

# Surface adsorption model calculation database and its application to activity prediction of heterogeneous catalysts

Masato Kobayashi <sup>a,b,\*</sup>, Haruka Onoda <sup>c</sup>, Yusuke Kuroda <sup>c</sup>, Tetsuya Taketsugu <sup>a,b</sup>

<sup>a</sup>Faculty of Science and WPI-ICReDD, Hokkaido University, Kita 10 Nishi 8, Sapporo 060-0810, Japan

<sup>b</sup>ESICB, Kyoto University, 1-30 Goryo-Ohara, Nishikyo-ku, Kyoto 615-8245, Japan

<sup>c</sup>Graduate School of Chemical Sciences and Engineering, Hokkaido University, Kita 10 Nishi 8, Sapporo 060-0810, Japan

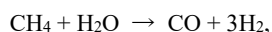
\**e-mail: k-masato@sci.hokudai.ac.jp*

(Received: December ##, 2019; Accepted for publication: ## ##, ####; Online publication: ## ##, ####)

The surface adsorption model calculation database, which we are constructing from systematic quantum chemical calculations, is reported. The database can be utilized to analyze the heterogeneous catalytic reactions. As an application, we predicted the experimental catalytic activity for the methane steam reforming reaction with this database by using the sparse modeling techniques.

**Keywords:** Calculation database, catalytic reaction, methane steam reforming, conversion rate, sparse modeling

Recent advances in the data-science methods have drawn attention to high-throughput quantum chemical calculations. In this study, we will report the surface model calculation database, which we are constructing by systematically carrying out the calculations of various chemical species adsorbed on metal surface models and storing their results, for heterogeneous catalytic reaction analyses. As an application example of this database, we attempted to evaluate the catalytic activity (namely, conversion rate) obtained from experiments for the following methane steam reforming (MSR) reaction,



by analyzing the information from the database and to specify its activation factor.

Our script program realized semi-automatic and systematic quantum chemical calculations with SIESTA program [1]. The

current database collects the results for the three-layer 3×3 fcc(111), (100), and (211) surface slab models (with 15 Å vacuum layer) and these surfaces with adsorbates at GGA-PBE/DZP level of calculations. Doping effect can be considered by substituting the central atom in the first layer with dopant. For each surface-adsorbate pair, various adsorption structures were considered by randomly generating 100 initial structures except for single-atom adsorption cases. Up to now, total 32,454 calculation results are collected in this database.

For the analysis of MSR reaction, we recollected total 1,028 calculation results, namely, those for clear surface and adsorption structures of H, C, O, H<sub>2</sub>, CH, OH, C<sub>2</sub>, and CO, for each metal surface from the database. As the statistical analysis, we utilized sparse modeling methods, where the following eq. (1) for the least square estimation with the regularization term  $p$  is minimized to perform the descriptor selection and regression simultaneously:

$$l(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + p(\boldsymbol{\beta}). \quad (1)$$

Here,  $\mathbf{y}$  represents the objective variable vector (i.e., methane conversion rate),  $\mathbf{X}$  is the descriptor matrix, and  $\boldsymbol{\beta}$  is the regression coefficient vector for descriptors. Namely, the following regularization terms (2) and (3) are used for LASSO [2] and MC+ [3] regressions, respectively:

$$p(\boldsymbol{\beta}) = \lambda \sum_j |\beta_j|, \quad (2)$$

$$p(\boldsymbol{\beta}) = \lambda \sum_j \int_0^{|\beta_j|} \left\{ 1 - \frac{x}{v\lambda} \right\}_+ dx, \quad (3)$$

where  $\{x\}_+ = \max(x, 0)$ . The regularization parameter  $\lambda$  was determined with the leave-one-out cross validation. These regressions were performed with the `ncvreg` package [4] in R language [5].

As the objective variable, the experimental conversion rates of methane on clean metal surface at around 1120 K were taken from Ref. [6]. Because the rates were obtained from the single article, the experimental conditions other than metal were the same. Table 1 lists the descriptors used in the sparse modelling regression taken from the surface adsorption model database. Since many adsorption structures are collected for adsorption calculations, we took maximum, minimum, average (avg), and standard deviation (sd) of these properties as the descriptors. Figure 1 shows the comparison of the methane conversion rates predicted by LASSO and MC+ methods with the experimental value. Particularly the MC+ prediction shows good agreement with the experimental value. Non-zero regression coefficients by MC+ method are listed in Table 2. It was found that the  $C_2$  adsorption energy on (111) surface negatively correlates with the experimental conversion rate, although  $C_2$  does not directly appear in the MSR reaction. Note that a negative adsorption energy represents the stabilization by the adsorption and that the  $C_2$  adsorption energy on (111) surface was selected also by LASSO regression with the largest negative coefficient. This correlation may reflect the deactivation by carbon deposition.

