

令和 5 年 6 月 21 日現在

機関番号：13601

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K05207

研究課題名（和文）多環芳香族の接触分解に影響する原料因子の抽出と反応予測モデルの構築

研究課題名（英文）Extraction of feedstock factors affecting catalytic cracking of polycyclic aromatic hydrocarbons and development of a reaction prediction model

研究代表者

嶋田 五百里（Shimada, Iori）

信州大学・学術研究院繊維学系・講師

研究者番号：40708187

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：残油流動接触分解（RFCC）プロセスにおいて、多環芳香族炭化水素は共存成分との相互作用によって分解し、高付加価値な単環芳香族炭化水素に転換される。しかし、RFCCプロセスの原料組成は極めて複雑であり、個別の成分の反応性を考慮した反応モデルの構築は困難であった。そこで本研究では、原料組成と反応条件から生成物組成を予測する機械学習モデルを構築するため、線形回帰と物理化学に基づく特徴量エンジニアリングを組み合わせることを検討した。これにより、ブラックボックスな非線形回帰モデルよりも高い予測精度を達成するとともに、構築したモデルを解析することで重要な反応因子を抽出できる可能性を示すことができた。

研究成果の学術的意義や社会的意義

本研究では、複雑な成分組成を持つ原料の触媒反応を対象として、原料組成と反応条件から生成物組成を予測する機械学習モデルを構築した。さらに、構築したモデルの解析から重要な反応因子を抽出できる可能性を示した。これらの成果は重質油の接触分解反応に限らず、例えばバイオマス資源や廃プラスチック油などの極めて複雑な組成を持つ原料の反応予測にも適用することができる。重要な反応因子を抽出することで、反応条件の最適化や優れた触媒設計につながる知見を得ることができ、低炭素社会や循環型社会の構築に貢献する優れたプロセスの開発につながることを期待できる。

研究成果の概要（英文）：In the residue fluid catalytic cracking (RFCC) process, polycyclic aromatic hydrocarbons are decomposed by interaction with coexisting components and converted to high-value monocyclic aromatic hydrocarbons. However, the feedstock composition of the RFCC process is extremely complex, and it is difficult to construct a reaction model that takes into account the reactivity of individual components. Therefore, in this study, we investigated the combination of linear regression and physics-based feature engineering to construct a machine learning model that predicts the product composition from the feedstock composition and reaction conditions. LASSO model with physics-based feature engineering achieved higher prediction accuracy than black-box nonlinear regression models, and also showed the possibility of extracting important reaction factors by analyzing the constructed model.

研究分野：反応工学

キーワード：石油精製 接触分解 機械学習 特徴量エンジニアリング

## 1. 研究開始当初の背景

重質油を分解し付加価値の高い軽質成分に転換する技術は石油のノーブルユースの観点から重要である。現状の石油精製では、重質油の分解には残油流動接触分解 (RFCC) プロセスが用いられる。RFCC プロセスは外部水素の導入を必要とせず、ゼオライト触媒を用いて分解を進行させる安価なプロセスである。しかし、重質油に含まれる多環芳香族は難分解性であり、触媒上でのコーク生成の原因物質となるため、その有効な転換技術の開発が求められている。一般に、多環芳香族の分解には水素化処理が用いられるが、水素化には多量の水素を必要とするためプロセスコストが高い。一方、RFCC プロセス内で多環芳香族を直接分解できると、重質油の安価かつ効率的な有効活用につながる。

我々は、接触分解反応場で進行する水素移行反応 (2 分子間での水素の授受反応) に着目し、多環芳香族の分解への応用を試みてきた。その中で、水素移行反応活性の高い希土類担持 USY 型ゼオライトを触媒に用い、3 環芳香族とパラフィンを混合したモデル原料の接触分解に取り組んだところ、3 環芳香族の 60%以上を分解し、選択率 25%程度で単環芳香族へ転換できることを見出した (図 1)。これに対し、3 環芳香族を単独で接触分解した場合の転化率は 20%程度と報告されており、水素供与体の共存が 3 環芳香族の分解を促進することが示唆された。

