault compensation in conformal phased arrays, addressing both complete element failure and phase shifter faults. By combining pre-tuned initialization with non-dominated sorting, PT-PSO achieved efficient SLL and gain recovery with faster

convergence and lower computational complexity than [25] and convex methods. Validation on a 1×8 cylindrical conformal array showed successful SLL restoration to -30 dB and minimal gain loss under severe central failures. The method is scalable to larger arrays and adaptable to conformal geometries. Its primary limitation is increased SLLs outside the constrained angular region in complete failure scenarios. Future work includes hybrid software-hardware SLL suppression using RIS, extension to complex conformal surfaces (e.g., spherical arrays), and compensation for amplifier-related nonlinearities and gain loss. These enhancements position PT-PSO as a scalable and fault-resilient solution for advanced beamforming applications in radar, communication, and medical imaging systems.

ACKNOWLEDGMENTS

This work is funded by the Higher Education Commission (HEC) Pakistan via Project No.20-17554/NRPU/R&D/HEC/2021.

REFERENCES

[1] G. Fabrizio, F. Colone, P. Lombardo, and A. Farina, "Adaptive beamforming for high-frequency over-the-horizon passive radar," *IET Radar, Sonar & Navigation*, vol. 3, no. 4, pp. 384-405, 2009.

[2] Y. Xu, X. Shi, A. Wang, and J. Xu, "Design of sum and difference patterns with common nulls and low SLLs simultaneously in the presence of array errors," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 2, pp. 934-944, 2018.

[3] Y. Albagory, "An efficient conformal stacked antenna array design and 3D-beamforming for UAV and space vehicle communications," *Sensors*, vol. 21, no. 4, p. 1362, 2021.

[4] P. Alinezhad and S. R. Seydnejad, "Broadband adaptive beamforming of conformal arrays for wireless communications based on generalized sidelobe canceller," *Wireless Pers. Commun.*, vol. 96, pp. 1131-1143, 2017.

[5] H. Schippers, J. Verpoorte, P. Jorna, A. Hulzinga, A. Meijerink, C. Roeloffzen, R.G. Heideman, A. Leinse, and M. Wintels, "Conformal phased array with beam forming for airborne satellite communication," in *Proc. Int. ITG Workshop on Smart Antennas*, Darmstadt, Germany, pp. 343-350, Feb. 2008.

[6] Y. Yu, Z. H. Jiang, H. Zhang, Z. Zhang, and W. Hong, "A low-profile beamforming patch array with a cosecant fourth power pattern for millimeter-wave synthetic aperture radar applications," *IEEE Trans. Antennas Propag.*, vol. 68, no. 9, pp. 6486-6496, Sep. 2020.

[7] F. Wang, T. Arslan, and G. Wang, "Breast cancer detection with microwave imaging system using wearable conformal antenna arrays," in *Proc. IEEE Int. Conf. Imaging Systems and Techniques* (*IST*), Beijing, China, pp. 1-6, Oct. 2017.

[8] H. Munsif, R. A. B. Saleem, A. A. Shah, S. Khattak, A. I. Najam, B. D. Braaten, and I. Ullah, "Beamforming array failure correction for mm-wave synthetic aperture radar applications," in *Next Generation Wireless Communication: Advances in Optical, mm-Wave, and THz Technologies,* M. El Ghzaoui, S. Das, V. Samudrala, and N. R. Medikondu, Eds. Cham: Springer Nature Switzerland, pp. 105-129, 2024.

[9] O. P. Acharya and A. Patnaik, "Antenna array failure correction," *IEEE Antennas and Propagation Magazine*, vol. 59, no. 6, pp. 106-115, June 2017.

[10] H. Munsif, B. D. Braaten, and I. Ullah, "Sidelobe level and nulls control of defected linear antenna arrays using an optimization algorithm," in *19th International Bhurban Conference on Applied Sciences and Technology* (*IBCAST*)*,* pp. 984-989, 2022.

[11] H. Munsif, B. D. Braaten, and I. Ullah, "Element failure correction techniques for phased array antennas in future terahertz communication systems," in S. Das, N. Anveshkumar, J. Dutta, and A. Biswas, Eds. *Advances in Terahertz Technology and its Applications*, Singapore: Springer Singapore. pp. 19-46, 2021.

[12] H. A. Malhat, A. S. Zainud-Deen, M. Rihan, and M. M. Badway, "Elements failure detection and radiation pattern correction for time-modulated linear antenna arrays using particle swarm optimization," *Wireless Personal Communications*, vol. 125, no. 3, pp. 2055-2073, 2022.

[13] B. Cao, Z. Xu, J. Yao, and T. Jiang, "Conformal array element failure correction based on PSO-CS algorithm," in *2021 International Applied Computational Electromagnetics Society* (*ACES-China*) *Symposium*, pp. 1-2, 2021.

[14] V. Djigan and V. Kurganov, "Antenna array calibration algorithm based on phase perturbation," in *IEEE East-West Design & Test Symposium* (*EWDTS*), pp. 1-5, 2019.

[15] V. I. Djigan and V. V. Kurganov, "Antenna array calibration algorithm without access to channel signals," *Radioelectronics and Communications Systems*, vol. 63, no. 1, pp. 1-14, 2020.

[16] Z. Hamici, "Fast beamforming with fault tolerance in massive phased arrays using intelligent learning control," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 7, pp. 4517-4527, 2019.

[17] J. Xiaochao, P. Xu, and T. Jiang, "Comparison between NU-CNLMS and GA for antenna array failure correction," in *2018 IEEE International*

Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting, pp. 2205-2206, 2018.

[18] N. S. Grewal, M. Rattan, and M. S. Patterh, "A linear antenna array failure correction using improved bat algorithm," *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 27, no. 7, p. e21119, Sep. 2017.

[19] L. A. Greda, A. Winterstein, D. L. Lemes, and M. V. T. Heckler, "Beamsteering and beamshaping using a linear antenna array based on particle swarm optimization," *IEEE Access*, vol. 7, pp. 141562-141573, 2019.

[20] A. M. Engroff, L. A. Greda, M. P. Magalhães, A. Winterstein, L. S. Pereira, A. G. Girardi, and M. V. T. Heckler, "Comparison of beamforming algorithms for retro-directive arrays with faulty elements," in *2016 10th European Conference on Antennas and Propagation* (*EuCAP*), pp. 1-5, 2016.

[21] Y.-S. Chen and I.-L. Tsai, "Detection and correction of element failures using a cumulative sum scheme for active phased arrays," *IEEE Access*, vol. 6, pp. 8797-8809, 2018.

[22] N. Boopalan, A. K. Ramasamy, F. Nagi, and A. A. Alkahtani, "Planar array failed element(s) radiation pattern correction: A comparison," *Applied Sciences*, vol. 11, no. 19, p. 9234, 2021.

[23] R. A. B. Saleem, A. A. Shah, H. Munsif, A. I. Najam, S. Khattak, and I. Ullah, "Pattern compensation of a planar phased array with centre elements phase malfunctioning using a genetic algorithm," *Progress in Electromagnetics Research Letters*, vol. 122, pp. 21-28, 2024.

[24] R. A. B. Saleem, A. A. Shah, H. Munsif, A. I. Najam, S. Khattak, and I. Ullah, "Radiation pattern correction of faulty planar phased array using genetic algorithm," *Advanced Electromagnetics*, vol. 13, no. 2, pp. 23-31, 2024.

[25] H. Munsif and I. Ullah, "Malfunctioning conformal phased array: Radiation pattern recovery with particle swarm optimisation," *IET Microwaves, Antennas & Propagation*, vol. 18, no. 9, pp. 654-666, 2024.

[26] S. Boyd, *Convex Optimization*. *Cambridge: Cambridge University Press*, 2004.

[27] I. Ullah, H. Munsif, S. Razzaq, and A. I. Najam, "Cylindrical phased array with adaptive nulling using eigen-correlation technique," *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 32, no. 2, p. e22969, 2022.

[28] R. L. Haupt, *Antenna Arrays: A Computational Approach*. Hoboken, NJ: John Wiley & Sons, 2010.

[29] M. Li, S.-L. Chen, Y. Liu, and Y. J. Guo, "Wide-angle beam scanning phased array antennas: A review," *IEEE Open Journal of Antennas and Propagation*, vol. 4, pp. 695-712, 2023.

APPENDIX

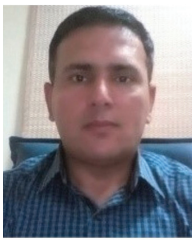
Table A1: Optimized excitation weights (amplitude/phase and phase-only) for complete and partial element failures using PT-PSO and CVX algorithms

Broadside Pattern with -30 dB FSLL						**Element Phase Failure Corrected by PT-PSO**				
Original		**Corrected CVX**		**Corrected PT-PSO**		**Single Center (4th)**			**Adjacent Center (4th and 5th)**	
Amp	**Phases (°)**	**Amp**	**Phases (°)**	**Amp**	**Phases (°)**	**Original (°)**	**Defected (°)**	**Corrected (°)**	**Defected (°)**	**Corrected (°)**
0.2710	186.15	0.3208	-167.32	0.4012	91.03	215	215	52	215	30
0.4926	148.33	0.0523	-23.54	0.0063	233.94	176	176	0	176	3
0.8128	127.10	0.8960	144.29	1.0000	37.33	155	155	12	155	321
1.0000	114.56	0	0	0	0	142	142+180	142+180	142+180	142+180
1.0000	113.32	1.0000	131.34	1.0000	24.97	142	142	331	142+180	142+180
0.8321	128.09	0.0231	-100.02	0.0697	168.56	155	155	1	155	321
0.5275	148.96	0.4592	158.01	0.4362	53.41	176	176	29	176	3
0.2946	187.01	0.0466	-27.32	0.0412	233.58	215	215	53	215	30

Amp=amplitude

Red numbers represent the amplitudes/phases of failed/faulty elements

Hina Munsif holds a bachelor's degree in Electrical (Communication) Engineering in 2015 from the University of Engineering and Technology Peshawar, KPK, Pakistan, and a master's degree in Electrical Engineering in 2020 from COMSATS University Islamabad, Abbottabad Campus, Pakistan. She is currently pursuing a Ph.D. degree in Electrical Engineering from COMSATS University Islamabad, Abbottabad Campus. Her research interests include antenna arrays for 6G-and-beyond communication and IoT.



Irfan Ullah received a Ph.D. degree in Electrical and Computer Engineering from North Dakota State University, Fargo, ND, USA, in 2014. He is an Associate Professor in Electrical Engineering Department at COMSATS University Islamabad, Abbottabad Campus, Pakistan. His research interests include beamforming arrays, machine learning in antenna arrays, electromagnetic metamaterials and topics in EMC.

Shahid Khattak received Dr.-Ing degree from Technische Universität Dresden, Germany, in 2008. He is a distinguished academic and researcher specializing in electrical and computer engineering, with a focus on advanced communication systems, signal processing, and applied electromagnetics. He has contributed extensively to both theoretical and applied research, with numerous publications in reputed international journals. Khattak is known for his dedication to academic excellence and mentoring, and he plays a pivotal role in advancing interdisciplinary research in engineering and technology.

Shafqat Ullah Khan received M.S. and Ph.D. degrees in Electronic Engineering from International Islamic University Islamabad and ISRA University, Pakistan, in 2008 and 2015, respectively. He was a Post Doc Fellow at the Faculty of Electrical Engineering, University Technology Malaysia from 2016 to 2017. He is an Associate Professor at the University of Buner. His research interest includes RF & microwave, antenna arrays and evolutionary algorithms.

Metamaterial-loaded Circularly Polarized Quad-band SIW MIMO Antenna with Improved Gain for Sub-6 GHz and X-band Applications

R. Anandan[1], Sathyasri Balasubramanian[2], Ravi Kumar Sanapala[3], and Prakash Kamisetti[4]

[1]Department of ECE, Dhanalakshmi Srinivasan College of Engineering and Technology
Mamallapuram, Chennai, Tamilnadu, India
anandandscet@gmail.com

[2]Department of ECE, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
Avadi Chennai, Tamilnadu, India
dr.sathyasrib@gmail.com

[3]Department of ECE, St. Martin's Engineering College
Dhulapally, Secunderabad, Telangana, India
sravikumarece@smec.ac.in

[4]Department of Electrical and Electronics Engineering, Vaagdevi College of Engineering (UGC Autonomous),
Bollikunta, Warangal, Telangana, India
prakashkams@gmail.com

Abstract **–** This work presents a quad-band metamaterial-loaded cavity-backed substrate integrated waveguide (SIW) MIMO antenna engineered for sub-6 GHz communication standards such as 5G and WLAN, as well as X-band applications. The use of a cavity-backed SIW architecture enables reduced fabrication complexity and straightforward integration with planar circuits, supporting compact and efficient antenna design. The antenna structure incorporates a modified rectangular split ring resonator (RSRR) slot along with an open-ended rectangular slot embedded within the SIW cavity-backed radiator. This configuration generates four resonant bands operating at 2.4, 3.3, 5.0, and 7.0 GHz. To enhance radiation characteristics, modified RSRR-based metamaterial unit cells are placed along the y-direction in front of the radiating elements. These cells contribute significantly to gain enhancement and enable circular polarization at the designated frequencies. The proposed antenna demonstrates realized gains of 5 dB, 8 dB, 6 dB, and 5 dB at the respective bands, supported by a consistent radiation efficiency of approximately 88%. The antenna also exhibits a stable unidirectional radiation pattern across all operating frequencies, making it suitable for directional multi-port MIMO configurations. To suppress inter-element interference, a cavity-backed parasitic structure is introduced, effectively reducing mutual coupling between radiators. Comprehensive MIMO performance analysis is carried out using standard metrics, including envelope correlation coefficient (ECC), total active reflection coefficient (TARC), channel capacity loss (CCL), and mean effective gain (MEG), confirming strong isolation and diversity capability. Experimental validation aligns closely with simulation results, establishing the proposed antenna's reliability and potential for use in high-performance, multi-band wireless systems.

Index Terms **–** Cavity-backed SIW, gain enhancement, metamaterial-loaded antenna, quad-band antenna, sub-6 GHz communication.

I. INTRODUCTION

In the current era of wireless communication technology and 5G, where wireless devices are rapidly evolving, there is a significant demand for enhanced data transmission capacity and link reliability [1]. Researchers have suggested multiple approaches to address demand, and MIMO technology has emerged as one of the most effective ways to improve data rates [1]. MIMO antennas, which stand for multiple-input multiple-output, are able to boost channel capacity and link reliability without using additional resources [2]. On the other hand, circularly polarized (CP) antennas effectively battle multipath fading, which results in minimal propagation losses [3]. Because of this, circularly polarized MIMO antennas are essential for the operation of contemporary wireless networks.

Submitted On: December 31, 2024
Accepted On: May 14, 2025

https://doi.org/10.13052/2025.ACES.J.400805

Circularly polarized wideband MIMO antennas have been developed [3]; nevertheless, circularly polarized multi-band antennas are more appropriate for contemporary communication systems. Numerous methodologies for attaining circular polarization have been reported in the literature. Circular polarization is attained by employing variously shaped slots and the deliberate positioning of a patch on these slots in [4–6]. The circular polarization is attained via a metamaterial superstate in [7, 8]. Further, to get minimal mutual coupling, high gain, and compact dimensions, numerous design issues must be addressed. In recent years, researchers have introduced various MIMO antennas without circular polarization and with circular polarization in the literature [9–26].

The substrate-integrated waveguide (SIW) cavity technology has emerged in the last decade, offering various applications and innovative strategies for existing uses. The incorporation of the SIW enables the simultaneous design of an antenna and its corresponding circuitry on a single plane, hence improving dependability relative to conventional circuitry, which occupies limited board space. To enhance the bandwidth of the SIW cavity-backed slot antenna, researchers have explored several ways over several years [27, 28].

In environments with significant dispersion, such as industries, a high-gain antenna is essential. The antenna's gain is improved by the utilization of multilayer substrates [29], metamaterials [30], and dielectric resonators [31], as reported in the literature. The integration of metamaterials into various antennas substantially enhances antenna gain performance [32]. The metamaterial array is situated in front of the antenna to boost the gain [32, 33]. Furthermore, the antenna gain is enhanced by situating the metamaterial array above the antenna in [34]. The antennas above demonstrate various methods for circularly polarized MIMO antenna implementation, mutual coupling reduction, bandwidth enhancement, and gain enhancement, but few sub-6 GHz MIMO antennas offer a wider bandwidth and higher gain in a compact planar configuration. A metamaterial layer superstate improved gain in stated antennas, but also made them bulky. This paper proposes a sub-6GHz circularly polarized metamaterial-loaded quad-band cavity-backed substrate integrated waveguide (SIW) MIMO antenna for 5G, WLAN, and X-band applications to address these issues. section II covers MIMO antenna design and development, section III covers experimental results, and section IV covers diversity performance. The proposed antenna is compared to existing works in section-V and concluded in section-VI.

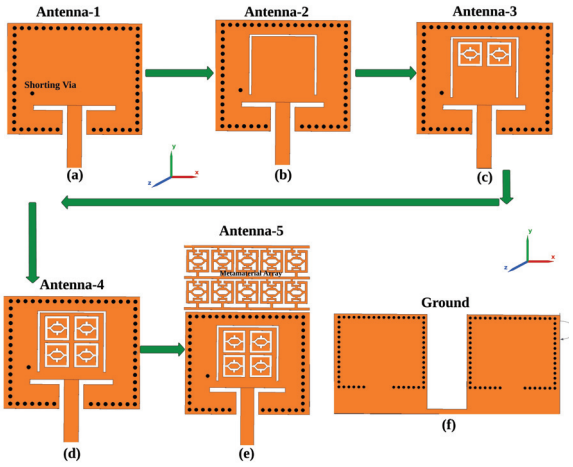


Fig. 1. (a)-(f) Development of the proposed SIW cavity backed radiator.

II. MIMO ANTENNA STRUCTURE AND ITS EVOLUTION

A. SIW cavity-backed radiator evolution

The configuration of the proposed SIW cavity-backed slot antenna is depicted in Fig. 1 (a)-(f). The design consists of a square-shaped SIW cavity, formed by the arrangement of metal vias in a square configuration, as illustrated in Fig. 1 (a). The cavity dimensions are determined based on equation (1) and (2), which induce first-order resonance at approximately 5 GHz as depicted in Fig. 2.

$$f_{SIW} = \frac{c}{2\sqrt{\varepsilon_r}}\sqrt{\left(\frac{1}{L_{\text{eff}}}\right)^2 + \left(\frac{1}{W_{\text{eff}}}\right)^2}, \tag{1}$$

$$L_{\text{eff}} = L_s - \frac{d^2}{0.95s}, \quad W_{\text{eff}} = W_s - \frac{d^2}{0.95s}. \tag{2}$$

The SIW cavity is created by embedding metal posts within the dielectric substrate, which serves as the lateral conducting walls, as depicted in Fig. 1 (a). The diameter (d) and pitch of metallic vias (S) are determined with the criteria of $d/s \geq 0.5$ and $d/\lambda_0 \leq 0.1$ to ensure minimal energy loss from the cavity sidewalls. To enhance the bandwidth of the SIW cavity radiator, a shorting pin is positioned 3 mm from the left side of the cavity and 8 mm from the top wall of the cavity, as depicted in Fig. 1 (a). The shorting pins create a short circuit in the current path, resulting in resonance. The enhancement in bandwidth is accomplished by stimulating two hybrid modes: resonance from the SIW cavity and shorting pins positioned in close proximity to each other. Initially, an inverted C-shaped radiating slot has been removed from the center of the SIW cavity, positioned 2 mm from the upper sidewalls and 5 mm from the left and right sidewalls of the cavity, as illustrated in Fig. 1 (b). The positioning of the inverted C-slot has a significant impact

on the current path of the first-order resonance mode, resulting in a shift of the corresponding resonant frequency towards the lower end, approximately 3.1 GHz as depicted in Fig. 2. Furthermore, a second-order resonance is generated at 4.1 GHz due to the reactive loading effect of the C-slot, as illustrated in Fig. 2. The integration of the C-slot leads to a decrease in series inductance; however, it simultaneously introduces series capacitance within the cavity circuit.

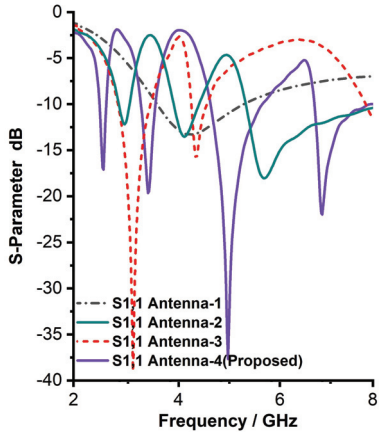


Fig. 2. Reflection coefficient of the proposed series-fed dipole radiator.

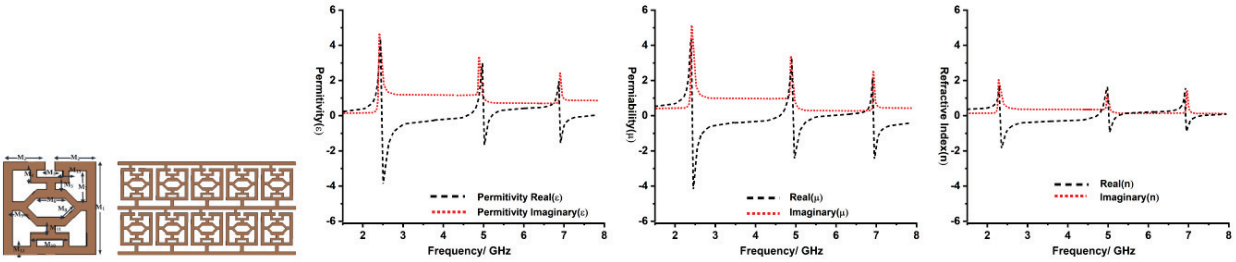
Further, the two modified SSRR-shaped radiating slots (modified SSRR slot-1 and slot-2) have been removed from the centre of the SIW cavity-backed radiator, positioned at a distance of 2mm from the upper sidewalls of the inverted C-shaped slot and 5mm from the left and right sidewalls of the inverted C-shaped slot, as depicted in Fig. 1 (c). The placement of the C-slot significantly affects the current path for first and second order resonance, resulting in altering the resonant frequencies of 2.9 GHz and 4.4 GHz. The third-order resonance occurs at 5.5 GHz due to the reactive loading effect of the modified SSRR slot-1 and 2 placement, as illustrated in Fig. 2.

Additionally, two modified SSRR-shaped radiating slots (modified SSRR slot-3 and slot-4) have been removed from the centre of the SIW cavity backed radiator, positioned 8 mm from the upper sidewalls of the inverted C-shaped slot and 5 mm from the left and right sidewalls, as depicted in the Fig. 1. The arrangement of the third and fourth modified SSRR-shaped slots significantly change the current pathways of the first, second, and third order mode resonances. Consequently, the resonant frequencies for these modes are adjusted to 2.4 GHz, 3.3 GHz, and 5 GHz, respectively, while the fourth-order resonance occurs at 7 GHz due to the reactive loading effect of the modified SSRR slots 3 and 4, as illustrated in Fig. 2. Ultimately, the proposed radiator, featuring an inverted C-shaped slot and modified SSRR slots, generates four resonance modes at 2.4 GHz, 3.3 GHz, 5 GHz, and 7 GHz, suitable for WLAN, sub-6 GHz 5G, and X-band applications as depicted in Fig. 2.

B. Metamaterial for gain enhancement and circular polarization conversion

The proposed modified SSRR metamaterial units are loaded in-front of the unit cell radiator to improve gain and provide polarization conversion. The proposed modified SSRR unit cell and rectangular connected modified SSRR metamaterial 5×2 array are depicted in Fig. 3 (a). The proposed metamaterial unit cell consists of a hexagonal shaped ring stub, which is placed inside the the modified SSRR split-ring stub as depicted in Fig. 3 (a). Further, the outer SSRR and inner hexagonal-shaped ring connected through the rectangular-shaped stub as depicted in Fig. 3 (a). The metamaterial used in the proposed design consists of 5 × 2 modified SSRR unit cells, and all the unit cells are connected using horizontal and vertical rectangular stubs. Further, 5 × 2 modified SSRR unit cells are printed in front of the antenna in end-fire directions, as illustrated in Fig. 1 (e). The 0.5 mm periodic distance maintained in x- and y-direction between the two modified SSRR unit cells. Performance analysis of the metamaterial is carried out by placing the metamaterial unit cell between two waveguide ports, and different boundary conditions (perfect electric conductor (PEC) and perfect magnetic conductor (PMC)) are applied in x- and y-axis. Further, entire characterization is carried out in CST Microwave Studio software. The y-axis is the direction in which an electromagnetic wave that is propagating normally is directed. On the y-axis, the typical trajectory of the electromagnetic wave's incidence is aligned with the y-axis. The electromagnetic interactions that take place within the modified SSRR metamaterial unit cell, which are driven by the incident waves, are responsible for the generation of resonance in both the transmitted and reflected waves. The y-axis is responsible for the facilitation of the excitation of the modified SSRR metamaterial unit cell. This is because the electromagnetic wave starts from the SIW cavity-supported antenna and propagates into the unit cells in that direction (y).

The permeability, permittivity, and refractive index are critical parameters for precisely evaluating the performance characteristics of the proposed metamaterial unit cell. The permeability, permittivity, and refractive index of the proposed metamaterial unit cell are depicted in Figs. 3 (b)-(d). The figure illustrates that the suggested metamaterial exhibits a negative real near-zero refractive index (NZRI) and epsilon-negative (ENG) characteristics at 2.4, 3.3, 5, and 7 GHz frequency ranges. The findings indisputably illustrate that the proposed SSRR metamaterial exhibits near-zero properties within



(a) (b) (c) (d)

Fig. 3. Analysis of the proposed SSRR metamaterial unit cell: (a) Unit-cell structure and metamaterial array, (b) permittivity, (c) permeability, and (d) refractive index.

the operational bands. Consequently, this specific operating frequency zone (2.4, 3.3, 5, and 7 GHz) can be utilized to boost gain and modify the polarization of the proposed antenna.

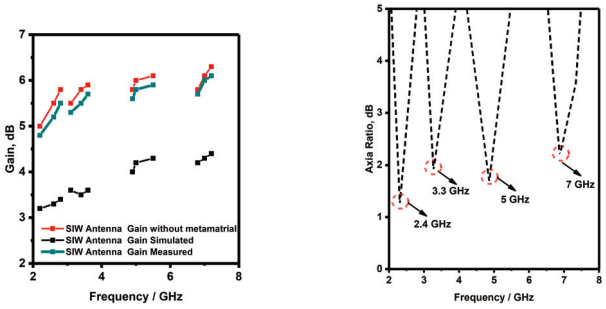


Fig. 4. Comparison of (a) Gain and (b) Axial Ratio of the proposed radiator with and without metamaterial.

The gain of the proposed antenna, both with and without metamaterial, is depicted in Fig. 4 (a). Figure 4 indicates that the proposed antenna exhibits gains of 3.3 dB, 2.8 dB, 3.1 dB, and 4.1 dB at frequencies of 2.4 GHz, 3.3 GHz, 5 GHz, and 7 GHz, respectively. The proposed antenna exhibits a gain of 5.5 dB, 8 dB, 6 dB, and 8 dB at frequencies of 2.4 GHz, 3.3 GHz, 5 GHz, and 7 GHz, respectively, subsequent to the incorporation of the metamaterial unit cells. The gain of the proposed antenna with SSRR metamaterial loading has been substantially enhanced without altering the resonant frequencies, as shown in Fig. 4 (a). In addition, metamaterial, in general, has the ability to alter the polarization. The proposed metamaterial unit cells are positioned in front of the antenna in the y-direction, which results in the transformation of the linear polarization into the circular polarization. In addition, the proposed metamaterial is capable of performing the function of a linear-to-circular polarization converter at the 2.4, 3.3, 5, and 7 GHz band, as shown by the axial ratio characteristic depicted in Fig. 4 (b). The axial of ratio of the proposed antenna is greater than three decibels at all of the working frequencies (2.4, 3.3, 5, and 7 GHz), as can be seen in Fig.4 (b).

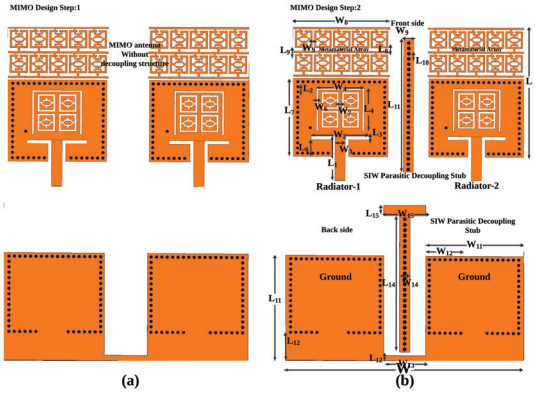


Fig. 5. Proposed MIMO antenna (a) without cavity backed parasitic stub decoupling structure, (b) With cavity backed parasitic stub decoupling structure.

C. Two-port MIMO antenna design

The antenna should have a high channel capacity, a high gain, and a good link reliability in order to facilitate communication in an environment that has a significant pathloss owing to scattering. As a result, we suggested the two-port circularly polarized metamaterial-loaded quad-band SIW MIMO antenna for use in sub-6GHz 5G, WLAN, and X-band applications in industries. The proposed circularly polarized MIMO antenna is designed by placing two SIW based metamaterial loaded radiators adjacent to each other, sharing a common ground plane, as illustrated in Figs. 5 (a)-(b).

The dimensions of the proposed MIMO antenna are shown in Table 1. The proposed MIMO antenna with dimensions is depicted in Fig. 5. As depicted in Fig. 5 the side-by-side placement design alleviates the mutual coupling between radiators; however, it is still inadequate for MIMO operation. Therefore, a separate

Table 1: Dimension of the proposed MIMO antenna

P=D	P=D	P=D
L = 40	L_{11}=17	W_6=0.3
L_1=6.8	L_{12}=6.4	W_7=0.6
L_2=0.4	L_{13}=1.8	W_8=15.8
L_3=0.6	L_{14}=22.8	L_s=W_9=0.3
L_4=6.5	R=0.5	W_{10}=1.8
L_5=0.3	W=35	W_{11}=5.6
L_6=3	W_1=14.8	W_{12}=15.8
L_7=12.8	W_2=3	W_{13}=1.8
L_8=0.6	W_3=9.2	W_{14}=31
L_9=22.8	W_4=7.4	W_{15}=6
L_{10}=0.3	W_5=1	W_{16}=6.2

decoupling structure is required to diminish the mutual coupling between the radiators. Consequently, we proposed a SIW cavity-backed parasitic stub decoupling structure. The SIW-based cavity-backed parasitic rectangular stub is positioned on both sides of the substrate, situated between the SIW-based radiator as depicted in Fig. 5. The SIW-based cavity-backed parasitic decoupling structure comprises parasitic rectangular stubs on both sides of the substrate in parallel. The SIW cavity is created by positioning circular shaped vias in a vertical configuration, which connect the front and back conductors as depicted in Fig. 5. The parasitic stub SIW cavity functions as a stop band filter between the radiators of the MIMO antenna. Due to these characteristics, the SIW cavity obstructs the flow of surface current between two simultaneously stimulated radiators. The proposed antenna effectively minimizes mutual coupling between radiating elements through the incorporation of a cavity-backed substrate-integrated waveguide (SIW) decoupling structure. This integrated SIW cavity not only enhances isolation but also preserves compactness, making it highly suitable for multi-port MIMO configurations.

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. Reflection coefficient and mutual coupling

The VNA measurement configuration for the proposed MIMO antenna is illustrated in Fig. 6. The comparison of the simulated and measured reflection coefficient and mutual coupling of the proposed MIMO antenna is illustrated in Figs. 7 (a)-(b).

Figure 7 (a) illustrates that the proposed MIMO antenna exhibits a reflection coefficient (S11/S22) of less than -10 dB at frequencies of 2.4, 3.3, 5, and 7 GHz, both for the simulated and measured results. Similarly, Fig. 7 (b) illustrates that the proposed MIMO antenna exhibits mutual coupling (S12/S21) of less than -15 dB

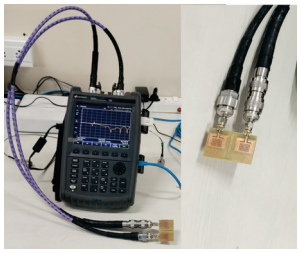
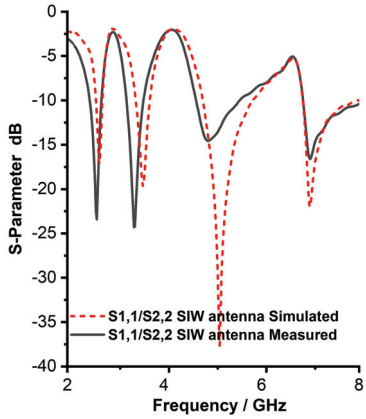
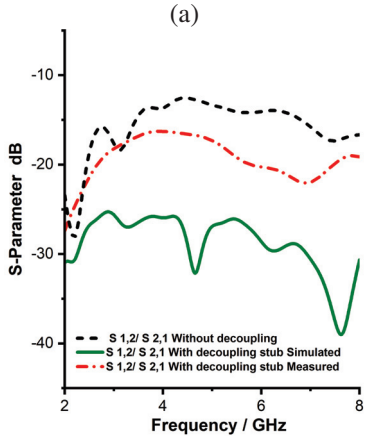


Fig. 6. Photograph of the fabricated antenna and measurement setup of the proposed MIMO antenna.



(a)



(b)

Fig. 7. Performance analysis of the proposed MIMO antenna: (a) Reflection coefficient, and (b) Mutual coupling.

for frequencies of 2.4, 3.3, 5, and 7 GHz, as evidenced by both the simulated and measured results. The simulated and measured results are in strong agreement.

B. Radiation pattern and efficiency

The radiation pattern measuring setup of the proposed MIMO antenna is illustrated in Fig. 8. The E- and H-plane 2D cross- and co-polarization radiation patterns of the proposed antenna at 2.4, 3.3, 5, and 7 GHz are

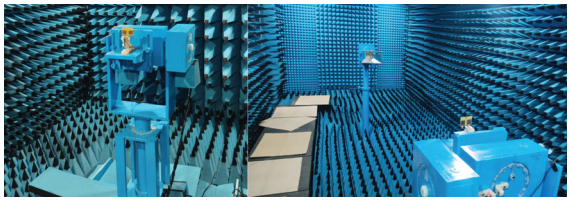


Fig. 8. Proposed MIMO antenna's radiation pattern measurement setup in an anechoic chamber.

depicted in Figs. 9 (a)-(h). Figures 9 (a)-(h) illustrates that the proposed antenna exhibits cross-polarization levels below -10 dB in both the E-plane and H-plane at the operational bands. Additionally, Figs. 9 (a)-(h) illustrate that the proposed antenna exhibits co-polarization levels below -8 dB in both the E-plane and H-plane at the working bands. The proposed MIMO antenna exhibits an omnidirectional radiation pattern across all operational bands (2.4, 3.3, 5, and 7 GHz), as illustrated in Figs. 9 (a)-(h).

The simulated and measured radiation patterns are in strong agreement. The simulated and measured efficiency of the proposed MIMO antenna is depicted in Fig. 10. Figure 10 illustrates that the proposed MIMO antenna exhibits above 90% efficiency across all operational bands (2.4, 3.3, 5, and 7 GHz).

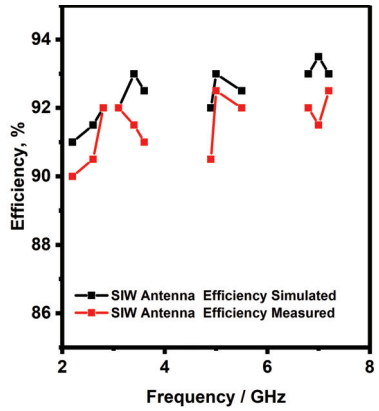


Fig. 10. Efficiency of the proposed MIMO antenna.

IV. DIVERSITY PERFORMANCE OF PROPOSED MIMO ANTENNA

The proposed system is assessed using several diversity performance indicators, including ECC, DG, CCL, TARC and MEG.

A. ECC

The envelope correlation coefficient, which is more commonly referred to as the ECC, is a statistical measure that is used to quantify the degree of correlation that occurs between the components of the MIMO antenna.

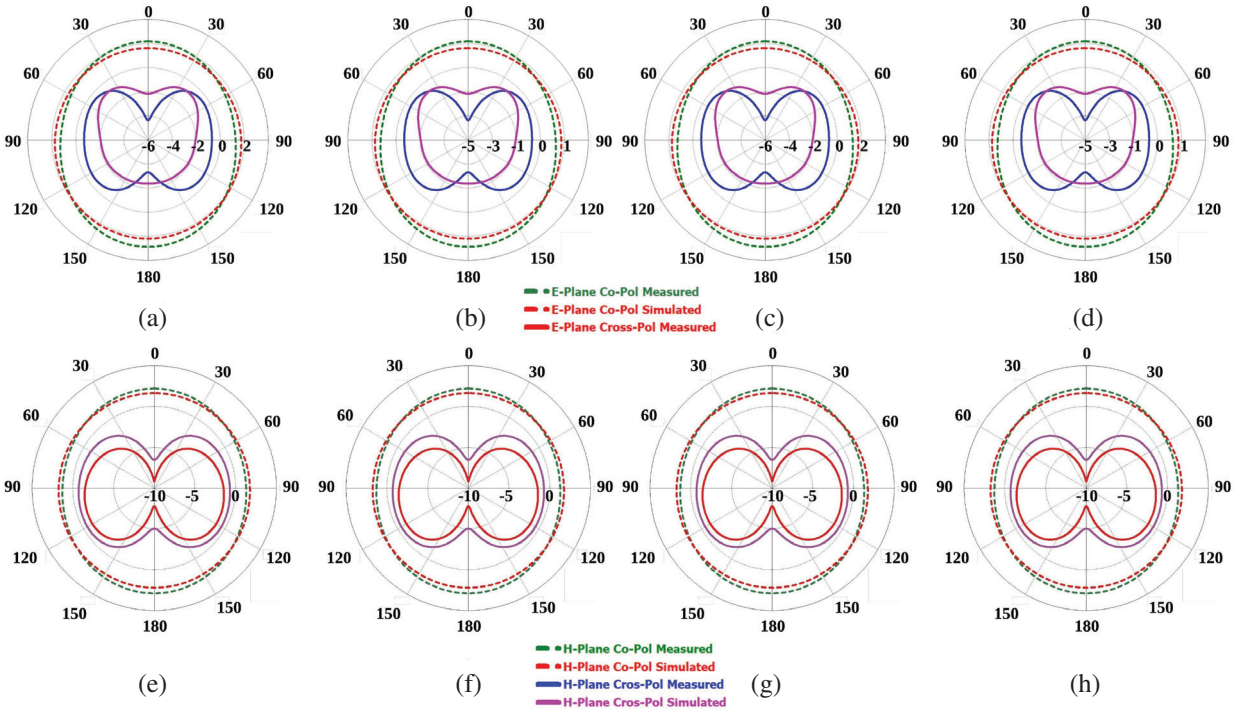
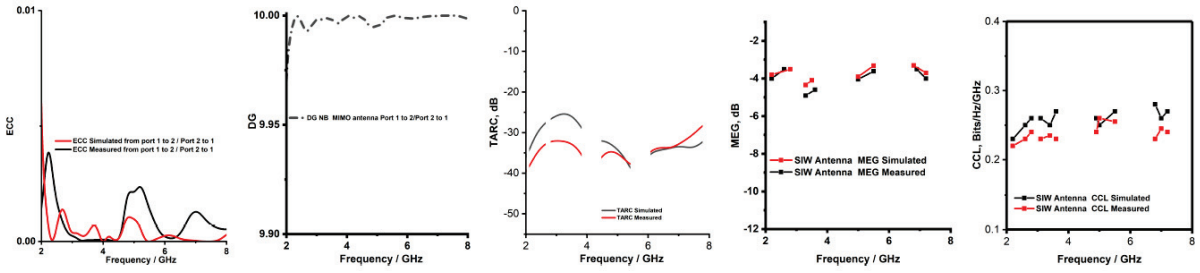


Fig. 9. Measured radiation patterns of the proposed antenna showing both co-polarization and cross-polarization components: (a) E-plane at 2.4 GHz, (b) at 3.3 GHz, (c) at 5 GHz, (d) at 7 GHz, (e) H-plane at 2.4 GHz, (f) at 3.3 GHz, (g) at 5 GHz, and (h) at 7 GHz.



(a) (b) (c) (d) (e)

Fig. 11. Diversity performance parameters of the proposed MIMO antenna: (a) ECC, (b) DG, (c) TARC, (d) MEG, and (e) CCL.

$$ECC = \frac{\left|\iint \overrightarrow{F_1}(\theta,\phi)\overrightarrow{F_2}(\theta,\phi)d\Omega\right|^2}{\iint \left|\overrightarrow{F_1}(\theta,\phi)\right| d\Omega^2 \iint \left|\overrightarrow{F_2}(\theta,\phi)\right| d\Omega^2}, \quad (3)$$

$$ECC = \frac{|S_{11}{}^*S_{12} + S_{21}{}^*S_{22}|^2}{\left(1-\left(|S_{11}|^2+|S_{21}|^2\right)\right)\left(1-\left(|S_{22}|^2+|S_{12}|^2\right)\right)}. \quad (4)$$

It is feasible to determine it using scattering parameters by using equation (3) in a lossless environment, which is characterized by the uniform distribution of power among antenna elements. However, this equation is only valid in situations when there is no loss. Therefore, the ECC can also be estimated in terms of the emitted far-field by employing Equation 4. In the case of uncorrelated MIMO antennas, the ideal value of ECC is zero. However, for actual MIMO antennas, the ECC value should be less than or equal to 0.5. The ECC proposed MIMO antenna is depicted in Fig. 11 (a). It can be observed from Fig. 11 (a) that the proposed MIMO antenna exhibits an ECC value of less than 0.04 at 2.4, 3.3,5 and 7 GHz.

B. Diversity gain (DG)

Diversity gain, a measure of the enhancement in system performance as a result of the diversity technique, is another essential metric and is computed by equation (5). An improved system's performance is shown by a larger diversity gain value. Its number should be around 10 for MIMO antenna performance that meets expectations.

$$DG = 10\sqrt{1-|ECC|^2}. \quad (5)$$

The DG of the proposed MIMO antenna is depicted in Fig. 11 (b). It can be observed from Fig. 11 (b) the proposed MIMO antenna exhibits a DG value of almost 10 at 2.4, 3.3, 5, and 7 GHz.

C. TARC

TARC is a statistical measure utilized to accurately evaluate the correlation among the elements of the MIMO antenna. The TARC of the two ports in a MIMO antenna system can be computed using equation (6). The TARC of the MIMO antenna must be below 0 dB for best performance.

$$TARC = \sqrt{\frac{|(S_{11}+S_{12}e^{j\theta})|^2+|(S_{21}+S_{22}e^{j\theta})|^2}{2}}. \quad (6)$$

The TARC of the MIMO antenna is illustrated in Fig. 11 (c). The Figure 11 (c) demonstrates that the suggested MIMO antenna has a TARC of less than -15 dB. Consequently, the suggested MIMO antennas exhibit less correlation among the radiators.

