Supplement of Geosci. Model Dev., 16, 6247–6266, 2023
https://doi.org/10.5194/gmd-16-6247-2023-supplement

Geoscientific Model Development
Open Access
EGU

Supplement of

A robust error correction method for numerical weather prediction wind speed based on Bayesian optimization, variational mode decomposition, principal component analysis, and random forest: VMD-PCA-RF (version 1.0.0)

Shaohui Zhou et al.

Correspondence to: Chloe Yuchao Gao (gyc@fudan.edu.cn)

Table S1. The best hyper-parameters of the models

Model	parameters
VMD-PCA-lightGBM	'max_depth' : 28, 'min_child_samples' : 30, 'n_estimators' : 436, 'num_leaves' : 287
VMD-PCA-XGBoost	'gamma' : 1, 'max_depth' : 19, 'min_child_weight' : 1, 'n_estimators': 408
VMD-PCA-RF	'max_depth' : 31, 'max_features' : 14, 'min_samples_leaf' : 28, 'min_samples_split' : 3, 'n_estimators' : 371
VMD-PCA-DBN	'input_length' : 20, 'output_length' : 1, 'loss_function' : 'MSE', 'optimizer' : 'Adam', 'hidden_units' : [400, 200], 'batch_size' :20000, 'epoch_pretrain' : 100, 'epoch_finetune' : 200
VMD-PCA-MLP	'batch_size' : 10114, 'hidden_layer_sizes' : 305, 'max_iter' : 386
lightGBM	'max_depth' : 21, 'min_child_samples' : 19, 'n_estimators' : 312, 'num_leaves' : 297
XGBoost	'gamma' : 0, 'max_depth' : 21, 'min_child_weight' : 9, 'n_estimators': 299
RF	'max_depth' : 40, 'max_features' : 12, 'min_samples_leaf' : 23, 'min_samples_split' : 2, 'n_estimators' : 440
DBN	'input_length' : 12, 'output_length' : 1, 'loss_function' : 'MSE', 'optimizer' : 'Adam', 'hidden_units' : [400, 200], 'batch_size' : 20000, 'epoch_pretrain' : 100, 'epoch_finetune' : 200
MLP	'batch_size' : 10232, 'hidden_layer_sizes' : 494, 'max_iter' : 311

Table S2. Table of evaluation indices of wind speed error predicted by 10 models in December 2021

Model	MAE（m s^{-1}）	RMSE（m s^{-1}）	rMAE（%）	rRMSE（%）	FA（%）	R
VMD-PCA-lightGBM	0.47	0.63	37.67	51.25	91.13	0.81
VMD-PCA-XGBoost	0.49	0.68	39.84	54.82	89.22	0.78
VMD-PCA-RF	0.46	0.62	37.36	50.39	91.79	0.82
VMD-PCA-DBN	0.53	0.75	43.32	61.13	87.93	0.71

VMD-PCA-MLP	0.53	0.72	43.04	58.47	87.2	0.75
lightGBM	0.49	0.67	39.59	54.16	89.68	0.79
XGBoost	0.51	0.70	41.51	56.64	87.9	0.77
RF	0.48	0.65	38.80	52.32	90.64	0.81
DBN	0.56	0.77	45.25	62.46	86.74	0.71
MLP	0.55	0.74	44.65	60.1	86.08	0.75

Table S3. Table of evaluation indices of wind speed error predicted by 10 models in January 2022

Model	MAE（m s^{-1}）	RMSE（m s^{-1}）	rMAE（%）	rRMSE（%）	FA（%）	R
VMD-PCA-lightGBM	0.45	0.63	34.75	48.65	91.49	0.78
VMD-PCA-XGBoost	0.47	0.66	36.31	51.01	90.23	0.76
VMD-PCA-RF	0.46	0.64	35.06	49.00	91.57	0.78
VMD-PCA-DBN	0.53	0.75	40.96	57.49	87.61	0.67
VMD-PCA-MLP	0.50	0.69	38.46	53.16	88.94	0.73
lightGBM	0.46	0.64	35.24	49.34	91.11	0.77
XGBoost	0.48	0.67	36.68	51.38	89.88	0.75
RF	0.46	0.64	35.18	49.13	91.36	0.78
DBN	0.53	0.74	40.97	56.86	87.71	0.68
MLP	0.49	0.68	37.83	52.26	89.57	0.74

