**1. Delhi Dataset (Remove RMSE)**

**A) Short Term prediction (Till 30 days)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Period of forecast | | | | | |
| 1-5 days | | 5-15 days | | 15-30 days | |
| MAE | RMSE | MAE | RMSE | MAE | RMSE |
| First order autoregressive ARIMA | 7.82 | 10.14 | **8.20** | 10.28 | 13.24 | 15.17 |
| **Damped-trend linear exponential smoothing** | 7.43 | 8.97 | 8.45 | 9.94 | **9.87** | 12.00 |
| Differenced first order autoregressive | 7.50 | 8.93 | 9.29 | 11.65 | 13.66 | 15.56 |
| SARIMA | **7.29** | 8.63 | 8.49 | 10.72 | 13.44 | 15.24 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Period of forecast | | | | | |
| 1-5 days | | 5-15 days | | 15-30 days | |
| MAE | RMSE | MAE | RMSE | MAE | RMSE |
| **Support Vector Regressor** | **9.15** | 10.17 | **9.24** | 10.67 | **9.95** | 12.68 |
| XGBoost Regressor | 8.96 | 10.59 | 9.61 | 12.38 | 13.61 | 18.28 |
| Decision Tree Regressor | 14.60 | 16.80 | 19.70 | 24.85 | 23.84 | 30.59 |
| Random Forest Regressor | 9.44 | 10.66 | 9.59 | 10.73 | 12.12 | 15.58 |

**A) Long Term prediction (Till 250 days)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Period of forecast | | | | | |
| 30-60 days | | 60-100 days | | 100-250 days | |
| MAE | RMSE | MAE | RMSE | MAE | RMSE |
| **First order autoregressive ARIMA** | 17.03 | 23.95 | **28.27** | 43.05 | **54.36** | 73.94 |
| Damped-trend linear exponential smoothing | 17.11 | 24.05 | 51.57 | 56.55 | 89.21 | 112.82 |
| Differenced first order autoregressive | 16.04 | 23.09 | 47.36 | 60.17 | 97.52 | 121.39 |
| SARIMA | **14.63** | 21.63 | 49.47 | 61.51 | 104.67 | 128.15 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Period of forecast | | | | | |
| 30-60 days | | 60-100 days | | 100-250 days | |
| MAE | RMSE | MAE | RMSE | MAE | RMSE |
| Support Vector Regressor | **14.54** | 20.61 | 21.58 | 31.15 | 30.53 | 44.48 |
| XGBoost Regressor | 16.61 | 22.42 | 22.10 | 29.72 | 29.17 | 42.50 |
| Decision Tree Regressor | 25.01 | 31.70 | 38.78 | 62.68 | 40.19 | 58.55 |
| **Random Forest Regressor** | 15.28 | 20.72 | **21.22** | 31.26 | **26.46** | 39.29 |

**2. Mumbai dataset**

**A) Short Term prediction (Till 30 days)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Period of forecast | | | | | |
| 1-5 days | | 5-15 days | | 15-30 days | |
| MAE | RMSE | MAE | RMSE | MAE | RMSE |
| First order autoregressive ARIMA | 9.47 | 11.04 | 19.27 | 25.05 | 21.45 | 25.11 |
| Damped-trend linear exponential smoothing | 14.61 | 16.98 | 21.14 | 25.66 | 37.70 | 44.03 |
| Differenced first order autoregressive | 10.78 | 15.12 | 17.47 | 23.31 | 34.55 | 40.58 |
| SARIMA | 9.80 | 13.44 | 17.58 | 23.29 | 34.40 | 40.15 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Period of forecast | | | | | |
| 1-5 days | | 5-15 days | | 15-30 days | |
| MAE | RMSE | MAE | RMSE | MAE | RMSE |
| Support Vector Regressor | 13.87 | 15.98 | 14.54 | 18.83 | 15.98 | 21.56 |
| XGBoost Regressor | 7.27 | 8.08 | 14.54 | 20.37 | 12.98 | 18.83 |
| Decision Tree Regressor | 10.50 | 12.51 | 16.00 | 19.70 | 17.97 | 22.98 |
| **Random Forest Regressor** | **11.54** | **12.17** | **11.73** | **17.31** | **12.93** | **18.78** |

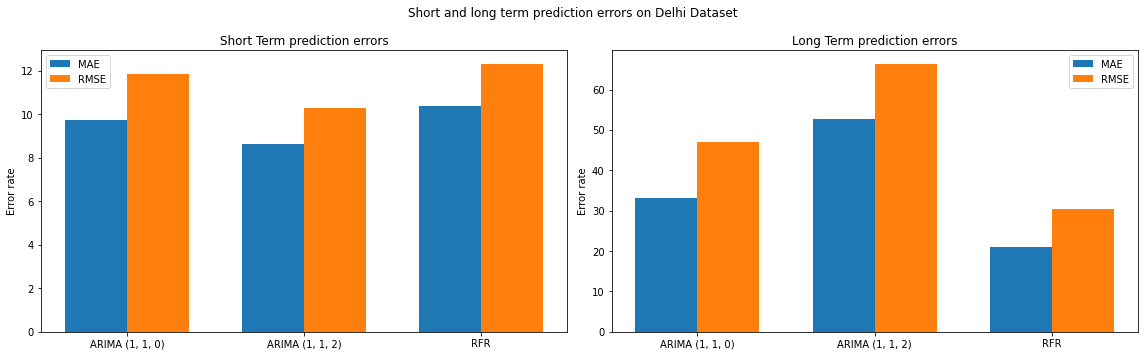
**A) Long Term prediction (Till 250 days)**

