

Data-Driven Identification of the Reaction Network in Oxidative Coupling of the Methane Reaction via Experimental Data

Itsuki Miyazato,* Shun Nishimura, Lauren Takahashi, Junya Ohyama, and Keisuke Takahashi



Cite This: *J. Phys. Chem. Lett.* 2020, 11, 787–795



Read Online

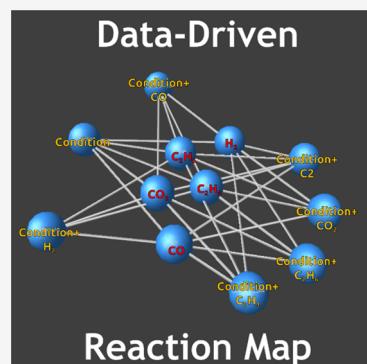
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Identifying details of chemical reactions is a challenging matter for both experiments and computations. Here, the reaction pathway in oxidative coupling of methane (OCM) is investigated using a series of experimental data and data science techniques in which data are analyzed using a variety of visualization techniques. Data visualization, pairwise correlation, and machine learning unveil the relationships between experimental conditions and the selectivities of CO, CO₂, C₂H₄, C₂H₆, and H₂ in the OCM reaction. More importantly, the reaction network for the OCM reaction is constructed on the basis of the scores provided by machine learning and experimental data. In particular, the proposed reaction map not only contains the chemical compound but also contains experimental conditions. Thus, data-driven identification of chemical reactions can be achieved in principle via a series of experimental data, leading to more efficient experimental design and catalyst development.



Understanding the details of chemical reactions has been a great mystery in chemistry as chemical reactions are involved with the reactions of numerous intermediate compounds whose details are invisible.^{1,2} As a result, the complexities of reaction networks are limited when trying to understand the reaction mechanism and apply it to catalyst design. Intermediate compounds that appear within a reaction are gradually becoming clearer with the recent development of operando techniques such as X-ray absorption spectroscopy, but the details of the whole reaction network via experiments remain unclear.^{3–5} On the other hand, automated reaction pathway searching in combination with first-principles calculations has become possible, with every detail of the reaction network being computationally estimated.^{6–8} Additionally, various methods have been employed in an attempt to determine a reaction path such as kinetic modeling.^{9–13} However, it is particularly challenging to control the selectivity of various reactant products as the reaction network is not only complicated but also greatly affected by experimental conditions, where any change in conditions can greatly influence the paths taken during a reaction.^{14–16} Here, data science is implemented to reveal the reaction mechanism of oxidative coupling of methane (OCM) based on trends and patterns within OCM experimental data.

The OCM reaction aims to produce C₂H₄ and H₂O from the reaction between CH₄ and O₂.^{14,17,18} Although the OCM reaction has been extensively investigated since 1982, OCM has not been applied to commercial applications due to its low ethylene yields.^{14,19} The difficulty of OCM research originates in its complicated reaction mechanism, which consists of complicated sequences of surface and gas-phase reactions.²⁰

Given these factors, catalyst and process designs are used in an attempt to obtain a higher ethylene yield in the OCM reaction. Process design focuses on reducing detrimental effects caused within a reactor such as decreasing the oxygen concentration.^{21–24} Catalyst candidates for the OCM reaction have also been designed and evaluated where Mn–Na₂WO₄/SiO₂ can be considered to be the standard catalyst used for the OCM reaction due to its high performance.^{25,26} The data-driven approach has also demonstrated previously that complex experimental conditions with Mn–Na₂WO₄/SiO₂ can be well-mapped for predicting the C₂ yield in the OCM reaction.¹⁶ Here, the relationships between the experimental conditions and the selectivities of CO, CO₂, C₂H₄, C₂H₆, and H₂ in the OCM reaction with Mn–Na₂WO₄/SiO₂ are investigated using data science to reveal how each selectivity behaves within the OCM reaction.

One hundred ninety-six OCM experimental data containing experimental conditions and the selectivities of CO, CO₂, C₂H₄, C₂H₆, and H₂ are investigated. The data were previously constructed by running a series of 196 OCM experiments with different experimental conditions over Mn–Na₂WO₄/SiO₂ where the selectivities of CO, CO₂, C₂H₄, C₂H₆, and H₂ and the conversions of CH₄ and O₂ has been previously reported, which proved to be successful in a machine learning application.¹⁶ Thus, the quantity and quality of the data are

Received: December 12, 2019

Accepted: January 15, 2020

Published: January 15, 2020

qualified against the machine learning study. The 196 data consisting of four experimental conditions, one yield, five selectivities, and two conversions are listed in **Table 1**, while the details of the 196 data are available in the *Supporting Information*.

Table 1. Details and Abbreviations of 196 Experimental OCM Data^a

experimental condition	abbreviation
weight of the catalyst (mg)	weight
introduced amount of CH ₄ and O ₂ gases (concentration %)	CH ₄ + O ₂
total flow amount of the introduced gas	total flow
ratio of the introduced CH ₄ and O ₂ gases	CH ₄ /O ₂
temperature of the reaction tube (K)	temperature
reaction output	abbreviation
O ₂ conversion (%)	O ₂ _conv
CH ₄ conversion (%)	CH ₄ _conv
H ₂ selectivity (%)	H ₂ _selec
CO selectivity (%)	CO_selec
CO ₂ selectivity (%)	CO ₂ _selec
C ₂ H ₆ selectivity (%)	C ₂ H ₆ _selec
C ₂ H ₄ selectivity (%)	C ₂ H ₄ _selec
C ₂ yield (%)	C ₂ _y

^aAll data variables are available in the *Supporting Information*.

Data science techniques are implemented to reveal the relationship between experimental conditions and selectivities within OCM reactions. The data visualization technique (violin plot) is used to understand the data structure and the behavior of the OCM reaction.²⁷ Pairwise correlation with the Pearson coefficient is calculated to evaluate the 1:1 correlation weight between experimental conditions and reaction output. Supervised machine learning is applied to understand the relations between experimental conditions and each selectivity within the OCM reaction through supervised machine learning fitting. Scikit-learn (version 0.17.0-4)²⁸ is used for machine learning. In particular, these six machine learning models are chosen and evaluated: linear regression (LR), support vector machine (SVM),²⁹ least absolute shrinkage and selection operator (LASSO),³⁰ random forest regressor (RFR),³¹ extra tree regressor (ETR),³² and support vector regression (SVR).²⁹ The hyperparameters of each model are tuned by changing parameters manually [i.e., C = 0, 1, 10, and 100, and γ = 1, 0.1, 0.01, and 0.001 in the case of SVR(rbf)] and taken where the cross validation score is the highest. The tuned hyperparameters are listed in **Table 2**. Cross validation is used to evaluate the algorithms where the data are randomly divided into 20% test and 80% trained data. Scores of 10 random test and trained data are taken and evaluated by obtaining scores of mean square error, and averages are taken for evaluation. Additionally, the median and standard deviation within 10 random CV scores are obtained to evaluate the accuracy of each training in a certain model (see the result shown in **Tables 3** and **4**). This CV process has been ensured to be unbiased; consequently, it has been accepted in our previous works.^{15,33–35} It is noted that H₂O formation is observed after the OCM reaction experiment; however the determination of the H₂O production rate is difficult at present. Therefore, gas quantities calculated as the conversion, yield, and selectivity values are used as the variables for machine learning in this study.

