

新規固体触媒候補の提案に関する手法の開発

○ 植村圭祐, 荒川正幹, 船津公人

東京大学大学院工学系研究科(〒113-8656 東京都文京区本郷7-3-1)

1. 緒言

近年ますます重要になってきている環境調和型化学プロセスの実現のためには、高活性で高選択性の触媒の開発が必要である。しかし触媒作用は非常に複雑であり、触媒性能を支配する因子が十分に解明されていないことが多いため、性能の高い触媒を効率よく開発することは容易ではない。そこでケモインフォマティクス(化学情報学)による触媒設計支援手法の開発が望まれる。

本研究の目的は、触媒とその性能との間の定量的関係について予測精度の高いモデルを構築し、そのモデルを用いて信頼性の高い新規触媒候補を提案する手法を開発することである。ケーススタディとして、プロピレンの酸素酸化によるプロピレンオキシド製造用固体触媒について特許から収集したデータの解析を行い、新規触媒候補の提案を行った。

2. 手法

図1に解析の流れを示す。特許から収集したデータを用いて、触媒とその性能との間のモデリングを行い、モデルの逆解析により新規触媒候補を探索する。

モデリングとは、対象に影響を与える因子 X (説明変数)によって性能や特性を評価する変数 Y (目的変数)の数学的な近似を行なうことである。今回用いたPLS(Partial Least Squares)[1]は、サンプル数よりも変数の数の方が多いデータや多重共線性のあるデータでも、安定した線型モデルを構築できる手法である。また、データ中の重要な情報とノイズを分離できるためノイズに強いという特徴がある。

一般に説明変数の数が多いほどモデルの適合度 R^2 値(説明分散)は高くなる傾向があるが、適切でない説明変数が含まれているとモデルの予測性 Q^2 値(Leave-One-Outによる予測的説明分散)が低くなる。従って必要な説明変数のみを選択することでモデルの予測精度が高くなる。GAPLS法[2]は、遺伝的アルゴリズム(GA:Genetic algorithm)を用いてPLSの Q^2 値が最大になるような変数の組を見つけ出す変数選択法である。

逆解析とは、構築されたモデルを用いて目的変数 Y が最適になるような説明変数 X を求めることである。今回用いたグリッドサーチは、各説明変数の定義域を一定間隔で分割して全ての場合について目的変数を予測し最適解を探す手法である。

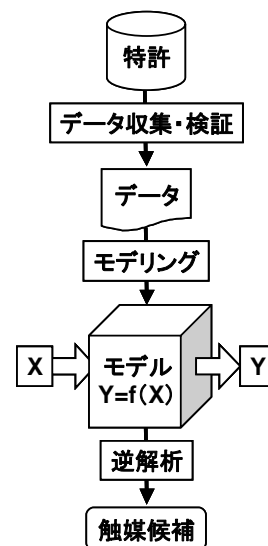


図1 解析の流れ

3. 結果と考察

3.1 解析データ

プロピレンからプロピレンオキシドへの気相酸素酸化触媒に関する特許8件からデータを収集した。反応温度が低いデータは収率が非常に低かったため、反応温度が200℃以上のデータを解析に用いた。また担体は CaCO_3 のものが最も性能がよくデータ数も多かったため、担体が CaCO_3 のデータを用いた。収集したデータから外れ値を除いた結果データ数は75件になった。データの一部を表1に示す。説明変数には触媒の組成、反応条件(温度、圧力、空間速度、原料気体の組成)、触媒前処理(有1、無0)を用いた。また目的変数には、プロピレン転化率、プロピレンオキシド選択率を用いて解析を行なった。

表1 データ例(収率の高い順)

Patent	触媒組成										反応条件							触媒前処理			触媒性能		
	Ag wt. %	K wt. %	Mo wt. %	W wt. %	Re wt. %	Au wt. %	Cl wt. %	Mg ppm	Fe ppm	Temp. °C	P psig	GHSV hr ⁻¹	C ₃ = mol %	O ₂ mol %	NO _x ppm	EtCl ppm	CO ₂ mol %	NO _x 0/1	EtCl 0/1	CO ₂ 0/1	Conv. %	Sel. %	Yield %
特表2001-526279	40	2	0	4.7	0	0	0	0	0	258	40	1200	10	5	200	50	10	0	0	0	12	48	5.76
特表2001-526279	40	2	0	4.7	0	0	0	0	0	258	40	1200	10	5	200	50	5	0	0	0	13	44	5.72
特表2001-526279	50	2	0.5	0	0	0	0	0	0	260	40	1200	10	5	200	50	0	0	0	1	13	44	5.72
特表2001-526279	50	2	0.5	0	0	0	0	0	0	260	40	1200	10	5	200	50	0	0	0	0	13	42	5.46
特表2001-526279	40	2	0	4.7	0	0	0	0	0	258	40	1200	10	5	200	50	25	0	0	0	10	54	5.40
特表2002-504911	51	2.1	0.51	0	0	0	0	0	0	250	30	1200	4	8	0	0	0	0	1	1	10	54	5.40
特表2001-526279	40	2	0	4.7	0	0	0	0	0	258	40	1200	10	5	200	50	0	0	0	1	13	41	5.33

3.2 モデリング

はじめに線型モデリング手法である PLS を用いて転化率と選択率それぞれのモデリングを行なったが、精度の高いモデルを構築することはできなかった。これはデータに非線型性があるためと考えられる。そこで2乗項と交差項を説明変数に加えた。変数の数は230個になり、データ数(75件)と比べて多すぎるため GAPLS を用いて変数選択を行なった。その結果、