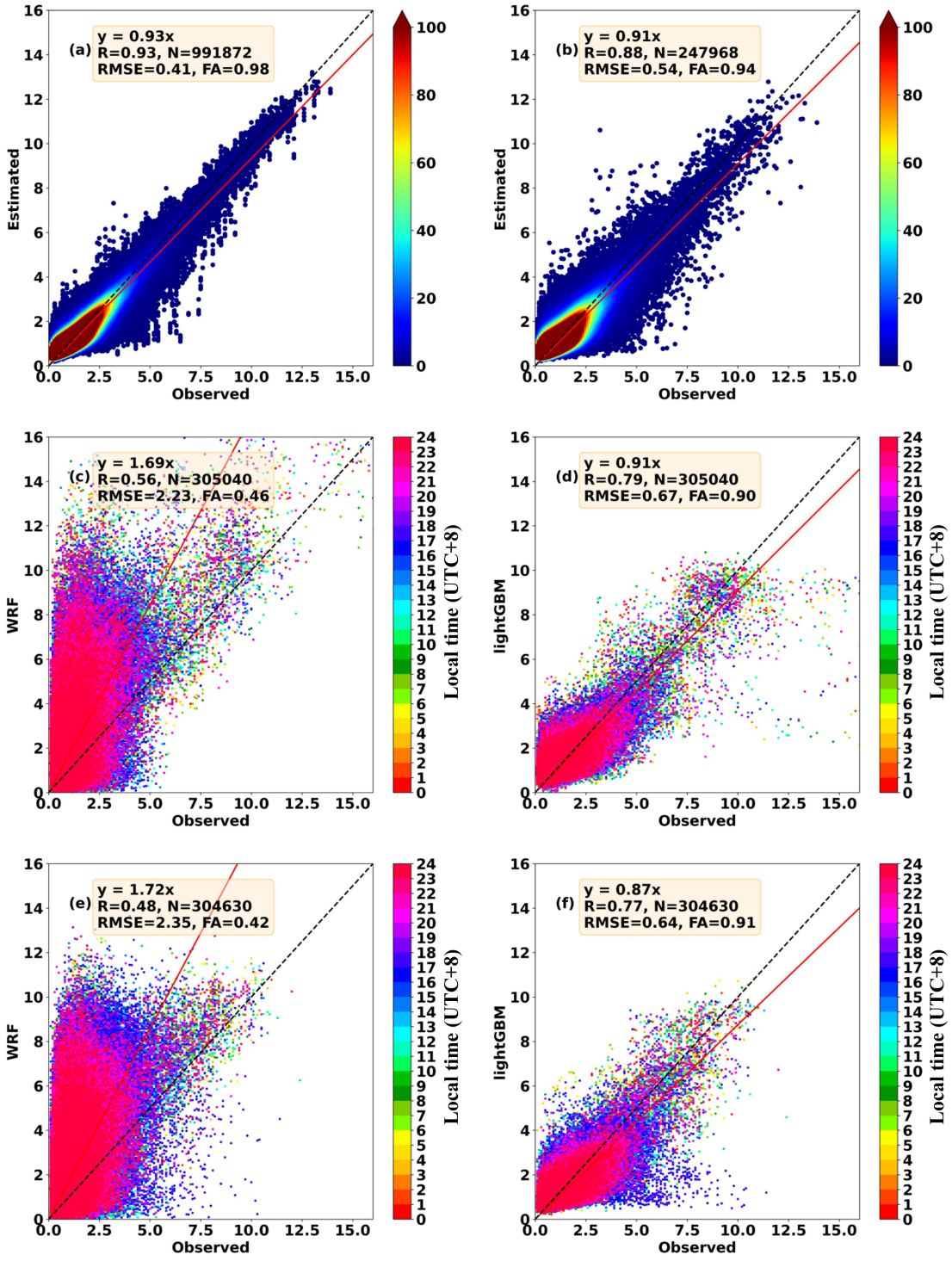

Figure S1. The scatter density map compared with the actual 10-meter wind speed: (a) 10-fold cross-validation training set of lightGBM model in February 2022, (b) 10-fold cross-validation validation set of lightGBM model in February 2022. The 24-hour scatter map compared with the actual 10-meter wind speed: (c) WRF forecasts in December 2021, (d) lightGBM model forecasts in December 2021, (e) WRF forecasts in January 2022, and (f) lightGBM model forecasts in January 2022.

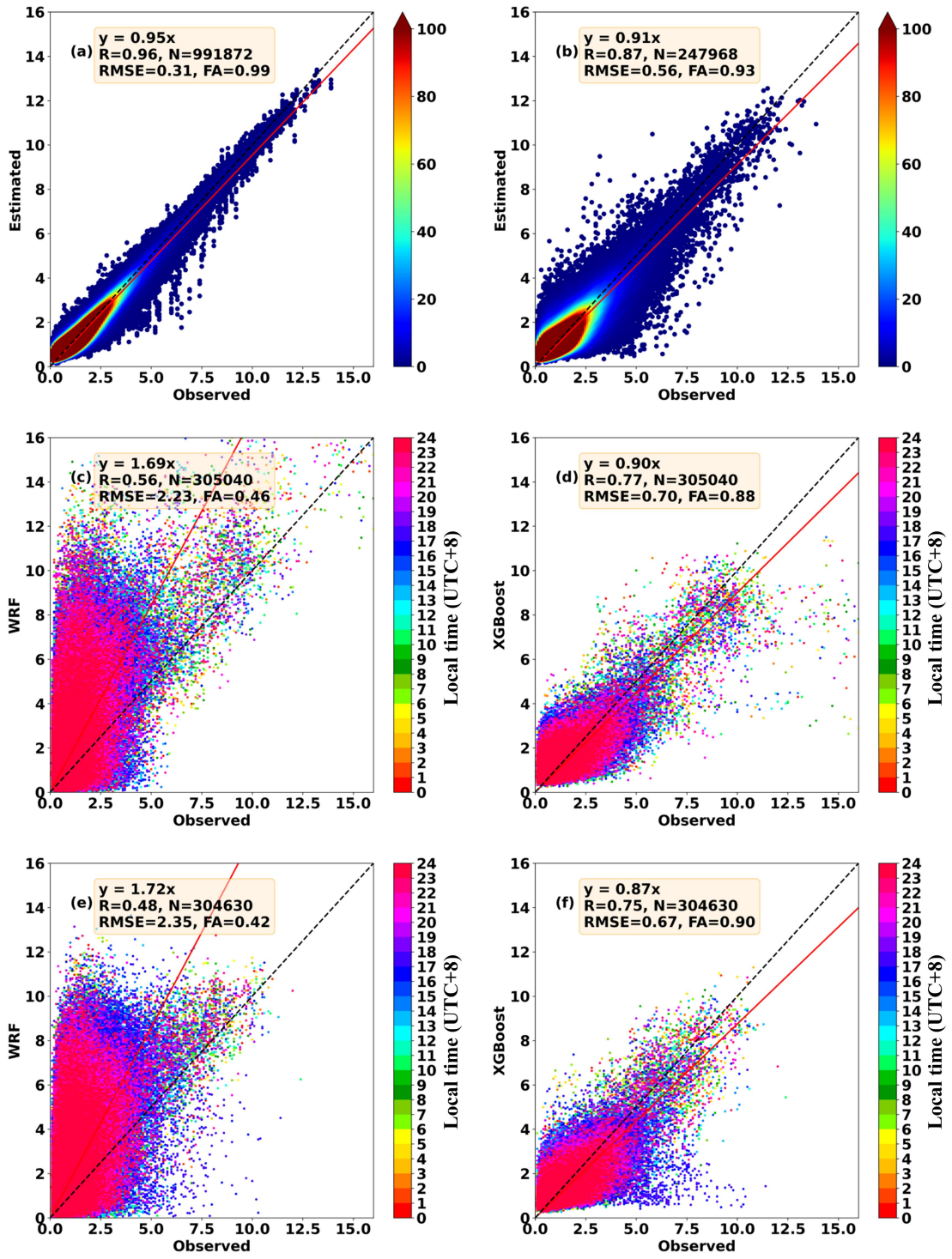

Figure S2. The scatter density map compared with the actual 10-meter wind speed: (a) 10-fold cross-validation training set of XGBoost model in February 2022, (b) 10-fold cross-validation validation set of XGBoost model in February 2022. The 24-hour scatter map compared with the actual 10-meter wind speed: (c) WRF forecasts in December 2021, (d) XGBoost model forecasts in December 2021, (e) WRF forecasts in January 2022, and (f) XGBoost model forecasts in January 2022.

