

Prediction of Corrosion Rate Using Big Data Analytics

Suryaprakash Samudrala¹, Suresha Talanki², Shoba M^{3*}, Sachin⁴, Varsha S⁵ and Jeet Roy⁶

¹Senior Engineer in Inspection & Corrosion Division, Kuwait Oil Company, Kuwait

^{2,3,4,5,6}Sri Venkateshwara College of Engineering, Bengaluru, Karnataka, India

ABSTRACT

This chapter provides an overview of corrosion of metals. It describes the general factors that influences metal corrosion. Over the years, there have been numerous studies on the rates of corrosion of metals in sea water. From various studies, it is now possible to identify the major factors that affect metal corrosion. These factors are metal composition, water composition, temperature, marine growth, seabed composition, and extent of water movement. The combined effect of all these complex and often interrelated factors is that each object must be considered individually when attempting to evaluate its corrosion history or when considering its recovered condition. Till date the available literature on corrosion factor prejudiced by the bacteria, which was predominantly focused on Sulfate Reducing Bacteria (SRB) that usually reside on sulfate (terminal electron acceptor) because SRB are often found at pitting sites.

KEY WORDS: CORROSION, BIG DATA, MLR ALGORITHM, HEAT MAPS, COUPONS.

INTRODUCTION

The corrosion is a natural process of degradation that converts a refined metal into a more chemically stable form which is its natural state. The very serious effects of the corrosion process generate problems of global significance. Many industries, such as energy, automotive, food, chemical, household appliances, brewery, suffer huge economic losses (Morsillo and Chico, 2013) because of the corrosion damages. Corrosion harms assets; forming pits and decreasing mechanical properties in water rotating propellers or cavitation, leaks in pipelines transporting fluids like water or petroleum, corrosive effects of acid rains on metallic and/or non-metallic surfaces etc. Some of the effects of corrosion include a significant

deterioration of natural and historic monuments as well as increase the risk of catastrophic equipment failures.

This paper summarizes and proves the theoretical concepts of corrosion with practical proofs that are embedded in the huge data-sets collected over the years from industry, and analyzing the same Using algorithms which can handle (Nishimura 2008) big data in different environments. Big Data is a term used to describe a collection of data that is huge in volume and yet growing exponentially with time. In-short such data is so large and complex that none of the traditional data management tools are able to store it or process it efficiently. Big data is a term that describes the large volume of data – both structured and unstructured – that inundates a business on a day-to-day basis. But it is not the amount of data that is important, it is what organizations do with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves.

Big Data helps the organizations to create new growth opportunities and entirely new categories of companies that can combine and analyze industry data. These companies have ample information about the products and services, buyers and suppliers, consumer preferences that can be captured and analyzed. While the term “big data” is relatively new, the act of gathering and storing

ARTICLE INFORMATION

*Corresponding Author: shoba0407@gmail.com
Received 9th Oct 2020 Accepted after revision 29th Dec 2020
Print ISSN: 0974-6455 Online ISSN: 2321-4007 CODEN: BBRBCA

Thomson Reuters ISI Web of Science Clarivate Analytics USA and Crossref Indexed Journal



NAAS Journal Score 2020 (4.31)
A Society of Science and Nature Publication,
Bhopal India 2020. All rights reserved.
Online Contents Available at: <http://www.bbrc.in/>
Doi: <http://dx.doi.org/10.21786/bbrc/13.13/33>

large amounts of information for eventual analysis is ages old. The concept gained momentum in the early 2000s when industry analyst Doug Laney articulated the now-mainstream definition of big data as the three Vs: Why is big data important? The importance of big data does not revolve around how much data a company has but how a company utilizes the collected data. Every company uses data in its own way; the more efficiently a company uses its data, the more potential it has to grow. The company can take data from any source and analyze it to find answers which will enable: Cost Savings: Some tools of Big Data like Hadoop and Cloud-Based Analytics can bring cost advantages to business when large amounts of data are to be stored and these tools also help in identifying more efficient ways of doing business. Time Reductions: The high speed of tools like Hadoop and in-memory analytics can easily identify new sources of data which helps businesses analyzing data immediately and make quick decisions based on the learnings.

