

Decision Tree Forecasting Report - 西条

$$\hat{y}(x) = \frac{1}{N_j} \sum_{x_i \in R_j} y_i \quad \text{for } x \in R_j$$

$$\mathcal{L} = \sum_{j=1}^J \sum_{x_i \in R_j} (y_i - \hat{y}_j)^2$$

Each node partitions the input space into regions R_j .

In each region, prediction \hat{y}_j is the average of training targets.

The tree is built by recursively minimizing the total squared error \mathcal{L} .

Splits are chosen to maximize the reduction in variance of target values.

The depth and leaf size control overfitting vs generalization.

A Decision Tree splits the input space into regions (R_1, R_2, \dots, R_j) by recursively partitioning the data along the features that most reduce the variance of the target variable.

For each region R_j , the model predicts the average of training samples within that region.

The cost function minimized at each split is the total squared error \mathcal{L} .

Tree depth and leaf size control model complexity, balancing bias and variance.

Prefecture code	38
Station code	38206050
Station name	西条
Target item	Ox(ppm)
Forecast horizon (hours)	24
Number of training samples	15735
Number of test samples	6744
Number of features used	140
Model	DecisionTreeRegressor
Max depth	5
Min samples split	2
Min samples leaf	1
Elapsed time	0 min 11 sec

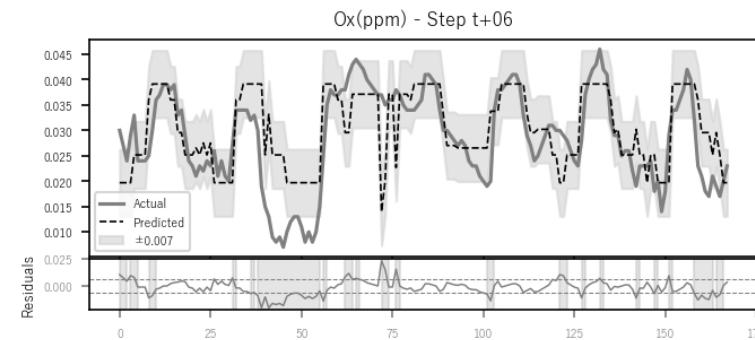
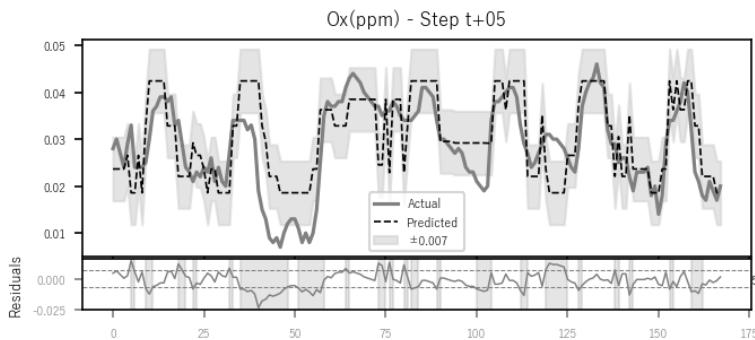
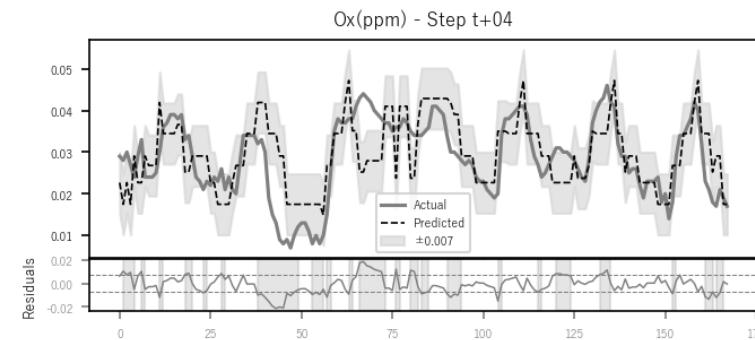
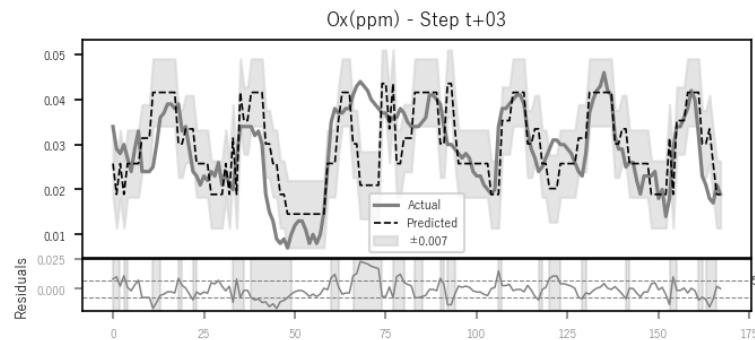
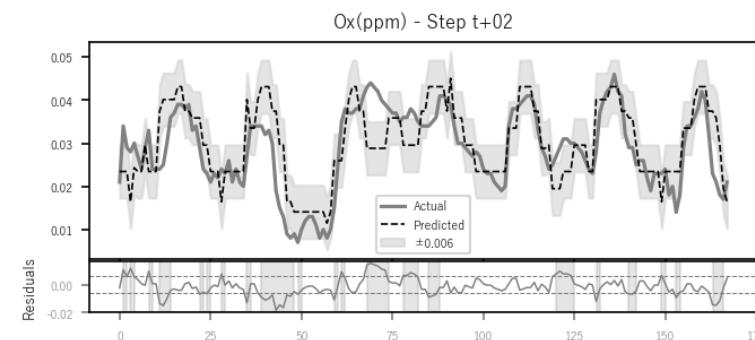
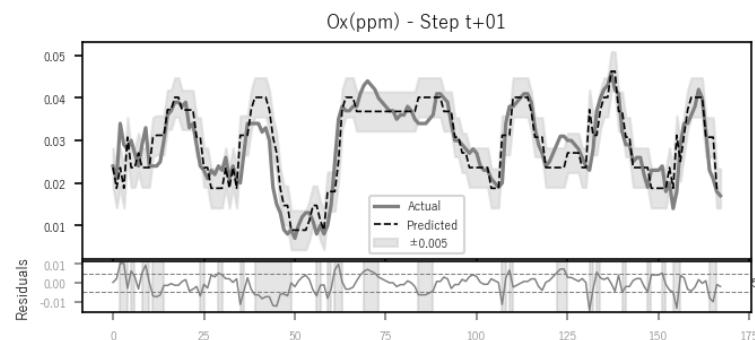
Features used for prediction

