Supplemental Information for

Assessment of Aerosol Iron Solubility using Global Dataset, Part II: Machine Learning and Deep Neural Network Coupled with SHapley Additive exPlanation Combined with Independent Component Analysis (SHAP-ICA).

Kohei Sakata[1*], Minako Kurisu[2], and Yoshio Takahashi[3]

[1]Materials Science and Engineering, Graduate School of Engineering, Tokyo Denki University, 5 Senjyu-Asahi-Cho, Adachi-ku, Tokyo 120-8551, Japan.

[2]Atmosphere and Ocean Research Institute, The University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa, Chiba 277-8564, Japan.

[3]Graduate School of Science, The University of Tokyo, 7-3-1, Hongo Bunkyo-ku, Tokyo 113-0033, Japan.

Email: kohei.sakata.33@mail.dendai.ac.jp

ORCID: 0000-0002-0103-9631

This PDF file includes:

Supporting text

Figures S1 to S14

Supplemental Text

S1. Machine environment

Data processing was performed using XGBoost (version 2.1.2) and TensorFlow (version 2.12.0) on a computing cluster equipped with an Intel 14th generation Core™ i7-14700F CPU, 32 GB of DDR4-3200 memory (16 GB × 2), and an NVIDIA GeForce RTX 4070 SUPER GPU (12 GB VRAM, 7168 CUDA cores). The NVIDIA driver, CUDA, and cuDNN versions used were 552.22, 12.4, and 8.2.1, respectively. Machine learning and deep learning processes were conducted within a Linux virtual environment (Ubuntu 22.04.5 LTS), using a Jupyter Notebook with a Python 3 kernel.

S2. Feature engineering of the input features for XGBoost and DNN

Feature engineering is crucial for developing highly accurate Fe_{sol}% prediction models. The predictors listed in Table 1 did not exhibit severe multicollinearity (all VIFs < 5, Figure S3a). Therefore, the adverse effects of multicollinearity on the model's accuracy and stability are minimal. To better capture the Fe_{sol}% variations associated with mineral dust concentration and aerosol acidification of Fe-bearing aerosols, the T-Al concentration (an indicator of mineral particles) and the [nss-SO_4^{2-}]/[T-Fe] ratio (an indicator of the acidity of Fe-containing aerosols) were considered for use as features, in addition to those from Table 1. However, these parameters were not employed as features because they negatively affected both the stability and accuracy of both XGBoost and DNN models (Figure 2). To identify the factors responsible for the deterioration in model accuracy, the model performances of XGBoost and DNN was validated by incorporating either T-Al concentrations or the [nss-SO_4^{2-}]/[T-Fe] ratio.

When T-Al concentration was added to the features listed in Table 1, multicollinearity arose due to its strong correlation with T-Fe concentration, which was reflected in VIF values higher than 5.0 for both T-Al and T-Fe (Figure S3b). In addition, even when ICA was performed with five components, T-Fe and T-Al concentrations could not be completely separated into distinct components, indicating that these variables were not independent. Even though, the performance of XGBoost and DNN deteriorated slightly due to multicollinearity, as reflected by a slight rise in their loss functions of MAE and RMSE (Figures 2a and 2c). In contrast, multicollinearity is likely to have a more detrimental impact on the SHAP evaluation of feature importance. Even though T-Fe and T-Al concentrations exhibit a strong correlation, the SHAP value of T-Fe declined with increasing concentration, whereas that of T-Al rose as its concentration increased (Figure S3c). This suggested that the contribution of T-Fe was distorted, effectively being canceled out by the opposing effect of T-Al. Indeed, T-Fe concentration showed negligible influence on Fe_{sol}% in the model that excluded T-Al (details are discussed in Section 3.2.1). Therefore, we judged that the interpretability of each feature's contribution to the predicted Fe_{sol}% was compromised by multicollinearity when T-Al was included; thus, T-Al was excluded from the feature set.

Given that prior research suggested aerosol Fe_{sol}% correlates with [nss-SO_4^{2-}]/[T-Fe], [NO_3^-]/[T-Fe], and [$C_2O_4^{2-}$]/[T-Fe] ratios (Zhu et al., 2020, 2022; Zhang et al., 2023; Sakata et al., 2023, 2025), the [nss-SO_4^{2-}]/[T-Fe] ratio was considered for inclusion in the feature set as an indicator of aerosol acidity for the purpose of evaluating its impact. Indeed, Shi et al. (2022) has employed sulfate, nitrate, and oxalate concentrations. However, reproducibility of Fe_{sol}% was decreased when [nss-SO_4^{2-}]/[T-Fe] was added to the input features (Figures 2a and 2d). This result is

considered to be due to the decrease in the number of available samples to 240, which led to a lack of training data. Therefore, this study adopted the parameters in Table 1 as features, as they secured the maximum number of samples and were not affected by multicollinearity.

Supplemental Figure

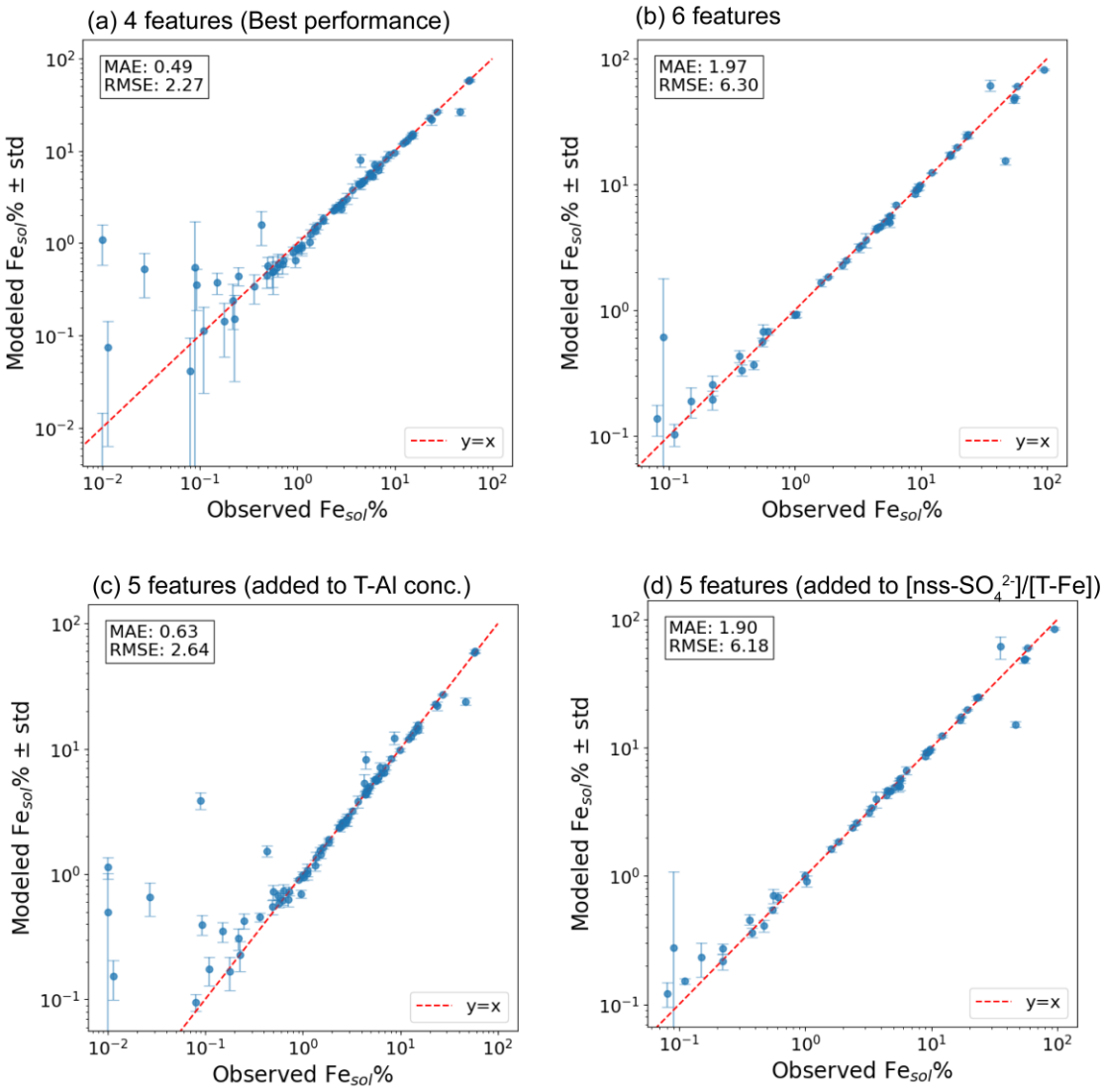


Figure S1. Correlation matrices and VIF values for the two feature sets: (a) the dataset of East Asian aerosol including T-Al concentration, and (b) the dataset excluding T-Al concentration. Colar bar shows Pearson's correlation factor (r).