Table 2. Implemented Machine Learning Models and the Corresponding Hyperparameters

model	hyperparameters
LR	fit_intercept = true, normalize = false, copy_X = true, n_jobs = none
SVR (linear)	degree = 3, γ = "auto_deprecated", C = 1.0
LASSO	α = 0.1, fit_intercept = true, normalize = false, precompute = false, copy_X = true, max_iter = 1000, warm_start = false positive = false, selection = "cyclic"
ETR	random_state = 3, n_estimators = 100, min_samples_split = 2, min_samples_leaf = 1, min_weight_fraction_leaf = 0.0, max_features = "auto"
RFR	random_state = 3, n_estimators = 100, min_samples_split = 2, min_samples_leaf = 1, min_weight_fraction_leaf = 0.0, max_features = "auto"
SVR (rbf)	kernel = "rbf", degree = 3, γ = 0.001, C = 10.0

Table 3. Cross Validation Scores (obtained as the mean square error) of Six Machine Learning Models for Predicting Each Selectivity with Different Models^a

target	model	average	median	standard deviation
H_2 selectivity	LR	0.39	0.47	0.17
	LASSO	0.40	0.47	0.17
	SVR (linear)	0.36	0.39	0.07
	RFR	0.71	0.67	0.14
	ETR	0.62	0.53	0.17
	SVR (rbf)	0.56	0.52	0.20
CO selectivity	LR	0.70	0.72	0.08
	LASSO	0.70	0.72	0.08
	SVR (linear)	0.65	0.67	0.12
	RFR	0.80	0.83	0.12
	ETR	0.72	0.79	0.24
	SVR (rbf)	0.64	0.65	0.15
C_2H_6 selectivity	LR	0.07	0.09	0.11
	LASSO	0.08	0.10	0.11
	SVR (linear)	-0.19	-0.18	0.22
	RFR	0.83	0.91	0.12
	ETR	0.81	0.85	0.11
	SVR (rbf)	0.32	0.39	0.23
CO_2 selectivity	LR	0.63	0.65	0.11
	LASSO	0.63	0.65	0.11
	SVR (linear)	0.60	0.62	0.11
	RFR	0.77	0.76	0.06
	ETR	0.74	0.74	0.09
	SVR (rbf)	0.69	0.77	0.13
C_2H_4 selectivity	LR	0.22	0.24	0.18
	LASSO	0.22	0.24	0.18
	SVR (linear)	0.16	0.19	0.15
	RFR	0.77	0.77	0.10
	ETR	0.65	0.72	0.25
	SVR (rbf)	0.47	0.47	0.12

^aStandard deviation in 10 random states of cross validation.

Violin plot visualization is performed to explore how each selectivity behaves against the change in temperature where the violin plot represents that distribution of data. The selectivities of CO, CO₂, C₂H₄, C₂H₆, and H₂ against temperature are shown in **Figure 1**, where the data with a CH₄/O₂ ratio of 3 are used. With the implementation of the violin plot, similarities within the selectivities become visible. In particular, the behaviors of C₂H₄ selectivity and CO₂ selectivity are similar to each other, where C₂H₄ selectivity and CO₂ selectivity have

Table 4. Cross Validation Scores (obtained as mean square error) for Predicting C₂H₄ Selectivity Using Five Descriptor Variables + X, Where X Is H₂ Selectivity, CO Selectivity, or CO₂ Selectivity^a

model	X	average	median	standard deviation
SVM (rbf)	none	0.47	0.47	0.12
	H ₂ selectivity	0.51	0.50	0.12
	CO selectivity	0.59	0.58	0.11
	C ₂ H ₆ selectivity	0.57	0.57	0.14
	CO ₂ selectivity	0.47	0.48	0.13
RFR	none	0.77	0.77	0.10
	H ₂ selectivity	0.80	0.79	0.06
	CO selectivity	0.76	0.74	0.07
	C ₂ H ₆ selectivity	0.80	0.79	0.07
	CO ₂ selectivity	0.83	0.84	0.08

^aStandard deviation in 10 random states of cross validation.

a peak temperature of 1048 K and are then found to decrease as shown in panels c and e of Figure 1. Although C₂H₆ selectivity is slightly increased around 1048 K, the change in the peak is smaller than that in the case of C₂H₄ selectivity. This tendency also corresponds to CH₄ conversion and C₂ yield as shown in panels g and i of Figure 1. Thus, one can consider that the optimal temperature of Mn–Na₂WO₄/SiO₂ catalysts in OCM is considered to be 1048 K. On the other hand, CO selectivity is linear with temperature as shown in Figure 1b. This suggests that the production of CO is strongly coupled with temperature. Hence, the violin plot of each selectivity gives insight into the reaction mechanisms.

Pairwise correlation with the Pearson coefficient is evaluated as a heat map to find the correlation between experimental conditions and reaction output as shown in Figure 2. Overall, Figure 2 suggests that the OCM reaction is a temperature dominant reaction as there is a strong correlation between temperature and reaction outputs. In particular, CO selectivity and O₂ conversion are indicated to strongly correlate with temperature with correlation scores of 0.80 and 0.85, respectively. CH₄ conversion and H₂ selectivity are also found to strongly correlate with temperature with correlation scores of 0.69 and 0.66, respectively. This relation is also reflected in the corresponding violin plots of Figure 1, where CH₄ conversion increases from 973 to 1048 K and H₂ selectivity increases from 1098 to 1173 K. In contrast, there is a negative correlation between CH₄ + O₂ and CO₂ selectivity, suggesting that CO₂ selectivity can be decreased as CH₄ + O₂ selectivity increases. Thus, pairwise correlation provides insight into the overall correlation between experimental conditions and reaction outputs.

To visualize the complexities of experimental conditions against each selectivity, parallel coordinate visualization is implemented.³⁶ Given that Figure 2 demonstrates an inverse proportion correlation between CH₄ + O₂ and CO₂ selectivity, parallel coordinates are implemented to illustrate the relationship between CO₂ selectivity and the experimental conditions shown in Figure 3. The parallel coordinates show that large CH₄ + O₂ selectivity results in a low CO₂ selectivity, showing good agreement with the inverse proportion correlation shown in Figure 2. This suggests that CO₂ selectivity can be controlled via CH₄ + O₂ selectivity.