NO(ppm)	NO2(ppm)	U	V	Ox(ppm)_lag1
Ox(ppm)_lag2	Ox(ppm)_lag3	Ox(ppm)_lag4	Ox(ppm)_lag5	Ox(ppm)_lag6
Ox(ppm)_lag7	Ox(ppm)_lag8	Ox(ppm)_lag9	Ox(ppm)_lag10	Ox(ppm)_lag11
Ox(ppm)_lag12	Ox(ppm)_lag13	Ox(ppm)_lag14	Ox(ppm)_lag15	Ox(ppm)_lag16
Ox(ppm)_lag17	Ox(ppm)_lag18	Ox(ppm)_lag19	Ox(ppm)_lag20	Ox(ppm)_lag21
Ox(ppm)_lag22	Ox(ppm)_lag23	NO(ppm)_lag1	NO(ppm)_lag2	NO(ppm)_lag3
NO(ppm)_lag4	NO(ppm)_lag5	NO(ppm)_lag6	NO(ppm)_lag7	NO(ppm)_lag8
NO(ppm)_lag9	NO(ppm)_lag10	NO(ppm)_lag11	NO(ppm)_lag12	NO(ppm)_lag13
NO(ppm)_lag14	NO(ppm)_lag15	NO(ppm)_lag16	NO(ppm)_lag17	NO(ppm)_lag18
NO(ppm)_lag19	NO(ppm)_lag20	NO(ppm)_lag21	NO(ppm)_lag22	NO(ppm)_lag23
NO2(ppm)_lag1	NO2(ppm)_lag2	NO2(ppm)_lag3	NO2(ppm)_lag4	NO2(ppm)_lag5
NO2(ppm)_lag6	NO2(ppm)_lag7	NO2(ppm)_lag8	NO2(ppm)_lag9	NO2(ppm)_lag10
NO2(ppm)_lag11	NO2(ppm)_lag12	NO2(ppm)_lag13	NO2(ppm)_lag14	NO2(ppm)_lag15
NO2(ppm)_lag16	NO2(ppm)_lag17	NO2(ppm)_lag18	NO2(ppm)_lag19	NO2(ppm)_lag20
NO2(ppm)_lag21	NO2(ppm)_lag22	NO2(ppm)_lag23	U_lag1	U_lag2
U_lag3	U_lag4	U_lag5	U_lag6	U_lag7
U_lag8	U_lag9	U_lag10	U_lag11	U_lag12
U_lag13	U_lag14	U_lag15	U_lag16	U_lag17
U_lag18	U_lag19	U_lag20	U_lag21	U_lag22
U_lag23	V_lag1	V_lag2	V_lag3	V_lag4
V_lag5	V_lag6	V_lag7	V_lag8	V_lag9
V_lag10	V_lag11	V_lag12	V_lag13	V_lag14
V_lag15	V_lag16	V_lag17	V_lag18	V_lag19
V_lag20	V_lag21	V_lag22	V_lag23	Ox(ppm)_roll_mean_3
Ox(ppm)_roll_std_6	NO(ppm)_roll_mean_3	NO(ppm)_roll_std_6	NO2(ppm)_roll_mean_3	NO2(ppm)_roll_std_6
U_roll_mean_3	U_roll_std_6	V_roll_mean_3	V_roll_std_6	Ox(ppm)_diff_1
Ox(ppm)_diff_2	Ox(ppm)_diff_3	NO(ppm)_diff_3	NO2(ppm)_diff_3	U_diff_3
V_diff_3	hour_sin	hour_cos	dayofweek	is_weekend

