A Structure-Informed DNN Model for Welding Flux Viscosity Prediction

Developing a hierarchical deep learning approach to integrate composition-structure-viscosity relationships for enhanced predictive performance

BY H. LIU, Y. ZHANG, X. SONG, Z. LI, AND C. WANG

Abstract

Viscosity plays a pivotal role in the performance of welding fluxes by directly affecting the weld pool and the resultant weld metal, both of which are largely determined by heat and mass transfer during the welding process. Accurate viscosity prediction is essential for developing high-quality welding fluxes. However, existing models struggle to capture the complex nonlinear relationships between viscosity, composition, and temperature, which may significantly compromise robustness and generalization. To address this challenge, we incorporate NBO/Si (the average number of non-bridging oxygens per network former Si^{4+}) as an intrinsic structural descriptor and propose a structure-informed deep neural network (S-DNN) that hierarchically models the composition-structure-viscosity relationship, effectively capturing nonlinear viscosity features. The S-DNN demonstrates excellent predictive performance (coefficient of determination [R^2] = 0.9053, mean absolute error [MAE] = 0.0633 Pa·s, and root mean square error [RMSE] = 0.1101 Pa·s). Cross-validation employing multiple experimental datasets further validates the generalization and engineering applicability. Incorporating structural descriptors improves model performance by providing physical constraints. Compared to a basic deep neural network that relies solely on composition and temperature, S-DNN offers significantly higher accuracy. This work establishes a reliable, data-driven framework for the intelligent design of welding fluxes.

https://doi.org/10.29391/2026.105.019

Keywords

- Welding Flux
- Viscosity
- Deep Neural Network (DNN)
- Prediction Model
- Generalization Capability

Introduction

Due to superior strength, toughness, and weldability, high-strength low-alloy (HSLA) steels are widely used in shipbuilding, offshore engineering, and oil and gas sectors (Refs. 1–3). Submerged arc welding, electro-gas welding, and electroslag welding are the primary techniques for HSLA steel fabrication, offering advantages such as high deposition efficiency, cost effectiveness, and process stability (Refs. 4–7). However, an overly high heat input can result in grain coarsening, which downgrades the mechanical properties of the weld metal (WM), posing a serious threat to structural reliability (Ref. 8). Notably, welding fluxes play a decisive role in such welding processes and are regarded as a key enabler for improving the overall WM performance (Refs. 9, 10). It has been proven that fluxes can alter alloying element transfer, stabilize the arc, dictate weld bead morphology, and improve thermal efficiency, all of which are closely linked to flux physicochemical properties, particularly viscosity (Refs. 11–13).

A substantial body of literature has demonstrated that the viscosity of welding fluxes governs mass and heat transfer during welding, directly impacting formability of the WM (Refs. 14, 15). On the one hand, undesirably high viscosity can hinder the diffusion of certain elements and heat sources, leading to unwanted welds with deep penetration but narrow width, which increases the risk posed by inherent pores and trapped slags (Refs. 16, 17). On the other hand, insufficient viscosity could fail to protect the weld pool and maintain arc

stability, potentially resulting in arc blowout and severe safety issues (Ref. 18). Therefore, fine-tuning viscosity is essential for the design and optimization of welding fluxes, which further determines ultimate weld performances (Ref. 19). However, it is widely acknowledged that viscosity is intrinsically governed by molten flux structures (Ref. 20). A consensus has been reached in silicate melts that the relationship between viscosity and various structural parameters is commonly quantified by the degree of polymerization (DOP), typically expressed by the parameter NBO/Si (the average number of non-bridging oxygens per network former Si^{4+}) (Refs. 21, 22). A decrease in NBO/Si could increase the structural hindrance to ionic migration, thereby increasing viscosity. These findings indicate that NBO/Si serves as an intrinsic structural parameter for evaluating viscosity variations in silicate melts, providing a valuable approach for predicting viscosity. In addition, researchers have incorporated the contribution of non-bridging oxygens around network formers and developed an empirical triangular model that could improve the predictive accuracy of the viscosity for certain slags, including blast furnace slags and mold fluxes (Refs. 23, 24). However, the application of such empirical models is limited due to factors such as composition, particularly in systems containing fluorides.

In recent years, machine learning (ML) has emerged as a powerful tool in materials design due to its salient advantages in solving complex nonlinear problems (Ref. 25) and in providing valuable insights into viscosity prediction of slags (Refs. 26, 27). Nevertheless, most existing viscosity models rely on end-to-end frameworks that employ only composition and temperature as inputs, while treating the intrinsic relationship between slag structure and viscosity as a black box. Such a strategy constrains the ability to capture complex nonlinear correlations, thereby limiting robustness, generalization, and predictive accuracy (Ref. 28). For example, Jiang et al. (Ref. 29) developed a viscosity prediction model for blast furnace slags using principal component analysis and k-nearest neighbor regression. The prediction accuracy on industrial data is 83%, but since the model has been trained on the industrial dataset, it limits the generalization ability of the model. Similarly, Yan et al. (Ref. 30) developed viscosity and melting temperature prediction models for mold fluxes using four algorithms and identified that gradient boosting regression trees yielded the most optimal performance. However, the model relied heavily on limited datasets and lacked validation of generalization. Given that NBO/Si provides intrinsic structural descriptors of viscosity, effective integration into predictive models is expected to enhance the representation of fundamental viscosity-structure relationships, thereby enabling the framework to better capture nonlinear viscosity features and improve generalization performance. However, accurately quantifying NBO/Si typically requires advanced experimental techniques, such as Raman spectroscopy and magic-angle-spinning nuclear magnetic resonance, which further undermine the engineering applicability of such models.

In this regard, the present data-driven study introduces a structure-informed deep neural network (S-DNN) designed to hierarchically model composition-structure-viscosity relationships. By incorporating NBO/Si as an intrinsic structural descriptor, the framework establishes a quantitative, interpretable link between melt structure and macroscopic properties, thereby improving the robustness of model generalization and its engineering relevance. The core process involves first predicting the NBO/Si value, then integrating it with composition and temperature inputs to predict viscosity. This approach provides a precise, intelligent pathway for the design and optimization of welding fluxes.

Research Methodology

Figure 1A presents the overall research flowchart of the present study, which comprises four stages: data collection, data preprocessing, model establishment, and experimental validation.

Database Collection

We curated a viscosity database for welding fluxes by aggregating peer-reviewed literature and rigorously uniform experimental records. During curation, compositions of corresponding welding fluxes were CaF_2, CaO, SiO_2, MgO, TiO_2, Al_2O_3, and other ancillary components. Then, the viscosity measurement methods and experimental conditions were standardized, and structural information was extracted in parallel. In addition, various measurement techniques have been reported in literature, among which the rotating cylinder method has been most widely adopted, and is generally regarded as a reliable approach for determining the viscosity of slags (Refs. 28, 31). Nevertheless, potential errors caused by spindle or crucible contamination should be considered. In detail, when employing molybdenum crucibles and spindles, measurements must be performed under inert or reducing atmospheres to ensure reliability (Ref. 32). To establish a high-quality viscosity database, the following criteria were implemented: (i) viscosity data should be obtained using the rotating cylinder method, (ii) experiments should be conducted under controlled atmospheres, and (iii) the apparatus should be calibrated with castor oil before testing. For studies reporting results solely in graphical form, numerical data were carefully extracted using Origin 2024. Details of the viscosity measurement procedure are further discussed in the Experimental Validation section. The database was divided into three mutually exclusive subsets: 70% of the samples were used for model training, 15% were reserved for validation to monitor learning performance, and the remaining 15% formed the testing set for independent evaluation. The data were linearly normalized before being fed into the network.

Data Preprocessing

The data collection process was carried out manually to minimize the risk of missing potential records and to eliminate abnormal entries. Samples with a total slag composition lower than or exceeding 100 wt-% were excluded. To ensure consistency across different data sources, the unit scales were standardized during data acquisition, with chemical compositions being expressed in wt-%, viscosity values unified to Pa·s, and temperature represented in K.