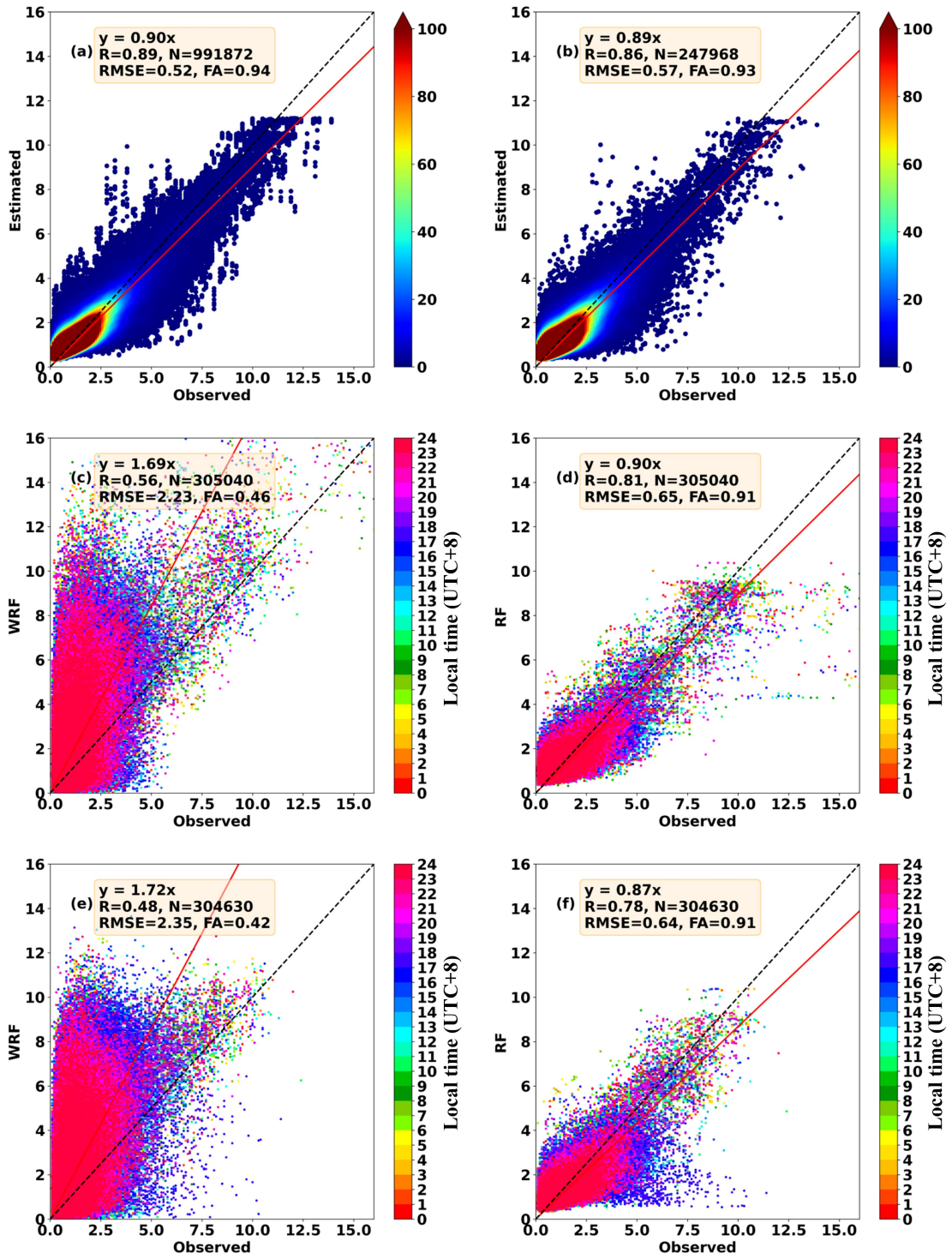

Figure S3. The scatter density map compared with the actual 10-meter wind speed: (a) 10-fold cross-validation training set of RF model in February 2022, (b) 10-fold cross-validation validation set of RF model in February 2022. The 24-hour scatter map compared with the actual 10-meter wind speed: (c) WRF forecasts in December 2021, (d) RF model forecasts in December 2021, (e) WRF forecasts in January 2022, and (f) RF model forecasts in January 2022.