D. MEG

MEG is another critical statistic for evaluating the performance of MIMO antennas, as it estimates the average received signal strength for each radiator. MEGi denotes the ratio of the mean power received by the ith element to the mean power incident on the jth element from the identical source. MEG can be computed using equation (7), and the MEG value should fall between the range of -3 to 3 for optimal MIMO performance.

$$\mathrm{MEG_i} = 0.5\left[1-\sum_{i=1}^{N}\left|\mathrm{S_{ij}}\right|^2\right]. \quad (7)$$

The MEG of the MIMO antenna is illustrated in Fig. 11 (c). Figure 11 (c) demonstrates that the suggested MIMO antenna has acceptable MEG values over the operating band.

E. CCL

MIMO antennas are designed to enhance channel capacity. Nevertheless, the correlation among the antenna elements results in losses. CCL, or channel capacity loss, is a metric used to estimate the maximum feasible limit for signal transmission while minimizing significant loss. The calculation of CCL can be performed using equation (8). The CCL should be maintained below 0.4 Bits/Hz/Sec in practical

Table 2: Comparison of the proposed work with existing work

Ref.	Total Size (mm^3)	Operating Frequency (GHz)	Mutual Coupling (dB)	Peak Gain (dBi)	Efficiency (%)	Decoupling Method	ECC
[10]	48 × 48 × 1.6	2.25–2.41, 3.36–3.65, 4.7–6.25	≤ -15	1.5–2.9	80–88	Self-isolated	<0.0086
[11]	25 × 25 × 1.57	4.36–6.90	≤ -20	2	>90	Self-isolated	<0.08
[12]	25 × 25 × 1.57	3.4–3.6, 5.15–5.85	≤ -15	2	>86	Parasitic stripline	<0.11
[13]	35 × 35 × 1.57	3.1–5.2	≤ -15	2.6	>90	Self-isolated	<0.05
[35]	75 × 150 × 1.6 (8 elements)	3.4–3.6	≤ -15	2.5	60–70	Parasitic structure	<0.012
[15]	40 × 35 × 1.6 (2 elements)	3.5, 5.5, 6.5	≤ -15	3.49	>80	Polarization diversity	<0.05
Proposed	40 × 35 × 1.6 (2 elements)	2.4, 3.3, 5, 7	≤ -15	**6.2**	>90	**SIW cavity backed parasitic stub**	<0.04

applications.

$$C_{\text{loss}} = -\log_2 \det\left(\psi^R\right),$$

$$\left(\psi^R\right) = \begin{bmatrix} \rho_{ii} & \rho_{ij} \\ \rho_{ji} & \rho_{jj} \end{bmatrix}, \tag{8}$$

$$\rho_{ii} = 1 - \left(|S_{ii}|^2 + |S_{ij}|^2\right), \tag{9}$$

$$\rho_{ij} = -\left(\mathrm{S}_{\mathrm{ii}}^{*}\,\mathrm{S}_{\mathrm{ij}} + \mathrm{S}_{\mathrm{ji}}^{*}\,\mathrm{S}_{\mathrm{ij}}\right).$$

Figure 11 (e) illustrates a representation of the CCL of the MIMO antenna. Figure 11 (e) demonstrates that the proposed MIMO antenna exhibits a CCL of less than 0.4 Bits/Hz/Sec at 2.4, 3.3, 5, and 7 GHz.

V. COMPARISON OF THE PROPOSED WORK WITH EXISTING WORK

A comparison of the performance of the proposed MIMO antenna with that of current MIMO antennas is shown in Table 2. Based on the data presented in Table 2, it is evident that the suggested MIMO antenna is superior to the existing antennas in terms of size, new approach, and gain.

VI. CONCLUSION

This study presents a metamaterial-loaded quad-band cavity-backed substrate integrated waveguide (SIW) MIMO antenna for 5G, WLAN, and X-band applications in the sub-6 GHz region. SIW cavity-backed slot antennas are justified by their reduced fabrication complexity and smooth integration with planar circuits. Four operational bands are created using a modified-RSRR slot and an open-ended rectangular slot on the SIW cavity-backed rectangular radiator. Loading modified rectangular split-ring resonator (RSRR) unit cells in front of the radiators in the y-direction improves the MIMO antenna's gains by 5 dB, 8 dB, 6 dB, and 5 dB for frequencies of 2.4, 3.3, 5, and 7 GHz. Additionally, loading a modified modified-RSRR produces circular polarization. The cavity-backed antenna creates a unidirectional radiation pattern and reaches 88% efficiency at 2.4, 3.3, 5, and 7 GHz. The proposed MIMO antenna uses a cavity-backed parasitic element decoupling construction to achieve mutual coupling below -15 dB. MIMO characteristics include ECC $<$ 0.04, TARC $<$ -20 dB, CCL $<$ 0.4 bits/Hz/sec, and MEG between -3 dB and 3 dB at 2.4, 3.3, 5, and 7 GHz. Experimental validation of the system confirms its intended functionality, with simulated and measured outcomes in agreement. The proposed antenna is suitable for sub-6 GHz 5G, WLAN, and X-band applications.

REFERENCES

[1] A. Kumar, C. S. Rai, and M. K. Khandelwal, "Realization of miniaturized triple-Band four-Port stacked MIMO antenna for WLAN applications at 2.9 GHz, 5.0 GHz, and 5.9 GHz," *AEU – International Journal of Electronics and Communications*, vol. 150, p. 154216, 2022.

[2] M. Kaushik, J. K. Dhanoa, and M. K. Khandelwal, "Partially omnidirectional and circularly polarized MIMO antenna covering sub-6-GHz band for 5G fast plan," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 13, no. 9, pp. 1443–1450, 2023.

[3] M. K. Khandelwal, "Metamaterial-based circularly polarized four-port MIMO diversity antenna embedded with slow-wave structure for miniaturization and suppression of mutual coupling," *AEU - International Journal of Electronics and Communications*, vol. 121, p. 153241, 2020.

[4] U. Ullah, M. Al-Hasan, S. Koziel, and I. B. Mabrouk, "Series-slot-fed circularly polarized multiple-input–multiple-output antenna array enabling circular polarization diversity for 5G 28 GHz indoor applications," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 9, pp. 5607–5616, 2021.

[5] S. Virothu and M. Satya Anuradha, "Flexible CP diversity antenna for 5G cellular Vehicle-to-Everything applications," *AEU - International Journal of Electronics and Communications*, vol. 152, p. 154248, 2022.

[6] U. Sharma, G. Srivastava, and M. K. Khandelwal, "Quad-band two-port MIMO antenna serving for sub-7 GHz frequency with integrated circularly polarized bands," *AEU - International Journal of Electronics and Communications*, vol. 160, p. 154503, 2023.

[7] M. Tiwari, K. Afroz, and M. Fatima, "Wideband metasurface loaded circularly polarized MIMO microstrip antenna with high isolation," *Microwave Review*, vol. 28, pp. 22–27, 2022.

[8] J. Wang, Y. Cheng, H. Luo, F. Chen, and L. Wu, "High-gain bidirectional radiative circularly polarized antenna based on focusing metasurface," *AEU - International Journal of Electronics and Communications*, vol. 151, p. 154222, 2022.

[9] M. Y. Li, Z. Q. Xu, Y. L. Ban, C. Y. D. Sim, and Z. F. Yu, "Eight-port orthogonally dual-polarised MIMO antennas using loop structures for 5G smartphone," *IET Microw. Antennas Propag*, vol. 11, pp. 1810–1816, 2017.

[10] P. V. Naidu, M. B. Dhanekula, K. M. Almustafa, A. Kumar, K. A. Meerja, and S. H. Akkapanthula, "Design and performance analysis of MAZE shaped quad port ACS fed tri band MIMO antenna for V2V and multi band applications," *AEU-Int. J. Electron. Commun*., vol. 134, 2021.

[11] A. Ali, M. E. Munir, M. Marey, H. Mostafa, Z. Zakaria, A. J. A. Al-Gburi, and F. A. Bhatti, "A compact MIMO multiband antenna for 5G/WLAN/WIFI-6 devices", vol. 14, p. 1153, 2023.

[12] P. V. Naidu, M. B. Dhanekula, K. M. Almustafa, A. Kumar, K. A. Meerja, and S. H. Akkapanthula, "Design and performance analysis of MAZE shaped quad port ACS fed tri band MIMO antenna for V2V and multi band applications," *AEU-Int. J. Electron. Commun*, vol. 134, pp. 30–44, 2021.

[13] D. Wu, Y. Qiu, G. Yu, R. Guo, G. Wu, J. Wang, Y. Zhang, M. Zhu, and H. M. Zhou, "Decoupling technique using ferrite-film loading for 5G MIMO applications," *Int. J. Antennas Propag*., vol. 45, pp. 1–12, 2022.

[14] A. K. Singh, A. K. Dwivedi, K. N. Nagesh, V. Singh, and R. S. Yadav, "Compact 4-Port planar MIMO antenna with enhanced isolation for WLAN/WiMAX applications," *Sādhanā*, vol. 47, no. 3, p. 138, 2022.

[15] S. B. Paiva, A. G. D. Junior, V. P. S. Neto, and A. G. Assunção, "A new compact dual-polarized MIMO antenna using slot and parasitic element decoupling for 5G and WLAN applications," *Electronics*, vol. 11, no. 12, p. 1943, 2022.

[16] J. G. D. Oliveira, A. G. D'assunção Junior, V. P. S. Neto, and A. G. Assunção, "New compact MIMO antenna for 5G, WiMAX and WLAN technologies with dual polarisation and element diversity," *IET Microw. Antennas Propag*, vol. 15, pp. 415–426, 2021.

[17] A. Ali, M. E. Munir, M. Marey, H. Mostafa, Z. Zakaria, A. J. A. Al-Gburi, and F. A. Bhatti, "A compact MIMO multiband antenna for 5G/WLAN/WIFI-6 devices," *Micromachines*, vol. 14, no. 6, p. 1153, 2023.

[18] U. Sharma, G. Srivastava, M. K. Khandelwal, and R. Roges, "Shorting pins-based triple band circularly polarized modified monopole compact dual-port MIMO antenna for sub-6 GHz wireless applications," *AEU - International Journal of Electronics and Communications*, vol. 176, p. 155162, 2024.

[19] M. Moharana and B. Dwivedy, "Circularly polarized planar antennas with enhanced characteristics for contemporary wireless communication use cases: A review," *IEEE Access*, vol. 12, pp. 134594–134613, 2024.

[20] H. Luo, Q. Zhang, G. Sun, H. Yu, and D. Niyato, "Symbiotic blockchain consensus: Cognitive backscatter communications-enabled wireless blockchain consensus," *IEEE/ACM Transactions on Networking*, vol. 32, no. 6, pp. 5372–5387, Dec. 2024.

[21] M. Dai, G. Sun, H. Yu, S. Wang, and D. Niyato, "User association and channel allocation in 5G mobile asymmetric multi-band heterogeneous networks," *IEEE Transactions on Mobile Computing*, vol. 24, no. 4, pp. 3092–3109, Apr. 2025.

[22] H. Luo, G. Sun, C. Chi, H. Yu, and M. Guizani, "Convergence of symbiotic communications and blockchain for sustainable and trustworthy 6G wireless networks," *IEEE Wireless Communications*, vol. 32, no. 2, pp. 18–25, Apr. 2025.

[23] Y. Wang, R. Xiao, N. Xiao, Z. Wang, L. Chen, Y. Wen, and P. Li, "Wireless multiferroic memristor with coupled giant impedance and artificial synapse application," *Advanced Electronic Materials*, vol. 8, no. 10, p. 2200370, 2022.

[24] Y. Yao, F. Shu, X. Cheng, H. Liu, P. Miao, and L. Wu, "Automotive radar optimization design in a spectrally crowded V2I communication environment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 8, pp. 8253–8263, 2023.

[25] X. Zhang, H. Zhang, L. Liu, Z. Han, H. V. Poor, and B. Di, "Target detection and positioning aided by reconfigurable surfaces: Reflective or holographic?," *IEEE Transactions on Wireless Communications*, vol. 23, no. 12, pp. 19215–19230, 2024.

[26] H. Chu, X. Pan, J. Jiang, X. Li, and L. Zheng, "Adaptive and robust channel estimation for IRS-aided millimeter-wave communications," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 7, pp. 9411–9423, July 2024.

[27] S. Mukherjee and A. Biswas, "Design of planar high-gain antenna using SIW cavity hybrid mode," *IEEE Trans. Antennas Propag*, vol. 66, no. 2, pp. 972–977, 2018.

[28] G. Q. Luo, Z. F. Hu, W. J. Li, X. H. Zhang, L. L. Sun, and J. F. Zheng, "Bandwidth-enhanced low-profile cavity-backed slot antenna by using hybrid SIW cavity modes," *IEEE Trans. Antennas Propag*, vol. 60, no. 4, pp. 1698–1710, 2012.

[29] H. Liu, H. Tian, L. Liu, and L. Feng, "Co-design of wideband filtering dielectric resonator antenna with high gain," *IEEE Trans. CircuitsSyst. II Express Briefs*, vol. 69, pp. 1064–1068, 2022.

[30] C. Shi, J. Zou, J. Gao, and C. Liu, "Gain enhancement of a dual-band antenna with the FSS," *Electronics*, vol. 11, no. 18, p. 2882, 2022.

[31] D. C. Lugo, R. A. Ramirez, J. Wang, and T. M. Weller, "Multilayer dielectric end-fire antenna with enhanced gain," *IEEE Antennas Wirel. Propag. Lett*., vol. 17, pp. 2213–2217, 2018.

[32] B. A. F. Esmail, S. Koziel, and D. Isleifson, "Metamaterial-based series-fed antenna with a high gain and wideband performance for millimeter-wave spectrum applications," *Electronics*, vol. 12, no. 23, p. 4836, 2023.

[33] B. A. F. Esmail, S. Koziel, and A. Pietrenko-Dabrowska, "Wideband high-gain low-profile series-fed antenna integrated with optimized metamaterials for 5G millimeter wave applications," *Sci Rep*, vol. 14, p. 185, 2024.

[34] A. Gorai, A. Deb, and J. R. Panda, "Millimeter wave/5G multiband SIW antenna with metasurface loading for circular polarization and bandwidth enhancement," *J Infrared Milli Terahz Waves*, vol. 43, pp. 366–383, 2022.

[35] R. Khan, W. T. Sethi, and W. A. Malik, "Enhancing gain and isolation of a quad-element MIMO antenna array design for 5G sub-6 GHz applications assisted with characteristic mode analysis," *Sci Rep*, vol. 14, p. 11111, 2024.

R. Anandan Associate Professor and Head, Department of Electronics and Communication Engineering, Dhanalakshmi Srinivasan College of Engineering and Technology, ECR, Mamallapuram Chennai, completed his Ph.D. in the Faculty of Information and Communication Engineering from Anna University Chennai in May 2023. His interest includes antenna, optical communication, satellite communication, digital electronics, and wireless communication. He has published seven papers in international conferences and national conferences and four papers in international journals with good impact factors.

Sathyasri Balasubramanian is currently working as Professor in Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Avadi, Chennai, Tamil Nadu, India. She has a teaching experience of about 21 years in various engineering colleges. Her research interest includes IoT, Health Care, Network Security and Wireless Communication Networks. She has published more than 35 papers in reputed Journals and Conferences. She is a life member of professional organizations such as Indian Society for Technical Education (ISTE) and Senior member of IEEE.

Ravi Kumar Sanapala is currently working as an Associate Professor in St. Martin's Engineering College, Department of Electronics and Communication Engineering. He has completed M.Tech in Electronics and Communication Engineering from Andhra University College of Engineering, Visakhapatnam, Andhra Pradesh, India in the year 2005. He has successfully completed and awarded his Ph.D. in Electronics and Communication Engineering from Jawaharlal Nehru Technological University, Hyderabad, Telangana. India in 2023. He has 18+ years of teaching experience. He has published more than 30 National, International Conferences and Journals including Scopus and SCI. He has published 10 patents and books in different fields. His area of interest includes Wireless Communication, Radar and Microwave Engineering.

Prakash Kamisetti is working as Professor, Electrical and Electronics Engineering, Vaagdevi College of Engineering, Warangal, India. He has 25 years of experience in teaching and research. He did his B.E. from University of Madras and obtained his MTech and Ph.D. from National Institute of Technology, Warangal. His areas of research are ANN, Meta-Heuristic Techniques, AIML, Smart Grids, Distribution Automationand Distribution System studies. He has published 18 papers in national and international journals. He is a life member of professional organizations such as Indian Society for Technical Education (ISTE) and Senior member of IEEE.

Design, Simulation, and Experimental Investigation of a 7.24 GHz Pattern Reconfigurable Monopole Antenna for Enhanced Wireless Communication

Emine Ceren Gözek[1,2], Fikret Tokan[1], Muharrem Karaaslan[3], and Fatih Özkan Alkurt[3]

[1]Department of Electronics and Communications Engineering
Yıldız Technical University, Esenler, 34220 Istanbul, Turkey
ftokan@yildiz.edu.tr

[2]Department of Electrical and Electronic Engineering
Kahramanmaras Sutcu Imam University, 46000 Kahramanmaras, Turkey
eminecerengozek@ksu.edu.tr

[3]Department of Electrical and Electronic Engineering
Iskenderun Technical University, 31200 Hatay, Turkey
muharrem.karaaslan@iste.edu.tr, fozkan.alkurt@iste.edu.tr

Abstract – This study presents a novel pattern reconfigurable antenna structure designed for wireless communication systems. The proposed antenna operates at a center frequency of 7.2 GHz, which is particularly significant for 5G and future 6G communication systems. This frequency band is suitable for high-speed data transmission and enhances user experience by improving signal quality. The antenna can provide various beam steering capabilities, making it adaptable for multiple applications, including mobile communication systems configurations. The design methodology involves using a circularly grounded monopole antenna. The equivalent circuit models for the RF switches in both open and closed states are provided, facilitating the analysis of the antenna's electrical characteristics. Simulation and measurement results demonstrate that the antenna maintains a reflection coefficient ($|S_{11}|$) below -10 dB across a bandwidth of approximately 500 MHz, confirming its wideband performance. The findings indicate that the proposed antenna structure is not only efficient but also adaptable to various operational conditions, making it a promising candidate for future wireless communication applications.

Index Terms – Antenna beamforming, pattern reconfigurability, RF switches, UWB antenna, wireless communications.

I. INTRODUCTION

Reconfigurable antennas play a significant role in wireless communication systems since they can dynamically alter their radiation patterns, frequencies, and polarizations [1, 2]. These antennas achieve reconfigurability through various electronic solutions, including PIN diodes [3], varactor diodes [4], RF-MEMS switches [5], phase shifters, Microwave Liquid Crystal (MLC) technology [6], metasurfaces [7], and optical switches [8].

The advantages of reconfigurable antennas in terms of radiation patterns can be categorized as follows: flexibility, space efficiency, cost-effectiveness, improved signal quality [9], enhanced capacity and coverage [10], and high data transmission rates [11]. These advantages are critically important for meeting the requirements of modern communication systems. One of the fundamental benefits of these antennas is their ability to adjust the radiation pattern according to specific communication scenarios dynamically. This feature enables the antenna to direct signals and avoid interference in crowded environments [12].

When examining the literature on pattern reconfigurable antennas, studies have investigated the ability of antennas designed using RF-MEMS switches to dynamically change their frequency and radiation patterns. Deng et al. state that the Ka-band microstrip patch antenna designed using RF-MEMS switches can achieve reconfigurable radiation patterns. This study demonstrates that RF-MEMS switches can effectively alter the antenna's radiation pattern [13]. Zhang et al. introduced an innovative approach by presenting a pattern reconfigurable planar array antenna characterized by a digital coding method, which allows for flexible modifications of the antenna's radiation pattern [14]. Abdulhameed et al. developed a technique to enhance antenna performance by utilizing EBG (Electromagnetic Band Gap) structures and PIN diodes to control the radiation pattern of a microstrip antenna in both elevation and azimuth planes

Submitted On: January 22, 2025
Accepted On: July 29, 2025

https://doi.org/10.13052/2025.ACES.J.400806

[15]. Jmai et al. noted that their MMIC (Monolithic Microwave Integrated Circuit) antenna design, integrated with RF-MEMS switches, provides reconfigurability for both frequency and radiation patterns [16].

In this study, a novel circularly grounded monopole antenna concept capable of controlling the radiation pattern for wireless communication systems has been designed using RF-MEMS technology, analyzed in a simulation program and experimentally tested. The pattern reconfigurability of the radiation pattern was achieved using 32 RF switches. A microcontroller was employed to monitor the logical states of the electronic components, specifically logic 1 and logic 0, while the surface current density was altered to facilitate the control of the antenna's radiation pattern. The proposed antenna structure proves the beam steering capabilities of the proposed antenna at 7-7.5 GHz frequency bands.

II. ANTENNA DESIGN METHODOLOGY FOR PATTERN RECONFIGURABLE MONOPOLE ANTENNA

The configuration of the suggested circular monopole antenna is illustrated in Fig. 1, showing both front and side views. The antenna incorporates copper, characterized by an electrical conductivity of $5.8\text{x}10^7$ S/m, for its metallic reflectors and ground plane. In contrast, FR-4, a dielectric material with a relative permittivity of 4.3, is utilized as the substrate for the antenna design. The primary purpose of utilizing a circular structure in the design is to take advantage of the fundamental symmetry conditions. The circularly symmetry geometry brings the advantage of direction independence, versatile communication capabilities, wide bandwidth, and polarization diversity. In the circular structure, the central ground is surrounded by

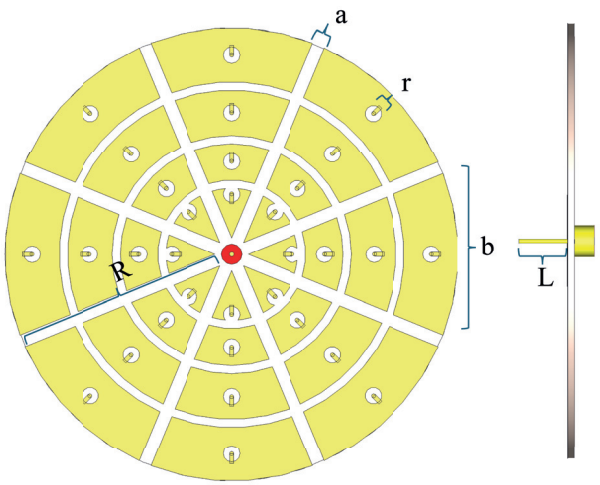


Fig. 1. Front and side views of the circular-shaped grounded monopole antenna.

32 identical patch reflectors to control the radiation beams. The beam control is achieved through the logical states of the RF switches, specifically logic 0 and logic 1, which correspond to the on/off states. These antenna dimensions were chosen by considering many factors such as the operating frequency of the antenna and the desired performance characteristics. These dimensions are provided in Table 1.

Table 1: Design parameters of recommended antenna

Parameter	R	L	r	a	b
Dimension (mm)	55	18.8	4	3.06	39.25

III. SWITCHING TECHNIQUE TO PROVIDE PATTERN RECONFIGURABILITY

The RF switches (BGS 1 2WN6) operate up to 9 GHz. In Fig. 2 (a), the switching network located in the substrate of the proposed antenna structure is illustrated, showing the connection configuration of 32 RF switches between the reflector and RF ground. Figures 2 (b) and (c) provide detailed connection points for a single RF switch. The design of the antenna allows for the RF current paths in the substrate to change under various switching conditions, depending on the on/off states of the RF switches utilized in the antenna design.

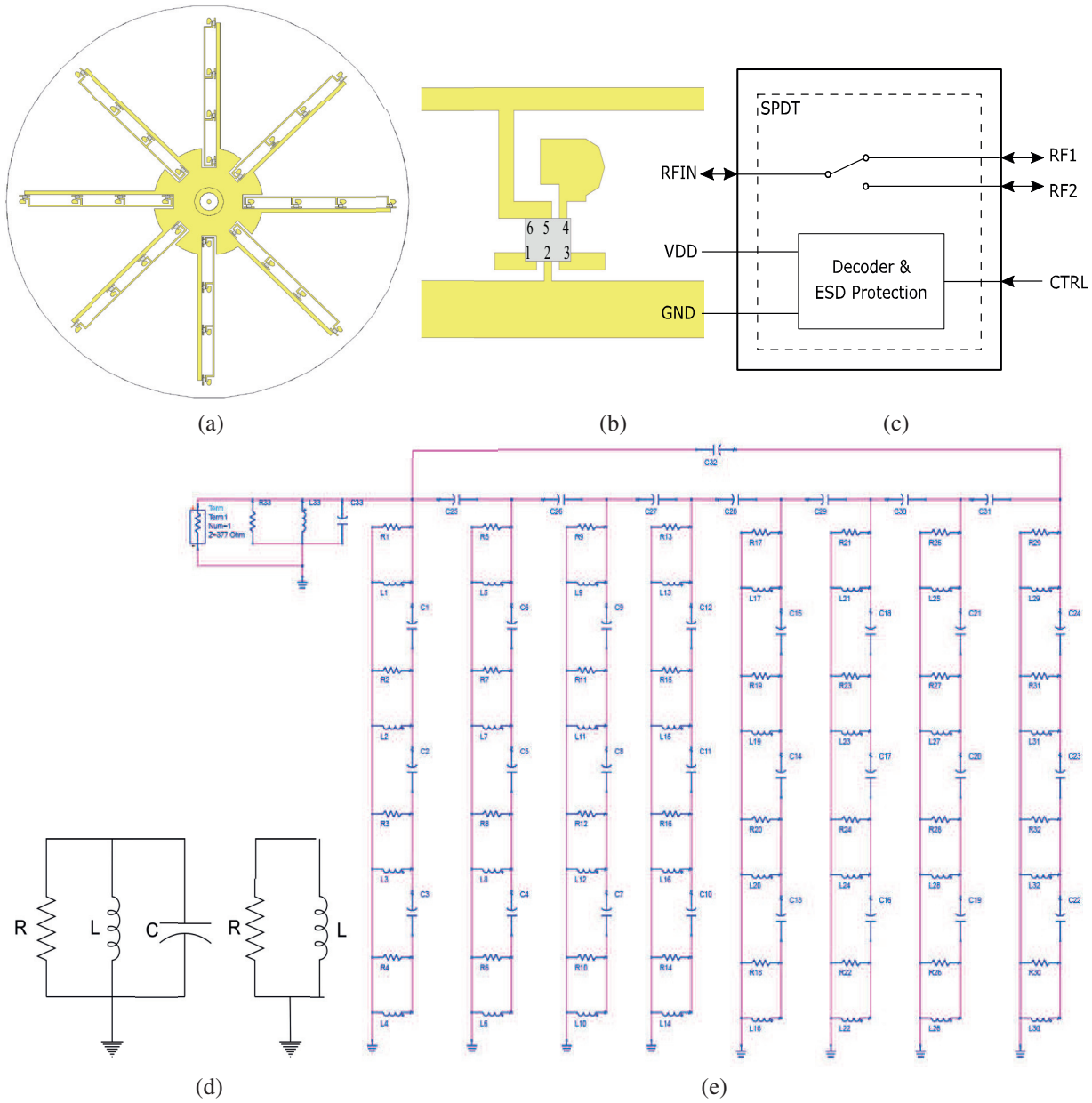
The equivalent circuits for the on/off states of the RF switch used in the proposed antenna structure are shown in Fig. 2 (d). The equivalent circuit model of the complete antenna structure is presented in Fig. 2 (e). These models represent the electrical characteristics of the antenna through a simplified circuit model, which facilitates the analysis of complex antenna structures. This approach enables a better understanding of performance, optimizes the design process, and allows for behavior analysis under different operating conditions. To clarify the derivation of the equivalent circuit model presented in Fig. 2 (e), the RLC (Resistance (R), Inductance (L), and Capacitance (C)) component values were calculated based on the electromagnetic behavior and geometrical configuration of the antenna's physical structure.

The inductance (L) represents the current paths along the metallic segments of the antenna, and it was approximated using standard inductance expressions for microstrip or wire segments, such as:

$$L \approx \frac{(\mu_0 \cdot l)}{(2\pi) \cdot [\ln(2l/r) - 1]}. \quad (1)$$

The capacitance (C) reflects the coupling between the patch reflectors and the ground plane, and it was estimated using the parallel-plate capacitance model:

$$C = \frac{(\varepsilon_r \cdot \varepsilon_0 \cdot A)}{d}, \quad (2)$$



(a) (b) (c) (d) (e)

Fig. 2. (a) The switching network located in the substrate of the proposed antenna structure, (b) RF switch connection diagram, (c) RF switch block diagram, (d) equivalent circuit model of an RF switch in the on/off-state, and (e) electrical equivalent circuit model of the antenna.

where A denotes the effective coupling area and d the vertical separation.

The resistance (R) accounts for conductor losses and is calculated based on the surface resistance R_s. The surface resistance is defined as:

$$R_s = \frac{1}{(\sigma \cdot \delta)}, \tag{3}$$

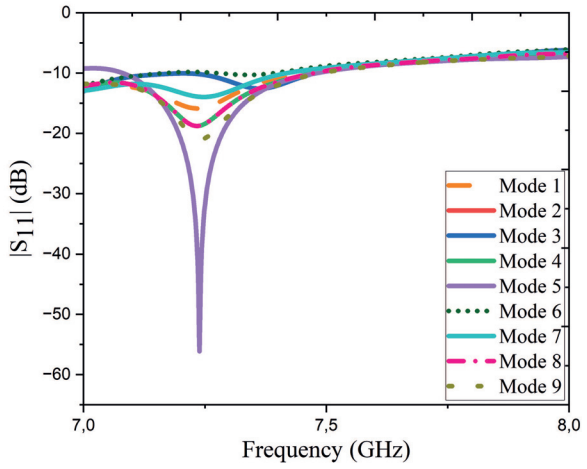
where the skin depth δ is given by:

$$\delta = \sqrt{\left(\frac{2}{(\omega \cdot \mu \cdot \sigma)}\right)}. \tag{4}$$

This completes the estimation of the RLC parameters based on physical and material properties.

IV. NUMERICAL ANALYSIS FOR THE PROPOSED ANTENNA STRUCTURE

Figure 3 shows the antenna's reflection coefficient $|S_{11}|$ and the table of switching states when the RF switches are in various switching states. For all switching states, the antenna appears to have a constant $|S_{11}|$ value below -10 dB in the 7-7.5 GHz band. The bandwidth of the antenna structure is 500 MHz.



(a)

Case	Sw1	Sw2	Sw3	Sw4	Sw5	Sw6	Sw7	Sw8
Mode1	1	1	1	1	1	0	0	0
Mode2	1	1	0	0	1	1	0	0
Mode3	0	0	1	1	0	0	1	1
Mode4	1	1	0	1	1	0	0	1
Mode5	1	1	1	1	1	1	1	1
Mode6	0	0	0	1	1	0	0	0
Mode7	0	0	0	1	1	1	1	0
Mode8	1	0	0	1	1	1	1	1
Mode9	1	1	0	1	1	1	1	1

(b)

Fig. 3. (a) Reflection coefficient of the antenna ($|S_{11}|$) under various switching conditions characteristics and (b) switching states.

The working principle of the pattern reconfigurable circular monopole antenna is based on controlling the far-field radiation patterns in the desired direction by dynamically regulating the electromagnetic current paths of the antenna through RF switches. The circular geometry of the antenna consists of multiple segments, which are turned on and off in certain combinations with the utilization of RF switches. Combinations of RF switches modulate the current density and thus the radiation pattern. Depending on the on and off states of the RF switches, the main lobe is formed in a specific (θ_0, $\emptyset_0$) direction and ensures that the radiation is limited in that direction.

Figure 4 presents various switching conditions of the antenna structure, showing the antenna visuals and simulated 3D and polar radiation patterns for each switching condition. The radiation patterns in polar form were obtained at 7.2 GHz in the θ=45° plane, where maximum radiation occurs. Across all switching conditions, maximum radiation is directed toward the θ=45° plane where the RF switches are closed. This indicates that, by adjusting the necessary switching conditions, it is possible to steer the radiation towards the desired direction.

The total electric field pattern corresponding to each switching state can be generally expressed as:

$$\mathrm{E}(\theta,\varphi)=\sum_{n=1}^{N} S_n\cdot F_n(\theta,\varphi), \tag{5}$$

where $\mathrm{E}(\theta,\varphi)$ is total electric field of the antenna considering to the switching states, S_n= *[S1, S2, S3, S_N]* is switching state (S_n=1 element active, S_n=0 element passive). The radiation function for the antenna structure can be expressed as:

$$F_n(\theta,\emptyset)=A_n\cdot\cos^m(\theta)e^{jkd_n\sin(\theta=45^\circ)\cdot\cos(\Phi-\phi_n)}, \tag{6}$$

where A_n is amplitude coefficient of the n element, m determines the magnitude of the main lobe, d_n is a distance of n element from the antenna center.

The total electric field pattern for the complex radiation function is expressed by:

$$\mathrm{E}(\theta,\varphi)=\sum_{n=1}^{N}\mathrm{S_n}\cdot\ \mathrm{A}_n\cdot\mathrm{e}^{\mathrm{j}\Phi_\mathrm{n}}\cdot\mathrm{f_n}(\theta,\varphi)\cdot\Gamma(\theta,\varnothing). \tag{7}$$

The term $\Gamma(\theta,\varnothing)$ accounts for non-idealities due to environmental effects and is defined as the sum of three main contributions:

$$\Gamma(\theta,\varnothing)=\gamma_g(\theta,\varnothing)+\gamma_m(\theta,\varnothing)+\gamma_r(\theta,\varnothing), \tag{8}$$

where $\gamma_g(\theta,\varnothing),\gamma_m(\theta,\varnothing)$, and $\gamma_r(\theta,\varnothing)$, and represent the ground reflection and surface wave effects, mutual coupling effects among antenna elements, and reflection and scattering effects from surrounding objects.

Alternatively, this term can be approximated in a simplified form:

$$\Gamma(\theta,\emptyset)=\eta(\theta,\emptyset)\cdot e^{\{j\psi(\theta,\emptyset)\}}, \tag{9}$$

where $\eta(\theta,\varnothing)$ denotes overall attenuation factor ($0<\eta\leq 1$) and $\psi(\theta,\varnothing)$ represents the net phase shift from all effects. We define phase angle as:

$$\Phi_\mathrm{n}=k\left(x_n\cdot\sin\theta\cdot\cos\emptyset+y_n\sin\theta\sin\emptyset+z_n\cos\theta\right),$$

$$k=\frac{2\pi}{\lambda}. \tag{10}$$

Finally, the individual radiation pattern of each element is defined as:

$$f_n(\theta,\emptyset)=G_n(\theta,\emptyset)P_n(\theta,\emptyset). \tag{11}$$

The antenna's ability to direct radiation toward any desired point in different switching states enhances the performance of communication systems, ensures energy efficiency, and increases system flexibility. Considering these advantages, this study enables modern communication systems to operate more effectively and efficiently. Observing the radiation patterns in various switching states, it is evident that the antennas can operate in single, dual, or multiple radiation modes. A single beam allows the antenna to focus in a specific direction, facilitating high-capacity data transfer with maximum energy. In dual-beam steering, the antenna radiates in two different directions, providing a wide coverage area and reducing signal interference, which offers significant advantages.

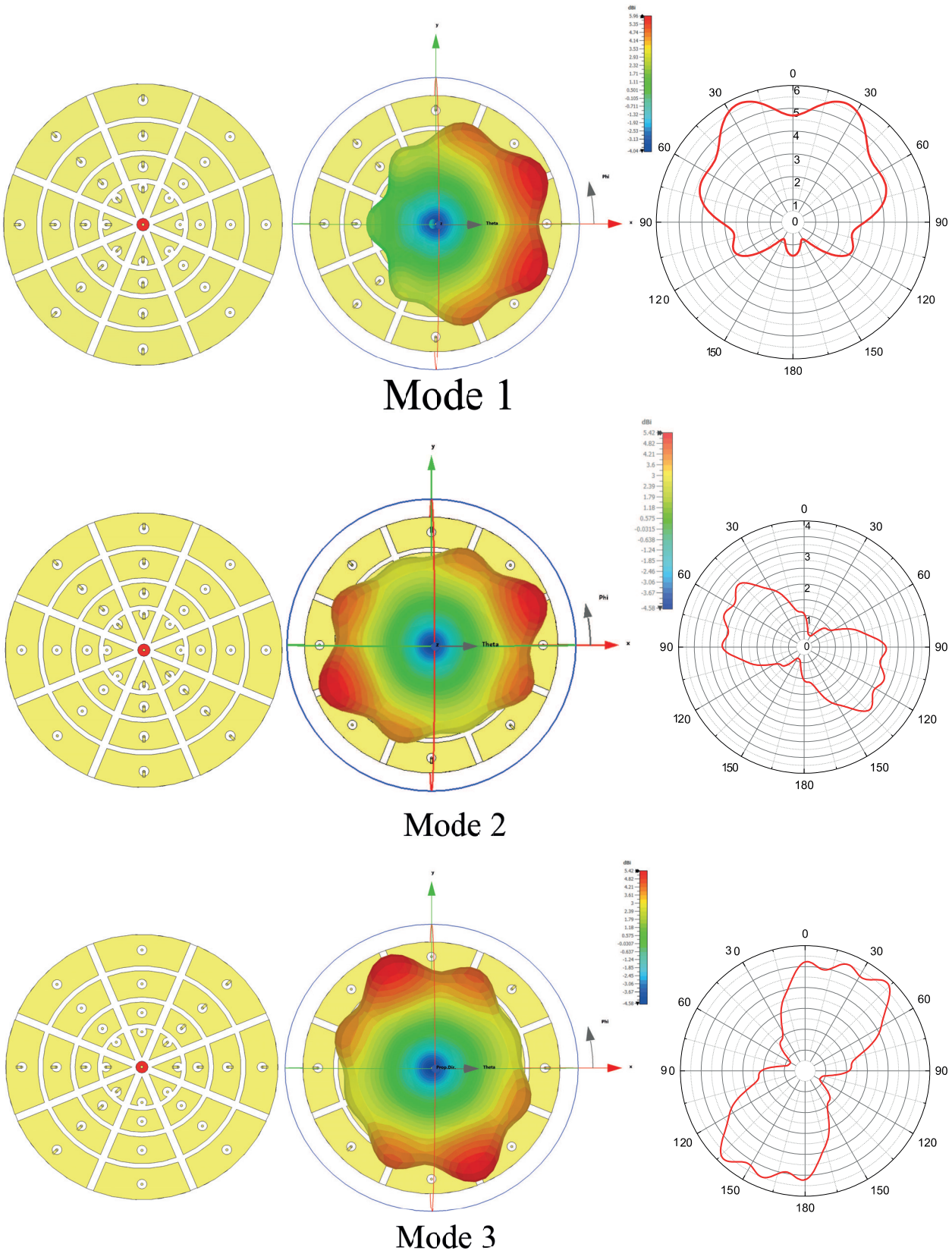
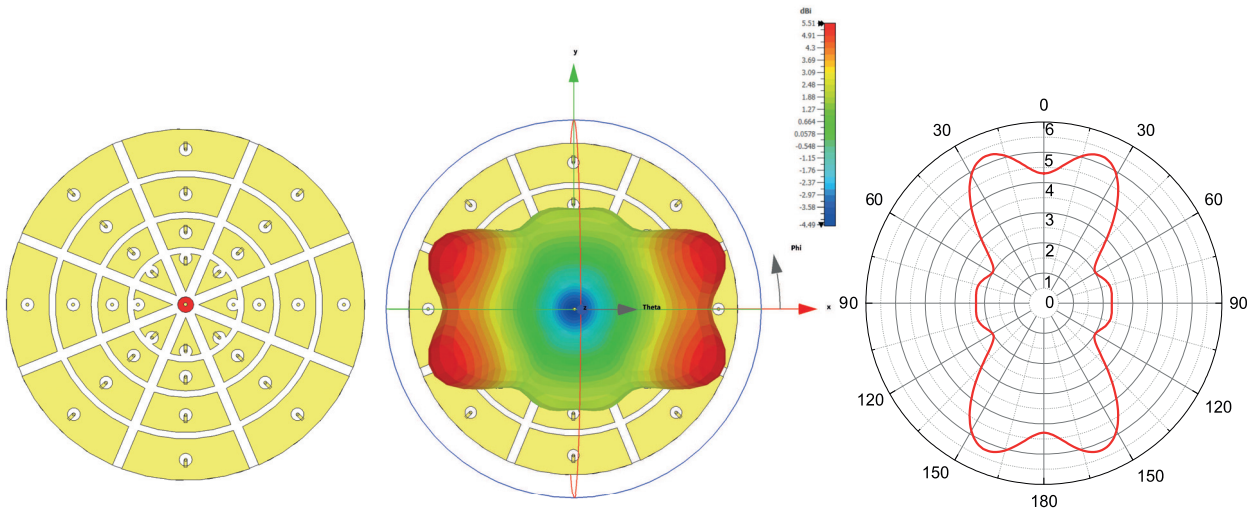
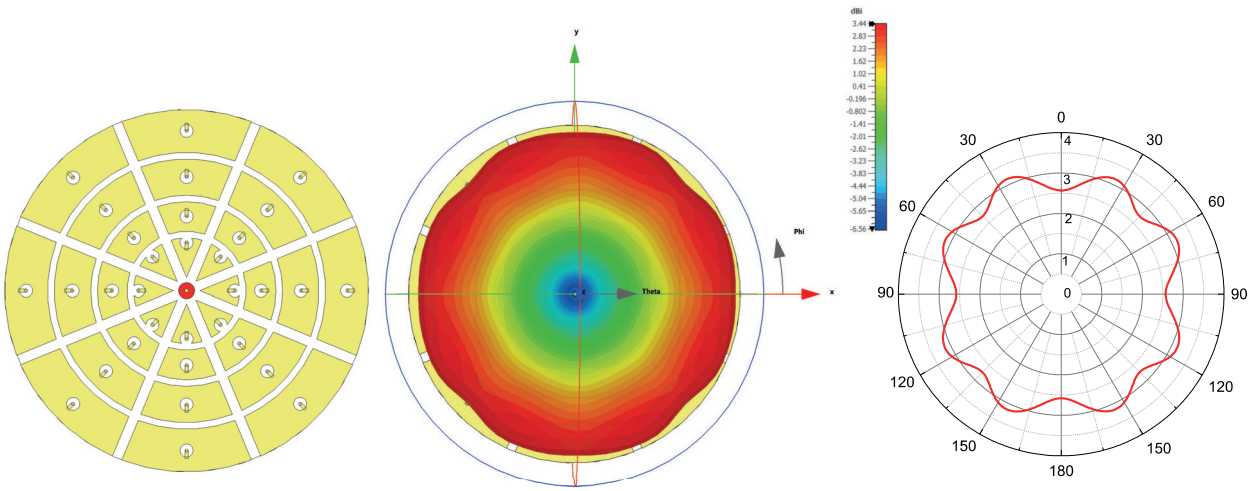