In addition, a feature selection method based on Recursive Feature Elimination with Cross-Validation (RFECV) was applied to assess the contribution of each compositional input feature to the target viscosity value (Ref. 33). This approach integrates recursive feature elimination with cross-validation to automatically identify the optimal feature subset. Specifically, the procedure commences by training the base model on the complete feature set and ranking variables by their coefficient magnitudes (i.e., feature importances). The least informative feature is then iteratively removed, while the predictive performance of the reduced subset is evaluated through cross-validation at each step. The subset with the lowest cross-validation error is ultimately selected as the optimal solution, thereby ensuring predictive accuracy while effectively reducing redundancy.

Model Establishment

In the present study, a detailed architecture for viscosity prediction jointly driven by composition, temperature, and structural information was established, as illustrated in Fig. 1B. By incorporating structural information, the model effectively captures nonlinear relationships between composition and viscosity, thereby enhancing the generalization capability of data-driven models. Meanwhile, the hierarchical framework allows preliminary prediction of structural features without relying on experimental measurements, which contributes to high-accuracy viscosity prediction. In the first stage, a structural proxy model was trained. Compositions were used as inputs, and structural parameters were used as outputs. The optimal fold number *K* was automatically selected, and *K*-fold cross-validation was performed to ensure in-fold training and out-of-fold prediction, thereby strictly preventing data leakage. This process enables the prediction of NBO/Si, represented as Z_1 (Eq. 1), while σ_1 (uncertainties of Z_1) was defined using out-of-fold RMSE (Eqs. 2–3). To further capture the mapping relationship from composition to structure, a bagging ensemble of *M* proxy networks, denoted as g_m, was trained. In the second stage, the primary viscosity prediction model was constructed. Inputs included compositions of fluxes, the temperature T, the structural variable Z, and the corresponding uncertainty σ. To enhance prediction accuracy and efficiency, feature selection was used to remove low-impact compositional features. To enhance generalization, a noise injection mechanism was adopted, in which Gaussian perturbations proportional to σ_1 were added to the structural variable, thereby generating Z_{noisy} (Eq. 4). This mechanism functions as adaptive regularization, enlarging the input neighborhood in high-uncertainty regions, and thus enforcing smoother and more conservative mappings, thereby improving both generalization capability and extrapolation robustness.

During validation and testing, two scenarios were considered. If external samples provided measured NBO/Si, these were normalized using global scaling factors from training and directly fed into the viscosity model. If structural parameters were unavailable, the bagging ensemble of *M* proxy models generated predictions, and the mean of these predictions was taken as the structural variable Z_2 (Eq. 5). The associated uncertainty σ_2 (Eq. 6) was estimated from the square root of the ensemble variance combined with the maximum overall out-of-fold RMSE. In this way, each data instance to be verified was endowed with a complete (z, σ_z) representation, which was subsequently incorporated into the viscosity prediction network, enabling robust prediction of welding flux viscosity. In addition, to further evaluate the contribution of structural features, we ran an ablation study in which we trained an end-to-end deep neural network (DNN) model with all else fixed using only composition and temperature. All computations were conducted in MATLAB 2019 using the Deep Learning Toolbox. The pseudocode of S-DNN model is provided in the Appendix (Table A1).

$$Z_1 = f_k(X_i), i \in \nu_k \tag{1}$$

$$\sigma_1 = RMSE_K + \varepsilon, i \in \nu_k \tag{2}$$

$$RMSE_K = \sqrt{\frac{1}{|\nu_K|}\sum_{i\in\nu_K}(f_k(X_i) - u_i)^2} \tag{3}$$

$$Z_{noisy} = clip(z + \chi.\sigma_Z, 0,1) \tag{4}$$

$$Z_2 = \frac{1}{M}\sum_{m=1}^{M} g_m(X) \tag{5}$$

$$\sigma_2 = \max\left(\text{std}\{g_m(X)\}_{m=1}^{M}, \frac{1}{K}\sum_{k=1}^{K} RMSE_k\right) + \varepsilon \tag{6}$$

where X_i denotes the compositional vector of the i-th sample, f_K represents the predictive model trained on the *K*-th fold, ν_K refers to the validation set of the *K*-fold cross-validation, ε is a numerical stability term, u_i denotes the true value of NBO/Si for the i-th sample, M is the number of models in the bagging ensemble, g_m indicates the m-th bagging model, and λ is a random perturbation factor.

Experimental Validation

Commercial fluxes based on CaF_2-Al_2O_3-MgO-SiO_2 systems are widely applied and investigated due to their inborn ability to protect the weld pool from atmospheric exposure, enhance WM refinement through element transfer, and minimize heat loss by providing thermal insulation (Ref. 34). Therefore, this system was selected for validating the generalization capability of the model. In parallel, the MgO/SiO_2 molar ratio was adjusted to achieve favorable melting points and adequate

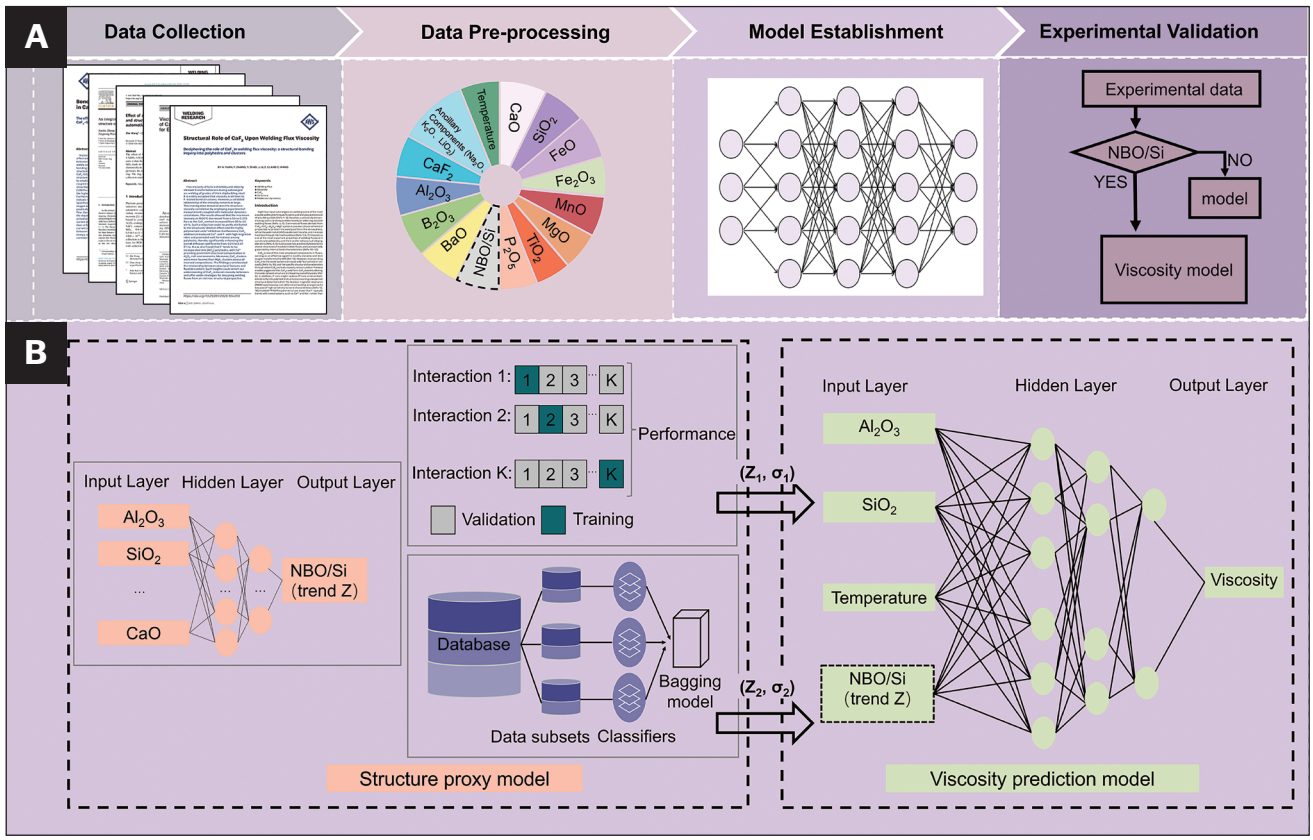

Fig. 1 — S-DNN viscosity prediction model: A — Flowchart; B — detailed architecture.