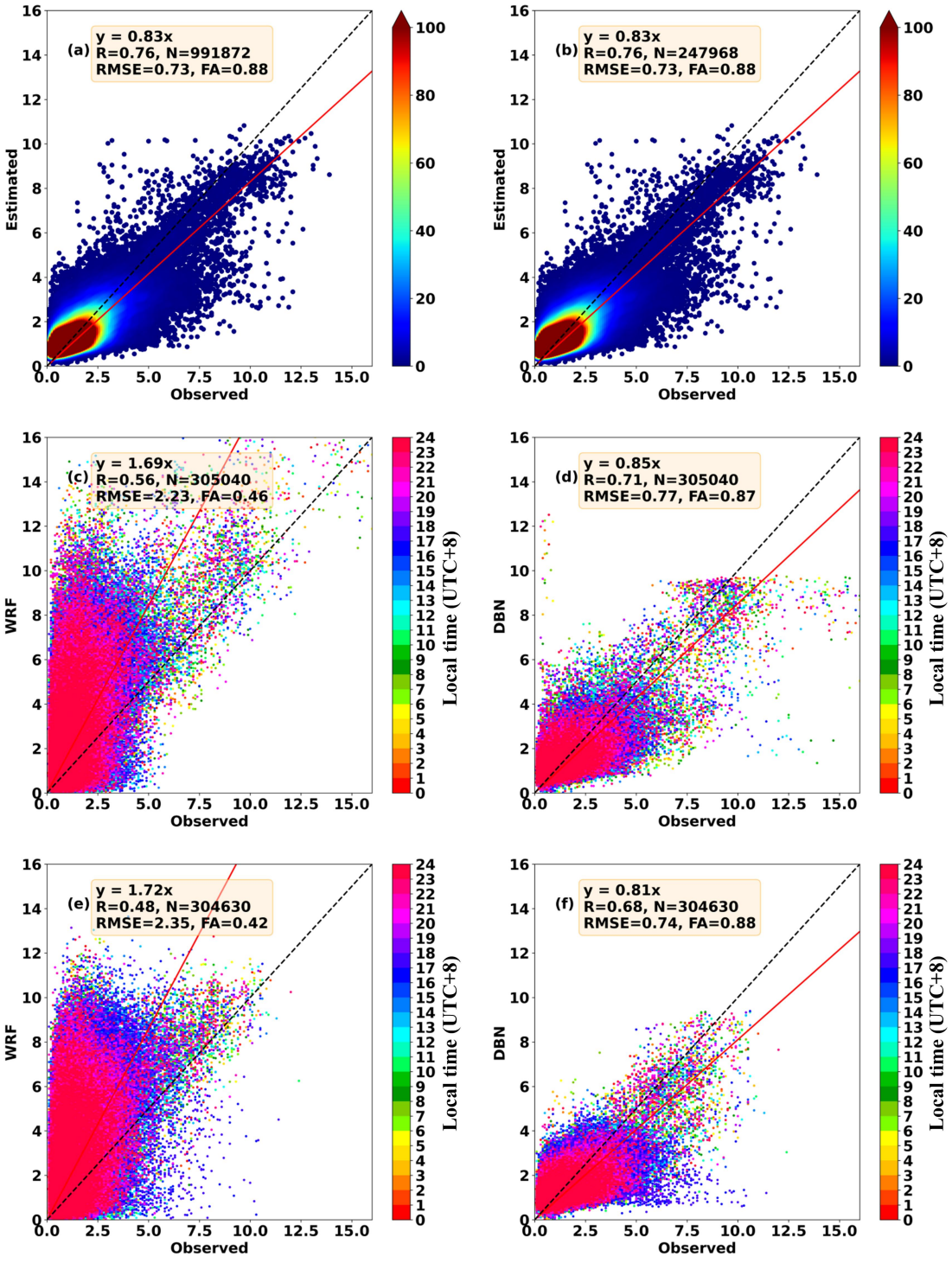

Figure S4. The scatter density map compared with the actual 10-meter wind speed: (a) 10-fold cross-validation training set of DBN model in February 2022, (b) 10-fold cross-validation validation set of DBN model in February 2022. The 24-hour scatter map compared with the actual 10-meter wind speed: (c) WRF forecasts in December 2021, (d) DBN model forecasts in December 2021, (e) WRF forecasts in January 2022, and (f) DBN model forecasts in January 2022.

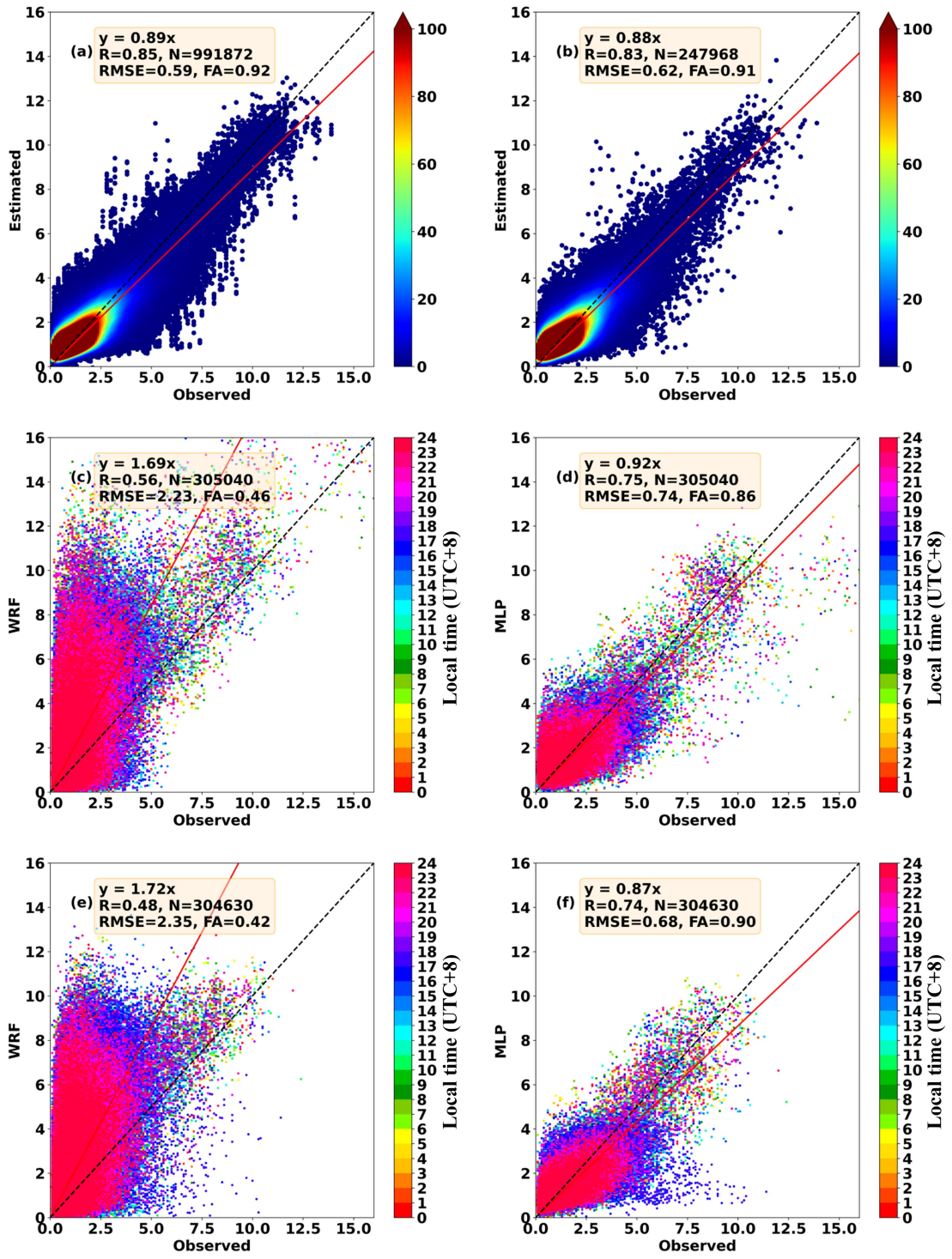

Figure S5. The scatter density map compared with the actual 10-meter wind speed: (a) 10-fold cross-validation training set of MLP model in February 2022, (b) 10-fold cross-validation validation set of MLP model in February 2022. The 24-hour scatter map compared with the actual 10-meter wind speed: (c) WRF forecasts in December 2021, (d) MLP model forecasts in December 2021, (e) WRF forecasts in January 2022, and (f) MLP model forecasts in January 2022.

