

# ARIMA parameter selection Report - 西条

ARIMA( $p, d, q$ ) :

$$(1 - \sum_{i=1}^p \phi_i L^i)(1 - L)^d y_t = c + (1 + \sum_{i=1}^q \theta_i L^i) \varepsilon_t$$

$$\hat{y}_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j}$$

$y_t$ : observed value at time  $t$

$\hat{y}_t$ : predicted (fitted) value at time  $t$

$\phi_i$ : autoregressive (AR) coefficients, capturing dependence on past values

$\theta_j$ : moving average (MA) coefficients, capturing dependence on past errors

$d$ : differencing order for stationarity (number of times data are differenced)

$L$ : lag operator ( $Ly_t = y_{t-1}$ )

$\varepsilon_t$ : white noise (random shock) at time  $t$

$c$ : constant or drift term

Model objective = minimize residual variance  $\sigma_\varepsilon^2$  to fit observed series.

Prefecture code	38
Station code	38206050
Station name	西条
Target item	Ox(ppm)
Number of training samples	8760
Number of testing samples	720
Model	ARIMA
ARIMA order	(1, 0, 1)
Parameter Grid (tested) p	[0, 1, 2]
Parameter Grid (tested) d	[0, 1, 2]
Parameter Grid (tested) q	[1, 2, 3]
Best Parameters (found)	p=1, d=0, q=1
Predictions mean	0.03231042252207946
Predictions std	0.0019354250868405425
Real mean	0.02472777777777778
Real std	0.016101356242239538
Ljung-Box residuals autocorrelation, Prob(Q)	7.316784858687175e-138
Residuals skew	0.8080506945569366
Residuals kurtosis	2.688304328532534

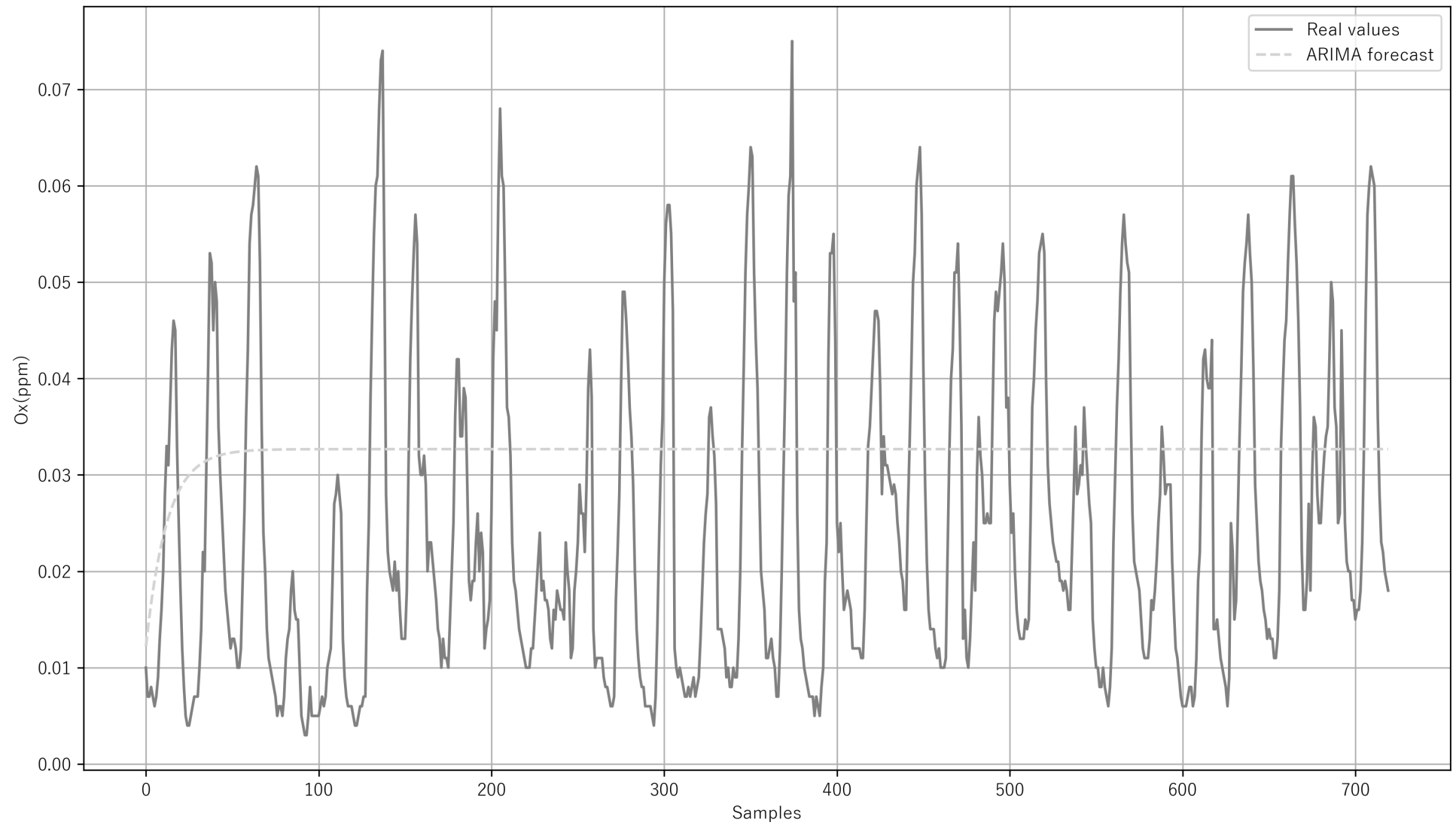
Features used for prediction

Ox(ppm)

