Taylor & Francis
Taylor & Francis Group

Science and Technology of Advanced Materials: Methods

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tstm20

Predicting the surface roughness of an electrodeposited copper film using a machine learning technique

Ryo Tamura, Ryuichi Inaba, Mami Watanabe, Yutaro Mori, Makoto Urushihara, Kenji Yamaguchi & Shoichi Matsuda

To cite this article: Ryo Tamura, Ryuichi Inaba, Mami Watanabe, Yutaro Mori, Makoto Urushihara, Kenji Yamaguchi & Shoichi Matsuda (2024) Predicting the surface roughness of an electrodeposited copper film using a machine learning technique, Science and Technology of Advanced Materials: Methods, 4:1, 2416889, DOI: 10.1080/27660400.2024.2416889

To link to this article: https://doi.org/10.1080/27660400.2024.2416889

SCIENCE AND TECHNOLOGY OF ADVANCED MATERIALS: METHODS
2024, VOL. 4, NO. 1, 2416889
https://doi.org/10.1080/27660400.2024.2416889

Taylor & Francis
Taylor & Francis Group

Predicting the surface roughness of an electrodeposited copper film using a machine learning technique

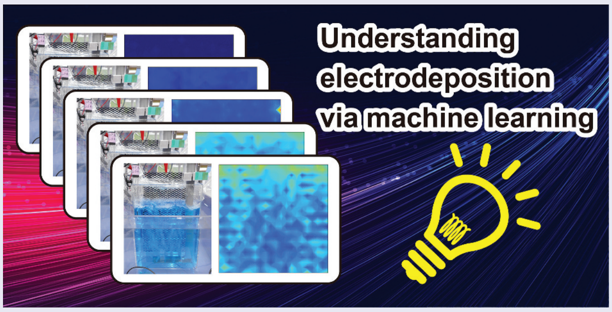
Ryo Tamura[a,b], Ryuichi Inaba[c], Mami Watanabe[c], Yutaro Mori[c], Makoto Urushihara[c], Kenji Yamaguchi[c] and Shoichi Matsuda[d]

[a]Center for Basic Research on Materials, National Institute for Materials Science, Tsukuba, Ibaraki, Japan; [b]Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba, Japan; [c]Innovation Center, Mitsubishi Materials Corporation, Naka-shi, Ibaraki, Japan; [d]Research Center for Energy and Environmental Materials (GREEN), National Institute for Materials Science, Tsukuba, Ibaraki, Japan

ABSTRACT
Electrodeposition-based metal coating techniques are used to manufacture various industrial products and rely on the quantitative control of the physical properties of the coating layers, such as electrical conductivity, surface roughness, and hardness. To clarify the experimental conditions required to realize the desired physical properties of metal coating layers and shed light on the complex mechanism of the involved reactions, we prepared a custom-built experimental dataset (60 conditions) on the surface roughness of electrodeposited thin copper films and submitted it to an open-access data repository. Data-driven analysis revealed that surface roughness is strongly affected by the deposition temperature, current, and interelectrode distance.

ARTICLE HISTORY
Received 10 April 2024
Revised 8 September 2024
Accepted 9 October 2024

KEYWORDS
Electrodeposition; machine learning; copper film; electrochemistry; surface roughness



IMPACT STATEMENT
To understand the relationship between the experimental conditions and surface roughness of electrodeposited copper thin films, the custom-built dataset for electrodeposition was prepared and data-driven approach was applied.

1. Introduction

Electrodeposition is a metal coating technique widely used to manufacture various industrial products, such as electronic devices and automobiles [1]. In particular, complicated metal coating techniques based on electrodeposition methods have been extensively researched owing to the increasing demand in the semiconductor industries [2,3]. In such applications, the ability to quantitatively control the physical properties of metal coating layers, such as electrical conductivity, surface roughness, and hardness, is crucial. The physical properties of the electrodeposited metal layer are largely influenced by the electrolyte composition, current density, temperature, stirring conditions, and other parameters [4]. Consequently, the electrodeposition mechanism is complex and characterized by an interplay of numerous factors.

To shed light on the above mechanism, considerable attention has been directed at process simulations based on physicochemical reaction models [5,6]. However, these simulations are limited by the use of model systems where the interaction of each element is simplified to minimize calculation costs. Thus, the applicability of the simulation model – based

CONTACT Ryo Tamura tamura.ryo@nims.go.jp Center for Basic Research on Materials, National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan; Kenji Yamaguchi kyam@mmc.co.jp Innovation Center, Mitsubishi Materials Corporation, 1002-14, Mukohyama, Naka-Shi 311-0102, Japan; Shoichi Matsuda matsuda.shoichi@nims.go.jp Research Center for Energy and Environmental Materials (GREEN), National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan
Supplemental data for this article can be accessed online at https://doi.org/10.1080/27660400.2024.2416889

approach to practical electrodeposition processes with numerous experimental variables is limited. Actually, most previous studies relied on the trial-and-error method to clarify the specific experimental condition required to realize metal coating layers with the desired physical properties.

