Volume 147
Number 7
7 April 2022
Pages 1247-1518

Analyst

rsc.li/analyst

ISSN 0003-2654

ROYAL SOCIETY OF CHEMISTRY

PAPER
Yuya Nagai and Kenji Katayama
Prediction of the photoelectrochemical performance of hematite electrodes using analytical data

Analyst

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Analyst*, 2022, **147**, 1313

Prediction of the photoelectrochemical performance of hematite electrodes using analytical data

Yuya Nagai and Kenji Katayama *

Machine learning (ML) has been extensively utilized in various fields of chemistry, such as molecular design and optimization of the fabrication parameters of the material. However, there is still a difficulty in applying ML for devices/materials fabricated in a lab because plenty of data for accurate calculation are difficult to obtain due to the limited number of samples. As a promising energy-harvesting material, we have studied hematite electrodes for photocatalytic water splitting. Herein, we have examined the critical factors affecting the photoelectrochemical (PEC) performance by applying ML for a limited number of fabricated electrodes to reveal the origin of poor reproducibility of the performance. To find the dominant factors affecting the PEC performance, the feature values were directly extracted from analytical data such as X-ray diffraction, Raman, UV/vis and photoelectrochemical impedance spectroscopy (PEIS) measurements. The dominant factors for the performance were identified from the prediction analysis of the performance by ML. Two types of descriptors were examined; all the analytical data were included and those without the PEIS data, which had a high correlation with the photocurrent. The determination coefficients (R^2) of the prediction accuracy were >0.8 in both cases and the dominant features were identified for the improvement of PEC performance without any prior knowledge.

Received 7th February 2022,
Accepted 3rd March 2022

DOI: 10.1039/d2an00227b

rsc.li/analyst

Introduction

Machine learning (ML) and deep learning (DL) have been extensively utilized for molecular design, prediction of chemical reactions, and generation of crystal structures have been demonstrated,[1–5] and datasets, descriptors, and methods are summarized in a calculation framework.[6] Furthermore, generative models have been recently applied for predicting the structure or generating the data for learning, such as 3-D inorganic crystal structures[7] and microstructures of thin films based on the composition and synthetic conditions.[8] By combination with a robotic operation, the synthesis process has been automated.[9] ML has been applied for the prediction of chemical reactions and physical properties based on a large database. Furthermore, a large number of chemical structures are virtually generated and connected to the properties calculated by quantum chemical calculations. However, there is still difficulty in applying ML/DL for the improvement/optimization of chemical reactions, material preparation, and device development in which actual experimental data are necessary for training, but a limited number of the data are available for practical purposes.

In the development of materials/devices such as thermoelectric, optoelectric, and photoelectric devices, it takes a long time to fabricate them in practice. This is because the number of devices is often limited to 10^2 on the lab scale, which makes it difficult to apply ML for the prediction/optimization of physical/chemical properties. Furthermore, the device performance is determined not only by the chemical compositions but also by the thickness, roughness, quality, *etc.* These are varied by the experimental operation parameters (temperature, concentration, flux, *etc.*). For these reasons, a limited number of applications using ML are found in the development processes of actual devices.

DL and neural network approaches with plenty of data are not helpful in many cases for material/device optimization to understand the origin of the physical properties because it is difficult to find out the crucial descriptors in the network, especially from a limited number of data. General ML strategy is more appropriate to understand the relationship between the descriptors and the target values, and various selection methods from many types of descriptors have been proposed.[10] Various optimization techniques have been studied using the Bayesian optimization process *via* Gaussian regression for the data not based on a mathematical

Department of Applied Chemistry, Chuo University, Tokyo 112-8551, Japan.
E-mail: kkata@kc.chuo-u.ac.jp; Tel: +81-3-3817-1913

model.[11,12] For example, a cross-coupling reaction was optimized in terms of concentrations and catalysts with a small number of synthetic data.[13] Recently, Tamura *et al.* demonstrated the optimization of material properties using molecular descriptors and the experimentally obtained analytical data.[14] Also, the solubility data were also predicted with a combination of analytical data and molecular descriptors.[15] These ML approaches for a small number of data could provide understandable insights into the development of materials/devices.

We have studied promising energy-harvesting materials/devices. Photocatalytic devices are used to decompose contaminants and to clean the atmosphere, and they are also promising devices for water splitting into oxygen and hydrogen using solar energy.[16] A hematite (iron oxide) electrode is known as an oxygen-evolution photocatalyst with visible light absorption, and there have been many studies to improve the photoelectrochemical performance (PEC) by fabrication and modification of surfaces with passivation and addition of cocatalysts.[17,18] However, one of the major drawbacks of the hematite electrode is poor reproducibility of the performance, and the underlying reasons for this are often unclear.