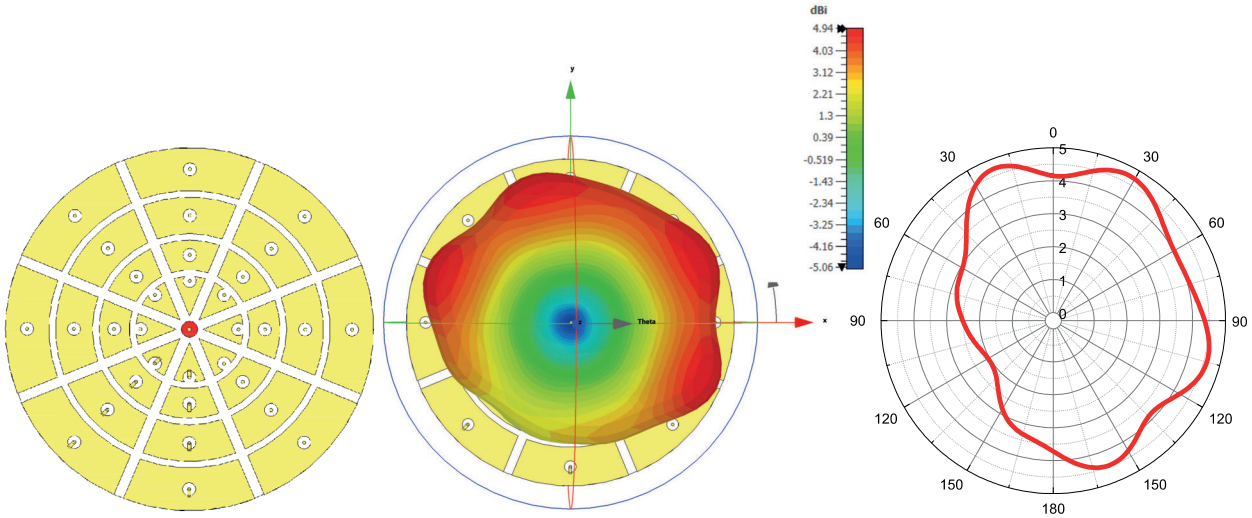
Fig. 4. *Continued.*



Mode 4



Mode 5



Mode 6

Fig. 4. *Continued.*

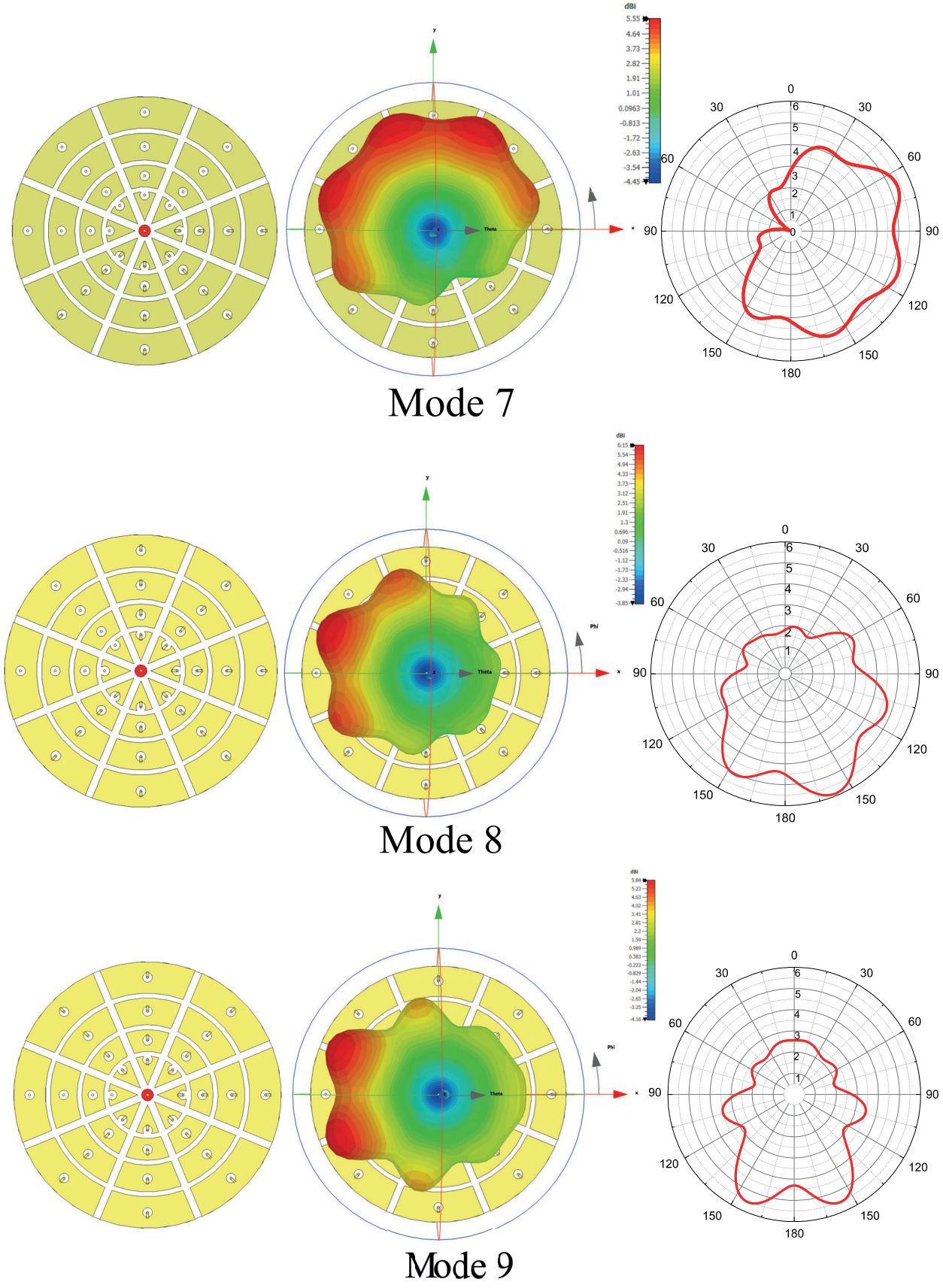
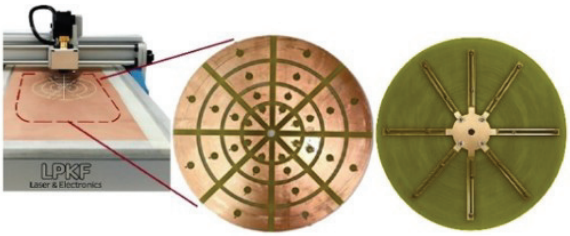


Fig. 4. Antenna switching conditions, simulated 3D and polar radiation patterns at θ=45° plane.

V. FABRICATION AND EXPERIMENTAL INVESTIGATION

Fabrication of the antenna was carried out in a microwave laboratory using the LPKF-E33 CNC-based printed circuit device shown in Fig. 5 (a). The dielectric material employed was double-sided copper-clad FR4 dielectric material. The monopole structure was fabricated from copper wire. The top and back layers of the produced antenna structure are illustrated in Figs. 5 (b) and (c).

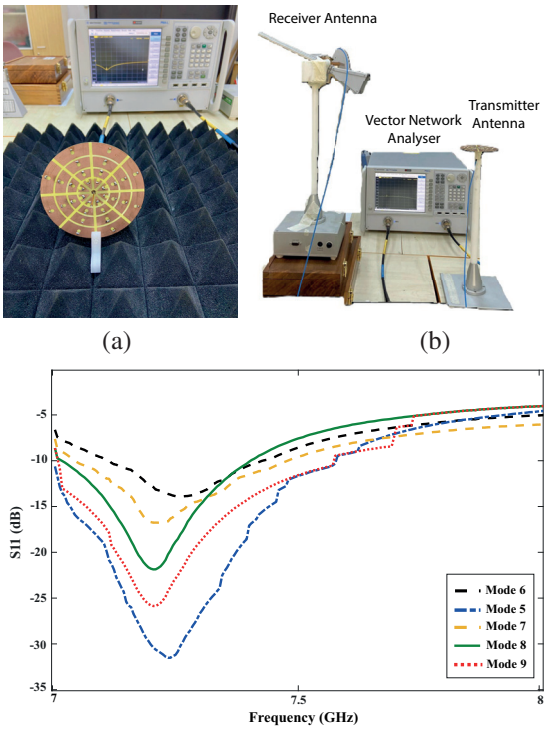
During the simulation phase, the beam steering capability of the antenna was achieved by controlling the on/off states of the RF switches. Each RF switch was positioned between the reflector elements located around the antenna and the ground plane. The on or off state of these switches altered the distribution of surface electromagnetic currents, thereby enabling the formation of beam modes with different radiation directions. As a result, the antenna could produce single, dual, or multi-beam radiation patterns. However, during the fabrication phase, due to the tiny dimensions of the RF switches (approximately 0.7×1.1 mm^2), it was not feasible to directly integrate them into the physical antenna structure. Instead, thin metallic connections were soldered at specific points to emulate the open or closed states of the switches. Through this approach, the switching logic used in the simulation was successfully replicated in the experimental prototype.



(a) (b) (c)

Fig. 5. Fabricated antenna structure (a) on LPKF, (b) top view and (c) back view.

$|S_{11}|$ measurements were performed using the Agilent PNA-L vector network analyzer with an operating frequency range of up to 43.5 GHz, as shown in Fig. 6 (a). The reflection coefficient and radiation model measurement setup of the proposed antenna structure are illustrated in Fig. 6 (b). As shown in the radiation model measurement setup, a reference horn antenna was utilized as the receiving antenna, while the fabricated antenna structure served as the transmitting antenna. The reflection coefficient plot of the fabricated proposed antenna structure is presented in Fig. 6 (c) for various switching states. The measured variations with frequency for all combinations show a broadband characteristic of the antenna at the 7.2 GHz center frequency.



(a) (b) (c)

Fig. 6. Measurement set up (a) reflection coefficient, (b) radiation pattern, and (c) measured reflection coefficient characteristics.

The measured radiation patterns of the fabricated antenna are presented together with the simulated ones for five randomly selected switching states in Fig. 7. These radiation patterns are almost in good agreement for all selected switching states. However, there are some ignorable mismatches which are caused by laboratory imperfections and using metallic connections in experiments that create insertion loss.

A detailed comparison of the proposed antenna with recent state-of-the-art designs is presented in Table 2, highlighting its distinctive advantages in terms of reconfigurability and design efficiency.

As presented in Table 2, the proposed antenna structure achieves a 16° beam steering capability at 7.2 GHz using RF-MEMS switches. Additionally, through various combinations of 32 individually switchable reflector elements, the antenna can generate single, dual, and multi-beam radiation patterns, which enhance energy efficiency, coverage area, and interference mitigation. The most significant distinction of this work from previous studies lies in the antenna's ability to precisely steer the beam toward a desired direction, which is not only electronically controlled but also mathematically mod-

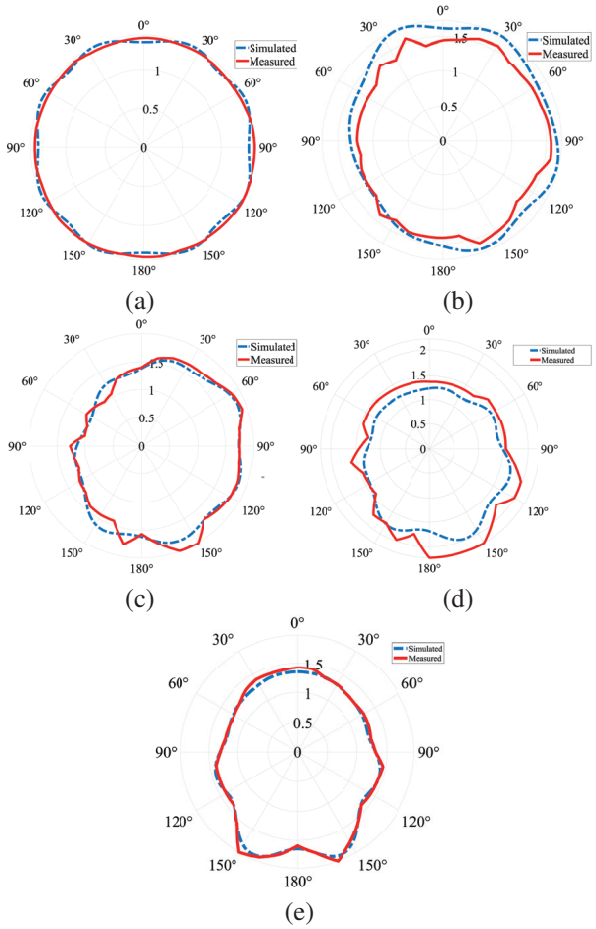


Fig. 7. Normalized radiation pattern of polar forms at (a) Mode 5, (b) Mode 6, (c) Mode 7, (d) Mode 8, and (e) Mode 9.

Table 2: Literature comparison

Ref.	Switching Element	Frequency (GHz)	Beam Steering Range	Reconfigurability Type
[13]	RF-MEMS	35	16°	Pattern
[15]	PIN Diode	6	40°	Pattern
[16]	RF-MEMS	26-30.7	-	Frequency
[17]	PIN Diode	1.57-2	84°	Pattern
This Work	RF-MEMS	7.2	16°	Pattern

elled and validated. This feature significantly improves the beam steering performance and facilitates integration into modern wireless communication systems.

VI. CONCLUSION

This study comprehensively examines the significance and potential of the novel circularly grounded monopole antenna in wireless communication systems. The proposed antenna structure designed using RF switches offers a critical advantage for meeting the requirements of modern communication systems by enabling dynamic adjustment of radiation patterns. The findings indicate that at a frequency of 7.2 GHz, the maximum radiation intensity occurs at an θ=45° direction when all RF switches are in the off position. This demonstrates the capacity of the proposed antenna to steer radiation toward desired directions by adjusting the switching conditions.

The antenna, operating in different switching states, exhibits single, dual, and multi-beam radiation modes, enhancing communication system performance and ensuring energy efficiency. The ability to form a single beam allows the antenna to focus on a specific direction, offering high data transfer capacity. In contrast, dual-beam steering provides significant advantages such as broad coverage and reduced signal interference.

In conclusion, the designed pattern reconfigurable antenna, with its flexibility, cost-effectiveness, and high data transmission speed, can significantly improve communication system performance and enhance user experience. The proposed antenna can be applied in 5G and 6G wireless communication systems and the Internet of Things. Future research should focus on frequency and polarization reconfigurability.

REFERENCES

[1] T. A. Elwi, F. Taher, B. S. Virdee, M. Alibakhshikenari, I. G. Zuazola, A. Krasniqi, A. S. Kamel, N. T. Tokan, S. Khan, and N. O. Parchin, "On the performance of a photonic reconfigurable electromagnetic band gap antenna array for 5G applications," *IEEE Access*, vol. 12, pp. 60849-60862, 2024.

[2] N. Ojaroudi Parchin, H. Jahanbakhsh Basherlou, Y. I. Al-Yasir, R. A. Abd-Alhameed, A. M. Abdulkhaleq, and J. M. Noras, "Recent developments of reconfigurable antennas for current and future wireless communication systems," *Electronics*, vol. 8, no. 2, p. 128, 2019.

[3] E. C. Gözek, F. Tokan, and M. Karaaslan, "Frequency reconfıgurable lıbra shape antenna for mmW 5G communıcatıons," *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 46, no. 7, pp. 1-16, 2025.

[4] Y. Mu, J. Han, D. Xia, X. Ma, H. Liu, and L. Li, "The electronically steerable parasitic patches for dual-polarization reconfigurable antenna using varactors," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 37, no. 1, pp. 58-67, 2022.

[5] W. M. Abdulkawi, A.-F. A. Sheta, W. A. Malik, S. U. Rehman, and M. S. Alkanhal, "RF MEMS

switches enabled H-shaped beam reconfigurable antenna," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 34, no. 9, pp. 1312-1319, 2019.

[6] S. Ma, X.-N. Li, Z.-D. Li, and J.-J. Ding, "Electronic beam steering metamaterial antenna with dual-tuned mode of liquid crystal material," *Sensors*, vol. 23, no. 5, p. 2556, 2023.

[7] G. Liu, C. Ju, Z. Li, Z. Yang, X. Qu, N. Liu, and W.-H. Zong, "A low-profile wideband pattern reconfigurable antenna with metasurface," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 39, no. 10, pp. 868-875, 2024.

[8] L. Silva, A. Alves, and A. Cerqueira Sodré Jr, "Optically controlled reconfigurable filtenna," *International Journal of Antennas and Propagation*, vol. 2016, no. 1, p. 7161070, 2016.

[9] A. H. Moore, S. Hafezi, R. R. Vos, P. A. Naylor, and M. Brookes, "A compact noise covariance matrix model for MVDR beamforming," *IEEE/ACM Transactions on Audio, Speech, and Language Processing,* vol. 30, pp. 2049-2061, 2022.

[10] M. Rihan, T. A. Soliman, C. Xu, L. Huang, and M. I. Dessouky, "Taxonomy and performance evaluation of hybrid beamforming for 5G and beyond systems," *IEEE Access*, vol. 8, pp. 74605-74626, 2020.

[11] J. Oh and K. Sarabandi, "Compact, low profile, common aperture polarization, and pattern diversity antennas," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 2, pp. 569-576, 2013.

[12] K. Li, Y. M. Cai, Y. Yin, and W. Hu," A wideband E-plane pattern reconfigurable antenna with enhanced gain," *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 29, no. 2, p. e21530, 2019.

[13] Z. Deng, Y. Wang, and C. Lai, "Design and analysis of pattern reconfigurable antenna based on RF MEMS switches," *Electronics*, vol. 12, no. 14, p. 3109, 2023.

[14] X. G. Zhang, W. X. Jiang, H. W. Tian, Z. X. Wang, Q. Wang, and T. J. Cui, "Pattern-reconfigurable planar array antenna characterized by digital coding method," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 2, pp. 1170-1175, 2019.

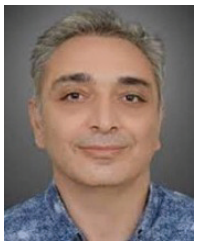
[15] M. Abdulhameed, M. Isa, Z. Zakaria, I. Ibrahim, and M. K. Mohsen, "Radiation pattern control of microstrip antenna in elevation and azimuth planes using EBG and pin diode," *International Journal of Electrical & Computer Engineering*, vol. 9, no. 1, p. 332, 2018.

[16] B. Jmai, S. Gahgouh, and A. Gharsallah, "A novel reconfigurable MMIC antenna with RF-MEMS resonator for radar application at K and Ka bands," *International Journal of Advanced Computer Science and Applications,* vol. 8, no. 5, pp. 468-473, 2017.

[17] W. M. Abdulkawi, W. A. Malik, A. F. A. Sheta, and M. A. Alkanhal, "A compact dual circular patch pattern reconfigurable antenna," *Microwave and Optical Technology Letters*, vol. 60, no. 11, pp. 2762-2768, 2018.

Emine Ceren Gözek is currently pursuing a doctoral education in the Communication Department of the Electronics and Communication Engineering program at Yıldız Technical University, Istanbul, Turkey. Simultaneously, she is a research assistant at Kahramanmaraş Sütçü Imam University. Her research interests include reconfigurable antenna, applications of metamaterials, electromagnetic waves and propagation.

Fikret Tokan received his Ph.D. degree from Yıldız Technical University, Istanbul, Turkey, in Communication Engineering in 2010. From October 2011 to October 2012, he was a Postdoctoral Researcher in the EEMCS Department of Delft University of Technology. From October 2012 to May 2013, he was a Postdoctoral Fellow at the Institute of Electronics and Telecomm. (IETR), University of Rennes 1, France. Since September 2002, he has worked in the Electromagnetic Fields and Microwave Technique Section of the Electronics and Communications Engineering Department of Yıldız Technical University. He is currently working as a professor in that department. His current research interests are UWB antenna design, dielectric lens antennas, reflector systems, electromagnetic waves, propagation, antenna arrays, scattering, and numerical methods.



Muharrem Karaaslan received the Ph.D. degree in physics from the University of Cukurova, Adana, Turkey, in 2009. He has co-authored almost 200 scientific contributions published in journals and conference proceedings. His research interests include the applications of metamaterials, analysis and synthesis of antennas, and waveguides.

Fatih Özkan Alkurt received the Ph.D. degree in electrical and electronics engineering from the Iskenderun Technical University, Turkey, in 2024. He was awarded the 2024 Excellence in Electrodynamics Award. His research interests include antenna designs, microwave transmission lines and metamaterials.

Numerical Simulation of Melt-wave in Electromagnetic Launcher

Kefeng Yang, Gang Feng, Shaowei Liu, Xiaoquan Lu, Xiangyu Du, and Tianyou Zheng

Air Defense and Antimissile School
Air Force Engineering University, Xi'an 710051, China
kkebee21@163.com, 329276300@qq.com, lsw3721@163.com,
dxy051603@163.com, 10951446520@qq.com, 15633041891@163.com

***Abstract* –** To accurately characterize the erosion phenomenon of the armature in electromagnetic railgun launches, a two-dimensional magneto-thermal-mechanical coupling model for melt-wave was developed. For the first time, a fully implicit finite volume method was employed for equation discretization, and an alternating direction implicit method was used for coupling calculations to obtain both steady-state and transient erosion characteristics of the armature. The results demonstrate that the velocity skin effect concentrates significant current at the armature tail, driving the propagation of the melt-wave. The erosion rate remains constant initially but increases significantly when variations in electrical conductivity are considered. After applying an external current, the erosion distance increases sharply with current amplitude before leveling off, and changes in the duration of current amplitude also significantly influence the erosion distance. This study provides a clear understanding of the armature's erosion behavior, offering a solid theoretical foundation for further research on armature transition phenomenon.

***Index Terms* –** Electromagnetic launcher, erosion depth, melt-wave, transition, velocity skinning effect.

I. INTRODUCTION

Electromagnetic railgun, a novel launcher that propels projectiles out of the chamber using the Lorentz force, holds great promise for both civilian applications and military advancements [1, 2]. Nevertheless, the entire launching process occurs in an extremely complex environment. Besides being affected by pulsed high current, under the velocity skinning effect, a large amount of Joule heat accumulates at the tail of the armature, leading to thermal melting of the armature [3–5]. As the armature moves, the ablation spreads towards the head, forming a melt-wave, which undergoes a transition as it traverses the armature [6, 7]. Consequently, exploring the melting characteristics of the armature serves as the foundation for preventing armature transition.

Owing to the hazardous experimental environment and intricate coupling scenarios, numerical simulations are commonly employed to study melt-wave. In prior research, Parks integrated the velocity skinning effect into the study of armature transition and derived an analytical solution for the melt-wave erosion depth through two distinct methods, revealing its correlation with the armature velocity. Parks was the first to introduce the concept of the melt-wave to predict transition, offering valuable inspiration for subsequent studies [8]. Stefani respectively utilized the finite difference method and the finite element analysis software EMAP3D to investigate the one-dimensional melt-wave erosion depth and proposed a pseudo melting algorithm to simulate ablation [9]. Gong conducted numerical simulations of one-dimensional, two-dimensional, and three-dimensional armature melt-wave using the finite difference method to comprehensively explore the variation patterns of the melt-wave [10]. Tang discovered that the melt-wave and electromagnetic force contribute to the armature's transition by modeling the three-dimensional melt-wave [11]. Tan et al. [12] and Sun et al. [13] adopted the finite element method with the Galerkin scheme. Tan et al. applied the nodal ablation method, while Sun et al. utilized the expanded magnetic boundary method to calculate the erosion rate of the moving armature. The above-mentioned literature predominantly relies on the finite difference method and the finite element method. In contrast to these two methods, the finite volume method offers significant advantages, as it can ensure both local and global energy conservation and directly handle the fluxes at the control volume boundaries [14, 15].

This paper presents, for the first time, a melt-wave computational model based on the finite volume method. By employing the approach of updating the magnetic field boundary conditions, this model simulates the actual melting process. Moreover, it investigates the impacts of the velocity skinning effect, initial magnetic excitation, and velocity on the erosion distance, taking into account the temperature dependent conductivity.

Submitted On: December 17, 2024
Accepted On: May 30, 2025

https://doi.org/10.13052/2025.ACES.J.400807

Additionally, the effects of different current waveforms on armature melting are explored.

II. MODEL CONSTRUCTION AND NUMERICAL DISCRETIZATION

A. Theoretical model

The two-dimensional electromagnetic railgun model usually has symmetry, in which copper is selected as the material for the rail and the armature material is 7075Al, which slides along the rail in the positive direction of the x-axis with the velocity v. Therefore, half of the armature and one side of the rail are selected as the calculation area, as shown in Fig. 1. In the calculation model, in order to simplify the calculation, it is assumed that the armature remains motionless and the rail slides in the negative direction of the x-axis with the speed v.

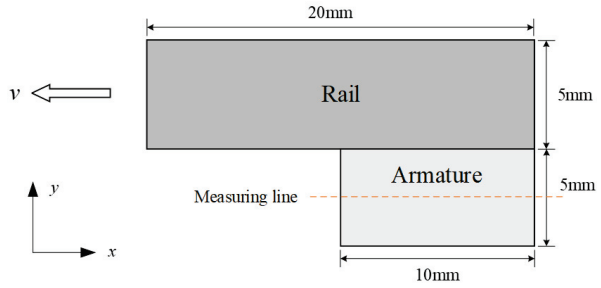


Fig. 1. Computational model.

The two-dimensional electromagnetic railgun model typically exhibits symmetry. In this model, copper is chosen as the material for the rail, while the armature is made from 7075 aluminum. The armature moves along the rail in the positive direction of the x-axis at a velocity v. For the purposes of analysis, we focus on half of the armature and one side of the rail, as illustrated in Fig. 1. To simplify the calculations, we assume that the armature remains stationary while the rail moves in the negative direction of the x-axis at speed v.

B. Discretization of the magnetic diffusion equation

After derivation of Maxwell's system of equations, the magnetic diffusion equation for the electromagnetic launcher can be expressed as [16]:

$$\frac{1}{\mu_0}\left(\nabla\times(\nabla\times\vec{B})\right)=\sigma\left(\frac{\partial\vec{B}}{\partial t}+\nabla\times\left(\vec{U}\times\vec{B}\right)\right), \quad (1)$$

where t is time, U is velocity, σ is conductivity, B is magnetic induction and μ_0 is vacuum permeability.

In this two-dimensional electromagnetic launcher model, the magnetic induction intensity in the x and y directions is neglected, and the magnetic permeability is assumed to be independent of temperature. Consequently, the magnetic diffusion equations are derived separately for the rail and the armature as follows [23]:

$$\frac{1}{\mu_0\sigma_1}\left(\frac{\partial}{x}\left(\frac{\partial B_1}{\partial x}\right)+\frac{\partial}{y}\left(\frac{\partial B_1}{\partial y}\right)\right)+v\frac{\partial B_1}{\partial x}=\frac{\partial B_1}{\partial t}, \quad (2)$$

$$\frac{1}{\mu_0\sigma_2}\left(\frac{\partial}{x}\left(\frac{\partial B_2}{\partial x}\right)+\frac{\partial}{y}\left(\frac{\partial B_2}{\partial y}\right)\right)=\frac{\partial B_2}{\partial t}, \quad (3)$$

whereB_1 and B_2 represent the magnetic induction in the z direction experienced by the rail and the armature, respectively. Additionally, σ_1 and σ_2 denote the conductivity of the rail and armature, respectively.

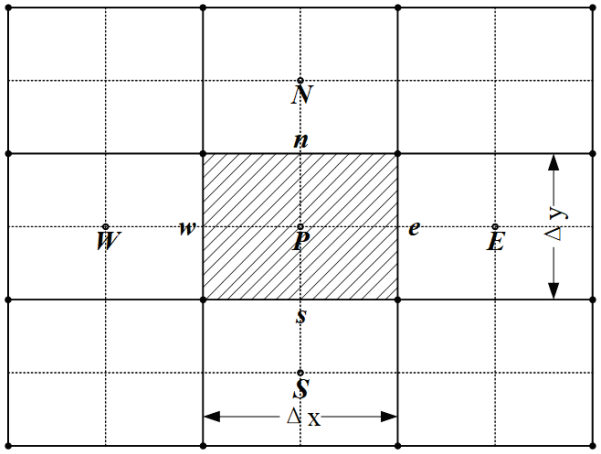


Fig. 2. Finite volume method calculation grid.

The finite volume method partitions the two-dimensional computational domain of the rail cannon into control volume units, allowing for the application of conservation integrals to each unit volume [22]. Figure 2 illustrates the two-dimensional magnetic diffusion control volume grid, where point P, marked as nodes, corresponds to the control volume units. The four nodes located in the southwest, southeast, northwest, and northeast regions are represented by the points E, S, W, and N. The corresponding four unit interfaces are denoted as e, s, w, and n.

By integrating equation (2), the value of B_1 at point P can be further expressed as follows:

$$\begin{aligned}&\frac{B^{n+1}_{1i,j}-B^{n}_{1i,j}}{\Delta t}\Delta V+(vB_1A)_e-(vB_1A)_w=\\&\left(\left(D_1A\frac{\partial B_1}{\partial x}\right)_e-\left(D_1A\frac{\partial B_1}{\partial x}\right)_w\right)+\\&\left(\left(D_1A\frac{\partial B_1}{\partial y}\right)_n-\left(D_1A\frac{\partial B_1}{\partial y}\right)_s\right)\end{aligned}, \quad (4)$$

where $D_1=1/\mu_0\sigma_1$, A is the area of the control volume and $A_e=A_w=\Delta y$, $A_n=A_s=\Delta x$. Control volume $\Delta V=\Delta x\Delta y$. The values at points W, E, N, and S are $B^n_{1i-1,j}$, $B^n_{1i+1,j}$, $B^n_{1i,j+1}$, and $B^n_{1i,j+1}$, respectively. Denoting the values of $\partial B_{-1}/\partial x$ in w, e, n, and s by $B^n_{1i-1,j}$, $B^n_{1i+1,j}$,

$B^n_{1i,j+1}$, and $B^n_{1i,j+1}$, respectively, equation (4) becomes:

$$\begin{aligned}\left(\frac{\Delta x\Delta y}{\Delta t}+2D_1\frac{\Delta y}{\Delta x}+2D_1\frac{\Delta x}{\Delta y}\right)B^n_{1i,j}-\frac{\Delta x\Delta y}{\Delta t}B^{n-1}_{1i,j}=\\ \left(D_1\frac{\Delta y}{\Delta x}-v\frac{\Delta y}{2}\right)B^n_{1i+1,j}+\left(D_1\frac{\Delta y}{\Delta x}+v\frac{\Delta y}{2}\right)B^n_{1i-1,j}+\\ D_1\frac{\Delta x}{\Delta y}B^n_{1i,j+1}+D_1\frac{\Delta x}{\Delta y}B^n_{1i,j-1}.\end{aligned} \tag{5}$$

The format of the implicit discrete equation that satisfies the heat conduction conditions can be expressed as follows:

$$\begin{aligned}a_pB^n_{1i,j}-a^0_pB^{n-1}_{1i,j}=a_eB^n_{1i+1,j}+a_wB^n_{1i-1,j}+\\ a_nB^n_{1i,j+1}+a_sB^n_{1i,j-1}\end{aligned}, \tag{6}$$

$$\begin{cases}a^0_p=\frac{\Delta x\Delta y}{\Delta t},\\ a_e=D_1\frac{\Delta y}{\Delta x}+v\frac{\Delta y}{2},\ a_w=D_1\frac{\Delta y}{\Delta x}-v\frac{\Delta y}{2}\\ a_n=a_s=D_1\frac{\Delta x}{\Delta y}\\ a_p=a^0_p+a_e+a_w+a_n+a_s\end{cases}. \tag{7}$$

For the armature, the discretized result is given by:

$$\begin{cases}a^0_p=\frac{\Delta x\Delta y}{\Delta t}\\ a_e=a_w=D_2\frac{\Delta y}{\Delta x},a_n=a_s=D_2\frac{\Delta x}{\Delta y}\\ a_p=a^0_p+a_e+a_w+a_n+a_s\end{cases}. \tag{8}$$

C. Discretization of the heat conduction equation

The heat conduction equation derived from Fourier's law can be formulated as follows:

$$\kappa\left(\nabla\times(\nabla\times\vec{T})\right)=\rho c\left(\frac{\partial\vec{T}}{\partial t}+\nabla\times\left(\vec{U}\times\vec{T}\right)\right)-\frac{J^2}{\sigma}, \tag{9}$$

where $\kappa$is thermal conductivity, T is temperature, ρ is density, and c is specific heat. J can be expressed as:

$$J=\sqrt{J_x^2+J_y^2}=\sqrt{(\frac{1}{\mu_0}\frac{\partial B}{\partial y})^2+(-\frac{1}{\mu_0}\frac{\partial B}{\partial x})^2}. \tag{10}$$

The heat transfer equations for the armature and the rail can be expressed as follows:

$$\frac{\kappa_1}{\rho_1c_1}\left(\frac{\partial}{x}\left(\frac{\partial T_1}{\partial x}\right)+\frac{\partial}{y}\left(\frac{\partial T_1}{\partial y}\right)\right)+v\frac{\partial T_1}{\partial x}+\frac{J_1^2}{\sigma_1}=\frac{\partial T_1}{\partial t}. \tag{11}$$

$$\frac{\kappa_2}{\rho_2c_2}\left(\frac{\partial}{x}\left(\frac{\partial T_2}{\partial x}\right)+\frac{\partial}{y}\left(\frac{\partial T_2}{\partial y}\right)\right)+\frac{J_2^2}{\sigma_2}=\frac{\partial T_2}{\partial t}. \tag{12}$$

Let $K_i=\kappa_i/\rho_ic_i$. The discretized heat conduction equations for the rail and the armature can be formulated as:

$$\begin{aligned}a_pT^n_{i,j}-a^0_pT^{n-1}_{i,j}=a_eT^n_{i+1,j}+a_wT^n_{i-1,j}+\\ a_nT^n_{i,j+1}+a_sT^n_{i,j-1}+\frac{J^2}{\sigma}\end{aligned}, \tag{13}$$

$$\begin{cases}a^0_p=\frac{\Delta x\Delta y}{\Delta t}\\ a_e=K_1\frac{\Delta y}{\Delta x}-v\frac{\Delta y}{2},\ a_w=K_1\frac{\Delta y}{\Delta x}+v\frac{\Delta y}{2}\\ a_n=a_s=K_1\frac{\Delta x}{\Delta y}\\ a_p=a^0_p+a_e+a_w+a_n+a_s\end{cases}, \tag{14}$$

$$\begin{cases}a^0_p=\frac{\Delta x\Delta y}{\Delta t}\\ a_e=a_w=K_2\frac{\Delta y}{\Delta x},\ a_n=a_s=K_2\frac{\Delta x}{\Delta y}\\ a_p=a^0_p+a_e+a_w+a_n+a_s\end{cases}. \tag{15}$$

D. Boundary conditions and simulation methods

Once the external current is applied, the magnetic fields produced by the two rails counteract each other due to their opposing current directions. In regions where the armature does not extend, no current flows, resulting in a magnetic induction intensity of 0. Within the armature, the current predominantly flows along the shortest path between the two rails, concentrating at the armature's tail. To prevent sudden changes in the magnetic field gradient for the rails that the current has traversed, the boundary conditions for the armature's magnetic field are established as depicted in Fig. 3. Regarding the temperature field, it is assumed that, initially, the temperature distribution within the armature is uniform at 300K, representing room temperature. It is further specified that all heat dissipates solely through internal conduction and not through the boundaries. Additionally, the normal temperature gradient along all edges is set to 0 [25].

The erosion treatment of the armature involves phase transformation, with the current focusing on the armature tail. This leads to a rapid temperature increase, causing the material to start melting upon reaching its melting point. Once the temperature surpasses the melting point it stabilizes, and any additional heat absorbed is stored. Upon reaching the latent heat of melting, the phase transition concludes, signifying the transformation of the armature into a liquid state. Following exposure to electromagnetic force, the liquid component separates from the solid armature. When the armature on the grid completely melts, assuming the molten metal separates entirely from the armature due to the high-speed armature action, the armature and the rail lose contact. The space between them is considered as air, with no current passing through the armature, resulting in a current density of zero. Consequently, the magnetic field boundary condition is then updated to reflect the unmelted region.

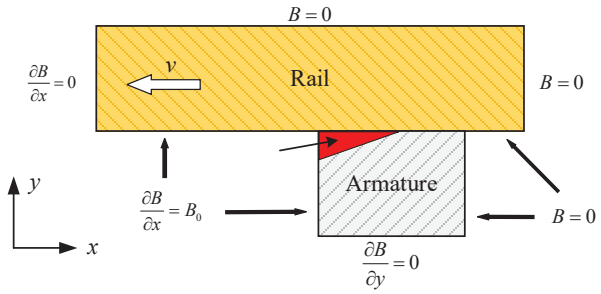


Fig. 3. Setting of magnetic field boundary conditions.

The model is partitioned into 0.1 mm square grids with a time step of $0.1\mu s$. The magnetic field boundary is updated using a method to simulate liquid armature separation, and the melt-wave calculation is conducted

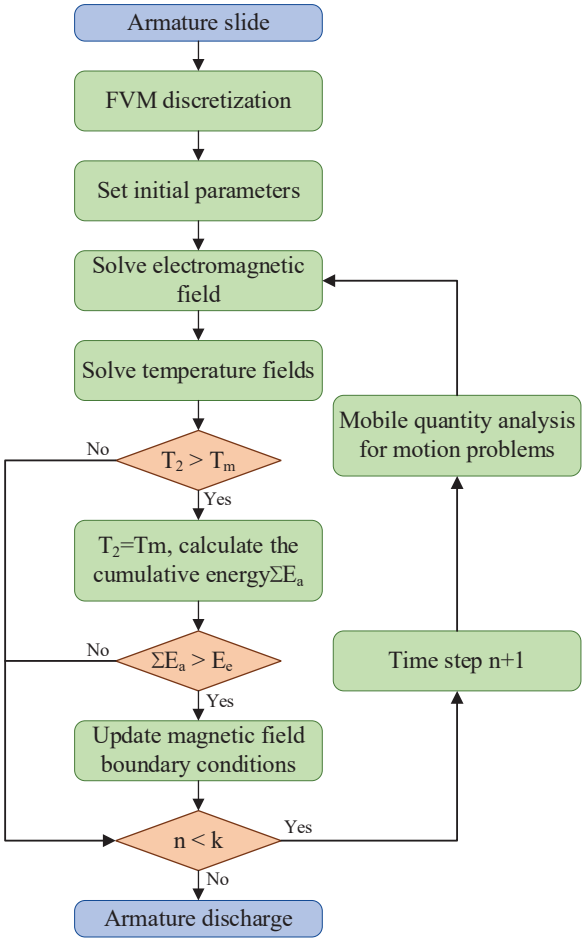


Fig. 4. Overall calculation process.

using MATLAB. The calculation process is illustrated in Fig. 4. By discretizing the magnetic diffusion and heat conduction equations for the armature and rail, these are transformed into linear equations. To calculate the steady melt-wave, initial parameters are set, maintaining the initial magnetic excitation or velocity while varying another parameter to solve for armature and rail temperatures. For the transient melt-wave calculation, the current output of each time step is computed to determine the motion characteristics and magnetic field excitation of the armature at that point. Subsequently, the magnetic field and temperature distributions are resolved using the finite volume method.

III. CALCULATIONS AND ANALYSIS

A. Calculation of steady state melt-wave

The magnetic induction intensity distribution of the armature and rail is primarily influenced by the initial magnetic field excitation and the speed of the armature, as evident from equation (1) and Fig. 3. These factors significantly impact the current density. Current density, as indicated in equation (9), plays a crucial role in altering the temperature distribution [24]. Therefore, it is essential to investigate the correlation between magnetic field excitation, armature speed, and the melt-wave. This study examines the impact of the velocity skin effect on the melt-wave and analyzes the temperature distribution of the armature with and without velocity. The erosion distance at various time points is investigated considering the influence of temperature on conductivity and compared to that under constant conductivity. The entire investigation is conducted based on $B_0 = 32$T.

Figure 5 (a) illustrates the progression of melt-waves at time intervals of 0.1 ms, 0.3 ms, and 0.5 ms with the armature speed set at 0. Initially, elevated temperatures are primarily localized at the region where the armature tail interfaces with the track. Over time, the melt-wave gradually extends towards the armature head, albeit at a relatively sluggish pace. The temperature distribution adjacent to the armature tail exhibits a wave-like pattern, diminishing towards the outer regions of the orbit.

Figure 5 (b) illustrates the propagation of melt-waves at various time points with an armature velocity of 100 m/s. In comparison to stationary armature, the melt-wave exhibits a notably broader spread over time, with armature erosion reaching 0.53 mm at 0.5 ms, whereas, at velocity, the armature erosion region forms a triangular shape, with erosion extending only to 0.13 mm at 0.5 ms. The advancement of the melt-wave at 100 m/s towards the armature head resembles a knife-like progression. This phenomenon is attributed to the speed skin effect, as depicted in Fig. 6. The rapid motion of the armature results in current concentration at the armature tail, leading to accelerated temperature rise and increased armature melting speed, thereby propelling the

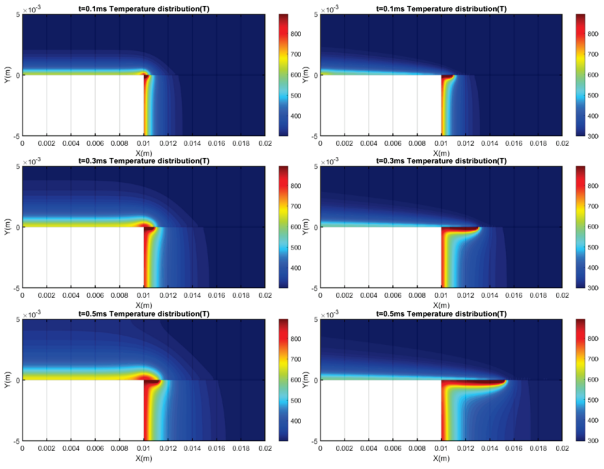


Fig. 5. Comparison of melt-wave: (a) Distribution of melt-waves when the armature is at rest and (b) Distribution of melt-waves at an armature speed of 100 m/s.

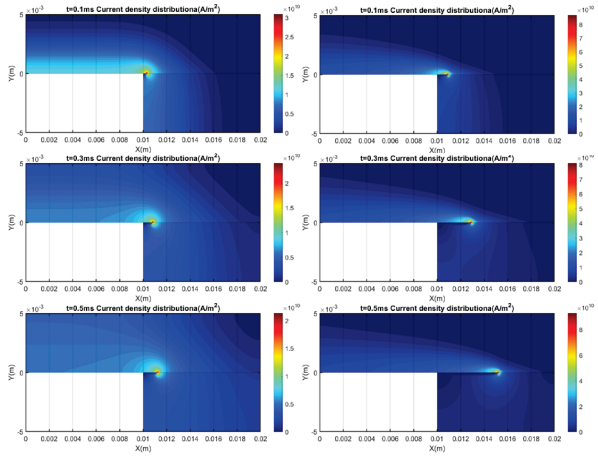


Fig. 6. Comparison of current density: (a) Distribution of melt-waves when the armature is at rest and (b) Distribution of melt-waves at an armature speed of 100 m/s.

melt-wave towards the armature head. Consequently, the temperature distribution along the trajectory gradually decreases as the armature moves, eventually approaching room temperature (300K) at the conclusion of the melt-wave.