Understand the market conditions: By analyzing big data you can get a better understanding of current market conditions. For example, by analyzing customers' purchasing behaviors, a company can find out the products that are sold the most and produce products according to this trend. By this, it can get ahead of its competitors. Control online reputation: Big data tools can do sentiment analysis. Therefore, you can get feedback about who is saying what about your company. If you want to monitor and improve the online presence of your business, then, big data tools can help in all this. Using Big Data Analytics to Boost Customer Acquisition and Retention: The customer is the most important asset any business depends on.

There is no single business that can claim success without first having to establish a solid customer base. However, even with a customer base, a business cannot afford to disregard the high competition it faces. If a business is slow to learn what customers are looking for, then it is very easy to begin offering poor quality products. In the end, loss of clientele will result, and this creates an adverse overall effect on business success. The use of big data allows businesses to observe various customer related patterns and trends. Observing customer behavior is important to trigger loyalty. Using Big Data Analytics to Solve Advertisers Problem and Offer Marketing Insights: Big data analytics can help change all business operations. This includes the ability to match customer expectation, changing company's product line and of course ensuring that the marketing campaigns are powerful. Big Data Analytics as a Driver of Innovations and Product Development: Another huge advantage of big data is the ability to help companies innovate and redevelop their products.

Corrosion being a continuous process and data generated to document corrosion in terms of its measurements in various fluids and under different process conditions is huge, Big Data analysis concepts were applied to get an insight into the relationship of corrosion (in

terms of corrosion rates) with different environments. The literature review was done in order to establish an Understanding of the state-of-the-art research on corrosion. From literature survey (Diaz 2013; Hoerle 2004; Panchenko 2008; Wang 2013) it was found that various forms of bacteria create different localized environments and hence the corrosion rates with various bacteria forms in different environments was sought in analyzing the huge corrosion data base. In this paper using the real-time data, we extracted the information using big data analysis and machine learning concepts and hence proved which bacteria is causing what effect in which environment and hence giving proper justification for the corrosion problem which is theoretical existing and matching with the correlated practical proofs to the industry sector.

METHODOLOGY

The Multiple Linear Regression (MLR) algorithm is used to find the correlation of the corrosion rate and pitting rate with other parameters in the data. This algorithm takes multiple inputs and generates the output based on the parameters to be predicted. Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of multiple linear regression (MLR) is to model the linear-relationship between the explanatory (independent) variables and response (dependent) variable.

The MLR algorithm we used to find the co relation of the corrosion rate and pitting rate with other parameters in the data. This algorithm takes multiple inputs n gives the output based on the parameter to be predicted. Heat maps are used as a means of pictorial representation to visualize the data in an efficient and understandable manner. This provides the clear picture about the correlation between the parameters used in the algorithm. According to our dataset, the maps will represent the corrosion rate, pitting rate based on the impact of bacteria and environment choice.

RESULTS AND DISCUSSION

Following steps were implemented to get the desired results, Classification of the dataset according to the environment, application of MLR algorithm to each set of the dataset obtained, representation of graphs using Heat Map Concepts. Dataset is been trained using MLR algorithm and visualized it graphically using heat maps for every dataset model. Figure 1 and Table 1 shows the results and action performed by the bacteria on dry-crude.

Dry Crude (CRD)

The effect of SRB is predominant and it constitutes to corrosion rate. Hence it is proved that SRB is most dangerous, if present then preventive measures need to be taken. When bacteria colony occurs, SRB is the most

dangerous bacteria. Figure 3 represents the coupon of dry crude. Theoretically the concept exists, and now hence proved with analysis as well

its existence in high chloride water and GAnB causes very severe pitting corrosion rate.