Even though hematite electrodes contain 'Fe_2O_3' in their chemical composition, each photoelectrode possesses a variety of PEC performance. Thus, we used analytical data as an indicator for each photoelectrode, which could have unintended/unnoticed information of each sample. We used various analytical measurements such as X-ray diffraction, Raman, UV/vis, and photoelectrical impedance. The feature values were extracted directly from the analytical data to find out the dominant factors for the PEC performance by ML.

Analytical data such as spectral data have been used in combination with ML/DL; spectral data can be converted into material properties or *vice versa*. It was demonstrated that X-ray analytical data could be converted to a crystalline structure.[19] For example, XANES spectra were used to obtain the atomic arrangement.[20] This indicates that the analytical data could include the structural and physical features of materials. Many feature extraction techniques have also been developed, such as determinant estimation from the three-dimensional data (such as a hyperspectral image) with a combination of matrix decomposition.[21,22] Thus, we examined the critical factors that affect the PEC performance of hematite data with a variety of performances by applying ML using a limited number of fabricated photoelectrodes in combination with various analytical data.

Methods

The spectral/pattern data were preprocessed by the noise/background removal method, such as spline fitting and frequency filtering for the removal of high-frequency noise. The intensities, positions, widths and areas of peaks were extracted as feature values of analytical data without any prior (biased) information. The irrelevant features were determined by a comparison between the standard deviation of each feature and the noise amplitude, and they were eliminated to avoid overfitting. After selecting the descriptors from the spectra/patterns, they were listed as a table (X). The target value (Y) was a photocurrent value at a specific bias voltage. Y data were predicted using the descriptors, X using various ML models. The number of descriptors was reduced by understanding the dominant descriptors, based on the cosine-similarity map,[23] least absolute shrinkage and selection operator (LASSO), partial least squares (PLS), and stepwise regression. An ML model was constructed by support vector regression (SVR), Gaussian process regression (GPR), decision tree regression (DT), and random forest regression (RF). Five-fold cross-validation was used to avoid overfitting in the fitting process.

Hematite electrodes were fabricated by a solution-derived method.[24] An FTO substrate (~7 Ω sq^{-1}, Sigma-Aldrich or SOLARONIX) was cut into 2 × 3 cm pieces, and a piece of the substrate was immersed in a precursor solution of hematite except for the top region (*ca.* 0.5 cm) for wiring (0.15 M iron(III) chloride hexahydrate ($FeCl_3 \cdot 6H_2O$, 99.9%, Wako) and 1 M sodium nitrate ($NaNO_3$, 99.9%, Wako) at 100 °C for one hour and then sintered in an electric furnace to obtain a thin film of α-Fe_2O_3. The surface area was *ca.* 5 cm^2. Twenty-eight samples were prepared for the analyses as stated in the first paragraph of the Results and discussion section. They were prepared at different sintering temperatures ranging from 600 to 750 °C. In the following section, additional 47 samples were prepared at a sintering temperature of 650 °C. Seventy-five samples were used in total. A three-electrode setup was used for the photoelectrochemical measurements, and the hematite samples were measured with a platinum wire as the counter electrode and an Ag/AgCl electrode as the reference electrode, and the potential was converted into reversible hydrogen electrode (RHE). Linear-sweep voltammetry was performed in a KOH solution (pH = 13.61) at a scan rate of 0.01 V s^{-1} using a three-electrode setup under 1 sun conditions (100 mW cm^{-2}) to obtain the PEC performance. The photocurrent density at 1.23 V (*vs.* RHE) was used as a target photocurrent. The samples were analyzed by UV/vis spectroscopy (USB2000+, Ocean optics), photoelectrochemical impedance spectroscopy (PEIS) (Model 660A, BAS, and 1260A, Solatron), X-ray diffraction (XRD) (Ultima IV, Rigaku), and Raman spectroscopy (NRS-3000, JASCO or Lamda Vision, excitation wavelength: 532 nm). In the PEIS measurements, electrodes were measured at 0.83 V (*vs.* RHE) with an AC voltage amplitude of 5 mV in the frequency range from 0.001 to 10 000 Hz. The data number was 12 in one order of frequency. These analytical data were used for the explanation of the target value.