Parallel coordinates are used to reveal the patterns based on the degree of C₂ yield against experimental conditions shown in Figure 4. C₂ yield is classified into the following groups:

entire range of C₂ yields, C₂ yield of <5, and C₂ yield of ≥15 (Figure 4). At a C₂ yield of <5, one can clearly see that a low temperature is observed as shown in Figure 4b. On the other hand, a relatively high temperature is required for a high C₂ yield as seen in Figure 4c. More importantly, a high CH₄ conversion and a high O₂ conversion appear to be crucial factors for achieving a high C₂ yield. It is also interesting to note that when C₂H₆ selectivity is high, C₂H₄ selectivity tends to be low while a low C₂H₆ selectivity results in a high C₂H₄ selectivity. Thus, parallel coordinates unveiled the hidden patterns within the OCM reactions.

Supervised machine learning is implemented to reveal the relationship between each selectivity and experimental condition; the trend of multivariate data with simple visualization techniques is difficult to capture due to its complexity. The following selectivities are chosen as objective variables: CO selectivity, C₂H₆ selectivity, C₂H₄ selectivity, and H₂ selectivity. C₂ yield has also been previously chosen as the target variable for evaluating benchmark parameters for catalyst function.¹⁶ The following experimental conditions are set as descriptor variables: “weight”, “CH₄ + O₂”, “total flow”, “CH₄/O₂”, and “temperature”. Six supervised machine learning models are evaluated for predicting each selectivity from experimental conditions. The cross validation scores for each machine learning technique are listed in Table 3. Additionally, the importance of the descriptor variables for RFR is evaluated and shown in Figure 5; the text is colored to evaluate how each descriptor influences the machine learning process in the random forest algorithm.

Of all of the machine learning models chosen, RFR results in the highest cross validation scores as can be seen in Table 3. Thus, one can consider that the relationship between experimental conditions and the selectivity of CO, CO₂, C₂H₄, C₂H₆, and H₂ is a nonlinear matter. However, CO selectivity predictions appear to behave differently as relatively high cross validation scores are observed with the LR and SVM (linear) models. This can be explained by the strong linear relationships seen in Figures 1b and 2, in which the selectivity of CO and temperature are strongly linearly correlated. The importance of the descriptor variables for CO selectivity is also evaluated as shown in Figure 5a, where the effect of temperature has a major impact. Thus, linear machine learning models can be applicable for predicting CO selectivity alongside nonlinear machine learning models.

The importance of descriptors as seen in Figure 5 shows that temperature has a large impact on CO and H₂ selectivity while temperature appears to have less impact on CO₂, C₂H₄, and C₂H₆ selectivities.

Meanwhile, Figure 5b shows that CH₄ + O₂ selectivity, rather than temperature, has a large impact on CO₂ selectivity. This corresponds to the trends seen in Figure 2, where CO₂ selectivity is seen to have an inverse relationship with CH₄ + O₂ selectivity where CO₂ selectivity decreases as CH₄ + O₂ selectivity increases. This also shows good agreement with the pairwise correlation map presented in Figure 2, where CO₂ selectivity is also seen to strongly correlate with CH₄ + O₂ selectivity. On the other hand, panels d and e of Figure 5 show that selectivities of C₂H₄ and C₂H₆ are the result of complex experimental conditions. Thus, one can see that each of the experimental conditions contributes a different degree to determining each selectivity in the OCM reaction.

RFR is found to have a good ability to predict each selectivity from the experimental conditions listed in Table 1.

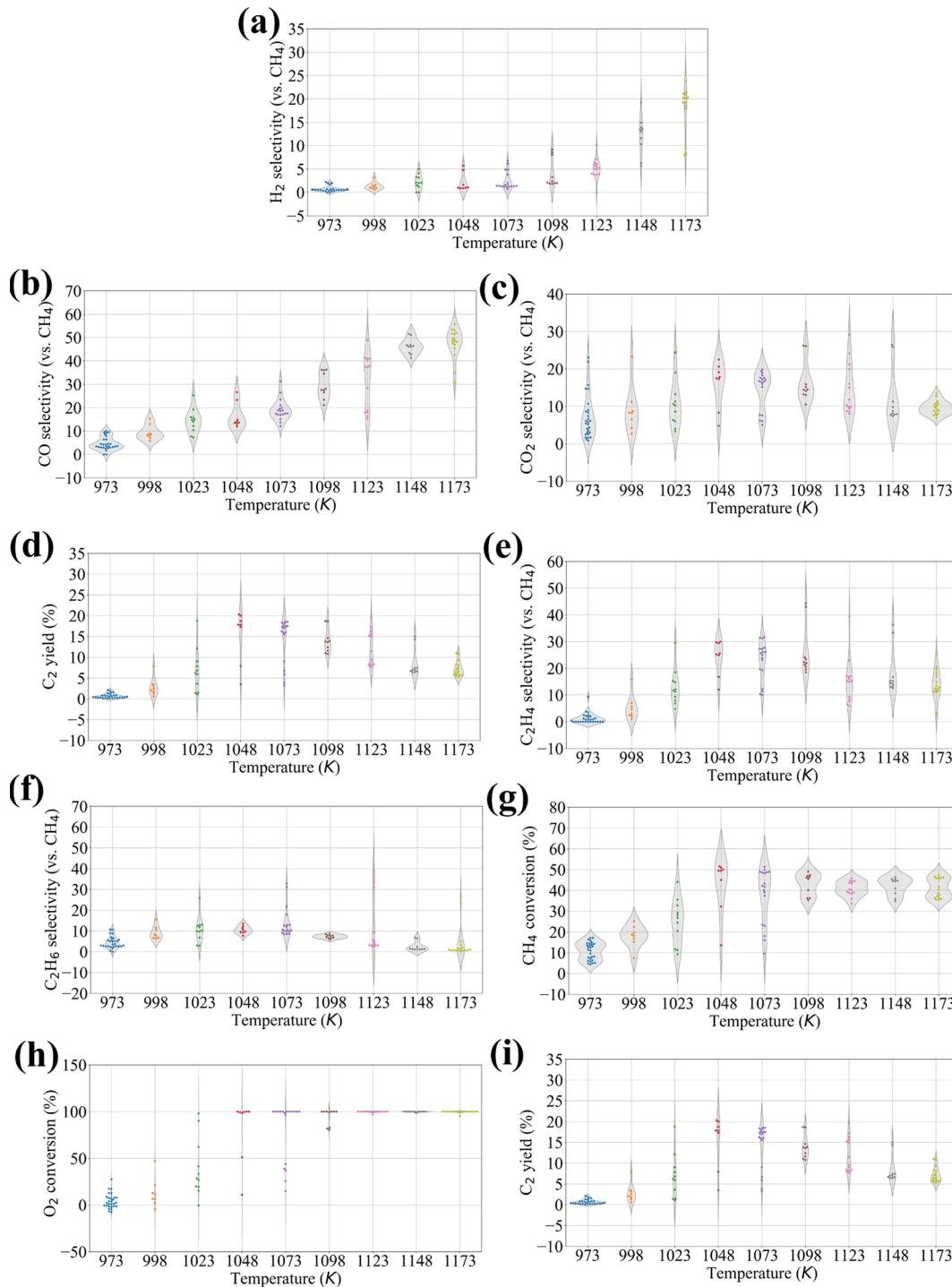


Figure 1. Violin plot of the selectivity of (a) H₂, (b) CO, (c) CO₂, (d) C₂, (e) C₂H₄, and (f) C₂H₆ and the conversion of (g) CH₄ and (h) O₂ and (i) yield of C₂ vs temperature. Note that applied data are fixed at a CH₄/O₂ ratio of 3 to fix the effect of reaction source variety.