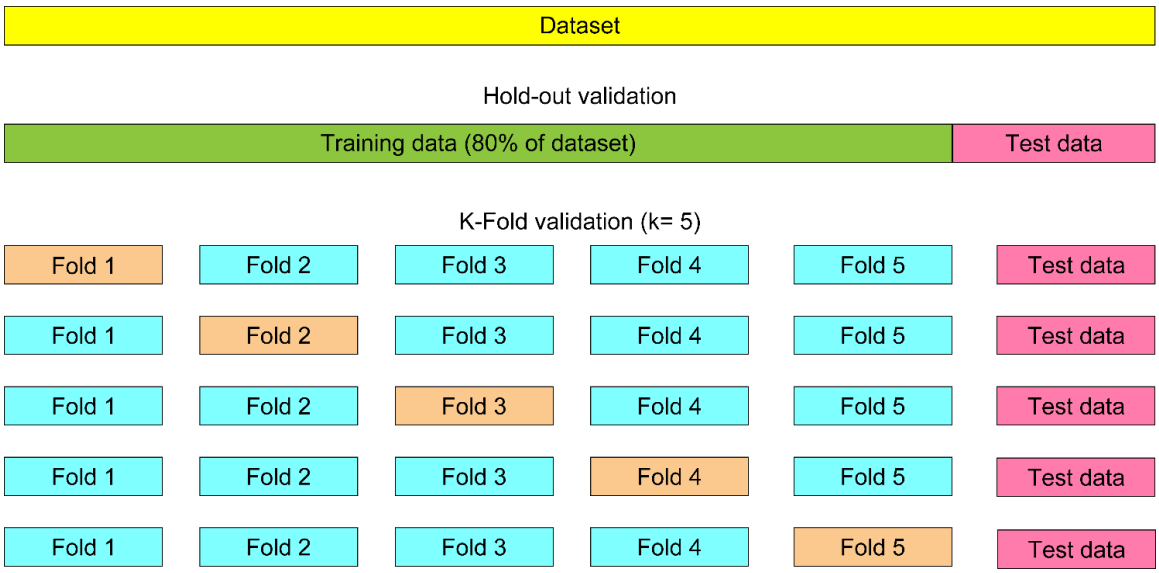


Figure S2. Schematic of hold-out and K-fold validations. Hold-out validation involves splitting a dataset into a training set and a test set. The model training is performed using training data, and then its performance is evaluated on test data. K-fold validation involves splitting a dataset into k equal-sized folds (in example, k = 5). The model was trained using four folds, and the remaining fold is used for model validation, which process was repeated 5 times with a different fold used for validation. Consequently, all samples within the training dataset are used exactly once as validation data. Regression analysis of the test data was performed with these models, and the average results were employed as the final result.

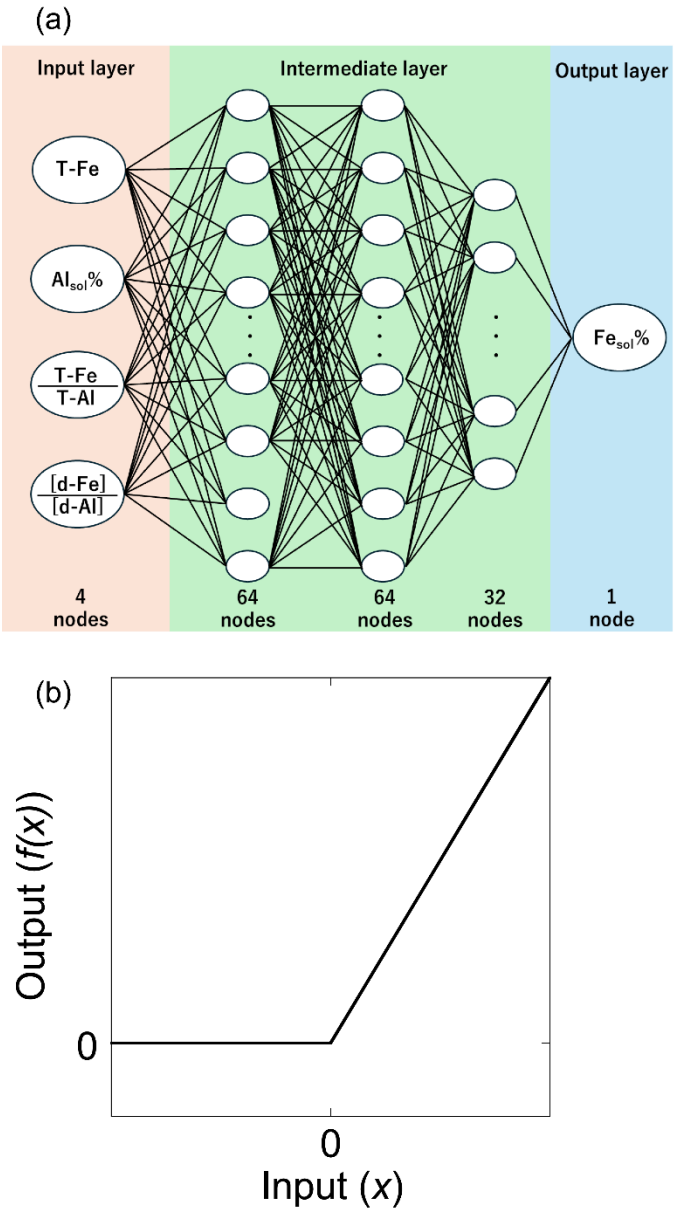


Figure S3. (a) Structure of DNN, (b) A figure of the rectified linear unit (ReLU). "If the input value (x) is less than 0, return 0 as f(x). If x is greater than 0, then f(x) = x.

