Supplementary Information

Accelerating Active Catalyst Discovery: A Probabilistic Prediction-Based Screening Methodology with Applications in Dry Reforming of Methane

Hyundo Park [a,b,1], Jiwon Roh [a,b,1], Hyungtae Cho [a], Insoo Ro [c,*], Junghwan Kim [b,*]

[a] Green Materials and Processes R&D Group, Korea Institute of Industrial Technology, 55 Jongga-ro, Ulsan 44413, Republic of Korea

[b] Department of Chemical and Biomolecular Engineering, Yonsei University, 50 Yonsei-ro, Seoul 03722, Republic of Korea

[c] Department of Chemical and Biomolecular Engineering, Seoul National University of Science and Technology, 232 Gongneung-ro, Seoul 01811, Republic of Korea

[1] These authors contributed equally to this work.

[*] Corresponding authors: Insoo@seoultech.ac.kr (I. Ro) and kjh24@yonsei.ac.kr (J. Kim)

1. Experimental setup for DRM reaction

In order to investigate the DRM reaction, an experimental setup for DRM reaction was constructed as shown in Figure S1. The experimental setup consists of three main parts: feed, reaction, and analysis part. The feed part consisted of Ar, CO_2, CH_4, H_2, and N_2 gas cylinders, and mass flow controllers (MFC) for both reactant gases. The main obligation of the feed part is to supply the components of interest including Ar, CO_2, CH_4, H_2, and N_2 to the reaction part, where a quartz-tbue fixed bed reactor was mounted vertically inside a furnace.

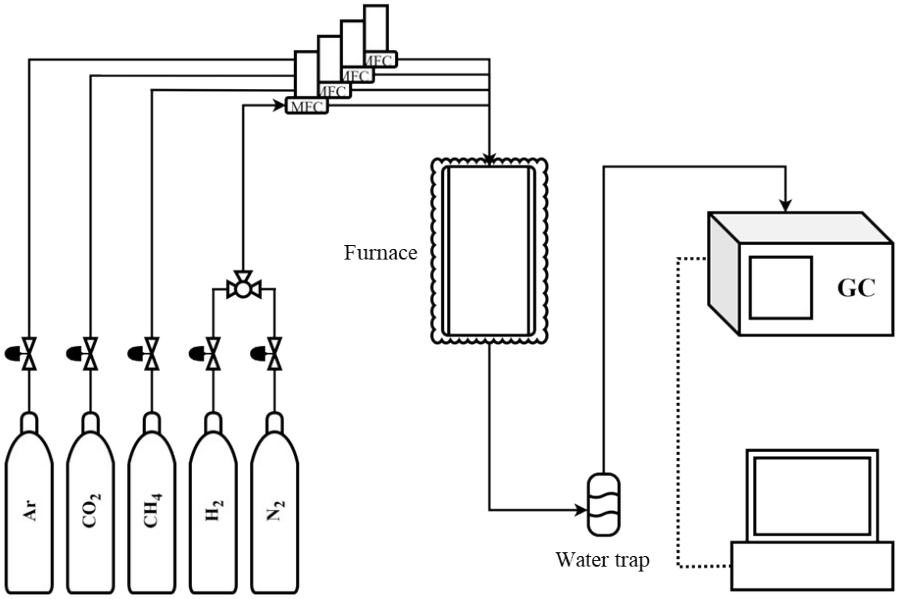

Fig. S1. Schematic of the experimental setup used for dry reforming of methane.

2. Probabilistic prediction model validation

We conducted N-fold cross-validation to validate the model. The cross-validation was performed for both 5 and 10 folds. Fig S2 shows the results of cross-validation using 5 and 10 folds.