Data-driven techniques have a leading position in materials science, enabling the extraction of the relationships between experimental parameters as explanatory variables and physical properties as objective functions [7–9]. This approach has also been successfully used in electrochemistry, particularly in the development of energy storage devices and electrocatalysts [10–13] and establishment of correlations between the thickness and hardness of thin electrodeposited metal films [14,15].

Herein, we focused on the surface roughness of electrodeposited metal films (Figure 1) in view of its importance for device fabrication (surface roughness reduction results in improved filling processes [16] and reduced electrical losses [17,18]). The related data-driven research has focused on alloy systems from a manufacturing viewpoint [19,20], whereas no datasets on the surface roughness of electrodeposited films are freely available. To bridge this gap, we prepared a custom-built experimental dataset of the surface roughness parameters (SRPs) of thin copper films electrodeposited under 60 conditions and published it in an open-access data repository (MDR, https://doi.org/10.48505/nims.4473). A data-driven approach with correlation analysis, logistic regression, and random forest regression models were used to extract important factors for understanding the nature of the surface roughness of electrodeposited thin films.

2. Experimental

2.1. Electrodeposition

Electrodeposited copper films were obtained via constant-current electrolysis in a $CuSO_4$ (140–220 g/L)/H_2

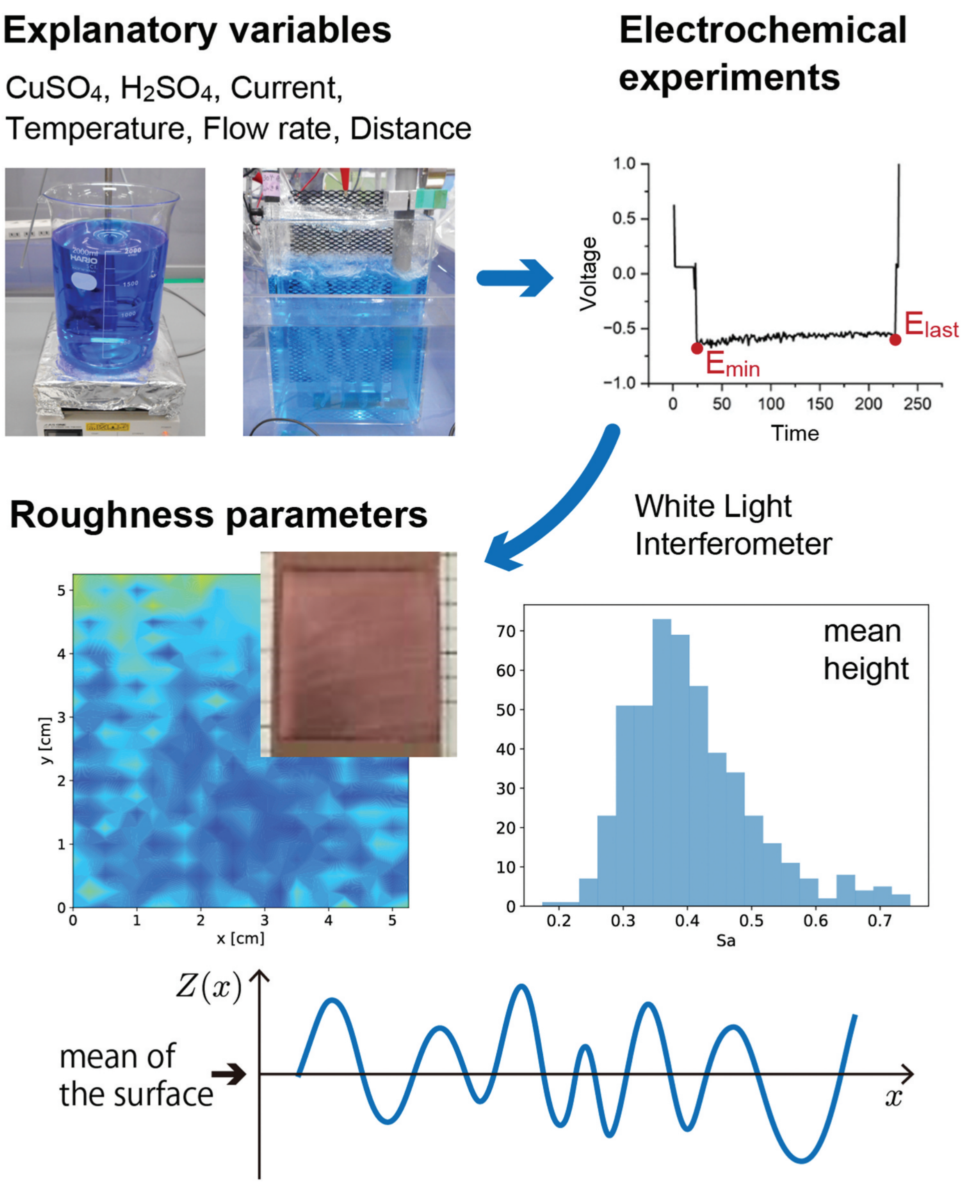


Figure 1. Schematic of the setup used to electrodeposit thin copper films and surface roughness measurements.