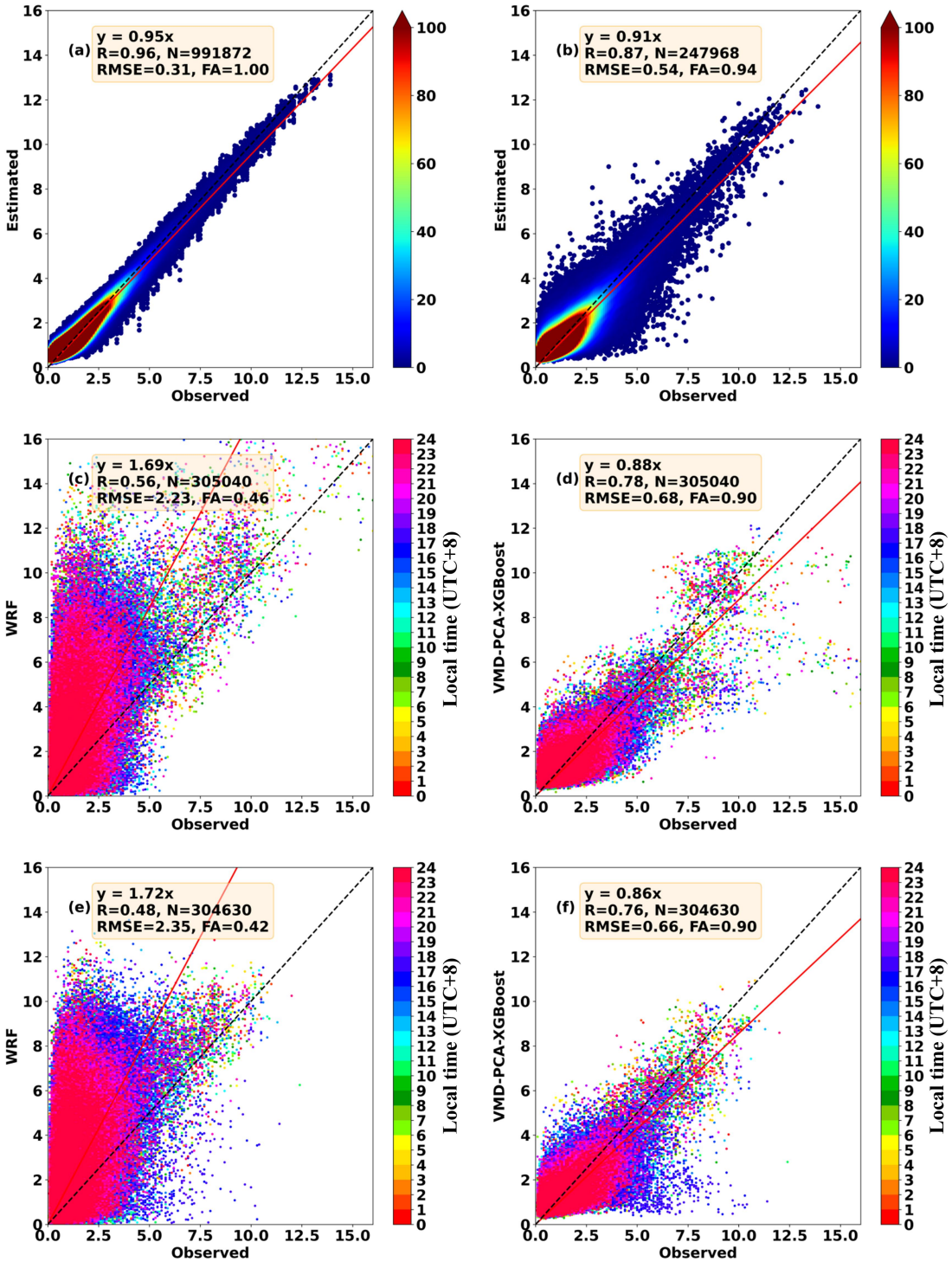

Figure S6. The scatter density map compared with the actual 10-meter wind speed: (a) 10-fold cross-validation training set of VMD-PCA-XGBoost model in February 2022, (b) 10-fold cross-validation validation set of VMD-PCA-XGBoost model in February 2022. The 24-hour scatter map compared with the actual 10-meter wind speed: (c) WRF forecasts in December 2021, (d) VMD-PCA-XGBoost model forecasts in December 2021, (e) WRF forecasts in January 2022, and (f) VMD-PCA-XGBoost model forecasts in January 2022.