The present report shows the usefulness of our surface adsorption model calculation database for the analysis of

heterogeneous catalytic reactions. Note that this database can be used not only for the specific MSR reaction but also for wide varieties of heterogeneous catalytic reactions that proceed on metal surfaces. The application of the present method to the experimental results from multiple articles is now under consideration.

Table 1. Descriptor candidates taken from the surface adsorption model database to regress the methane conversion rate.

From metal surface calculation
- Relaxation energy for first layer
- Average charge of first-layer surface
From adsorption calculation for adsorbate $X$
- Adsorption energy (max, min, avg, sd)
- Average charge of surface (max, min, avg, sd)
- Charge of adsorbate (max, min, avg, sd)
- Height of adsorbate (max, min, avg, sd)
- Distance b/w adsorbed atoms (max, min, avg, sd)

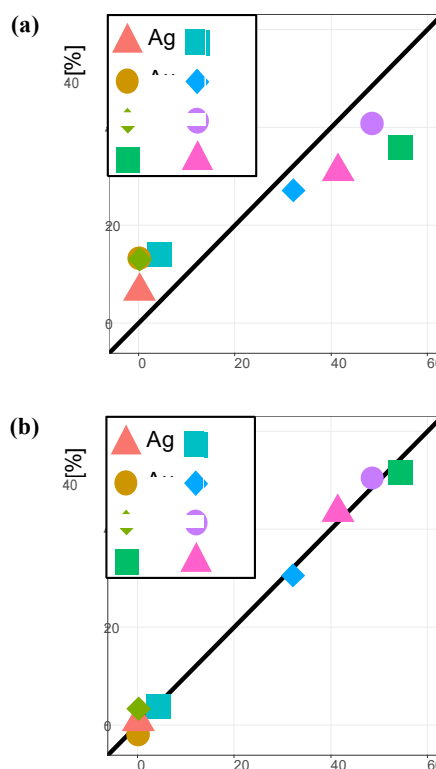


Figure 1. Comparison of the methane conversion rates predicted by (a) LASSO and (b) MC+ methods with the experimental value [3].

Table 2. Descriptors selected by MC+ method and their regression coefficients (in brackets) for methane conversion rate.

Max of C <sub>2</sub> adsorption energy on (111) surface [-26.54]
Min of C <sub>2</sub> adsorption height on (111) surface [8.83]
SD of C charge for CH adsorption on (111) surface [-1.40]

We are grateful for financial support from the programs of the Ministry of Education, Culture, Sports, Science and Technology (MEXT, Japan) “Elements Strategy Initiative to Form Core Research Center” (Grant Number JPMXP0112101003) and “Priority Issue on Post-K computer” (Development of new fundamental technologies for high-efficiency energy creation, conversion/storage and use). M.K. was supported by JST for PRESTO (Grant Number JPMJPR15N3). Some of the present calculations were performed using the computer facilities at Research Center for Computational Science, Okazaki and at Research Institute for

Information Technology, Kyushu University, Japan.

## References

- [1] J.M. Soler, E. Artacho, J.D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, *J. Phys.: Cond. Mat.* **14**, 2745 (2002).
- [2] R. Tibshirani, *J. Roy. Stat. Soc. B* **58**, 267 (1996).
- [3] C.-H. Zhang, *Ann. Stat.* **38**, 894 (2010).
- [4] P. Breheny and J. Huang, *Ann. Appl. Stat.* **5**, 232 (2011).
- [5] R Core Team, “R: A Language and Environment for Statistical Computing,” Foundation for Statistical Computing, Vienna, Austria, 2017.
- [6] D. Mei, V.-A. Glezakou, V. Lebarbier, L. Kovarik, H. Wan, K.O. Albrecht, M. Gerber, R. Rousseau, and R.A. Dagle, *J. Catal.* **316**, 11 (2014).