SO_4 (20–180 g/L) bath. Standard additives (JCU, CU-BRITE_RF) and 50 ppm chloride ions were introduced into the bath to ensure the stable formation of electrodeposited copper films. A 70 mm × 70 mm masked copper plate was used as the working electrode, and a Ti-coated mesh was used as the counter electrode. The electrodeposition time was controlled in such a way so as to achieve a total capacity of 8.1 C/cm^2. The current density, bath temperature, flow rate of the bubbled gas, and interelectrode distance were varied. Higher current densities generally result in faster film formation, and higher temperatures improve the electrodeposition uniformity by increasing conductivity [21]. In addition, the flow rate affects the nature of the stirring process and, hence, the mode of ion attachment to the electrodes. Finally, the interelectrode distance affects the shape of the current distribution between the electrodes and affects the uniformity of electrodeposition [22]. As all of these parameters are relevant to film formation, an understanding of their correlations with surface roughness is vital for comprehending electrodeposition.

2.2. Measurement of surface roughness

The surface roughness of the electroplated copper films was measured using a white light interferometer (AMETEK, USA) for areas of 0.79 mm × 0.79 mm using a 22 × 22 mesh grid with a 2.5 mm interval. In total, 484 (= 22 × 22) points were obtained for each sample. The difference in the height of each point compared with the surface mean was defined as $Z(x, y)$ [μm], where x and y are the two-dimensional Cartesian coordinates in a 0.79 mm × 0.79 mm area (Figure 1). By definition, $Z(x, y)$ can be positive or negative.

2.3. Data-driven analytical techniques

Pearson's correlation coefficient (r) was used for correlation analysis to check one-by-one correlations, with $r > 0$ and $r < 0$ indicating positive and negative linear correlations, respectively. Generally, moderate correlations were observed at $|r| > 0.3$. In addition, logistic regression and random forest regression models were adopted as machine learning models. A classification model for predicting the presence or absence of film formation and regression model for predicting SRPs using experimental parameters were developed. The use of random forest models allows one to assess the importance of each feature in the prediction process. Data standardization was performed during model training. The logistic regression and random forest regression models were trained using the Scikit-learn package [23].

3. Results and discussion

3.1. Conditions for the (non)adherence of electrodeposited copper films

The concentrations of H_2SO_4 and $CuSO_4$ in the electrolyte (five each) were set as explanatory variables (Table 1). Other explanatory variables included process parameters, namely the current density (three levels), temperature (six levels), gas bubbling flow rate (six levels), and interelectrode distance (three levels) (Table 1). Based on these variables, 60 experimental conditions were randomly selected for electrodeposition experiments. In 13 cases, the electrodeposited copper film did not adhere to the substrate; thus, we addressed the conditions for (non)adherence using the data-driven approaches discussed in this section.

To verify the prediction accuracy of the machine learning model determining the presence or absence of film formation and extract the corresponding relationship with the explanatory variables, we used a logistic regression for the binary classification of adherence/nonadherence. Leave-one-out cross-validation was performed to evaluate prediction accuracy, and in each validation calculation, the regularization hyperparameter of the logistic regression was determined by five-fold cross-validation, that is, nested cross-validation was used. The confusion matrix is shown in Table 2 (accuracy = 0.833). The coefficients for each explanatory variable obtained when all data were used as training data are summarized in Figure 2. The accuracy for training data equaled 0.95. The results revealed that the $CuSO_4$ concentration was the most important feature for binary classification and had to be high for successful copper film formation. The experimental conditions resulting in the nonadhesion of the electrodeposited copper films are shown in Table 3. Among the 13 experimental conditions, 11 had the lowest concentration of $CuSO_4$ (20 g/L). In the remaining two conditions (ID 48, 49), the flow rate was zero, although the $CuSO_4$ concentration was sufficient. The logistic regression analysis identified the flow rate as the factor with the second largest positive

Table 1. Experimental parameters used as explanatory variables.

$CuSO_4$ g/L	H_2SO_4 g/L	Current A/dm^2	Temperature °C	Flow rate L/min	Distance cm
20	140	2	19	0	5
60	180	4	22	5	7.5
80	200	6	25	7.5	10
180	220		28	10	
			30	12.5	
			31	15	

Table 2. Confusion matrix for logistic regression – based predictions of film (non)production obtained using the leave-one-out cross-validation.