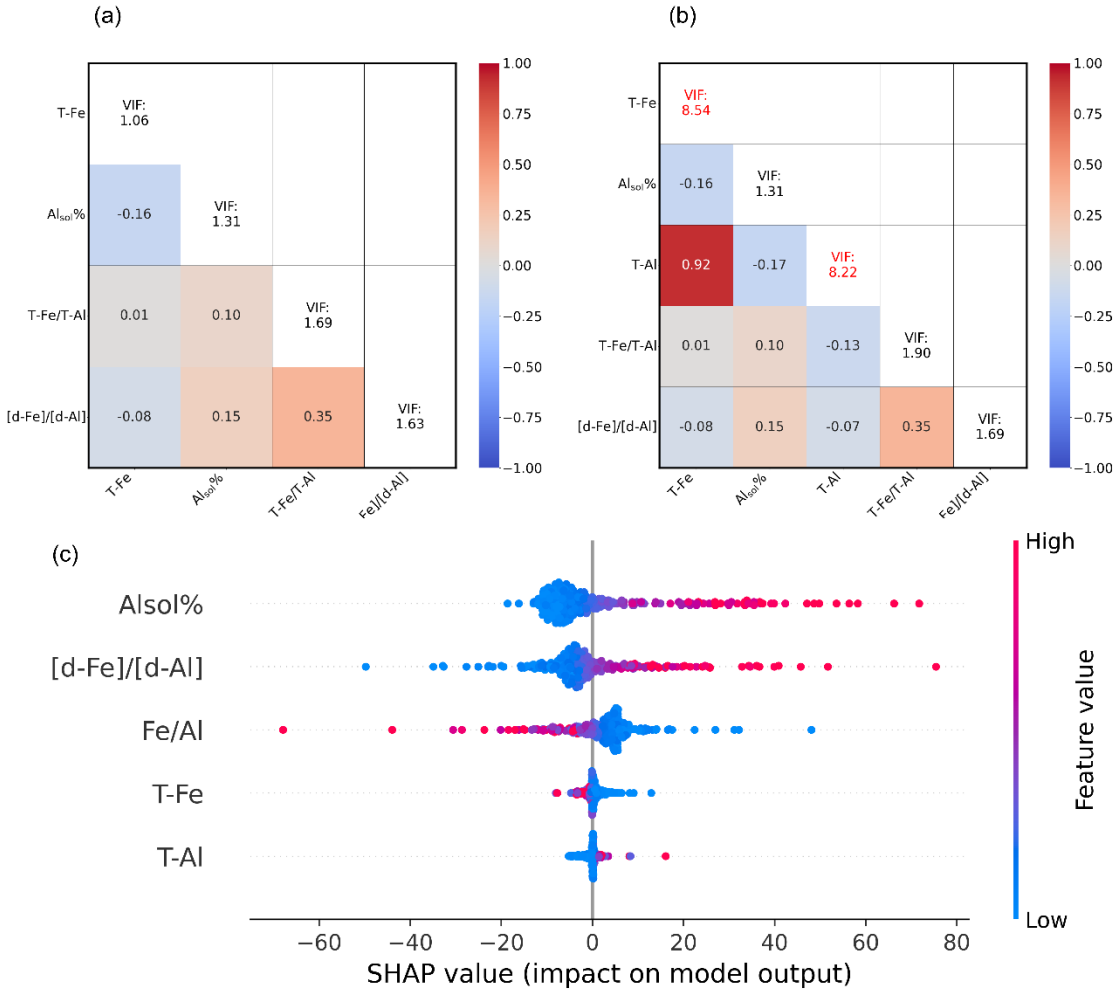
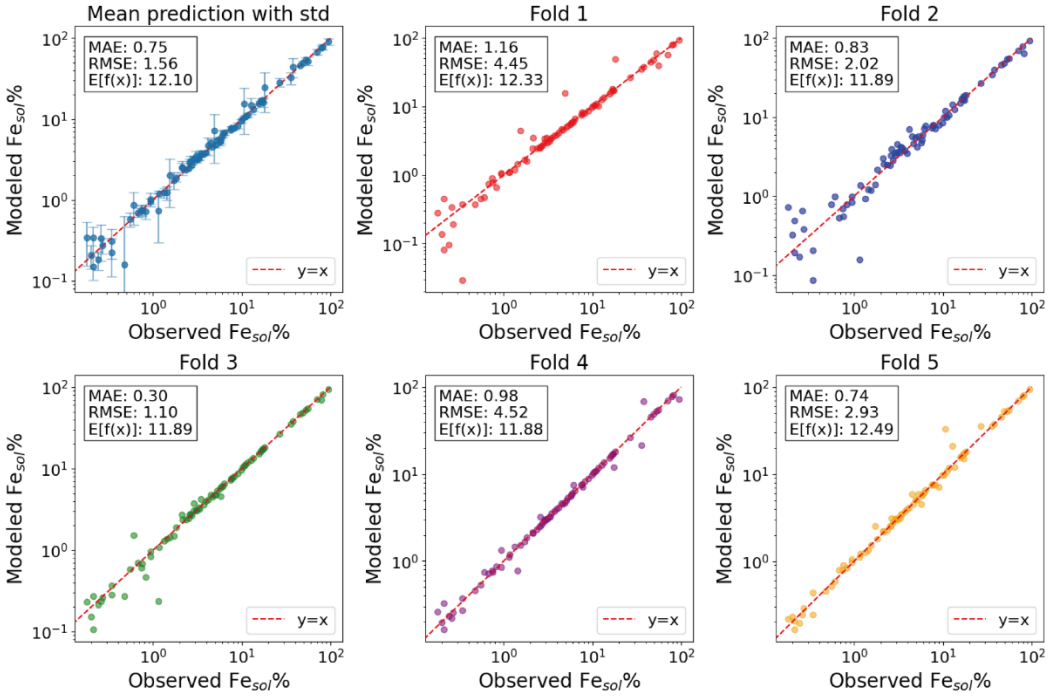


Figure S3. Correlation matrices and VIF values for the two feature sets: (a) the dataset excluding T-Al concentration (four features) and (b) the dataset of East Asian aerosol including T-Al concentration (five features). Colar bar shows Pearson's correlation factor. (c) SHAP value for each feature of Fe_{sol}% reproducing when T-Al is used as the feature.

(a) XGBoost



(b) DNN

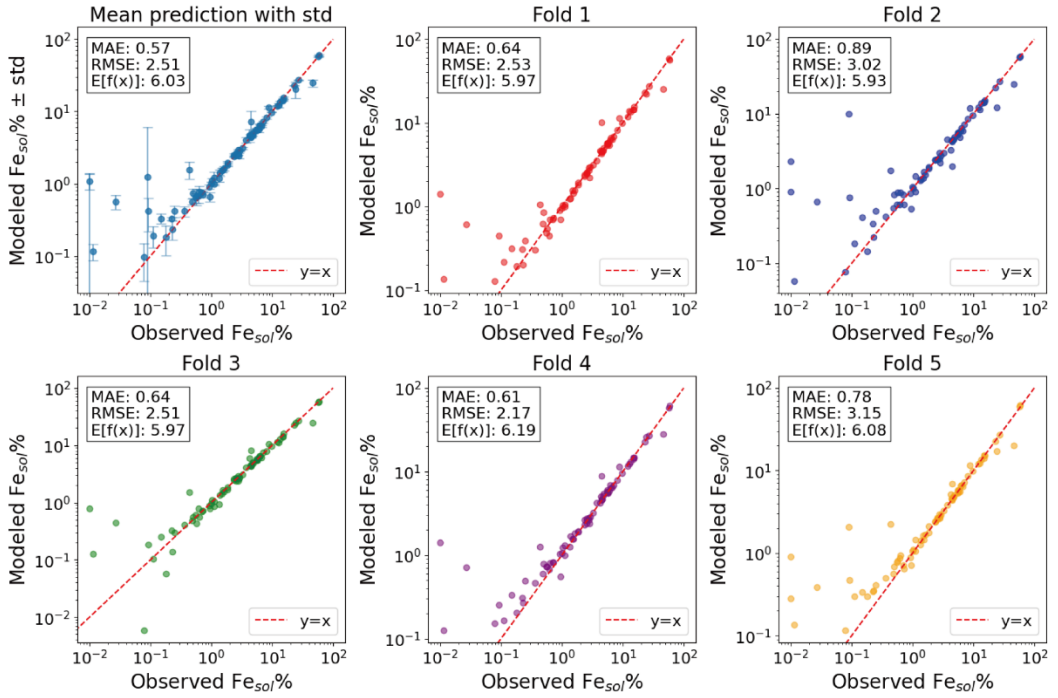


Figure S4. Mean and fold-specific reproduced values of Fe_{sol}% in East Asia from (a) XGBoost, (b) from DNN.