Figure 7 displays the conductivity-temperature profiles of 7075Al and copper, derived through data fitting based on literature [17]. Subsequently, Fig. 8 illustrates the propagation of the melt-wave at a velocity of 100 m/s corresponding to the conductivity alteration. The erosion rate is observed to be higher when conductivity changes, primarily attributed to the conductivity variation accelerating magnetic diffusion. The temporal evolution of magnetic induction intensity along the measurement line

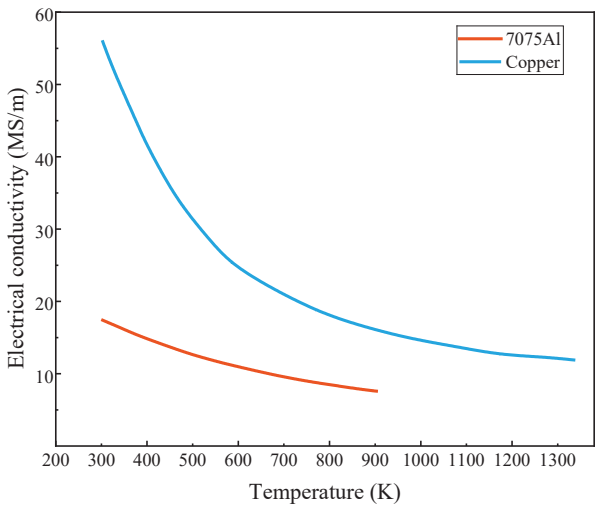


Fig. 7. Conductivity-temperature curve.

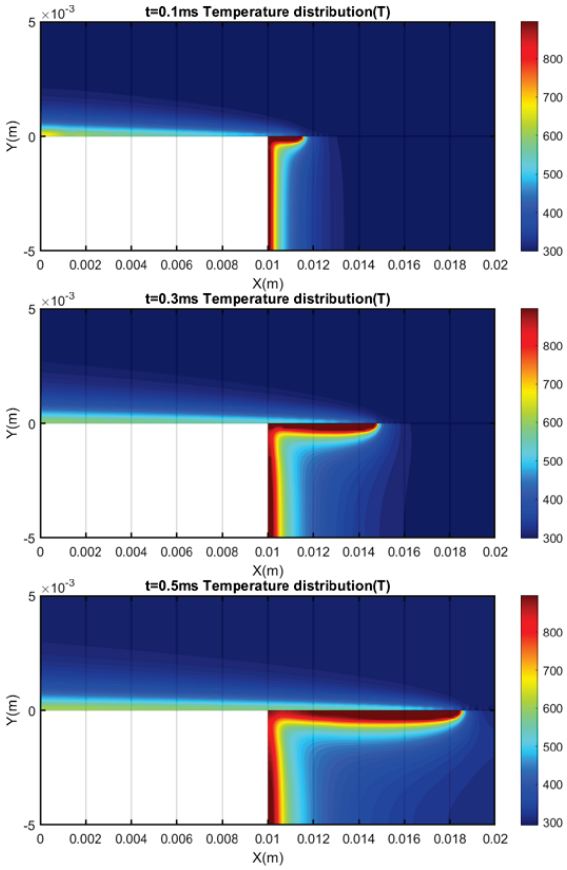


Fig. 8. Melt-wave distribution of conductivity change at armature speed 100 m/s.

is depicted in Fig. 9. Notably, under identical conditions, the magnetic induction intensity is greater with conductivity variation compared to its absence, with the impact

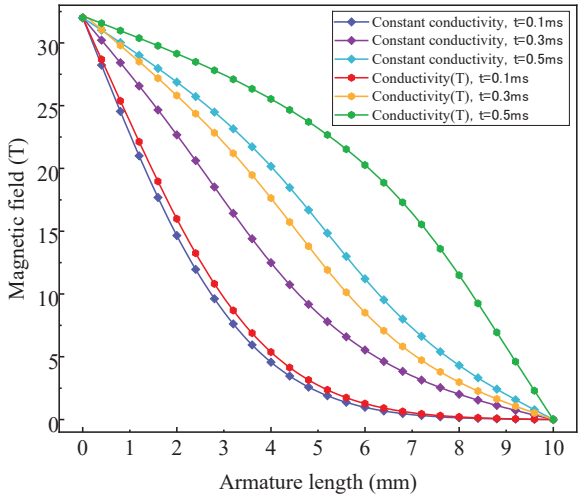


Fig. 9. Magnetic flux density distribution along the measurement line under different conditions.

of conductivity alteration on magnetic diffusion intensifying over time.

Table 1 displays the erosion rates calculated by the model in this study using mesh sizes of 0.05 mm and 0.1 mm, respectively, in comparison with findings from the literature [8, 11, 13, 18]. When v=100 m/s, the erosion velocities calculated by the two mesh sizes are identical. In the remaining scenarios, the erosion velocity discrepancy is a mere 0.1 mm. Given that the 0.1 mm mesh provides sufficient calculation accuracy while conserving computational resources, the results align with those of Parks and other researchers, falling within a reasonable range. Consequently, the proposed finite volume method for melting wave calculation in this study is deemed viable.

Figure 10 illustrates the correlation between melt-wave erosion depth and time under various conditions. The graph demonstrates a linear increase in erosion depth over time, with a consistent slope representing erosion velocity. This observation aligns with Parks' conclusion in the literature [18, 19] that the erosion velocity formula $v_m = \rho_1 \mu I^2 v / \rho_2 \omega^2 E_m$ is time-independent.

Figure 11 illustrates a non-linear variation in erosion depth in relation to the external magnetic field excitation at $v_m = 150$ m/s. An increase in magnetic excitation results in a deeper penetration of the melting wave into the armature head, leading to a larger growth amplitude. The erosion distance discrepancy between conductivity change and constancy is merely 0.1 mm at $B_0 = 20$T, whereas it escalates to 2.8 mm at $B_0 = 40$T.

Figure 12 illustrates the impact of armature speed on erosion depth at $B_0 = 40$T. An increase in armature speed results in a corresponding increase in erosion depth; however, this escalation gradually dimin-

Table 1: Comparison of erosion distances

Model	**Velocity of Erosion**		
	$B_0 = 32$T $v = 100$m/s	$B_0 = 40$T $v = 100$m/s	$B_0 = 40$T $v = 150$m/s
Parks [8]	10.1m/s	15.8m/s	23.7m/s
Barber and Dreizin [18]	8.6m/s	13.4m/s	20.0m/s
Tang et al. [11]	7.5m/s	12.2m/s	18.8m/s
Sun et al. [13]	6.4m/s	13.64m/s	16.68m/s
This paper (0.1 mm)	9.9m/s	14.8m/s	21.1m/s
This paper (0.05 mm)	9.8m/s	14.8m/s	21.0m/s

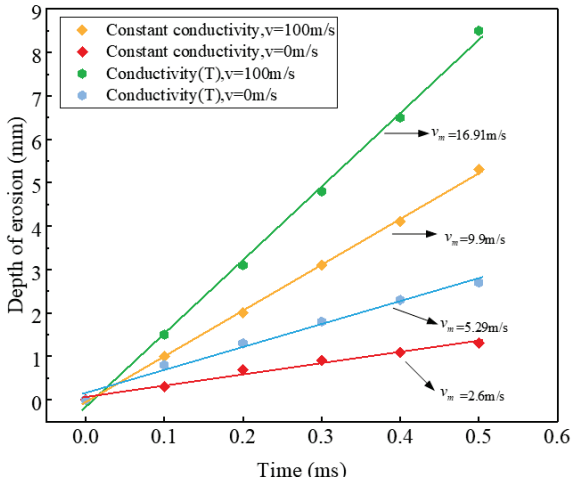


Fig. 10. Comparison of erosion rates.

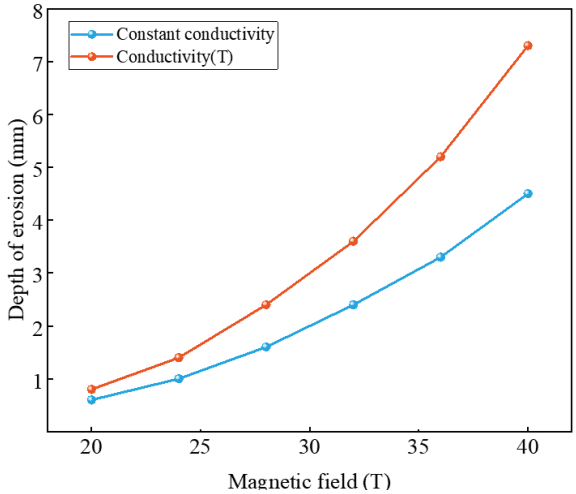


Fig. 11. Erosion distance in different magnetic fields.

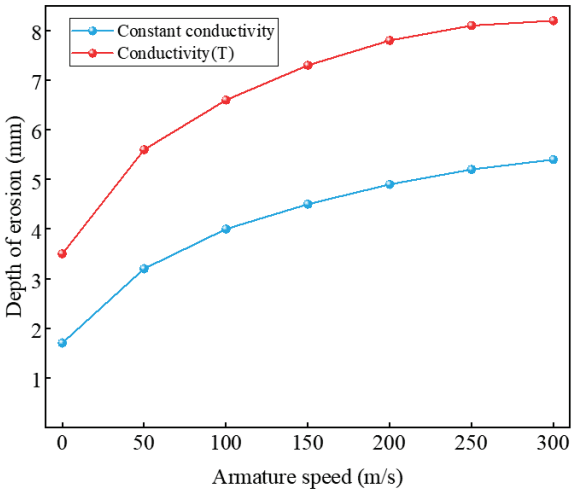


Fig. 12. Erosion distance at different armature speeds.

ishes. With higher velocities, the disparity in melting distances between the two conditions remains relatively constant.

B. Calculation of transient melt-wave

During the actual firing process, the armature is driven by a pulsed high current, which can be expressed as follows [20]:

$$I = \begin{cases} I_0 \sin(\pi t/(2t_1)), & t < t_1 \\ I_0 e^{-(t-t_1)/t_2}, & t \geq t_1 \end{cases}. \tag{16}$$

Given $I_0 = 0.3$MA, $t\_1 = 0.4$ ms, $t_2 = 0.6$ ms, and the inductance gradient is L= 0.5μH/m, the electromagnetic force F, velocity v, and the displacement l on the armature in motion can be expressed as follows [21]:

$$F = 0.5LI^2, \tag{17}$$

$$v = \int \frac{0.5LI^2}{m} dt, \tag{18}$$

$$l = \int v dt. \tag{19}$$

During the launching process, the initial excitation of the magnetic field and the speed of the armature are dictated by the external input current. Figure 13 illustrates the correlation between the external current, armature motion parameters, and time.

Figure 14 depicts the temporal evolution of armature temperature distribution at critical time instants (0.2 ms, 0.36 ms, and 0.8 ms) post current initiation. At the initial phase ($t = 0.2$ ms) corresponding to armature acceleration onset, the measured maximum localized temperature reaches 401K, remaining substantially below the bulk material's melting threshold. The subsequent thermomechanical response reveals surface degradation initiation at $t = 0.36$ ms, characterized by incipient phase-change phenomena. Progressive joule heating culminates in a solidified melt-wave penetration depth of 1.5 mm at $t = 0.8$ ms, demonstrating time-dependent electrothermal coupling effects.

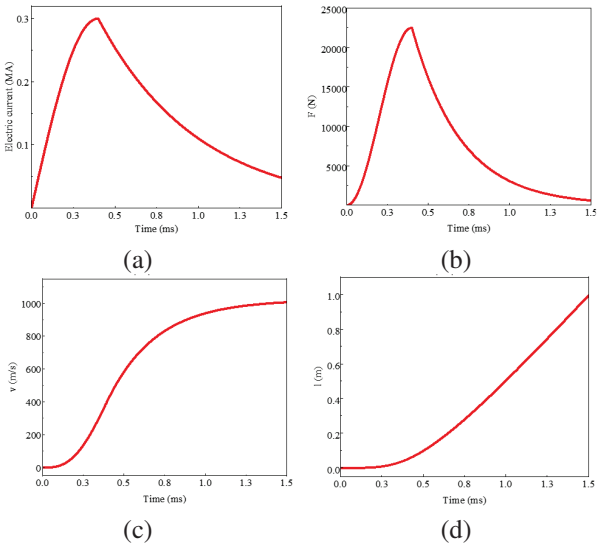


Fig. 13. Current and armature motion parameter curves: (a) External pulse high current waveform, (b) Electromagnetic force curve on the armature, (c) Armature speed curve, and (d) Armature displacement curve.

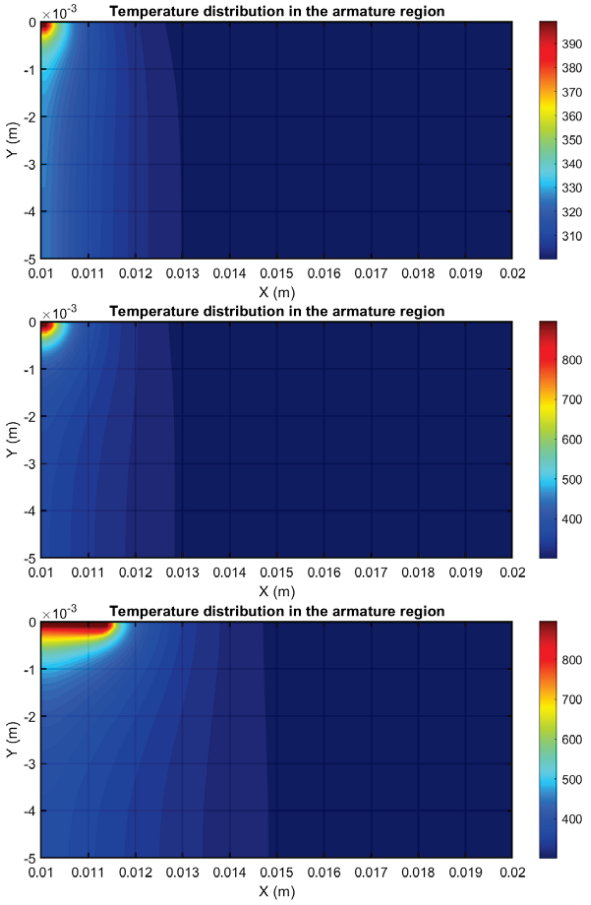


Fig. 14. Distribution of armature melt-waves at different moments after current connection.

This study investigates the influence of external current on melt-wave behavior by systematically varying the peak current magnitude and its temporal profile. Figure 15 presents the time-dependent depth of erosion evolution for currents of 0.24MA, 0.30MA, and 0.36MA. As shown in Fig. 13 (a), the current initially increases before gradually decaying. Correspondingly, the melt depth in Fig. 15 exhibits an initial growth phase followed by stabilization after 1 ms, resulting from a dynamic equilibrium between increasing armature velocity and decreasing current density. Post-stabilization depth of erosion stabilizes at 0.5 mm, 1.7 mm, and 3.2 mm for 0.24MA, 0.30MA, and 0.36MA,

respectively, indicating significant suppression of thermal erosion with current reduction. However, excessive current reduction compromises armature muzzle velocity, limiting practical applicability. Comparison with literature [13] confirms consistent temporal trends and maximum depth of erosion (1.7 mm) at 0.30MA, validating the experimental methodology.

The concurrent increase in current and armature velocity synergistically enhances thermal erosion. Figure 16 demonstrates that shorter armature acceleration durations correlate with earlier erosion initiation and reduced stabilization distances. This correlation arises because rapid current rise rates intensify instantaneous joule heating, accelerating phase-change processes. As evidenced in Fig. 15, an optimization strategy

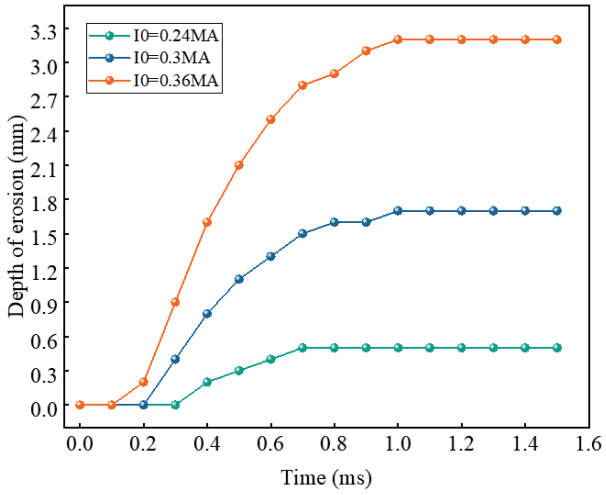


Fig. 15. Distance of erosion at different peak currents.

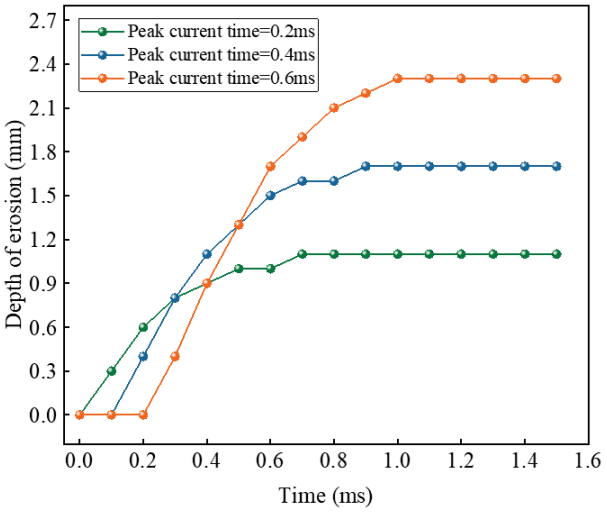


Fig. 16. Erosion distance for different rise times.

involves elevating current amplitude while compressing rise time to minimize melt formation without compromising armature launch precision. This approach leverages the competing effects of enhanced electromagnetic propulsion and reduced thermal exposure duration.

IV. CONCLUSION

The penetration of the melt-wave through the armature surface creates a significant gap between the armature and the rail, leading to armature transition. Therefore, this paper proposes, for the first time, a two-dimensional numerical model of the melt-wave in electromagnetic railgun launches based on the finite volume method. By comparing the results with those from four previous studies, the model demonstrates consistency, confirming the applicability of the finite volume method for investigating armature characteristics in electromagnetic railgun systems.

Steady-state calculations of the melt-wave reveal that high temperatures initially concentrate at the armature tail and propagate toward the head, driven by the velocity skin effect. Under constant conditions, the erosion distance increases linearly over time. When considering the temperature dependence of electrical conductivity, the erosion distance increases more significantly. Additionally, increases in initial magnetic excitation and armature velocity also lead to greater erosion distances.

Upon applying an external current, changes in initial magnetic excitation and armature velocity are determined by the current waveform. As the launch time increases, the erosion distance initially rises sharply before stabilizing. Reducing the current rise time or magnitude effectively inhibits the progression of erosion.

This study provides a theoretical foundation for improving the launch lifespan of electromagnetic railguns. To better approximate real-world launch conditions, future research should consider the effects of non-ideal contact on armature melting.

REFERENCES

[1] J. Lydia, R. Karpagam, and R. Murugan, "A novel technique for dynamic analysis of an electromagnetic rail launcher using FEM coupled with Simplorer," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 37, pp. 229-237, 2022.

[2] J. Zhang, J. Lu, S. Tan, B. Li, and Y. Zhang, "Research status of surface damage in rails for electromagnetic launchers," *Acta Armamentarii*, vol. 44, pp. 1908-1919, 2023.

[3] X. Wan, J. Lu, D. Liang, and J. Lou, "Thermal analysis in electromagnetic rail launcher taking friction heat into account under active cooling condition," *IEEE Access*, vol. 8, pp. 84720-84740, 2020.

[4] B. Gu and C. Yang, "A review of the computational investigation of three-dimensional transient magneto-thermal coupling characteristics of electromagnetic railgun armature and rail in high-speed sliding electrical contact," *J. Phys.: Conf. Ser.*, vol. 2478, no. 8, p. 082016, June 2023.

[5] S. Ren, G. Feng, and S. Liu, "Study on wear of electromagnetic railgun," *IEEE Access*, vol. 10, pp. 100955-100963, 2022.

[6] S. Ma, S. Lu, H. Ma, H. Wang, A. Nong, D. Ma, C. Yan, and C. Liu, "Investigation on the spatial-temporal distribution of electromagnetic gun rail temperature in single and continuous launch modes," *IEEE Trans. Plasma Sci,* vol. 50, pp. 2270-2278, July 2022.

[7] L. Chen, X. Xu, Z. Wang, J. Xu, P. You, and X. Lan, "Melting distribution of armature in electromagnetic rail launcher," *IEEE Trans. Plasma Sci,* vol. 51, pp. 234-242, Jan. 2023.

[8] P. B. Parks, "Current melt-wave model for transitioning solid armature," *Journal of Applied Physics*, vol. 67, no. 7, pp. 3511-3516, Apr. 1990.

[9] F. Stefani, R. Merrill, and T. Watt, "Finite element analysis of melt wave ablation in electromagnetic rail launcher armatures," *IEEE Trans. Magn.*, vol. 41, no. 1, pp. 437-441, Jan. 2005.

[10] F. Gong. "Study of erosion mechanism at the armature-rail contact interface in railgun," Thesis, Nanjing University of Science and Technology, Nanjing, Jiangsu, China, 2014.

[11] B. Tang, Y. Xu, Q. Lin, and B. Li, "Synergy of melt-wave and electromagnetic force on the transition mechanism in electromagnetic launch", *IEEE Trans. Plasma Sci,* vol. 45, pp. 1361-1367, July 2017.

[12] S. Tan, J. Lu, X. Zhang, X. Guan, and X. Long, "Two-dimensional numerical simulation of melt-wave erosion in solid armatures," *J. Xi'an Jiaotong Uni.*, vol. 50, no. 3, pp. 106-111, Mar. 2016.

[13] J. Sun, J. Cheng, Q. Wang, L. Xiong, Y. Cong, and Y. Wang, "Numerical simulation of melt-wave erosion in 2-D solid armature," *IEEE Trans. Plasma Sci.*, vol. 50, no. 4, pp. 1032-1039, Apr. 2022.

[14] J. Cen and Q. Zou, "Deep finite volume method for partial differential equations," *Journal of Computational Physics*, vol. 517, p. 113307, July 2024.

[15] B. Li, J. Lu, S. Tan, Y. Jiang, and Y. Zhang, "Application of finite volume method in analyzing sliding electrical contact problem," *Journal of Naval Univ. of Engineering*, vol. 31, no. 6, pp. 23-28, Dec. 2019.

[16] Y. Yang, Q. Yin, C. Li, H. Li, and H. Zhang, "Simulation and experimental verification of magnetic field diffusion at the launch load during electromagnetic launch," *Sensors*, vol. 23, no. 18, p. 8007, 2023.

[17] K. T. Hsieh and B. K. Kim, "International railgun modelling effort," *IEEE Trans. Magn.,* vol. 33, no. 1, pp. 245-248, Jan. 1997.

[18] J. P. Barber and Y. A. Dreizin, "Model of contact transitioning with 'realistic' armature-rail interface," *IEEE Trans. Magn.*, vol. 31, no. 1, pp. 96-100, Jan. 1995.

[19] T. Watt and F. Stefani, "The effect of current and speed on perimeter erosion in recovered armatures," *IEEE Trans. Magn.*, vol. 41, no. 1, pp. 448-452, Jan. 2005.

[20] J. D. Powell and B. K. Kim, "Observation and simulation of solid armature railgun performance," *IEEE Trans. Magn.*, vol. 24, no. 1, pp. 84-89, Jan. 1999.

[21] A. Guo, X. Du, X. Wang, and S. Liu, "Calculation of armature melting wear rate based on contact surface heat distribution," *IEEE Trans. Plasma Sci.*, vol. 51, no. 7, pp. 2981-2990, July 2024.

[22] C.-C. Wu, D. Völker, S. Weisbrich, and F. Neitzel, "The finite volume method in the context of the finite element method," *Materials Today: Proceedings*, vol. 62, pp. 2679-2683, 2022.

[23] G. Liao, W. Wang, B. Wang, Q. Chen, and X. Liu, "Transient mixed-lubrication and contact behavior analysis of metal liquid film under magneto-thermal effect," *International Journal of Mechanical Sciences*, vol. 271, p. 109142, June 2024.

[24] J. Yao, L. Chen, S. Xia, J. He, and C. Li, "The effect of current and speed on melt erosion at rail-armature contact in railgun," *IEEE Trans. Plasma Sci.*, vol. 47, no. 5, pp. 2302-2308, May 2019.

[25] G. Fei and W. Chunsheng, "Two-dimensional numerical simulation of melt-wave erosion in solid armatures," *J. Nanjing Univ. Sci. Technol.*, vol. 36, no. 3, pp. 487-491, 2012.

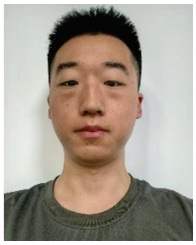
Kefeng Yang received the B.S. degree in weapons launch engineering from Air Force Engineering University, Xi'an, China, in 2023. He is currently a graduate student in mechanics at Air Force Engineering University, where his interests include theory and technology of weapon systems, especially about the electromagnetic launch technology.

Gang Feng received the M.S. degree in weapons from Air Force Engineering University, Xi'an, China. He is currently a Professor at the Air and Missile Defense College, Air

Shaowei Liu received the B.S. and Ph.D. degrees in weapons from the Air Force Engineering University, Xi'an, China, in 2005 and 2008, respectively. He is currently an Associate Professor at the Air Defense and Anti-Missile School, Air Force Engineering University. His current research interests include the theory and technology of weapon systems.

Xiaoquan Lu received the B.S. degree in weapons launch engineering from Air Force Engineering University, Xi'an, China, in 2022. He is currently a graduate student in mechanics at Air Force Engineering University, where his interests include theory and technology of weapon systems.

Xiangyu Du received the B.S. degree in weapons in 2020 and the M.S. degree in mechanics in 2023 from the Air Force Engineering University, Xi'an, China, where he is currently pursuing the Ph.D. degree in weapons. His current research interests include theory and technology of weapon systems, especially about electromagnetic launch technology.



Tianyou Zheng received the B.S. degree in weapons launch engineering from Air Force Engineering University, Xi'an, China, in 2023. He is currently a graduate student in mechanics at Air Force Engineering University, where his interests include contact characteristics of railguns under pulsed high current conditions.

Design of GNSS/INS Coupled Navigation Algorithm Using Adaptive Neuro-Fuzzy Inference Systems

Chen Zerui, Xiao Yanhong, Wu Xin, Hu Houpeng, Xiao Jian, Yang Shang, and Gao Zhenghao

Guizhou Power Grid Co. Ltd.
Electric Power Science Research Institute, Guizhou, Guiyang 550002, China
chen_ze011@outlook.com, xiao_yanhong18@outlook.com, wu_xin96@outlook.com, hu_houpeng85@outlook.com
xiao_jian25@outlook.com, yang_shang18@outlook.com, gao_zhenghao96@outlook.com

***Abstract* –** Traditional GNSS/INS (Global Navigation Satellite Systems/Inertial Navigation Systems) coupled navigation algorithms often struggle with accuracy in GNSS-denied or challenging environments. This paper introduces a novel adaptive fusion algorithm leveraging an Adaptive Neuro-Fuzzy Inference System (ANFIS) that dynamically adjusts sensor weightings based on real-time signal quality and system performance. The core innovation lies in the real-time integration of fuzzy logic and neural network learning, enabling the system to continuously adapt and optimize its decision-making rules for navigation accuracy. A comprehensive, dynamic error source model is constructed incorporating GNSS atmospheric delays, orbit errors, and INS drift to enhance the learning-driven weight adjustment mechanism. The resulting ANFIS-based fusion strategy shows significant superiority over traditional Kalman-based methods, achieving over 90% robustness across harsh scenarios with an average execution time of 0.69 seconds, demonstrating improved adaptability, learning capability, and fault resilience in dynamic environments.

***Index Terms* –** Adaptive neuro-fuzzy inference system, data fusion, GNSS/INS coupled navigation, signal loss, system robustness.

I. INTRODUCTION

GNSS (Global Navigation Satellite Systems) and INS (Inertial Navigation Systems) are widely used in modern navigation technology, especially in autonomous driving, aerospace, and drones. GNSS provides global positioning information but is limited by signal shielding, multipath effects, and atmospheric delays, especially in environments such as high-rise buildings and tunnels in cities, which can easily lead to positioning errors or signal loss. INS provides high-frequency dynamic positioning through inertial sensors (accelerometers and gyroscopes), but its errors accumulate over time, gradually decreasing accuracy. GNSS/INS (Global Navigation Satellite Systems/Inertial Navigation Systems) coupled navigation technology came into being to overcome the shortcomings of GNSS and INS. EGNSS refers to the European Global Navigation Satellite System, developed by the European Union and primarily represented by the Galileo program. It complements other systems like GPS to improve positioning accuracy and reliability, especially in Europe. Traditional methods such as Kalman filtering have poor robustness in harsh environments and achieving high-precision and stable navigation is difficult.

Many researchers have applied different optimization schemes in response to the challenges of GNSS/INS coupled navigation technology. Abdolkarimi and Mosavi [1] proposed an improved neuro-fuzzy system to handle the prediction of INS positioning errors during long-term GPS outages. Cong et al. [2] found that the random error characteristics of inertial sensors and the instability caused by GNSS signal interruption threaten MEMS-based GNSS/INS vehicle navigation systems and improved positioning accuracy through the data fusion method of support vector machines. Traditional extended Kalman filters (EKF) and unscented Kalman filters (UKF) have been widely used in GNSS and INS data fusion. Traditional GNSS/INS fusion methods like EKF and UKF struggle in complex environments such as tunnels and urban canyons due to their poor robustness under GNSS signal degradation or loss [3–5]. Their reliance on linearity and Gaussian noise assumptions limits effectiveness in nonlinear and noisy real-world conditions. Fixed parameters also reduce adaptability to dynamic changes [6]. Moreover, cumulative INS errors, particularly drift, remain uncorrected during GNSS outages [7]. Conventional models cannot effectively address nonlinear errors like multipath and atmospheric disturbances [8]. Even fuzzy logic and machine learning approaches often lack sufficient real-time adaptability in dynamic scenarios.

Submitted On: January 2, 2025
Accepted On: June 11, 2025

https://doi.org/10.13052/2025.ACES.J.400808

Given the limitations of traditional GNSS/INS data fusion methods, the research community has been committed to exploring the fusion application of advanced technologies such as fuzzy logic and neural networks in recent years. Neural networks have significant advantages in dealing with complex nonlinear problems, while fuzzy logic can effectively deal with system uncertainty. Early research shows that fuzzy inference systems can improve robustness by adapting GNSS/INS fusion weights [9–10]. However, most of these studies are based on static models for analysis and research, and many also face difficulties: lack of real-time adaptability and difficulty adapting to complex and changeable dynamic environments [11].

This study uses the GNSS/INS coupled navigation algorithm of Adaptive Neuro-Fuzzy Inference Systems (ANFIS) to integrate the adaptive characteristics of fuzzy logic and the learning ability of neural networks to achieve the function of quickly responding to environmental changes. Enhancing ANFIS for GNSS/INS navigation involves integrating electromagnetic signal propagation models that account for diffraction, reflection, refraction, and attenuation in complex environments. These models provide key variables like multipath effects, signal delays, and local electromagnetic field variations, which can be used as inputs for ANFIS. By including these propagation-aware parameters, ANFIS can more accurately detect signal degradation and distinguish between different types of errors. This allows for proactive adjustment of GNSS and INS weightings in the fusion process, maintaining accuracy even during signal obstruction or interference. Real-time implementation can utilize ray-tracing, urban environment maps, or machine learning-based estimators. Overall, this integration boosts robustness, precision, and adaptability of the navigation system in dynamic and harsh conditions.

This paper introduces a GNSS/INS coupled navigation algorithm using ANFIS, which integrates fuzzy logic and neural networks to handle nonlinear and uncertain environments. It achieves real-time data fusion and dynamic weight adjustment based on environmental conditions, enhancing accuracy and robustness. The approach models and corrects complex error sources such as signal loss, sensor drift, and atmospheric interference. With an average runtime of just 0.69 seconds, the algorithm meets high-performance computing demands. Its robust design parallels advanced numerical modeling in mechanics and offers transferable values across domains. Applications include autonomous navigation, structural monitoring, and real-time system control. By constructing a precise error source model and designing an adaptive data fusion strategy and dynamic adjustment mechanism, the efficient fusion of GNSS and INS data is completed. Particularly when the GNSS signal is unstable or interrupted, ANFIS can instantly adjust the INS weight to ensure system continuity and accuracy [12]. Experimental results show that the algorithm has a robustness of over 90% in various harsh environments and a running time of 0.69 seconds. Compared with traditional algorithms, its performance is significantly improved, overcoming the robustness challenge in complex environments.

Traditional GNSS/INS-enabled navigation systems struggle to maintain accuracy and resilience in complex and difficult environments such as urban canyons, tunnels, and signal interference or multipath zones. GNSS suffers from signal loss and degradation, whereas INS errors build over time. Because of their linear assumptions and set parameter configurations, traditional data fusion approaches such as UKF and EKF lack adaptability and robustness in such environments. Furthermore, current intelligent technologies, such as static fuzzy logic and neural network models, frequently fail to respond in real time to dynamic environmental changes.

This paper proposes a novel GNSS/INS coupled navigation algorithm based on ANFIS. The main contributions include:

1. Developing a comprehensive GNSS/INS error model accounting for GNSS signal degradation (e.g., atmospheric delay, multipath) and INS drift and bias.
2. Designing an adaptive neuro-fuzzy system combining fuzzy logic and neural networks to dynamically adjust GNSS/INS data fusion.
3. Implementing a strategy to automatically shift reliance between GNSS and INS based on signal quality and environmental changes.
4. Achieving over 90% robustness and maintaining low positioning errors even under signal loss, interference, and multipath effects.
5. Demonstrating fast processing with an average running time of 0.69 seconds, outperforming EKF and UKF methods in both speed and accuracy.

The primary aim of this study is to develop a robust, high-precision, real-time adaptive GNSS/INS coupled navigation system by leveraging the capabilities of ANFIS. This system is intended to overcome the limitations of traditional fusion methods and provide reliable navigation performance under dynamic and harsh environmental conditions.

II. RELATED WORK

Wu [13] enhanced a GPS/INS ultra-tightly-coupled system using a neural-fuzzy adaptive Kalman filter, where a fuzzy controller adjusts measurement noise

based on innovations and a neural network optimizes its parameters. This ANFIS-based approach improves accuracy and robustness in dynamic noise conditions, supporting reliable GNSS/INS navigation. Lai et al. [14] addressed high uncertainty in tightly-coupled GNSS/INS systems using an adaptive fuzzy neural network-aided progressive Gaussian filter. They jointly estimated step size and measurement noise via a variational Bayesian approach. A fuzzy algorithm adapted measurement noise, while a neural network refined state error covariance. This formed an ANFIS-based filtering framework. The approach improved estimation robustness and accuracy. Experiments showed superior performance over existing filters.

Yue et al. [15] developed a robust GNSS/INS navigation algorithm for urban and GNSS-denied environments. A two-tier fusion scheme was used to enhance accuracy. SVR-AKF adaptively fuses GNSS and INS data by tuning Kalman filter parameters, ANFIS models and corrects INS drift during GNSS outages. This approach effectively handled sensor nonlinearities and uncertainties. Road tests confirmed improved positioning performance.

Wang et al. [16] used a centralized Kalman filter with loose coupling for efficient GNSS/INS integration. Feedback correction maintained accuracy over long durations. The system is simple, reliable, and suitable for low-power UAVs. It offers a solid baseline for intelligent methods like ANFIS. ANFIS can adapt to nonlinear errors and improve performance in dynamic conditions. This foundation supports the development of an adaptive neuro-fuzzy navigation algorithm.

Mahdi et al. [17] proposed a system integrating RISS with GNSS using ANFIS for accurate navigation in GNSS-degraded areas. ANFIS calibrated sensor errors, enhancing reliability during outages. It outperformed traditional methods in 2D RMSE and maximum error. Real-world tests confirmed its robustness. The method is scalable to low-cost IMUs. This makes it suitable for diverse navigation applications.

Wang et al. [18] addressed GNSS/INS integration challenges using tight coupling, non-holonomic constraints, and outlier rejection to enhance robustness and accuracy. These methods are effective during GNSS outages and sensor drift. ANFIS can further improve this by learning nonlinear system behaviors and adapting in real-time. Its data-driven approach outperforms traditional filters in dynamic conditions. Thus, Wang et al. provided a strong foundation for integrating ANFIS into navigation algorithms.

Gudivaka and Thanjaivadivel [19] presented an IoT-based adaptive signal processing framework to enhance robotic navigation through intelligent sensor fusion. Their adaptive fusion strategy is leveraged in this research by integrating GNSS and INS data using ANFIS to handle nonlinearities and sensor drift. This integration improves navigation accuracy, robustness to signal loss, and adaptability in dynamic environments.

Sun et al. [20] demonstrated effective GNSS/compass fusion using ANFIS for real-time vehicle positioning and car-following status detection. The method accurately interpreted dynamic parameters with a low false alarm rate. This parallels GNSS/INS integration, where INS can enhance state estimation. ANFIS proved capable of handling nonlinear, uncertain sensor data. Their results validated the feasibility of real-time intelligent inference. This supports the development of an ANFIS-based GNSS/INS navigation algorithm.

The MCPP method by Zhao and Yang [21] supports ANFIS-based GNSS/INS design by providing satellite fault probabilities as adaptive inputs. These enhance ANFIS's ability to model nonlinear errors and handle GNSS faults. Both aim to improve urban positioning accuracy. MCPP's probabilistic approach complements ANFIS's fuzzy logic, avoiding fixed thresholds. Its use in tightly-coupled systems ensures compatibility. Real-world tests validate its effectiveness for intelligent navigation fusion.

Jwo et al. [22] highlighted the importance of GNSS/INS integration for robust navigation in GNSS-denied environments. They emphasize AI methods, particularly ANNs, for modeling without detailed physical system knowledge. Their support for ANN-based fusion aligns with the shift toward intelligent systems like ANFIS. ANFIS combines neural learning with fuzzy logic for adaptive performance. The study's focus on performance evaluation supports assessing ANFIS-based models. Overall, their findings justify developing advanced GNSS/INS fusion algorithms using ANFIS. The work by Wu et al. [23] enhances GNSS/INS integration using adaptive deep learning to correct INS errors from nonlinear noise. It reflects ANFIS principles by modeling without strict noise assumptions and combining learning with filtering for interpretability. The model emphasizes efficiency and generalization. Experimental results under GNSS loss show significant accuracy gains, supporting ANFIS-based adaptive navigation algorithm design.

Zhang and Wang [24] proposed the DGP-MLP method to enhance GNSS/INS navigation during outages, aligning with ANFIS-based approaches. Their hybrid model supports the integration of fuzzy logic and learning, with adaptive training strategies similar to ANFIS tuning. Using past GNSS data for prediction also informs ANFIS input design. Real-world results validate its effectiveness, supporting the development of ANFIS-

Table 1: Comparative analysis of ANFIS-based GNSS/INS integration methods

Reference	Method	Advantage	Limitation
Wu [13]	Neural-fuzzy adaptive Kalman filter with ANFIS	Adapts to dynamic noise; improves accuracy and robustness	Complex tuning; not real-time
Lai et al. [14]	ANFIS + progressive Gaussian filter + variational Bayes	Joint estimation of step size and noise; high robustness	Computationally heavy
Yue et al. [15]	SVR-AKF + ANFIS to model INS drift	Robust in urban/GNSS-denied areas; two-level fusion	SVR may not generalize well
Wang et al. [16]	Centralized Kalman filter (loose coupling) with ANFIS potential	Low-power, simple system; suitable for UAVs	Loose coupling limits integration accuracy
Mahdi et al. [17]	RISS-GNSS-ANFIS integration	High accuracy in degraded areas; scalable to low-cost IMUs	Focused on 2D; needs expansion to 3D/6DOF
Wang et al. [18]	Tight coupling + non-holonomic constraints + outlier rejection	Resilient to sensor drift and GNSS loss	Relies heavily on constraints and threshold tuning
Gudivaka & Thanjaivadivel [19]	IoT-based ANFIS sensor fusion for robots	Adaptive to environment; robust to drift	Limited to robotics context; generalization may vary
Sun et al. [20]	GNSS/Compass + ANFIS	Real-time, low false alarms; dynamic adaptation	Less applicable to full GNSS/INS integration
Zhao & Yang [21]	MCPP + ANFIS	Probabilistic satellite fault handling; better error modeling	May need high computational load for MCPP
Jwo et al. [22]	ANN fusion framework	Supports learning without system model	Lacks interpretability; ANN-only
Wu et al. [23]	Adaptive deep learning for INS correction	Generalizable, interpretable; no strict noise assumptions	Deep models may overfit or need large data
Zhang & Wang [24]	DGP-MLP hybrid; uses GNSS history + fuzzy logic	Supports ANFIS input design; real-world tested	Not inherently real-time ANFIS
Guo & Tu [25]	LSTM + EMD/Wavelet + EKF	Predictive during GNSS outage; interpretable	LSTM training is time-intensive; sensitive to outliers

based navigation systems. Guo and Tu [25] used LSTM to predict pseudo-GNSS data during outages from IMU and INS inputs, a role ANFIS can also fulfill. EMD and wavelet filtering improve data quality. ANFIS offers interpretable, adaptive learning. Like LSTM, it trains with GNSS and predicts during loss. EKF fuses predictions to correct INS drift, supporting ANFIS-based GNSS/INS design.

Table 2 provides a comparative overview of recent research efforts integrating ANFIS into GNSS/INS navigation algorithms. It summarizes each study's methodology, highlights the key advantages that ANFIS brings (such as improved robustness, adaptability, and accuracy in complex or GNSS-denied environments) and outlines the main limitations or challenges faced, including computational complexity, tuning difficulty, and real-time adaptability. This comparison helps to identify strengths and gaps in current approaches, guiding future developments in intelligent navigation systems.