Table 1 — Chemical Compositions of the Fluxes (mol %)

Fluxes No.	CaF_2	Al_2O_3	MgO	SiO_2	MgO/SiO_2
A1	30.00	20.00	14.29	35.71	0.40
A2	30.00	20.00	17.74	32.26	0.55
A3	30.00	20.00	20.59	29.41	0.70

WM impact strength (Ref. 35). Reagent-grade CaF_2, Al_2O_3, MgO, and SiO_2 (Sinopharm Chemical Reagent Co. Ltd.) were employed as raw materials for flux preparation, with corresponding compositions being detailed in Table 1. Quaternary CaF_2-Al_2O_3-MgO-SiO_2 fluxes were synthesized through conventional melting and quenching. Initially, the mixed powders were preheated at 1173 K for 5 h in a resistance furnace to eliminate residual moisture, followed by melting at 1823 K for 30 min. in a silicon-molybdenum resistance furnace under a high-purity argon atmosphere (0.3 L/min). The molten flux was subsequently quenched in water (273 K) and dried at 473 K for 3 h. The obtained samples were ground and sieved to particle sizes below 74 µm.

Viscosity measurements were performed using the rotating-cylinder method in accordance with YB/T 185-2017. The experiments were conducted on a melt physical property comprehensive testing system (VDR-16000, Chongqing University, China) equipped with a Brookfield digital viscometer (DV2T, Brookfield Engineering Laboratories, USA). Before testing, the viscometer was calibrated with standard silicone oil (0.499 Pa·s) at 298 K. For each trial, 100 g of flux powder was placed in a molybdenum crucible (100 mm in height) and positioned in the uniform-temperature zone of a $MoSi_2$ resistance furnace. The furnace temperature was calibrated using a reference B-type thermocouple and maintained within ± 2 K. To prevent crucible oxidation, high-purity argon (99.999 vol-%) was continuously introduced at a flow rate of 0.3 L/min. The samples were heated to 1823 K at 5 K/min and held for 1 h to ensure chemical homogeneity. A molybdenum spindle (Φ15 mm × 20 mm; taper angle: 120 deg; shaft diameter: 3 mm) was immersed in the molten slag and positioned 20 mm above the bottom of the crucible. Measurements were carried out over the temperature range of 1523 to 1773 K, with each setpoint being held for 20 min to secure thermal

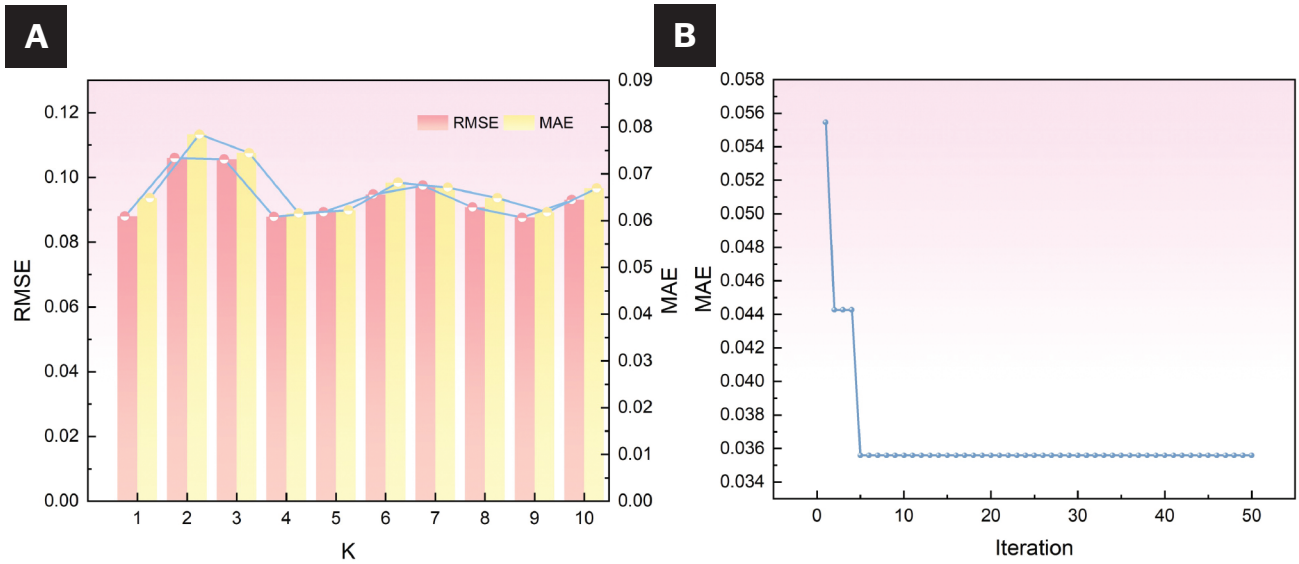

Fig. 2 — Model optimization process: A — Performance comparison across different K in K-fold cross-validation; B — optimization trace of Bayesian Optimization.

Table 2 — List of Hyperparameters in the Model

Hyperparameters	Range
Number of Hidden Layers	(1, 3)
Hidden Layer Size	(1, 500)
Epochs	(500, 2000)
Optimization Tolerance	($1e^{-6}$, $1e^{-3}$)
Bayesian-Optimization Iterations	(10, 200)
Learning Rate	($1e^{-4}$, $1e^{-2}$)
Number of Bagging Base Learners	(5, 30)
L2 Regularization Factor	($1e^{-4}$, $1e^{-1}$)
Cross-validation folds	(1, 10)

equilibrium. At each temperature, viscosity was determined as the average of three rotation speeds (10, 20, and 30 rpm), and the deviations from the mean remained within 3%.

Optimization Algorithm and Model Validation

To alleviate the inherent stochastic bias potentially incurred by single-pass partitioning of the dataset into training and testing subsets, we employed a *K*-fold cross-validation scheme to assess the generalization capability of the model rigorously (Ref. 36). Concretely, the dataset was stratified into *K* mutually exclusive subsets, wherein each fold sequentially served as the validation set while the remaining *K*-1 folds were designated for training. The process was cyclically repeated *K* times, and the final performance metric was computed by aggregating and averaging all validation outcomes. To determine the optimal value of *K*, a global search strategy was implemented. As illustrated in Fig. 2A, the model attained peak robustness when $K = 9$, which was consequently adopted as the optimal fold parameter.

To further enhance the robustness of the structural prediction model, a bagging ensemble strategy was implemented (Ref. 37). Bootstrap sampling generated multiple training subsets, and base learners were trained independently on each subset. During the prediction phase, the final prediction was obtained by averaging the outputs of all base

Table 3 — Feature Ranges of Variables in the Dataset

	Database	Min (wt-%)	Max (wt-%)
Compositions	Al_2O_3	2.28	63.83
	CaF_2	1.00	34.18
	CaO	1.50	59.20
	SiO_2	2.31	56.40
	FeO	2.95	38.6
	Fe_2O_3	0.02	20.00
	MgO	1.00	33.00
	MnO	5.51	25.00
	TiO_2	1.25	55.00
	P_2O_5	2.82	3.00
	BaO	1.00	18.89
	B_2O_3	1.00	12.00
	Ancillary Components (Na_2O, K_2O, Li_2O)	7.71	14.89
	Temperature	1288.95 K	1923 K
	Viscosity	0.011 Pa·s	1.943 Pa·s
	NBO/Si	1.058	3.798

learners. This approach effectively reduces model variance, alleviates overfitting, and enhances generalization. In addition, prediction deviations across models were employed to quantify uncertainty, thereby offering more reliable confidence assessments for subsequent analyses.