Results and discussion

Photocurrent prediction of photoanodes from the analytical data including the photoelectrochemical impedance data

Twenty-eight hematite photoanode samples were prepared. They were sintered at four different temperatures from 600 to

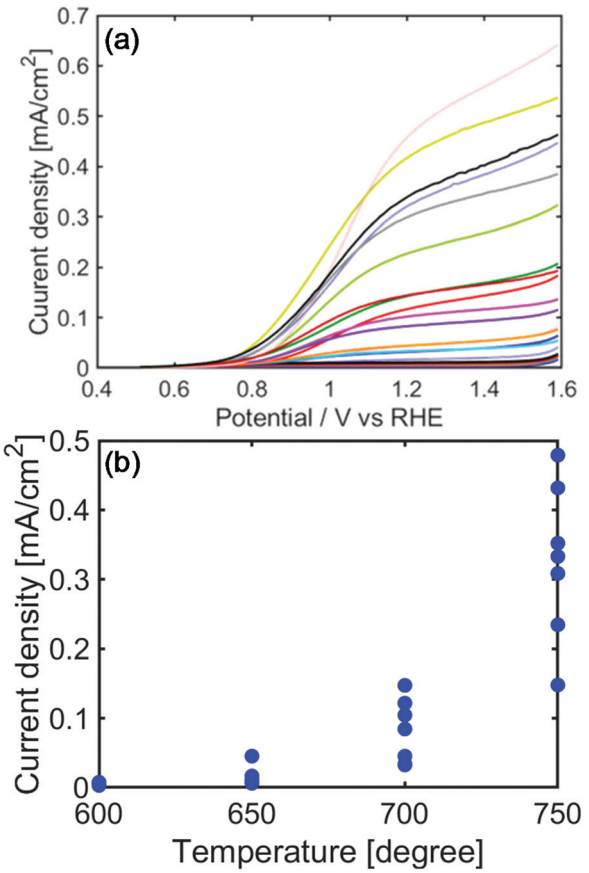

Fig. 1 (a) The current density – voltage curves for 28 samples of hematite photoanodes. The potential is given *versus* reversible hydrogen electrode (RHE). (b) The photocurrent densities at 1.23 V *vs.* RHE at different calcination temperatures.

750 °C. The current–potential curves for them are shown in Fig. 1(a). The photocurrent density at 1.23 V *vs.* RHE was used as a metric of the photoelectrochemical (PEC) performance (water oxidation potential). The photocurrent values of the photoanodes widely varied in a range from 2.8 to 450% from the average photocurrent. Although there was a general tendency that the photocurrent densities increased with the calcination temperature, they still showed a large variation even at the same calcination temperature, as shown in Fig. 1(b). All the samples were analyzed with PEIS, UV/vis spectroscopy, Raman spectroscopy, and XRD to represent the difference in the physical properties of the photoanodes.

The feature values were extracted from the analytical data after the removal of noise and background using general data processing such as smoothing and spline approximation for the background. The baseline was not removed for the UV/vis spectra because it had a correlation with the photocurrent density. After preprocessing, the intensities, positions, areas, and widths of the peaks were selected as feature values. Initially, we did not remove any specific peaks and included as many peaks as possible, and unnecessary features were automatically removed by the following descriptor selection processes by ML predictions. The features used in the analytical data are shown in Fig. 2.

Next, the feature values were refined because the features contributing to the photocurrent are limited. The irrelevant features to the photocurrent were removed if the standard deviation normalized by the average value was smaller than 1%. When they were smaller than the noise amplitude, they simply affect the calculation accuracy inversely by overfitting the prediction to the noise fluctuation, and it is better to eliminate them. For this reason, the number of features was reduced from 99 to 72. The excluded features were mainly from the location of the XRD and Raman peaks.

Then, the determinant factors for the photocurrent were investigated to reveal the origin of the photocurrent variation. In this process, the inverse values of the descriptors were added to the original descriptors to evaluate the inverse relationship between the descriptors and the performance. Several descriptor selection methods were tested, and the results by the stepwise regression are shown in Table 1. Linear regression models were constructed by changing the combination of descriptors and searching for the optimal combination based on the squared sum of the residual error. The response plot is shown in Fig. 3, where the predicted values by the model are plotted against the target values. The prediction accuracy was sufficiently high (determination coefficient: R^2 = 0.98), and clearly, the result indicates that the selected descriptors worked as the determining factors for the photocurrent.

For the refinement of the model, the nonlinear regression method was utilized to predict the performance with the selected five descriptors. As for a nonlinear prediction method, GPR, SVR, DT and RF regressions were examined, and GPR provided the highest prediction rate. The data derived from twenty-eight samples were divided into two parts with a ratio of 8 : 2. The former corresponds to the training data for making a prediction model, and the latter was used for model validation. Five-fold cross-validation was used; five different combinations of the training and test datasets were used to avoid overfitting to a specific dataset. Fig. 4 shows a response plot for the training and test data, where the caption (1st, 2nd, 3rd, 4th, and 5th) of the plots indicates all the different training and test datasets in the cross-validation. The average R^2 values of the five training and test data were 0.91 and 0.91, respectively. The same level of the determination coefficients for the test data with the training data assures accuracy without overfitting. Therefore, the model constructed by ML was sufficiently accurate only by the features extracted from the analytical data, even for a small number of samples.