Here, interpolation filling using the trained RFR model is performed to predict CO, CO₂, C₂H₄, C₂H₆, and H₂ selectivities. The following experimental conditions are given to the trained RFR for each selectivity: weight = 50, CH₄ + O₂ selectivity = 90, total flow = 30, CH₄/O₂ ratio = 1–5, and temperature = 974–1172 K. Note that weight, CH₄ + O₂ selectivity, and total flow are fixed while temperature and CH₄/O₂ ratio are set as variables.

Interpolation filling is carried out, and the three-dimensional surface plots of CH₄/O₂ ratio and temperature versus

predicted CO, CO₂, C₂H₄, C₂H₆, and H₂ selectivities are shown in Figure 6, where the corresponding training data are plotted as gray dots. Figure 6 immediately illustrates that each selectivity behaves differently with temperature and CH₄/O₂ ratio. CO and H₂ selectivity as seen in panels a and c, respectively, of Figure 6 are found to increase as the temperature increases, while H₂ selectivity is low until the temperature reaches 1120 K. Alternatively, the selectivities of CO₂, C₂H₄, and C₂H₆ (panels b, d, and e, respectively, of Figure 6) are found to have mountain-like plots where high

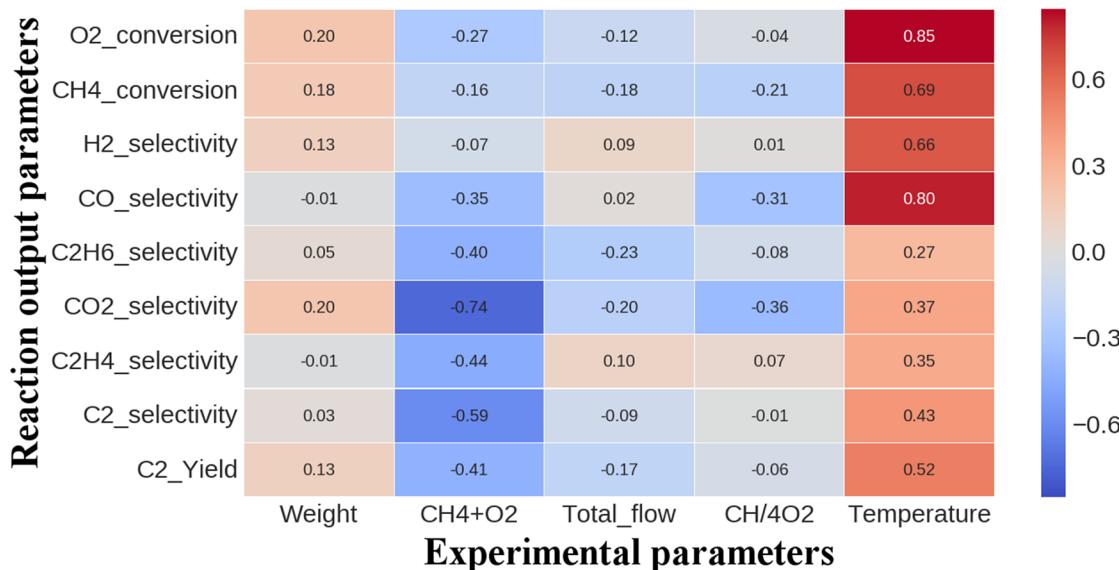


Figure 2. Heat map of pairwise correlation between reaction outputs and experimental parameters.

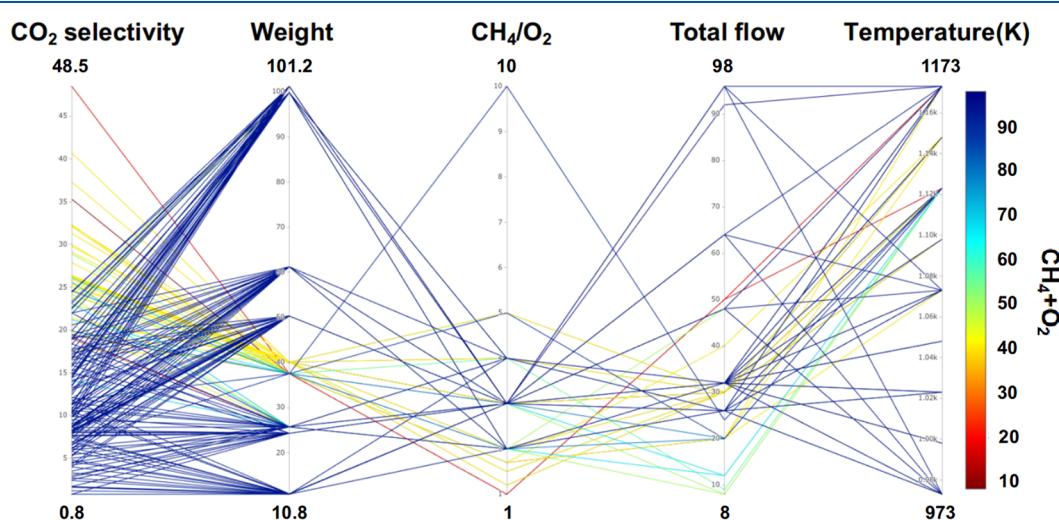


Figure 3. Parallel coordinates for CO₂ selectivity. CO₂ selectivity appears to concentrate where CH₄ + O₂ selectivity is higher. This tendency corresponds to the strong negative pairwise correlation between CH₄ + O₂ and CO₂ selectivity shown in Figure 2. Note that all variables are normalized.

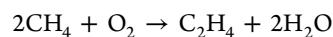
selectivities are observed around 1080 K. Furthermore, the effect of CH₄/O₂ ratio seems to play an important role in CO, CO₂, C₂H₄, and C₂H₆ selectivities; a small CH₄/O₂ ratio tends to induce high CO, CO₂, C₂H₄, and C₂H₆ selectivities. These results thus demonstrate that machine learning can be used to help predict and map out selectivities in relation to experimental conditions, making it a useful tool when attempting to design experiments.