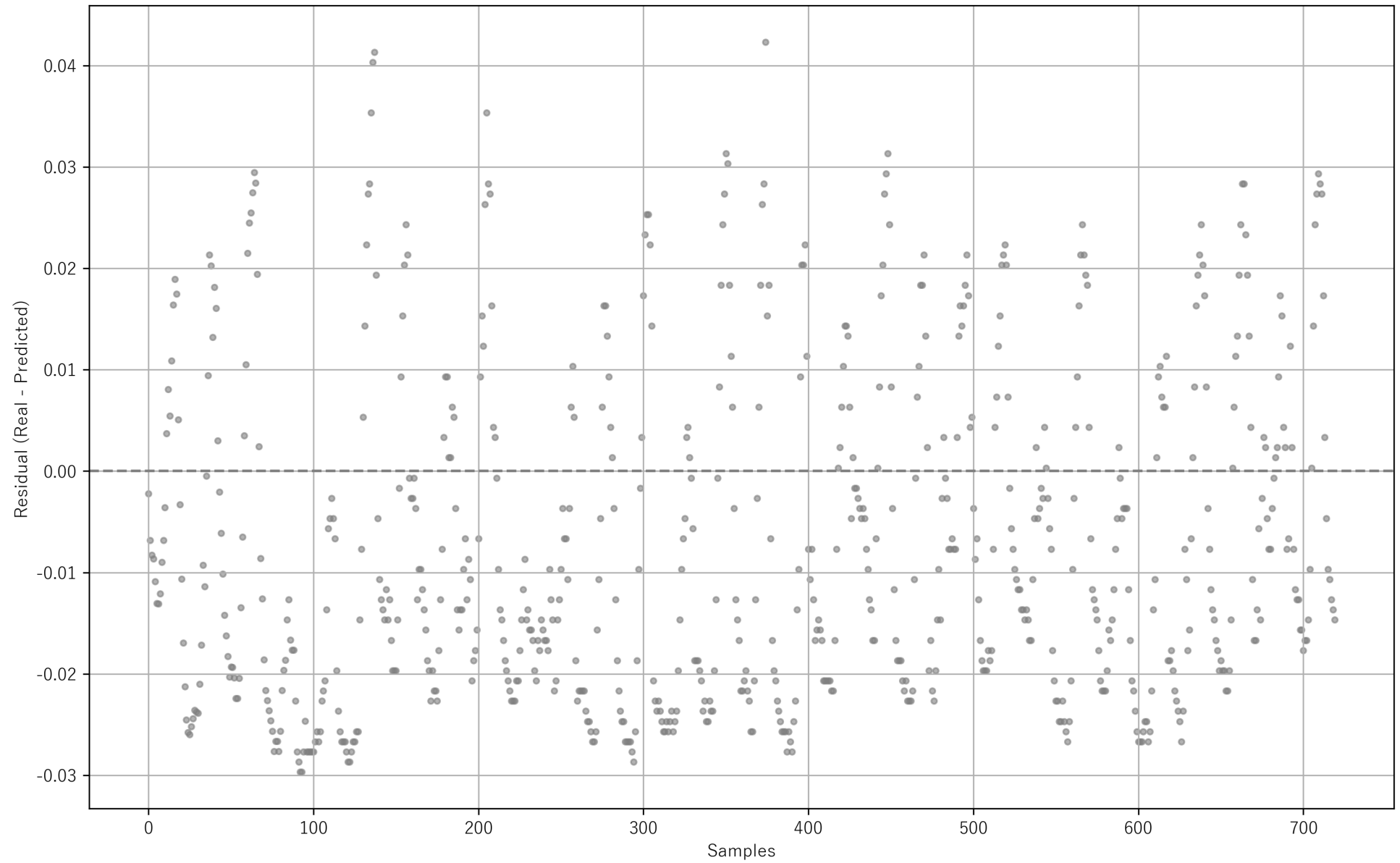
Model accuracy

Target	R <sup>2</sup>	MAE	RMSE
Ox(ppm)	-0.2123	0.0157	0.0177

ARIMA(1, 0, 1)  
 $R^2: -0.21233$



Residuals



Histogram of Residuals – Distribution & Central Tendency

