**Supplementary Information**

**Machine Learning Approach for Predicting Atmospheric Corrosion Rate of Zinc**

Jasmita Nur Azizah,a Vani Virdyawan,b Afriyanti Sumboja,a Asep Ridwan Setiawan,a\*

a Material Science and Engineering Research Group, Faculty of Mechanical and Aerospace Engineering, Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia.

b Mechanical Production Engineering Research Group, Faculty of Mechanical and Aerospace Engineering, Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia.

\*Email:[asep.ridwans@material.itb.ac.id](mailto:asep.ridwans@material.itb.ac.id)

|  |  |
| --- | --- |
|  |  |
| Accuracy metric (R2 and RMSE) of GB regressor. R2=1 and RMSE=0 for perfect prediction. | |
|  |  |
| Accuracy metric (R2 and RMSE) of Bagging regressor. R2=1 and RMSE=0 for perfect prediction. | |
|  |  |
| Accuracy metric (R2 and RMSE) of XGB regressor. R2=1 and RMSE=0 for perfect prediction. | |

Figure S1. Train-test-spliting ratio iteration to find the best train-test split ratio that gives higher accuracy.

Diagram

Description automatically generated

Figure S2. The flowchart of typical cross validation workflow in model training.



Figure S3. K-fold cross validation scheme

The corrosion data was collected from previous published journals, so information extraction was carried out on relevant information (Table S1). [1-6]

|  |  |  |  |
| --- | --- | --- | --- |
| Table S1. Input parameters | | | |
| No. | Parameter | Information | Source |
| 1. | T | Measurement temperature | Research journal |
| 2. | RH | Relative Humidity |
| 3. | TOW | Time of Wetness |
| 4. | Prec | Precipitation |
| 5. | SO2 | SO2 deposition |
| 6. | H+ | [H+] |
| No. | Property Target | Information |  |
| 1. | CR | Corrosion rate (CR) | Research journal |

The model was tested with new input corrosion data from Panama, Rio Janeiro, Iguazu, Sao Paulo, Bauta, Labastida, dan Mexico (Table S2). [7-13]

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Table S2. New data to be used as validation | | | | | | | |
| Location | T (oC) | RH (%) | TOW (Hours/a) | Prec (mm/a) | SO2 (µg/m3) | [H+] (mg/L) | CR (g/m2) |
| Panama | 26 | 71 | 3000 | 279 | 21.7 | 0 | 7.568 |
| Rio Janeiro | 21.4 | 80 | 4800 | 1252 | 43.5 | 0 | 10.496 |
| Iguazu | 21.2 | 75 | 5500 | 1000 | 0 | 0 | 9.710 |
| Sao Paulo | 19.6 | 75 | 5500 | 1622 | 57.8 | 0 | 8.639 |
| Bauta | 24 | 81 | 5000 | 1590 | 18 | 0 | 8.711 |
| Labastida | 12 | 73 | 5500 | 550 | 3.6 | 0 | 2.213 |
| Mexico | 15.4 | 64 | 1000 | 709 | 18.9 | 0 | 5.855 |

Predictions were made on the corrosion rate data in Jakarta, Bandung, Valparaiso, and Barcelona in Table S3.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table S3. New data on corrosion rates at several test locations | | | | | | |
| Location | T (oC) | RH (%) | TOW (Hours/a) | Prec (mm/a) | SO2 (µg/m3) | [H+] (mg/L) |
|
| Jakarta, Indonesia | 31 | 65 | 5500 | 100 | 71 | 0 |
| Bandung, Indonesia | 24 | 81 | 3500 | 999 | 17 | 0 |
| Valparaiso, Chile | 17 | 83 | 5600 | 0.1 | 23.6 | 0 |
| Barcelona, Spain | 26 | 78 | 500 | 640 | 1.5 | 0 |
| Tokyo, Japan | 22 | 73 | 29.1 | 1530 | 12.4 | 0 |

EDA was conducted on initial dataset, contains statical information and data quality (Table S4)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Table S4. Information on research datasets used in machine learning | | | | | | | |
|  | **T** | **RH** | **TOW** | **Prec** | **SO2** | **[H+]** | **CR** |
| **count** | 171 | 171 | 171 | 171 | 171 | 171 | 171 |
| **mean** | 11.04 | 72.39 | 3121.36 | 797.73 | 12.31 | 0.01 | 7.25 |
| **std** | 7.96 | 9.54 | 1807.60 | 461.34 | 16.29 | 0.02 | 4.95 |
| **min** | -16.60 | 33 | 0.00 | 17.00 | 0.00 | 0.00 | 0.65 |
| **25%** | 7.00 | 67.50 | 2111 | 525.10 | 3.00 | 0.00 | 4.07 |
| **50%** | 10.70 | 75 | 3074 | 679.60 | 6.20 | 0.00 | 6.62 |
| **75%** | 16.35 | 79 | 4378.50 | 974.85 | 13.70 | 0.03 | 9.07 |
| **max** | 27.00 | 98 | 8760 | 2624.00 | 83.30 | 0.19 | 27.00 |