		Predicted	
		Nonadhesion	Adhesion
Actual	Nonadhesion	8	5
	Adhesion	5	42

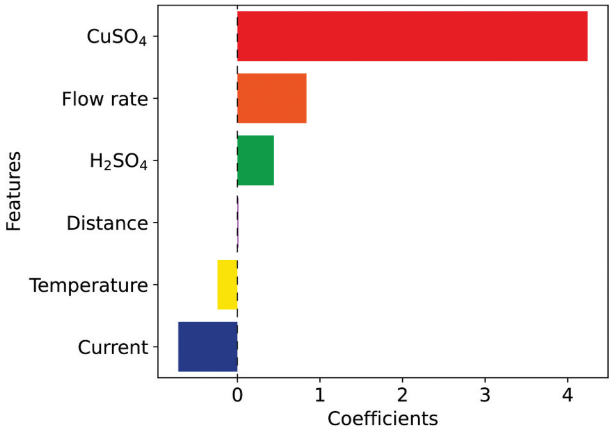


Figure 2. Coefficients of the logistic regression predicting film production/nonproduction. All data are used as training data, and the accuracy for the training data is 0.95.

coefficient. These results revealed that the adhesion of electrodeposited copper films requires a $CuSO_4$ concentration above the minimal value and a nonzero flow rate. This conclusion, obtained by machine learning analysis, was essentially consistent with the current state of knowledge in electrodeposition. Specifically, the high concentration of $CuSO_4$ and electrolyte flow are crucial for preserving a high concentration of Cu^{2+} at the electrode surface and thus realizing efficient electrodeposition. Note that the same conclusion was reached using the random forest regression model (Figure S1).

3.2. Relationship between experimental parameters and mean height of roughness

Subsequently, we considered the relationship between the experimental parameters and SRPs by focusing on the 47 cases of the electrodeposited films adhering to the substrate.

The mean height (S_a) was defined as

$$S_a = \frac{1}{A}\int\int_A |Z(x,y)|dxdy, \tag{1}$$

where A is the area ($0.79\,\text{mm} \times 0.79\,\text{mm}$). Given that $Z(x,y)$ can take both positive and negative values, the absolute values were considered to evaluate the mean height. Note that other SRPs can be used to evaluate surface roughness, with the results of the corresponding analysis summarized in Supplementary Note A. In total, 484 S_a values were calculated, as summarized in Figure 3a. Here, the IDs are labelled in the order of decreasing median S_a. The two-dimensional maps of S_a for each sample are summarized in Figure S2, and all data can be obtained from https://doi.org/10.48505/nims.4473. In the ideal case, the electrodeposited film should have a low median S_a with a small variance and no outliers. The standard deviation (std) and maximum (max) values of S_a were used to account for variance and outliers. For example, large variances were observed for samples ID46 and 47, and large outliers were observed for samples ID27 and 43. These samples were not favorable from the viewpoint of uniform coating formation.

Correlation analysis was performed to deeply understand the relationship between S_a and the experimental parameters, with the obtained Pearson's correlation coefficients summarized in Figure 3b and the distributions between the S_a values and experimental parameters summarized in Figure S3. The results revealed that among the six kinds of experimental parameters, current and interelectrode distance showed a positive correlation with median S_a, suggesting that the surface became smoother (lower S_a) when these parameters decreased. In addition, S_a was negatively correlated with temperature, i.e. high temperatures had a smoothing effect on the film surface. The results of statistical analysis were consistent with the knowledge of electrochemistry. Namely, the use of low currents, short interelectrode distances, and high temperatures can increase the uniformity of the Cu^{2+} concentration on the electrode

Table 3. List of experimental conditions resulting in film nonadhesion.

ID	$CuSO_4$	H_2SO_4	Current	Temperature	Flow rate	Distance
48	100	180	6	25	0	10
49	60	200	6	25	0	10
50	20	220	6	25	12.5	10
51	20	220	6	25	5	7.5
52	20	220	4	28	15	5
53	20	220	4	31	12.5	5
54	20	200	4	31	12.5	5
55	20	180	6	30	12.5	7.5
56	20	140	6	25	10	10
57	20	140	6	22	12.5	10
58	20	180	6	22	7.5	5
59	20	180	4	19	12.5	10
60	20	220	4	22	10	10