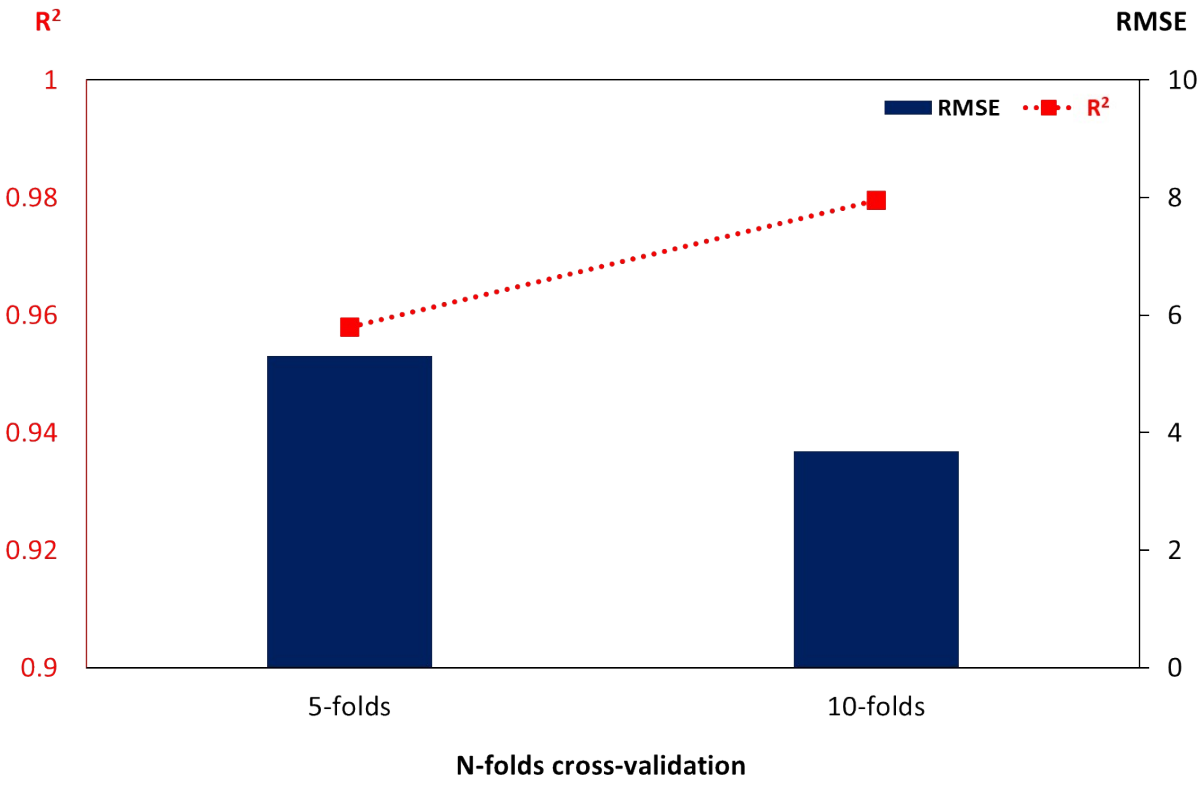

Fig. S2. The results of N-folds cross-validation using 5-folds and 10-folds.

3. Feature importance results

Table S1. Feature importance value

Domain	Component	Value
Design variables (Active metal)	Ni	0.1168
	Co	0.0623
	Pt	0.0373
	Pd	0.0197
	La	0.0190
	K	0.0158
	Rh	0.0121
	Ce	0.0105
	Cu	0.0101
	Zr	0.0101
	Mg	0.0086
	Ru	0.0080
	Y	0.0060
	Ir	0.0047
	Fe	0.0047
	Ca	0.0036
	Li	0.0021
	BN	0.0016
	Mn	0.0011
	Sr	0.0008
Design variables (Support)	ZrO_2	0.0453
	CeO_2	0.0170

