Enhancing Predictive Accuracy for Residual Tensile Strength of GFRP Bar in Coastal Environments: An Integrated Bagging-Stacking Ensemble Model

M. M. Rahman*, Go Chaeyeon*, Kwag Shinyoung**
*Master's course, Dept. of Civil Engineering, Hanbat National University
**Associate Professor, Dept. of Civil and Environmental Engineering, Hanbat National University
e-mail: skwag@hanbat.ac.kr

해안 환경에서 GFRP 바의 잔류 인장 강도에 대한 예측 정확도 향상: 통합 배깅-스태킹 앙상블 모델

M. M. Rahman*, 고채연*, 곽신영**
*한밭대학교 토목공학과 석사과정
**한밭대학교 건설환경공학과 부교수

Abstract

Glass-fiber reinforced polymer (GFRP) bars have emerged as a viable alternative to traditional rebar in coastal environments, offering advantages such as high strength and corrosion resistance. Nevertheless, their tensile strength can deteriorate significantly in harsh conditions characterized by strong alkalinity, high salinity, and humidity. Accurate prediction of GFRP tensile strength under various degradation conditions is challenging. To address this, we developed a machine learning model that combines ensemble techniques for precise residual tensile strength prediction of GFRP. Our study employed a dataset of 379 data points, divided into training and test sets. We evaluated the performance of models obtained from individual and ensemble machine learning methods, comparing them with bootstraps vs sampling ratio and four performance indexes.

1. Introduction

Conventional steel-reinforced concrete structures within marine environments often require frequent maintenance due to steel corrosion. To overcome this problem, corrosion resistance, lightweight, durability, and high strength of GFRP bars drive their demand in offshore structures, but alkaline conditions adversely affect their durability [1].

The durability of GFRP rebars in severe environmental conditions remained a rising interest of a wide range of researchers [2],[3],[4]. Furthermore, Iqbal et al. [5] and Iqbal et al. [6] observed that larger-diameter GFRP rebars exhibited superior tensile strength (TS) compared to smaller ones under similar environmental conditions. Another study [7] discussed the influence of temperature and fiber volume fraction (Vf) on GFRP bar TS, noting Vf's negligible impact at 11-25°C but significant effect at higher temperatures. Bars with Vf $\geq$ 60% or $<$ 50% affected GFRP TS, with Vf influencing water absorption and, consequently, TS degradation with increasing temperature.

Recently, Kabiru et al. [8] and Jeyasehar & Sumangala [9] utilized machine learning models to create machine learning models that predict TSR, such as the artificial neural network (ANN), genetic algorithms (GAs), genetic programming (GP), and genetic expression programming (GEP), and the adaptive neuro-fuzzy inference system (ANFIS) method was used. Afterward, Go et al. [10] proposed a combined bagging and stacking ensemble (CBSE) model to improve model accuracy. However, when determining the number of bootstraps in a prediction model, Go et al. [10] did not provide a clear basis for choosing a specific number of bootstraps or a sampling rate that was most appropriate for the data set, but used an arbitrary selection. Therefore, in this study, we identify a sampling ratio and the number of bootstraps for both of training and testing data.

2. Data and Methodology

This study utilizes 379 TSR data points from several literature sources, each representing GFRP bar TSR within a specific range of input variables: bar diameter, volume fraction, pH, environmental temperature, and conditioning duration. Table 1 provides the maximum values (Max), minimum values (Min), mean values (Mean), and standard deviations (SD) for both the input and output variable values in the study. In this paper, 80 % of the data is allocated for training, and the other 20 % of the data is used for testing.

[Table 1] Analysis of Statistical Characteristics for Input and Output Variables in GFRP Bar TSR Dataset.

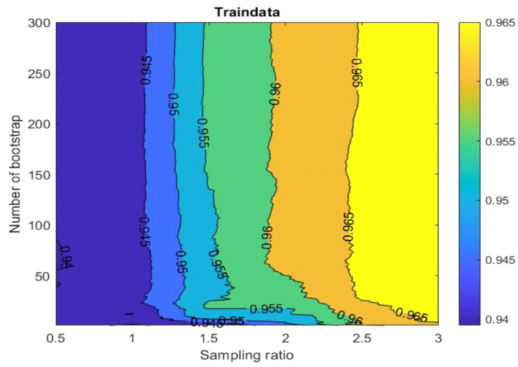
Variables	Max	Min	Mean	SD
Dia (mm)	16.00	6.00	11.81	2.20
vf (%)	0.84	0.45	0.58	0.09
pH	13.60	12.00	12.86	0.48
Temp (℃)	80.00	11.00	41.48	18.03
Time (days)	540.00	17.00	149.17	118.79
TSR (%)	102.00	22.50	82.57	14.54