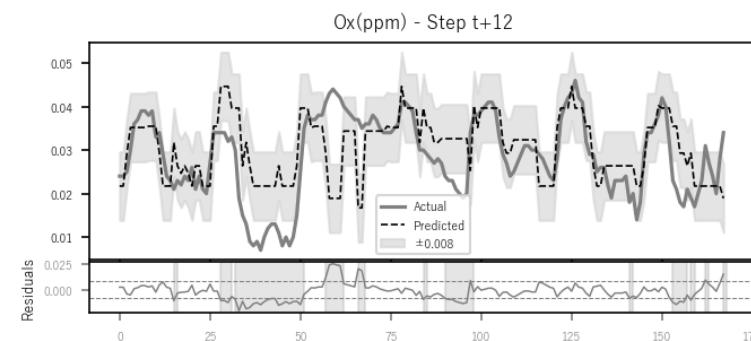
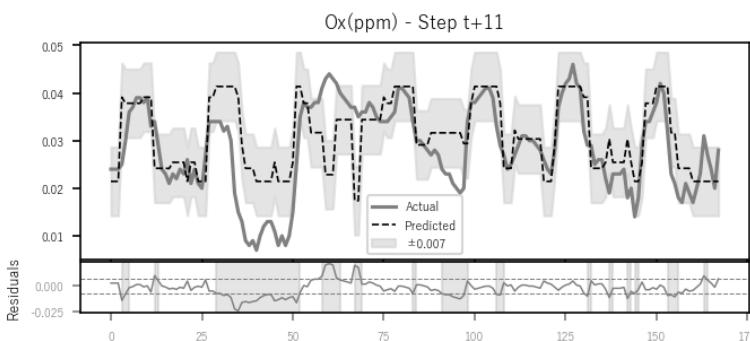
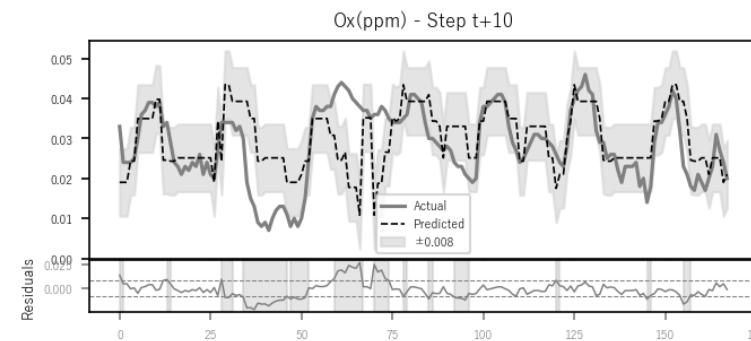
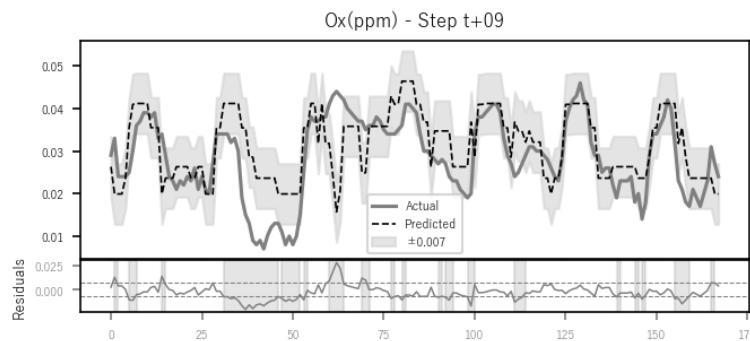
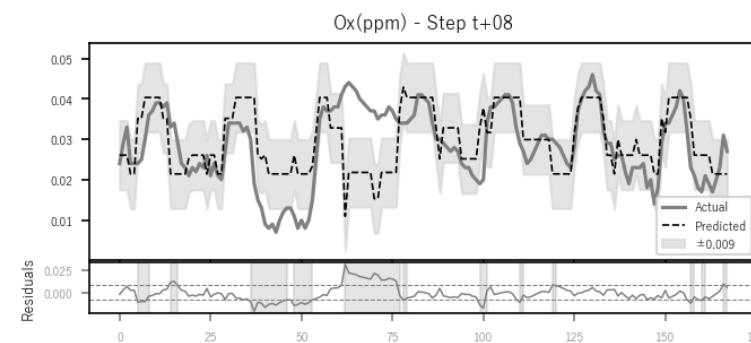
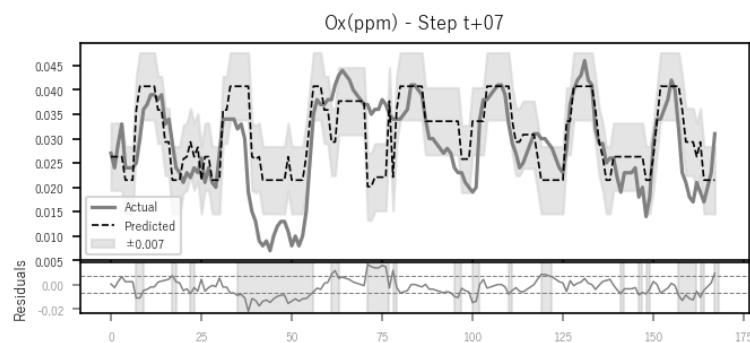
Model accuracy