	Y_2O_3	0.0153
	Al_2O_3_gamma	0.0149
	CaO	0.0129
	ZrO_2_nanocrystalline	0.0129
	MgO_nanocrystalline	0.0115
	SiO_2	0.0111
	MgO	0.0101
	ZSM_5	0.0097
	MnO	0.0082
	TiO_2	0.0069
	La_2O_3	0.0060
	PrO_2	0.0045
	SBA_15	0.0044
	Al_2O_3_alpha	0.0037
	$MgAl_2O_4$_nanocrystalline	0.0022
	Al_2O_3_mesoporous	0.0021
	$MgAlO_x$	0.0005
	Si_3N_4	0.0003
	MCM_41	0.0002
	V_2O_5	0.0001
Pretreatment variables	Calcination_Temperature	0.0507
	Reduction_ Temperature	0.0274
	Preparation	0.0189
	Calcination_time	0.0183
	Reduction_time	0.0155
	Reduction_H_2	0.0148
Operating variables	Reaction_ Temperature	0.0855

GHSV	0.0621
Time_on_stream	0.0538
Reaction_CO_2	0.0453
Reaction_CH_4	0.0278
Reaction_Ar	0.0114
Reaction_He	0.0082
Reaction_N_2	0.0058

37

38

4. The analysis of experimental validation results

Among the 3958 training data points, the distribution based on the active metal is as follows: Ni: 3044, Co: 553, Pt: 513, Rh: 130, Ru: 128, Pd: 52, Ir: 25, Ar: 21, Mn: 20, and Cu: 2. And the distribution based on the -ZrO_2 support is as follows: ZrO_2: 303, CaO-ZrO_2: 75, Y_2O_3-ZrO_2: 73, SiO_2-ZrO_2: 70, CeO_2-ZrO_2: 46, MgO-ZrO_2: 39, Al_2O_3-ZrO_2: 7, La_2O_3-ZrO_2: 2.