In addition, to achieve optimal hyperparameter tuning for the viscosity prediction model, the Bayesian Optimization (BO) strategy was adopted. Due to the inherent balance between global exploration and local exploitation in high-dimensional search spaces, BO has been widely recognized for effectiveness in deep learning hyperparameter tuning (Refs. 38, 39). As shown in Fig. 2B, the objective function converged rapidly after approximately 5 iterations, indicating stabilization of the search trajectory. Compared with random search paradigms, the BO architecture yielded not only substantial improvements in prediction accuracy but also a marked reduction in computational burden. In addition, the model's specific hyperparameter configuration is shown in Table 2. The BO process is documented in the Appendix (Table A2).

To ensure a comprehensive and rigorous assessment of the predictive capabilities of the model, three performance functions — MAE, RMSE, and R2 — were employed. MAE quantifies the average magnitude of the absolute difference between predicted and experimental values, thereby reflecting the model's global prediction bias. In contrast, RMSE is more sensitive to individual large deviations, rendering it a stricter criterion for error evaluation. As for R2, it is designed to characterize the proportion of variance in the observed data that can be explained by the predicted outputs (Ref. 40). A value approaching unity indicates a higher level of agreement and overall fitting accuracy. Taken together, these evaluation indices provide a multidimensional perspective on the model's reliability, robustness, and generalization performance across different datasets.

$$MAE = \frac{1}{N}\sum_{1}^{N}\left|y_t - y_p\right| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{1}^{N}|y_t - y_p|} \quad (8)$$

$$R^2 = 1 - \frac{\sum_1^N (y_t - y_p)^2}{\sum_1^N (y_t - \bar{y_t})^2} \quad (9)$$

Results and Discussion

Data and Features

A total of 1383 datasets were collected, each containing composition, temperature, structural features, and corresponding viscosity values. Based on these data, a systematic welding flux database was established, with statistical characteristics being summarized in Table 3. The distribution of the datasets was further visualized using violin plots, as shown in Fig. 3. During data collection, it was noted that the available literature on welding flux viscosity remains limited, and the sample diversity is also limited. To address this deficiency, the present work, while strictly adhering to the compositional scope of welding flux, was moderately expanded to include relevant data from electroslag, refining, and mold slag systems. The major components in the database are CaF_2, SiO_2, Al_2O_3, MgO, etc. The inclusion of these diverse systems relies on a fundamental commonality. Their viscosity is governed by the DOP within the silicate network, a relationship universally quantified by the NBO/Si parameter. This shared theoretical foundation ensures structural comparability. To further guarantee methodological consistency, the analysis incorporated only data obtained through standardized measurement.

The analysis of the correlation between composition and viscosity was conducted using the RFECV method, and the results are presented in Fig. 4. The analysis revealed that SiO_2 exerts the most significant influence on viscosity, which is highly consistent with previous findings (Ref. 41). In addition, within the defined correlation coefficient range, several other influential components were identified, including CaF_2, Al_2O_3, CaO, FeO, MgO, and Fe_2O_3. Based on the established database, certain components were excluded primarily due

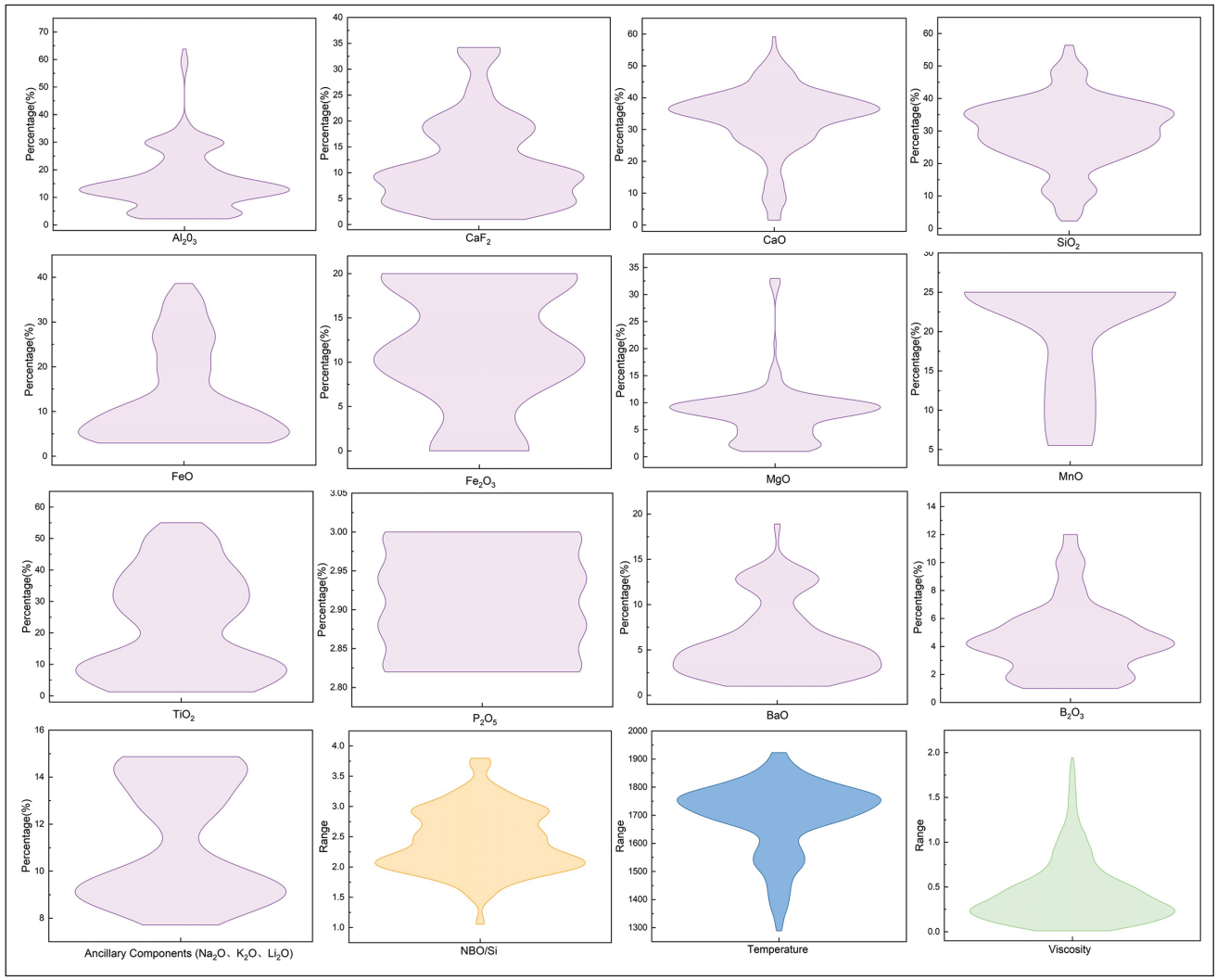

Fig. 3 — Violin and box plots of data distribution for each feature.

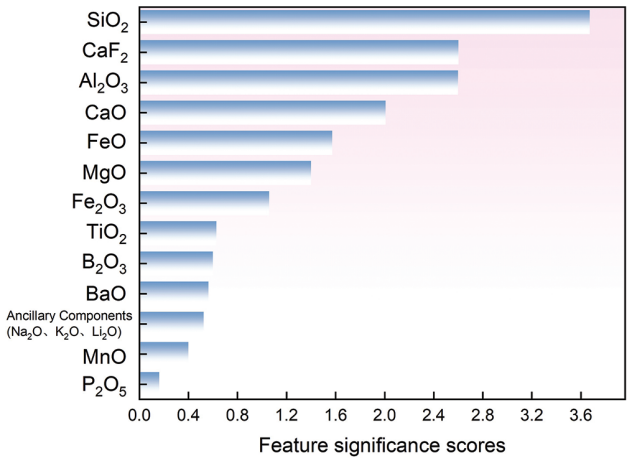

Fig. 4 — Correlation analysis of independent features.

to insufficient data quality and limited application in welding flux systems. Through this screening process, critical input components for the viscosity prediction model were identified, thereby enhancing the model's interpretability while effectively avoiding feature redundancy.