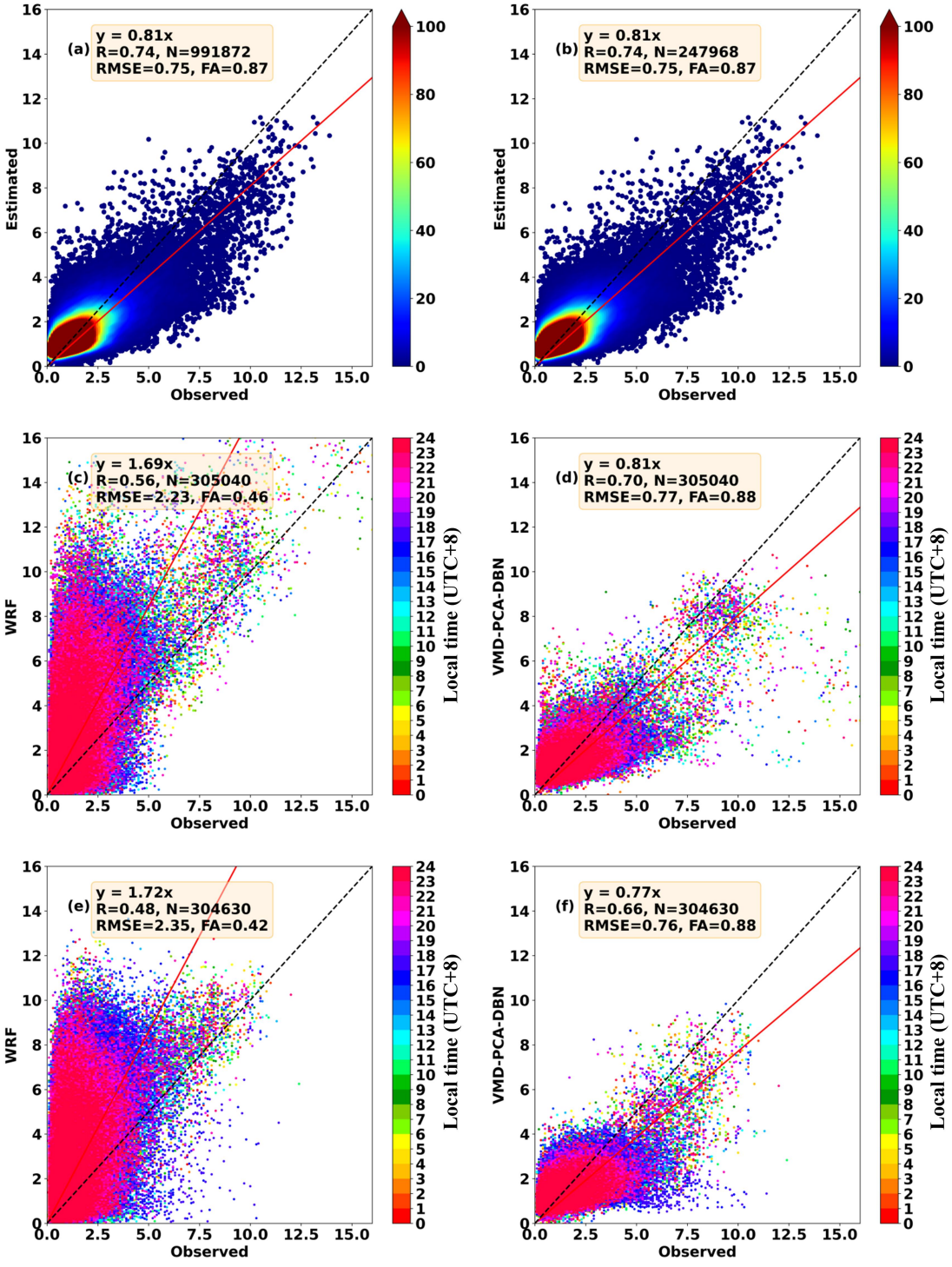

Figure S7. The scatter density map compared with the actual 10-meter wind speed: (a) 10-fold cross-validation training set of VMD-PCA-DBN model in February 2022, (b) 10-fold cross-validation validation set of VMD-PCA-DBN model in February 2022. The 24-hour scatter map compared with the actual 10-meter wind speed: (c) WRF forecasts in December 2021, (d) VMD-PCA-DBN model forecasts in December 2021, (e) WRF forecasts in January 2022, and (f) VMD-PCA-DBN model forecasts in January 2022.

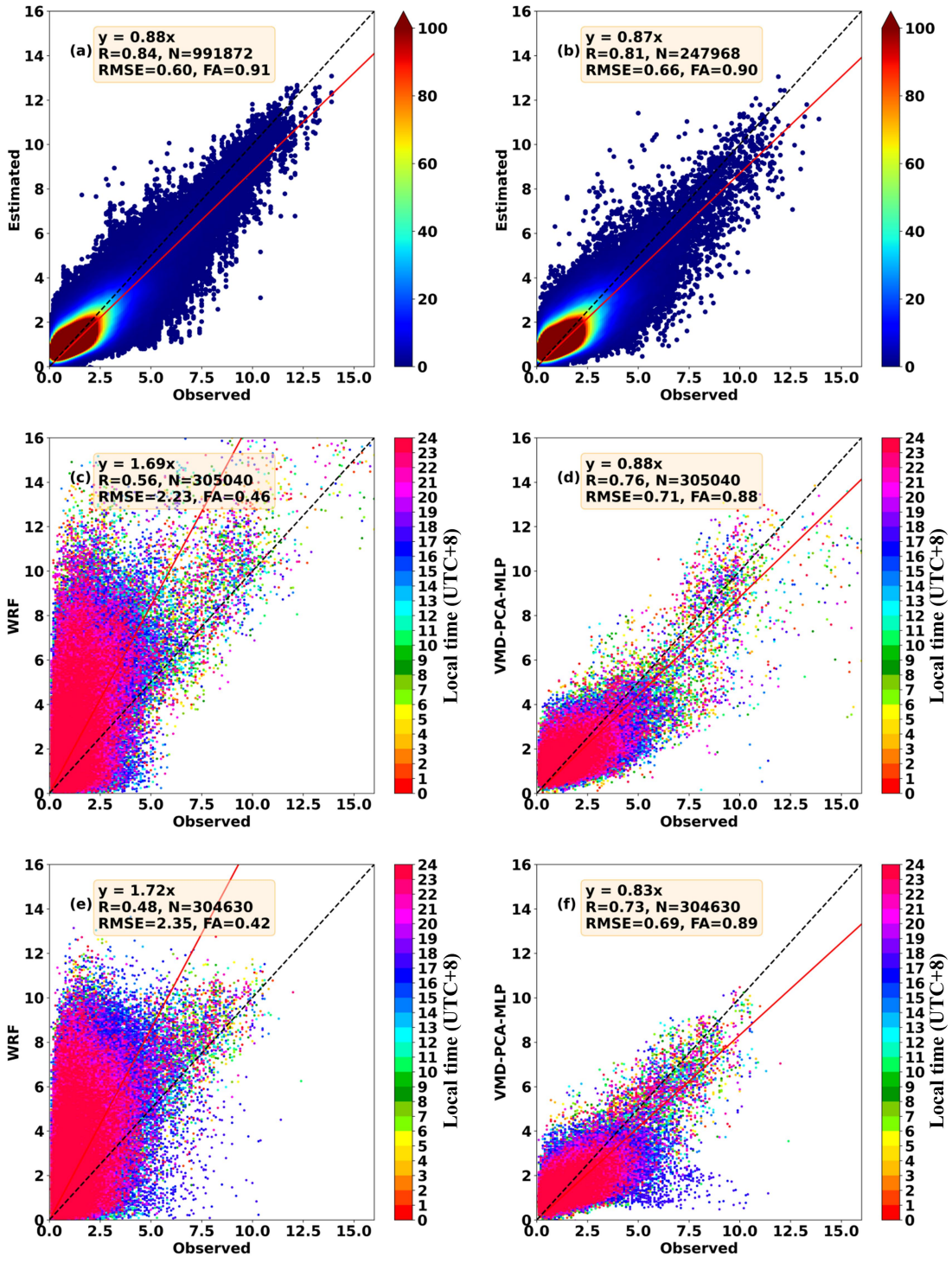

Figure S8. The scatter density map compared with the actual 10-meter wind speed: (a) 10-fold cross-validation training set of VMD-PCA-MLP model in February 2022, (b) 10-fold cross-validation validation set of VMD-PCA-MLP model in February 2022. The 24-hour scatter map compared with the actual 10-meter wind speed: (c) WRF forecasts in December 2021, (d) VMD-PCA-MLP model forecasts in December 2021, (e) WRF forecasts in January 2022, and (f) VMD-PCA-MLP model forecasts in January 2022.