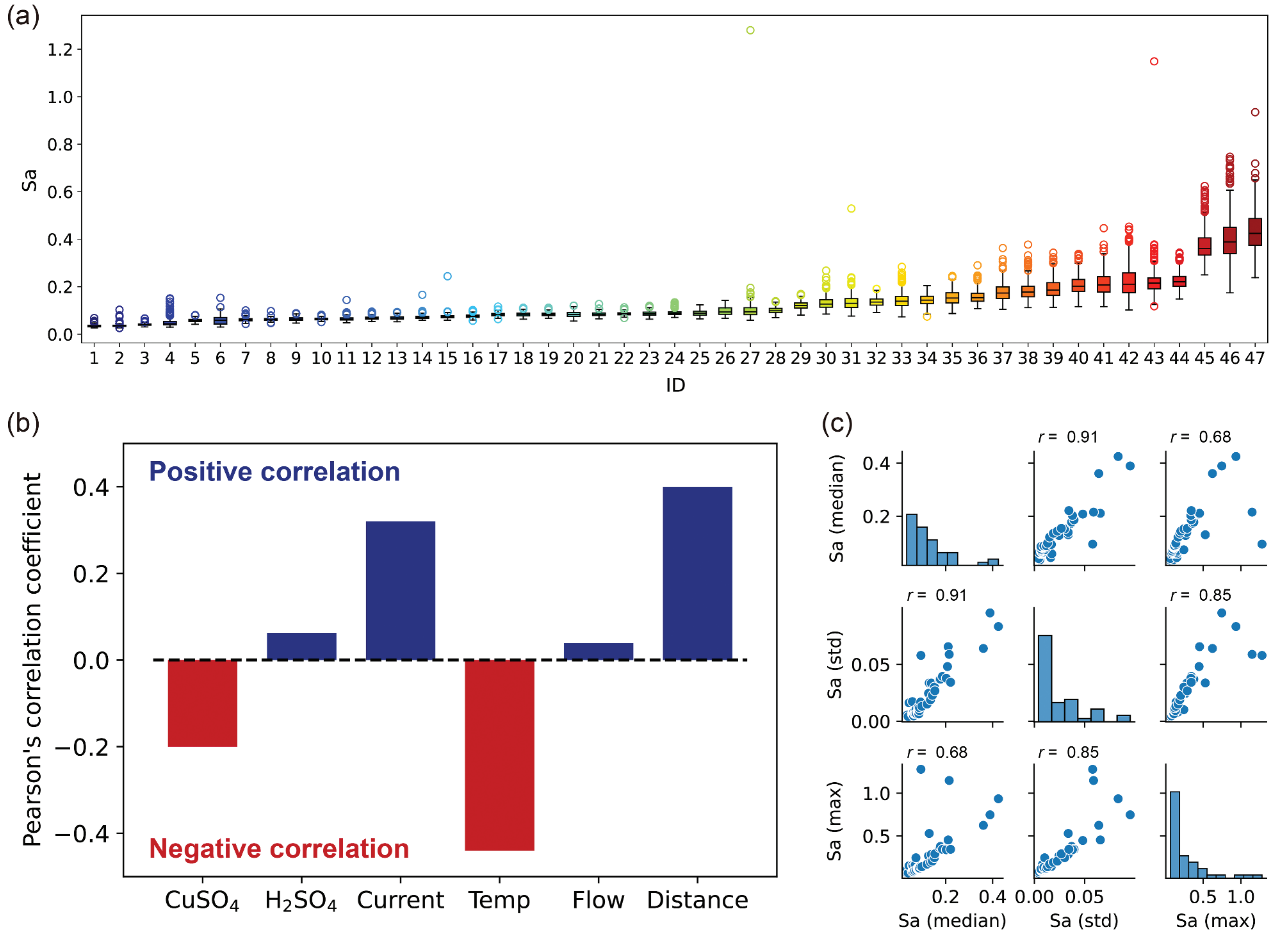


Figure 3. (a) Distribution of S_a for each sample. The IDs are listed in order of decreasing median S_a for 484 datapoints. (b) Pearson's correlation coefficients between the median S_a and experimental parameters. (c) Relationship between S_a statistics. r is the Pearson's correlation coefficient.

surface, which is a key parameter determining the surface roughness of the electrodeposited film.

We also performed statistical analysis for the std and max values of S_a, revealing similar relationships with the experimental parameters (Figure S4). Besides, strong positive correlations were confirmed between the S_a statistics (Figure 3c), that is, when the median value was minimized, a small variance and no outliers could be achieved simultaneously.

Next, we examined whether S_a could be predicted using the experimental parameters. For this purpose, the median, std, and max values of S_a were determined from the experimental parameters using a random forest regressor. To evaluate the prediction performance, the coefficient of determination (R^2) for the predicted and actual values was calculated using leave-one-out cross-validation. Figure 4a shows the results obtained when each S_a statistic on the normal and logarithmic scales was objective. Here, larger R^2 values indicate more accurate predictions. The results reveal that the prediction performance on a logarithmic scale is superior to that on a normal scale. Figure 4b presents the median values of S_a predicted using cross-validation. Compared with the distribution on the normal scale (Figure 3c), the data are more uniformly distributed on the logarithmic scale. These results reveal that the prediction accuracy on the logarithmic scale was superior to that on the standard scale and was therefore preferable for roughness parameter estimation.