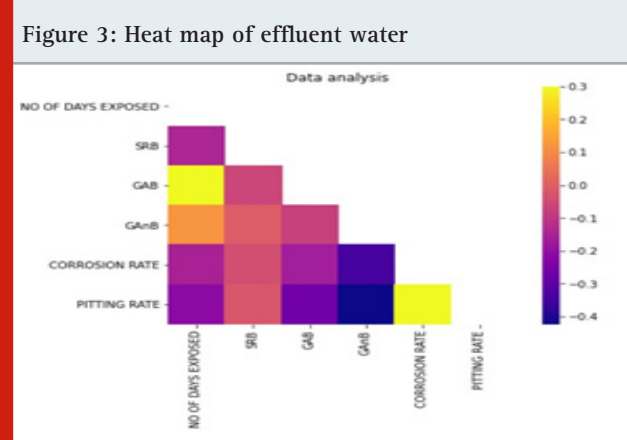
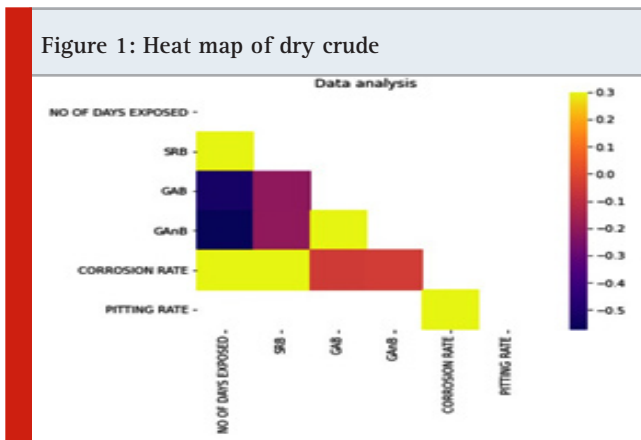
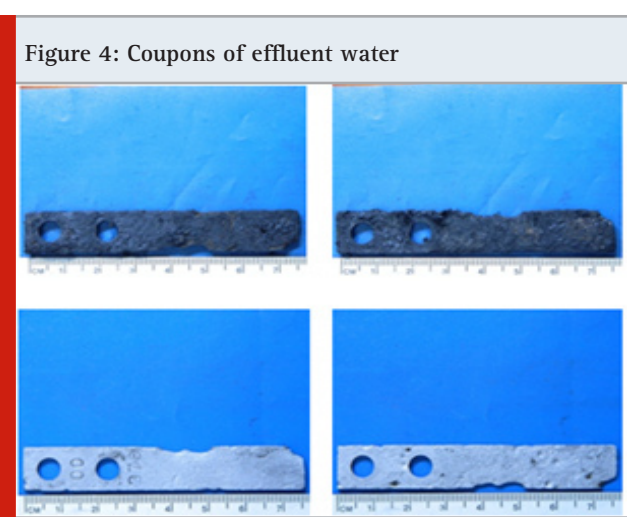


Table 1. Dry Crude Corrosion/Pitting Rate

Location Description	Corrosion Rate (mpy)	Pitting Rate (mpy)	SRB (Counts/cm ²)	GAB (Counts/cm ²)	GAnB (Counts/cm ²)
Dry Crude Outlet from Dual Tank to Export Pump	2	10	3800	380	3800
Dry Crude Outlet from Dual Tank	21	36	3800	380	3800
Dry Crude Outlet from Dual Tank	9	8	3800	0	0

Table 2. Effluent water Corrosion/Pitting Rate

Location Description	Corrosion Rate (mpy)	Pitting Rate (mpy)	SRB (Counts/cm ²)	GAB (Counts/cm ²)	GAnB (Counts/cm ²)
Effluent Water Outlet from 1st Stage Desalter Train-A to Heat Exchanger	5	269	17	17000	17000
Balance Tank Outlet	7	515	17	17000	17000
Effluent Water Outlet from Transfer Pump	4	213	27	27000	27000
Effluent Water Inlet to Dispatch Pumps from Effluent Water Tank	7	25	27	27000	27000



Effluent Water (EFW)

The heat map of effluent water is given in Figure 3 and the table comprising of corrosion and pitting rate is given in Table 2. The table shows the values of various bacteria as well. Figure 4 depicts the coupons of effluent water. Generally effluent water and produced water have high chlorides, due to high “chlorides”, presence of SRB is minimum or doesn’t exist. GAB is the main cause and next is GAnB, whereas SRB is minimum as it cannot have

Wet Crude (CRW)

Presence of GAnB causes pitting rate of corrosion where SRB is minimum in wet crude. This concept was not clear in the literature available, now it is proved and the concept is clear. The heat map of wet crude is given in Figure 5 and the table comprising of corrosion and pitting rate is given in Table 3. The table shows the values of various bacteria as well. Figure 6 depicts the coupons of wet crude.