The model introduces a novel approach, CBSE, incorporating both bagging and stacking techniques to enhance the predictive accuracy of the GFRP bar's TSR model. This method basically uses a meta-model, that needs the training data containing predicted values such as nonlinear regression (NLR), support vector machine (SVM), artificial neural network (ANN), Gaussian process regression (GPR), and bagging method, and a meta-model utilizes the GPR method to train. From this meta-model, the final predicted value is derived.

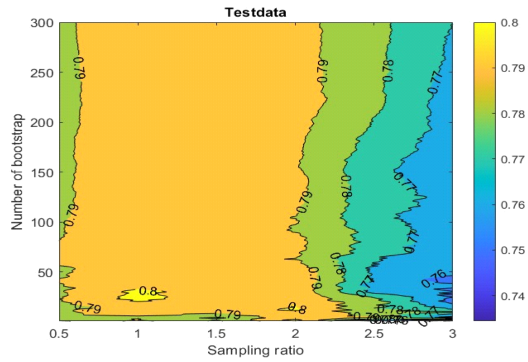
3. Result and discussion

In this study, the evaluation utilizes four performance metrics: R^2 (Coefficient of determination), Root Mean Squared Error (RMSE), VAF (Variance Account For), and MAPE (Mean Absolute Percentage Error), demonstrating the model's accuracy. Specifically, the R^2 of the CBSE model was analyzed according to the number of bootstraps from a minimum of 1 to a maximum of 300 and the ratio of sampling with replacement from a minimum of 0.5 to a maximum of 3.0 (See Fig.2 and Fig.3). Regarding the training data, Fig.2 shows that the greater the number of bootstraps and the ratio of sampling with replacement, the higher the accuracy R^2 value of the training model. With respect to the testing data, Fig.3 shows that the highest performance is shown when the number of bootstraps is 25 and the ratio of sampling with replacement is 1.0. Based on these results, it can be seen that as the number of bootstraps and the ratio of sampling with replacement increases, the meta-model tends to overfit.

To confirm its efficacy, the generated model having the best estimates of bootstrap number and ratio of sampling with replacement identified from the above stage is compared with others using identical training and testing data. Table 2 presents an overview of the prediction accuracy of the generated models, revealing that the model identified from this study exhibits the highest performance.



[Fig. 2] Response Contour Graph for training data.



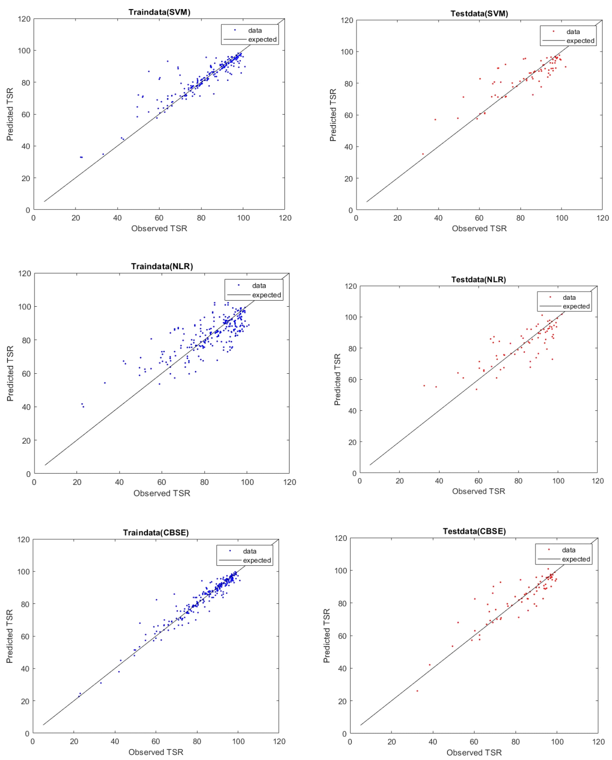
[Fig. 3] Response Contour Graph for testing data.

[Table 2] Summary of performance index for each model's training and testing dataset.