III. SYSTEM DESIGN AND IMPLEMENATION

A. Problem modeling and error sources

1. GNSS error model

GNSS signal errors mainly come from atmospheric delay, satellite orbit error, multipath effect, and signal shielding. Traditional GNSS/INS fusion navigation often simplifies GNSS errors into white or Gaussian noise while ignoring the uncertainty in the actual environment. Especially under multipath and shielding conditions, the navigation positioning accuracy is significantly reduced. To more accurately describe GNSS errors, this study constructs a dynamic error model, which comprehensively considers the influence of satellite geometry, multipath effect, and environmental shielding. By applying satellite geometry error, shielding condition, and reflected signal strength as input, a dynamic weight adjustment strategy is adopted to achieve real-time error estimation and correction. In response to multipath or

shielding problems, this study combines satellite orbit and position error models for further optimization to improve positioning accuracy in complex environments:

$$E_{\text{GNSS}} = \alpha \cdot \Delta\theta + \beta \cdot \Delta\text{SNR} + \gamma \cdot \text{Multipath} + \delta \cdot \text{Blockage} + \varepsilon. \quad (1)$$

E_{GNSS} is the GNSS positioning error in meters (m). $\Delta\theta$ is the satellite geometric error, indicating the change in the satellite geometric structure, which is usually related to the number and position of satellites. $\triangle$SNR is the signal strength change, the signal-to-noise ratio change. Multipath is the multipath effect, which indicates the error caused by signal reflection or refraction. Blockage is the signal occlusion, indicating environmental factors' impact on signal reception. ε is random noise or other unmodeled error sources. α, β, γ, and δ are weight factors which indicate the contribution of different error sources to the positioning error.

2. INS error model

The main sources of error in the INS system include sensor offset, zero drift, calibration error, and inertial integration error, especially the drift problem of the accelerometer and gyroscope. Since the error of INS is cumulative, the error gradually amplifies during long-term operation, affecting the system's accuracy. This study analyzes the noise characteristics of INS sensors and constructs the error propagation model of the accelerometer and gyroscope. The Kalman filter framework is used to dynamically correct INS's zero drift and sensor offset. For the drift error of the accelerometer and gyroscope, the study applies an adaptive calibration method based on real-time sensor feedback and model prediction, continuously correcting the error gain during the navigation process, thereby reducing the cumulative error of INS. Further, through the periodic correction of inertial data and the dynamic optimization of the error gain matrix, the accuracy stability of INS in long-term operation is enhanced, ensuring that it can provide relatively accurate positioning information even without GNSS support. The error gain update equation is:

$$\mathbf{K}_{\text{adapt}} = \frac{1}{\lambda} \sum_{i=1}^{n} \left(\mathbf{e}_i \mathbf{e}_i^T \right), \quad (2)$$

where $\mathbf{K}_{\text{adapt}}$ is the adaptive error gain, $\mathbf{e}_i$ is the error at the first moment, λ is the forgetting factor.

To solve the problem of GNSS and INS error coupling, a comprehensive error source model is constructed and integrates the geometric, multipath, and occlusion errors of GNSS and the drift and integral errors of INS. Through in-depth analysis of the error characteristics, fuzzy logic is used to dynamically allocate the weights of each error source. When the GNSS signal quality decreases, the system intelligently increases the INS weight to stabilize positioning. When the signal is stable, the GNSS weight is strengthened to improve accuracy. The model optimizes data fusion based on the real-time transmitted data and environmental factors to complete the real-time correction of error source parameters.

B. ANFIS architecture design

1. Fuzzy logic and neural network

The fuzzy logic component of ANFIS is specifically designed to handle the uncertainty factors in the GNSS/INS fusion process. It first converts the two continuously changing indicators (GNSS signal quality and INS error) into fuzzy sets through membership functions. The system derives the corresponding weight adjustment factor and error correction amount using the IF-THEN fuzzy inference mechanism. Next, using a fuzzy rule base, the system flexibly adjusts the weight allocation between GNSS and INS based on the current signal quality and error characteristics. This mechanism ensures that, when the GNSS signal is unstable or missing, the system can maintain highly accurate and stable positioning performance by augmenting the INS.

The neural network in the ANFIS system concentrates on optimizing the rules and parameters of the affiliation function in fuzzy logic to enhance the system's ability to adapt to the dynamic environment. The system accurately adjusts the center value and width of the affiliation function. It optimizes the weights and bias parameters to minimize the error using the feed-forward neural network structure and the backpropagation algorithm. Even in extreme cases, such as high occlusion of the GNSS signal, the system still maintains a stable and reliable performance. The process is displayed in Fig. 1.

Fuzzy logic and neural networks enhance the error source model for GNSS/INS integration via the ANFIS. ANFIS effectively models GNSS errors like atmospheric delay, multipath effects, and signal blockage as well as INS errors, including sensor drift. Fuzzy logic addresses uncertainty and nonlinearity by assigning weights based on signal and environmental conditions, while neural networks optimize fuzzy membership functions and rules through learning. Fuzzy logic transforms real-time variables such as GNSS signal quality into fuzzy sets and utilizes IF-THEN rules for weight allocation between GNSS and INS, ensuring stable navigation. The neural network component refines fuzzy rules using backpropagation and gradient descent, allowing adaptation to real-time changes. ANFIS dynamically adjusts weightings, improving navigation accuracy and system robustness, achieving over 90% reliability with a 0.69-second average response time.

2. ANFIS model training and optimization

ANFIS uses supervised learning combined with the gradient descent method and backpropagation algorithm to optimize the neural network's fuzzy rules and various parameters. IF-THEN logic statements used in

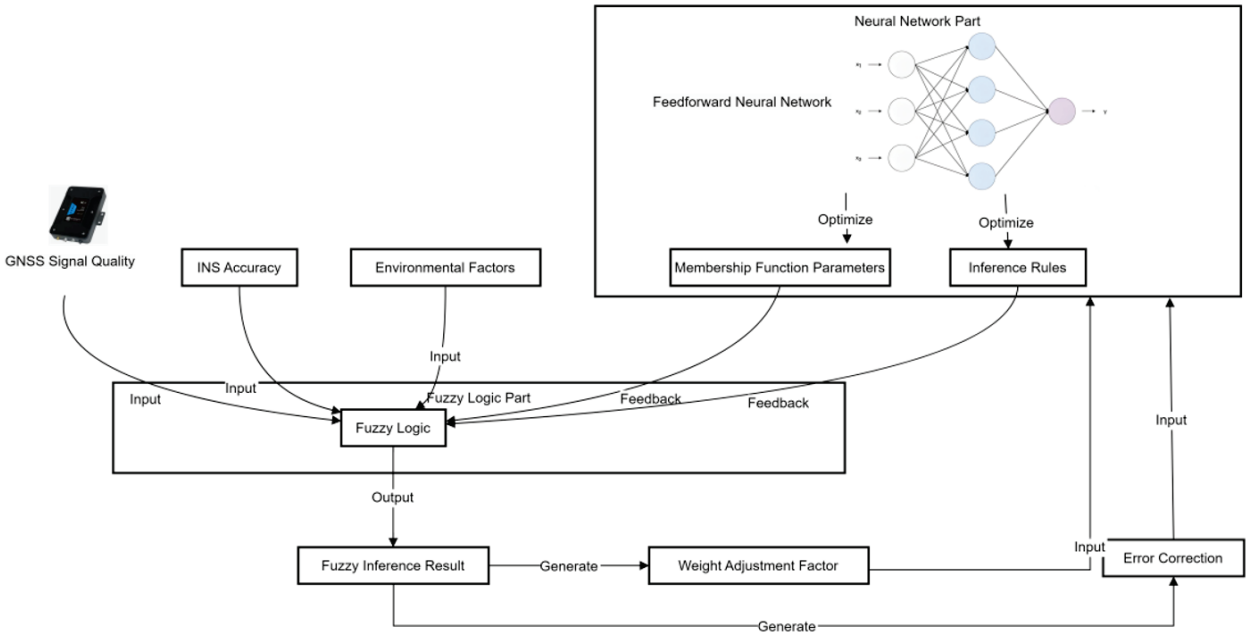


Fig. 1. ANFIS process in the GNSS/INS coupled navigation system.

fuzzy systems to handle uncertainty. They enable adaptive decision-making in GNSS/INS systems by adjusting weights based on conditions like signal quality. In the training process, the backpropagation algorithm minimizes the prediction error. It gradually adjusts the weight of the neural network to optimize the membership function parameters in the fuzzy rule base. At the same time, the output weight in the fuzzy reasoning process is adjusted by the gradient descent method to improve the data fusion strategy. Through multiple training cycles, the model gradually adapts to different environmental changes, adjusts the fusion ratio of GNSS and INS in real-time, enhances the role of INS when the GNSS signal is unstable, and strengthens the weight of GNSS when the GNSS signal is strong, ensuring that the system always maintains the best navigation accuracy. Finally, the trained ANFIS model automatically adjusts and optimizes the data fusion strategy in various environments, improving the accuracy and robustness of the GNSS/INS coupled navigation system. The ANFIS model utilizes supervised learning, merging gradient descent and backpropagation to refine fuzzy rules and network parameters. The backpropagation algorithm minimizes prediction errors by adjusting neural network weights, enhancing membership function parameters. Simultaneously, output weights in the fuzzy inference process are updated via gradient descent to improve data fusion. Through repeated training, the model adapts to environmental changes, adjusting the GNSS and INS fusion ratio. It increases INS weight in weak GNSS signals and favors GNSS in high-quality conditions. After training, ANFIS autonomously optimizes data fusion in real time, enhancing the accuracy and robustness of GNSS/INS navigation systems. In GNSS/INS integration using ANFIS, gradient descent and backpropagation are employed to optimize data fusion under varying environmental conditions. Gradient descent minimizes prediction errors by adjusting fuzzy system parameters, while backpropagation fine-tunes neural network weights and membership functions. This enables real-time adaptation to signal quality and sensor errors, allowing the system to shift reliance between GNSS and INS based on reliability. Unlike traditional methods, this approach effectively handles nonlinearities and signal disruptions, maintaining over 90% robustness and ensuring accurate navigation even in harsh environments.

C. GNSS and INS data fusion strategy

1. Data fusion strategy design

To give full play to the complementary effect of GNSS and INS, this study is based on the adaptive data fusion strategy of ANFIS. When the GNSS signal is stable and accurate, the system increases the weight of the GNSS to provide accurate positioning information. Once the GNSS signal is interfered with or lost, the system automatically switches to INS mode and adjusts the weight of INS to maintain continuous positioning. ANFIS adjusts the weight distribution of GNSS and INS based on data to ensure that high-precision navigation can still be achieved using INS when the GNSS signal fluctuates.

2. Implementation of the fusion process

ANFIS uses the quality of real-time signals, historical parameters, and the status of various environments to adjust the weights of GNSS and INS in data fusion. The GNSS weight is reduced and the INS weight is increased to deal with the influence of GNSS signals being blocked and multipath effects. When the signal is good, the GNSS weight is dynamically enhanced. With the help of fuzzy rules generated by training, ANFIS effectively handles the nonlinear relationship between GNSS and INS, balances the positioning results of the two through fuzzy inference, and avoids single sensor errors. Even when the GNSS signal is missing or disturbed, the system can adjust the INS weight.

Figure 2 shows the dynamic adjustment of weights based on real-time signal quality and environmental factors in the GNSS/INS coupling system. In the upper part of Fig. 2, as the GNSS signal quality improves, the weight of GNSS increases, and the weight of INS decreases. When the GNSS signal quality is 0.9, the GNSS weight is close to 1 and the INS weight drops to around 0. The lower part of Fig. 2 shows the impact of environmental factors on the weight. When environmental factors increase, the GNSS weight decreases and the INS weight increases, ensuring that the INS can provide stable navigation data when the GNSS signal weakens. Figure 2 shows that the ANFIS model automatically adjusts the weight under different conditions, thereby improving the accuracy and robustness of the system.

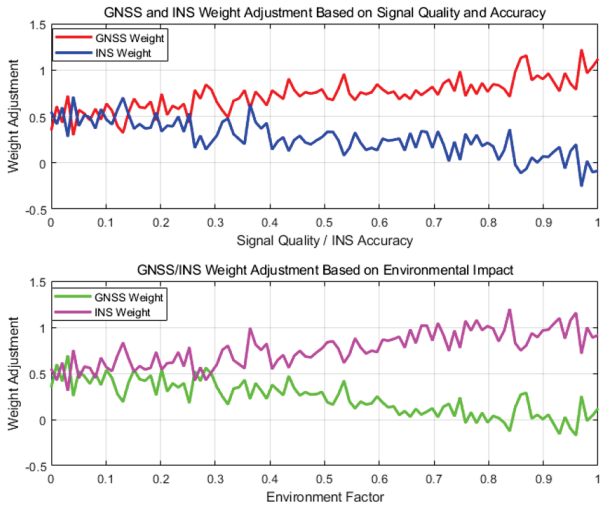


Fig. 2. Dynamic adjustment of weights in the GNSS/INS coupling system.

3. Fusion strategy optimization and application

To make the fusion strategy of GNSS and INS more efficient, this paper applies a self-optimization mechanism based on ANFIS. During the experiment, the system continuously optimizes the parameters of fuzzy rules and neural networks based on real-time data and error feedback. Through continuous training and adjustment, ANFIS learns rules more suitable for the current environment and flexibly adjusts the data fusion strategy according to environmental changes. When the GNSS signal is lost or the quality is reduced, ANFIS increases the weight of INS to ensure a smooth transition of the system. Ehen the GNSS signal returns to normal, the system quickly resumes its dependence on GNSS and reduces the weight of INS.

Various error sources can be identified and targeted for adjustment during ANFIS model training. To address the multipath effect of GNSS, ANFIS adjusts the weights of the input variables to weaken their influence on localization and reinforces the role of INS in case of signal occlusion to prevent accuracy from being impaired. This optimization ensures that GNSS/INS data fusion provides stable accuracy in real-time applications while maintaining system robustness under extreme conditions.

The self-optimization mechanism using ANFIS enhances the performance of GNSS/INS coupled navigation systems in challenging environments. By integrating fuzzy logic with neural networks, ANFIS dynamically adjusts the weight between GNSS and INS data based on real-time signal quality and environmental conditions. It is trained on actual data to recognize error patterns and optimize fusion strategies. When GNSS signals are degraded or lost, ANFIS increases reliance on INS to maintain positioning accuracy. The system also continuously refines its rules using error feedback, ensuring robust and adaptive navigation. Experiments show over 90% robustness and an average processing time of 0.69 seconds, confirming the system's effectiveness in real-world applications.

D. Dynamic adjustment mechanism

1. GNSS signal loss processing

When the GNSS signal is missing or strongly interfered with, the system quickly switches to INS mode to ensure the continuity of navigation and high positioning stability. The ANFIS model automatically increases the weight of INS through its flexible weight adjustment mechanism, thereby ensuring that INS can continue to play a core role when GNSS is unavailable. ANFIS continuously monitors the GNSS signal strength and the number of available satellites, and immediately and adaptively increases the weight of INS once a signal loss is detected. With INS's continuous position estimation capability, this mechanism prevents the deviation of positioning information.

Furthermore, ANFIS optimizes the INS error correction strategy tasks to maintain high positioning accuracy. As the operating time of INS grows, its positioning

accuracy may gradually decrease. For this reason, ANFIS utilizes real-time data to accurately correct the errors of INS to avoid accumulation of errors. ANFIS implements effective error compensation in response to environmental changes by dynamically adjusting the parameters of the accelerometers and gyroscopes, ensuring that the system maintains high accuracy during periods of GNSS signal loss. This mechanism allows the system to quickly adapt to changes in the external environment and reduces the reliance on static weight configurations.

ANFIS enhances robustness in GNSS environments facing signal loss or multipath interference by integrating mechanisms that evaluate signal strength, satellite geometry, and multipath indicators. It adjusts fusion weights between GNSS and INS data dynamically, relying more on INS when GNSS quality declines. By combining fuzzy logic with neural networks, ANFIS optimizes rules and functions through real-world data training, facilitating real-time adaptation to environmental changes. It constructs detailed error models for both GNSS and INS to correctly identify and mitigate inaccuracies in harsh conditions. In practice, ANFIS lowers GNSS dependence during interference while improving INS compensation, swiftly reverting to GNSS when conditions stabilize, achieving over 90% robustness in challenging scenarios like urban canyons and tunnels. The GNSS/INS coupled navigation system utilizes ANFIS to maintain accuracy during GNSS signal loss or interference. It begins with dynamic error modeling addressing atmospheric delays, satellite geometry, multipath effects, and INS sensor drift. Real-time corrections use environmental data. The fusion strategy adapts GNSS weight based on signal conditions, enhancing GNSS during strong signals and relying on INS when GNSS degrades. ANFIS combines fuzzy logic for uncertainty management and neural networks for optimizing fuzzy rules. Continuous monitoring of GNSS parameters assesses signal reliability. In outages, it compensates for INS drift with adaptive calibration, achieving over 90% robustness and low positioning errors with a 0.69-second execution time.

2. Signal quality judgment and dynamic adjustment

ANFIS measures signal reliability by monitoring signal-to-noise ratio, satellite spatial distribution, and multipath effects.

Figure 3 shows that the signal strength fluctuates between -180 dBm and -120 dBm. The number of satellites varies between 5 and 15, with lower numbers of satellites leading to reduced positioning accuracy. The SNR value is in the fluctuating range of 0-50 dB. A higher SNR corresponds to better localization accuracy, while a lower SNR may lead to an increased error. The fluctuation of GNSS signal quality directly affects the accuracy and stability of the positioning system.

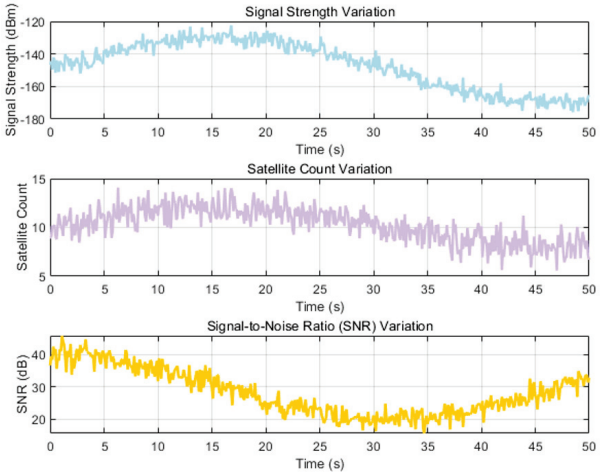


Fig. 3. Signal strength, number of satellites, and signal-to-noise ratio over time.

ANFIS uses fuzzy logic to evaluate GNSS signal quality, uses input variables to determine signal reliability, and dynamically adjusts data fusion strategies based on the evaluation results. If the signal quality is low, ANFIS relies on INS to provide continuous positioning information and gradually reduce its reliance on GNSS data. When the signal returns to normal, ANFIS increases its reliance on GNSS data and further optimizes the positioning results. This mechanism ensures that the system is adaptively adjusted according to actual circumstances under various environmental conditions to maximize the accuracy and stability of the navigation system.

ANFIS enhances GNSS/INS navigation by adaptively adjusting fuzzy logic rules, neural network parameters, and data fusion strategies in real time. It uses environmental inputs like signal strength and sensor errors to determine optimal GNSS and INS weightings, increasing reliance on INS during GNSS signal loss. A neural network continuously learns and optimizes the fuzzy system for improved adaptability. This dynamic approach ensures accurate, stable navigation by compensating for errors such as multipath effects and sensor drift. ANFIS achieves over 90% robustness in harsh conditions, while maintaining fast processing, making it suitable for real-time navigation in complex environments.

IV. EXPERIMENT AND PERFORMANCE EVALUATION

A. Performance evaluation method

In the performance evaluation, this paper simulates the lateral and longitudinal errors of complex environments such as urban high-rise buildings and tunnels under different geographical conditions by using actual GNSS and INS sensor data sets and the stability and robustness of experimental scenes under various harsh

conditions. The performance of the ANFIS optimization algorithm under these harsh conditions is tested through these experimental data. To comprehensively evaluate the performance of the ANFIS algorithm, this study conducts a comparative analysis with traditional EKF and UKF.

B. Positioning error

This paper evaluates the algorithm's accuracy by calculating the positioning error (lateral error and longitudinal error).

Figure 4 shows the error performance of ANFIS, EKF, and UKF in various geographical environments. ANFIS has the best performance in open areas, with lateral and longitudinal errors of 0.66 meters and 0.65 meters, respectively. In complex environments, such as urban canyons and tunnels, ANFIS error is still low, while the errors of EKF and UKF increase, especially in tunnels where the lateral error of EKF exceeds 0.9 meters. Overall, ANFIS shows high accuracy and stability, highlighting its advantages in complex environments.

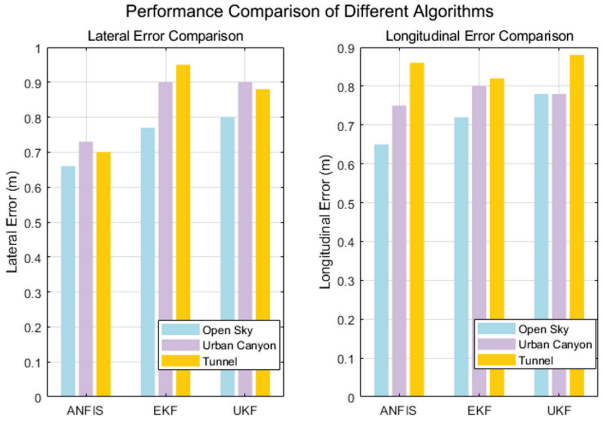


Fig. 4. Lateral and longitudinal errors of the three algorithms ANFIS, EKF, and UKF in different geographical environments.

C. Robustness in different harsh scenarios

To comprehensively evaluate the performance of the GNSS/INS coupled navigation algorithm of ANFIS in harsh environments, this paper designs a variety of experimental scenarios, covering signal loss, multipath interference, and high dynamic environments. Table 2 shows the system's positioning error, stability, and robustness indicators in these scenarios. The robustness and stability of the algorithm in complex environments are further confirmed through comparative verification.

Under normal GNSS signals, the positioning error is 0.5 meters, the system stability is good, and the robustness is as high as 95%. However, when the GNSS signal is lost, the positioning error increases to 1.2 meters, and the system stability and robustness decrease slightly. In the face of multipath interference, the positioning error is 0.8 meters and the robustness is 90.6%. Overall, the robustness in each environment is higher than 90%.

Table 2: Positioning error, stability, and robustness of the system under various harsh conditions

Environmental Condition	Positioning Error (m)	System Stability (Maximum Deviation, m)	Robustness (Stability Score)
Normal GNSS Signal	0.5	0.2	95%
GNSS Signal Loss	1.2	0.5	90.5%
Multipath Interference	0.8	0.3	90.6%
High-Dynamics Environment	1.5	0.6	91%
Strong Electromagnetic Interference	1	0.4	94%
Short-Term Signal Interference	0.7	0.25	92%

D. Algorithm running time

By evaluating the algorithm running time, its computational burden is compared with the performance difference of EKF and UKF.

Figure 5 compares the execution time of ANFIS, EKF, and UKF in different environments. In the GNSS

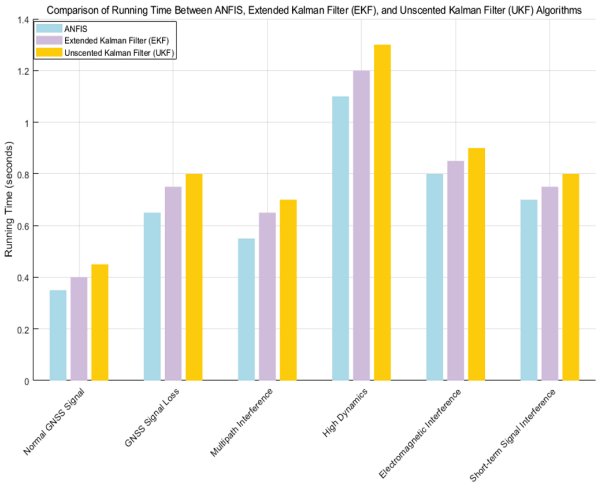


Fig. 5. Running time of the three algorithms in different environments.

signal loss scenario, ANFIS outperforms EKF's 0.75 seconds and UKF's 0.80 seconds with a running time of 0.65 seconds, demonstrating its fast-processing capability. In the face of multipath interference, ANFIS's 0.55-second running time takes the lead again, demonstrating excellent robustness. In a highly dynamic environment, UKF takes the longest time (1.30 seconds), while ANFIS (1.10 seconds) and EKF (1.20 seconds) are shorter, highlighting that UKF is computationally intensive. Overall, ANFIS has good adaptability under complex conditions, with an average running time of 0.69 seconds.

V. CONCLUSION

This study presents a novel ANFIS-based GNSS/INS coupled navigation algorithm that substantially improves positioning accuracy and system robustness in challenging environments characterized by GNSS signal loss, multipath interference, and dynamic conditions. By integrating a detailed error source model with a real-time adaptive fusion strategy, the system effectively adjusts the weighting between GNSS and INS inputs, ensuring continuous and reliable navigation even when satellite signals are degraded or unavailable. The self-optimizing inference engine further enhances performance by dynamically evolving fuzzy rules based on sensor feedback and environmental changes. Experimental results demonstrate that the proposed approach consistently achieves over 90% robustness with significantly reduced positioning errors compared to traditional EKF and UKF, while maintaining low computational latency suitable for real-time applications. These advancements highlight the ANFIS-based fusion method as a superior alternative, offering enhanced stability, adaptability, and accuracy in GNSS-challenged scenarios.

DECLARATIONS

Funding: Research on Key Technologies of Transmission line condition Monitoring Based on Multi-Source Data Fusion based on Beidou high-precision Positioning Technology supported by China Southern Power Grid Co. Ltd. (GZKJXM20222325)

Conflict of interests: Authors do not have any conflicts.

Data Availability Statement: No datasets were generated or analyzed during the current study.

Code availability: Not applicable.

Authors' Contributions: Chen Zerui, Xiao Yanhong, and Wu Xin are responsible for designing the framework, analyzing performance, validating the results, and writing the article. Hu Houpeng, Xiao Jian, Yang Shang, and Gao Zhenghao are responsible for collecting the information required for the framework, providing software, conducting critical reviews, and administering the process.

REFERENCES

[1] E. S. Abdolkarimi and M. R. Mosavi, "A modified neuro-fuzzy system for accuracy improvement of low-cost MEMS-based INS/GPS navigation system," *Wirel. Pers. Commun*., vol. 129, no. 2, pp. 1369-1392, 2023.

[2] L. Cong, S. Yue, H. Qin, B. Li, and J. Yao, "Implementation of a MEMS-based GNSS/INS integrated scheme using supported vector machine for land vehicle navigation," *IEEE Sensors J*., vol. 20, no. 23, pp. 14423-14435, 2020.

[3] Y. Xiao, H. Luo, F. Zhao, F. Wu, X. Gao, Q. Wang, and L. Cui, "Residual attention network-based confidence estimation algorithm for non-holonomic constraint in GNSS/INS integrated navigation system," *IEEE Trans. Veh. Technol*., vol. 70, no. 11, pp. 11404-11418, 2021.

[4] R. Sun, G. Wang, Z. Fan, T. Xu, and W. Y. Ochieng, "An integrated urban positioning algorithm using matching, particle swarm optimized adaptive neuro fuzzy inference system and a spatial city model," *IEEE Trans*. *Veh*. *Technol*., vol. 69, no. 5, pp. 4842-4854, 2020.

[5] E. S. Abdolkarimi and M. R. Mosavi, "Impact assessment of efficient denoising techniques in AI-based low-cost INS/GPS integration during blockage of GPS satellites," *Arab J. Sci. Eng.*, vol. 47, no. 11, pp. 14583-14600, 2022.

[6] W. Jiang, Y. Yu, K. Zong, B. Cai, C. Rizos, J. Wang, S. Shangguan, and W. Shangguan, "A seamless train positioning system using a LiDAR-aided hybrid integration methodology," *IEEE Trans. Veh. Technol*., vol. 70, no. 7, pp. 6371-6384, 2021.

[7] Y. Yuan, Y. Wang, W. Gao, and F. Shen, "Vehicular relative positioning with measurement outliers and GNSS outages," *IEEE Sensors J*., vol. 23, no. 8, pp. 8556-8567, 2023.

[8] Y. Li, R. Chen, X. Niu, Y. Zhuang, Z. Gao, X. Hu, and N. El-Sheimy, "Inertial sensing meets machine learning: Opportunity or challenge?" *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 9995-10011, 2021.

[9] B. Zhu, X. Tao, J. Zhao, M. Ke, H. Wang, and W. Deng, "An integrated GNSS/UWB/DR/VMM positioning strategy for intelligent vehicles," *IEEE Trans. Veh. Technol*., vol. 69, no. 10, pp. 10842-10853, 2020.

[10] M. N. Cahyadi, T. Asfihani, H. F. Suhandri, and R. Erfianti, "Unscented Kalman filter for a low-cost GNSS/IMU-based mobile mapping application

under demanding conditions," *Geodesy Geodyn.,* vol. 15, no. 2, pp. 166-176, 2024.

[11] C. Zhang, X. Zhao, C. Pang, Y. Wang, L. Zhang, and B. Feng, "Improved fault detection method based on robust estimation and sliding window test for INS/GNSS integration," *J. Navig*., vol. 73, no. 4, pp. 776-796, 2020.

[12] A. Siemuri, K. Selvan, H. Kuusniemi, P. Valisuo, and M. S. Elmusrati, "A systematic review of machine learning techniques for GNSS use cases," *IEEE Trans. Aerosp. Electron. Syst*., vol. 58, no. 6, pp. 5043-5077, 2022.

[13] J. Wu, "An innovative neural-fuzzy adaptive Kalman filter for ultra-tightly-coupled GPS/INS integrated system," in *AOPC 2017: Space Optics and Earth Imaging and Space Navigation*, vol. 10463, pp. 75-80, Oct. 2017.

[14] X. Lai, S. Tong, and G. Zhu, "Adaptive fuzzy neural network-aided progressive Gaussian approximate filter for GPS/INS integration navigation," *Measurement*, vol. 200, p. 111641, 2022.

[15] S. Yue, L. Cong, H. Qin, B. Li, and J. Yao, "A robust fusion methodology for MEMS-based land vehicle navigation in GNSS-challenged environments," *IEEE Access*, vol. 8, pp. 44087-44099, 2020.

[16] G. Wang, Y. Han, J. Chen, S. Wang, Z. Zhang, N. Du, and Y. Zheng, "A GNSS/INS integrated navigation algorithm based on Kalman filter," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 232-237, 2018.

[17] A. E. Mahdi, A. Azouz, A. Noureldin, and A. Abosekeen, "A novel machine learning-based ANFIS calibrated RISS/GNSS integration for improved navigation in urban environments," *Sensors*, vol. 24, no. 6, p. 1985, 2024.

[18] D. Wang, Y. Dong, Z. Li, Q. Li, and J. Wu, "Constrained MEMS-based GNSS/INS tightly-coupled system with robust Kalman filter for accurate land vehicular navigation," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 5138-5148, 2019.

[19] B. R. Gudivaka and M. Thanjaivadivel, "IoT-driven signal processing for enhanced robotic navigation systems," *Int. J. Eng. Technol. Res. Manag*., vol. 4, no. 5, 2020.

[20] R. Sun, D. Xue, and Y. Zhang, "Rear-end collision detection based on GNSS/compass fusion and adaptive neuro fuzzy inference system," in *2017 Forum on Cooperative Positioning and Service (CPGPS)*, pp. 273-277, May 2017.

[21] H. Zhao and Z. Yang, "A novel fault detection and exclusion method for applying low-cost INS/GNSS integrated navigation system in urban environments," *IEEE Trans. Intell. Transp. Syst*., vol. 26, pp. 143-156, 2024.

[22] D. J. Jwo, A. Biswal, and I. A. Mir, "Artificial neural networks for navigation systems: A review of recent research," *Appl. Sci*., vol. 13, no. 7, p. 4475, 2023.

[23] F. Wu, H. Luo, F. Zhao, L. Wei, and B. Zhou, "Optimizing GNSS/INS integrated navigation: A deep learning approach for error compensation," *IEEE Signal Process. Lett*., vol. 31, pp. 3104-3108, 2024.

[24] Y. Zhang and L. Wang, "A hybrid intelligent algorithm DGP-MLP for GNSS/INS integration during GNSS outages," *J. Navig*., vol. 72, no. 2, pp. 375-388, 2019.

[25] C. Guo and W. Tu, "A novel self-learning GNSS/INS integrated navigation method," in *Proc. 34th Int. Tech. Meeting Satellite Div. Inst. Navig*. (*ION GNSS+*), pp. 168-179, Sep. 2021.

Chen Zerui received his M.E. degree in Computer Science and Technology from Guizhou University, China, in 2018. He currently works at the Power Science Research Institute of Guizhou Power Grid Co. Ltd. His research interests include algorithmic game theory, blockchain, information security, and AI security.

Xiao Jian , born in 1993, is currently an Engineer at the Electric Power Research Institute, Guizhou Power Grid Company Ltd., Guiyang, China. He received his Ph.D. in Information and Communication Engineering from Harbin Engineering University. His research interests include power artificial intelligence and power big data analytics.

Yang Shang was born in 1997 in Tianzhu, Guizhou, China. He received his master's degree from North China Electric Power University and is currently working at China Southern Power Grid Company Limited. His research interests include communication technology, artificial intelligence, and big data analysis.

Hu Houpeng was born in 1992 and is a postgraduate student. His main research direction is electric vehicle charging and replacement technology.

Gao Zhenghao , male, born August 1979, is a senior engineer. He graduated in 2002 from the Department of Computer Science at Guizhou University, China, majoring in Computer Applications. Since July 2002, he has been working at the IoT Application Technology Research Center of the Electric Power Research Institute, focusing on information technology, power monitoring system development, network security, and communication technology research.

Wu Xin , born October 1976, is an engineer at the Electric Power Science and Technology Research Institute of Guizhou Power Grid Co. Ltd. He holds a bachelor's degree and is of Han nationality, originally from Jinhua, Zhejiang. His research interests include information technology and big data applications.

Xiao Yanhong , born April 1976, is a senior engineer at the Electric Power Science Research Institute of Guizhou Power Grid Co. Ltd. She is of Han nationality and originally from Fuquan, Guizhou, China. Her research interests include network security testing and online monitoring technology for power grids.

Effect of Corrosion on Electromagnetic Shielding Effectiveness of Enclosures with Gasketed Seams

Xin He[1], Gang Zhang[2], Yin Shi[2], Lixin Wang[1], Zhongliang Du[2], and Zhitian Wang[2]

[1]Department of Robotics and Advanced Manufacture
Harbin Institute of Technology, Shenzhen 518000, China
21B953013@stu.hit.edu.cn, wlx@hit.edu.cn

[2]Department of Electrical Engineering and Automation
Harbin Institute of Technology, Harbin 150000, China
Zhang_hit@hit.edu.cn, 1425978482@qq.com, 23S106100@stu.hit.edu.cn, wzt20120811@gmail.com

***Abstract* –** Gasketed seams are a primary pathway for electromagnetic leakage in enclosures, and corrosion at these seams can significantly worsen the leakage. Therefore, this work investigates the degradation of shielding effectiveness (SE) in enclosures with gasketed seams when exposed to corrosive environments. Since corrosion products typically exhibit low electrical conductivity, a simulation analysis is first conducted to examine how variations in seam conductivity affect the SE of the enclosure. Subsequently, experiments are conducted on the enclosure with gasketed seams to evaluate changes in its SE after being exposed to a corrosive environment. Experimental results show that enclosures with gasketed seams exhibit a significant decline in SE when exposed to corrosive environments. For improving the SE of enclosures and mitigating the adverse effects of corrosion, installing electrically conductive rubber gaskets is a more effective solution than using finger stock gaskets.

***Index Terms* –** Corrosion, electromagnetic interference gasket, enclosure, shielding effectiveness.

I. INTRODUCTION

Installing shielded enclosures is a common and effective solution for preventing electromagnetic leakage and interference in electronic devices. Ideally, an enclosure should consist of a continuous metal casing. However, seams are inevitable during manufacturing, which can lead to electromagnetic leakage and a reduction in shielding effectiveness (SE) [1]. Installing electromagnetic interference (EMI) gaskets or applying techniques such as riveting and welding can improve the electrical continuity of the enclosure, thereby enhancing its SE [2]. The enclosure is vulnerable to electrochemical corrosion at the seams when exposed to corrosive environments [3]. This degradation can compromise the electrical continuity at the seams, thus diminishing the enclosure's immunity to conducted and radiated disturbances. Therefore, it is essential to investigate the impact of corrosion on the SE of enclosures with gasketed seams.

The EMI gaskets typically offer good conductivity and sealing capability, and they are installed in the seams to minimize electromagnetic leakage. Various EMI gaskets are available to meet different practical requirements, among which conductive rubber and finger stock gaskets are the most commonly used. Conductive rubber and finger stock gaskets are widely used in the electromagnetic compatibility (EMC) design of enclosures. In [4], it is demonstrated that conductive silicone rubber gaskets filled with silver-plated glass particles exhibit high stability under severe vibration and intense simulated electromagnetic pulse (EMP) conditions, recommending their application in critical military applications. In [5] and [6], a combination of modified Bethe's small aperture theory and filtering theory is employed to investigate the impact of corrosion on the performance of finger stock gaskets. Although these studies examined the characteristics of two different types of EMI gaskets and their effectiveness in enhancing the SE of enclosures, they did not compare the performance of the two gasket types under identical operating conditions, such as a corrosive environment.

Enclosures are highly susceptible to electrochemical corrosion when exposed to atmospheric, marine, and similar environments for extended periods. Galvanized steel is commonly used as a material for shielding enclosures. In marine atmospheric environments, factors such as sodium chloride (NaCl) solution concentration and exposure duration significantly influence the corrosion behavior of galvanized steel [7]. When exposed to electrolyte solutions, enclosures undergo galvanic and electrolytic corrosion. These corrosion processes occur at the joint interfaces of the enclosure, thereby compromising

Submitted On: May 18, 2025
Accepted On: July 29, 2025

https://doi.org/10.13052/2025.ACES.J.400809

electrical continuity and reducing SE [8]. Open circuit potential and electrochemical impedance measurements can be used to predict the corrosion characteristics of EMI gaskets [9].

To evaluate the electromagnetic shielding performance of the enclosure, the most straightforward method is to measure the SE. The SE testing methods are outlined in [10], which are commonly used for enclosures and shielding materials [11–14]. However, measuring SE is often complex and costly. Study [15] demonstrates a one-to-one correspondence between the SE of the enclosure and the transfer impedance of the gasket. Based on this conclusion, related studies investigate the SE of enclosures by measuring the transfer impedance of gaskets [16–19]. Although measuring the gasket's transfer impedance simplifies the testing process, it does not enable accurate determination of the enclosure's SE and fails to directly reflect its electromagnetic shielding capability.

In summary, understanding the degradation mechanism of enclosures in corrosive environments and enhancing the reliability of enclosure design are key topics in EMC and environmental reliability. However, research relevant to this aspect is insufficient. This work primarily focuses on a comparative analysis of three types of seam structures: filling with conductive rubber gaskets, finger stock gaskets, and metal-metal contact (without EMI gaskets). This work examines the effects of corrosion on seam structures and their impact on SE through both simulation and experimental studies. Continuous corrosion tests are conducted on various seam configurations, and the SE of enclosures is measured and compared across different exposure durations. This work also examines the impact of various gasket materials on SE under corrosive conditions, offering valuable guidance for engineers in selecting gaskets suitable for such environments.

The rest of this work is organized as follows. Section II analyzes the mechanisms by which corrosion affects the SE of enclosures. Section III presents simulation studies that examine the impact of seam conductivity on the SE of enclosures. In section IV, experimental measurements are carried out to determine the SE variations in enclosures with three types of gasketed seams after exposure to corrosive environments. Finally, section V concludes this work.

II. THEORETICAL BACKGROUND

A. Corrosion mechanisms of seam structures

As categorized in [1], corrosion can be classified into eight types based on their characteristics and underlying mechanisms: uniform corrosion, galvanic corrosion, crevice corrosion, pitting corrosion, intergranular corrosion, selective corrosion, erosion corrosion, and stress corrosion. For the gasketed seam investigated in this work, the presence of joint seams promotes the accumulation of corrosive solutions, making it highly vulnerable to crevice corrosion. Additionally, if the gasket and enclosure are made of different materials, galvanic corrosion is likely to occur once a conductive solution enters the seam. Among various corrosion types, crevice corrosion and galvanic corrosion are the most prominent mechanisms affecting gasketed seams. Therefore, this study mainly focuses on these two types of corrosion.

1. Crevice corrosion

Enclosures assembled using riveting, welding, or bolted connections typically contain seams at their joints. Corrosive media can accumulate in these seams and react with the metal surfaces, resulting in crevice corrosion. This type of localized corrosion can occur in all metals and alloys, particularly on surfaces of metals that tend to form passive films. Metals susceptible to crevice corrosion develop irregular pits of varying depths within crevices covered by corrosion products, leading to the formation of enclosed galvanic cells. The self-catalyzing nature of these cells is the primary factor accelerating the progression of crevice corrosion.

To mitigate the risks of crevice corrosion, enclosure materials should preferably be selected from metals and alloys with a reduced tendency for self-passivation, such as stainless steels with higher molybdenum (Mo), chromium (Cr), or nickel (Ni) content. Additionally, the design should minimize the presence of crevices and dead-end areas. When seamless construction is impractical, filling crevices with solid materials can help maintain electrical continuity across the enclosure surface.

2. Galvanic corrosion

Seam structures filled with EMI gaskets are often prone to galvanic corrosion. This type of corrosion occurs when dissimilar metals with different electrochemical potentials come into contact in the presence of an electrolyte. The metal with the lower corrosion potential serves as the anode, losing electrons through oxidation reactions and undergoing accelerated local corrosion. In contrast, the metal with the higher corrosion potential functions as the cathode, gaining electrons through reduction reactions and thus being protected from corrosion. Although galvanic corrosion improves the corrosion resistance of the cathodic material, it also speeds up the corrosion of the anodic material. As a result, the overall corrosion rate of the joint structure rises, reducing its corrosion resistance.

Installing finger stock gaskets in enclosures raises the risk of galvanic corrosion at contact joints between dissimilar metals in the seams, potentially accelerating metal degradation. In conductive rubber gaskets,

most metal particles are encapsulated within the rubber matrix, thereby limiting their direct contact with the enclosure. Consequently, when exposed to a corrosive environment, galvanic corrosion is confined to the few metal particles present on the gasket surface that are in contact with the enclosure. Therefore, compared to finger stock gaskets, conductive rubber gaskets pose a lower risk of galvanic corrosion at the gasketed seams.

EMI gaskets are typically made from conductive materials with relatively high electrochemical potentials, while enclosures are often constructed from highly malleable metals with lower potentials. This potential difference drives galvanic corrosion at the interface of contact. To mitigate galvanic corrosion in enclosures with gasketed seams, it is important to minimize the electrochemical potential difference between contacting materials.

B. Shielding effectiveness

SE is primarily used to describe the electromagnetic shielding effect of a shield. For electric fields, SE is expressed as:

$$\mathrm{SE_e} = 20\log_{10}\left|E_0/E_\mathrm{s}\right|, \tag{1}$$

where E_0 and E_s are the electric field strengths at the same position before and after shielding. For magnetic fields, SE is expressed as:

$$\mathrm{SE_m} = 20\log_{10}\left|H_0/H_\mathrm{s}\right|, \tag{2}$$

where H_0 and H_s are magnetic field strengths at the same position before and after shielding. In the far-field region, $\mathrm{SE_e} = \mathrm{SE_m}$ holds.