The predicted values of H₂, C₂H₄, and CO₂ selectivity are found to be slightly lower than the C₂H₆ and CO selectivity as shown in Table 3, suggesting that other factors could be involved. Here, addition of selectivities as descriptor variables is proposed to improve the accuracy for predicting H₂, C₂H₄, and CO₂ selectivity. The prediction score is improved where the highest average cross validation score is returned when X is equal to CO₂ selectivity, H₂ selectivity, or C₂H₆ selectivity (0.83, 0.80, or 0.80, respectively). The results therefore suggest that C₂H₄ selectivity is impacted by CO₂, H₂, and C₂H₆,

allowing one to hypothetically assume that C₂H₄ is a secondary reaction in OCM.

Further studies of OCM catalysts suggest a multiphase reaction such as the cases reported for the formation of ethylene from ethane.^{17,37} This allows one to consider that the improvement of the cross validation score is a reflection of a multiphase reaction. With the inclusion of experimental output as learning features, it thus becomes possible to include the impact of a multiphase reaction when applying machine learning.

Data visualization, pairwise correlations, and machine learning help provide details regarding the relationships between each selectivity and experimental conditions in the OCM reaction. A reaction route and mechanisms of the OCM are proposed on the basis of analyses of these data. The OCM reaction is commonly described as follows:³⁸



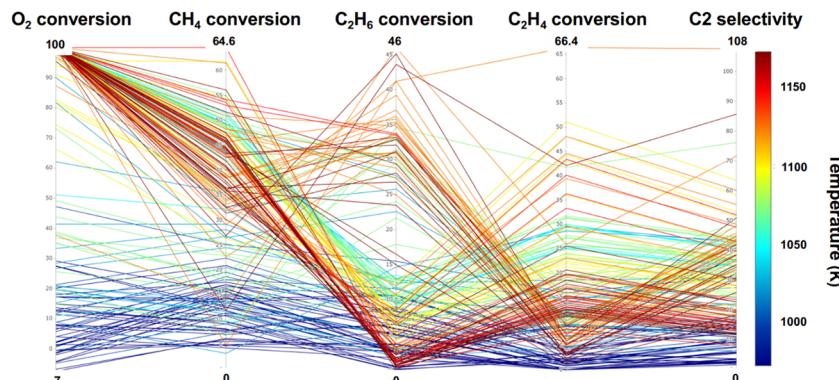
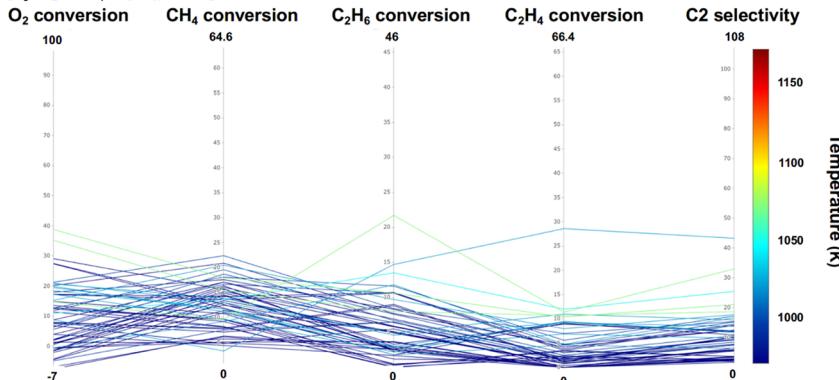
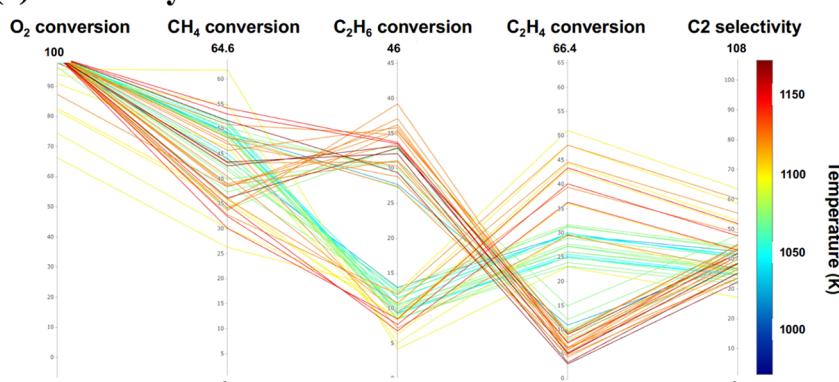
(a) All range of C₂ yield(b) C₂ yield < 5(c) 15 ≤ C₂ yield

Figure 4. Parallel coordinates for C₂ yield of (a) the entire range of C₂ yields, (b) C₂ yields of <5, and (c) C₂ yields of >15. Black, red, blue, and purple arrows indicate the “selected C₂ range”, the “index of increase against lower C₂ yield”, the “index of decrease against lower C₂ yield”, and the “index of separation against lower C₂ yield”, respectively. As a reference, the parallel coordinates with whole parameters are available in the Supporting Information. Note that all variables are normalized.

However, Table 1 shows that the OCM reaction is involved with CO, CO₂, C₂H₄, C₂H₆, and H₂ selectivities, making the OCM reaction more complex than it was initially thought to be. Tables 2 and 4 also reveal a direct relationship between experimental conditions and C₂H₆ selectivity, while C₂H₄ selectivity requires H₂ and C₂H₆ in addition to experimental conditions. This suggests that C₂H₆ is the initial reaction while C₂H₄ is possibly a secondary reaction and thus helps clarify the manner in which the OCM reaction occurs. This agrees well with the previously proposed reaction mechanism.^{13,39,40} This provides more opportunities to draw the OCM reaction network on the basis of scores in cross validation of machine learning. In particular, the reaction network shown in Figure 7 is constructed by predicting each CO, CO₂, C₂H₄, C₂H₆, and

H₂ selectivity using the following set of descriptor variable: condition + X (X = none, CO, CO₂, C₂H₄, C₂H₆, and H₂ selectivity), where condition is “weight”, “CH₄ + O₂”, “total flow”, “CH₄/O₂”, and “temperature”. In this way, one can consider that a high score in cross validation would reveal how each selectivity is correlated with other selectivities along experimental conditions.

