

CATALYST CHARACTERIZATION USING MACHINE LEARNING ANALYSIS OF TRANSIENT KINETIC DATA

M. Ross Kunz, Yixiao Wang, Rebecca Fushimi

Catalysis and Transient Kinetic Laboratory

Idaho National Laboratory

Idaho Falls, USA

rebecca.fushimi@inl.gov

Compared to conventional kinetic characterization conducted under steady-state conditions, transient kinetic experiments provide a volume of data that enables robust analysis using machine learning methods. Low-pressure pulse response experiments of an ammonia decomposition probe reaction were conducted on a series of metallic catalysts. Unique experimental conditions enable a linear relationship between concentration and reaction rate. Time-dependent rate and concentration data were analyzed using feature selection to extract a catalyst 'fingerprint' that removes bias from estimation of the reaction model.

Keywords: Transient Kinetics, Temporal Analysis of Products, Catalyst, Machine Learning

INTRODUCTION

At present, nearly 90% of chemical processes rely on a catalyst at some point in the manufacturing process [1]. Catalytic technologies are a key driver for sustainability as they control the productivity, selectivity and energy intensity in the production of essential molecules such as ammonia, benzene, ethylene, etc. that are the cornerstone of fertilizers, fuels and consumer goods supporting a prosperous society.

A catalyst controls chemical transformation, e.g. hydrogen and nitrogen to ammonia, through a series of elementary reaction steps orchestrated by unique structures, active sites, on the surface. A traditional microkinetic model describes the stoichiometry of an elementary reaction network along with the energetics of each step. This information provides is used for discriminating different catalyst formulations or for predicting how a catalyst will perform in different reactor settings. These models typically begin with experience and chemical intuition to create a meaningful network to which one compares experimental data. In general, microkinetic models are formulated using density function theory or surface science experiments of well-defined surface features, e.g. the Pt(111) surface termination. As such, these models can be difficult to extend to industrial catalysts which are complex, multicomponent supported metals or mixed metal oxides where the surface features are difficult to characterize.

The Rate-Reactivity Model (RRM) was proposed by Yablonsky et al. as a robust phenomenological model for systematic characterization of industrial catalysts using transient kinetic data [2]. This model describes how the reaction rate is influenced by a linear combination of gas and surface concentrations as well as their pairwise products:

$$R_i(t) = \beta_0 + \sum_k \beta_{i,k} C_k(t) + \sum_j \beta_{i,j} Z_j(t) + \sum_k \sum_j \beta_{i,k,j} C_k(t) Z_j(t) + \sum_j \sum_l \beta_{i,k,l} Z_j(t) Z_l(t) + \varepsilon$$

where $R_i(t)$ is the time-dependent transformation rate of gas i , β_0 is the intercept of the rate/concentration dependence and can be used to calculate the number of active sites, $C_i(t)$ is the gas concentration of species i and $Z(t)$ is the surface concentration or uptake time dependence of species i and ε is random noise. These variables are preceded by linear coefficients, β , referred to as 'reactivities' which convey clear physicochemical meaning and fully characterize the catalyst state. Unlike traditional microkinetic models, the RRM can be used to characterize a reaction network without the need for assumptions about the underlying mechanism.

FORMULATION OF THE PROBLEM

Compared to steady-state experiments, which only provide coarse kinetic data related to the slow reaction step, transient experiments significantly expand the volume of detailed elementary kinetic data that can be used to characterize an industrial catalyst. The unique physical characteristics of the TAP (Temporal Analysis of Products) pulse response experiment renders gas/solid interactions as first order and the time dependence of the rate, gas and surface concentration are directly observed [3]. The RRM can be implemented with exclusive use of TAP data and the application of robust statistics and variable selection methods enable a top-down approach to a novel form of microkinetic industrial catalyst characterization.

In this work, TAP pulse response experiments were conducted using ammonia decomposition as a probe reaction to distinguish a series of mono- and bimetallic catalysts. The Y-Procedure analytical method put forth by Yablonsky [4], was used to first transform pulse response exit flux data to time-dependent rate and concentration. This pure kinetic data recorded for both reactants and products was then analyzed according to the linear RRM which describes all potential interactions of gas and surface concentrations on the reaction rate. A combination of covariance estimation and feature selection was used to identify and quantify the influence of each interaction on the rate. The RRM reaction network of experimental data were expressed and interpreted using Elastic Net regression [5]. These initial steps advance towards our goal of providing a quantitative kinetic fingerprint to understand *how* each catalytic surface orchestrates the reaction mechanism and *why* certain materials perform better in the global catalytic process.