Figure 5: Heat map of wet crude

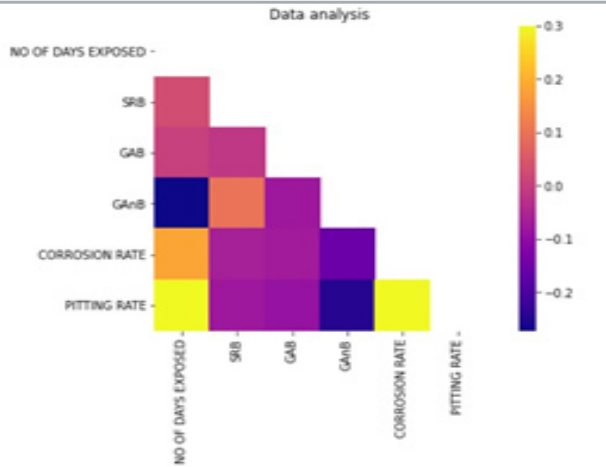


Table 3. Wet Crude Corrosion/Pitting Rate

Location Description	Corrosion Rate (mpy)	Pitting Rate (mpy)	SRB (Counts /cm2)	GAB (Counts /cm2)	GAnB (Counts /cm2)
Crude Outlet from Wet Tank T-001 to Train C	120	266	0	170	17000
Crude Outlet from Wet Tank T-001 to Train C	143	180	38	170	17000
Crude Outlet from Wet Tank	117	52	38	380	38000
Crude Inlet to Dual Tank	117	55	38	380	38000

Figure 6: Coupons of wet crude



Wet Gas (HWG)

Due to droplet formation, deposits are not formed on the coupons as water get condenses. Observations from the plot reveal existence of all types of bacteria. SRB effect is moderate in presence of high number of colonies of GAB and GAnB, which cause very severe pitting corrosion. Minimal amounts of solid exist in gas, therefore SRB, GAB and GAnB traces are found in the wet gas. The heat map of wet gas is given in Figure 7 and the table comprising of corrosion and pitting rate is given in Table 4. The table shows the values of various bacteria as well. Figure 8 depicts the coupons of wet gas.

Figure 7: Heat map of wet gas

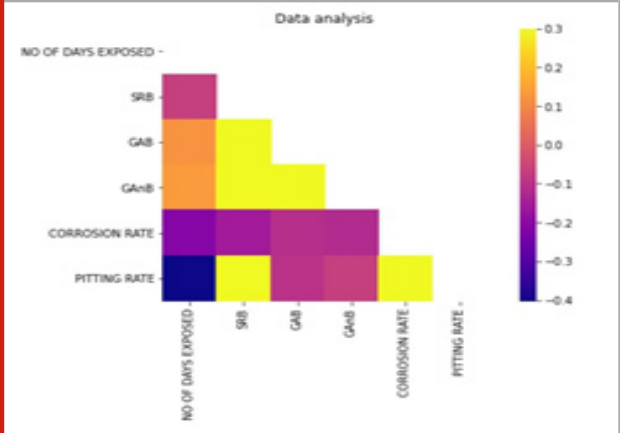
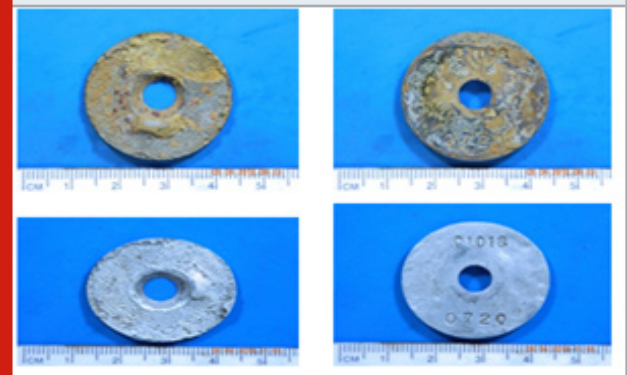