The MAE values ​​from Gradient Boosting Regressor, Bagging Regressor, and XGBoost Regressor. (Table S5)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Table S5. The results of the comparison of predictive atmospheric corrosion rate values with the experimental value | | | | | | | |
| Location | CR actual | CR Validation (g/m2) | | |  | | |
| BR | GB | XGB | BR | GB | XGB |
| Panama | 7.568 | 10.167 | 8.77 | 9.565 | 2.599 | 1.202 | 1.997 |
| Rio Janeiro | 10.496 | 10.52 | 8.941 | 9.19 | 0.024 | 1.555 | 1.306 |
| Iguazu | 9.710 | 14.108 | 15.913 | 20.155 | 4.398 | 6.203 | 10.445 |
| Sao Paulo | 8.639 | 9.141 | 9.522 | 9.612 | 0.502 | 0.883 | 0.973 |
| Bauta | 8.711 | 9.731 | 8.132 | 10.674 | 1.02 | 0.579 | 1.963 |
| Labastida | 2.213 | 4.41 | 4.795 | 3.496 | 2.197 | 2.582 | 1.283 |
| Mexico | 5.855 | 7.314 | 7.819 | 6.761 | 1.459 | 1.964 | 0.906 |
|  | | | | | 12.199 | 14.968 | 18.873 |
| MAE | | | | | 1.743 | 2.138 | 2.696 |

**REFERENCE**

[1] Y. M. Panchenko and A. I. Marshakov, “Prediction of first-year corrosion losses of carbon steel and zinc in continental regions,” *Materials (Basel).*, 2017, doi: 10.3390/ma10040422.

[2] Y. M. Panchenko, A. I. Marshakov, L. A. Nikolaeva, V. V Kovtanyuk, T. N. Igonin, and T. A. Andryushchenko, “Comparative estimation of long-term predictions of corrosion losses for carbon steel and zinc using various models for the Russian territory,” *Corros. Eng. Sci. Technol.*, 2017, doi: 10.1080/1478422X.2016.1227024.

[3] Y. M. Panchenko and A. I. Marshakov, “Long-term prediction of metal corrosion losses in atmosphere using a power-linear function,” *Corros. Sci.*, 2016, doi: 10.1016/j.corsci.2016.04.002.

[4] S. Pintos, N. V Queipo, O. Troconis de Rincón, A. Rincón, and M. Morcillo, “Artificial neural network modelling of atmospheric corrosion in the MICAT project,” *Corros. Sci.*, vol. 42, no. 1, pp. 35–52, 2000, doi: 10.1016/S0010-938X(99)00054-2.

[5] D. de la Fuente, J. G. Castaño, and M. Morcillo, “Long-term atmospheric corrosion of zinc,” *Corros. Sci.*, vol. 49, no. 3, pp. 1420–1436, 2007, doi: 10.1016/j.corsci.2006.08.003.

[6] S. Feliu, M. Morcillo, and S. Feliu, “The prediction of atmospheric corrosion from meteorological and pollution parameters—I. Annual corrosion,” *Corros. Sci.*, vol. 34, no. 3, pp. 403–414, 1993, doi: 10.1016/0010-938X(93)90112-T.

[7] climatestotravel.com, “Climate - Panama.” [Online]. Available: https://www.climatestotravel.com/climate/panama.

[8] Climate-data.org, “Rio De Janeiro Climate Summary.” [Online]. Available: https://en.climate-data.org/south-america/brazil/rio-de-janeiro/rio-de-janeiro-853/.

[9] weatherandclimate.com, “Climate in Puerto Iguazú (Misiones).” [Online]. Available: https://weather-and-climate.com/average-monthly-Rainfall-Temperature-Sunshine,puerto-iguazu,Argentina.

[10] W. W. & C. Information, “Average Monthly Snow And Rainfall In Sao Paulo (Sao Paulo State) In Millimeter.” 2021, [Online]. Available: https://weather-and-climate.com/average-monthly-precipitation-Rainfall,sao-paulo,Brazil.

[11] W. W. & C. Information, “Bauta Monthly Climate Averages.” [Online]. Available: https://www.worldweatheronline.com/bauta-weather-averages/la-habana/cu.aspx.

[12] weatherandclimate.com, “Average Monthly Snow And Rainfall In Labastida (Basque Country) In Millimeter.” [Online]. Available: https://weather-and-climate.com/average-monthly-precipitation-Rainfall,labastida-basque-country-es,Spain.

[13] Climatemps.com, “Rainfall/Precipitation in Mexico City Mexico.” 2017, [Online]. Available: http://www.mexico-city.climatemps.com/precipitation.php.