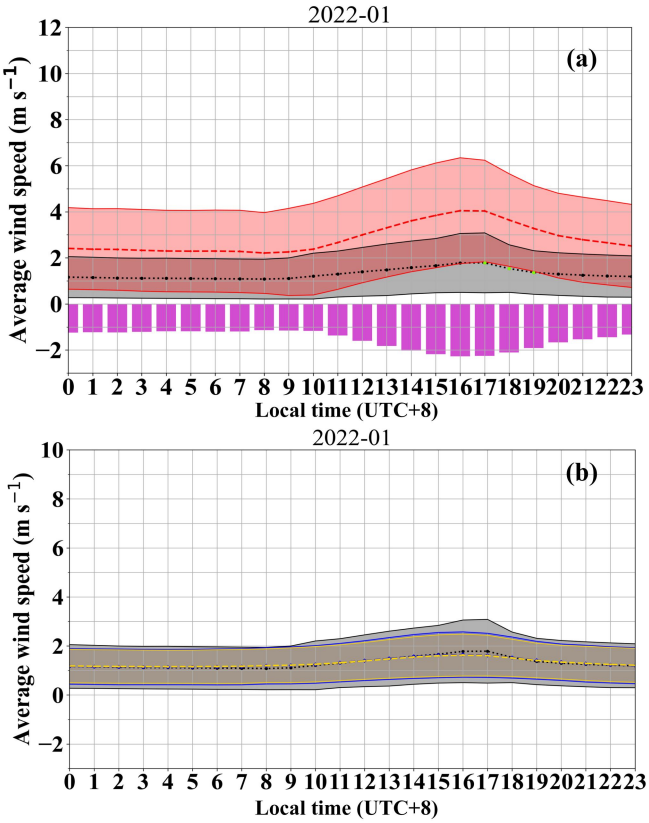

Figure S9. VMD-PCA-lightGBM, VMD-PCA-RF, and WRF daily variation of predicted and actual wind speeds in January 2022. (The shading areas represent an interval of 1 standard deviation, which is a 68% confidence interval.)

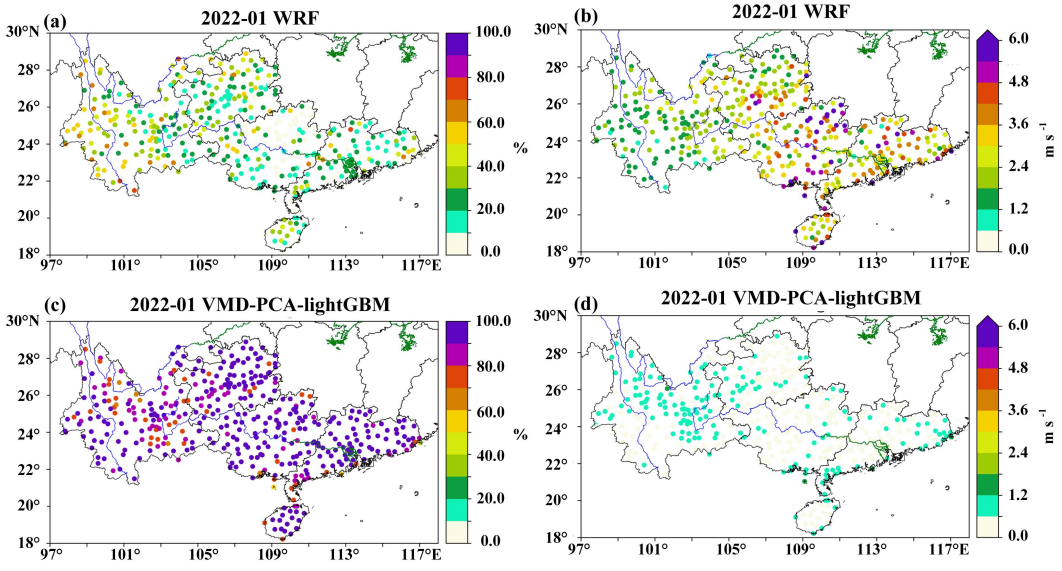

Figure S10. FA ((a), (c)) and RMSE ((b), (d)) distribution maps of VMD-PCA-lightGBM and WRF models on 410 sites in five southern provinces in January 2022.