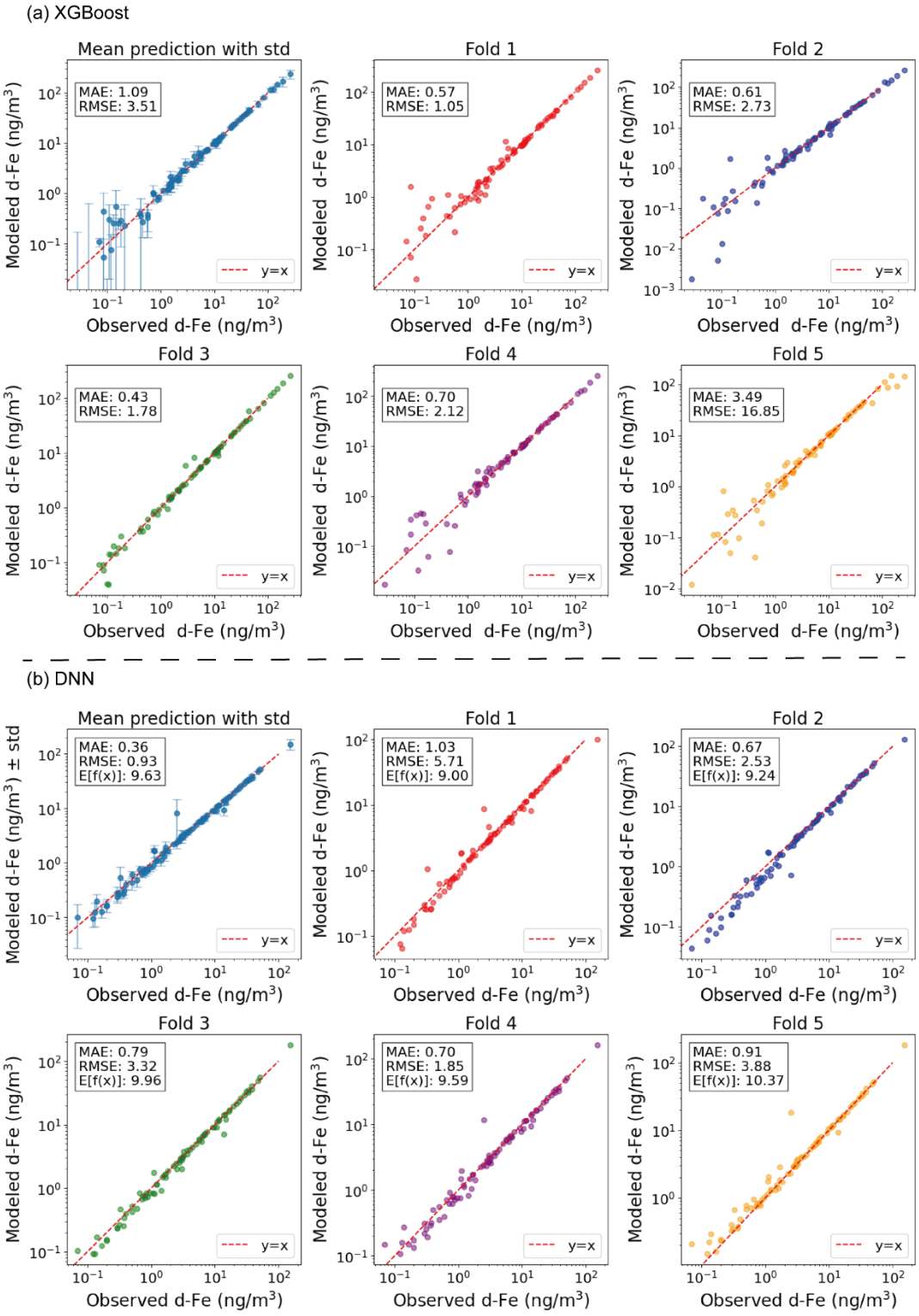


Figure S5. Mean and fold-specific reproduced values of d-Fe concentration in East Asia from (a) XGBoost, (b) from DNN.

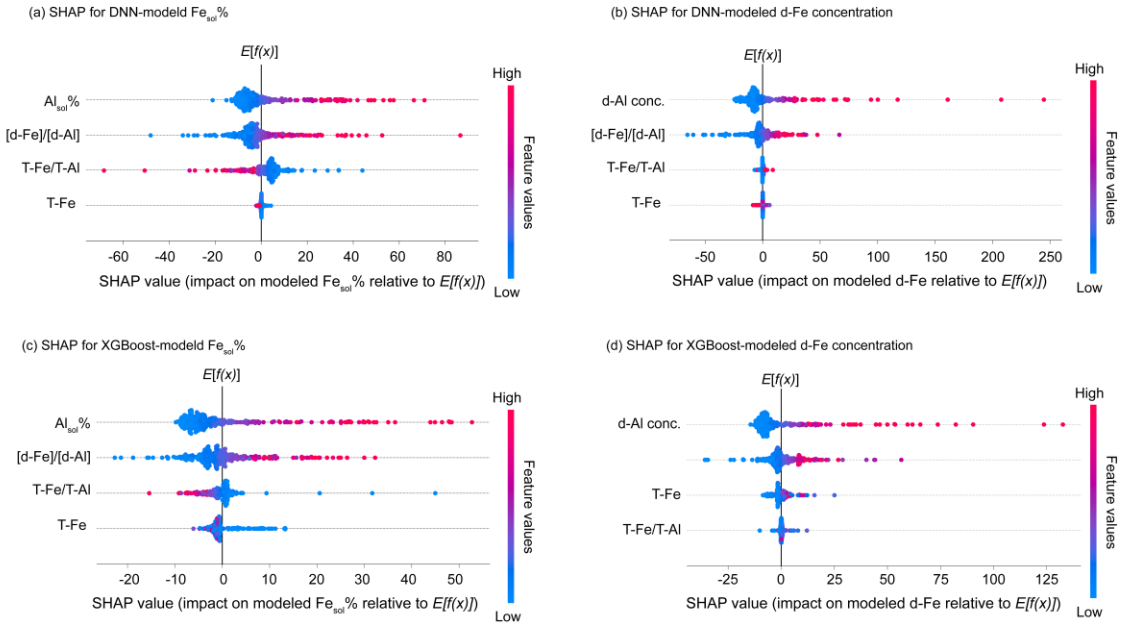


Figure S6. SHAP values of Fe_{sol}% in the test dataset for (a) DNN and (b) XGBoost, and d-Fe concentration for (c) DNN and (d) XGBoost.