A shield with seams can be considered a combined wall comprising seams and metal panel regions. According to the research in [21], the overall SE of the shield can be expressed as:

$$\mathrm{SE_e} = -20\log_{10}\left[\sum_{i=1}^{N}\left(10^{-\frac{SE_i}{20}}\frac{w_i}{W}\right)\right]. \tag{3}$$

in which SE_i represents the SE of the ith region, w_i is the width of the ith region, W signifies the total width of the N regions, meaning the width of the shield. Equation (3) describes the relationship between the overall SE of the shield and the SE contributions from its sub-regions. According to (3), regions that provide sufficiently high SE contribute minimally to the overall SE because they show little electromagnetic leakage. In contrast, regions with low SE dominate the overall shielding performance, since they are the primary sources of electromagnetic leakage. For a shield that includes structures like gaskets and apertures, it can be divided into regular regions and weaker regions containing these features. Because of the lower SE in these weak regions, they mainly determine the overall SE of the shield. This shows that improving the SE of weak regions, such as gasketed seams and apertures, can significantly enhance the shield's overall performance.

When incident EM waves encounter a shield, they undergo reflection, transmission, and absorption processes, as illustrated in Fig. 1. The SE of the shield is affected by these processes. The SE in decibels of the shield can be expressed as [22]:

$$\mathrm{SE} = R + A + M, \tag{4}$$

where R, A, and M represent reflection loss, absorption loss, and multiple reflection factor.

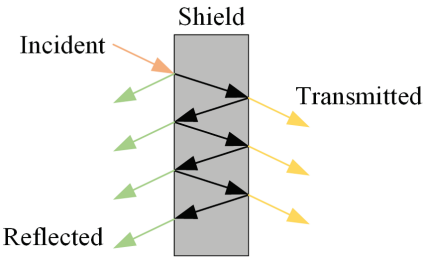


Fig. 1. Propagation behavior of EM waves upon encountering a shield.

Assuming the incident EM waves originate sufficiently far from the shield, they can be approximated as uniform plane waves. For a shield with a thickness of t, the following equations hold [23]:

$$\begin{aligned} R &= 20\log_{10}\left|\tfrac{(Z_0+Z_\mathrm{S})^2}{4Z_0Z_\mathrm{S}}\right|, \\ A &= 20\log_{10}\left(e^{t/\delta}\right), \\ M &= 20\log_{10}\left|1-\left(\tfrac{Z_0-Z_\mathrm{S}}{Z_0+Z_\mathrm{S}}\right)^2 e^{-2t(1+j)/\delta}\right|. \end{aligned} \tag{5}$$

$Z_0 = 120\pi\ \Omega$ is the wave impedance in vacuum, δ is the skin depth, and Z_S is the intrinsic impedance of shield. Z_S and δ can be obtained through:

$$\begin{aligned} Z_\mathrm{S} &= \sqrt{\tfrac{j\omega\mu_0\mu_r}{\sigma+j\omega\varepsilon_0\varepsilon_r}}, \\ \delta &= \sqrt{\tfrac{1}{\pi f\mu_0\mu_r\sigma}}, \end{aligned} \tag{6}$$

in which f is the frequency of the EM waves, $\omega = 2\pi f$ denotes the angular frequency, σ is the conductivity of the shielding material, $\mu_0 \approx 4\pi\times10^{-7}$ H/m represents the permeability of free space, $\varepsilon_0 \approx 8.854\times10^{-12}$ F/m is the permittivity of free space, ε_r and μ_r are the relative permittivity and relative permeability of the shielding material.

Since the value of ε_0 is very small and the shield typically has high electrical conductivity, it follows that $\sigma \gg \omega\varepsilon_0\varepsilon_r$. Then, Z_S can be written as:

$$Z_\mathrm{S} = \sqrt{\frac{j\omega\mu_0\mu_r}{\sigma}}. \tag{7}$$

In (7), the value of μ_0 is relatively small. Thus, for a shield with high electrical conductivity, $Z_\mathrm{S} \ll Z_0$ holds. Finally, the SE of the shield can be approximately expressed as:

$$\mathrm{SE} = 20\log_{10}\left|\frac{Z_0}{4Z_\mathrm{S}}\right| + 20\log\left|e^{t/\delta} - e^{-t(1+2j)/\delta}\right|, \tag{8}$$

and it can be deduced from (3) to (8) that the SE of the shield is positively correlated with σ.

III. SIMULATION AND RESULTS

As mentioned in section II, the seam structures of a shield are prone to severe corrosion when exposed to a corrosive environment. This corrosion degrades the electrical continuity at the seams. Specifically, the corrosion products (mainly oxides and metal salts with low conductivity) accumulate on the contact surfaces, reducing the conductivity of the seam area. For a seam structure, its SE is positively correlated with its conductivity, as can be derived from (8). Therefore, once corrosion occurs, the SE of the seam structure decreases. In contrast, the metal panel regions outside the seams are typically thick enough that corrosion affects only the surface layer. As a result, these panels maintain good electrical continuity and can still provide high SE. Consequently, the seams represent the weakest point of the shield and the overall SE of a shield with seams is heavily depends on the SE of the seam structures, as indicated by (3).

In summary, the conductivity of the seam structure has a significant impact on the overall SE of the shield. To demonstrate this supposition, this section investigates the effect of seam conductivity on the shield's overall SE through simulation.

A. Simulation setup

A 3D simulation model of the shielding enclosure is constructed in Ansys HFSS, based on the physical structure shown in Fig. 2 (a) and clarified in Fig. 2 (b). The enclosure is made of iron and has a cubic shape with a side length of 310 mm and a wall thickness of 5.6 mm. The test plate is a square iron plate measuring 380 mm per side. A square test window (200 mm per side) is cut into the test plate and an iron cover plate with a side length of 260 mm is mounted over the window. The gasket is installed at the joint of the cover plate and the test plate, with a total length of 480 mm.

As described in [21], the actual gasketed seam can be modeled as a uniform and isotropic structure. By adjusting the conductivity of the model, the gasketed seam can be made equivalent to the actual one. Therefore, in this section, the complex actual gasketed seam is simplified into a uniform and isotropic model. As mentioned earlier, corrosion degrades the electrical continuity of the actual gasketed seam. In the simulation, this effect is represented by reducing the conductivity of the gasket model. However, it should be noted that, since the impact of corrosion on gasket conductivity cannot be accurately quantified, the simulations in this work can only provide a qualitative analysis of its effect on SE.

In the simulation model, the seam width between the two contacting cover plates is set to 2 mm. The initial conductivity of the seam material is set to 1000 S/m and is gradually reduced to 100 S/m, 10 S/m, and 1 S/m to simulate the progressive material degradation caused by corrosion. A tetrahedral mesh with adaptive refinement is used to accurately capture critical geometric and electromagnetic features. A plane wave with an amplitude of 1 V/m is incident normally on the test panel. An electric field probe is positioned at the geometric center of the enclosure to measure the internal electric field strength. Based on these measurements, the SE of the enclosure is calculated. In the HFSS simulation, the boundary condition is set to 'Radiation,' and adaptive meshing is employed. The solver's maximum delta energy is specified as 0.01 to ensure convergence. A discrete frequency sweep is conducted over the range of 0.5-1.5 GHz with a step size of 0.01 GHz. The simulation results are shown in Fig. 3.

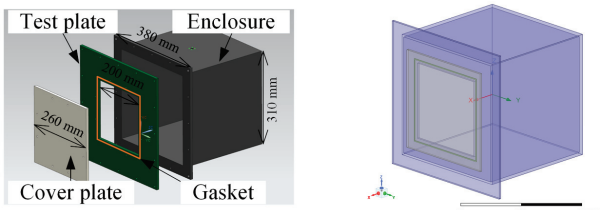


Fig. 2. Enclosure model: (a) physical structure of the enclosure and (b) corresponding simulation model.

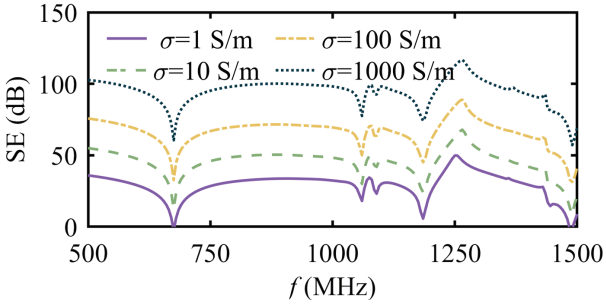


Fig. 3. SE simulation results for different seam conductivities.

B. Analysis of simulation results

As shown in Fig. 3, the conductivity of the seam material significantly impacts the SE of the enclosure. In actual corrosive environments, both the conductivity and the geometric characteristics of the contact surfaces undergo considerable changes due to corrosion. These changes substantially degrade the electrical continuity at the seams, resulting in increased electromagnetic leakage and a corresponding reduction in SE.

Conductivity of the gasket model is set to 1 S/m, and distributions of the electric field, magnetic field, and

surface current of the enclosure at 1 GHz are illustrated in Figs. 4, 5, and 6.

As shown in Figs. 4 and 5, the enclosure provides effective shielding against both electric and magnetic fields. Aside from the edges of the cover, the electric field and magnetic field intensity peaks mainly at the seam regions, indicating that these seams serve as the primary pathways for electromagnetic leakage into the interior of the enclosure. The surface current distribution on the enclosure is illustrated in Fig. 6. It can be seen that the surface directly exposed to the incident wave shows higher surface current density. As these surface currents cross the seam with decreased electrical continuity, a voltage drop is created at the seam. This voltage drop acts as a secondary electromagnetic source, re-emitting EM

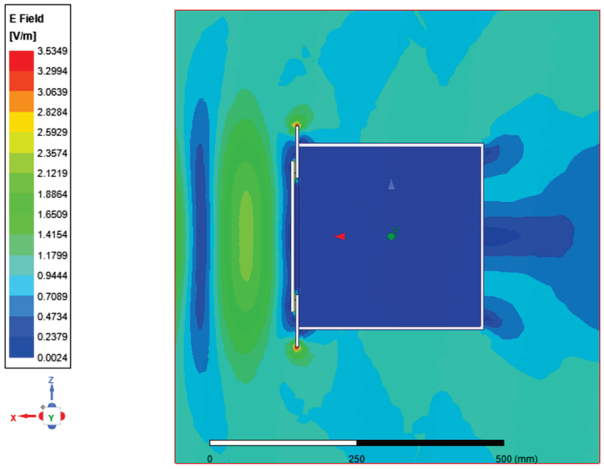


Fig. 4. Electric field distribution in the XZ-plane within the solution region.

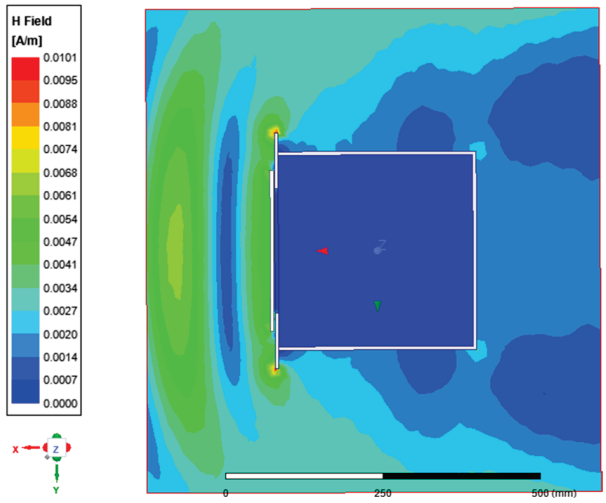


Fig. 5. Magnetic field distribution in the XY-plane within the solution region.

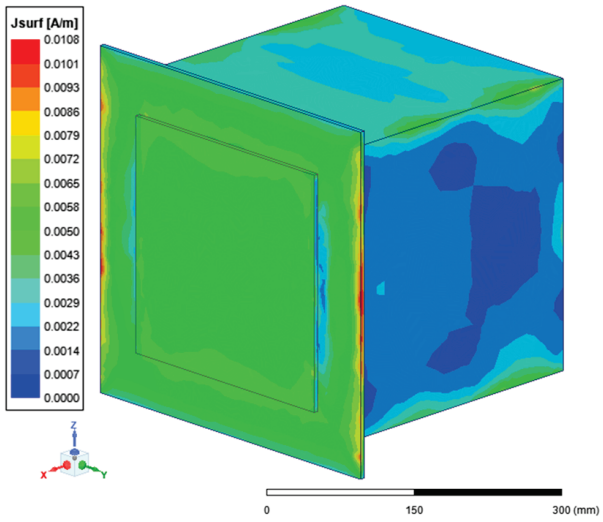


Fig. 6. Surface current density of the enclosure.

waves into the interior of the enclosure. Consequently, the SE of the enclosure decreases.

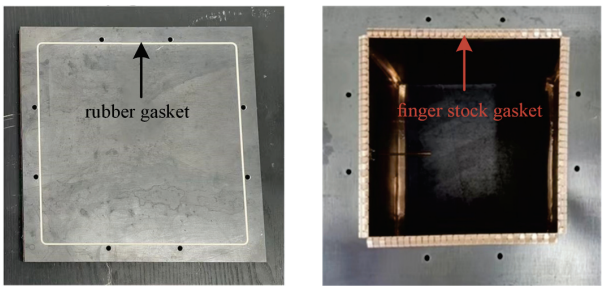
Simulation and analysis confirm that the seams of the enclosure are a primary factor contributing to degradation of SE. Corrosion-induced degradation at the seams reduces their electrical continuity which, in turn, increases their susceptibility to electromagnetic leakage. Consequently, the seams become more effective pathways for electromagnetic waves to penetrate the enclosure, resulting in a substantial decline in overall shielding performance.

IV. EXPERIMENTAL AND RESULTS

A. Corrosion degradation experiment

To conduct the experiments, a steel shielding enclosure, with the structure and dimensions shown in Fig. 2 (a), is manufactured according to SAE ARP1173 standard [20]. The test seam refers to the interface formed by the connection of the two square cover plates. In the experiments, the test seam configurations are classified into three types: (1) installation of a conductive rubber gasket (Fig. 7 (a)), (2) installation of a finger stock gasket (Fig. 7 (b)) and (3) metal-to-metal overlap without any gasket.

The gasket and two cover plates are assembled into the experimental sample and fastened using bolts. A torque wrench is utilized to ensure a uniform tightening force across all bolts. Since the experimental sample will later be assembled with the enclosure for SE testing, the contact area between the larger cover and the enclosure is covered with Teflon tape to prevent corrosion. The experimental samples are then placed in an acid salt spray test chamber for corrosion degradation testing, as shown in Fig. 8. The experimental setup follows the acetic acid salt



(a) (b)

Fig. 7. Plates with EMI gaskets installed: (a) conductive rubber gasket and (b) finger stock gasket.

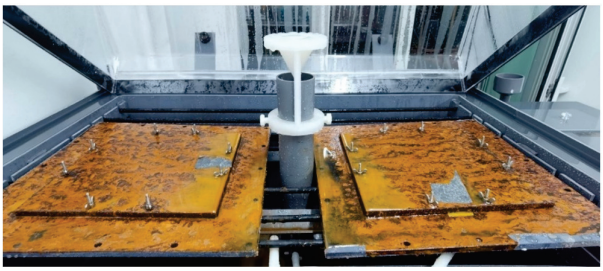
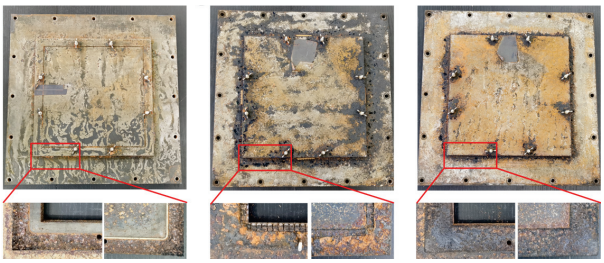


Fig. 8. Samples undergoing corrosion testing in the acetic acid salt spray chamber.

spray test specified in the ISO 9227 standard. The test is conducted at a temperature of 35°C using a 5% (by mass) NaCl solution. The pH of the solution is adjusted to 3.0 by adding glacial acetic acid. Each test cycle lasted 16 hours. After each cycle, the samples are removed, rinsed, dried, and measured for SE before being returned to the test chamber for the next cycle. In this work, each sample underwent five experimental cycles, resulting in a total test duration of 80 hours.

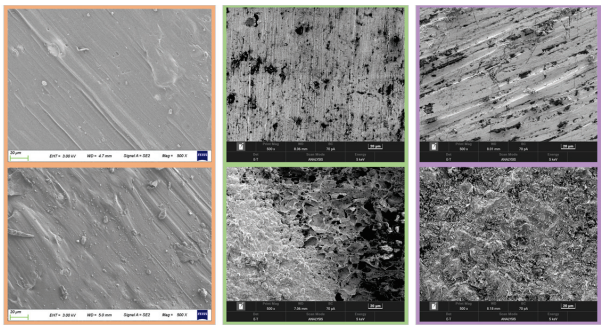
After the corrosion test, the appearance of the experimental samples is shown in Fig. 9. Severe corrosion is observed at the seams, accompanied by the accumulation of corrosion products. Based on visual inspection among the three samples, the one with the finger stock gasket exhibited the most severe corrosion, while the one with the conductive rubber gasket showed the least corrosion. This result may be attributed to the conductive rubber gasket providing better sealing, which reduces the accumulation of corrosive solution in the gasketed seam and thereby delays corrosion in that area. In contrast, the installation of a finger stock gasket introduces numerous micro gaps within the seam, allowing corrosive liquid to accumulate and accelerate corrosion in the region.

The surfaces of each gasket material before and after corrosion are examined with a scanning electron microscope (SEM) to reveal the microstructural changes at contact surfaces. The SEM observation results are shown in Fig. 10, captured at a magnification of 500×. As



(a) (b) (c)

Fig. 9. Experimental samples after corrosion degradation testing: (a) conductive rubber gasket, (b) finger stock gasket, and (c) no gasket (metal-to-metal contact).



(a) (b) (c)

Fig. 10. SEM images of gasket material surfaces before and after corrosion: (a) conductive rubber, (b) finger stock, and (c) iron.

shown in Fig. 10, after 80 hours of corrosion, the surface of the conductive rubber remains mostly intact, whereas significant corrosion deposits are observed on the surfaces of the finger stock and iron. These deposits, primarily composed of metal oxides and metal salts, have low electrical conductivity and contribute to the loss of electrical continuity. The transfer impedance of the gasket material, both before and after corrosion, is measured according to the SAE ARP-1705C standard, and the results are presented in Fig. 11. Results show that the transfer impedance of the conductive rubber gasket did not increase significantly after corrosion, whereas the transfer impedance of the finger stock gasket and the iron increased substantially. Since SE is inversely related to transfer impedance, it can be concluded that the samples with finger stock gaskets and those without gaskets will experience noticeable decreases in SE.

The results shown in Figs. 9–11 indicate that when corrosion occurs at gasketed seams, corrosion products tend to accumulate on the contact surface. These deposits, mainly consisting of metal oxides and metal salts, have low electrical conductivity. As a result, surface conductivity at the contact interface decreases,

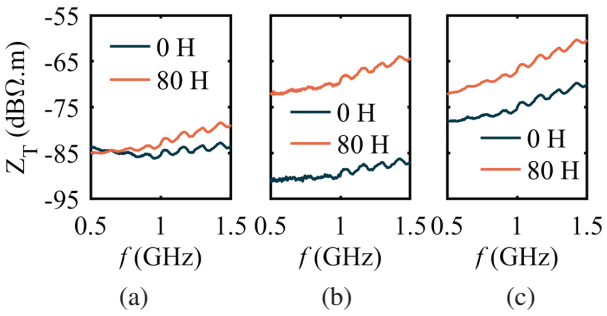


Fig. 11. Transfer impedance of gasket material before and after corrosion: (a) conductive rubber, (b) finger stock, and (c) iron.

transfer impedance of the gasketed seam increases, and electrical continuity is weakened. When surface currents induced by incident electromagnetic waves flow across the gasketed seam, the increased transfer impedance causes a notable voltage drop. This voltage drop acts as a secondary radiation source, introducing electromagnetic energy into the enclosure's interior. As a consequence, electromagnetic leakage rises, which can lead to a reduction in SE.

B. SE test of the enclosure

The samples shown in Fig. 9 are assembled with the box body to form a complete shielding enclosure. The SE test is then conducted in accordance with the standard specified in [10]. The SE test setup is shown in Fig. 12. The test system consists of an R&S® ZVL3 2-port vector network analyzer (VNA), a power amplifier (PA), a double-ridged horn antenna, a half-wave dipole antenna, and connecting cables. In this work, SE is tested in the frequency range of 500 MHz to 1.5 GHz.

In Fig. 12, the excitation signal generated by Port 1 of the VNA is amplified by the PA and transmitted to the double-ridged horn antenna, which radiates EM waves toward the cover plate of the enclosure. Inside the enclosure, a half-wave dipole antenna serves as the receiving antenna. It captures the EM waves that penetrate the enclosure and delivers the received signal to Port 2 of the VNA. The SE of the enclosure is then calculated based on the measured scattering parameter S21. Specifically, the measurement result obtained without the enclosure is recorded as $S_{21,B}$, and the result obtained after the enclosure is installed is recorded as $S_{21,T}$. SE of the enclosure, expressed in decibels (dB), is then calculated using:

$$\mathrm{SE} = S_{21,\mathrm{B}} - S_{21,\mathrm{T}}. \tag{9}$$

C. Experimental results and analysis

Following the SE testing procedure, the samples with the three different seam configurations undergo SE measurements. The SE test results of the enclosure with conductive rubber gasketed seams are shown

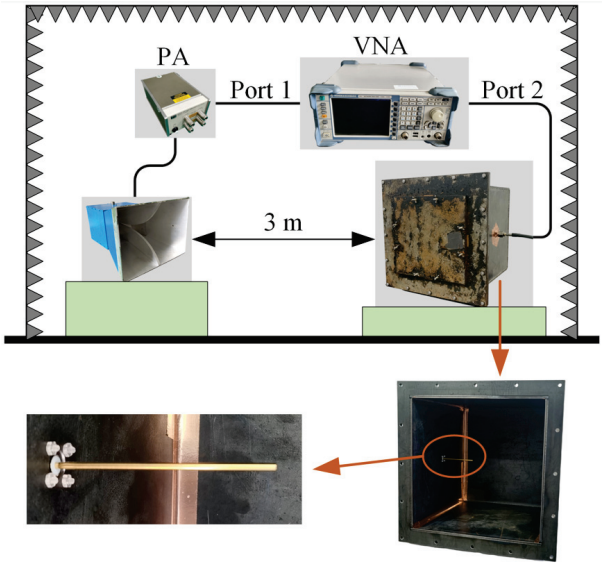


Fig. 12. SE test setup.

in Fig. 13. It can be observed that SE fluctuates only slightly throughout the entire test period. This result indicates that the conductive rubber gasket is not significantly affected by corrosion, demonstrating that the gasketed seam, where the conductive rubber gasket is installed, exhibits strong resistance to corrosive degradation.

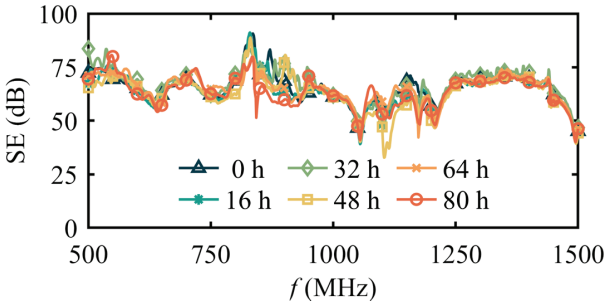


Fig. 13. SE enclosure with conductive rubber gasketed seams.

SE test results of the enclosure with finger stock gasketed seams are shown in Fig. 14. It is evident that SE progressively declines with longer corrosion durations. After 80 hours, SE decrease exceeds 20 dB, highlighting that the finger stock gasket fails to offer adequate corrosion protection, resulting in significant deterioration at the seams. A major factor is the presence of multiple gaps in the finger stock gasket, which lacks the sealing and water resistance characteristics that conductive rubber gaskets have. Consequently, corrosive solution accumulates within the gasketed seam, accelerating the corrosion process. The resulting corrosion byproducts,

which are poor conductors, hinder the electrical continuity of the gasketed seam and exacerbate EM leakage, ultimately diminishing the overall SE.

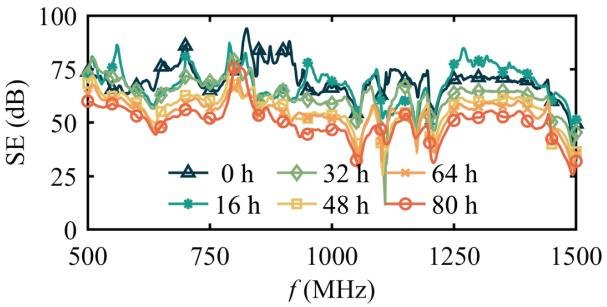


Fig. 14. SE of the enclosure with finger stock gasketed seams.

SE test results of the enclosure with metal-metal contact structure at different degradation durations are shown in Fig. 15. These results are similar to those observed with finger stock gasketed seams, showing a decrease in SE with increasing corrosion time. However, the reduction in SE is less pronounced. This result can be attributed to the fact that the two contacting plates are made of the same metal, which eliminates the potential difference required for galvanic corrosion. Only crevice corrosion occurs in the metal-metal contact structure. Consequently, the degradation of the metal-to-metal contact structure is less severe than that observed with finger stock gaskets.

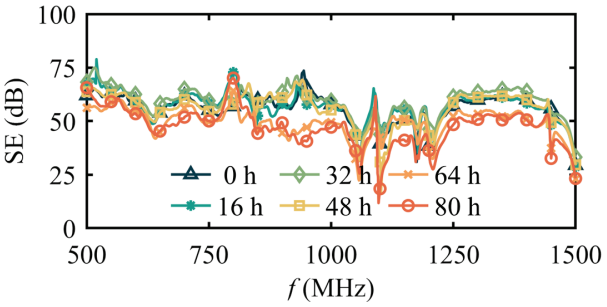


Fig. 15. SE of the enclosure with metal-metal contact structure.

Since the conductivity of the gasketed seam in the simulation model differs from that in the actual enclosure, the magnitude of the simulated and measured SE values is not consistent. Compared to the smooth simulation curve shown in Fig. 3, the measured SE curve exhibits fluctuations. This is mainly because the leakage signal level detected by the antenna inside the enclosure is very low, making the measurements vulnerable to instrument background noise. Additionally, the simulation model simplifies the structure by omitting the fastening bolts present in the real enclosure, leading to differences in resonant frequencies between simulation and measurement. Nonetheless, the overall trend of change in both simulation and measurement remains consistent.

SE of the three gasketed seam configurations in their initial uncorroded state is compared and shown in Fig. 16. It can be seen that the sample without any gasket has noticeably lower SE, while the samples with conductive rubber and finger stock gaskets show significantly higher SE with similar performance. This supports the idea that installing EMI gaskets improves electrical continuity at the seams, thereby enhancing the overall electromagnetic shielding performance of the enclosure.

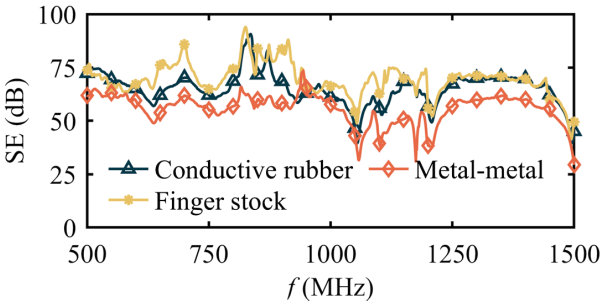


Fig. 16. Comparison of the initial SE for the three seam types.

This work calculated the average SE of samples with a gasketed seam at various stages of corrosion, as summarized in Table 1. The results in Table 1 indicate that among the three tested seam configurations, the seam with the conductive rubber gasket exhibited the least susceptibility to corrosion, while the seam with the finger stock gasket was the most affected. The superior performance of conductive rubber is attributed to its excellent deformability, which allows it to create a tight seal with the enclosure at the seam. This close contact offers some water-blocking ability, reducing the accumulation of corrosive solution on the contact surface. As a result, corrosion progresses more slowly, and the deterioration of SE remains minimal. In contrast, gasketed seams using a finger stock gasket (or those without any gasket) tend to have many small gaps where corrosive solution can accumulate, leading to crevice corrosion and a significant decrease in SE. Additionally, material mismatch between the finger stock gasket and the enclosure can cause galvanic corrosion at the contact surface, further speeding

Table 1: Average SE (dB) of three gasketed seams

Sample	0 h	16 h	32 h	48 h	64 h	80 h
Conductive Rubber	65.1	64.1	66.5	62.8	64.0	63.4
Finger Stock	69.5	68.1	62.8	56.1	54.1	49.7
Metal-Metal	55.7	56.8	57.1	55.9	49.1	47.5

up degradation. Consequently, the sample with the finger stock gasket shows the most notable SE degradation.

To improve the durability of enclosures exposed to corrosive environments, it is essential to design seam structures with water-resistant features that prevent the accumulation of corrosive solutions at contact interfaces. When choosing gasket materials, conductive rubber is preferred due to its superior sealing and corrosion resistance. If finger stock gaskets need to be used, they should be made from the same material as the enclosure or from materials with similar corrosion potentials to reduce galvanic corrosion. These design strategies can greatly enhance the corrosion resistance of the gasketed seam and slow the loss of SE over time.

V. CONCLUSION

This work investigated the degradation of electromagnetic shielding effectiveness (SE) of EMC gaskets caused by corrosion, using an integrated approach that combines experimental corrosion testing, SE measurements, and full-wave electromagnetic simulations in HFSS. Experimental results show that when exposed to a corrosive environment, the contact surface between the enclosure and the electromagnetic interference (EMI) gasket experiences corrosion. Low-conductivity corrosion products accumulate at the interface, increasing transfer impedance across the gasketed seam and decreasing electrical continuity. These changes lead to greater electromagnetic leakage and a corresponding reduction in SE.

The main contributions of this work are as follows: (a) Unlike most studies that evaluate the shielding performance of degraded EMI gaskets in isolation, this work installs the EMI gaskets on an enclosure before subjecting them to corrosion and measuring the SE. This approach better reflects real-world applications and the degradation behavior of EMI-gasketed seams. (b) While existing studies primarily rely on experimental measurements, this work combines both experimental testing and full-wave electromagnetic simulations conducted in HFSS. Shielding degradation is modeled by adjusting the conductivity of the gasket in the simulation model. Once a relationship between measurable physical quantities (such as transfer impedance) and simulated gasket conductivity is established, SE degradation can be accurately predicted. (c) By comparing the corrosion-induced degradation characteristics of different EMI gasket types, this study offers practical recommendations for material selection and seam design in enclosures exposed to corrosive environments. These insights can help mitigate corrosion effects and enhance shielding reliability in real-world applications.

Future work will focus on developing efficient and accurate electromagnetic simulation methods to model EMI gasket degradation caused by corrosion.

REFERENCES

[1] G. J. Gooch and J. K. Daher, *Electromagnetic Shielding and Corrosion Protection for Aerospace Vehicles*. Berlin: Springer, 2007.

[2] B. Wilmot, "EMI gaskets and their correct application," in *1992 Regional Symposium on Electromagnetic Compatibility*, Tel-Aviv, Israel, pp. 4.2.3/1-4.2.3/2, 1992.

[3] S. K. Das, J. Nuebel, and B. Zand, "An investigation on the sources of shielding degradation for gaskets with zinc coated steel enclosures," in *IEEE 1997, EMC, Austin Style. IEEE 1997 International Symposium on Electromagnetic Compatibility. Symposium Record*, Austin, TX, USA, pp. 66-71, 1997.

[4] J. F. Walther, "Electrical stability during vibration and electromagnetic pulse survivability of silver-plated glass bead filled EMI shielding gaskets," in *National Symposium on Electromagnetic Compatibility*, Denver, CO, USA, pp. 40-45, 1989.

[5] D. Pouhè and G. Mönich, "Assessment of shielding effectiveness of gaskets by means of the modified Bethe's coupling theory," *IEEE Transactions on Electromagnetic Compatibility*, vol. 50, no. 2, pp. 305-315, May 2008.

[6] D. Pouhè and P. Tcheg, "Study on corrosion-induced shielding-degradation of EMI gaskets," *IEEE Transactions on Electromagnetic Compatibility*, vol. 58, no. 4, pp. 1052-1059, Aug. 2016.

[7] X. Yan, Y. Cong, G. Wang, and Z. Zhang, "Effects of various accelerated corrosion conditions on corrosion behavior of galvanized steel in marine atmospheric environment," in *2022 IEEE Sustainable Power and Energy Conference (iSPEC)*, Perth, Australia, pp. 1-5, 2022.

[8] J. Xie, M. G. Pecht, S. K. Das, J. Nuebel, and B. Zand, "Corrosion of a zinc-coated steel enclosure at the contact interfaces of gasket joints," *IEEE Transactions on Components and Packaging Technologies*, vol. 23, no. 1, pp. 136-142, Mar. 2000.

[9] A. R. Pawlowych, R. J. Thibeau, and M. R. Lambert, "Galvanically compatible elastomeric gasketing material for EMI shielding applications," in *2004 International Symposium on Electromagnetic Compatibility (IEEE Cat. No.04CH37559)*, Silicon Valley, CA, USA, vol. 3, pp. 841-845, 2004.

[10] *IEEE Standard Method for Measuring the Effectiveness of Electromagnetic Shielding Enclosures*, IEEE Std 299-2006, Feb. 2007.

[11] Z. Yan, F. Qin, J. Cai, S. Zhong, and J. Lin, "Shielding performance of materials under the excitation of high-intensity transient electromagnetic pulse," *IEEE Access*, vol. 9, pp. 49697-49704, 2021.

[12] Z. Yan, F. Qin, J. Cai, S. Zhong, and J. Lin, "Study on the characterization of shielding effectiveness of

materials under wide band electromagnetic pulse," in *2020 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO)*, Hangzhou, China, pp. 1-3, 2020.

[13] M. Pavlík, A. Gladyr, and J. Zbojovský, "Comparison of measured and simulated data of shielding effectiveness, reflection and absorption of electromagnetic field," in *2020 IEEE Problems of Automated Electrodrive. Theory and Practice (PAEP)*, Kremenchuk, Ukraine, pp. 1-4, 2020.

[14] P. Lessner and D. Inman, "Quantitative measurement of the degradation of EMI shielding and mating flange materials after environmental exposure," in *1993 International Symposium on Electromagnetic Compatibility*, Dallas, TX, USA, pp. 207-213, 1993.

[15] J. P. Quine and A. J. Pesta, "Shielding effectiveness of an enclosure employing gasketed seams-relation between SE and gasket transfer impedance," in *Proceedings of International Symposium on Electromagnetic Compatibility*, Atlanta, GA, USA, pp. 392-395, 1995.

[16] H. W. Denny and K. R. Shouse, "EMI shielding of conductive gaskets in corrosive environments," in *IEEE International Symposium on Electromagnetic Compatibility*, Washington, DC, USA, pp. 20-24, 1990.

[17] D. Pissoort, J. Catrysse, T. Claeys, F. Vanhee, B. Boesman, and C. Brull, "Towards a stripline setup to characterise the effects of corrosion and ageing on the shielding effectiveness of EMI gaskets," in *2015 IEEE International Symposium on Electromagnetic Compatibility (EMC)*, Dresden, Germany, pp. 7-12, 2015.

[18] A. N. Faught, "An introduction to shield joint evaluation using EMI gasket transfer impedance data," in *1982 IEEE International Symposium on Electromagnetic Compatibility*, Santa Clara, CA, USA, pp. 1-7, 1982.

[19] R. J. Mohr, "Evaluation technique for EMI seams," in *1987 IEEE International Symposium on Electromagnetic Compatibility*, Atlanta, GA, USA, pp. 1-4, 1987.

[20] *Test procedure to measure the RF shielding characteristics of EMI gaskets,* SAE-ARP 1173-2004, 2004.

[21] G. Zhang, X. He, L. Wang, D. Yang, Y. Shi, H. Zhang, and A. Duffy, "Equivalent simulation approach for the shielding effectiveness of enclosures with gasketed seams," *IEEE Transactions on Electromagnetic Compatibility*, pp. 1-10, 2025.

[22] S. Celozzi, R. Araneo, P. Burghignoli, and G. Lovat, *Electromagnetic Shielding: Theory and Applications*. Hoboken, NJ: John Wiley & Sons, 2023.

[23] S. Kovar, J. Valouch, and H. Urbancokova, "Calculation of shielding effectiveness of materials for security devices," *MATEC Web of Conferences*, vol. 125, p. 02036, 2017.

Xin He was born in Wenshan, China, in 1996. He received the B.S. and M.Sc. degrees in electrical engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 2019 and 2021, respectively. He is currently working toward the Ph.D. degree in electrical engineering with HIT, Shenzhen. His research interests include cable fault detection and location, and electromagnetic compatibility analysis.

Gang Zhang was born in Tai'an, China, in 1984. He received the B.S. degree in electrical engineering from the China University of Petroleum, Dongying, China, in 2007, and the M.S. and Ph.D. degrees in electrical engineering from the Harbin Institute of Technology (HIT), Harbin, in 2009 and 2014, respectively. He is currently a professor in electrical engineering with HIT, and a visiting professor with the University of L'Aquila, L'Aquila, Italy. His research interests include electrical contact theory, uncertainty analysis of electromagnetic compatibility, and the validation of CEM.

Yin Shi was born in Harbin, China, in 1999. He received the B.S. and M.Sc. degrees in electrical engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 2022 and 2025, respectively. His research interests include electrical contact theory, shielding effectiveness, and finite element simulation.

Lixin Wang received the B.S. degree in electrical engineering from Nankai University, Tianjin, China, in 1988, and the M.S. and Ph.D. degrees in electrical engineering from the Harbin Institute of Technology (HIT), Harbin, in 1991 and 1999, respectively. He is currently a professor of power electronics and electric drives with HIT, Shenzhen. He conducts research with Faults Online Monitoring and Diagnosis Laboratory, HIT, on a wide variety of topics including electromagnetic compatibility at the electronic system level, aircraft electromechanical fault diagnosis expert system, and prediction and health management (PHM) of Li-ion battery.

Zhongliang Du was born in Mudanjiang, China, in 2000. He received the B.S. degree in electrical engineering from Shandong University, Jinan, China, in 2023. He is currently pursuing the M.Sc. degree in electrical engineering with Harbin Institute of Technology (HIT). His research interests include electromagnetic compatibility technology of power electronic devices.

Zhitian Wang was born in Changchun, China, in 2002. He received the B.S. degree in electrical engineering from Harbin Institute of Technology (HIT), Harbin, China, in 2023, where he is currently pursuing the M.Sc. degree. His research interests include modeling of conducted interference, design of active EMI filter, and finite element simulation.

Wear Analysis of Transmission Gear Tooth of Coal Mining Machine using the Finite Element Method

Shuilin Wang[1,2], Fanping Meng[3], Zhimin Cao[2], and Zhitao Liang[2]

[1]Xuhai College, China University of Mining and Technology
Xuzhou, 221008, China

[2]Institute of Intelligent Manufacturing and Smart Transportation, Suzhou City University
Suzhou, 215104, China

[3]Henan Energy Group Co. Ltd
Zhengzhou, 450008, China
fpingmeng@163.com, fanping_meng58@outlook.com

Abstract – This research employs finite element analysis for simulation and calculation to investigate the causes of tooth wear in coal mining machine transmission gears and how it affects gear transmission performance. Utilising Hertz theory, we calculated the maximum contact stress and Hertz half-width during the gear transmission process. A three-dimensional geometric model of the coal mining machine's right rocker transmission system was created using Pro/E software, which was subsequently analyzed with ANSYS/LS-DYNA for gear meshing. The wear quantity on the gear tooth surface was determined using the Archard wear model, disregarding the effects of lubricant and gear tooth temperature rise. Our simulations revealed the wear distribution and changes in wear amount across various tooth surface wear models under different operating conditions. Notably, we found that wear quantity is directly correlated with tooth wear range and contact stress, with significant wear occurring at the top and root of the tooth. Furthermore, we conducted laser cladding remanufacturing experiments, optimizing process parameters to enhance wear resistance and fatigue strength. The microstructure of the remanufactured tooth face exhibited homogeneity and a lower friction factor than new surfaces. This study offers novel insights into the wear mechanisms of coal mining machine gears. It demonstrates the effectiveness of laser cladding technology in enhancing gear performance, providing practical implications for the design and maintenance of gear systems in harsh operating environments.

Index Terms – Archard wear model, coal mining machine gear transmission, finite element analysis, Hertz contact theory, tooth wear.

I. INTRODUCTION

Gearing is a typical and crucial component of transmission systems, which are essential to the proper operation of coal mining apparatus and equipment [1]. This paper aims to identify the key elements that lead to gear tooth degradation in transmission systems of coal mining machinery and examines how these elements can be assessed quantitatively to improve gear functionality. Utilizing finite element analysis, the study examines the factors contributing to tooth wear, analysing stress distribution and wear patterns across various operational conditions to enhance understanding of the wear mechanisms. The research offers three significant contributions: firstly, it employs advanced finite element modelling to replicate the gear meshing process, yielding important insights into contact stress and wear patterns; secondly, it utilizes Hertz theory to determine the peak contact stress, providing a theoretical basis for grasping gear interactions; and thirdly, it proposes practical suggestions for enhancing gear design and remanufacturing techniques, ultimately seeking to boost the reliability and efficiency of coal mining equipment. The efficiency and safety of coal mining are closely linked to the reliability of the transmission gear, which is the main component of the coal mining equipment. The transmission gears are prone to tooth wear due to the coal mining machine's complex operating environment and severe force conditions. Understanding Gear tooth wear in coal mining machines is essential for improving efficiency, safety, and cutting maintenance costs. Wear can diminish performance, escalate energy usage, and result in gear failures, which pose safety risks. By identifying the causes of wear, more effective maintenance strategies can be implemented, while advancements in technology

Submitted On: September 10, 2024
Accepted On: July 28, 2025

https://doi.org/10.13052/2025.ACES.J.400810

improve gear design and durability, ultimately leading to enhanced overall performance. In addition to lowering transmission efficiency, tooth wear can raise noise levels, induce mechanical vibrations, and potentially result in gear failure. The appearance of these issues will significantly increase equipment maintenance costs and cause more downtime, negatively impacting the productivity and financial benefits of coal mining machinery [2]. Thus, researching and analyzing the aspects that contribute to gear tooth wear, as well as its causes, is crucial from a practical standpoint.

In the subject of mechanical transmission, the analysis of gear tooth wear has long been significant. In actual use, contact tension, tooth roughness, material characteristics, and lubrication conditions all work together to cause gear wear [3]. A theoretical foundation for gear design is provided by the Hertz theory-based gear contact stress calculation method, which utilizes the maximum contact stress formula to accurately predict the stress distribution of gears during the meshing process. Hertz theory, however, ignores other intricate considerations such as nonlinear variations in tooth wear and dynamic friction characteristics. As a result, researchers have been progressively developing numerical analytical techniques based on the finite element method in recent years [4]. These techniques enable a more accurate simulation of the stress state and deformation behaviour of gears under various complex operating conditions.

