## **Electronic supplementary information**

## Developing catalytic materials for the oxidative coupling of methane through statistical analysis of literature data

Evgenii V. Kondratenko<sup>a,\*</sup>, Michael Schlüter<sup>a</sup>,

Manfred Baerns<sup>b</sup>, David Linke<sup>a</sup>, Martin Holena<sup>a, \*</sup>

<sup>a)</sup> Leibniz-Institute for Catalysis at the University of Rostock, Albert-Einstein-Strasse 29a, 18059 Rostock (Germany); <u>martin.holena@catalysis.de</u> (MH), <u>evgenii.kondratenko@catalysis.de</u> (EVK)

<sup>b)</sup> Department of Inorganic Chemistry Fritz-Haber Institute of Max-Planck Society Faradayweg 4-6, 14195 Berlin (Germany)

## Kinds of empirical models considered in our research

i. The traditional quadratic response surface model,

$$f(x,w) = w_0 + \sum_{j=1}^d (w_j x_j + w_{j,j} x_j^2) + \sum_{j=2}^d \sum_{k=1}^{j-1} w_{j,k} x_j x_k.$$
(1)

Here, as in the main text of the paper, d is the dimension of input data x, and p is the dimension of model parameters w. If the number n of training pairs was not sufficient to learn all p parameters of the model (2), then only the purely quadratic model was learned instead:

$$f(x,w) = w_0 + \sum_{j=1}^d (w_j x_j + w_{j,j} x_j^2).$$
(2)

ii. Clustered RBF network with diagonal Gaussian radial basis functions centered in data:

$$f(x,w) = w_{k,0} + \sum_{i=1}^{B_k} \sum_{j \in D_k} w_{k,(i-1)d_k+j} e^{-\frac{\left(x_j - x_j^{(k,i)}\right)^2}{w_{k,(B_k+i-1)d_k+j}}} if x \in C_k, k = 1, \dots, n_C, \quad (3)$$

where  $n_c$  is the number of clusters,  $\mathcal{M} \subset \bigcup_{k=1}^{n_c} C_k, C_k \cap C_{k'} = \emptyset$  for  $k \neq k'$ ,

 $D_k$  is the set of  $d_k$  dimensions to which the modeling in the cluster  $C_k$  is

restricted, 
$$\emptyset \neq D_k \subset \{1, \dots, n\}$$
 for  $k = 1, \dots, n_c$ 

 $w_{k,(B_k+i-1)d_k+j}>0 \text{ for } i=1,\ldots,B_k, j\in D_k, k=1,\ldots,n_C,$ 

 $B_k$  is the number of basis functions used for modeling in the cluster  $C_k$ 

$$x^{(k,1)}, \dots, x^{(k,B_k)} \in \mathcal{M} \cap C_k \text{ for } k = 1, \dots, n_C, \text{ and } x^{(k,1)}, \dots, x^{(k,B_k)} \text{ fulfill}$$

$$E\left(w_{k,0} + \sum_{j \in D_k} w_{k,j} e^{-\frac{\left(x_j - x_j^{(k,1)}\right)^2}{w_{k,d_k+j}}}, \mathcal{M} \cap C_k\right)$$

$$= \min_{\xi \in \mathcal{M} \cap C_k} E\left(w_{k,0} + \sum_{j \in D_k} w_{k,j} e^{-\frac{\left(x_j - \xi_j\right)^2}{w_{k,d_k+j}}}, \mathcal{M} \cap C_k\right).$$
(4)

If  $B_k > 1$ , then for  $l = 1, \dots B_k$ ,  $\mathcal{M}_k^l = \mathcal{M} \cap C_k \setminus \{x^{(k,1)}, \dots, x^{(k,l)}\}$ , also

$$E\left(w_{k,0} + \sum_{i=1}^{l+1} \sum_{j \in D_{k}} w_{k,(i-1)d_{k}+j} e^{-\frac{\left(x_{j} - x_{j}^{(k,i)}\right)^{2}}{w_{k,(B_{k}+i-1)d_{k}+j}}}, \mathcal{M}_{k}^{\ell}\right)$$
$$= \min_{\xi \in \mathcal{M}_{k}^{\ell}} E\left(w_{k,0} + \sum_{j \in D_{k}} w_{k,ld_{k}+j} e^{-\frac{\left(x_{j} - \xi_{j}\right)^{2}}{w_{k,2ld_{k}+j}}}\right)$$
$$+ \sum_{i=1}^{l} \sum_{j \in D_{k}} w_{k,(i-1)d_{k}+j} e^{-\frac{\left(x_{j} - x_{j}^{(k,i)}\right)^{2}}{w_{k,(B_{k}+i-1)d_{k}+j}}}, \mathcal{M}_{k}^{\ell}\right).$$