To extract the relationship between the process parameters and surface roughness, we highlighted the importance of the features for predicting the median value of S_a when all data were used as the training data (Figure 4c). The R^2 value for the training data in the case of the random forest regression was 0.8703. The obtained values were not outstanding but still indicated that the relationship between these multiple process parameters determines film roughness. A similar tendency was observed for the standard deviation and maximum values of S_a (Figure S5).

3.3. Relationship between roughness parameters and electrochemical data

The accuracy of the predictions based on the process parameters discussed in Section 3.2 was not sufficient.

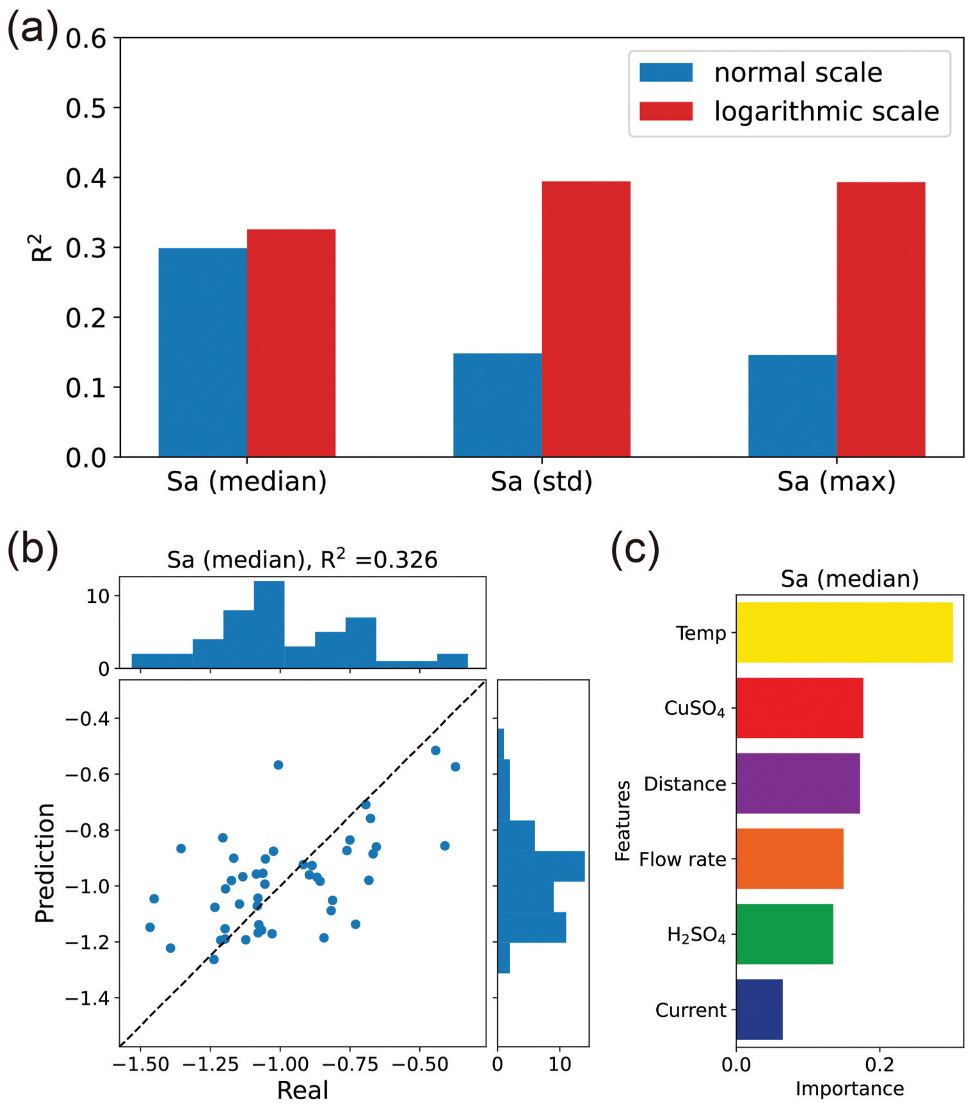


Figure 4. (a) Coefficients of determination (R^2) between predicted and true roughness parameters when the leave-one-out cross-validation is performed on normal and logarithmic scales. (b) Prediction results of median S_a obtained using a random forest regressor when the leave-one-out cross-validation is performed on logarithmic scales. (c) Feature importance of random forest regression models for logarithmic scales when median S_a values are predicted. To evaluate feature importance, the R^2 value for the training data was determined as 0.8703.