Table 4. Wet gas Corrosion/Pitting Rate

Location Description	Corrosion Rate (mpy)	Pitting Rate (mpy)	SRB (Counts /cm2)	GAB (Counts /cm2)	GAnB (Counts /cm2)
Gas Outlet from Dual Tank	18	286	3800	380	38000
TV Gas to LP Flare	31	56	3800	380	3800
TV Gas to LP Flare	31	78	380	380	3800

Figure 8: Coupons of wet gas



Final Effluent Water (FEFW)

Due to the presence of maximum chlorides, both corrosion rate and pitting rate is minimal. No traces of SRB present. Pitting rate is more due to below mentioned two reasons: i) Presence of high chlorides ii) Presence of GAnB Hence it can be observed that GAnB is a bad actor, and presence of it damages asset integrity. Recorded in literature and now proved practically. The heat map of final effluent water is given in Figure 9 and the table comprising of corrosion and pitting rate is given in Table 5. The table shows the values of various bacteria as well. Figure 10 depicts the coupons of final effluent water.

Figure 9 : Heat map of final effluent water

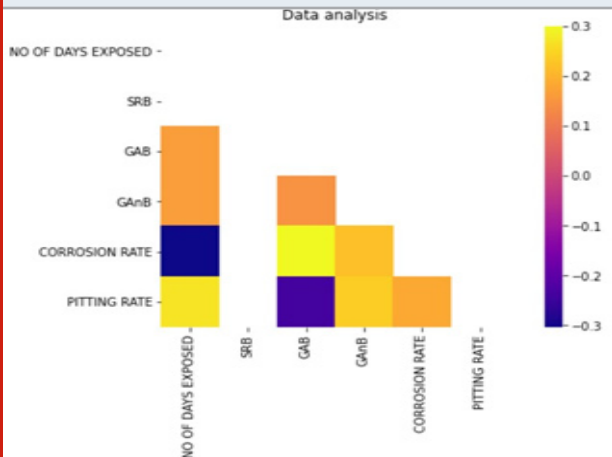
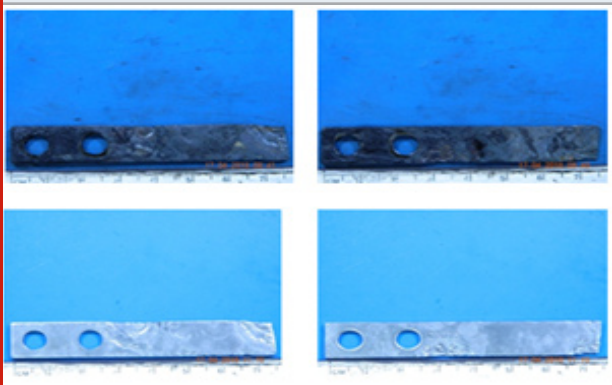


Table 5: Final effluent water Corrosion/Pitting Rate

Location Description	Corrosion Rate (mpy)	Pitting Rate (mpy)	SRB (Counts /cm2)	GAB (Counts /cm2)	GAnB (Counts /cm2)
Effluent Water Inlet to Dispatch Pump	3	72	0	27	27000
Effluent Water Inlet to Dispatch Pump	1	68	0	270	27000
Effluent Water Outlet from Dispatch Pump	12	49	0	270	27000

Figure 10: Coupons of final effluent water



Produced Water (PDW)

The heat map of produced water is given in Figure 11 and the table comprising of corrosion and pitting rate is given in Table 6. The table shows the values of various bacteria as well. Figure 12 depicts the coupons of produced water.

Recycled Water (RCW)

The heat maps of recycled water is given in Figure 13 and the table comprising of corrosion and pitting rate is given in Table 7. The table shows the values of

various bacteria as well. Figure 14 depicts the coupons of recycled water.