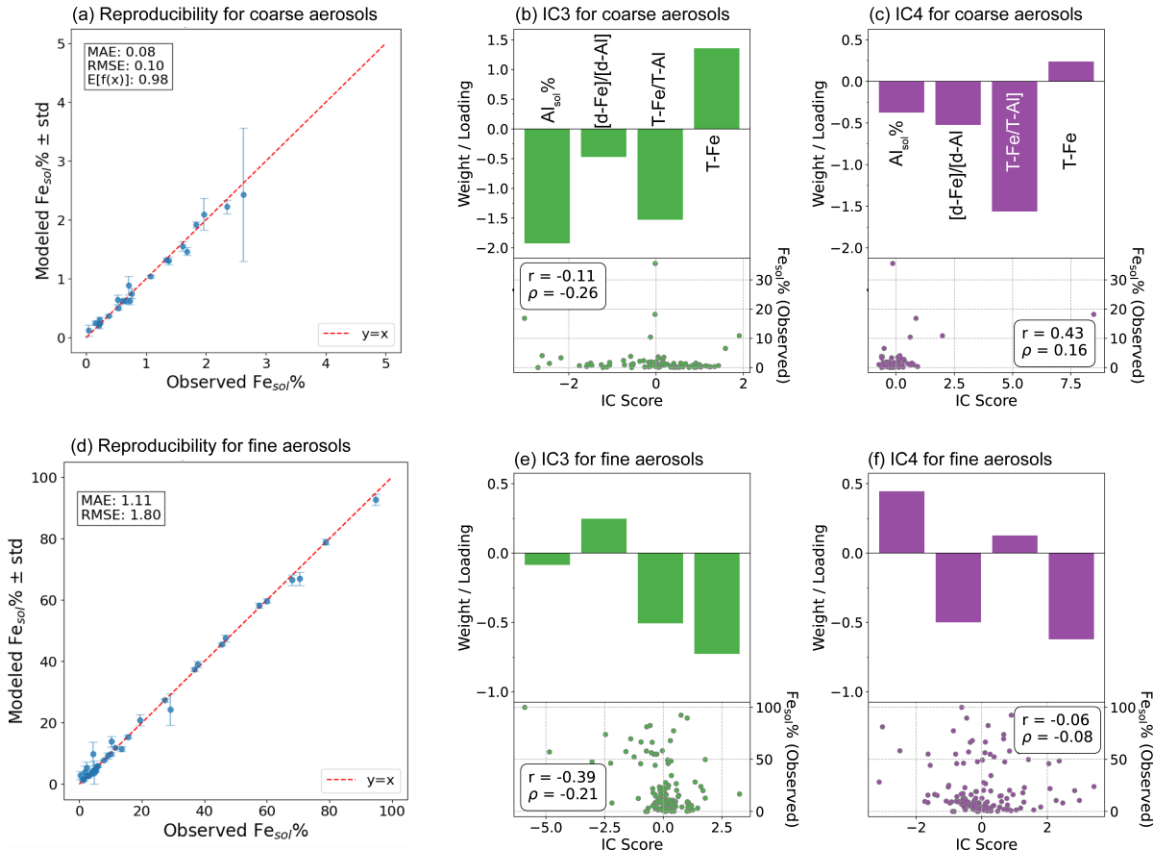


Figure S7. (a) Reproducibility of Fe_{sol}% in coarse aerosol particles by DNN. Characteristics of (b) IC3 and (c) IC4 derived from ICA of the SHAP value matrix for Fe_{sol}% in coarse aerosol particles. The upper bar plot indicates the feature loadings, and the lower scatter plot shows the relationship between IC score and observed Fe_{sol}%. Pearson's correlation coefficient (r) and Spearman's rank correlation coefficient (ρ) are shown in each panel. Due to low r and ρ, these IC components less influence on Fe_{sol}%. (d-f) same for fine aerosol particles.

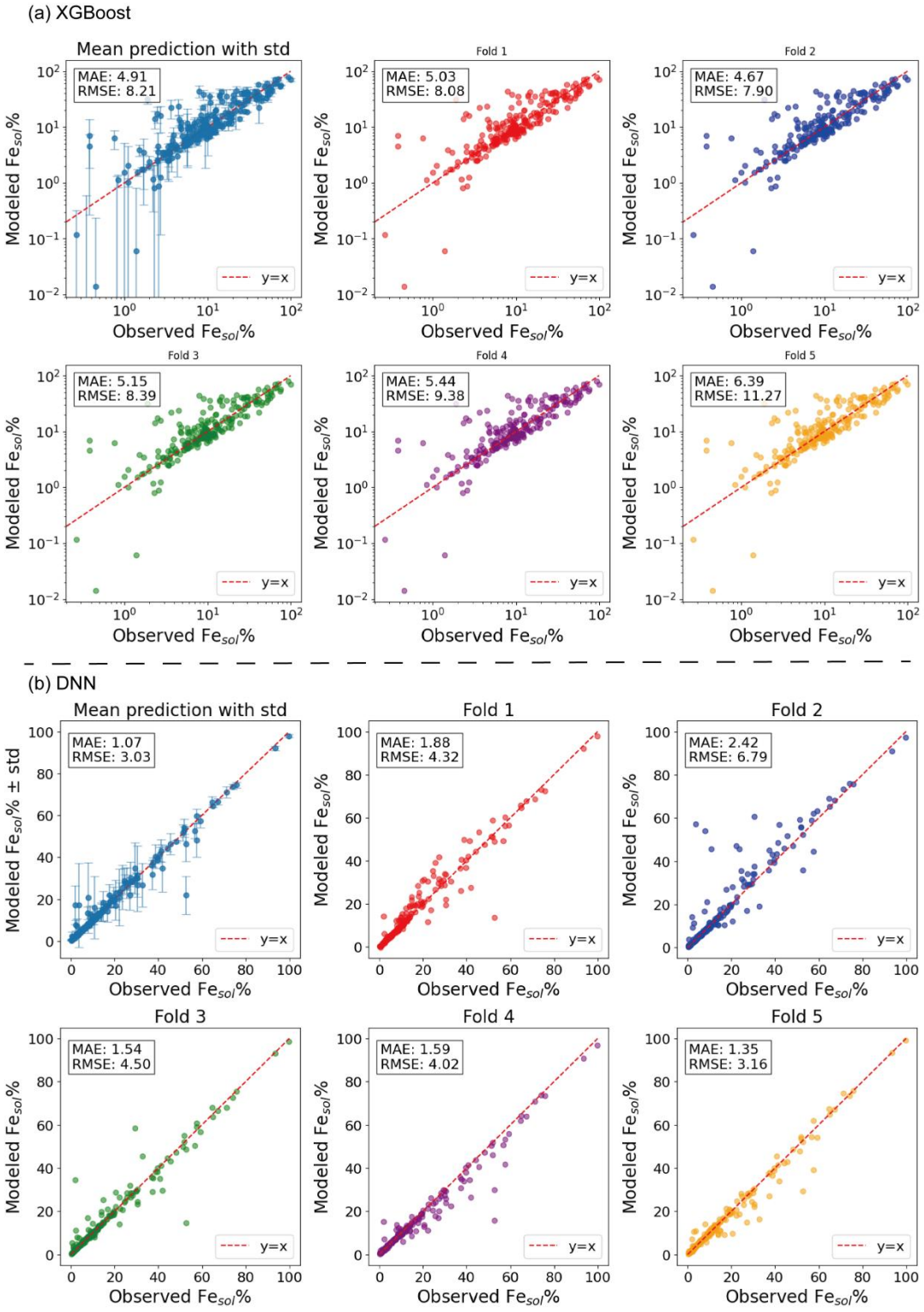


Figure S8. Mean and fold-specific reproduced values of Fe_{sol}% in the North Pacific aerosols from (a) XGBoost, (b) from DNN.