Therefore, in this section, we discuss the use of electrochemical data as a descriptor, which is not a parameter that can be controlled but is considered to reflect the experimental conditions. During electrodeposition, a constant current was applied, and the voltage was measured. As the voltage time course represents the electrochemical process, the information contained therein can be correlated with the surface roughness of the electrodeposited film. Based on this consideration, we investigated the relationship between the roughness parameters and electrochemical data (the lowest voltage during electrodeposition (E_{min}) and voltage at the end of the process (E_{last}); Figure 1). No correlation was observed between S_a and voltage (Figure 5a), as indicated by close-to-zero Pearson's correlation coefficients. In addition, we compared the prediction performance obtained using only six experimental parameters and those with E_{min} and E_{last}. Figure 5b shows the R^2 values for the correlation between the predicted and true roughness parameters when the leave-one-out cross-validation is performed with and without E_{min} and E_{last} on a logarithmic scale. Unfortunately, when electrochemical data were used as explanatory variables, the prediction performance decreased. Therefore, we concluded that the voltage value did not strongly contribute to the prediction of surface roughness.

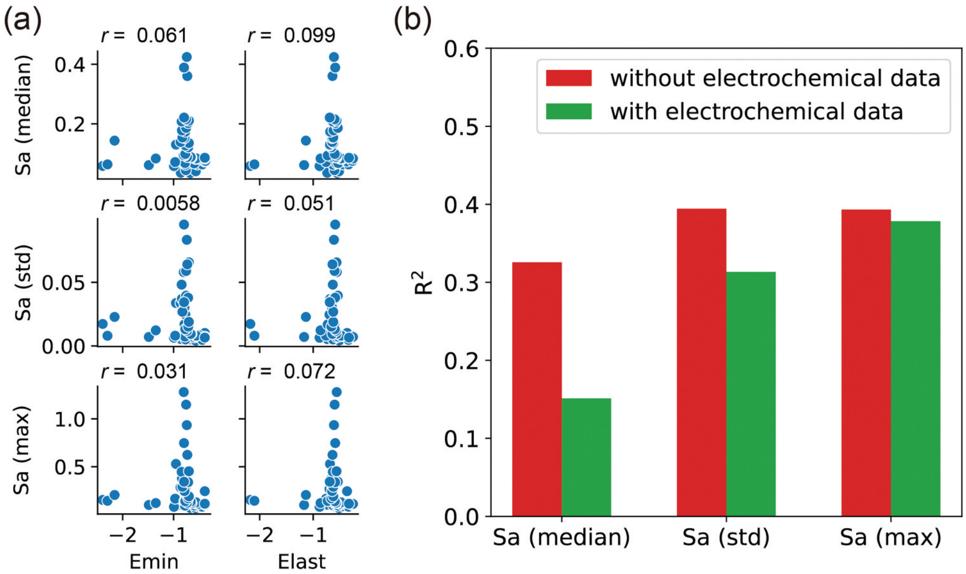


Figure 5. (a) Relationship between S_a and voltage. r is the Pearson's correlation coefficient. (b) R^2 between predicted and true roughness parameters when the leave-one-out cross-validation is performed with and without electrochemical data (E_{min} and E_{last}) on a logarithmic scale.

4. Conclusions

A comprehensive data-driven analysis was performed to identify the critical factors determining the surface roughness of electrodeposited thin copper films. Six experimental parameters were selected as explanatory variables, and a custom-built dataset was prepared for 60 experimental conditions. The $CuSO_4$ concentration and flow rate were found to strongly influence the adherence/nonadherence of the electrodeposited film to the substrate. In addition, the temperature, current, and interelectrode distance were strongly correlated with surface roughness. The use of a logarithmic scale instead of a normal one resulted in an enhanced prediction performance. To accumulate more data in the near future, based on the indications obtained herein, the use of a high-throughput experimental setup controlled by artificial intelligence may be effective [24,25], which is something that will be examined in our laboratory in due course. Obtaining additional data through this approach will allow more accurate information to be extracted using the methodology demonstrated herein. We believe that the present study provides a new direction for integrating electrodeposition with data-driven techniques.

Acknowledgements

This project was conducted at the Mitsubishi Materials Corporation (MMC) - National Institute for Materials Science (NIMS) Center of Excellence for Materials Informatics Research. The authors thank Kiyotaka Nakaya for helpful discussions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

The data that support the findings of this study are openly available in MDR at https://doi.org/10.48505/nims.4473.