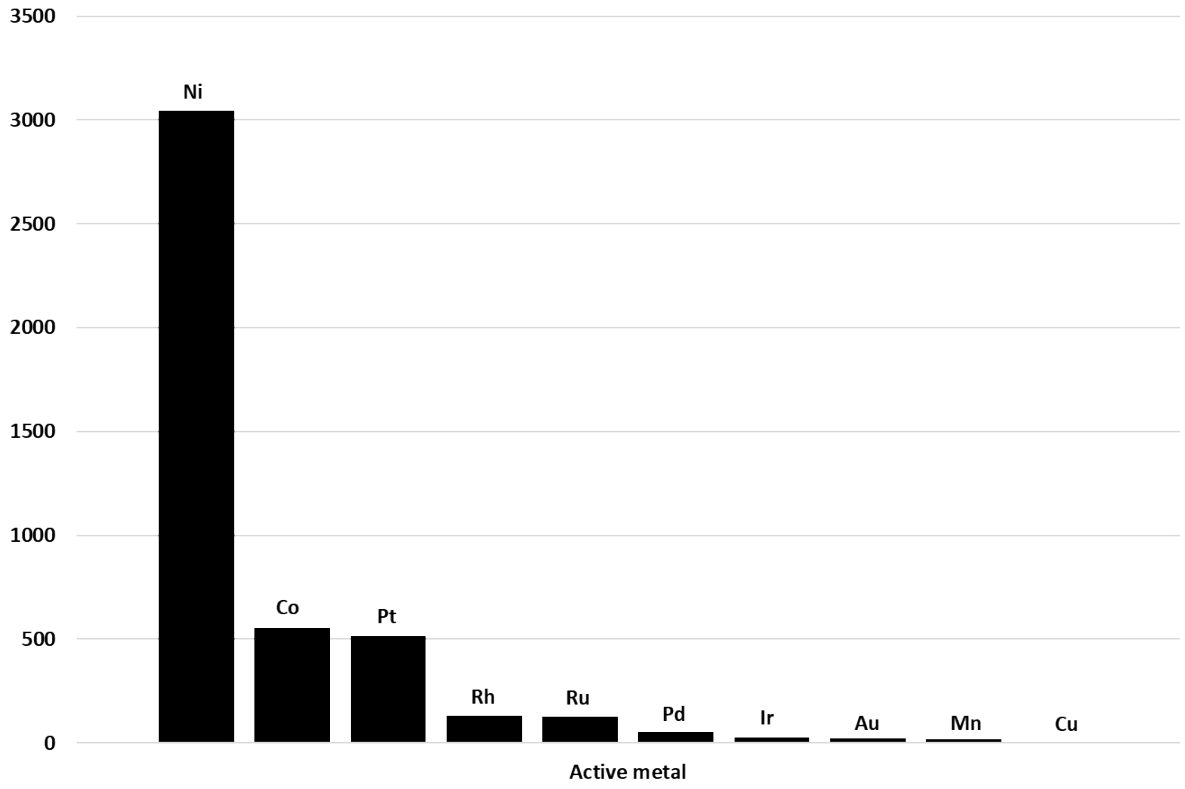

Fig. S3. Number of active metals in training data.

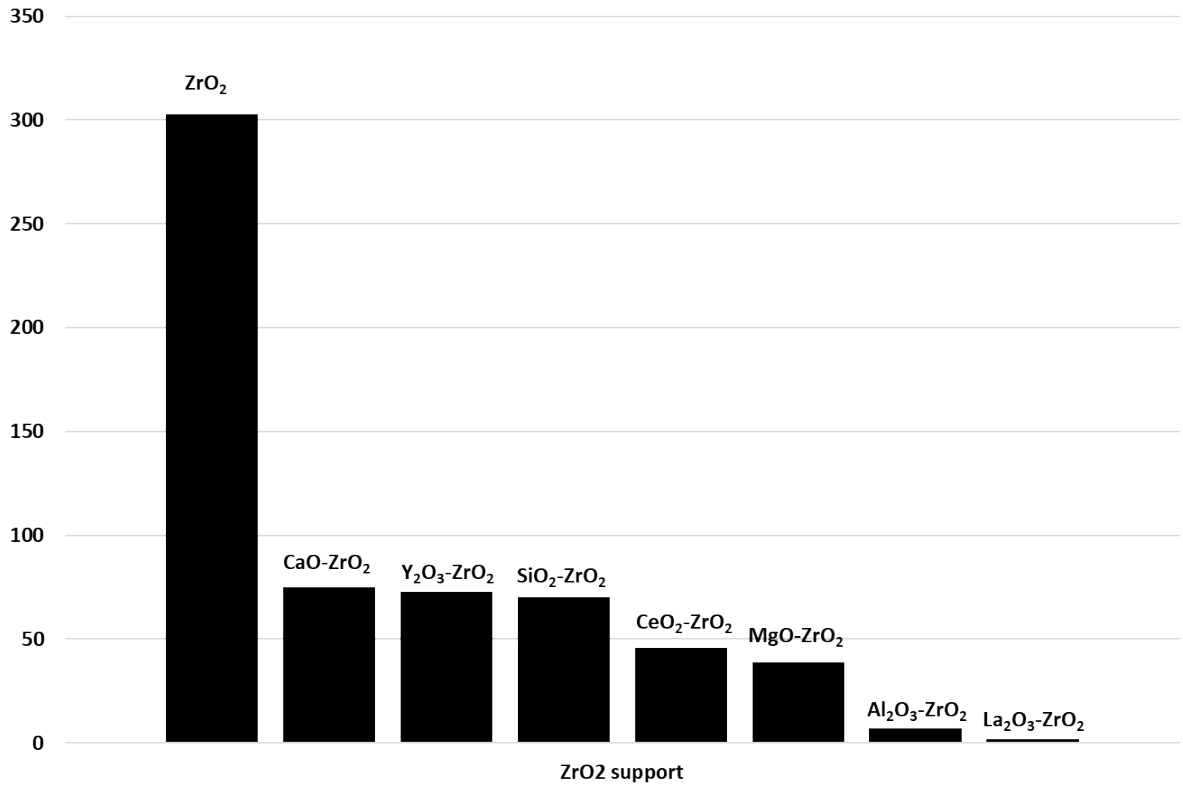

Fig. S4. Number of -ZrO_2 supported catalysts in training data.