Structural Proxy Model

Predictions from each fold were visualized in Fig. 5A, which shows that the predicted values align closely with the diagonal line y = x, indicating the model's ability to reliably reproduce the true structural parameters. To further evaluate stability, prediction errors for each fold were analyzed (Fig. 5B). The results indicate that the maximum RMSE was 0.0943 in fold 4, while the minimum was 0.0666 in fold 1, with an overall average RMSE of 0.0872, suggesting reasonable data partitioning and reliable model performance. Based on this, the training performance of the bagging ensemble was evaluated, as presented in Fig. 5C. The RMSE values ranged from 0.0463 (model 8) to 0.0689 (model 6). Although some variability was observed

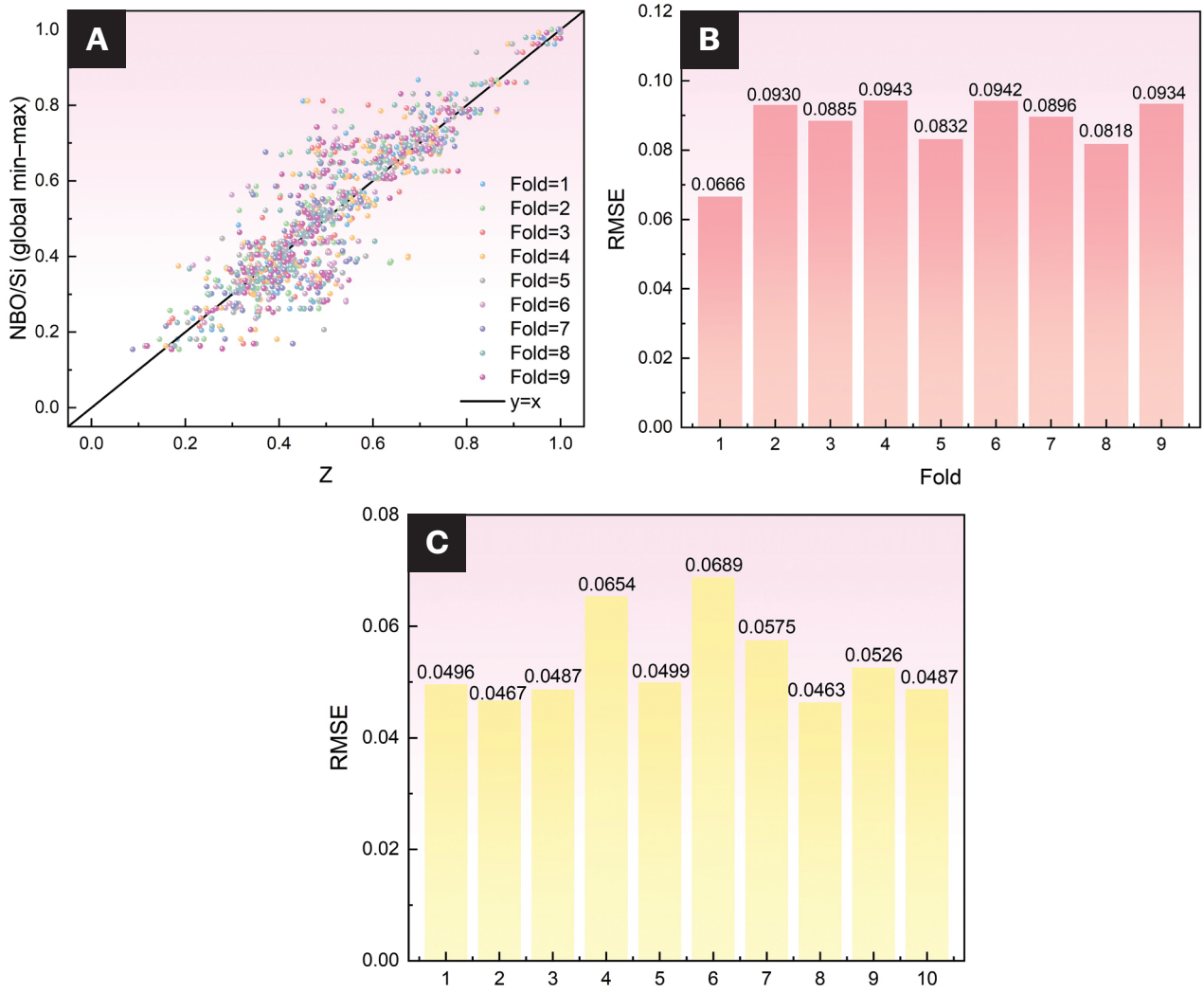

Fig. 5 — Performance evaluation of the structural proxy model: A — Normalized NBO/Si predictions; B — RMSE for each fold in K-fold cross-validation; C — RMSE of each base learner in the bagging ensemble.

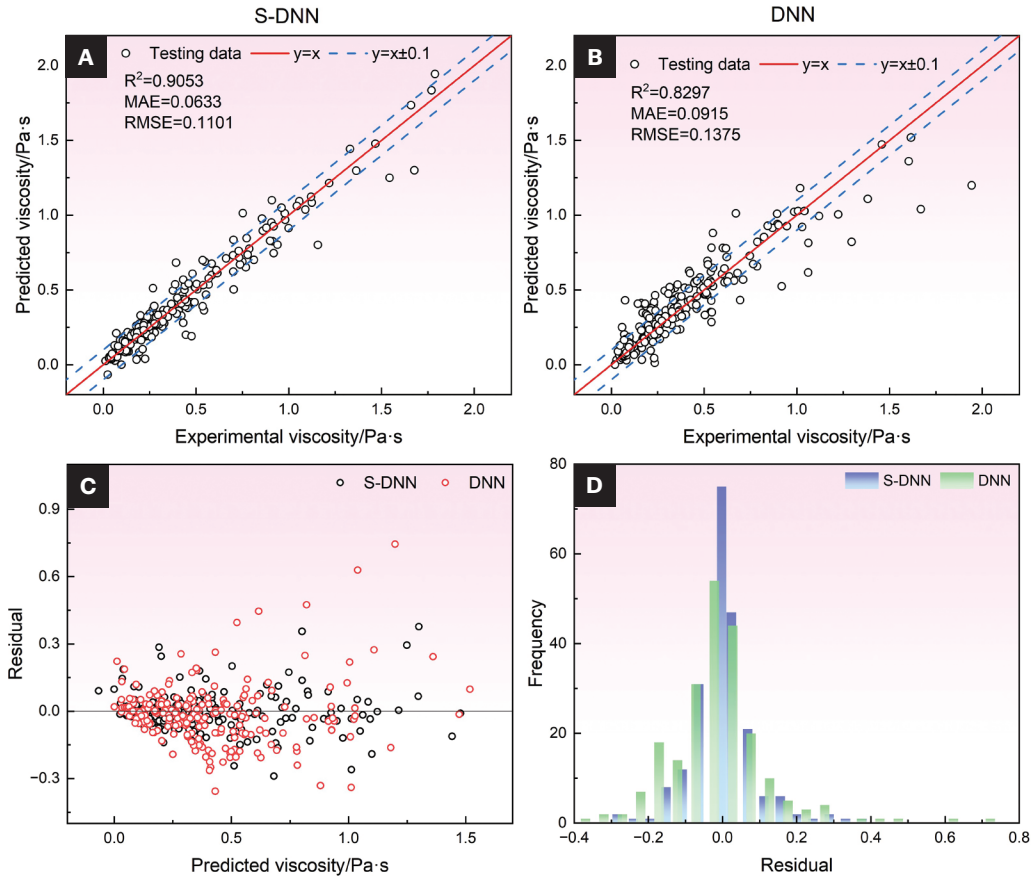

Fig. 6 — Testing performance of the viscosity prediction model: A — Parity plot for S-DNN; B — parity plot for DNN; C — residual distribution; D — residual histogram.

among individual base learners, the ensemble achieved an average RMSE of 0.0534, with significantly reduced variance, thereby substantially improving robustness. These results collectively demonstrate that the structural proxy model exhibits satisfactory accuracy and stability. Consequently, the predicted structural information was incorporated into the viscosity-prediction framework to further enhance the model's performance.