In the selected features, three PEIS and one XRD features were selected. The features, R1 and R3, in PEIS represent resistances between a solution and a hematite electrode and between an FTO and a hematite electrode, respectively.[25] It is obvious that the efficiency of the photoelectrodes was dominated by interfacial conditions. The contact between the FTO and hematite electrodes was a more serious factor to improve in these sample sets because the coefficients shown in Table 1 for the R3 descriptors had larger coefficients than those for

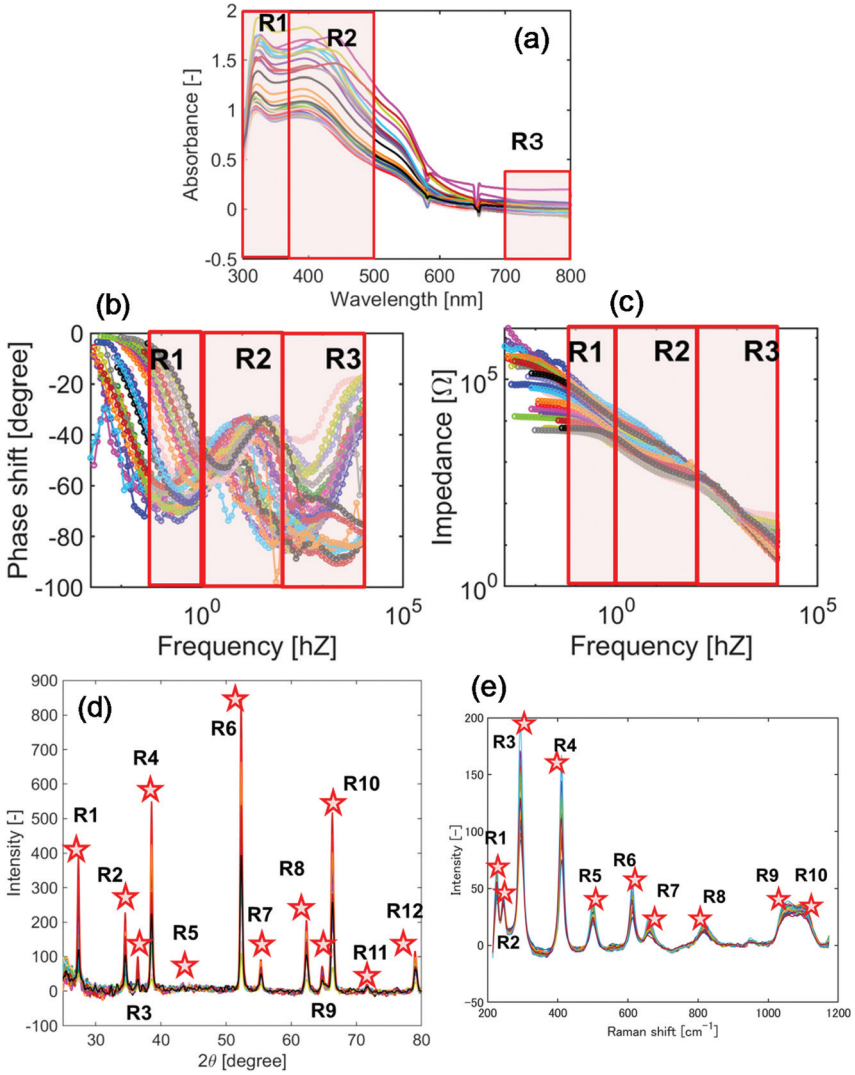

Fig. 2 The analytical data for hematite electrodes (28 samples) with an indication of feature values; (a) UV/vis spectra, (b) phase shift and (c) impedance in the PEIS data, (d) XRD patterns, and (e) Raman spectra. Each separated region is shown in red rectangles or stars with labels.

Table 1 Selected descriptors by the stepwise regression and the coefficient for each descriptor

Descriptors	Coefficient
inv_PEIS_phase_R1_max	−0.347
inv_PEIS_phase_R3_min	−1.063
inv_PEIS_imp_R1_max	0.410
inv_xrd_R3_pks	0.135
PEIS_phase_R3_min	−0.793

the R1 descriptors. Furthermore, we could find a minor correlation with a peak in the XRD, which corresponds to the (110) surface of hematite. This is consistent with a previous study indicating that the (110) surface is relevant to the hematite photocatalytic performance.[24] From these results, a reduction

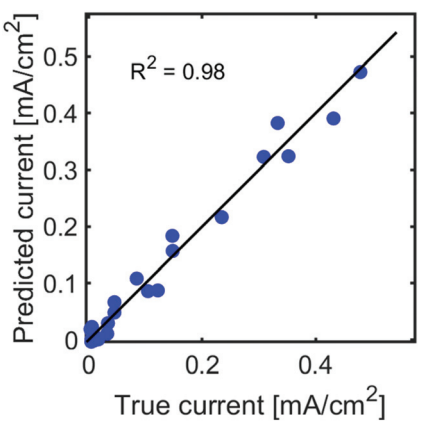

Fig. 3 The scatter plot for the true current and the predicted current.