Fatigue fractures in the gearbox housing of a coal mining machine's cutting part (CPGH) are analyzed by examining external coal cutting and internal gear meshing excitations, using finite element and modal analysis to identify stress concentrations and resonance frequencies. The impact of design parameters on stress is explored to recommend solutions for reducing stress concentration and preventing resonance failures. Finite element analysis predicts tooth root fatigue and Rolling Contact Fatigue (RCF) in spur gears made of steel, titanium, and coated materials. An innovative methodology for loaded tooth contact analysis (LTCA) of spiral bevel gears enhances misalignment predictions, thereby improving gear performance and durability [5-7]. Researching the issue of gear wear is now fraught with difficulties. Firstly, it is challenging to fully and effectively characterize the actual operating conditions of gears using classic analytical approaches due to the extremely nonlinear and time-varying properties of the stress distribution and wear behaviour of gear transmission systems. Second, the accuracy and computational efficiency of the present models are frequently insufficient to account for the multi-scale and multi-physical field coupling difficulties that arise during the gear meshing process, such as contact friction, heat transfer, and material damage [8]. Furthermore, simplified models are frequently utilized in traditional finite element analysis due to limitations in processing power and modelling complexity, which can cause the analysis results to deviate from real working conditions. The credibility of the results may be impacted, for instance, if the lubricant effect and temperature rise are not taken into account. In this scenario, the stress concentration area on the tooth surface determined by the study may differ somewhat from the real wear situation [9, 10]. Similar studies have explored wear in underground construction equipment such as disc cutters in shield tunnelling machines. For instance, one investigation under Baiyun Airport analysed the linear correlation between disc cutter wear and parameters like installation radius, cutting length, and mass-point velocity [11]. The study examined rock behaviour under repeated shield disc cutter impacts, revealing that internal damage accumulates over time, reducing rock strength. Key mechanical responses, including stress and modulus, changed with impact rate, enhancing understanding of the rock-breaking process in tunnelling [12]. In another domain, a fault diagnosis method combining Feature Mode Decomposition (FMD) with Multi-Scale (CNN) Convolutional neural network has been proposed to improve accuracy in helicopter gearbox fault detection. These cross-domain techniques highlight the growing trend of integrating advanced diagnostics and machine learning in wear analysis and machinery fault prediction [13].

Researchers have proposed various changes to address these difficulties. By determining the contact stress and wear range, the wear calculation approach based on Archard's wear equation is frequently used in gear wear prediction to estimate tooth wear. Archard's wear model is crucial in studying gear tooth wear, particularly in coal mining machinery, as it provides a theoretical framework for predicting wear rates based on contact mechanics. This model enables a quantitative wear analysis, allowing engineers to design more durable gears by considering material properties, surface treatments, and operating conditions. It also correlates wear with load, sliding distance, and material hardness, thereby aiding in the development of better maintenance strategies and design improvements. The Archard wear equation is:

$$V = k \cdot H \cdot F \cdot d. \tag{1}$$

Where:
V = Volume of material worn away (cubic units)
k = Wear coefficient (dimensionless)
F = Normal load (Newtons)
d = Sliding distance (meters or millimetres)
H = Hardness of the material (Pascals)

This equation helps to quantify wear and predict gear performance, making it essential for improving

design and durability. The nonlinear fluctuation of stress and deformation at each site during gear meshing is not fully taken into account, even though this method makes the wear calculation simpler [14, 15]. The results also depend on the choice of wear coefficients. Furthermore, the numerical simulation method of finite element analysis offers substantial support to the gear wear study. The computation accuracy can be significantly improved by carefully selecting factors such as cell type, mesh density, and material model. It remains a challenging task to address, nevertheless, how to strike a compromise between the model's correctness and computational efficiency while avoiding instability issues (such as the negative volume phenomenon) throughout the calculation process [16].

In this study, several novel contributions to gear wear analysis are introduced, including the integration of advanced finite element analysis (FEA) with local mesh encryption techniques, improving the accuracy of wear predictions by more precisely evaluating stress concentration areas on gear tooth surfaces. Unlike previous studies that used simplified wear models, this research combines the Archard wear model with detailed simulations to capture the complex interactions between contact stress, nonlinear stress fluctuations, and varying operational conditions. The real-world applications by validating its findings against empirical data from coal mining machinery and enhancing the credibility of the wear predictions are highlighted. Furthermore, it explores the impact of gear wear on remanufacturing processes, particularly through laser cladding, offering new insights into improving wear resistance and fatigue strength. Finally, the research takes a holistic approach, correlating wear with overall gear performance to provide actionable recommendations for strengthening gear design and maintenance strategies. With the aforementioned issues in mind, this research uses the finite element method to conduct a systematic investigation into the wear of the tooth surface of the transmission gears in coal mining machines [17-19]. This paper calculates the contact half-width based on Hertz theory and performs local mesh refinement in the tooth contact region to achieve more precise findings in tooth stress analysis. In addition to improving the accuracy of tooth stress distribution, local encryption can significantly reduce computation time compared to the conventional uniform meshing method [20, 21].

Several tooth wear loss models, including the unworn, full desensitized, and localized wear models, are developed in this paper. These models enable the simulation of the impacts of various wear levels on the gear transmission system's performance, leading to a more comprehensive understanding of wear behaviour. Gear wear from friction, pressure, and insufficient lubrication lowers efficiency, load capacity, and longevity. Common kinds include tooth flank wear, pitting, and scuffing. As wear progresses, friction, noise, vibration, and temperature rise, potentially leading to failure. Wear rates are influenced by factors such as load, speed, material, and environmental conditions. To minimize the impacts, adequate lubrication, long-lasting materials, an ideal design, and regular maintenance are necessary. Monitoring wear through inspections and analysis extends gear life and ensures system reliability. For the potential negative volume issue that arises throughout the computation procedure, corresponding remedies are also suggested. Computational instability can be efficiently prevented and the reliability of simulation results increased by adjusting the K-file control parameters.In LS-DYNA, the K-file refers to the keyword input file that contains the control parameters required for finite element analysis. This file defines the essential settings for the simulation, including time integration options, contact definitions, material models, and numerical stability controls. By adjusting the K-file parameters, it is possible to mitigate computational instabilities, such as negative volume errors, while also improving the overall accuracy and reliability of the simulation results.

II. HERTZ THEORY-BASED GEAR CONTACT CALCULATION

A. Maximum contact stress calculation in gearing

The maximal contact stress is determined by Hertz theory [22].

$$\sigma_H = \sqrt{\frac{4-\varepsilon}{3}}\sqrt{\frac{2}{\sin\alpha\cos\alpha}} \sqrt{\frac{1}{\pi\left(\frac{1-\mu_1^2}{E_1}+\frac{1-\mu_2^2}{E_2}\right)}}\sqrt{\frac{2KT_1}{bd_1^2}\cdot\frac{u+1}{u}}. \tag{2}$$

$$\varepsilon_\alpha = \frac{1}{2\pi}\left[z_1\left(\tan\alpha_{a1}-\tan\alpha\right)+z_2\left(\tan\alpha_{a2}-\tan\alpha\right)\right]. \tag{3}$$

The degree of overlap of the straight-toothed cylindrical gear transmission is expressed as ε. The pressure angle is denoted by α (°), while α_1 and α_2 represent the apex circle pressure angles of the master and slave gears, respectively. The number of teeth of the master and slave gears are denoted by Z_1 and Z_2. The modulus of elasticity of the material is represented by E (MPa), and the material's Poisson's ratio is ν. The load factor is indicated as K, and the real maximum driving torque of the main wheel is given in N·mm. The transmission ratio is denoted by a, the gear tooth width is represented by b, and d refers to the indexing circle diameter of the main wheel.

B. Hertz half-width

The Hertz contact hypothesis states that for correct analysis results, the mesh at the contact should be drawn more densely. When any point of the gear is involved in meshing, the edge length of the cell at the contact is typically 1/10 of the Hertz half-width or less, meaning that there are more than five contact cells [23]. The equation for Hertz half-width is

$$a = \sqrt{\frac{8T_1\left(\frac{1-\mu_1^2}{E_1}+\frac{1-\mu_2^2}{E_2}\right)}{d_1 L\pi\left(\frac{1}{\rho}+\frac{1}{\rho_2}\right)\cos\alpha}}. \tag{4}$$

III. ANALYSIS OF GEAR MESHING TRANSMISSION USING FINITE ELEMENTS

A. Estimating dental wear

Wear on the tooth surface was calculated using the wear equation developed by Archard [24].

$$h_i = k\sum_{j=1}^{m} p_{i,j}s_{i,j}, \tag{5}$$

where $p_{i,j}$ is the tooth contact stress in MPa, $s_{i,j}$ is the tooth wear range in mm; i is the tooth's position where the calculation point is in mm; j is the total number of wear events that the calculation point has experienced.

During the Δt period, the gear tooth profile experiences a grinding stroke that is proportional to the pressure angles $\alpha 1$ and $\alpha 2$.

$$\begin{aligned}\Delta = |\Delta_{s1} - \Delta_{s2}| = r_{v1}\tan\alpha_1\sec^2\alpha_1\Delta\alpha_1 \\ - r_{v2}\tan\alpha_2\sec^2\alpha_2\Delta\alpha_2.\end{aligned} \tag{6}$$

This is derived from the fact that, throughout the gear meshing process, the length of the meshing line remains constant.

$$r_{b1}\tan\alpha_1 + r_{b2}\tan\alpha_2 = a\sin\alpha. \tag{7}$$

The equation involves the following: Here, α is the pressure angle at the gear tooth node (°), a is the center distance of the gear pair (mm), r_b is the base circle radius of the master and slave gears (mm), and r_{b1} and r_{b2} are the base circle radii of the master and slave gears in mm, respectively.

B. Finite element modeling

Table 1 displays the geometric measurements and material characteristics of the final-stage gears of the easy-to-maintain SL1000 coal mining machine's right rocker arm drive system.

Pro/E parametric modelling was used to create a three-dimensional geometric model of the gear drive system, taking into account the system's large size and heavy mass, which resulted in a large number of complete gear model units. The analysis and solution time were excessively long, which had an impact on the solution's efficiency. The six-tooth active wheel and seven-tooth driven wheel types are established based on the Saint-Venant concept.

Table 1: Gear parameters and material properties

Main and Driven Gears Geometric Parameter	Value 1	Value 2	Unit/Note
Number of teeth	27	28	–
Modulus	28		mm
Pressure angle	23		°
Tooth width	312		mm
Gear material	18Cr2Ni4WA		
Elastic modulus E	1.97×10^5 MPa		
Poisson's ratio μ	0.288		
Mass density ρ	7.91×10^{-6} kg/mm^3		
Yield limit σ_s	1.5×10^3 MPa		

After importing the model into ANSYS/LS-DYNA, the Hertz formula was used to determine the Hertz half-width, which was found to be a = 4.32 mm. After choosing shell163 four-node tetragonal cells and solid164 eight-node hexahedral cells, the model was subjected to mapping and free-meshing using MESHTOOL [25]. To minimizes computation time and increase the accuracy of the solution, the portion of the mesh that is not involved in meshing can be divided into rough sections. For the meshing of the tooth surface, local mesh refinement is applied to the tooth root stress concentration, with an edge length of 4mm for the cells.

Tooth wear failure frequently occurs in the active wheel at the meshing teeth above the pitch circle, near the tooth top, and in the driven wheel at the meshing teeth below the pitch circle, near the tooth root [26]. This is evident from the coal mining machine rocker arm final-stage gears in the actual working process, as well as from the literature [27, 28].

Based on the Archard wear formula, j = 1,000, meaning that only the active wheel characteristics should be considered, as the difference in size between the slave and master gears is not significant. The wear quantity of the tooth at the root is calculated to be 0.0432 mm, and at the top, it is 0.00594 mm. Consequently, the simulation Model 1 shows no wear on the tooth surface; Model 2 shows full damage wear on the active wheel gear tooth surface; and Model 3 shows wear of 0.00594 mm on the active wheel gear tooth surface. Following the split of the finite element model, Figs 1 to 3 show Model 3, which represents the follower wheel gear tooth flank wear of 0.0432 mm among the three typical wheel tooth flank wear loss models.

Figure 3 illustrate different simulation results: the left shows the normalized wear depth distribution, while the right shows the corresponding convergence coefficient (CONV-HCOE).

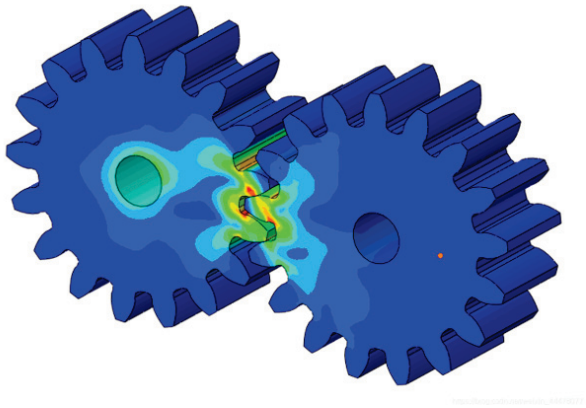


Fig. 1. Unworn gear tooth surface (Model 1).

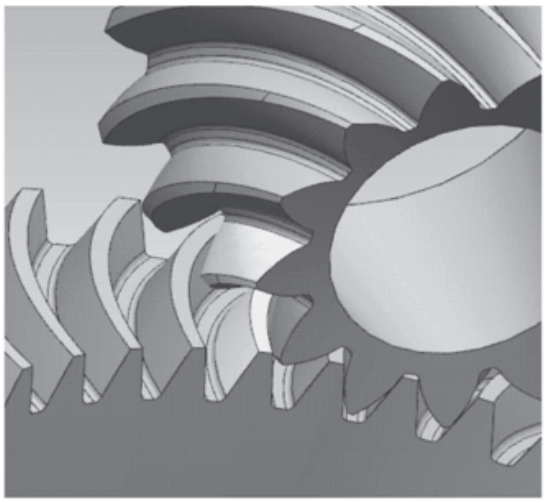


Fig. 2. Fully worn gear tooth surface on the active wheel (Model 2).

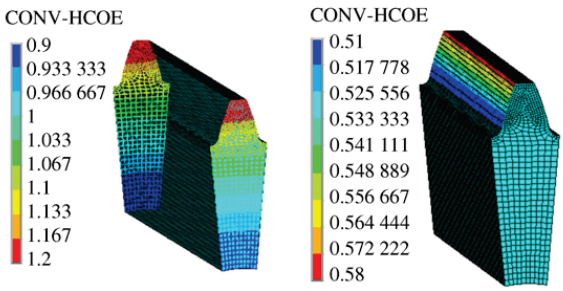


Fig. 3. Partially worn follower gear tooth surface with root and top wear (Model 3).

IV. EXPERIMENTATION

A. Preparing worn, remanufactured teeth

There are various procedures involved in preparing worn gear teeth for remanufacturing. First, an initial check evaluates the wear and damage to the gear. Sandblasting is then used to clean the gear's surface, eliminating pollutants and preparing it for subsequent treatment. Following that, cracks are identified using procedures such as optical inspection, liquid penetrant testing, magnetic particle testing, and ultrasonography. Once cracks and faults are discovered, a thorough inspection decides whether the gear can be fixed or replaced. If the gear is repairable, it undergoes additional preparation, such as removing fatigue layers and applying surface treatments to restore its functionality. Throughout the process, documentation is kept to chronicle discoveries and actions performed. The working surface hardness of the gear teeth on the low-speed, heavy-duty walking track gear steel grade 18Cr2Ni4WA was measured to be between HRC58.6 and 61.5, with a tooth core measuring HRC38.2. To simplify the experimental process and provide 48 mm, or half of the tooth width, for the study, a wire cutter will be used to cut out the block, depicted in Fig. 4, along the root of the wheel teeth. Figure 4 (a) depicts the pitting area following sandblasting to identify the fatigue pitting area, whereas Fig. 4 (b) shows the end face broken tooth drop block. The fatigue pitting region in Fig. 4 (a) is dispersed over the upper half of the tooth height and is wider in the vicinity of the tooth's end face. The upper portion of the tooth height and the top of the tooth are where the majority of the broken teeth are located in the Fig. 4 (b). The distribution of the fatigue region location was clarified after sandblasting, oil removal, and rust treatment. Slag removal and fault identification are necessary before laser cladding; otherwise, the slag and the original residual fractures will constitute the cause of cracks in the cladding layer.

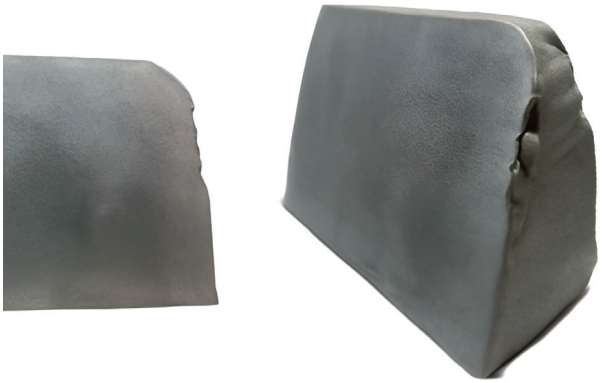


Fig. 4. Gear teeth following sandblasting: (a) Prolonged tensile deformation area; (b) Oblique tensile deformation area

In our study, we utilized a belt sander to "sandblast" the gear tooth work surface multiple times until the pitting craters are entirely gone to remove the fatigue

pitting layer. The gear tooth's working face was then liquid probed and kept sanded to check for any cracks. The thickness of the unilateral gear tooth's working face was measured to have decreased by 0.25 mm following the removal of the fatigue layer and the discovery of cracks.

B. Laser remanufacturing of tooth flanks

The powdered nickel-based fusion alloy and blended Ni60A: CeO2 = 99:1 (mass ratio) were ball-milled for 8 hours at a rotational speed of 650 r/min. The powder was then dried in a vacuum drying furnace at 150 °C for 1.5 hours to completely dehumidify it and prevent the formation of air holes caused by the failure of water vapour to be discharged from the metal during the fusion cladding process, due to the metal's high-speed solidification, as shown in Table 2.

Table 2: Parameters of the laser cladding method

Variable Parameters	Numerical Value
Power kW	3.3
Spot size mm^2	1×10
Spot velocity (mm/s)	9
Rate of powder delivery (r/min)	4.8
Rate of overlap between neighboring fusion trajectories %	42
Powder that contains gas	Ar
Coaxial shielding flow rate of argon gas (L/h)	590

Figure 5 displays the working surface morphology of the gear teeth following laser cladding.

(a) (b)

Fig. 5. Functional shape of wheel teeth after cladding: (a) Gear working surface; (b) Gear tooth end showing the start and end of laser cladding.

After welding, the rectangular laser spot has a surface that is nearly smooth and clean, similar to net moulding, as seen in Fig. 6. The rectangular spot has two major advantages over the circular spot, in addition to surface finishing: firstly, the surface interaction with the substrate is smooth; secondly, as illustrated in Fig. 6, the melting efficiency is high.

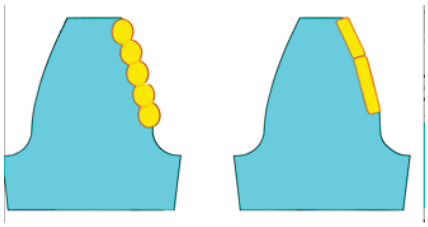


Fig. 6. Fused cladding bonding interface schematic diagram: (a) Circular polishing; (b) Short polishing.

Abrupt changes in stress at the bonding interface are easily produced because the cladding layer's material properties, such as Poisson's ratio and modulus of elasticity, differ from the substrate's. Interestingly, Laves phase, a brittle material at ambient temperature, is the material present at the bonding interface. Due to the Laves phase's brittleness, it is best to reduce both its content and overall area, as a large amount of Laves phase will weaken the mechanical qualities of the tooth surface.

C. Microstructure near the overlap zone's interface

As seen in Fig. 7, the upper surface of the matching coating is smooth, clean, and neat. The interface between the lower portion of the fused cladding layer and the substrate at the adjacent two laser laps is reasonably flat, with nearly no fluctuation or undulation. A "reasonably flat" interface between the cladding layer and the substrate in laser remanufactured coatings is essential for gear teeth' mechanical integrity and performance, especially in demanding environments like coal mining machinery. A flat interface promotes even load distribution, reducing localised stress concentrations that can cause premature failures such as fatigue, cracking, and spalling, thereby enhancing gear durability and longevity. The rapid geometry changes lead to high-stress points, increasing the risk of wear and failure. Improved bonding between the cladding and substrate is facilitated by a flat interface, which prevents delamination and ensures efficient load transfer. The wear resistance of the cladding layer benefits from a flat surface, which allows uniform material distribution and mitigates rapid wear. Additionally, a flat interface aids in thermal management by enhancing heat dissipation, thereby preventing localizeds overheating that could degrade material properties. From a microstructural perspective, a flat interface promotes a consistent microstructure, which is essential for achieving the desired hardness and toughness, thereby reducing wear inconsistencies. Overall, sustaining a "reasonably flat" interface is crucial for gear systems' longevity, wear resistance, and reliability, enhancing bonding quality and thermal management while minimising stress concentrations. The laser remanufactured fused cladding layer is free of defects such as

pores, fissures, and unfused particles, and it is uniformly and densely organised.

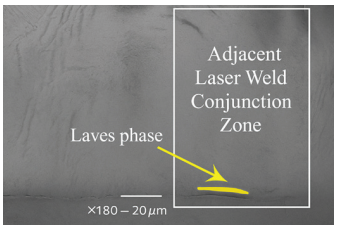


Fig. 7. Microstructure of the adjacent laser weld conjunction zone showing the presence of the Laves phase (magnification: ×180, scale: 20 μm).

Compared to the wavy binding interface of a circular spot, the gentle binding interface of a rectangular spot has a smaller Laves phase area.

D. Laser-remanufactured composite coatings' wear properties

Samples of wear performance are taken from the new gear block at the same location as the rectangular laser-coated tooth face block. It can be conclusively demonstrated that the process parameters used for manufacturing laser wear-resistant coatings are successful if the fatigue zone is chosen as the wear test location and its test performance surpasses that of the new gear tooth surface.

The rectangular area in Fig. 9 (a) of 10 × 20 mm^2 and 10 mm in height along the tooth surface downward was selected as the wear surface to be examined. The tooth surface was then smoothed with sandpaper and polished with a polishing machine to a surface roughness of 0. 2 μm.

A UMT-2 friction and wear testing machine was used to conduct the experiments. The following were the working conditions: room temperature of 23 °C, loaded with 25 N, reciprocating discs in dry wear mode without lubrication, single test duration of 7,200 s, grinding ball diameter of 5.6 mm, GCr15 steel ball material, and 5 mm abrasion mark length.

It is evident from the comparison of Figs. 8 (a) and 8 (b) that the newly created tooth flanks perform significantly worse than the laser-prepared remanufactured wear-resistant coatings. Following an identical wear performance test, the laser cladding surface exhibits nearly no wear markings at all. In contrast, the wear pattern of the fresh flank samples displays larger wear grooves in the middle. In Figs. 8 (c) and 8 (d), the laser nickel-based coating and the new tooth face grinding GCr15 beads on the wear surface demonstrate a relatively flat and smooth wear morphology, when compared to the grinding of the GCr15 beads. Meanwhile, the new tooth face cutting block, when subjected to the grinding of the GCr15 beads from both sides of the gradual protrusion, clearly displays a micro-ploughing morphology.

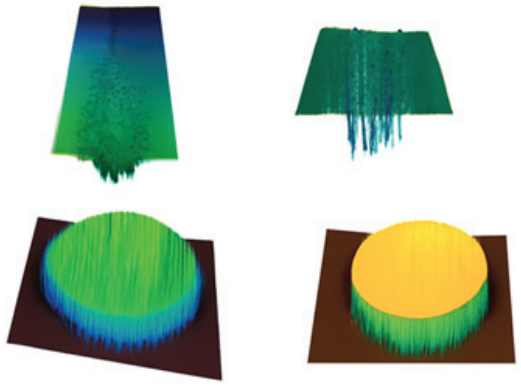


Fig. 8. 3D topographical surface plots of gear tooth surfaces after the abrasion test: (a) Least recently structured containment coating; (b) New producer tooth sonic sites; (c) Surface height variation up to 83 μm; (d) Surface height variation up to 40 μm.

E. Curve of the friction factor

The friction factor curve is a valuable tool for evaluating the performance of laser-remanufactured coatings compared to new tooth surfaces, particularly in gear systems. It measures the resistance to motion between two surfaces, taking into account surface roughness, material properties, and lubrication. The curve shows higher friction during the initial running-in phase, followed by a stable wear stage. By analyzing this curve, it is evident that laser-remanufactured coatings offer superior wear resistance, improved lubrication, and lower energy losses compared to new tooth surfaces, which tend to degrade more quickly. This makes remanufactured coatings more efficient and durable, making them ideal for applications that require long-term performance and lower operating costs. The curve also provides valuable insights for predictive maintenance, process optimisations, and material selection, helping reduce energy consumption, improve efficiency, and prevent unexpected failures. In summary, the friction factor curve is a crucial tool for evaluating the effectiveness of laser-remanufactured coatings in terms of wear resistance, efficiency, and gear performance. The friction factor of the laser remanufactured coating is consistently lower than the friction factor of the new tooth surface after the running-in process ends and the tooth surface enters the stable wear stage, as shown by the comparison between the two in Fig. 9. Additionally, the difference tends to continue growing.

Furthermore, there is a close relationship between wear performance and fatigue performance, as they are essentially the macroscopic expressions of the build-up

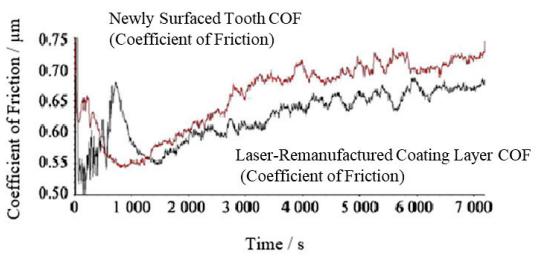


Fig. 9. Friction factor between a newly surfaced tooth and a laser-remanufactured coating.

of microcrystalline dislocations at the metal microcrystalline boundaries of the tooth surface, which eventually cause slippage and spalling. The finer grains in the fine and compact fine-grain reinforced wear-resistant coatings made possible by the addition of rare earths mean that, in theory, their fatigue performance will be higher than that of the tooth surface with coarser grains.

V. CONCLUSION

This study uses Hertz theory and finite element analysis to examine gear tooth surface wear in coal mining machinery, focusing on stress concentration at the tooth's crown and roots. Laser cladding remanufacturing develops wear resistance and fatigue strength. Still, the Archard wear model determines that tooth wear is directly related to sliding distance and contact stress, with wear increasing dramatically at the crown and root. This study enhances the gear wear mechanisms in coal mining machinery by providing empirical data on the performance of laser-remanufactured coatings compared to new tooth surfaces. The findings provide critical insights into optimizing gear design and remanufacturing processes, resulting in increased operational efficiency and reduced maintenance costs. The study's novelty is combining finite element analysis with experimental wear testing to estimate laser cladding performance. It helps innovation in gear material selection and surface treatments by correlating microstructural properties to wear behavior. Likewise, applying friction factor analysis as a predictive technique represents a big step forward in assessing gear systems under various operational conditions.

Declarations

Funding: This research was supported by the "Qinglan Project" of colleges and universities in Jiangsu Province, Grant Number: 2022.

Conflicts of interests: The authors declare no conflict of interest.

Data Availability Statement: No datasets were generated or analyzed during the current study.

Code availability: Not Applicable.

Authors' Contributions: Shuilin Wang, Zhimin Cao, Zhitao Liang: Responsible for designing the framework, analyzing the performance, validating the results, and writing the article. Fanping Meng: Responsible for collecting information, providing software, conducting a critical review, administering the process, and serving as the corresponding author.

REFERENCES

[1] X. Zhao, W. Fan, Z. Wang, Z. Wen, and P. Wang, "An explicit finite element approach for simulations of transient meshing contact of gear pairs and the resulting wear," *Wear*, vol. 523, Art. no. 204802, June 2023.

[2] A. G. Mustafayev and C. R. Nasirov, "A study of factors affecting wear and destruction of teeth in gear mechanisms," *Nafta-Gaz*, vol. 79, no. 9, pp. 604–610, Sept. 2023.

[3] J. Alam and S. Panda, "Multi-objective optimization with finite element analysis of profile shifted altered tooth sum spur gear," *Advances in Materials and Processing Technologies*, vol. 10, no. 2, pp. 1024–1051, Jan. 2024.

[4] Y. Yang, N. Hu, Y. Li, Z. Cheng, and G. Shen, "Dynamic modelling and analysis of planetary gear system for tooth fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 207, p. 110946, Jan. 2024.

[5] R. Zhang, Y. Zhang, L. Zhu, and J. Huang, "Fatigue failure mechanism of coal mining machine cutting gearbox housing," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 42, pp. 1–15, May 2020.

[6] S. Baragetti, "Finite element analysis and experiments for predicting fatigue and rolling contact fatigue behavior of spur gears," *Periodica Polytechnica Mechanical Engineering*, vol. 66, no. 4, pp. 304–313, Oct. 2022.

[7] Y. Liu, "A semi-analytical loaded contact model and load tooth contact analysis approach of ease-off spiral bevel gears," *Machines*, vol. 12, no. 9, p. 623, Sept. 2024.

[8] A. Czakó, K. Řehák, A. Prokop, J. Rekem, D. Lástic, and M. Trochta, "Static transmission error measurement of various gear-shaft systems by finite element analysis," *Journal of Measurements in Engineering*, vol. 12, no. 1, pp. 183–198, Dec. 2024.

[9] S. Mo, L. Wang, Q. Hu, G. Cen, and Y. Huang, "Coupling failure dynamics of tooth surface morphology and wear based on fractal theory," *Nonlinear Dynamics*, vol. 112, no. 1, pp. 175–195, Nov. 2024.

[10] M. Li, Y. Luo, L. Xie, C. Tong, and C. Chen, "Fatigue reliability assessment method for wind power gear system based on multidimensional finite element method," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 238, no. 3, pp. 540–558, Apr. 2024.

[11] B. Zou, Y. Chen, Y. Bao, Z. Liu, B. Hu, J. Ma, and X. Long, "Impact of tunneling parameters on disc cutter wear during rock breaking in transient conditions," *Wear*, vols. 560–561, Art. no. 205620, 2025.

[12] B. Zou, J. Yin, Z. Liu, and X. Long, "Transient rock breaking characteristics by successive impact of shield disc cutters under confining pressure conditions," *Tunnelling and Underground Space Technology*, vol. 150, Art. no. 105861, 2024.

[13] A. Wan, Z. Zhu, K. Al-Bukhaiti, X. Cheng, X. Ji, J. Wang, and T. Shan, "Fault diagnosis of helicopter accessory gearbox under multiple operating conditions based on feature mode decomposition and multi-scale convolutional neural networks," *Applied Soft Computing*, vol. 180, Art. no. 113403, 2025.

[14] D. Wu, C. Zhang, H. Wu, L. Ji, R. Ran, and Y. Xu, "Forest fire recognition based on feature extraction from multi-view images," *Traitement du Signal*, vol. 38, no. 3, pp. 775–783, June 2021.

[15] F. Andary, C. Heinzel, S. Wischmann, J. Berroth, and G. Jacobs, "Calculation of tooth pair stiffness by finite element analysis for the multibody simulation of flexible gear pairs with helical teeth and flank modifications," *Multibody System Dynamics*, vol. 59, no. 4, pp. 395–428, July 2023.

[16] Z. Ali, W. Imdad, Q. B. Jamali, I. A. Memo, F. A. Solangi, and S. Ahmed, "Tooth strength analysis of bevel gear using SolidWorks simulation tool," *Journal of Computing & Biomedical Informatics*, vol. 6, no. 1, pp. 230–237, Dec. 2023.

[17] Y. Wu, "Exploration of the integration and application of the modern new Chinese style interior design," *International Journal for Housing Science and its Applications*, vol. 45, no. 2, pp. 28–36, June 2024.

[18] W. Wang, "ESG performance on the financing cost of A-share listed companies and an empirical study," *International Journal for Housing Science and its Applications*, vol. 45, no. 2, pp. 1–7, June 2024.

[19] B. Pang, X. Wang, and T. Zhang, "Reliability analysis of thin-walled ring gear based on tooth surface fatigue and wear," *Journal of Mechanical Science and Technology*, vol. 38, no. 4, pp. 1985–1997, Apr. 2024.

[20] P. Sreekar, "Cost-effective cloud-based big data mining with K-means clustering: An analysis of Gaussian data," *International Journal of Engineering & Science Research*, vol. 10, no. 1, pp. 229–249, May 2020.

[21] N. S. Allur, "Big data-driven agricultural supply chain management: Trustworthy scheduling optimization with DSS and MILP techniques," *IEEE Access*, vol. 8, no. 4, pp. 1–16, Aug. 2020.

[22] S. Mo, L. Wang, M. Liu, Q. Hu, H. Bao, G. Cen, and Y. Huang, "Study of the time-varying mesh stiffness of two-stage planetary gear train considering tooth surface wear," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 238, no. 1, pp. 279–297, Apr. 2024.

[23] Y. Guan, J. Wang, Q. Wang, K. Ning, and B. Wang, "Three-arc gear coupling design analysis for high load-carrying applications," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 237, no. 24, pp. 5853–5864, 2023.

[24] C. Jia, G. Zhang, and G. Li, "Numerical analysis of tooth contact and wear characteristics of internal cylindrical gears with curved meshing line," *Applied Sciences*, vol. 14, no. 13, p. 5399, 2024.

[25] Z. Zhao, Y. Yang, H. Ma, H. Wang, H. Tian, and C. Han, "Meshing characteristics of spur gear pairs with tooth modification under different assembly errors and sensitivity analysis for impact factors," *Journal of Mechanical Science and Technology*, vol. 37, no. 1, pp. 149–162, Dec. 2023.

[26] L. Guo and Y. Sun, "Economic forecasting analysis of high-dimensional multifractal action based on financial time series," *International Journal for Housing Science and its Applications*, vol. 45, no. 1, pp. 11–19, 2024.

[27] K. Saqib, "Postmodernism, social dynamics, and e-commerce evolution," *International Journal for Housing Science and its Applications*, vol. 45, no. 1, pp. 20–24, 2024.

[28] I. M. Jamadar, R. Nithin, S. Nagashree, V. P. Prasad, M. Preetham, P. K. Samal, and S. Singh, "Spur gear fault detection using design of experiments and support vector machine (SVM) algorithm," *Journal of Failure Analysis and Prevention*, vol. 23, no. 5, pp. 2014–2028, Aug. 2023.

Shuilin Wang (born September 1983) is a female scholar of Han nationality from Bayannur, Inner Mongolia. She holds a Ph.D. and is an Associate Professor at Xuhai College, China University of Mining and Technology. Her main research interests include mining machinery design, fault diagnosis, and analysis.

Fanping Meng (born March 1981) is a male scholar of Han nationality from Weifang, Shandong Province. He holds a Master's degree and currently works in the Human Resources Department of Henan Energy Group Co., Ltd. His primary research focus is on the management of mining machinery and equipment.

Zhimin Cao male, holds a Master's degree and is affiliated with the Institute of Intelligent Manufacturing and Smart Transportation, Suzhou City University, Suzhou, China. His research focuses on the design and application of mining machinery as well as intelligent manufacturing technologies.

Zhitao Liang male, is affiliated with the Institute of Intelligent Manufacturing and Smart Transportation, Suzhou City University, Suzhou, China. His main research interests include mining machinery design, mechanical fault diagnosis, and intelligent manufacturing systems.

Analysis and Research on the Construction Stage of Suspension Bridge Steel Towers Using Midas Civil and ANSYS

Haodong Wang[1], Liu Jie[1], Fan Ji[2], Yongkai Zhou[1], and Liu Liang[2]

[1]Technical Center
Sichuan Road and Bridge East China Construction Co. Ltd.,
Chengdu, Sichuan Province 610000, China
657103053@qq.com, 281622164@qq.com, 121083920@qq.com

[2]Technical Center
Sichuan Highway and Bridge Construction Group Co. Ltd.,
Bridge Engineering Branch, Chengdu, Sichuan Province 610000, China
852772307@qq.com, 417232337@qq.com

***Abstract* –** Scholars have conducted extensive research on the anchoring methods and cable systems of suspension bridges, but there is relatively less research on the analysis of steel tower construction stages. As an important supporting part of the entire bridge, the stability of steel towers during installation directly affects the safety assurance during construction. In order to comprehensively understand the dynamic characteristics of steel towers during the construction stage, this paper presents a comprehensive analysis of the construction process of steel towers in large-span suspension bridges, focusing on the roles of active and cross braces in maintaining structural stability. Utilizing advanced finite element modelling techniques in both Midas Civil and ANSYS, the study evaluates the deformation and stress responses of the steel tower under various loading conditions, including self-weight and wind loads. The findings reveal critical insights into the maximum deformation behaviors and stress distributions at different construction stages, underscoring the importance of jacking operations and the strategic installation of braces. By comparing the performance of structures with and without these braces, the research demonstrates their essential role in enhancing the overall stability and safety of the tower during construction. Furthermore, considering the influence of wind loads and crane loads on the structure in Midas Civil, the paper analyzes the changes in structural strength and stiffness, providing useful references and guidance for this project and similar engineering endeavors.

***Index Terms* –** Construction stage, Midas Civil and ANSYS, steel towers, suspension bridge.

I. INTRODUCTION

As an important form of bridge structure, suspension bridges play a crucial role in modern transportation infrastructure. Their unique structural advantages provide reliable solutions for spanning natural obstacles, such as rivers and valleys. During the construction process of suspension bridges, the installation and control of bridge towers are critical stages. The steel tower, as an integral supporting part of the entire bridge, directly influences the safety and lifespan of the bridge. Scholars' research on suspension bridges tends to focus more on performance design, anchoring methods, and construction techniques. For instance, in the 1990s, countries such as the United States and European nations transitioned from safety factor methods to partial coefficient limit state methods in bridge design [1, 2]. Xiong et al. [3] based their work on load resistance and statistical characteristics of key design parameters to construct reliability-based limit state design equations for gravity-type anchorages in suspension bridges. Zhao et al. [4] conducted a numerical analysis of the entire bridge, including local anchorages, to investigate the safety factors of anchorages against sliding. A method was proposed for analyzing the construction process as a whole of a steel box girder ground-anchored suspension bridge using the finite element software Midas Civil [5]. The utilization of Midas Civil to establish a spatial finite element model of bridge towers, calculating and analyzing parameters such as pre-elevation, anti-push stiffness, anti-twist stiffness, and allowable deviation under three construction conditions — bare towers, suspended cables, and completed bridges — is employed [6]. The finite element software MIDAS/Civil was used

Submitted On: September 12, 2024
Accepted On: June 11, 2025

https://doi.org/10.13052/2025.ACES.J.400811

to simulate and analyze the asynchronous construction method of "tower-first, beam-later" during the construction phase [7]. Newton's Law was used to establish vibration control equations for main cable parameters, and studying the nonlinear parameter vibration issues during suspension bridge construction was also utilized [8]. Zhu et al. [9] employed MIDAS/Civil to establish a spatial analysis model for a suspension bridge during operation and utilized ANSYS to construct a solid finite element model for the concrete tower pylons for local analysis.

Although many countries worldwide have successfully constructed numerous large-span suspension bridges, deficiencies still exist in design theories and industry standards, especially in research on the analysis of steel tower construction stages [10]. Our proposed work integrates artificial neural networks, electrothermal inverter models, and finite element analysis, as demonstrated by Deevi, to enhance real-time thermal and structural analysis of steel towers during staged construction [11]. Incorporating intelligent FEA modelling enhances simulation fidelity, safety, and efficiency throughout the construction process. Kodadi employed cloud-based seismic management systems in machine learning and wavelet analysis for rapid data processing and accurate earthquake prediction [12]. Inspired by their cloud-based seismic management system, our research integrates the deployment of Midas Civil and ANSYS on a cloud-based high-performance computing (HPC) platform that facilitates real-time, large-scale structural and electromagnetic simulations. Cloud computing strategies, including load balancing, auto-scaling, and resource optimization, were presented by Dharma [13]. Our framework incorporates their Cloud computing strategies into our study to dynamically manage simulation workloads and maintain system stability through real-time monitoring, resulting in faster, cost-effective, and energy-efficient multiphysics simulations. Therefore, to comprehensively understand the stress and deformation states of steel towers during the construction stage, this paper, based on a specific Yangtze River bridge, utilizes Midas Civil for finite element analysis. The effectiveness of the analysis model is validated by comparing its results with those from the ANSYS finite element software. Furthermore, the construction stage analysis considers wind loads and crane working loads to investigate their effects on the stress and deformation of steel towers. The sensitivity of steel towers to various loads at different stages is analyzed, and the structural strength, stiffness, and overall stability are verified during the construction phase. The stress and deformation effects of various components in the main tower under different construction stages are also analyzed in detail, providing an important reference for subsequent construction activities.

This paper examines the construction stages of steel towers in large-span suspension bridges, with a focus on the impact of varying loading conditions on structural stability. It uses advanced finite element modelling techniques in Midas Civil and ANSYS to understand how active and cross braces affect deformation and stress distribution. The research provides a detailed examination of steel tower dynamics, emphasizing the importance of active braces for stability. It introduces innovative modelling approaches that incorporate variable cross-section designs, enhancing the reliability of the analysis. A comparative analysis between Midas Civil and ANSYS validates the effectiveness of simplified modelling techniques and offers insights into steel tower performance under various construction scenarios. The findings provide recommendations for engineering practices that improve safety and reliability in suspension bridge tower construction. This paper fills a significant gap in bridge engineering literature by focusing on steel tower construction.

This paper is organized as follows. Section II highlights key findings and gaps in current research. Section III presents the methodology used in this study, including a brief description of Midas Civil and ANSYS finite element model. Section IV presents the analysis of structural effects under different loads along with the analysis of calculation results under tower crane load. Finally, section V summarizes the key findings, conclusions, and recommendations.

II. PROJECT OVERVIEW

This article is based on the Zhangjinggao Yangtze River Bridge project. The bridge is located in the lower reaches of the Yangtze River, in the Chengtong section of the Sa Group in Rugao, China. The span of the bridge is 7,859 meters, with two main navigation bridges and three approach bridges (south, central, and north). The South Navigation Bridge, spanning 2,300 meters, is currently the world's largest span suspension bridge under construction. The precision requirements, deformation control, and automation welding rate for fabricating the Zhangjinggao Yangtze River Bridge steel towers are high. The vertical deviation of each segment must be controlled within 5 millimeters, and the planar deviation must be controlled within 2 millimeters. The north navigation bridge spans the Hanfu North Waterway in Rugao and features a main span of 1,208 meters, making it a single-span cable-stayed bridge. This article focuses on the analysis of the South Main Tower of the North Navigation Bridge. The south anchor tower adopts a gantry structure steel tower with a total height of 217 meters (including the tower crown). The tower crown is 9 meters high, with a height of 58.1 meters for the crossbeam and below and 149.9 meters above the crossbeam. The elevation of the tower crown is $+224.000$ meters,

the elevation of the tower top (excluding the crown) is +215.00 meters, and the elevation at the center of the tower bottom is +7.00 meters. Two crossbeams are set on the tower column, as shown in Fig. 1.