Viscosity Prediction Model

The generalization capability of both the S-DNN and DNN models was evaluated on the testing set, and the corresponding prediction results are illustrated in Figs. 6A and B. The results clearly demonstrate that S-DNN exhibits significantly superior predictive performance compared with DNN. In detail, the predicted scatter points of S-DNN are more tightly distributed along the $y = x$ line, with the majority being confined within the $y = \pm 0.1x$ range, whereas those of DNN appear more dispersed. Quantitative performance metrics further confirm this trend, as R^2, MAE, and RMSE values of S-DNN are 0.9053, 0.0633 Pa·s, and 0.1101 Pa·s, respectively, both outperforming the DNN model. A comparison of residuals between the two models, as shown in Fig. 6C reveals that the residuals of DNN are more widely scattered and exhibit markedly larger fluctuations, while those of S-DNN remain more compact and stable. Moreover, Fig. 6D presents histograms of absolute error distributions for the testing data. Although both models show relatively concentrated error distributions that approximate normality, the distribution of S-DNN is more convergent and primarily centered around 0.01. In summary, these results indicate that S-DNN surpasses DNN in terms of predictive reliability and stability, thereby demonstrating a more prominent generalization capability. The enhanced performance can be attributed to the incorporation of NBO/Si as a physically meaningful mediated structural information. On one hand, such structural information provides physical constraints and causal guidance to the model; on the other hand, it reframes the highly nonlinear viscosity prediction problem into lower-dimensional subproblems. This decomposition effectively reduces model variance and significantly improves generalization capability.

Evaluation and Prospects of the Prediction Model

CaF_2-Al_2O_3-MgO-SiO_2 welding flux systems were prepared, and composition, temperature, and viscosity data from three sample groups were used to form validation subsets. Through unseen validation subsets, the generalizability and engineer-

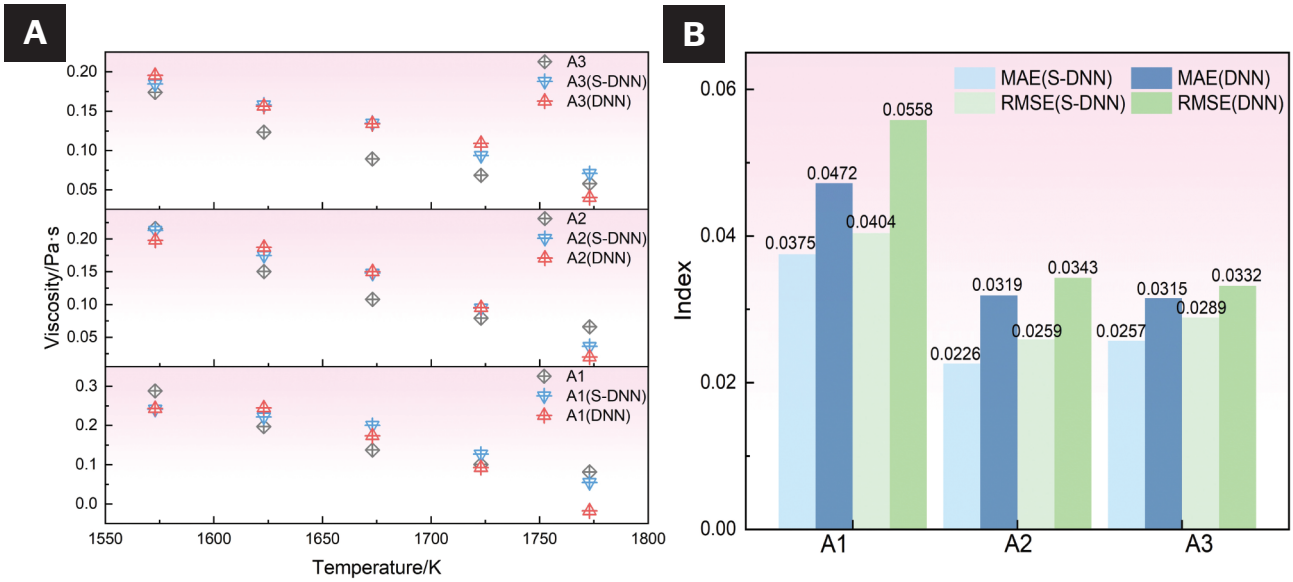

Fig. 7 — A — Comparison of experimental and predicted viscosity of A1-A3; B — MAE and RMSE between S-DNN and DNN models.

ing applicability of both the S-DNN and DNN models were evaluated. As shown in Fig. 7A, both models exhibit similar trends in viscosity variation with temperature. However, the S-DNN model's predictions consistently align more closely with measured values. To quantitatively assess model performance, the MAE and RMSE were calculated for samples A1-A3, as summarized in Fig. 7B. The S-DNN model achieved an MAE of 0.0375 Pa·s, 0.0226 Pa·s, and 0.0257 Pa·s, and the RMSE values are 0.0404 Pa·s, 0.0259 Pa·s, and 0.0289 Pa·s, respectively, demonstrating superior predictive accuracy compared to the DNN model. This enhancement can be attributed to the incorporation of structural information, which enables the S-DNN to better capture the complex, nonlinear relationships between input features and output viscosity. Notably, high-accuracy viscosity prediction can be achieved without additional experiments to obtain NBO/Si, thereby enhancing the engineering applicability of the S-DNN model. These results confirm the efficacy of the S-DNN framework within the compositional range represented by the present database. It should be noted, however, that the model's reliance on NBO/Si as a structural proxy is most physically justified for silicate-based flux systems, in which Si^{4+} serves as the primary network former. For systems devoid of SiO_2, such as Al_2O_3-based fluxes, or those with markedly different structural motifs, the predictive performance of the model may be limited. Future improvements can be achieved by incorporating additional structural descriptors and expanding the viscosity dataset.

This study establishes a dedicated database for welding flux viscosity prediction and introduces a novel structure-informed viscosity model that hierarchically models the composition-structure-viscosity relationship, leading to the development of a high-precision viscosity prediction model. The proposed model significantly facilitates the design and optimization of welding flux, thereby enhancing the service reliability of WM in HSLA steels. Additionally, this work offers a new strategy and lays a methodological foundation for applying neural networks to welding. Furthermore, intelligent control models for other key physicochemical properties of fluxes (e.g., electrical conduction, surface tension, and melting temperature) will be further curated and elucidated in future endeavors.

Conclusions

In this study, a model for predicting welding flux viscosity, driven by composition, temperature, and particularly structural information, was developed. This framework not only improves predictive accuracy but also significantly broadens engineering applicability. The main conclusions are summarized as follows:

- The structural proxy model reliably infers missing structural parameters, achieving an average RMSE of 0.0534.
- Incorporating structural information substantially improves the accuracy and generalization capability of the viscosity prediction model. On the test set, the S-DNN model outperformed the basic DNN model.
- Even without experimental structural data, the S-DNN model could accurately predict viscosity across multiple sets of experimental data, demonstrating robust generalization to unseen conditions.

Acknowledgments

The authors sincerely acknowledge the financial support from the National Natural Science Foundation of China (Grant Nos. W2411047 and 52404393), the Fundamental Research Funds for the Central Universities (Grant Nos. 2025GFZ018 and N25ZLL045), and the State Key Laboratory of Precision Welding & Joining of Materials and Structures (No. MSWJ-24M03).