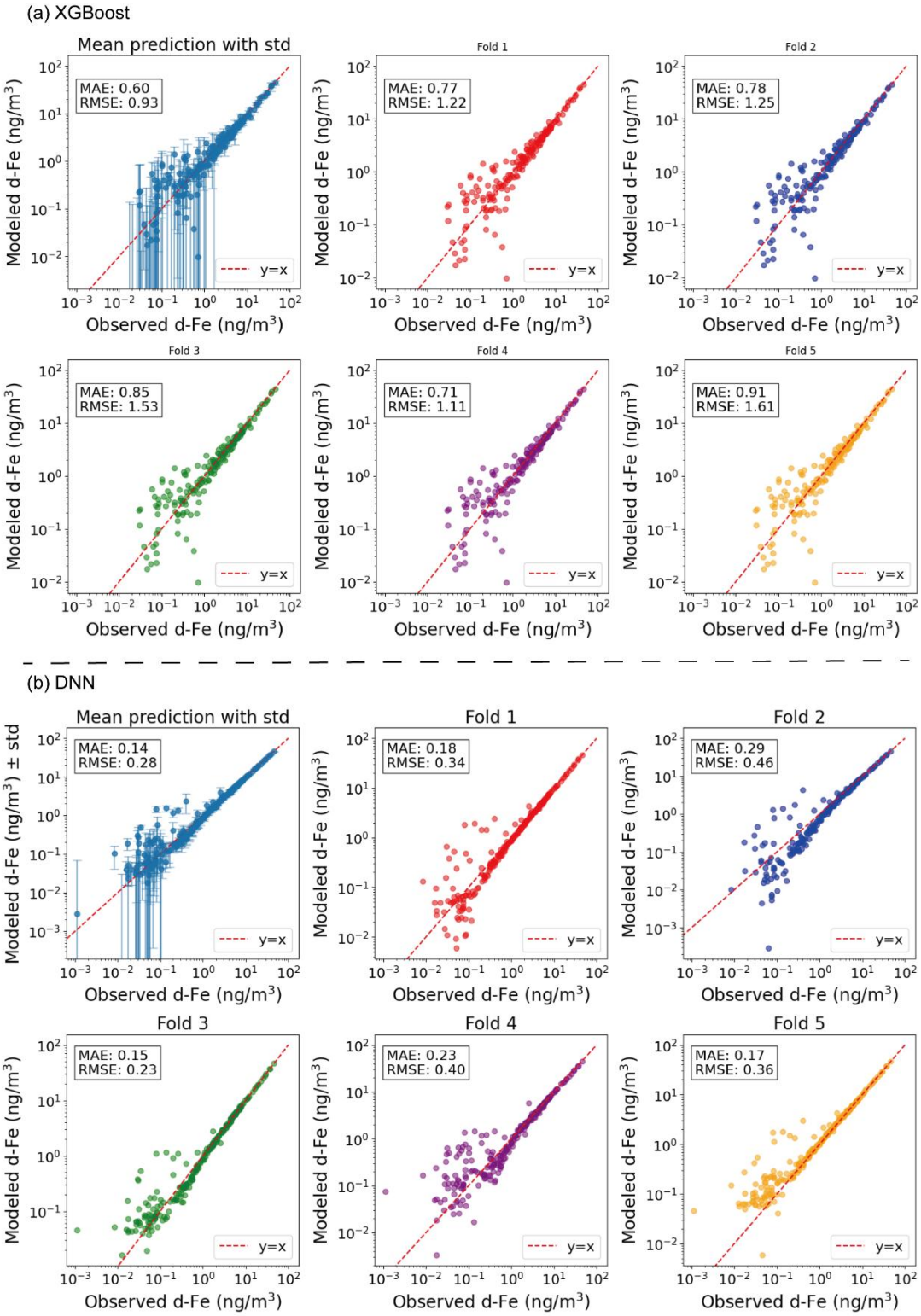


Figure S9. Mean and fold-specific reproduced values of d-Fe concentrations in the North Pacific aerosols from (a) XGBoost, (b) from DNN.

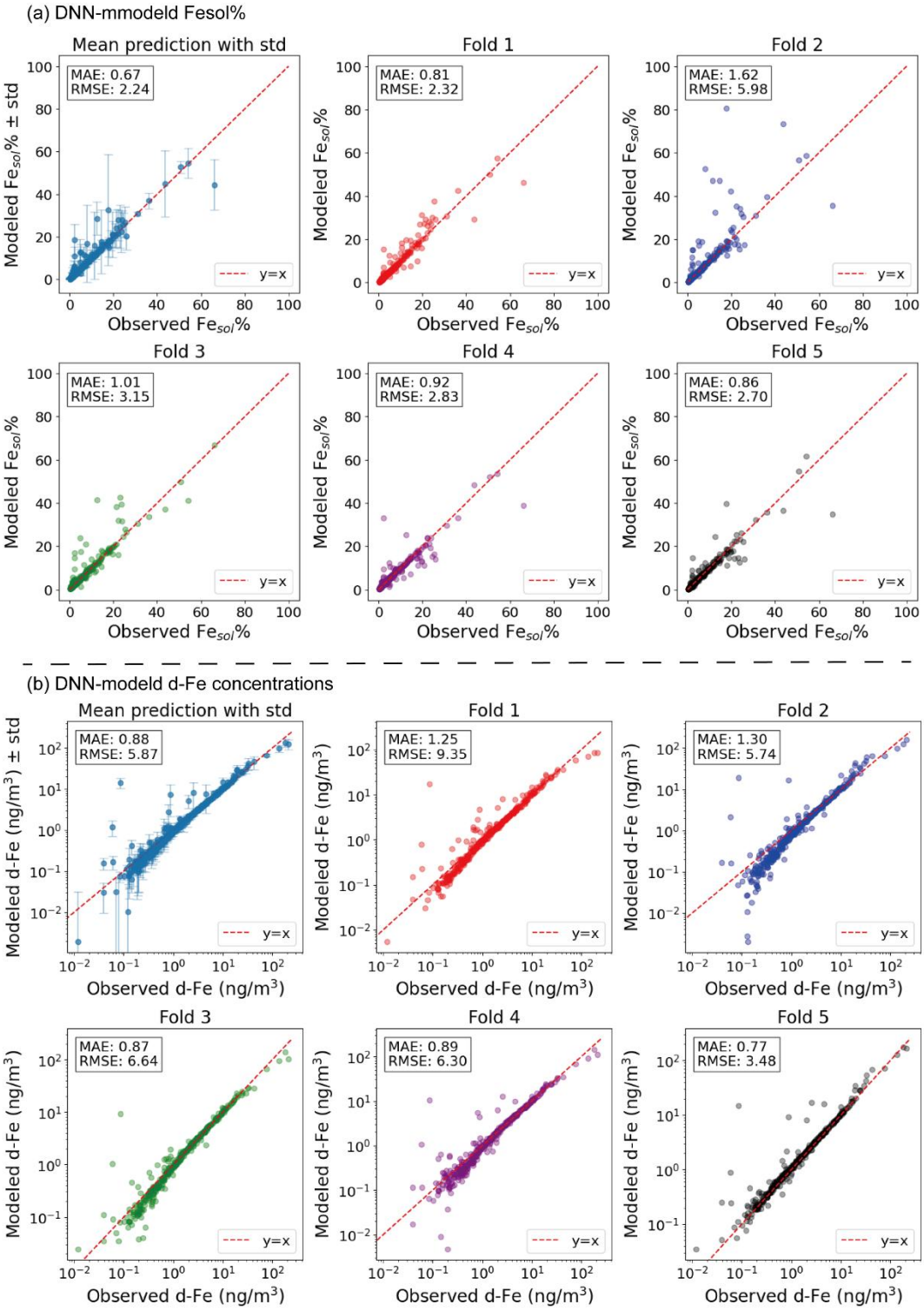


Figure S10. Mean and fold-specific reproduced values of (a) Fe_{sol}% and (b) d-Fe concentrations in Atlantic aerosols.