Overall, the proposed reaction map determined by machine learning (Figure 7) supports the previously reported reaction mechanism in OCM.^{13,39,40} Figure 7 shows that C₂H₆ and CO selectivity is directly connected from conditions, suggesting that C₂H₆ and CO can be the initial reaction in OCM. On the other hand, H₂, CO₂, and C₂H₄ selectivities are not directly connected to experimental conditions, suggesting that CO₂

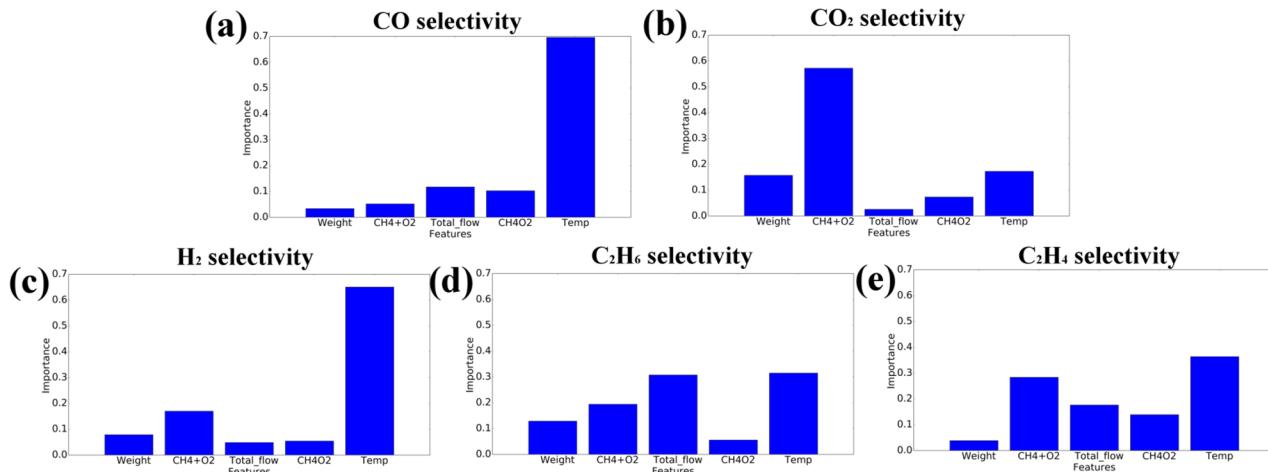


Figure 5. Importance of descriptors within the RFR model against (a) CO selectivity, (b) CO₂ selectivity, (c) H₂ selectivity, (d) C₂H₆ selectivity, and (e) C₂H₄ selectivity.

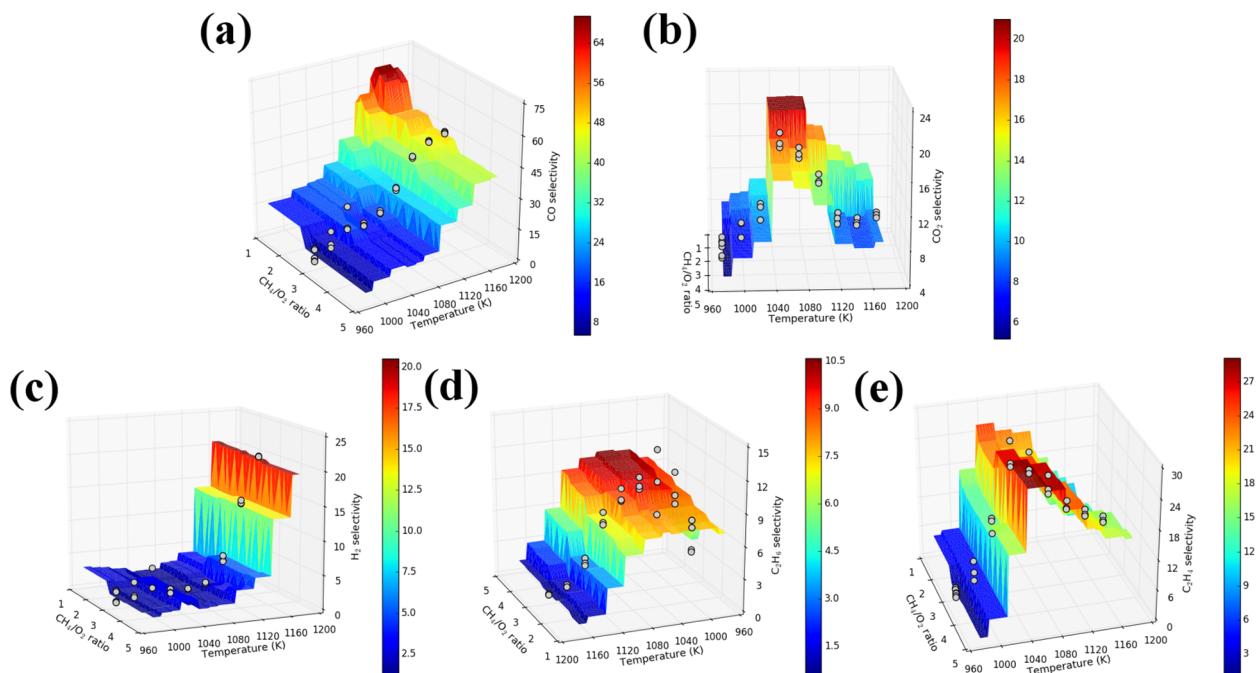


Figure 6. Three-dimensional surface plot of CH₄/O₂ and temperature vs predicted (a) CO, (b) CO₂, (c) H₂, (d) C₂H₆, and (e) C₂H₄ selectivity. Gray dots indicate the corresponding data points in trained data.

and C₂H₄ can be considered the second reaction. One can see that C₂H₄ selectivity is connected to the condition + X (C₂H₆ and H₂), suggesting that C₂H₄ is the product of dehydrogenation of C₂H₆. Similarly, CO selectivity is relatively strongly connected to byproduct gases (C₂H₄, C₂H₆, and H₂), suggesting that CO can be the final product of the OCM reaction. In the case of H₂ selectivity, it is connected to condition + CO₂, implying that H₂ can result from dehydrogenation of CH₄ to form CO₂. Thus, the reaction map created by machine learning could provide insight into the global reaction route of OCM reactions. Additionally, the proposed approach results in the reaction map containing both experimental conditions and corresponding compounds. Thus, one can consider that the reaction map constructed by data science can provide insight into the design of the multiphase reactor⁴¹ for dealing with complex reactions.