Target	R ²	MAE	RMSE
Ox(ppm)_t+01	0.8196	0.0056	0.0077
Ox(ppm)_t+02	0.7480	0.0068	0.0092
Ox(ppm)_t+03	0.6508	0.0080	0.0108
Ox(ppm)_t+04	0.5730	0.0088	0.0119
Ox(ppm)_t+05	0.5216	0.0094	0.0126
Ox(ppm)_t+06	0.5274	0.0095	0.0125
Ox(ppm)_t+07	0.5057	0.0096	0.0128
Ox(ppm)_t+08	0.5016	0.0096	0.0129
Ox(ppm)_t+09	0.5058	0.0095	0.0128
Ox(ppm)_t+10	0.5068	0.0096	0.0128
Ox(ppm)_t+11	0.4949	0.0097	0.0130
Ox(ppm)_t+12	0.4817	0.0099	0.0131
Ox(ppm)_t+13	0.4695	0.0099	0.0133
Ox(ppm)_t+14	0.4784	0.0098	0.0132
Ox(ppm)_t+15	0.4864	0.0097	0.0131
Ox(ppm)_t+16	0.4959	0.0097	0.0130
Ox(ppm)_t+17	0.4831	0.0098	0.0131
Ox(ppm)_t+18	0.4665	0.0100	0.0133
Ox(ppm)_t+19	0.4637	0.0100	0.0134
Ox(ppm)_t+20	0.4684	0.0100	0.0133
Ox(ppm)_t+21	0.4689	0.0099	0.0133
Ox(ppm)_t+22	0.4859	0.0098	0.0131
Ox(ppm)_t+23	0.4697	0.0099	0.0133
Ox(ppm)_t+24	0.4361	0.0103	0.0137