References

[1] Winand R. Electrodeposition of metals and alloys—new results and perspectives. Electrochim Acta. 1994;39(8–9):1091–1105. doi: 10.1016/0013-4686(94)E0023-S
[2] Giurlani W, Zangari G, Gambinossi F, et al. Electroplating for decorative applications: recent trends in research and development. Coatings. 2018;8(8):260. doi: 10.3390/coatings8080260
[3] Tao P, Chen Y, Cai W, et al. Effect of copper sulfate and sulfuric acid on blind hole filling of HDI circuit boards by electroplating. Materials. 2021;14(1):85. doi: 10.3390/ma14010085
[4] Gabrielli C, Moçotéguy P, Perrot H, et al. Mechanism of copper deposition in a sulphate bath containing chlorides. J Electroanal Chem. 2004;572(2):367–375. doi: 10.1016/j.jelechem.2004.01.025
[5] Chen Z, Liu S. Simulation of copper electroplating fill process of through silicon via. In: 2010 11th International Conference on Electronic Packaging Technology & High Density Packaging; China. 2010. p. 433–437.
[6] Leiden A, Kölle S, Thiede S, et al. Model-based analysis, control and dosing of electroplating electrolytes. Int J Adv Manuf Technol. 2020;111(5–6):1751–1766. doi: 10.1007/s00170-020-06190-0
[7] Rajan K. Materials informatics. Mater Today. 2005;8(10):38–45. doi: 10.1016/S1369-7021(05)71123-8

[8] Agrawal A, Choudhary A. Perspective: materials informatics and big data: realization of the “fourth paradigm” of science in materials science. APL Mater. 2016;4(5):053208. doi: 10.1063/1.4946894
[9] Terayama K, Sumita M, Tamura R, et al. Black-box optimization for automated discovery. Acc Chem Res. 2021;54(6):1334–1346. doi: 10.1021/acs.accounts.0c00713
[10] Ng M-F, Zhao J, Yan Q, et al. Predicting the state of charge and health of batteries using data-driven machine learning. Nat Mach Intell. 2020;2(3):161–170. doi: 10.1038/s42256-020-0156-7
[11] Mistry A, Franco AA, Cooper SJ, et al. How machine learning will revolutionize electrochemical sciences. ACS Energy Lett. 2021;6:1422–1431. doi: 10.1021/acsenergylett.1c00194
[12] Lombardo T, Duquesnoy M, El-Bouysidy H, et al. Artificial intelligence applied to battery research: hype or reality? Chem Rev. 2022;122(12):10899–10969. doi: 10.1021/acs.chemrev.1c00108
[13] Sakaushi K, Hoisang W, Tamura R. Human–machine collaboration for accelerated discovery of promising oxygen evolution electrocatalysts with on-demand elements. ACS Cent Sci. 2023;9(12):2216–2224. doi: 10.1021/acscentsci.3c01009
[14] Kilic A, Yildirim R, Eroglu D. Machine learning analysis of Ni/SiC electrodeposition using association rule mining and artificial neural network. J Electrochem Soc. 2021;168(6):062514. doi: 10.1149/1945-7111/ac0aaa
[15] Katirci R, Aktas H, Zontul M. The prediction of the ZnNi thickness and Ni % of ZnNi alloy electroplating using a machine learning method. Trans IMF. 2021;99(3):162–168. doi: 10.1080/00202967.2021.1898183
[16] Carpio R, Jaworski A. Review—management of copper damascene plating. J Electrochem Soc. 2018;166(1):D3072. doi: 10.1149/2.0101901jes
[17] Brist G, Hall S, Clouser S, et al. Non-classical conductor losses due to copper foil roughness and treatment. ECWC 10 Conference at IPC Printed Circuits Expo, S19-2-1; 2005.
[18] Hinaga S, Koledintseva M, Anmula PKR, et al. Effect of conductor surface roughness upon measured loss and extracted values of PCB laminate material dissipation factor. In: Proceedings of the IPC APEX EXPO; Las Vegas, NV; 2009.
[19] Elangovan M, Sakthivel NR, Saravanamurugan S, et al. Machine learning approach to the prediction of surface roughness using statistical features of vibration signal acquired in turning. Procedia Comput Sci. 2015;50:282–288. doi: 10.1016/j.procs.2015.04.047
[20] Zhang W. Surface roughness prediction with machine learning. J Phys: Conf Ser. 2021;1856(1):012040. doi: 10.1088/1742-6596/1856/1/012040
[21] Schlesinger M, Paunovic M. Modern electroplating. USA: John Wiley & Sons, Inc.; 2010.
[22] Matlosz M, Creton C, Clerc C, et al. Secondary current distribution in a hull cell: boundary element and finite element simulation and experimental verification. J Electrochem Soc. 1987;134(12):3015. doi: 10.1149/1.2100332
[23] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825–2830.
[24] Matsuda S, Lambard G, Sodeyama K. Data-driven automated robotic experiments accelerate discovery of multi-component electrolyte for rechargeable Li–O_2 batteries. Cell Rep Phys Sci. 2022;3(4):100832. doi: 10.1016/j.xcrp.2022.100832
[25] Tamura R, Tsuda K, Matsuda S. NIMS-OS: an automation software to implement a closed loop between artificial intelligence and robotic experiments in materials science. Sci Technol Adv Mater: Methods. 2023;3(1):2232297. doi: 10.1080/27660400.2023.2232297