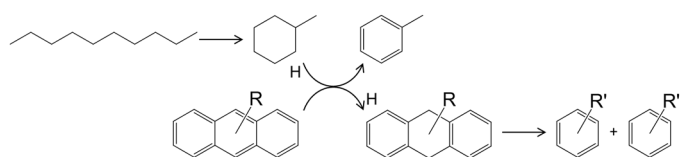


図 1 水素移行反応を利用した 3 環芳香族の接触分解反応機構

一方、重質油は非常に複雑な組成を持つ混合物であり、モデル物質を用いて得られた知見の活用には課題がある。従来、重質油の反応モデルには、沸点ごとに成分をまとめて取り扱うランプモデルが用いられたため、詳細な組成や分子構造の違いによる反応性の違いを反映することはできていなかった。それに対し、近年のペトリオミクス技術の発展により、重質油の組成を分子レベルで分析することが可能になってきた。しかし、膨大な成分数を含む原料の反応ネットワークは極めて複雑であり、特に水素移行反応のような 2 分子反応では組み合わせ数が発散するため、素反応レベルで網羅的に考慮することは困難である。そこで、原料の組成や含まれる分子構造の情報に基づいて反応を予測する技術の開発が求められる。これを達成することができると、原料の違いを反映した反応条件最適化や、多環芳香族の更なる分解促進に向けた触媒設計などにつながることを期待される。

## 2. 研究の目的

本研究では、多環芳香族の接触分解反応を対象として、複雑な原料組成を持つ反応における生成物組成を予測するモデルを構築する。さらに、構築したモデルを解析することで生成物組成を決定するのに重要な役割を担う反応因子を抽出することを目指す。

## 3. 研究の方法

### 3. 1. 研究の方針

膨大な組み合わせの反応の結果として起こる生成物組成の変化について解析をするため、本研究では機械学習の利用を検討した。固体触媒反応の予測に機械学習を導入した研究は近年多く報告されているが、報告されている反応系は C1 化学等の比較的単純なものが多く、重質油のような複雑な原料組成をもつ反応への適用方法については検討例が少ない。そこで本研究では、3 環芳香族を含む複数のモデル物質の混合物からなる反応原料を用いた反応系において、原料組成と反応条件から生成物組成を予測する機械学習モデルの構築に取り組んだ。

触媒反応の反応条件と反応生成物の関係は複雑であるため、既往研究では非線形モデルの利用により予測精度が向上することが報告されているが、反応条件を機械学習モデルに組み込む方法についての詳細な議論はされていない。ここで、機械学習では一般に予測精度と解釈可能性の間にトレードオフがあることが知られており、予測精度の高い非線形モデルはブラックボックスで解釈が困難なため重要な反応を抽出することは難しい。そこで、予測精度と解釈可能性を両立するモデルの構築が求められる。

予測精度と解釈可能性の両立に向けたアプローチの一つとして、物理と機械学習の融合がある。物理的なルールを機械学習モデルの構造中や損失関数に組み込むことで予測精度の向上や解釈可能性の改善が報告されている。また、物理法則を用いて特徴量を拡張する特徴量エンジニアリングも報告されている。しかし、これらの研究では非線形回帰手法が用いられており、モデルの解釈性は十分とは言えない。また、触媒分野での検討例は報告されていない。本研究では、解釈可能性の高い線形回帰を用いつつ、物理化学に基づいた特徴量エンジニアリングを行うことで生成物組成予測モデルの予測精度の向上を試みた。

### 3.2. 実験及び解析の方法

解析対象には反応温度、接触時間、原料組成を変更して行った 62 点の反応試験データを用いた。原料は *n*-ヘキサデカン（パラフィン）、9-メチルアントラセン（多環芳香族）、テトラリン（ナフテン）、トルエン（不活性物質）、9,10-ジヒドロアントラセン（反応中間体）を様々な比率で混合して調製した。触媒には残油流動接触分解プロセス用平衡触媒を用いた。反応試験は固定床流通式反応器を用いて行い、反応温度は 723 - 823 K、通油時間は 75 s、接触時間の指標である原料触媒比（Cat/oil）は 0.4 - 7.5 g/g とした。生成物中に含まれる単環芳香族（MAH）、縮合 2 環芳香族（c2PAH）、3 環・4 環芳香族及びコーク（3PAH/4PAH/coke）を定量し、生成物組成予測モデルにおける目的変数とした。

本研究で考慮した特徴量を表 1 に示す。基本特徴量となる反応温度  $T$ 、接触時間  $\tau$ 、各原料濃度  $X_i$  ( $i = 1 - 5$ ) に加え、原料間の相互作用を表現する濃度交差項を組み込んだ。さらに、特徴量エンジニアリングとして、温度をアレニウス型 ( $-\exp(-E_a/RT)$ ) に変換する、1 次及び 2 次反応の反応速度項を導入する、微分反応器モデルを仮定した濃度変化項を導入するなど、物理化学的な情報を取り入れた特徴量を作成した。最後に、分散が 0 となる特徴量を除外することで、55 個の特徴量からなるデータセットを作成した。また、比較のために、基本特徴量のすべての組み合わせの 2 次及び 3 次の項を追加する網羅的特徴量エンジニアリングも検討した。網羅的特徴量エンジニアリングでは 90 個の特徴量からなるデータセットを作成した。