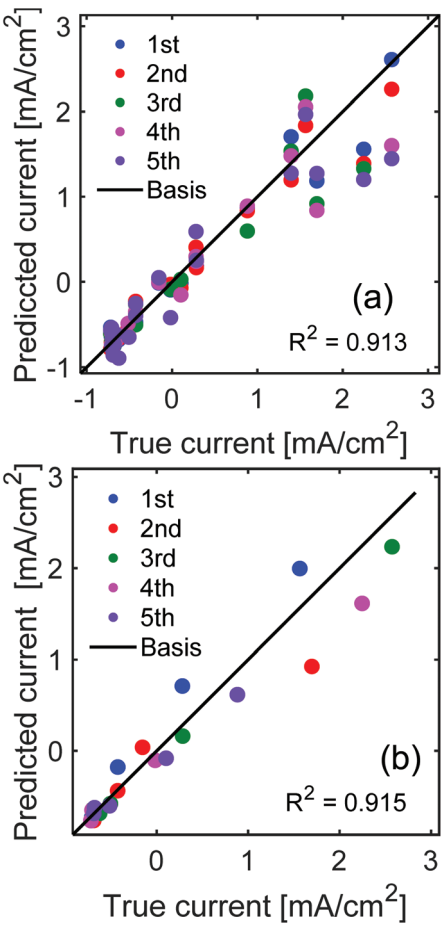

Fig. 4 The scatter plots of the true current densities *versus* the predicted values for hematite electrodes for the (a) training data and (b) test data. 1st, 2nd, 3rd, 4th, and 5th correspond to the datasets with different combinations of the training and test data in the cross-validation calculation.

of the resistivity at two interfaces (FTO/hematite and hematite/solution) are the key issues, and also it is preferred to have the (110) facet on the electrode. Thus, we could find the important features from these analytical data without prior knowledge.

However, the obtained result was straightforward for the electrode preparation because it includes PEIS, a similar measurement as the photocurrent measurement (PEC). Since the measurement setup of PEIS was the same as the current–voltage measurement and the photocurrent was measured using both the measurements, it is reasonable that many of the features were selected from the PEIS data. Except for the descriptors in PEIS, the analyses could only give single information on the structural properties. Thus, we studied the correlation between the photocurrent and the analytical data except for the PEIS data in the next step.

Photocurrent prediction of photoanodes using analytical data without the photoelectrochemical impedance data

For a valid prediction without the PEIS data, we increased the number of the samples to 75. Seventy-five hematite photoanode samples were prepared under the same experimental conditions, and the corresponding current–potential curves are shown in Fig. 5. As a metric of the PEC performance, the photocurrent density at 1.23 V *vs.* RHE was used again. The photocurrents were dispersed in a range from 1.2 to 350% from the average photocurrent, irrelevant to the sintering temperatures. All the samples were analyzed by UV/vis spectroscopy, Raman spectroscopy, and XRD to represent the difference between the photoanodes. The spectral data were preprocessed by background and noise elimination, the same as in the previous section, and the intensities and positions of the peaks were selected as feature values. The features of the peaks and the peak regions were labeled as shown in Fig. 6. The selected features of the peaks and the regions were a little different from those used in the previous section due to the S/N ratio of the analytical data. The peak intensities of the XRD patterns and the Raman spectra were normalized by one of the peaks to handle the intensity difference among the samples due to the day-to-day error conditions of the analytical equipment. In addition, the first derivatives of the spectra were considered for the UV/vis and Raman spectra to extract the shoulder peaks at around 550 nm for UV/vis and 650 cm^{-1} for Raman. Thirty-seven features were extracted as feature values. Similarly, as in the previous section, the descriptors with the normalized standard deviation <1% were removed. The number of features was reduced from 37 to 20.

Next, the determinant factors for the photocurrent were examined. Various methods were tested and the result of the stepwise regression was used to identify the important descriptors. The calculation selected 12 feature values. The extracted descriptors are shown in Table 2 with the coefficients in the prediction model function. Fig. 7 shows a scatter plot for the predicted and experimental values. As shown in Fig. 7, the prediction accuracy was sufficiently high according to the determination coefficient (0.894). We could successfully determine 12 descriptors to determine the photocurrent, even without the PEIS data.