Figure 11: Heat map of produced water

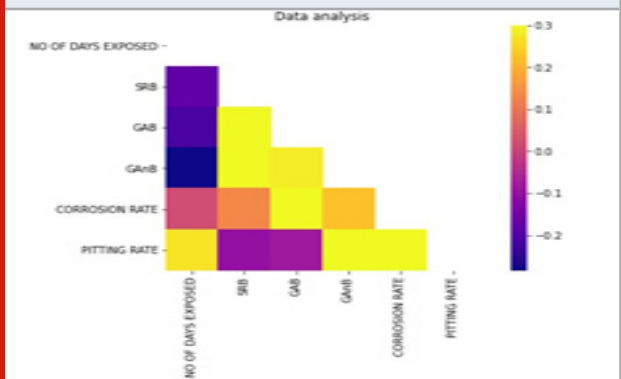


Table 6. Produced water Corrosion/Pitting Rate

Location Description	Corrosion Rate (mpy)	Pitting Rate (mpy)	SRB (Counts /cm2)	GAB (Counts /cm2)	GAnB (Counts /cm2)
Oily Water Outlet from LP Wet Separator	20	10	2700	1700	27
Oily Water Outlet from LP Wet separator	15	10	2700	1700	2
Balance Tank 07-D-801	16	10	2700	1700	27
Oily Water Outlet from New LP Wet separator (3-Phase)	23	7	2700	270	27

Figure 12: Coupons of produced water

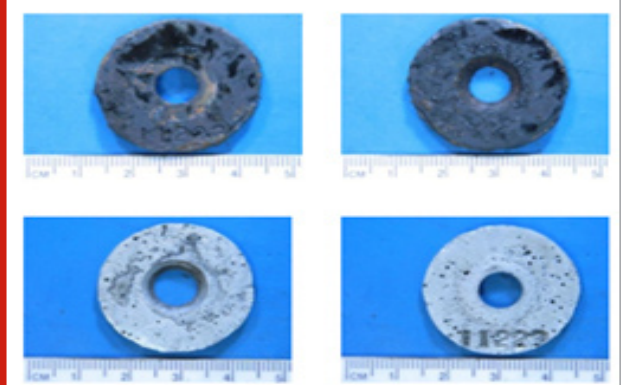


Table 7. Recycled water Corrosion/Pitting Rate

Location Description	Corrosion Rate (mpy)	Pitting Rate (mpy)	SRB (Counts /cm2)	GAB (Counts /cm2)	GAnB (Counts /cm2)
Recycle Water Outlet from 2 nd Stage Desalter Train-C	26	7	270	270	27
Recycle Water Outlet from 2 nd Stage Desalter Train-D to RCW Pump	109	6	270	270	27
Recycle Water Outlet from 2 nd Stage Desalter Train D to Recycled Water Pump	39	8	270	270	27
Desalter Train A Recycled Water Suction Pump	86	5	270	270	27

Figure 13: Heat maps of recycled water

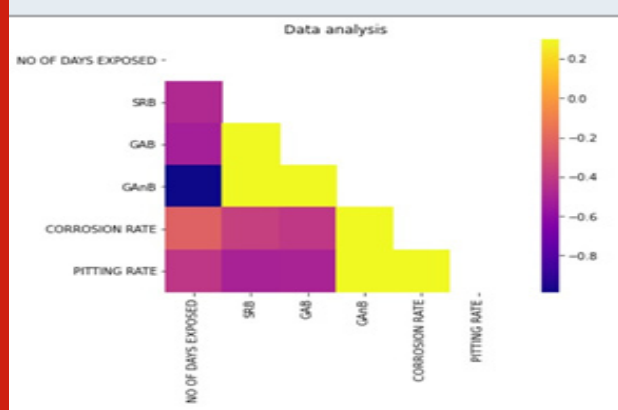
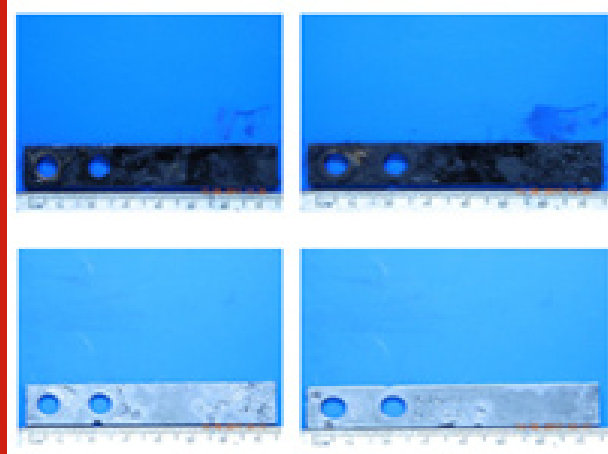


Figure 14: Coupons of recycled water



CONCLUSION

Corrosion analysis justifies that all analysis made are accurate with respect to the theory concepts, and the observations which are recorded are proved by the analysis. GAnb has the predominant effect on pitting rate and corrosion rate if present in large number in given environment. SRB contributes more to general corrosion rate and in presence of GAnB the effect of SRB can be minimized. If GAB is present along with SRB, the effect of GAnB is further reduced. The factors considered are humidity, chloride deposits with respect to corrosion and pitting rate. The effort in this paper is to prove the facts available in the literature.