In conclusion, the reaction route of oxidative coupling of methane (OCM) is identified using a data-driven approach by combining data science with a series of experimental data. Data visualization, pairwise correlation, and machine learning reveal trends and patterns hidden within the experimental data. In particular, details of the reaction route in the OCM reaction are proposed where the relationships between experimental conditions and the selectivities of CO, CO₂, C₂H₄, C₂H₆, and H₂ are identified. Thus, this data-driven approach proposes that the details of the chemical reaction route can be revealed, in principle, by combining data science techniques and a series of experimental data. Data science techniques used in this manner act as bridges between experiment and computation by revealing new outcomes through data visualization and machine learning. More importantly, data science techniques also shed light on phenomena that researchers have instinctively understood to occur after analyzing series of

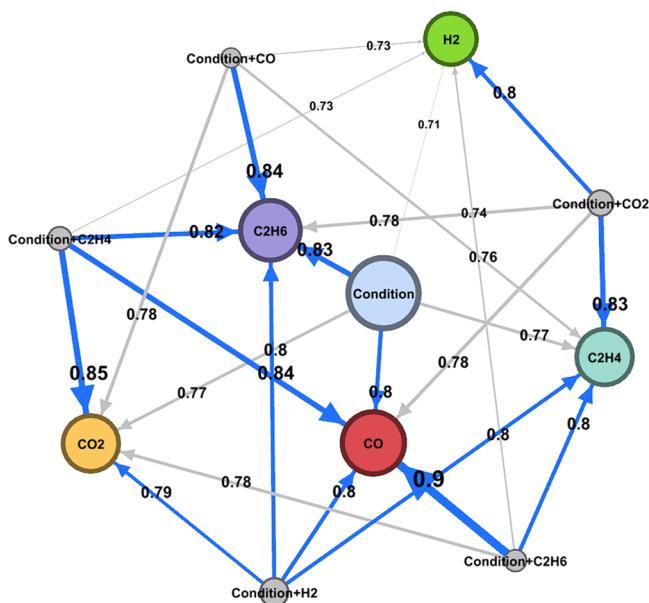


Figure 7. OCM reaction route map created by score in cross validation of RFR machine learning. CO, CO₂, C₂H₄, C₂H₆, and H₂ represent selectivity, while condition + X (X = none, CO, CO₂, C₂H₄, C₂H₆, and H₂ selectivity) represents the descriptor variable. Scores of >0.80 are colored blue. Edge labels denote cross validation scores, while edge thickness denotes weight where high score is set to heavier weight. Note that the data are available in the Supporting Information.

experimental results but have conventionally been excluded and not mentioned in scientific publications.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcllett.9b03678>.

Figures S-1 and S-2, Tables S-1 and S-2, and one additional reference ([PDF](#))

■ AUTHOR INFORMATION

Corresponding Author

Itsuki Miyazato — Hokkaido University, Sapporo, Japan, and National Institute for Materials Science (NIMS), Tsukuba, Japan; orcid.org/0000-0002-1533-9790; Email: miyazato@sci.hokudai.ac.jp

Other Authors

Shun Nishimura — Japan Advanced Institute of Science and Technology, Nomi, Japan; orcid.org/0000-0003-3084-1444

Lauren Takahashi — Hokkaido University, Sapporo, Japan, and National Institute for Materials Science (NIMS), Tsukuba, Japan; orcid.org/0000-0001-9922-8889

Junya Ohyama — Kumamoto University, Kumamoto, Japan, and Kyoto University, Kyoto, Japan; orcid.org/0000-0002-7438-5236

Keisuke Takahashi — Hokkaido University, Sapporo, Japan, and National Institute for Materials Science

(NIMS), Tsukuba, Japan; orcid.org/0000-0002-9328-1694

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jpcllett.9b03678>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This study was partly funded by the Materials Research by Information Integration Initiative (MI²I) project of the Support Program for Starting Up Innovation Hub of the Japan Science and Technology Agency (JST), JST CREST Grant JPMJCR17P2, and JSPS KAKENHI Grant-in-Aid for Young Scientists (B) Grant JP17K14803.

■ REFERENCES

- Ulissi, Z. W.; Medford, A. J.; Bligaard, T.; Nørskov, J. K. To Address Surface Reaction Network Complexity Using Scaling Relations Machine Learning and DFT Calculations. *Nat. Commun.* **2017**, *8*, 14621.
- Jaworski, W.; Szymkuć, S.; Mikulak-Klucznik, B.; Piecuch, K.; Klucznik, T.; Kaźmierowski, M.; Rydzewski, J.; Gambin, A.; Grzybowski, B. A. Automatic Mapping of Atoms Across Both Simple and Complex Chemical Reactions. *Nat. Commun.* **2019**, *10*, 1434.
- Yuan, N.; Pascanu, V.; Huang, Z.; Valiente, A.; Heidenreich, N.; Leubner, S.; Inge, A. K.; Gaar, J.; Stock, N.; Persson, I.; Martin-Matute, B.; Zou, X. Probing the Evolution of Palladium Species in Pd@ MOF Catalysts during the Heck Coupling Reaction: An Operando X-ray Absorption Spectroscopy Study. *J. Am. Chem. Soc.* **2018**, *140*, 8206–8217.
- Su, X.; Wang, Y.; Zhou, J.; Gu, S.; Li, J.; Zhang, S. Operando Spectroscopic Identification of Active Sites in NiFe Prussian Blue Analogues as Electrocatalysts: Activation of Oxygen Atoms for Oxygen Evolution Reaction. *J. Am. Chem. Soc.* **2018**, *140*, 11286–11292.
- Li, X.; Yang, X.; Zhang, J.; Huang, Y.; Liu, B. In Situ/Operando Techniques for Characterization of Single-Atom Catalysts. *ACS Catal.* **2019**, *9*, 2521–2531.
- Yang, M.; Zou, J.; Wang, G.; Li, S. Automatic Reaction Pathway Search via Combined Molecular Dynamics and Coordinate Driving Method. *J. Phys. Chem. A* **2017**, *121*, 1351–1361.
- Martínez, T. J. Ab initio Reactive Computer Aided Molecular Design. *Acc. Chem. Res.* **2017**, *50*, 652–656.
- Grambow, C. A.; Jamal, A.; Li, Y.-P.; Green, W. H.; Zador, J.; Suleimanov, Y. V. Unimolecular Reaction Pathways of a γ -Ketohydroperoxide from Combined Application of Automated Reaction Discovery Methods. *J. Am. Chem. Soc.* **2018**, *140*, 1035–1048.
- Arndt, S.; Otremba, T.; Simon, U.; Yildiz, M.; Schubert, H.; Schomäcker, R. Mn–Na₂WO₄/SiO₂ as Catalyst for the Oxidative Coupling of Methane. What is Really Known? *Appl. Catal., A* **2012**, *425–426*, 53–61.
- Sun, J.; Thybaut, J. W.; Marin, G. B. Microkinetics of Methane Oxidative Coupling. *Catal. Today* **2008**, *137*, 90–102.
- Bossche, M. V. d.; Gronbeck, H. Methane Oxidation over PdO Revealed by First-principles Kinetic Modeling. *J. Am. Chem. Soc.* **2015**, *137*, 12035–12044.
- Fleischer, V.; Simon, U.; Parishan, S.; Colmenares, M. G.; Görke, O.; Gurlo, A.; Riedel, W.; Thum, L.; Schmidt, J.; Risso, T.; Dinse, K.-P.; Schomäcker, R. Investigation of the Role of the Na₂WO₄/Mn/SiO₂ Catalyst Composition in the Oxidative Coupling of Methane by Chemical Looping Experiments. *J. Catal.* **2018**, *360*, 102–117.
- Karakaya, C.; Zhu, H.; Loebick, C.; Weissman, J. G.; Kee, R. J. A Detailed Reaction Mechanism for Oxidative Coupling of Methane