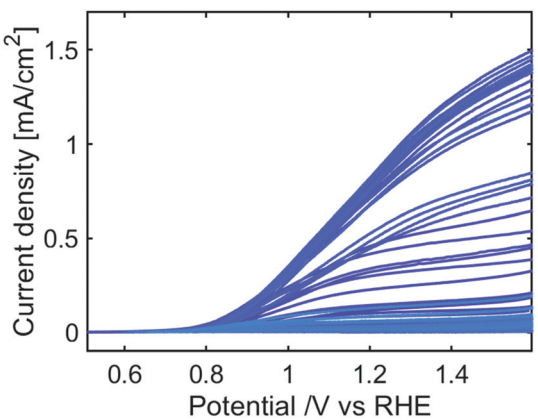

Fig. 5 The current–voltage curves for 75 samples of hematite photoanodes. The potential is given *versus* RHE.

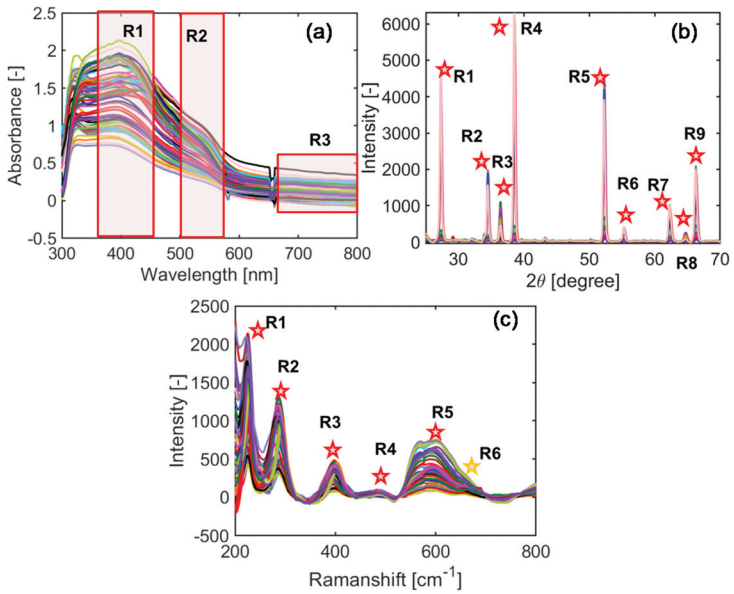

Fig. 6 The analytical data for the hematite electrodes; (a) UV/vis spectra, (b) XRD patterns, and (c) Raman spectra. Each separated region is shown in red rectangles or stars with labels.

Table 2 Selected descriptors obtained by the stepwise regression and the coefficients in the prediction model for the data without the PEIS data

Descriptors	Coefficient
UV_Vis_pks_abs	−0.939
UV_Vis_average_abs	0.400
Raman_pk2_int	0.587
Raman_pk3_int	−0.394
Raman_pk5_int	−0.876
Raman_pk2_loc	0.431
XRD_pk3_int	0.865
XRD_pk4_int	−0.724
XRD_pk6_int	−0.516
XRD_pk7_int	0.363
XRD_pk9_int	−0.637

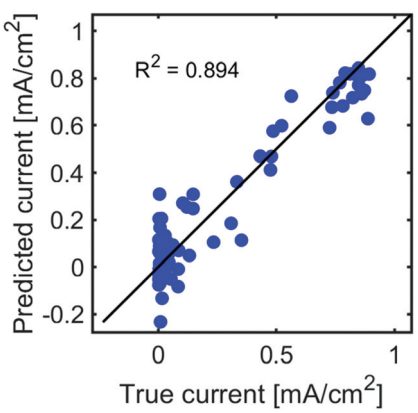

Fig. 7 The scatter plots of the true current and the predicted current.

For the improvement of the prediction model function, nonlinear regression methods were utilized to predict the performance with the selected 12 descriptors. The models were tested with 5-fold cross-validation, and GPR provided the best results. Fig. 8 shows a scatter plot for the predicted photocurrent *versus* the target values. As described in the previous section, five different combinations of datasets were indicated by the 1st, 2nd, 3rd, 4th, and 5th, and the average R^2 values are shown in Fig. 8. As shown in Fig. 8, the determination coefficients were 0.856 and 0.855 for the training and test data. The same level of accuracy for the training and test assures the validity of the calculation. Therefore, the model constructed by ML was accurate enough to predict the target values, even without the PEIS analytical data.