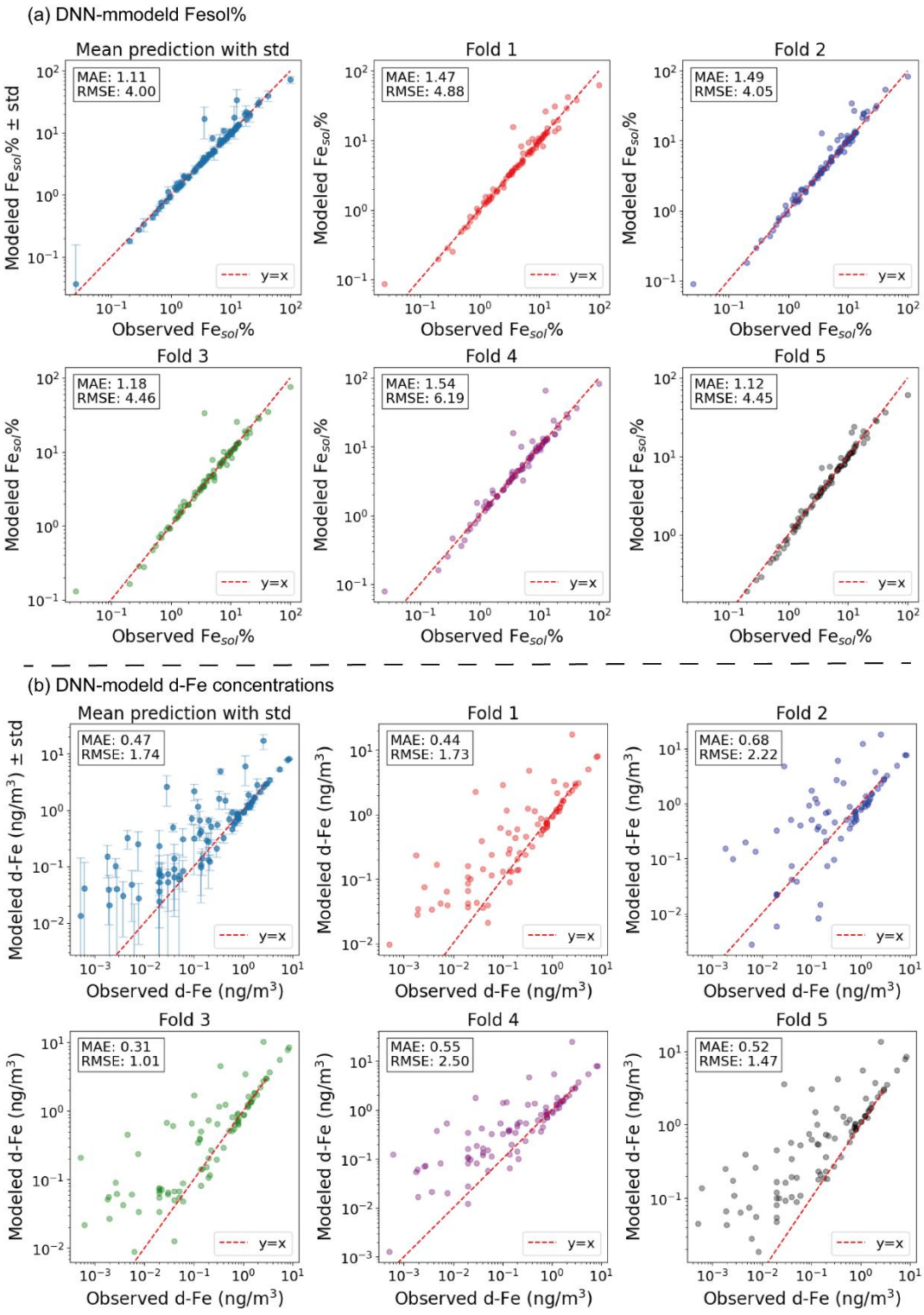


Figure S11. Mean and fold-specific reproduced values of (a) Fe_{sol}% and (b) d-Fe concentrations in the South Pacific aerosols.

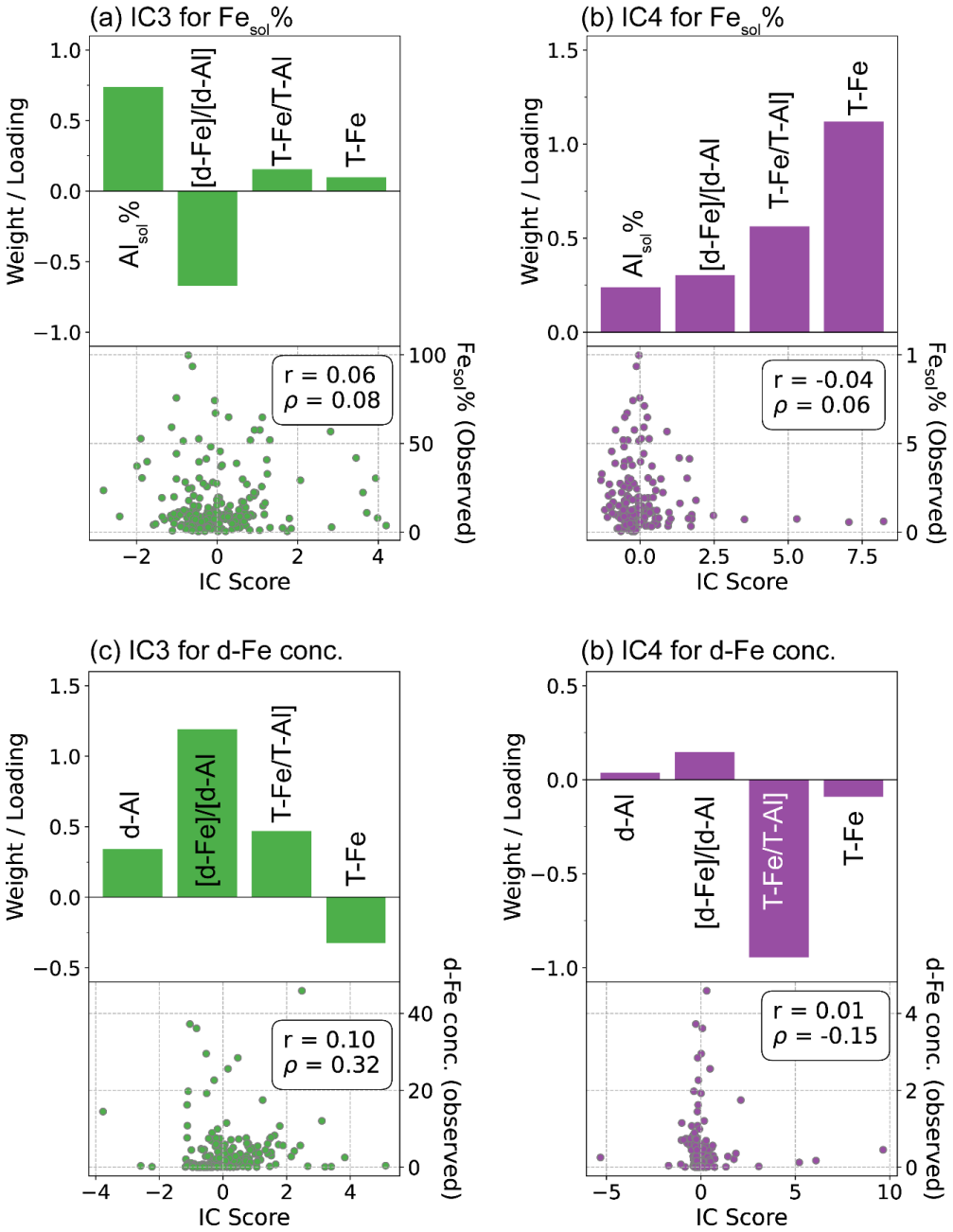


Figure S12. Characteristics of IC3 and IC4 derived from ICA of the SHAP value matrix for Fe_{sol}% and d-Fe concentrations in the North Pacific aerosols: (a) IC3 for Fe_{sol}%, (b) IC4 for Fe_{sol}%, (c) IC3 ford-Fe concentration, and (d) IC4 for d-Fe concentration. In each panel, the upper bar plot indicates the feature loadings, and the lower scatter plot shows the relationship between IC score and observed Fe_{sol}%. Pearson's correlation coefficient (r) and Spearman's rank correlation coefficient (ρ) are shown in each panel. Due to low r and ρ, these IC components less influence on Fe_{sol}%.