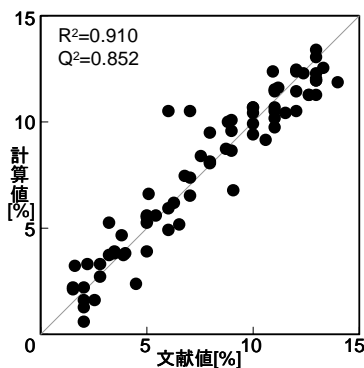


図2 転化率の文献値と計算値

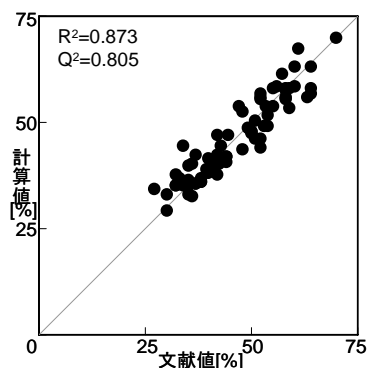


図3 選択率の文献値と計算値

転化率と選択率のそれぞれについて Q^2 値の高いモデルが得られた。文献値とモデルによる計算値の関係を図 2、3 に示す。どちらのグラフも点が対角線付近に集まっていることから、モデルの精度が高いことが確認できる。

3.3 逆解析

反応条件、触媒前処理をデータの中央値(表 2)に固定し、触媒組成について定義域(データの最小値と最大値の間)を5等分してグリッドサーチを行なった。ただし交差項についても定義域の範囲に収まるように制約条件を設けた。

表2 データの中央値・最小値・最大値

	触媒組成									反応条件							触媒前処理			触媒性能			
	Ag	K	Mo	W	Re	Au	Cl	Mg	Fe	Temp.	P	GHSV	C ₃ =	O ₂	NO _x	EtCl	CO ₂	NO _x	EtCl	CO ₂	Conv.	Sel.	Yield
中央値	44	2	0	0	0	0	0	0	0	250	30	1200	10	5	200	50	0	0	0	0	8	44.3	3.6
最小値	25	0.57	0	0	0	0	0	0	0	216	30	1200	4	5	0	0	0	0	0	0	1.5	27	0.8
最大値	54	2.3	1.8	5	0.5	4.6	0.7	7900	700	260	300	4800	10.7	8.15	380	500	52.1	1	1	1	14	70	5.76

一般には転化率が上がると選択率が下がり、選択率が上がると転化率が下がるという関係があるため、転化率と選択率を同時に最適化することはできない。そこで転化率と選択率に関するパレート最適解を求めたところ、14個の解が見つかった。これらの触媒候補の転化率と選択率の関係を図4に示す。十字印は特許から収集したデータを表し、丸印は逆解析により見つかった触媒候補を表している。グラフの右上に並んでいる候補は学習データよりも収率が高いと予想される。このうち収率の予測値が8%以上の触媒候補の組成を表3に示す。学習データ(表1)と比較して Ag、Re、Cl、Mg が多く含まれていることがわかる。

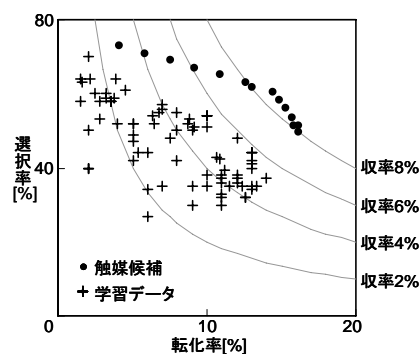


図4 触媒候補の転化率と選択率

4. 結言

化学情報学の手法を用いて、触媒組成と反応条件から触媒性能を予測するモデルを構築することができた。そして学習データの触媒より性能が高いと予想される触媒候補を提案することができた。この手法を用いることで、触媒機構の豊富な知識を必要とせず、短時間のうちに少ない実験回数でより性能の高い触媒を設計できるようになると期待される。

表3 触媒候補

	Ag	K	Mo	W	Re	Au	Cl	Mg	Fe	転化率	選択率	収率
1	54	0.57	0	5	0.5	0	0.56	7900	0	14.4	60.7	8.75
2	54	0.916	0	5	0.5	0	0.56	7900	0	14.8	58.4	8.66
3	54	1.262	0	5	0.5	0	0.56	7900	0	15.3	56.1	8.56
4	54	1.608	0	5	0.5	0	0.56	7900	0	15.7	53.8	8.43
5	54	1.954	0	5	0.5	0	0.56	7900	0	16.1	51.5	8.29
6	54	1.954	0	5	0.5	0	0.42	7900	0	15.8	51.6	8.14
7	54	0.57	0	5	0.5	0.92	0.56	7900	0	13.1	61.7	8.05
8	54	2.3	0	5	0.5	0	0.42	7900	0	16.1	49.6	8.00

謝辞

多くの助言を頂いた東京大学堂免教授とデータ収集にご協力頂いた堂免研究室の皆さまに感謝いたします。

参考文献

- [1] S. Wold et al. , Chemom. Intell. Lab. Syst. 58 (2001) 109-130.
- [2] K. Hasegawa et al. , J. Chem. Inf. Comput. Sci. 37 (1997) 306-310.