In the selected features, two UV/vis, four Raman and five XRD features were selected, which included reasonable and unexpected descriptors. In the UV/vis spectra, the intensities in two regions (R1 and R3) were selected; the former represents the light absorption near the band edge, which is understandable because it is relevant to the light absorption. However, the light absorption in the near-infrared region seems irrelevant intuitively, but this absorbance reflects the light scattering ability of the particulate electrodes due to the roughness of the surfaces. The surface roughness possibly enhanced the light absorption of particles. The (110) peak in the XRD pattern was selected, the same as in the previous section, and reasonably understood. However, the other four selected peaks in the XRD patterns correspond to the peaks of FTO, which have been ignored mostly, but they were selected possibly because the peak intensities of FTO are relevant to the sample thickness (the FTO peaks are reduced with an increase in the thickness of a hematite layer). Three Raman

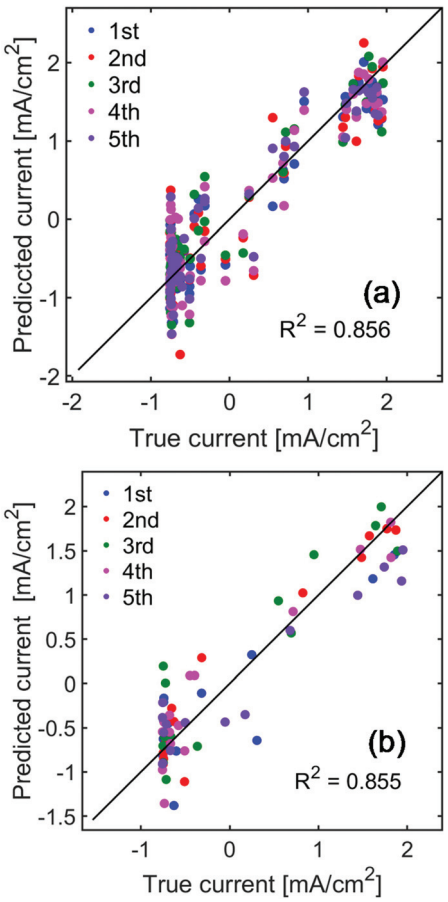

Fig. 8 The scatter plots of the true current densities *versus* the predicted values for hematite electrodes for the (a) training and (b) test data. 1st, 2nd, 3rd, 4th, and 5th correspond to the datasets with different combinations of the training and test data in the cross-validation calculation.

peaks (R2, R3 and R5) corresponding to E_g were extracted as important descriptors.[24,26] These structural orders are relevant to the photoelectrochemical performance. The result shows that the ML prediction could extract dominant descriptors without any prior knowledge, even without the photoelectrochemical data, and the information could be related to the necessary structural information of materials.

Conclusion

We applied ML techniques to predict the photocurrent densities of hematite photoelectrodes from various analytical data, such as UV/vis spectroscopy, Raman spectroscopy, XRD, and PEIS, and successfully found important factors for the improvement of photocurrents. The descriptors were selected directly from the analytical data, such as the peak intensities and positions, and all the features were used to determine the dominant factors in the analytical data by the stepwise regression, and finally the prediction function was refined using a GPR model. The prediction was provided at first using all the analytical data, and could reach a determination coefficient of ~0.9. Furthermore, the prediction was possible even without the PEIS data, which is a similar type of measurement method to the photocurrent measurement. The latter prediction provided more information on the structure and physical properties. From the analyses, the ML prediction could offer various important factors without any prior knowledge, even for the situation where a small number of data is available in the practical development of devices/materials. This methodology would work effectively for any material/device development with a target value since the descriptors were identified without any prior information about the sample system from the analytical data for dozens of samples. The significant descriptors about hematite were clarified in this study and this will lead the photocurrent enhancement by inductive searching and tuning the experimental parameters affecting the descriptors, and it is now under progress.

Data and software availability

All the source analytical data and the codes for the descriptor selection and performance prediction are openly available *via* GitHub; https://github.com/Katayama-ChuoU/PEC_performance_prediction.

The program was prepared using Matlab R2021a.

Author contributions

YN performed the experiments and analyses, and YN and KK discussed the results. YN and KK wrote the manuscript, and both of them reviewed it.

Conflicts of interest

The authors have no competing interests or other interests that might be perceived to influence the results and/or discussion reported in this article.

Acknowledgements

The research was financially supported by JST PRESTO (#JPMJPR1675), KIOXIA corporation, the Iron and Steel Institute of Japan, the Institute of Science and Engineering, Chuo University and JSPS KAKENHI (#JP21J21092).