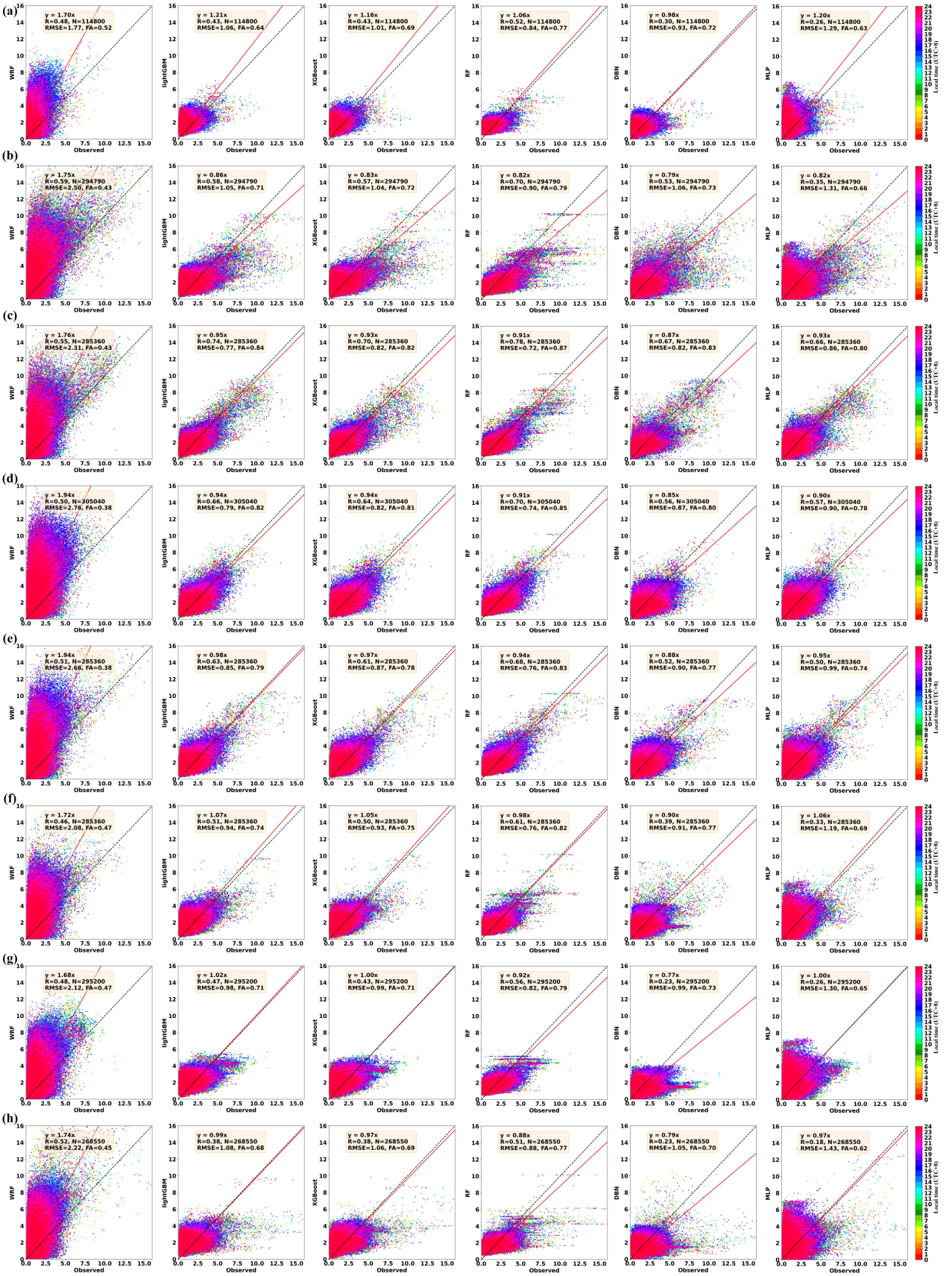

Figure S11. In Experiment 1, 24-hour scatter comparison of forecast wind speed and actual wind speed of WRF, lightGBM, XGBoost, RF, DBN and MLP models in different months respectively. ((a), (b), (c), (d), (e), (f), (g) and (h) respectively represent September 2021, October 2021, November 2021, March 2022, April 2022, May 2022, June 2022 and July 2022.)

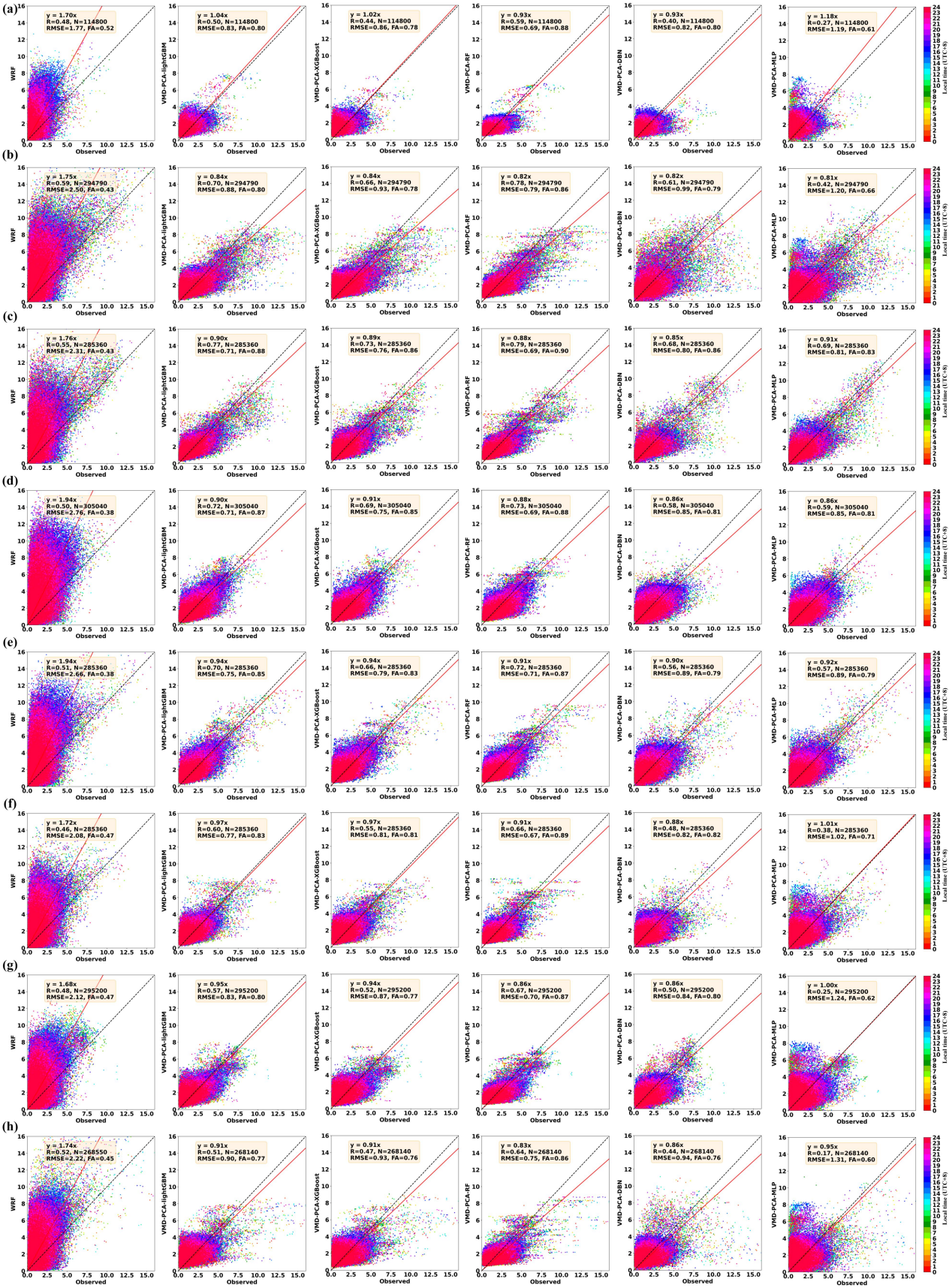

Figure S12. In Experiment 2, 24-hour scatter comparison of forecast wind speed and actual wind speed of WRF, VMD-PCA-lightGBM, VMD-PCA-XGBoost, VMD-PCA-RF, VMD-PCA-DBN and VMD-PCA-MLP models in different months respectively. ((a), (b), (c), (d), (e), (f), (g) and (h) respectively represent September 2021, October 2021, November 2021, March 2022, April 2022, May 2022, June 2022 and July

87 **2022.)**