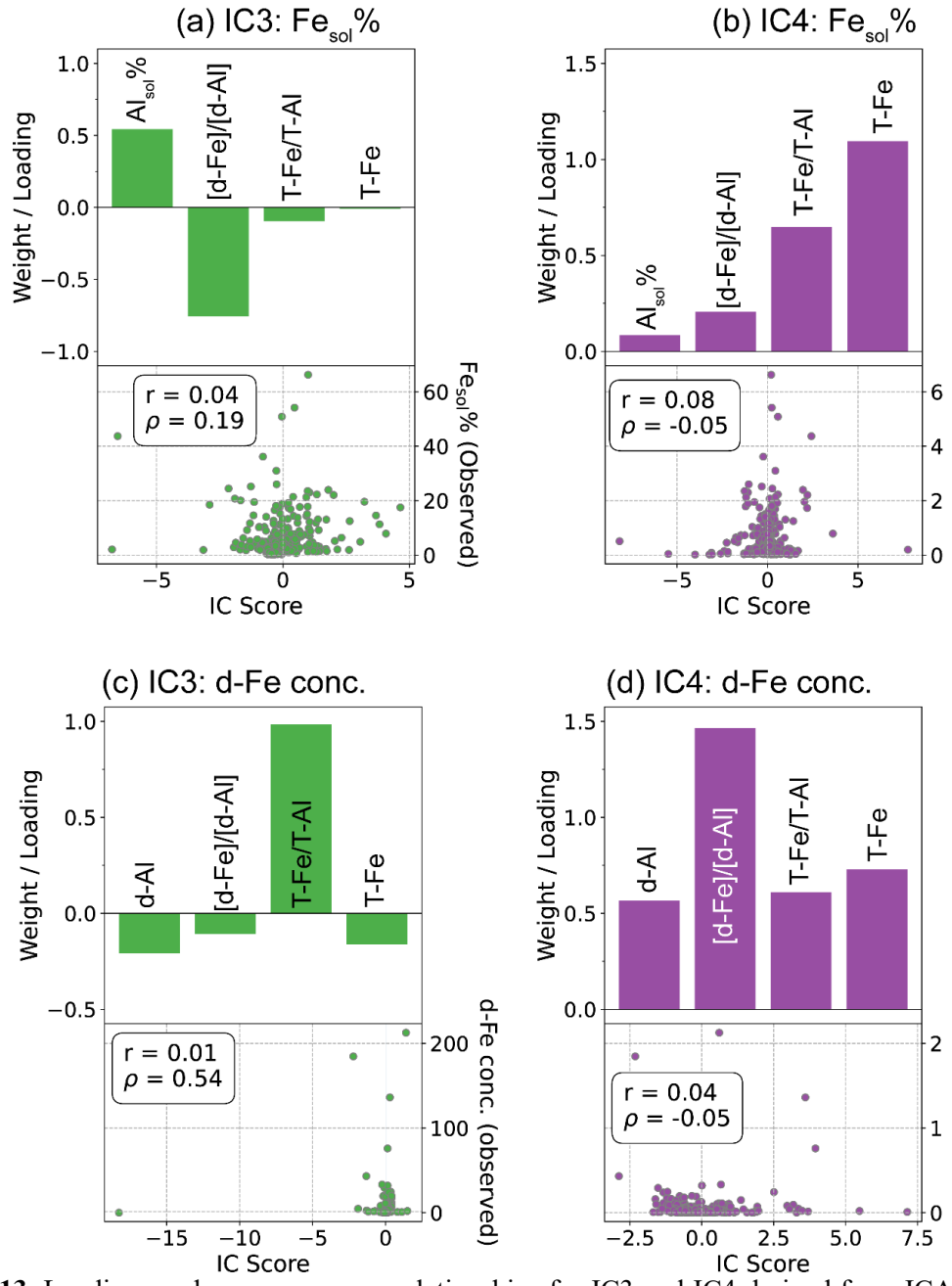


Figure S13. Loadings and score–response relationships for IC3 and IC4 derived from ICA of the SHAP value matrix for Fe_{sol}% and d-Fe concentrations in the Atlantic aerosols: (a) IC3 for Fe_{sol}%, (b) IC4 for Fe_{sol}%, (c) IC3 ford-Fe concentration, and (d) IC4 for d-Fe concentration. In each panel, the upper bar plot indicates the feature loadings (Al_{sol}%, [d-Fe]/[d-Al], T-Fe/T-Al, and T-Fe), and the lower scatter plot shows the relationship between IC score and observed Fe_{sol}%. Pearson's correlation coefficient (r) and Spearman's rank correlation coefficient (ρ) are shown in each panel. Due to low r and ρ, these IC components less influence on Fe_{sol}%.

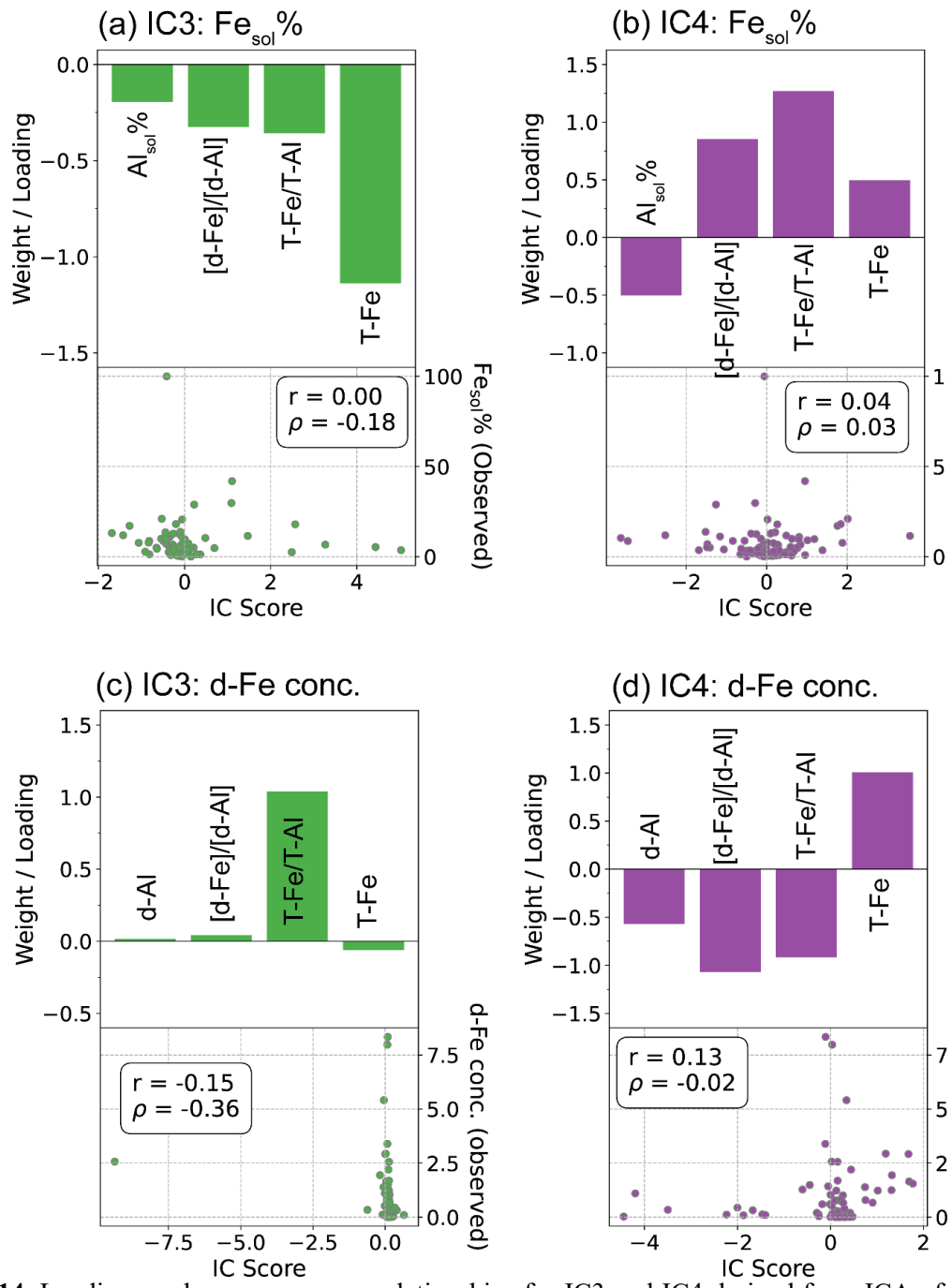


Figure S14. Loadings and score–response relationships for IC3 and IC4 derived from ICA of the SHAP value matrix for Fe_{sol}% and d-Fe concentrations in the South Pacific aerosols: (a) IC3 for Fe_{sol}%, (b) IC4 for Fe_{sol}%, (c) IC3 ford-Fe concentration, and (d) IC4 for d-Fe concentration. In each panel, the upper bar plot indicates the feature loadings (Al_{sol}%, [d-Fe]/[d-Al], T-Fe/T-Al, and T-Fe), and the lower scatter plot shows the relationship between IC score and observed Fe_{sol}%. Pearson's correlation coefficient (r) and Spearman's rank correlation coefficient (ρ) are shown in each panel. Due to low r and ρ, these IC components less influence on Fe_{sol}%.