REFERENCES

- Alcantara J, de la Fuente D, Chico B, et al. Marine atmospheric corrosion of carbon steel: a review. *Materials*. 2017;10(4):406.
- Cai YK, Zhao Y, Ma XB, et al. Application of hierarchical linear modelling to corrosion prediction in different atmospheric environments. *Corros Eng Sci Technol*. 2019;54(3):266–275.
- Cai Y, Zhao Y, Ma X, et al. Influence of environmental

- factors on atmospheric corrosion in dynamic environment. *Corros Sci*. 2018;137:163–175.
- Chico B, de la Fuente D, Diaz I, et al. Annual atmospheric corrosion of carbon steel worldwide. An integration of ISOCORRAG, ICP/UNECE and MICAT databases. *Materials*. 2017;10:601.
- Diaz I, Cano H, de la Fuente D, et al. Atmospheric corrosion of Ni-advanced weathering steels in marine atmospheres of moderate salinity. *Corros Sci*. 2013;76:348–360.
- Dong J, Han E, Ke W. Introduction to atmospheric corrosion research in China. *Sci Technol Adv Mat*. 2007;8(7–8):559–565.
- Hoerle S, Mazaudier F, Dillmann P, et al. Advances in understanding atmospheric corrosion of iron. II. Mechanistic modelling of wet-dry cycles. *Corros Sci*. 2004;46(6):1431–1465.
- Kamrunnihar M, Urquidi-Macdonald M. Prediction of corrosion behaviour of alloy 22 using neural network as a data mining tool. *Corros Sci*. 2011;53 (3):961– 967
- Lee S, Peng J, Shin D, et al. Data analytics approach for melt-pool geometries in metal additive manufacturing. *Sci Technol Adv Mat*. 2019;20(1):972–978.
- Lien LTH, San PT, Hong HL. Results of studying atmospheric corrosion in Vietnam 1995–2005. *Sci Technol Adv Mat*. 2007;8(7–8):552–558.
- Morcillo M, Chico B, Diaz I, et al. Atmospheric corrosion data of weathering steels. A review. *Corros Sci*. 2013;77:6–24.
- Natesan M, Selvaraj S, Manickam T, et al. Corrosion behavior of metals and alloys in marine-industrial environment. *Sci Technol Adv Mat*. 2008;9(4):045002.
- Nishimura T. Corrosion resistance of Si–Al-bearing ultrafine-grained weathering steel. *Sci Technol Adv Mat*. 2008;9(1):013005.
- Panchenko YM, Marshakov AI. Long-term prediction of metal corrosion losses in atmosphere using a powerlinear function. *Corros Sci*. 2016;109:217–229.
- Panchenko YM, Marshakov AI, Igonin TN, et al. Longterm forecast of corrosion mass losses of technically important metals in various world regions using a power function. *Corros Sci*. 2014;88:306–316.
- Pruksawan S, Lambard G, Samitsu S, et al. Prediction and optimization of epoxy adhesive strength from a small dataset through active learning. *Sci Technol Adv Mat*. 2019;20(1):1010–1021.
- Shin D, Lee S, Shyam A, et al. Petascale supercomputing to accelerate the design of hightemperature alloys. *Sci Technol Adv Mat*. 2017;18(1):828–838.
- Shi YN, Fu DM, Zhou XY, et al. Data mining to online galvanic current of zinc/copper internet atmospheric corrosion monitor. *Corros Sci*. 2018;133:443–450.
- Soares CG, Garbatov Y, Zayed A, et al. Influence of environmental factors on corrosion of ship structures in marine atmosphere. *Corros Sci*. 2009;51 (9):2014–2026.