RESEARCH ANALYSIS

The six-step mechanism for ammonia synthesis originally proposed by Temkin is microscopically equivalent to the same decomposition mechanism with sequential dehydrogenation described by Ertl. In the TAP experiment, only gas phase species, steps 1, 5 and 6, are directly observed. Three materials, polycrystalline iron, cobalt and a bimetallic preparation of CoFe were evaluated in the thin zone TAP reactor configuration [3]. Ammonia was pulsed at 550 °C in the Knudsen flow regime (where gas-gas collisions are insignificant) and the pulse response exit flux of ammonia, nitrogen and hydrogen were recorded. The Y-Procedure method was used to convert exit flux to rate and concentration

with the time-dependence intact [4]. The rate data, height normalized for temporal comparison, for reactants and products is shown in Fig. 16. The entire pulse response was recorded over 3 seconds but only 1 second of data is shown for clarity.

Table 9. Ammonia decomposition

1	$NH_3 + Z \leftrightarrow ZNH_3$
2	$ZNH_3 + Z \leftrightarrow ZNH_2 + ZH$
3	$ZNH_2 + Z \leftrightarrow ZNH + ZH$
4	$ZNH + Z \leftrightarrow ZN + ZH$
5	$2ZH \leftrightarrow H_2 + 2Z$
6	$2ZN \leftrightarrow N_2 + 2Z$

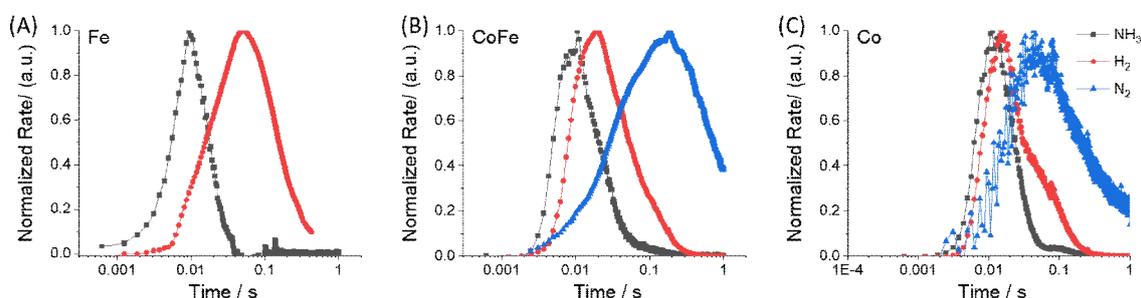


Fig. 16. Time-dependent transformation rate for ammonia, hydrogen and nitrogen when ammonia is pulsed over A) Fe, B) Co and C) CoFe catalysts at 550 °C in the thin zone TAP reactor

While nitrogen production was observed over the iron sample, the signal was too low for reliable rate calculation using the Y-Procedure. There is a distinct time delay between the conversion of ammonia and the production of hydrogen and nitrogen for each sample. In the case of CoFe and Co, the production of nitrogen is slower than that of hydrogen.

In the traditional approach, one would use the well-defined physics of the TAP reactor and the Temkin/Ertl microkinetic mechanism to simulate the reaction-diffusion system and regress rate parameters for steps 1-6. While the Temkin/Ertl mechanism is derived from extensive research and experimentation, it nonetheless excludes numerous possible steps based on this collected experience. Moreover, this approach is problematic as we do not directly observe reactions steps 2 – 4. As an alternative, we apply the generalized RRM which makes no *a priori* assumption of the kinetic model and only parses the rate as some linear combination of interactions of the gas and surface concentrations, including a linear dependence on their polynomial effects (which would account for phenomena where, for example, two surface species may influence the rate). While this approach does not describe the surface reactions explicitly, as in reaction steps 2 – 4, from the time-dependent rate information, the integral can be used to calculate the surface concentration or uptake/release of species during the pulse, $Z_i = \int_0^t R_i dt$. Thus, the rate of ammonia transformation may be described as:

$$R_{NH_3}(t) = k_{app} C_{NH_3}(t) (N - 2Z_{NH_3}(t) + Z_{N_2}(t) + 3Z_{H_2}(t)),$$

where k_{app} is the apparent rate constant and N is the total number of active sites. The concentration of surface species is some linear function of the uptake and release of molecules on the surface; this function will change according to different reaction mechanisms. This linear model, however, is restrictive in the sense that there is an assumed interaction of species based on stoichiometry and the model contains only one apparent rate constant for determining the rate. In the transient experiment where global stoichiometry is not observed, one can make a compromise using the total atom balance in the gas phase to estimate the quantity of N and H retained on the surface in different forms, *i.e.* NH_x . Each different stoichiometric and rate constant combination may be tested based on the RRM which, for ammonia decomposition, can be expressed as:

$$R_{NH_3}(t) = \beta_0 + \beta_1 C_{NH_3}(t) + \beta_2 C_{NH_3}(t) Z_{NH_x}(t) + \beta_3 C_{NH_3}(t) Z_N(t) + \beta_4 C_{NH_3}(t) Z_H(t) + \beta_5 Z_{NH_x}(t) Z_N(t) + \beta_6 Z_{NH_x}(t) Z_H(t) + \beta_7 Z_N(t) Z_H(t) + \varepsilon$$

where Z_N and Z_H represent the atomic surface concentration of nitrogen and hydrogen, respectively, for all surface species. Only the gas concentration of ammonia is considered since the TAP experiment is conducted far from equilibrium and backwards reactions of products are expected to be negligible. However, it is straightforward to analyze an exhaustive list of all 21 variables that could be considered for ammonia decomposition.

The RRM may be expressed in more compact notation if we let the dependent variable, y , be defined as time-dependent vector of reaction rates: $y \in \mathbb{R}^{n \times 1}$ where n is equal to the number of time-points collected. Then our predictor matrix, X , consists of a set of observations with a number of columns, p , for time-dependent vectors for the observed gas concentrations, $C_i(t)$, surface concentrations, $Z_i(t)$, their pairwise products, $C_i(t)Z_i(t)$, etc. according to the RRM: $X \in \mathbb{R}^{n \times p}$. We can then solve for the set of reactivity coefficients through optimization of β , *i.e.* minimizing the differences between the reaction rate and the predictor matrix according to the general form of the ordinary least squares (OLS) model:

$$\min_{\beta} \|y - X\beta\|_2^2$$

where $\|\cdot\|_2^2$ denotes the l_2 -norm or Euclidean distances and it is assumed that each variable is mean centered so that the intercept is zero. A key distinction should be made, in our approach the experimental data is not used for model validation but rather for model reduction. In other words, the regression model is used to correctly estimate the coefficient weights, or lack thereof, for a particular concentration to impact the rate of chemical transformation.

Since the reaction rates of reactants and products within a pulse are highly correlated, this can result in high variability and poorly determined reactivity coefficients using linear regression. Multicollinearity is especially prevalent since the surface concentrations are calculated through the integral of the reaction rate. This may be combatted through imposing two penalties, λ_1 and λ_2 , on the size of each coefficient through Elastic Net regression such that:

$$\|y - X\beta\|_2^2 + \lambda_1 \|\beta\|_2^2 + \lambda_2 |\beta|_1,$$

where $|\cdot|_1$ denotes the l_1 -norm enforcing sparsity on the coefficients and the additional l_2 -norm alleviates multicollinearity [5]. Table 10 summarizes the kinetic information determined from ammonia decomposition pulse response data collected for each catalyst using different analysis methods. The first method, *conversion*, takes only the ammonia rate and gas concentration (conversion) into account. Here, the time-dependent data is collapsed into a single summary statistic and represents the type of measurement that would result from conventional steady-state analysis. Both the number of active sites and apparent rate constant can be estimated but the correlation value is very low indicating there is additional information in the data that has not been accounted for by this simplistic model. The value of the apparent rate constant, k_{app} , for the iron catalyst is similar to values reported in the literature. However, the extreme value determined for cobalt using this method is likely a result of the almost complete conversion of ammonia which results in a concentration term close to zero.

Table 10. Summary information of active site and apparent rate constant estimates

Catalyst	Method	β_0 , (active sites)	β_1 , (\bar{k}_{app} , 1/s)	R^2
Fe	Conversion	1.87×10^{12}	126	0.08
	RRM	4.16×10^{11}	138	0.89
	EN-RRM	3.19×10^{11}	117	0.90
CoFe	Conversion	4.66×10^{12}	980	0.15
	RRM	2.43×10^{11}	650	0.92
	EN-RRM	6.85×10^{11}	633	0.95
Co	Conversion	2.18×10^{12}	17,100	0.32
	RRM	1.76×10^{11}	584	0.88
	EN-RRM	3.54×10^{11}	665	0.88

When using the RRM, the time-dependent reactant and product data is included in the analysis and while the correlation value increases, the number of active sites decreases by an order of magnitude and the apparent rate constant estimate for CoFe and cobalt are significantly low. The model including Elastic Net regression, *EN-RRM*, takes into account the co-linearity of the pulse response kinetic data and generates estimates for the number of active sites and apparent rate constant that agree with the observed conversion along with reliable correlation coefficients indicating that the majority of the data has contributed to the estimate.