References

1 A. C. Mater and M. L. Coote, *J. Chem. Inf. Model.*, 2019, **59**, 2545–2559.
2 D. Fooshee, A. Mood, E. Gutman, M. Tavakoli, G. Urban, F. Liu, N. Huynh, D. V. Vranken and P. Baldi, *Mol. Syst. Des. Eng.*, 2018, **3**, 442–452.

3 D. C. Elton, Z. Boukouvalas, M. D. Fuge and P. W. Chung, *Mol. Syst. Des. Eng.*, 2019, **4**, 828–849.
4 J. Schmidt, M. R. G. Marques, S. Botti and M. A. L. Marques, *npj Comput. Mater.*, 2019, **5**, 1–36.
5 K. Min and E. Cho, *J. Phys. Chem. C*, 2020, **124**, 14759–14767.
6 Z. Wu, B. Ramsundar, E. N. Feinberg, J. Gomes, C. Geniesse, A. S. Pappu, K. Leswing and V. Pande, *Chem. Sci.*, 2018, **9**, 513–530.
7 C. J. Court, B. Yildirim, A. Jain and J. M. Cole, *J. Chem. Inf. Model.*, 2020, **60**, 4518–4535.
8 L. Banko, Y. Lysogorskiy, D. Grochla, D. Naujoks, R. Drautz and A. Ludwig, *Commun. Mater.*, 2020, **1**, 1–10.
9 S. Steiner, J. Wolf, S. Glatzel, A. Andreou, J. M. Granda, G. Keenan, T. Hinkley, G. Aragon-Camarasa, P. J. Kitson, D. Angelone and L. Cronin, *Science*, 2019, **363**, eaav2211.
10 H. Kaneko, *Heliyon*, 2021, **7**, e07356.
11 Y. Wang, A. Iyer, W. Chen and J. M. Rondinelli, *Appl. Phys. Rev.*, 2020, **7**, 041403.
12 M. Harada, H. Takeda, S. Suzuki, K. Nakano, N. Tanibata, M. Nakayama, M. Karasuyama and I. Takeuchi, *J. Mater. Chem. A*, 2020, **8**, 15103–15109.
13 T. Ebi, A. Sen, R. N. Dhital, Y. M. A. Yamada and H. Kaneko, *ACS Omega*, 2021, **6**(41), 27578–27586.
14 R. Tamura, Y. Takei, S. Imai, M. Nakahara, S. Shibata, T. Nakanishi and M. Demura, *Sci. Technol. Adv. Mater. Methods*, 2021, **1**, 152–161.
15 A. Kurotani, T. Kakiuchi and J. Kikuchi, *ACS Omega*, 2021, **6**, 14278–14287.
16 Q. Wang and K. Domen, *Chem. Rev.*, 2020, **120**, 919–985.
17 B. Klahr, S. Gimenez, F. Fabregat-Santiago, J. Bisquert and T. W. Hamann, *J. Am. Chem. Soc.*, 2012, **134**, 16693–16700.
18 M. G. Ahmed, I. E. Kretschmer, T. A. Kandiel, A. Y. Ahmed, F. A. Rashwan and D. W. Bahnemann, *ACS Appl. Mater. Interfaces*, 2015, **7**, 24053–24062.
19 Z. Chen, N. Andrejevic, N. C. Drucker, T. Nguyen, R. P. Xian, T. Smidt, Y. Wang, R. Ernstorfer, D. A. Tennant, M. Chan and M. Li, *Chem. Phys. Rev.*, 2021, **2**, 031301.
20 Y. Liu, N. Marcella, J. Timoshenko, A. Halder, B. Yang, L. Kolipaka, M. J. Pellin, S. Seifert, S. Vajda, P. Liu and A. I. Frenkel, *J. Chem. Phys.*, 2019, **151**, 164201.
21 C. L. M. Morais, P. Giamougiannis, R. Grabowska, N. J. Wood, P. L. Martin-Hirsch and F. L. Martin, *Analyst*, 2020, **145**, 5915–5924.
22 C. L. M. Morais, P. L. Martin-Hirsch and F. L. Martin, *Analyst*, 2019, **144**, 2312–2319.
23 Y. Nagai and K. Katayama, *Analyst*, 2021, **146**, 5045–5054.
24 J.-W. Jang, C. Du, Y. Ye, Y. Lin, X. Yao, J. Thorne, E. Liu, G. McMahon, J. Zhu, A. Javey, J. Guo and D. Wang, *Nat. Commun.*, 2015, **6**, 7447.
25 T. Tokubuchi, R. I. Arbi, P. Zhenhua, K. Katayama, A. Turak and W. Y. Sohn, *J. Photochem. Photobiol., A*, 2021, **410**, 113179.
26 C. P. Marshall, W. J. B. Dufresne and C. J. Rufledt, *J. Raman Spectrosc.*, 2020, **51**, 1522–1529.