- over Mn/Na₂WO₄/SiO₂ Catalyst for Non-isothermal Conditions. *Catal. Today* **2018**, *312*, 10–22.
- (14) Zavyalova, U.; Holena, M.; Schlögl, R.; Baerns, M. Statistical Analysis of Past Catalytic Data on Oxidative Methane Coupling for New Insights into the Composition of High-performance Catalysts. *ChemCatChem* **2011**, *3*, 1935–1947.
- (15) Takahashi, K.; Miyazato, I.; Nishimura, S.; Ohyama, J. Unveiling Hidden Catalysts for the Oxidative Coupling of Methane based on Combining Machine Learning with Literature Data. *ChemCatChem* **2018**, *10*, 3223–3228.
- (16) Ohyama, J.; Nishimura, S.; Takahashi, K. Data Driven Determination of Reaction Conditions in Oxidative Coupling of Methane via Machine Learning. *ChemCatChem* **2019**, *11*, 4307–4313.
- (17) Keller, G.; Bhasin, M. Synthesis of Ethylene via Oxidative Coupling of Methane: I. Determination of Active Catalysts. *J. Catal.* **1982**, *73*, 9–19.
- (18) Farrell, B. L.; Igenegbai, V. O.; Linic, S. A Viewpoint on Direct Methane Conversion to Ethane and Ethylene Using Oxidative Coupling on Solid Catalysts. *ACS Catal.* **2016**, *6*, 4340–4346.
- (19) Zavyalova, U.; Holena, M.; Schlögl, R.; Baerns, M. Statistical Analysis of Past Catalytic Data on Oxidative Methane Coupling for New Insights into the Composition of High-Performance Catalysts. *ChemCatChem* **2011**, *3*, 1935–1947.
- (20) Lunsford, J. H. The Catalytic Oxidative Coupling of Methane. *Angew. Chem., Int. Ed.* **1995**, *34*, 970–980.
- (21) Baerns, M.; Kondratenko, E. V. *Handbook of Heterogeneous Catalysis*, 2nd ed.; Wiley-VCH Verlag GmbH & Co. KGaA, 2008; pp 3010–3023.
- (22) Makri, M.; Vayenas, C. G. Successful Scale up of Gas Recycle Reactor–separators for the Production of C₂H₄ from CH₄. *Appl. Catal., A* **2003**, *244*, 301–310.
- (23) Haag, S.; Bosomoiu, M.; van Veen, A. C.; Mirodatos, C. Oxidative coupling of methane in a catalytic membrane reactor: impact of the catalystmembrane interaction on the reactor performance. *Stud. Surf. Sci. Catal.* **2007**, *167*, 19–24.
- (24) Mleczko, L.; Baerns, M. Catalytic Oxidative Coupling of Methane—Reaction Engineering Aspects and Process Schemes. *Fuel Process. Technol.* **1995**, *42*, 217–248.
- (25) Simon, U.; Görke, O.; Berthold, A.; Arndt, S.; Schomäcker, R.; Schubert, H. Fluidized Bed Processing of Sodium Tungsten Manganese Catalysts for the Oxidative Coupling of Methane. *Chem. Eng. J.* **2011**, *168*, 1352–1359.
- (26) Arndt, S.; Otremba, T.; Simon, U.; Yildiz, M.; Schubert, H.; Schomäcker, R. Mn–Na₂WO₄/SiO₂ as Catalyst for the Oxidative Coupling of Methane. What is Really Known? *Appl. Catal., A* **2012**, *425*, 53–61.
- (27) Hintze, J. L.; Nelson, R. D. Violin Pots: a Box Plot-density Trace Synergism. *Am. Stat.* **1998**, *52*, 181–184.
- (28) Pedregosa, F.; et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.
- (29) Cortes, C.; Vapnik, V. Support-vector Networks. *Machine Learning* **1995**, *20*, 273–297.
- (30) Santosa, F.; Symes, W. W. Linear Inversion of Band-limited Reflection Seismograms. *SIAM J. Sci. Comput.* **1986**, *7*, 1307–1330.
- (31) Breiman, L. Random Forests. *Machine learning* **2001**, *45*, 5–32.
- (32) Geurts, P.; Ernst, D.; Wehenkel, L. Extremely Randomized Trees. *Machine learning* **2006**, *63*, 3–42.
- (33) Miyazato, I.; Tanaka, Y.; Takahashi, K. Accelerating the Discovery of Hidden Two-dimensional Magnets Using Machine Learning and First Principle Calculations. *J. Phys.: Condens. Matter* **2018**, *30*, 06LT01.
- (34) Takahashi, K.; Miyazato, I. Rapid Estimation of Activation Energy in Heterogeneous Catalytic Reactions via Machine Learning. *J. Comput. Chem.* **2018**, *39*, 2405–2408.
- (35) Miyazato, I.; Takahashi, I.; Takahashi, K. Automatic Oxidation Threshold Recognition of XAFS Data Using Supervised Machine Learning. *Mol. Syst. Des. Eng.* **2019**, *4*, 1014.
- (36) Inselberg, A.; Dimsdale, B. Parallel Coordinates: A Tool for Visualizing Multi-dimensional Geometry. *Proceedings of the First IEEE Conference on Visualization: Visualization '90*; 1990; pp 361–378.
- (37) Lee, J. S.; Oyama, S. Oxidative Coupling of Methane to Higher Hydrocarbons. *Catal. Rev.: Sci. Eng.* **1988**, *30*, 249–280.
- (38) Olah, G. A.; Molnár, Á. *Hydrocarbon Chemistry*; John Wiley & Sons, 2003.
- (39) Lee, M. R.; Park, M.-J.; Jeon, W.; Choi, J.-W.; Suh, Y.-W.; Suh, D. J. A Kinetic Model for the Oxidative Coupling of Methane over Na₂WO₄/Mn/SiO₂. *Fuel Process. Technol.* **2012**, *96*, 175–182.
- (40) Fleischer, V.; Steuer, R.; Parishan, S.; Schomäcker, R. Investigation of the Surface Reaction Network of the Oxidative Coupling of Methane over Na₂WO₄/Mn/SiO₂ Catalyst by Temperature Programmed and Dynamic Experiments. *J. Catal.* **2016**, *341*, 91–103.
- (41) Kitson, P. J.; Marie, G.; Francoia, J.-P.; Zalesskiy, S. S.; Sigerson, R. C.; Mathieson, J. S.; Cronin, L. Digitization of Multistep Organic Synthesis in Reactionware for On-demand Pharmaceuticals. *Science* **2018**, *359*, 314–319.