Higher order reactivity coefficients can be interpreted as well but are not considered here for brevity. These results indicate that the transient experiment, when accounting for inherent co-linearity, can provide significantly greater data for robust estimate of kinetic features. Moreover, from the detailed transient data, an experimentally determined microkinetic network connectivity can be examined using graph structure. Fig. 17, results from the inverse covariance matrix of all gas species within the system. The gas and surface species are indicated by *g* and *s*, respectively. For example, H(*g*) = gas phase hydrogen, N(*s*) = surface nitrogen, NH(*g*) = gas phase ammonia. The scalar values within the matrix, with a larger value indicating stronger correlation, are indicated by the edge intensity. Negative interactions (reducing the rate of ammonia conversion) are indicated in red and positive interactions are indicated in blue.

From the graph structure, one observes that for the iron catalyst, the formation of hydrogen is dependent on the ammonia gas and hydrogen surface concentration as might be expected from the conventional mechanism in. However, an additional process step is observed where nitrogen gas interacts with hydrogen on the surface. The formation of nitrogen gas is impacted by the same three processes as for hydrogen formation with additional contributions from both nitrogen gas and surface concentrations. A different analysis can be made for the CoFe and Co (not included here) catalysts. The connections between the gas and surface species found for the CoFe catalyst are remarkably distinct from Fe. This new analysis of a connectivity network, derived from experimental data, can be used as a *mechanistic fingerprint* for distinguishing the microkinetic features of different materials. Development and demonstrations of this new approach for different catalytic systems is ongoing.

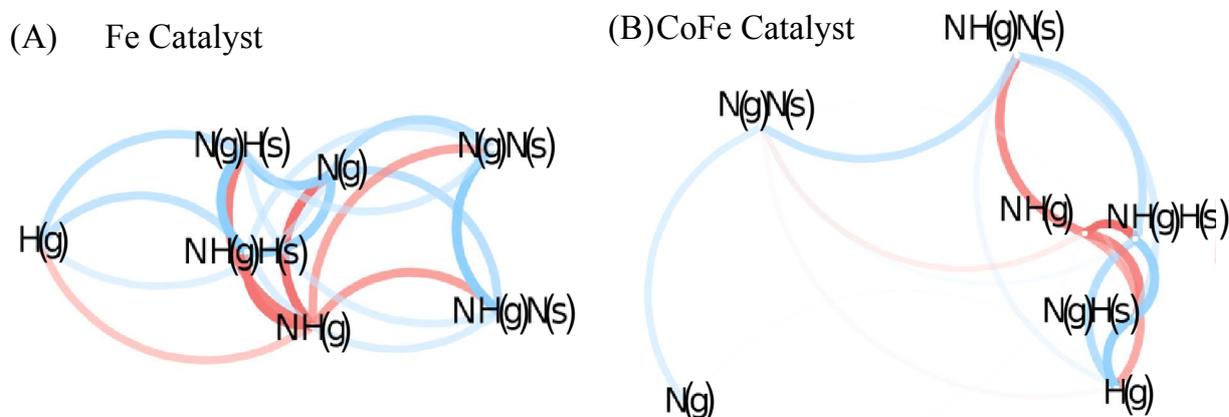


Fig. 17. Multidimensional clustering of gas and surface interactions different catalysts

CONCLUSIONS

Transient kinetic experiments provide a unique characterization of both catalytic properties and reaction network features. Compared to steady-state experiments, which provide only slow measurement of a singular kinetic state, the transient experiment offers rate/concentration kinetic data 1000-fold in a 3 second pulse and hence could be considered 'high-throughput kinetic sampling'. Quantitative connectivity from transient experiments, conveyed in a graph structure, presents a new approach to microkinetic characterization of reaction networks, screened by frugal variable selection tools, that arise solely from experimental data and contain limited *a priori* assumptions.

REFERENCES

- [1] B.L. Farrell, *et al.*, ACS Publications, 2016.
- [2] G.S. Yablonsky, *et al.*, International Journal of Chemical Kinetics, 48 (2016) 304-317.
- [3] K. Morgan, *et al.*, Catalysis Science & Technology, 7 (2017) 2416-2439.
- [4] G.S. Yablonsky, *et al.*, Chemical Engineering Science, 62 (2007) 6754-6767.
- [5] H. Zou, *et al.*, Journal of the Royal Statistical Society: Series B, 67.2 (2005) 301-320.