Model	Phase	R^2	RMSE	VAF	MAPE
ANN	Training	0.7169	5.0623	87.6517	4.9867
	Testing	0.7169	7.9650	71.1095	7.9142
SVM	Training	0.8585	5.4194	86.0755	4.6162
	Testing	0.7525	7.7820	73.3699	7.289
GPR	Training	0.8997	3.5818	93.8179	3.1969
	Testing	0.7737	6.7997	79.5591	6.2624
NLR	Training	0.6219	8.8577	62.1916	9.5171
	Testing	0.6606	8.7209	66.2262	9.006
Bagging (GPR)	Training	0.9382	3.5807	93.8217	3.2204
	Testing	0.7744	7.1103	77.4407	6.7574
Stacking (GPR)	Training	0.9400	3.5274	94.0042	3.0993
	Testing	0.7916	6.8334	79.3963	6.1954

CBSE (GPR) with Best estimates	Training	0.9452	3.3736	94.5156	2.9153
	Testing	0.8030	6.7792	79.5507	6.1731

Furthermore, below are visualization plots (Fig.4) showing a correlation between observed and predicted TSR values for both the training and testing datasets of the SVM, NLR, and CBSE models. These plots reveal that the highest accuracy is this study's model.



[Fig. 4] Demonstration of regression plots depicting the relationship between observed and predicted TSR values using several models: SVM, NLR, CBSE.

4. Conclusion

Upon reviewing the accuracy outputs of all the models, the numerical results can be summarized as follows.

(a) The proposed method of this study (CBSE model with the best estimates) obtained a model with higher accuracy (R^2 = 0.9452 for training and 0.8030 for testing) than individual ensemble models and the single GPR model.

(b) Among the individual machine learning models, the GPR model exhibited superior accuracy with R^2 values of 0.8997 for training and 0.7737 for testing.

Finally, this study's model outperforms single machine learning methods by attaining more accuracy by averaging of results from twenty different output sets with 25 bootstraps and a 1.0 ratio of sampling with replacement.

References

[1] Iqbal, M., Zhang, D., Jalal, F. E., & Javed, M. F. (2021). Computational AI prediction models for residual tensile strength of GFRP bars aged in the alkaline concrete environment. Ocean Engineering, 232, 109134.

[2] Robert, M., Cousin, P., & Benmokrane, B. (2009). Durability of GFRP reinforcing bars embedded in moist concrete. Journal of Composites for Construction, 13(2), 66-73.

[3] Robert, M., Cousin, P., & Benmokrane, B. (2009). Durability of GFRP reinforcing bars embedded in moist concrete. Journal of Composites for Construction, 13(2), 66-73.

[4] Wang, Z., Zhao, X. L., Xian, G., Wu, G., Raman, R. S., Al-Saadi, S., & Haque, A. (2017). Long-term durability of basalt-and glass-fibre reinforced polymer (BFRP/GFRP) bars in seawater and sea sand concrete environment. Construction and Building Materials, 139, 467-489.

[5] Iqbal, M., Zhang, D., Khan, M. I., Zahid, M., & Jalal, F. E. (2023). Effects of Rebar Size and Volume Fraction of Glass Fibers on Tensile Strength Retention of GFRP Rebars in Alkaline Environment via RSM and SHAP Analyses. Journal of Materials in Civil Engineering, 35(9), 04023318.

[6] Iqbal, M., Zhang, D., & Jalal, F. E. (2022). Durability evaluation of GFRP rebars in harsh alkaline environment using optimized tree-based random forest model. Journal of Ocean Engineering and Science, 7(6), 596-606.

[7] Duo, Y., Liu, X., Liu, Y., Tafsirojjaman, T., & Sabbrojjaman, M. (2021). Environmental impact on the durability of FRP reinforcing bars. Journal of Building Engineering, 43, 102909.

[8] Kabiru, O. A., Owolabi, T. O., Ssennoga, T., & Olatunji, S. O. (2014). Performance comparison of SVM and ANN in predicting compressive strength of concrete.

[9] Jeyasehar, C. A., & Sumangala, K. (2006). Damage assessment of prestressed concrete beams using artificial neural network (ANN) approach. Computers & structures, 84(26-27), 1709-1718.

[10] Go, C., Kwak, Y. J., Kwag, S., Eem, S., Lee, S., & Ju, B. S. (2023). On developing accurate prediction models for residual tensile strength of GFRP bars under alkaline-concrete environment using a combined ensemble machine learning methods. Case Studies in Construction Materials, 18, e02157.

Acknowledgment

This study was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (No. RS-2022-00144328).