iii. Clustered RBF network with diagonal Gaussian radial basis functions centered in data, combined with linear regression:

$$f(x,w) = w_{k,0} + \sum_{j \in D_k} w_{k,2B_k d_k + j} x_j$$

$$+ \sum_{i=1}^{B_k} \sum_{j \in D_k} w_{k,(i-1)d_k + j} e^{-\frac{\left(x_j - x_j^{(k,i)}\right)^2}{w_{k,(B_k + i-1)d_k + j}}} \text{ if } x \in C_k, k = 1, \dots, n_C,$$
(5)

where again the conditions (4) hold.



Fig. S 1 Time-on-stream methane conversion obtained over (a) La-Sr- (●-La<sub>2</sub>O<sub>3</sub>, □-LaSrCs, ○-LaSrNa, △-LaSrLi, ▽-LaSrMn, ◇-LaSrBa) and (b) Mg-Sr-based (■-MgO, ■-MgSrCs, ●-MgSrNa, ▲-MgSrLi, ▼-MgSrMn, ◆-MgSrBa) catalysts at 1073 K and τ of 0.0039 g min ml<sup>-1</sup> using a feed of 29% CH<sub>4</sub> in air.



Fig. S 2 Time-on-stream (a) methane conversion and (b) yield of C<sub>2</sub> hydrocarbons obtained over ●-La<sub>2</sub>O<sub>3</sub>, ●-LaBaLi, ●-LaBaNa, ●-LaBaCs, ●-LaBaMn at 1073 K and τ of 0.0039 g min ml<sup>-1</sup> using a feed of 29% CH<sub>4</sub> in air.



Fig. S 3 Time-on-stream (a) methane conversion and (b) yield of C<sub>2</sub> hydrocarbons obtained over ●-La<sub>2</sub>O<sub>3</sub>, ●-LaMgLi, ●-LaMgNa, ●-LaMgCs, ●-LaMgMn, ●-LaMgBa and ●-LaBaSr at 1073 K and τ of 0.0039 g min ml<sup>-1</sup> using a feed of 29% CH<sub>4</sub> in air.



Fig. S 4 Time-on-stream (a) methane conversion and (b) yield of C<sub>2</sub> hydrocarbons obtained over ●-La<sub>2</sub>O<sub>3</sub>, ●-LaNaLi, ●-LaNaCs, ●-LaNaMn, ●-LaLiMn, ●-LaCsLi and ●-LaCsMn at 1073 K and τ of 0.0039 g min ml<sup>-1</sup> using a feed of 29% CH<sub>4</sub> in air.



Fig. S 5 Time-on-stream (a) methane conversion and (b) yield of C<sub>2</sub> hydrocarbons obtained over ●-MgO, ●-MgBaLi, ●-MgBaCs, ●-LaBaNa, ●-LaBaMn at 1073 K and τ of 0.0039 g min ml<sup>-1</sup> using a feed of 29% CH<sub>4</sub> in air.



Fig. S 6 Time-on-stream (a) methane conversion and (b) yield of C<sub>2</sub> hydrocarbons obtained over ●-MgO, ●-MgLaLi, ●-MgLaNa, ●-MgLaCs, ●-MgLaMn, ●-MgLaBa and ●-MgLaSr at 1073 K and τ of 0.0039 g min ml<sup>-1</sup> using a feed of 29% CH<sub>4</sub> in air.



Fig. S 7 Time-on-stream (a) methane conversion and (b) yield of C<sub>2</sub> hydrocarbons obtained over ●-MgO, ●-MgNaLi, ●-MgNaCs, ●-MgNaMn, ●-MgLiMn, ●-MgCsLi and ●-MgCsMn at 1073 K and τ of 0.0039 g min ml<sup>-1</sup> using a feed of 29% CH<sub>4</sub> in air.