表 1 モデル構築に用いた特徴量

基本特徴量	
温度	$T$
接触時間	$\tau$
原料濃度	$X_i (i = 1 - 5)$
追加特徴量	
濃度交差項	$X_i X_j (i, j = 1 - 5)$
アレニウス型温度	$\exp(-E_a/RT)$
反応速度項	$\exp(-E_a/RT) \cdot X_i$ $\exp(-E_a/RT) \cdot X_i X_j$
濃度変化項	$\exp(-E_a/RT) \cdot X_i \tau$ $\exp(-E_a/RT) \cdot X_i X_j \tau$

回帰モデルの構築には解釈可能性の高い線形回帰の一種である LASSO を用いた。LASSO は重要度の低い特徴量の回帰係数をゼロにすることで重要な特徴量のみを選択する機能を持つ。また、比較のため、3 種類の線形回帰（最小二乗法（OLS）、部分最小二乗法（PLS）、Ridge 回帰（RR））及び 5 種類の非線形回帰（カーネルリッジ回帰（KRR）、サポートベクター回帰（SVR）、ランダムフォレスト（RF）、Extremely Randomized Trees（ERT）、勾配ブースティング回帰木（GBRT））を用いた回帰モデルも構築した。モデルの予測精度の検証では、訓練データとテストデータの分割を変えながら 10-fold クロスバリデーションを 100 回繰り返し行った。各分割におけるテストデータの二乗平均平方根誤差（RMSE）及び決定係数（ $R^2$ ）を算出し、その平均と分散を用いて予測精度を評価した。

## 4. 研究成果

### 4.1. 物理化学に基づく特徴量エンジニアリングの効果

物理化学に基づく特徴量エンジニアリングが LASSO 回帰の予測精度に与える影響を調査した。基本特徴量のみを用いた場合と物理化学に基づく特徴量エンジニアリングを施した場合の LASSO 回帰による予測値と実測値の関係を図 2 に示す。なお、ここでは活性化エネルギー  $E_a$  として 100 kJ/mol を仮定した。基本特徴量のみを用いた場合では、実測値が変化するにもかかわらず予測値が部分的に一定となることがあり、予測精度は低かった。一方、特徴量エンジニアリングを施すことで訓練データ及びテストデータともに良く予測できることが確かめられた。この結果から、物理化学に基づく特徴量エンジニアリングにより LASSO 回帰の予測精度を向上させることが確かめられた。

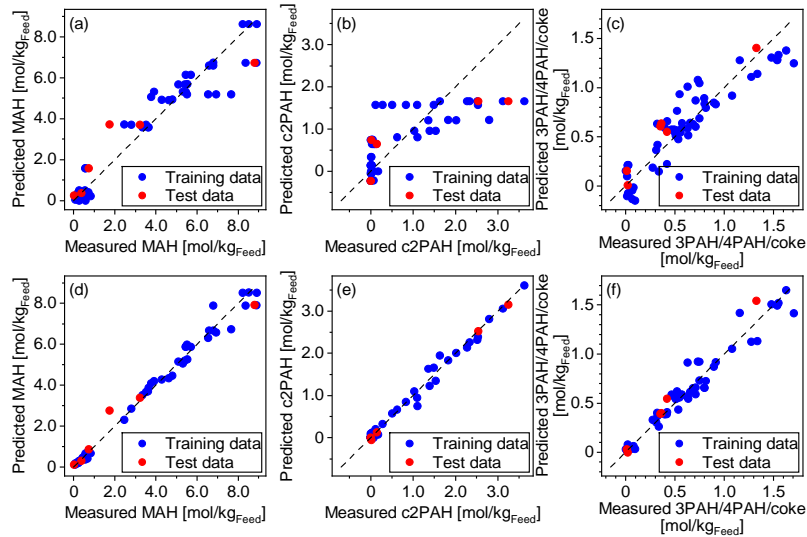


図2 異なる特徴量を用いた LASSO モデルによる生成物組成予測結果。(a)-(c)基本特徴量のみを用いた LASSO モデル；(d)-(f)物理化学に基づく特徴量エンジニアリングを施した LASSO モデル。(a)(d)MAH；(b)(e)c2PAH；(c)(f)3PAH/4PAH/coke の予測結果。