References

1. Wang, Z., Zhong, M., Liu, H., Zhang, Y., Imants, K., and Wang, C. 2025. Assessing element transfer behaviors in the weld metal of EH36 shipbuilding steel: A case study into CaF_2-MnO-CaO-SiO_2 fluxes. *Metallurgical and Materials Transactions B* 56(3): 2144–2149. DOI: 10.1007/s11663-025-03532-9

2. Wang, S., Chiang, C., and Chan, S. 2003. Effect of initial microstructure on the creep behavior of TMCP EH36 and SM490C steels. *Materials Science and Engineering A* 344(1–2): 288–295. DOI: 10.1016/S0921-5093(02)00425-2

3. Wu, Y., Yuan, X., Zhong, M., Ma, J., and Wang, C. 2023. Tailoring columnar austenite grains in the weld metal of EH36 shipbuilding steel via CaF_2-TiO_2 fluxes. *Metallurgical and Materials Transactions A* 54(12): 4573–4578. DOI: 10.1007/s11661-023-07199-2

4. Iyama, J., Matsumoto, Y., Ishii, T., Shimokawa, H., Nikaido, M., and Yamada, S. 2019. Fracture strength of electroslag welding joint with high-performance steel. *Journal of Constructional Steel Research* 153: 495–508. DOI: 10.1016/j.jcsr.2018.11.010

5. Kiran, D., Basu, B., and De, A. 2012. Influence of process variables on weld bead quality in two wire tandem submerged arc welding of HSLA steel. *Journal of Materials Processing Technology* 212(10): 2041–2050. DOI: 10.1016/j.jmatprotec.2012.05.008

6. Sengupta, V., Havrylov, D., and Mendez, P. F. 2019. Physical phenomena in the weld zone of submerged arc welding: A review. *Welding Journal* 98(10): 283S–313S. DOI: 10.29391/2019.98.025

7. Wang, G., Zhang, Y., Tian, H., Li, Z., and Wang, C. 2026. Pinpointing element transfer locations during submerged arc welding. *Welding Journal* 105(1): 18–26. DOI: 10.29391/2026.105.002

8. Sun, L., Li, H., Zhu, L., Liu, Y., and Hwang, J. 2020. Research on the evolution mechanism of pinned particles in welding HAZ of Mg treated shipbuilding steel. *Materials & Design* 192: 108670. DOI: 10.1016/j.matdes.2020.108670

9. Yuan, H., Zhang, Y., Liu, H., Li, Z., and Wang, C. 2025. Bond characteristic-dependent viscosity variations in CaF_2-SiO_2-Al_2O_3-MgO welding fluxes. *Welding Journal* 104(4): 107S–118S. DOI: 10.29391/2025.104.009

10. Han, C., Zhong, M., Zuo, P., and Wang, C. 2024. SiO_2-bearing fluxes induced evolution of γ columnar grain size. *Welding Journal* 103(12): 362S–371S. DOI: 10.29391/2024.103.031

11. Kurlanov, S., Potapov, N., and Natapov, O. 1993. Relationship of physical and welding-technological properties of fluxes for welding low-alloy steels. *Welding International* 7(1): 65–68. DOI: 10.1080/09507119309548346

12. Song, X., Zhang, Y., Yuan, H., Li, Z., and Wang, C. 2025. Exploring the role of ZrO_2 on viscosity and structure of CaF_2-SiO_2-CaO-ZrO_2 welding fluxes. *Metallurgical and Materials Transactions B* 56(5): 5671–5681. DOI: 10.1007/s11663-025-03723-4

13. Yuan, H., Zhang, Y., and Wang, C. 2026. Influence of MnO upon electrical conductive mechanisms of submerged arc welding fluxes: Insights from ab initio molecular dynamics simulations. *Acta Metallurgica Sinica* 62(1): 191–202. DOI: 10.11900/0412.1961.2025.00065

14. Schwemmer, D. D., Olson, D. L., and Williamson, D. L. 1979. The relationship of weld penetration to the welding flux. *Welding Journal* 58(5): 153S–160S.

15. Bai, H., Zhang, Y., Zhao, Y., Zhong, M., Li, Z., and Wang, C. 2025. Numerical analysis of slag viscosity effects mechanism in submerged arc welding pool. *Metallurgical and Materials Transactions B* 56(2): 1659–1671. DOI: 10.1007/s11663-025-03448-4

16. Akao, K., Ishihara, T., Kitada, T., Nishino, Y., Okuda, N., Wada, T., and Nagatani, K. 1986. Improvement of notch toughness and soundness in longitudinal seam weld of line pipe. *Transactions of the Iron and Steel Institute of Japan* 26(5): 379–386.

17. Bai, H., Zhang, Y., Yuan, H., and Wang, C. 2025. Identifying the role of slag viscosity upon submerged arc weld pool features: A computational fluid dynamics investigation. *Metallurgical and Materials Transactions B* 56(4): 4036–4048. DOI: 10.1007/s11663-025-03600-0

18. Zhang, Y., Yuan, H., Tian, H., Wang, Z., and Wang, C. 2023. Elucidating electrical conductive mechanisms for CaF_2-SiO_2-CaO-TiO_2 welding fluxes. *Metallurgical and Materials Transactions B* 54(6): 3023–3030. DOI: 10.1007/s11663-023-02885-3

19. Liu, H., Zhang, Y., Shi, S., Yuan, H., Li, Z., and Wang, C. 2025. Predicting viscosity of submerged arc welding fluxes: A deep neural network approach facilitated by data augmentation. *Metallurgical and Materials Transactions B* 56(4): 3927–3937. DOI: 10.1007/s11663-025-03606-8

20. Kim, J., Lee, T., and Sohn, I. 2018. Effect of compositional variation in TiO_2-based flux-cored arc welding fluxes on the thermo-physical properties and mechanical behavior of a weld zone. *Metallurgical and Materials Transactions A* 49(7): 2705–2720. DOI: 10.1007/s11661-018-4624-1

21. Haghdani, S., Tangstad, M., and Einarsrud, K. E. 2022. A Raman-structure model for the viscosity of SiO_2-CaO-Al_2O_3 system. *Metallurgical and Materials Transactions B* 53(3): 1733–1746. DOI: 10.1007/s11663-022-02483-9

22. Liu, H., Zhang, Y., Zhao, Y., Yuan, H., Basu, S., and Wang, C. 2025. Unveiling the amphoteric behavior of TiO_2 in fused CaF_2-TiO_2-MgO-SiO_2 submerged arc welding fluxes. *Metallurgical and Materials Transactions B* 56(1): 699–710. DOI: 10.1007/s11663-024-03383-w

23. Chou, K. C. 2022. General solution model and its new progress. International *Journal of Minerals, Metallurgy and Materials* 29(4): 577–585. DOI: 10.1007/s12613-022-2411-x

24. Zhang, G., Zhen, Y., and Chou, K. C. 2016. Viscosity and structure changes of CaO-SiO_2-Al_2O_3-CaF_2 with substituting Al_2O_3 for SiO_2. *Journal of Iron and Steel Research International* 23(7): 633–637. DOI: 10.1016/j.jisri.2016.05.007

25. Han, L., and Xiang, Z. 2025. Intelligent design and synthesis of energy catalytic materials. *Fundamental Research* 5(2): 624–639. DOI: 10.1016/j.fmre.2023.10.012

26. Chen, X., Dong, Y., Jiang, Z., and Liu, Y. 2025. Predicting the viscosity of CaF_2-based slag and reverse design of electroslag systems using explainable machine learning. *Metallurgical and Materials Transactions B* 56(3): 3125–3139. DOI: 10.1007/s11663-025-03484-0

27. Xuan, W., Zhao, C., Ma, Z., Liu, Z., and Zhang, J. 2025. Data-driven machine learning methods to predict the viscosity-temperature characteristics for gasification slags. *Fuel* 398: 135287. DOI: 10.1016/j.fuel.2025.135287