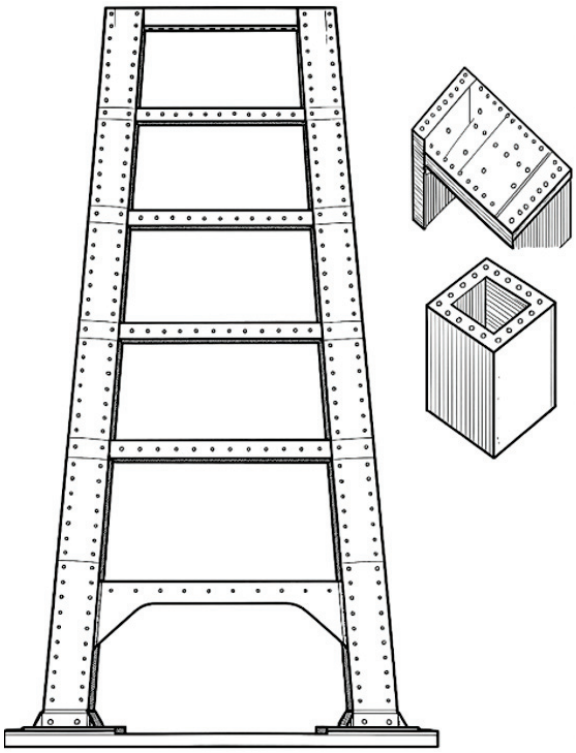


Fig. 1. Schematic diagram of the south tower of the North Channel Bridge.

III. FINITE ELEMENT MODEL VERIFICATION

A. Midas Civil finite element model

Midas Civil software is widely used in the engineering industry due to its high integration of functions and ease of modelling. In this paper, Midas Civil is employed for construction stage analysis. Due to the inherent inclination of the steel tower and the variable cross-section of each segment, the SPC section property calculator is used to import the steel tower sections, and the model is simulated using beam elements. The entire tower consists of 2,019 nodes and 3,563 components. The steel tower and crane base are fixed with constraints at the bottom. The active braces, upper beam supports, and lower beam supports are connected to the steel tower with varying degrees of freedom in elastic connections, depending on the construction stage. Temperature loads are used to simulate the jacking force of the jacks. The weight of each steel component is predicted based on its dimensions and the type of steel used, and its impact on the structure is considered throughout the entire construction plan [11]. Therefore, adjustments to the unit weight of each segment are made to match the design requirements.

During the gradual installation of bridge pier segments, they undergo gradual deformation due to their weight and other factors. Since the steel towers are installed during periods of relatively constant temperature and low wind speeds, during the comparative analysis stage, only the self-weight of the structure is considered, while factors such as wind loads and temperature loads are ignored. The primary objective of this analysis is to determine the overall deformation, stress, and the required jacking force for each brace of the steel tower, thereby facilitating comparison with the ANSYS finite element model to validate the effectiveness of this simplified model. Additionally, the calculation results are used to determine the most critical state of the structure during construction. The main construction stages and corresponding finite element models are shown in Table 1 and Figure 2.

Table 1: Main construction steps

Working Conditions	Construction Steps
1	Install T0-T7 segments
2	Install the lower beam bracket and push up the tower column
3	Locking jack and lower beam bracket:
4-5	Install lower beam segments 1-3 in sequence
6	Install the lower beam closing section and unlock it
7	Install sections T8-T11 and the first attachment of the tower crane
8-9	Install active cross brace 1 and push tower column
10	Locking jack and active cross brace 1, install segments T12-T14
11-12	Install active cross brace 2 and push the tower column
13	Locking jack and active cross brace 2
14	Install sections T15-T17 and install the remaining tower crane
15-16	Install active cross brace 3 and push the tower column
17-18	Locking jack and active cross brace 3, install segments T18-T22
19-20	Install the upper beam bracket and push up the tower column
21-23	Install the lower beam sections 1-3 to the closing section in sequence and unlock them
24	Remove the upper beam bracket and active cross brace

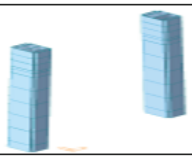
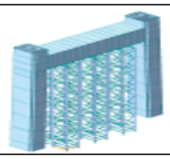
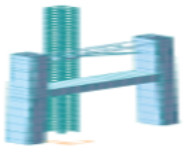
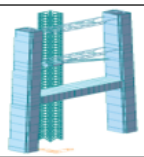
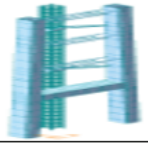
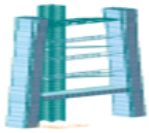
	
Working condition 2: Install T0-T7 section	Working condition 6: Install the lower beam closing section and unlock it.
	
Working condition 9: Active cross brace 1 pushes up the tower column	Working condition 11: Active cross brace 2 pushes up the tower column
	
Working condition 16: Active cross brace 3 pushes up the tower column	Working condition 20: Install the upper beam bracket and push up the tower column
	
Working condition 23: Install the closing section of the upper crossbeam and unlock it.	Working condition 24: Remove the upper crossbeam bracket and active cross brace

Fig. 2. Schematic diagram of the construction stage model.

B. ANSYS finite element model

This paper utilizes large-scale finite element general-purpose computational software, ANSYS, for detailed modelling and analysis, aiming to compare and validate the effectiveness of the simplified model in Midas Civil. Each segment of the steel tower in the literature is modelled using the shell element SHELL181, while the lower beam of the steel tower, three active braces, and upper beam supports are modelled using the beam element BEAM189. Additionally, shell element SHELL181 is used to simulate the external steel shell, vertical partitions, vertical ribs, and local joints of each segment of the steel tower, strictly following the design thickness specified in the drawings.

In the construction stage analysis, ANSYS functionality allows for the simulation of the construction process through techniques such as "killing" and "activating" elements, which are particularly suitable for simulating the step-by-step establishment of structures [12]. When elements are "killed" in the ANSYS program, their stiffness matrix is multiplied by a very small coefficient, typically defaulting to 1e-6. The results of the solution will not include the mass and energy of "dead" elements. The stiffness, mass, and load parameters return to their real state upon reactivation. After elements are "reactivated," there is no history of strain, and they are "annealed" through the process of being both dead and alive, where all stresses and strains are zero when they are alive. The ANSYS finite element model is shown in Fig. 3

Meanwhile, during the construction stage simulation process, the deflection deformation of the tower body mainly originates from its weight. Therefore, it is necessary to compare and verify the weight of each segment of the steel tower calculated from the model with the weights specified in the design drawings. This ensures the accurate application of the steel tower's self-weight load and verifies the accuracy of the model establishment through mass differences. This is summarized in Table 2.

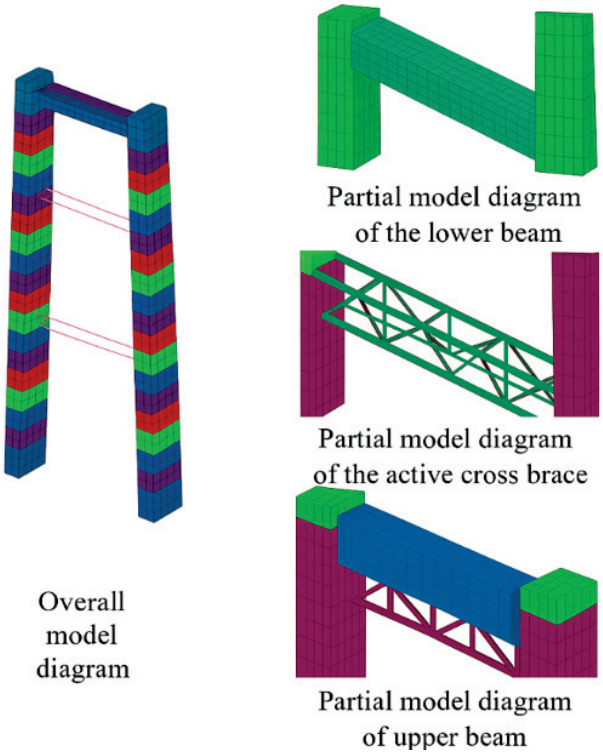


Fig. 3. Schematic diagram of ANSYS finite element model.

Table 2: Model quality error

Segment Number	Model Weight (t)	Design Weight (t)	Error (%)
1	225.7	264.8	14.7
2	202.3	232.6	13
3	202	232.3	13
4	201.7	231.9	13
5	201.4	232	13.2
6	217.4	262	17
7	209.3	253	17.3
8	217.6	262.9	17.2
9	200.3	230	12.9
10	200.1	229.7	12.9
11	207.7	237.2	12.4
12	239.9	280	14.3
13	239.2	278.4	14.1
14	238.6	277.7	14.1
15	238.3	277.1	14
16	238	276.5	13.9
17	237.4	275.9	13.9
18	236.9	275.3	13.9
19	225.6	270.1	16.5
20	238	288.5	17.5
21	225.7	279.6	19.3
22	311.4	332.3	6.3

Table 3: Thrust force and displacement comparison between Midas and ANSYS models

Location	Midas Thrust Force (t)	Midas Displacement (mm)	ANSYS Thrust Force (t)	ANSYS Displacement (mm)
Lower cross-member bracket	23	2	20	2.17
Active cross brace 1	62	4.74	66	5.87
Active cross brace 2	26	3.26	34	4.98
Active cross brace 3	18	2.14	22	3.38
Upper beam bracket	29	5.56	39	7.21
Vertex displacement		453		398

C. Comparison of finite element calculation results

During the construction process of the main tower, three temporary active braces are set up to support the toppling, located at the top push rods of the lower beam supports at +97.95 m (section T11), +131.341 m (section T14), and +164.932 m (section T17), as well as the top push rods of the upper beam supports. The top push force, top push displacement, and vertex displacement of each brace are shown in Table 3. Table 3 shows the largest differences in top push displacement and deformation between the two models in stages 2 and 3 of the active braces. However, the calculation results for the lower beam supports, upper beam supports, and vertex displacement are relatively consistent [13]. Overall, using Midas Civil for construction stage analysis proves to be effective.

IV. ANALYSIS OF STRUCTURAL EFFECTS UNDER DIFFERENT LOADS

A. Analysis of calculation results under self-weight state

1. Main tower strength and stiffness

The maximum stress and lateral deformation variations of the steel tower during the construction stage are depicted in Fig. 4. Figure 4 shows that the maximum combined stress of the steel tower decreases due to the application of the jacking force during tower installation, although the numerical fluctuations are small. However, the stress level shows a gradually increasing trend, indicating overall stability in structural strength. According to the analysis results, the maximum stress in the structure occurs at the base of the steel tower during each construction stage. The minimum stress is reached during the installation of the lower beam supports at 6.4 MPa, while the maximum stress is attained during the closing stage of the upper beam at 21.9 MPa. After removing the active braces and upper beam supports, the stress reaches 21.8 MPa. This analysis highlights the crucial role of jacking operations in maintaining structural stability and underscores the importance of monitoring and adjusting the jacking force [14].

Due to the structure's inclination, there is a tendency for the steel tower to lean inward during installation, which is a result of its weight. Therefore, this analysis focuses more on lateral deformation behavior. The maximum deformation of the steel tower occurs after the upper beam supports are installed, reaching 5.984 mm. The structure returns to its design shape under the influence of various jacking forces before installing the next section, with deformation reduced to 0.54 mm after removing the upper beam supports and active braces. According to the analysis results, the maximum deformation occurs at the top of the steel tower in each construction stage.

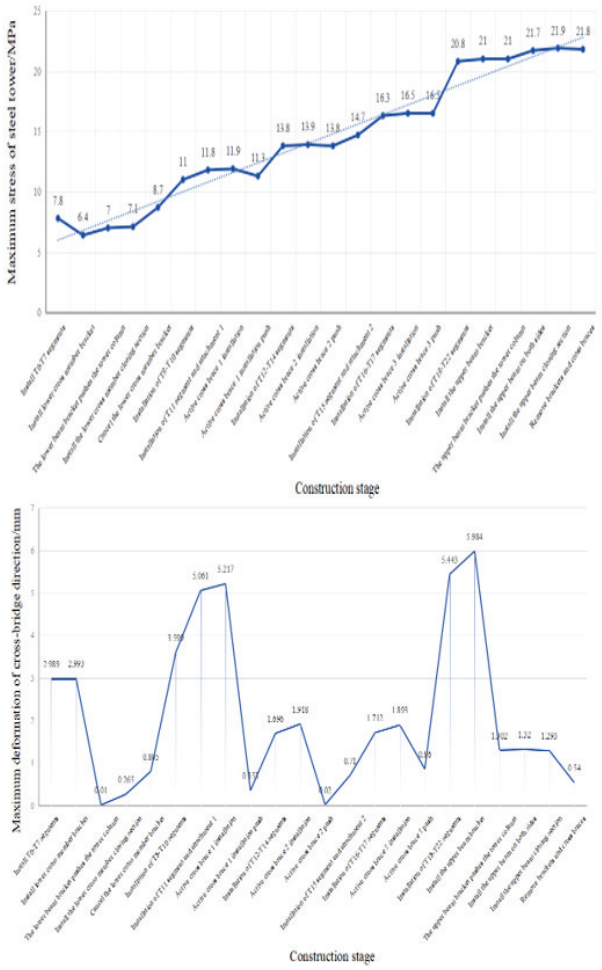


Fig. 4. Maximum stress and deformation development curve of the main tower during (unit: MPa).

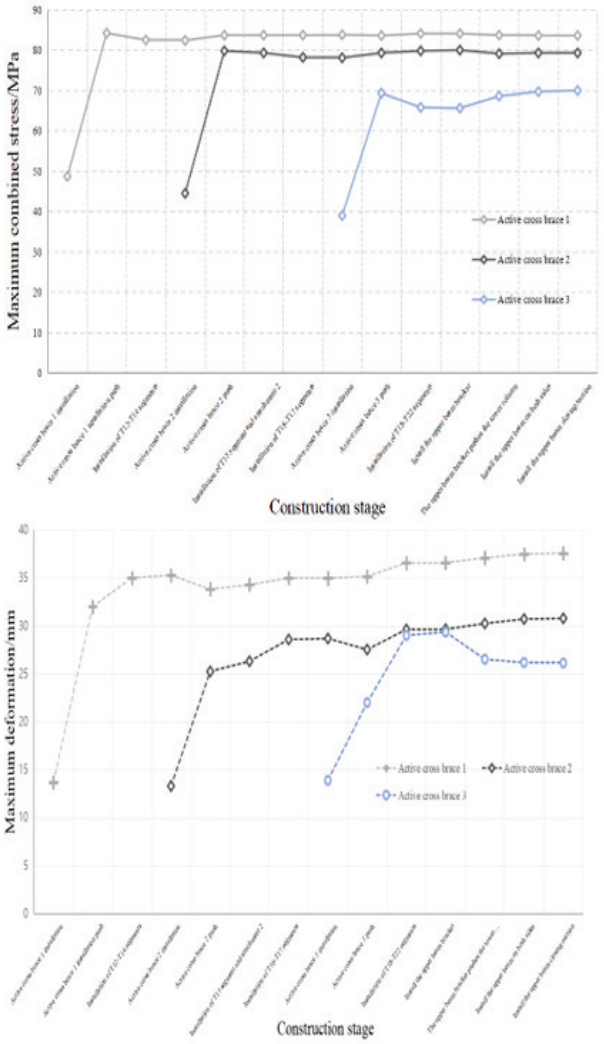


Fig. 5. Maximum stress and deformation development curve during the construction stage (unit: MPa).

2. Active cross brace strength and stiffness

The actual development of the maximum lateral deformation of the steel tower during the construction stage is illustrated in Fig. 5. Due to the structure's inclination, there is a tendency for the steel tower to lean inward during installation, which is a result of its weight. Therefore, this analysis focuses more on lateral deformation behavior. Figure 5 shows that the maximum deformation of the steel tower occurs after the installation of the upper beam supports, reaching 5.984 mm. Under various jacking forces, the structure returns to its design shape before installing the next section, with deformation reduced to 0.54 mm after removing the upper beam supports and active braces. According to the analysis results, the maximum deformation occurs at the top of the steel tower in each construction stage. Additionally, before installing the second active brace, the third active brace, and the upper beam supports, the structure is in a state of the maximum cantilever, with its stiffness at the lowest point, thus being in the most unfavorable condition.

B. Analysis of calculation results under wind load

Due to geographical factors, the steel tower may be subjected to strong winds when not under construction. A wind tunnel test was conducted on self-supporting bridge tower models to determine response patterns under various wind speeds and angles of attack [15]. A formula for estimating permissible oscillation amplitudes of bridge tower bending vortex-induced resonance based on tower height parameters was summarized by [16, 19]. The significant impact of wind angle of attack and average wind speed on the flutter response of self-supporting bridge towers, using a multi-channel coupled vibration frequency calculation method, was emphasized by [17, 20].

This paper analyses the stress under three maximum cantilever conditions based on the calculation results under self-weight conditions. The wind loads on both sides of the steel tower are considered to simultaneously bear the maximum wind loads in the longitudinal and transverse directions. Here, wind loads are not applied to the active braces, and the analysis examines the degree to which the steel tower is affected by wind loads on the active braces. The tower crane considers the loads on the crane head and wind loads on the tower body under non-working conditions, taking into account loading directions in the longitudinal, transverse, and attachment directions. According to the design wind resistance standards for cable-stayed towers, the return period for wind speed during construction is 30 years.

According to the "Code for Design of Highway Bridges for Wind Resistance" [18, 21], the bridge falls under the R2 region, with a ground category of A. Located in Nantong, the wind speed is U10 = 28.5 m/s. For calculations, the heights Z of active braces 1, 2, and 3 are 94.95 m, 128.538 m, and 161.932 m, respectively. The wind load on the steel tower is then calculated based on height variation from the tower base to the top. The wind load calculation formula is as follows:

$$U_{s10} = U_{10} \times k_e, \tag{1}$$

$$U_{s10} = U_{10} \times k_e, \tag{2}$$

$$U_d = k_f k_t k_h U_{10}, \tag{3}$$

$$U_d = k_f \left(\frac{Z}{10}\right)^{\alpha 0} U_{s10}, \tag{4}$$

$$U_g = G_v U_{sd}, \tag{5}$$

$$F_g = 0.5 \rho U_d^2 C_D A_n, \tag{6}$$

where U_{10} represents the basic wind speed; k_e is the terrain conversion coefficient, taken as 1.174; U_d represents the design basic wind speed, taken as the maximum of the two; $k_f$is the wind risk coefficient, taken as 1.02; k_t represents the terrain condition coefficient, taken as 1; k_h is the wind pressure height correction coefficient, taken by linear interpolation; $\alpha 0$ represents the surface roughness coefficient, taken as 0.12; G_v is the equivalent static coefficient, obtained by linear interpolation; ρ represents air density, taken as 1.25; C_D is the component drag coefficient; A_n represents the projected area of the component.

1. Maximum cantilever state of the first cross brace

Figure 6 shows the maximum total deformation and maximum combined stress of the steel tower and active brace 1 in the first cantilever condition. From Fig. 6, it can be observed that wind loads have a significant impact on the stiffness of the steel tower and active brace 1. Compared to the case without wind loads, the maximum deformation of the steel tower under transverse wind loads reaches 108.425 mm. In contrast, the maximum deformation of the active brace occurs under wind loads in the attachment direction, reaching 77.569 mm. Compared to changes in stiffness, the changes in strength for both the steel tower and active brace 1 are relatively small. Under transverse wind loads, the maximum combined stress in the steel tower reaches 42.1 MPa, while the maximum stress in the active brace occurs in the attachment direction, reaching 86.5 MPa.

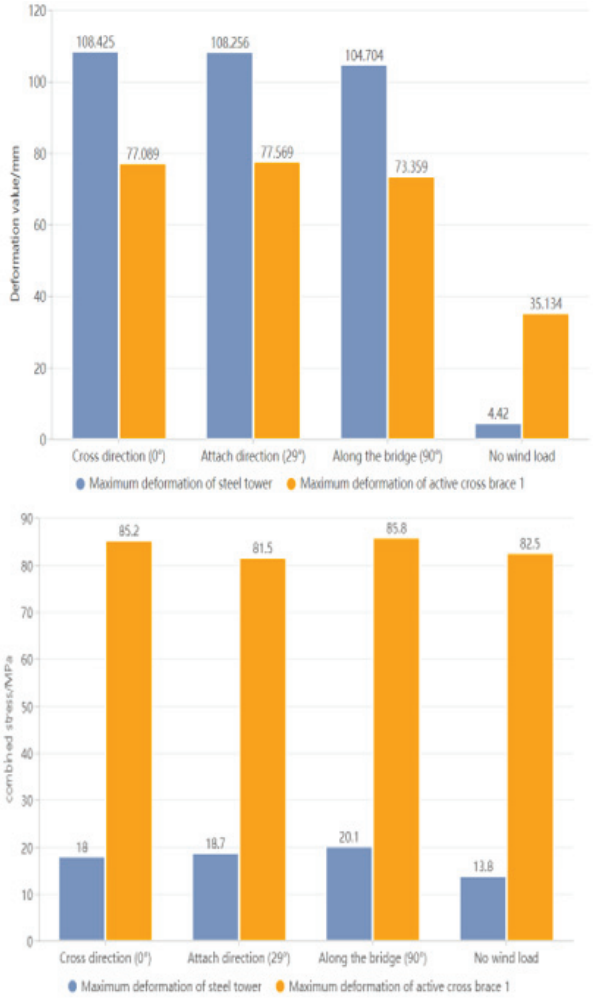


Fig. 6. Maximum stress and deformation development curves of wind load in different directions.

2. Maximum cantilever state of the second cross brace

In the second cantilever condition, the maximum deformation and maximum combined stress of the steel tower, active brace 1, and active brace 2 are shown in Fig. 7.

Compared to the first cantilever condition, it can be observed from Fig. 7 that as the height of the steel tower increases, the influence of wind loads on the stiffness of the steel tower also increases. Under transverse wind loads, the maximum deformation of the steel tower

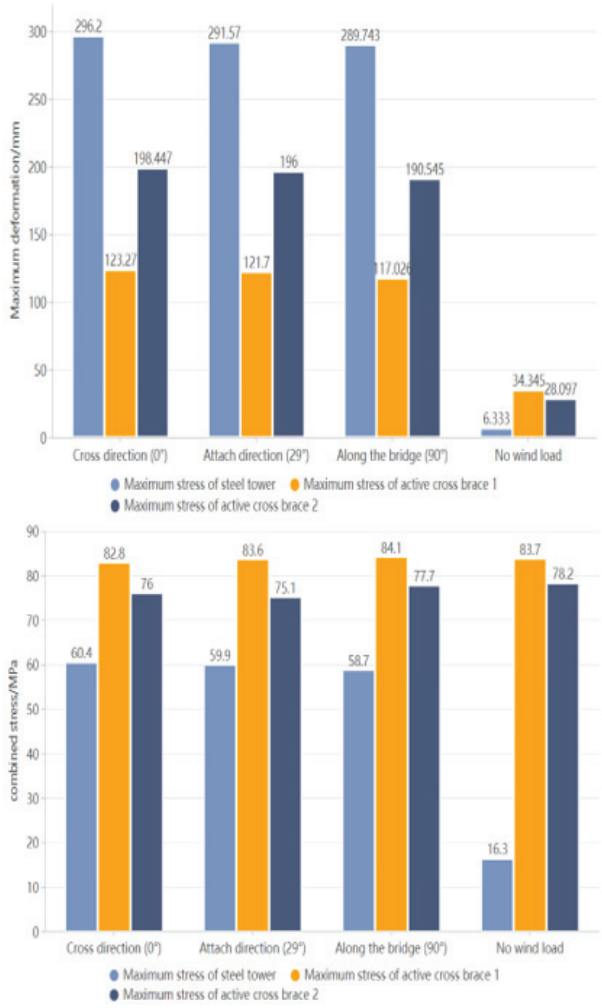


Fig. 7. Maximum stress and deformation development curves of wind load in different directions.

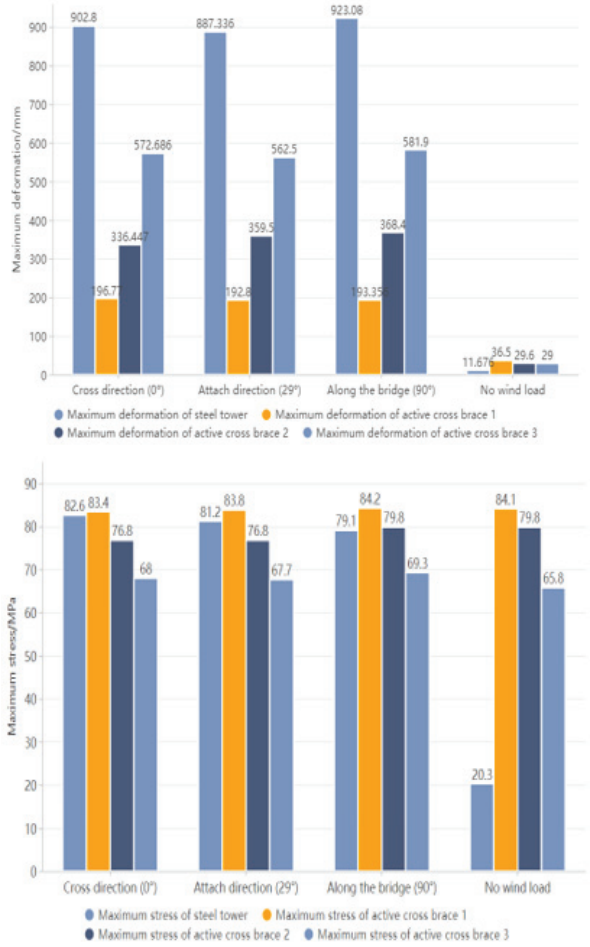


Fig. 8. Maximum stress and deformation development curves of wind load in different directions.

reaches 296.2 mm, the maximum deformation of active brace 1 is 123.27 mm, and the maximum deformation of active brace 2 is 198.447 mm. Despite changes in stiffness, the structural changes in strength remain relatively small. Under transverse wind loads, the maximum combined stress of the steel tower reaches 60.4 MPa. The maximum stress of active brace 1 occurs under wind loads in the attachment direction, reaching 83.6 MPa. The maximum stress of active brace 2 occurs under wind loads in the longitudinal direction, reaching 77.7 MPa.

3. Maximum cantilever state of the third cross brace

In the third cantilever condition, the maximum total deformation and maximum combined stress of the steel tower, active brace 1, active brace 2, and active brace 3 are shown in Fig. 8.

From Fig. 8, it can be seen that, due to the longer cantilever length, the influence of wind loads on the stiffness of the steel tower under this condition is more severe. Compared to the previous conditions, under longitudinal wind loads, the maximum deformation of the steel tower reaches 923.08 mm. The maximum deformation of active brace 1 under transverse wind loads is 196.77 mm. The maximum deformation of active brace 2 under longitudinal wind loads is 368.4 mm. The maximum deformation of active brace 3 reaches 581.9 mm. Despite changes in stiffness, the overall structural changes in strength remain relatively small. Under transverse wind loads, the maximum combined stress of the steel tower reaches 82.6 MPa; the maximum stress of active brace 1 under longitudinal wind loads reaches 86.5 MPa; and the maximum stresses of active brace 2 and active brace 3 reach 79.8 MPa and 69.3 MPa, respectively. According to the calculation results, the structure's overall strength and stiffness meet the requirements specified.

C. Analysis of calculation results under tower crane load

1. Maximum cantilever state of the first cross brace

During installation of the steel tower, it is necessary to consider the impact of the tower crane load on the structure in different directions. Figure 9 shows the maximum total deformation and maximum combined stress of the steel tower and active brace 1 in the first cantilever condition. Figure 9 shows that, compared to wind loads, the influence of the tower crane load on the strength and stiffness of the steel tower and active brace 1 is relatively small. Under longitudinal tower crane load, the maximum stress of the steel tower reaches 20.1 MPa with a deformation of 26.3 mm, while active brace 1 reaches 85.8 MPa with a deformation of 43 mm.

2. Maximum cantilever state of the second cross brace

Figure 10 shows the maximum total deformation and maximum combined stress of the steel tower, active brace 1, and active brace 2 in the second cantilever condition. From Fig. 10, it can be observed that, under the longitudinal tower crane load, the maximum stress in the steel tower reaches 21.4 MPa, accompanied by a maximum deformation of 64.6 mm in the attachment direction. Active brace 1 reaches a maximum stress of 85.4 MPa under both longitudinal and transverse tower crane loads, with a maximum deformation of 41 mm in the longitudinal direction. Active brace 2 reaches a maximum stress of 84.1 MPa under transverse tower crane load, with a maximum deformation of 53 mm in the longitudinal direction.

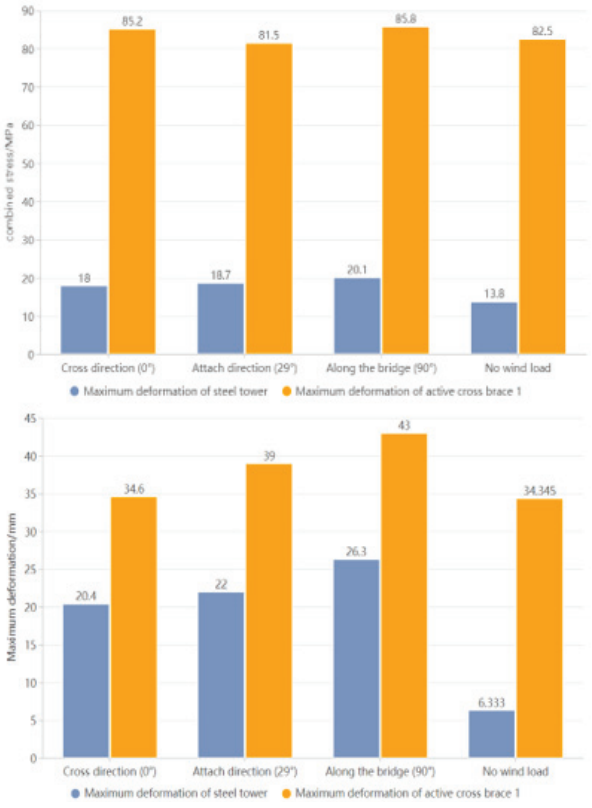


Fig. 9. Maximum stress and deformation development curve of tower crane load in different directions.

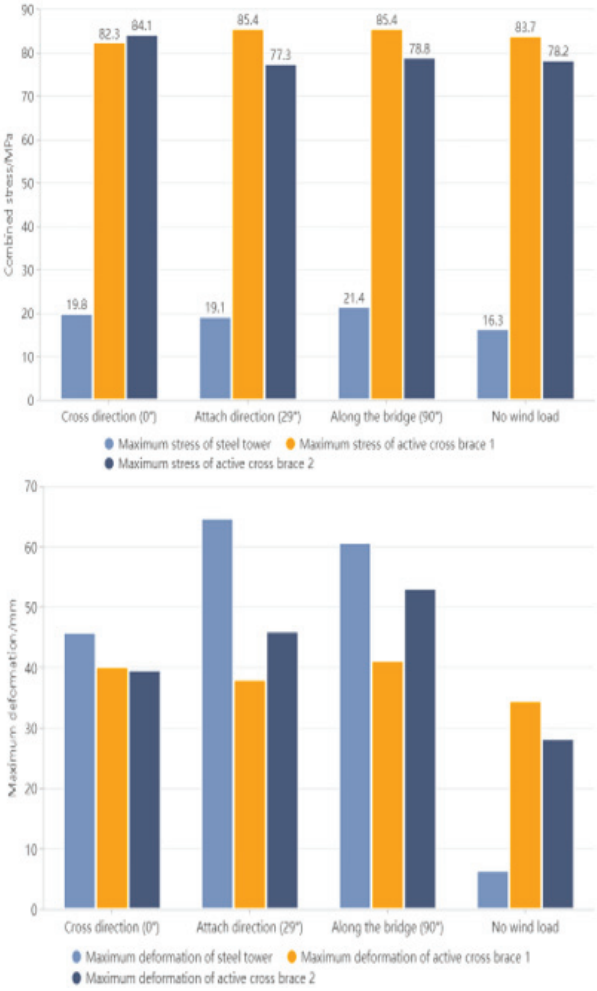


Fig. 10. Maximum stress and deformation development curve of tower crane load in different directions.

3. Maximum cantilever state of the third cross brace

In the third cantilever condition, the maximum total deformation and maximum combined stress of the steel tower, active brace 1, active brace 2, and active brace 3 are shown in Fig. 11. From Fig. 11, it can be observed that, under the longitudinal tower crane load, the maximum stress in the steel tower reaches 26.1 MPa, accompanied by a maximum deformation of 98 mm in the attachment direction. Active brace 1 reaches a maximum stress of 85.4 MPa with a maximum deformation of 43 mm in the longitudinal direction. Active brace 2 reaches a maximum stress of 82.5 MPa under transverse tower

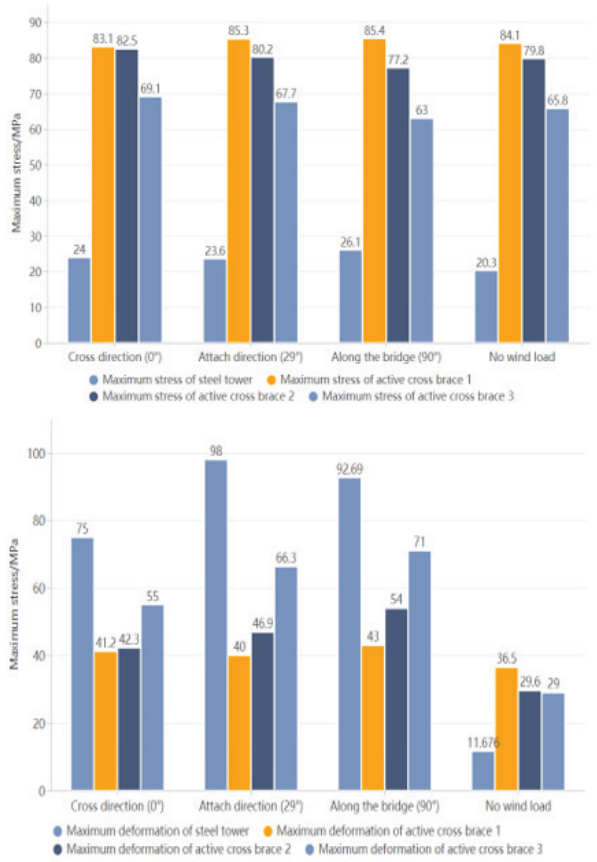


Fig. 11. Maximum stress and deformation development curve of tower crane load in different directions.

crane load, with a maximum deformation of 54 mm in the longitudinal direction. Active brace 3 reaches a maximum stress of 69.1 MPa under a transverse tower crane load, with a maximum deformation of 71 mm in the longitudinal direction.

V. CONCLUSION

This research examines two finite element modelling methods, Midas Civil and ANSYS, employed in assessing steel towers throughout the construction phases of suspension bridges. The results show that each software tool has unique but supportive functions in structural analysis. Midas Civil is especially useful for initial evaluations because of its intuitive interface and effective modelling features. It enables engineers to quickly model construction scenarios and assess the overall performance of the steel tower under various conditions. The results acquired from Midas Civil have been validated through comparison with those from ANSYS, demonstrating its reliability for initial design phases.

ANSYS provides a more detailed and comprehensive analysis, capturing complex interactions and localized impacts that may be vital for preserving structural integrity. Its sophisticated modelling features enable a thorough examination of stress distribution and deformation, rendering it a crucial instrument for ultimate validation and adherence to safety regulations. Nonetheless, it is essential to recognize the constraints of every method. Midas Civil, although efficient, might overlook some complex aspects of structural behavior, which could result in oversimplifications. While ANSYS offers greater accuracy, it requires considerable computational power and expertise, which may not be practical for every engineering team. Overall, integrating Midas Civil and ANSYS in the analysis process enhances understanding of steel tower dynamics during construction. Utilizing Midas Civil for initial assessments, followed by detailed validation with ANSYS, can lead to improved safety, reliability, and efficiency in the design and construction of suspension bridges. This dual approach not only aids in addressing the complexities of steel tower construction but also contributes valuable insights for future engineering practices.

Future research will enhance the analysis of steel towers in suspension bridge construction by implementing advanced modelling techniques, such as nonlinear analysis and dynamic simulations, to better understand their behavior under various loads. It will also develop real-time structural health monitoring systems to validate model predictions and assess performance. The study will be expanded to include comparative analyses of additional finite element software tools and investigate the impact of new materials, such as high-strength steel and composites, on structural performance.

REFERENCES

[1] AASHTO, "AASHTO LRFD Bridge Design Specifications," 8th ed., AASHTO, Washington, DC, 2017.

[2] T. L. L. Orr, "Selection of characteristic values and partial factors in geotechnical designs to Eurocode 7," *Computers and Geotechnics*, vol. 26, no. 3, pp. 263-279, Apr. 2000.

[3] W. Xiong, R. Wang, T. Wang, Y. Liu, and J. Zhang, "Design method for anti-skid safety of gravity anchor of suspension bridges with a span of over 2000 meters," *Chinese Journal of Highways*, vol. 37, no. 4, pp. 166-175, Apr. 2024.

[4] X. Q. Zhao, X. N. Gong, and P. P. Guo, "Caisson-bored pile composite anchorage foundation for long-span suspension bridge: Feasibility study and parametric analysis," *Journal of Bridge Engineering*, vol. 27, no. 12, p. 04022117, Oct. 2022.

[5] L. Ye, H. Ke, and Z. Chen, "Analysis of the construction stage of steel box main beam ground anchored suspension bridges based on

Midas/Civil," *Zhongwai Highway*, vol. 40, no. 6, pp. 140-145, 2020.
[6] S. Tan, G. Xiong, and Y. Yang, "Research on monitoring indicators for construction of large span suspension bridge towers," *China Municipal Engineering*, no. 3, pp. 55-58, 138, 2022.
[7] C. Lu, W. Wang, and Y. Liao, "Numerical simulation analysis of the construction process of a large-span suspension bridge tower," *Building Structure*, vol. 48, no. S2, pp. 959-962, 2018.
[8] J. Lv, H. Kang, and R. Wang, "Nonlinear parametric vibration of the main cable of a suspension bridge under longitudinal action of bridge towers," *J. Hunan Univ. Sci. Technol. (Natural Science Edition),* vol. 32, no. 3, pp. 41-46, July 2017.
[9] D. Zhu, X. Luo, and C. Zuo, "Performance analysis of the concrete main tower structure of an in-service suspension bridge based on ANSYS," *Highway and Motor Transport*, no. 2, pp. 112-118, 2021.
[10] X. Zhao, W. Zhan, X. Yan, and M. Liu, "Current status and future development prospects of suspension bridge anchorage research," *Journal of Geotechnical Engineering*, vol. 43, no. 2, pp. 150-153, 2021.
[11] D. P. Deevi, "Artificial neural network enhanced real-time simulation of electric traction systems incorporating electro-thermal inverter models and FEA," *Int. J. Eng. Sci. Res*., vol. 10, no. 3, pp. 36-48, 2020.
[12] S. Kodadi, "High-performance cloud computing and data analysis methods in the development of earthquake emergency command infrastructures," *Journal of Current Science*, vol. 10, no. 3, pp. 87-105, 2022.
[13] T. V. Dharma, "Optimizing cloud computing environments for big data processing," *Int. J. Eng. Sci. Res*., vol. 13, no. 2, pp. 1756-1775, 2023.
[14] Y. Niu, "Construction technology for positioning and installation of curved section steel towers in self-anchored suspension bridges," *Northern Transportation*, no. 2, pp. 11-15, 2023.
[15] L. Sun, Q. Wang, and Y. Zhang, "Design and local stress analysis of steel tower suspension wall for self-anchored suspension bridge," *Highway*, vol. 66, no. 1, pp. 171-175, 2021.
[16] X. Qu and C. Qian, "Multi-point constraint (MPC) method and finite element analysis of heat exchangers," *Pressure Vessel*, vol. 30, no. 2, pp. 54-58, 2013.
[17] F. Cao, Z. Liu, and H. Huang, "Innovative design of A-shaped bridge tower for single main cable suspension bridge," *Urban Road and Bridge and Flood Control*, no. 4, pp. 15, 75-77, 85, 2023.
[18] L. Zeng, R. Ma, and A. Chen, "Experimental study on wind resistance performance of self-supporting tower of large-span suspension bridge," *Hunan Transportation Technology*, vol. 49, no. 1, pp. 75-79, 2023.
[19] J. Li, G. Zhang, and B. Zhang, "Determination of allowable amplitude for the vortex-induced resonance of steel bridge towers," *Chinese Journal of Highways*, vol. 27, no. 3, pp. 45-50, 2014.
[20] J. Li, Z. Shen, S. Xing, and W. Tang, "Buffeting response of a free-standing bridge pylon in a trumpet-shaped mountain pass," *Wind & Structures*, vol. 30, no. 1, pp. 85-97, Jan. 2020.
[21] Ministry of Transport of the People's Republic of China, "Wind Resistant Design Code for Highway Bridges," JTG/T D60-01-2004, China Communications Press, Beijing, China, 2004.

Haodong Wang was born in Dazhou, Sichuan, P.R. China, in 1996. He received his master's degree from Qinghai University, P.R. China, in 2022. Currently, he works at the Technical Center of Sichuan Road and Bridge East China Construction Co., Ltd. His research interests include road and bridge construction and design, as well as seismic design of structures.

Liu Jie was born in Chengdu, Sichuan Province, China. He obtained his bachelor's degree in 2009 from Jincheng College, Sichuan University. Currently, he is a project manager at Sichuan Road and Bridge Group. His research focuses on large-scale bridge construction, and he has participated in the construction of several world-class bridges.

Fan Ji was born in Nantong, Jiangsu Province, China, in 1997. He received his master's degree from Southwest Jiaotong University in 2022. He is currently working at the Technical Center of Sichuan Highway and Bridge Construction Group Co. Ltd., Bridge Engineering Branch. His specialization is in the design and construction of underground railways and mountain tunnels.

Yongkai Zhou was born in Chengdu, Sichuan Province, China. He obtained his bachelor's degree from Xihua University in 2009. Currently, he works as the Deputy Director at the Technical Center of Sichuan Road and Bridge East China Construction Co., Ltd. His research focuses on large-scale bridge construction and design.

Liu Liang was born in Tianmen City, Hubei Province, China. He obtained his master's degree from Chang'an University in 2013. Currently, he serves as the Deputy Chief Engineer of the Bridge Engineering Branch at Sichuan Highway and Bridge Construction Group Co. Ltd. His research focuses on bridge engineering construction and management.