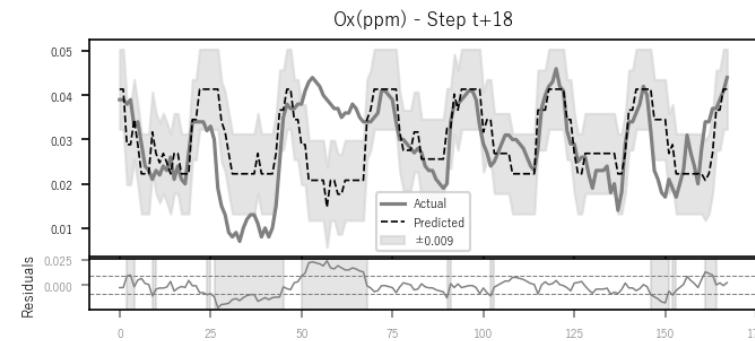
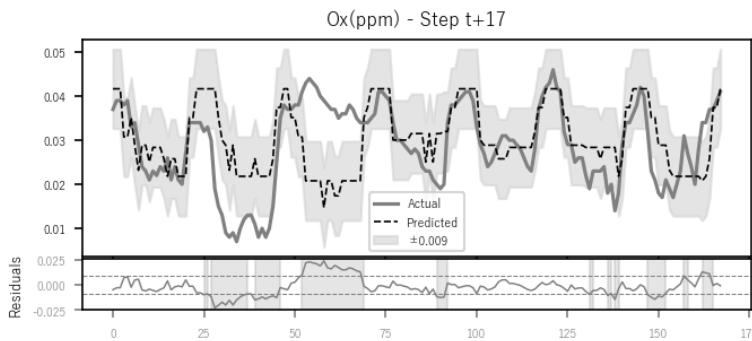
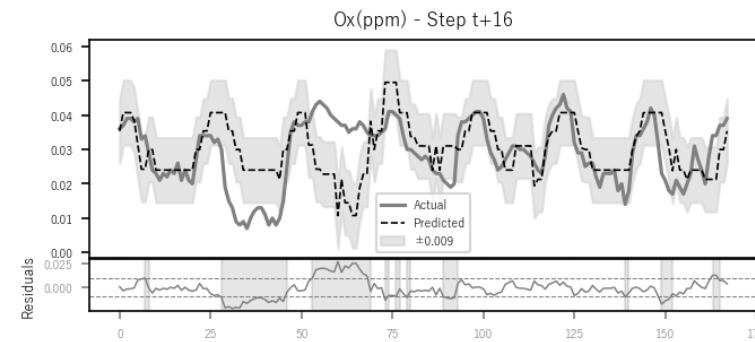
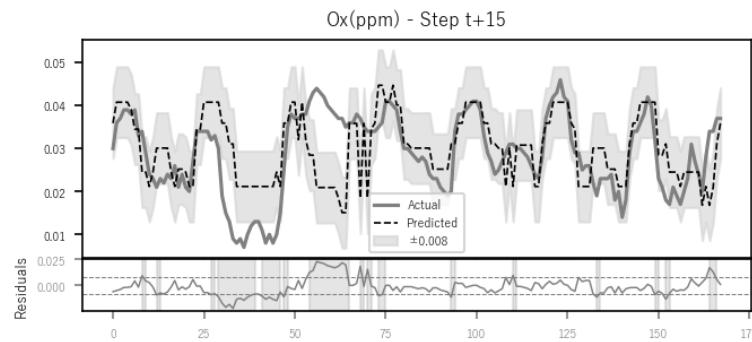
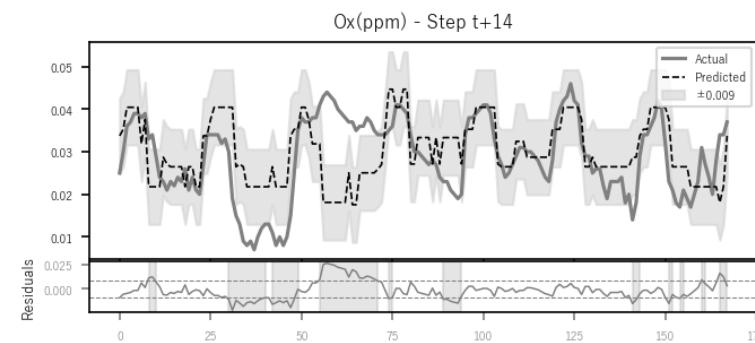
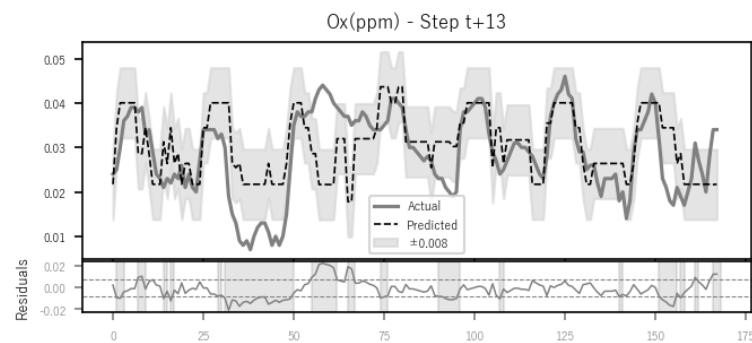
Comparison between actual and predicted values
with \pm Standard Deviation Bands



Comparison between actual and predicted values
with \pm Standard Deviation Bands



Comparison between actual and predicted values
with \pm Standard Deviation Bands



Comparison between actual and predicted values
with \pm Standard Deviation Bands