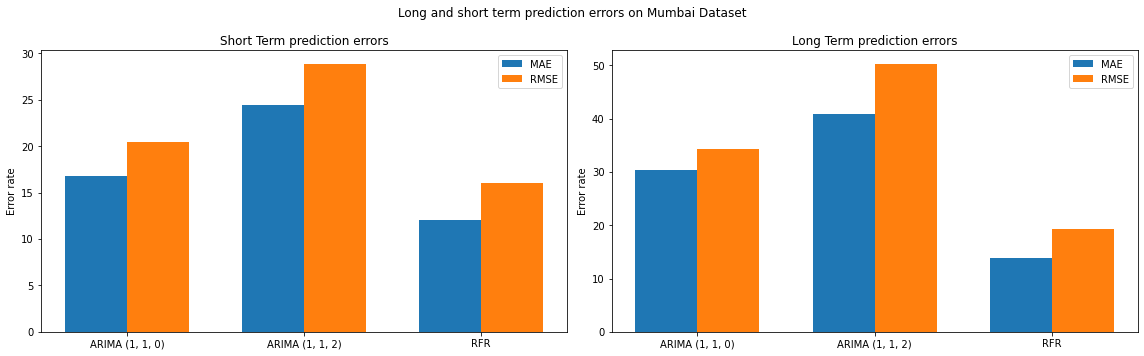
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Period of forecast | | | | | |
| 30-60 days | | 60-100 days | | 100-250 days | |
| MAE | RMSE | MAE | RMSE | MAE | RMSE |
| **First order autoregressive ARIMA** | **25.77** | **29.37** | **27.65** | **31.81** | **37.81** | **41.90** |
| Damped-trend linear exponential smoothing | 25.27 | 30.80 | 30.03 | 40.54 | 67.60 | 79.73 |
| Differenced first order autoregressive | 24.56 | 29.14 | 38.22 | 47.45 | 63.42 | 75.14 |
| SARIMA | 24.48 | 28.99 | 38.08 | 47.38 | 68.19 | 80.18 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Period of forecast | | | | | |
| 30-60 days | | 60-100 days | | 100-250 days | |
| MAE | RMSE | MAE | RMSE | MAE | RMSE |
| Support Vector Regressor | 15.13 | 20.03 | 16.44 | 22.43 | 15.21 | 20.69 |
| XGBoost Regressor | 13.65 | 18.81 | 15.62 | 21.66 | 13.74 | 19.52 |
| Decision Tree Regressor | 20.16 | 27.80 | 21.80 | 28.06 | 18.15 | 24.31 |
| **Random Forest Regressor** | **13.25** | **18.16** | **15.05** | **21.01** | **13.04** | **18.68** |

Average MAE and RMSE when short term and long-term values are predicted on Delhi AQI and Mumbai AQI dataset.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Dataset | Model | Short Term | | Long Term | |
| Avg. MAE | Avg. RMSE | Avg. MAE | Avg. RMSE |
| Delhi AQI | First order autoregressive ARIMA | 9.75 | 11.86 | 33.22 | 46.98 |
| Damped-trend linear exponential Smoothing | **8.62** | 10.30 | 52.63 | 64.47 |
| Random Forest Regressor | 10.38 | 12.32 | **20.98** | 30.42 |
| Mumbai AQI | First order autoregressive ARIMA | 16.73 | 20.4 | 30.41 | 34.36 |
| Damped-trend linear exponential Smoothing | 24.48 | 28.89 | 40.96 | 50.35 |
| Random Forest Regressor | **12.07** | 16.08 | **13.78** | 19.28 |





From the graph, Random Forest Regressor outperforms several statistical methods.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Period of forecast | | | | | |
| 1-5 hours | | 5-15 hours | | 15-30 hours | |
| MAE | RMSE | MAE | RMSE | MAE | RMSE |
| Random Forest Regressor | 9.35 | 10.49 | 13.91 | 21.84 | 10.73 | 17.17 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Period of forecast | | | | | |
| 30-60 hours | | 60-100 hours | | 100-250 hours | |
| MAE | RMSE | MAE | RMSE | MAE | RMSE |
| Random Forest Regressor | 8.33 | 13.23 | 9.14 | 13.51 | 11.06 | 15.51 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Short term | | Long term | |
| MAE | RMSE | MAE | RMSE |
| Random Forest Regressor | 11.33 | 16.5 | 9.51 | 14.08 |

Random forest has good long term prediction capabilities. Even though short-term prediction errors are higher than that of long-term prediction error, it is significantly low than the error rates found in literature.

Comparison with literature

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Zheng et.al | DAQFF | RFR short-term | RFR long-term |
| MAE | 23.70 | 25.01 | **11.33** | **9.51** |