28. Chen, Z., Wang, M., Meng, Z., Wang, H., Liu, L., and Wang, X. 2021. Development of structure-informed artificial neural network for accurately modeling viscosity of multicomponent molten slags. *Ceramics International* 47(21): 30691–30701. DOI: 10.1016/j.ceramint.2021.07.248

29. Jiang, D., Zhang, J., Wang, Z., Feng, C., Jiao, K., and Xu, R. 2020. A prediction model of blast furnace slag viscosity based on principal component analysis and k-nearest neighbor regression. *JOM* 72(11): 3908–3916. DOI: 10.1007/s11837-020-04360-9

30. Yan, W., Shen, Y., Chen, S., and Wang, Y. 2024. Viscosity and melting temperature prediction of mold fluxes based on explainable machine learning and Shapley additive explanations. *Journal of Non-Crystalline Solids* 636: 123037. DOI: 10.1016/j.jnoncrysol.2024.123037

31. Cui, L., Ni, P., Lv, W., and Li, Y. 2025. A review on viscosity and structure of high-temperature borosilicate melts. *Journal of Non-Crystalline Solids* 666: 123683. DOI: 10.1016/j.jnoncrysol.2025.123683

32. Hou, Y., Zhang, S., Guo, J., Zhou, H., and Lv, X. 2024. Effect of BaO/Al_2O_3 molar ratio on the viscosity and structure of CaO-SiO_2-MgO-Al_2O_3-BaO slag. *Journal of Non-Crystalline Solids* 626: 122815. DOI: 10.1016/j.jnoncrysol.2023.122815

33. Liu, W., and Wang, J. 2021. Recursive elimination-election algorithms for wrapper feature selection. *Applied Soft Computing* 113: 107956. DOI: 10.1016/j.asoc.2021.107956

34. Yuan, H., Zhang, Y., Wang, C., Zhao, Y., Li, J., Li, Z., and Wang, C. 2025. Structural role of CaF_2 upon welding flux viscosity. *Welding Journal* 104(5): 164S–174S. DOI: 10.29391/2025.104.013

35. Natalie, C. A., and Olson, D. L. 1986. Physical and chemical behavior of welding fluxes. *Annual Review of Materials Science* 16: 389–413. DOI: 10.1146/annurev.ms.16.080186.002133

36. Xu, P., Ji, X., Li, M., and Lu, W. 2023. Small data machine learning in materials science. *npj Computational Materials* 9(1): 42. DOI: 10.1038/s41524-023-01000-z

37. Liu, Y., Yang, Z., Zou, X., Ma, S., Liu, D., Avdeev, M., and Shi, S. 2023. Data quantity governance for machine learning in materials science. *National Science Review* 10(7): nwad125. DOI: 10.1093/nsr/nwad125

38. Mukherjee, S., Sankar, M. A., and Mantripragada, V. T. 2025. A robust machine learning framework for predicting slag eye formation in industrial steelmaking ladles. *Metallurgical and Materials Transactions B* 56(5): 4936–4950. DOI: 10.1007/s11663-025-03670-0

39. Ji, Y., Chen, J., Zhou, L., Wang, W., Qi, J., Liu, P., and Chen, K. 2025. Crystallization prediction and reverse engineering framework construction for mold flux based on machine learning methods. *Materials Today Communications* 45: 112426. DOI: 10.1016/j.mtcomm.2025.112426

40. Zhang, H., and Zhou, Y. 2023. AI-based modeling and data-driven identification of moving load on continuous beams. *Fundamental Research* 3(5): 796–803. DOI: 10.1016/j.fmre.2022.02.013

41. Yang, J., Wang, Z., and Sohn, I. 2022. Topological understanding of thermal conductivity in synthetic slag melts for energy recovery: An experimental and molecular dynamic simulation study. *Acta Materialia* 234: 118014. DOI: 10.1016/j.actamat.2022.118014

Appendix

Table A1 — Pseudocode of S-DNN Model

Algorithm 1 S-DNN for Viscosity Predictions

Require: Composition matrix $X_{comp} \in R^{N\times 13}$, Temperature vector T, Partial structural targets (NBO/Si) u, Number of folds K, Ensemble size M

Ensure: Trained S-DNN parameters θ^S, Predicted viscosity, $\hat{y}$ Uncertainty σ_z

Structure Proxy Modeling:

Initialize trend feature vector z ← NaN, uncertainty vector $\sigma_z \leftarrow$ NaN

Partition NBO-labeled subset $\mathcal{D}_u$ into K disjoint folds $\{F_1, \ldots, F_K\}$

for each fold $k = 1$ to K **do**

  Train a local regressor f_k on $\{\mathcal{D}_u \setminus F_k\}$ mapping $X_{comp} \rightarrow u$.

  for samples in F_k **do**

    Generate out-of-fold prediction: $(\hat{z_k}) = f_k(X_{comp})$

    Compute fold-specific residual $\epsilon_k = \mathrm{RMSE}(\hat{z_k}, u_k)$

    Update $z[F_k] \leftarrow \hat{z_k}$ and $\sigma_z[F_k] \leftarrow \epsilon_k$

  end for

end for

Bagging Ensemble Completion:

Train an ensemble of M diverse networks $\{B_1, \ldots, B_M\}$ on the full set $\mathcal{D}_u$

for each sample $i \notin \mathcal{D}_u$ **do**

  Compute mean μ_i and standard deviation s_i from ensemble predictions $\{B_m(X_i)\}_{m=1}^{M}$

  $z_i \leftarrow \mu_i$, $\sigma_{z,i} \leftarrow \max(s_i, \mathrm{mean}(\epsilon))$

Table A1 — continued

17:	**end for**
18:	**S-DNN Training with Stochastic Regularization:**
19:	**Feature Selection:**
20:	Extract optimized feature subset $X_{sub} \in R^{N\times7}$ from X_{comp} via RFECV
21:	Construct input tensor $\widetilde{X} = [X_z, z, \sigma_z, T]$
22:	**Bayesian Architecture Search:**
23:	Objective: $\min_{arch} Loss_{val}$ (DNN (arch; ($\widetilde{X}_{train}$), Y_{train}))
24:	**Stochastic Noise Injection**
25:	$z_{noisy} = clip(z + \zeta \cdot \sigma_z, 0, 1)$, where $\zeta \sim N(0,1)$
26:	Identify optimal hidden layer configuration $\{L_1, L_2, L_3\}$
27:	**Final Training:**
28:	Train S-DNN θ^S using optimal architecture on augmented input $\tilde{X}$
29:	**return:** θ^S, predicted $\hat{y}$, and physical uncertainty σ_z

Table A2 — The Bayesian Optimization Process

Hidden Layer Size	Number of Iterations	Training Set			Testing Set		
		MAE	RMSE	R^2	MAE	RMSE	R^2
(47,485,22)	1	0.09372	0.1454	0.8455	0.1135	0.1791	0.7492
(35,498,13)	2	0.08524	0.1228	0.8898	0.09613	0.1493	0.8256
(8,161,3)	3	0.08524	0.1228	0.8898	0.09613	0.1493	0.8256
(9,14,6)	4	0.03901	0.05649	0.9767	0.09195	0.2234	0.6066
(10,18,5)	5	0.03782	0.05307	0.9794	0.0633	0.1101	0.9053
(10,18,5)	10	0.03782	0.05307	0.9794	0.0633	0.1101	0.9053
(10,18,5)	20	0.03782	0.05307	0.9794	0.0633	0.1101	0.9053
(10,18,5)	50	0.03782	0.05307	0.9794	0.0633	0.1101	0.9053

HONGYU LIU, YANYUN ZHANG, XIAOXI SONG, and **CONG WANG** (*wangc@smm.neu.edu.cn*) are with the Key Laboratory for Ecological Metallurgy of Multimetallic Mineral (Ministry of Education), School of Metallurgy, Northeastern University, Shenyang, China. **YANYUN ZHANG** is also with the State Key Laboratory of Precision Welding & Joining of Materials and Structures, Harbin, China. Lastly, **ZUSHU LI** is with WMG, University of Warwick, Coventry, UK.