#### 4.2. 活性化エネルギー $E_a$ の影響

特徴量エンジニアリングにおける活性化エネルギー  $E_a$  の影響について調査した  $E_a$  を 10 - 150 kJ/mol の範囲で変更して特徴量エンジニアリングを行い、LASSO 回帰によるテストデータの予測精度を調べた結果を図3に示す。活性化エネルギーが変化することでなだらかに RMSE や  $R^2$  が変化し、特に  $E_a$  が 50 - 60 kJ/mol において最も RMSE が小さく、 $R^2$  が高くなった。従来の研究で、多環芳香族炭化水素の接触分解反応では水素移行反応とクラッキング反応が重要な影響を持つことが確かめられている。このうち、水素移行反応の活性化エネルギーは 50 - 70 kJ/mol 程度の報告例があり、今回予測精度が高くなった活性化エネルギーの値と一致する。この結果より、実際に進行している反応に適した活性化エネルギーの値を選択することができると予測精度が向上することが示唆された。あるいは、実験データから予測モデルを構築することで、重要な反応の活性化エネルギーを推察できる可能性も期待できる。

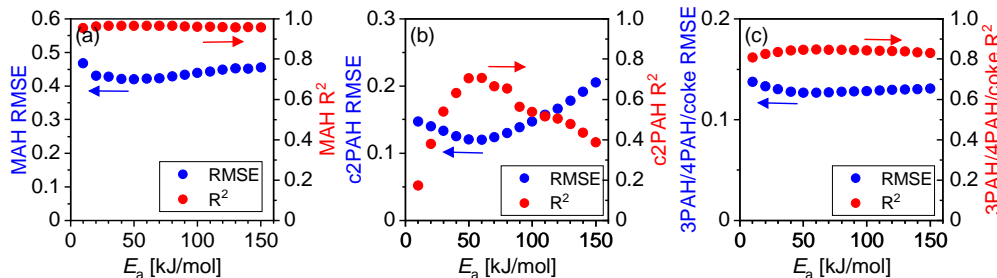


図3 異なる活性化エネルギー  $E_a$  の値を用いたときの生成物組成予測結果の RMSE と  $R^2$ 。(a) MAH；(b)c2PAH；(c)3PAH/4PAH/coke の予測結果。

#### 4.3. モデルの解釈

線形回帰モデルは解釈可能性が高い回帰手法であり、係数の大きさから重要な因子を抽出することが可能である。図4に MAH、c2PAH、3PAH/4PAH/coke の生成量を目的変数にした LASSO 回帰の主な標準回帰係数を示す。MAH の生成量に対して正の大きな係数を持つものとして  $x_4$  (トルエン濃度) や  $x_3$  (テトラリン濃度) があり、これは原料中に MAH が多くと生成物中の MAH も多くなることを示している。同様に、3PAH/4PAH/coke 生成量に対して  $x_2$  (9-メチルアントラセン濃度) が極めて大きい正の標準回帰係数を示したが、これも原料中に 3 環芳香族が多い場合に生成物中にも 3 環芳香族またはさらに重質な生成物が多く含まれることを示唆する。一方、 $\exp(-E_a/RT) * x_3$  や  $\exp(-E_a/RT) * x_3 * t$  は MAH に対して負の標準回帰係数を持つ一方で c2PAH に対しては正の標準回帰係数を持つ。これらはテトラリンの 1 次反応速度やそれによる濃度変化を表す項であり、テトラリンが脱水素してナフタレンになることで MAH が減少し c2PAH が増加することを示していると解釈できる。同様に、 $\exp(-E_a/RT) * x_5 * t$  が 3PAH/4PAH/coke に正の標準回帰係数を持つのも、 $x_5$  (9,10-ジヒドロアントラセン) の脱水素反応によって 3 環芳香族が生成することを示唆する。一方、3PAH/4PAH/coke の減少 (重質成分の分解) に重要な因子としては、 $\exp(-E_a/RT) * x_2 * x_3$  が挙げられる。これはテトラリンと 9-メチルアントラセンの間の 2 次反応で

あるが、既往研究で検討された水素移行反応による 3 環芳香族の水素化を示唆していると考えられる。このように、物理化学に基づく特徴量エンジニアリングと LASSO 回帰を組み合わせたモデルでは、標準回帰係数から重要な反応因子を抽出できる可能性が示された。

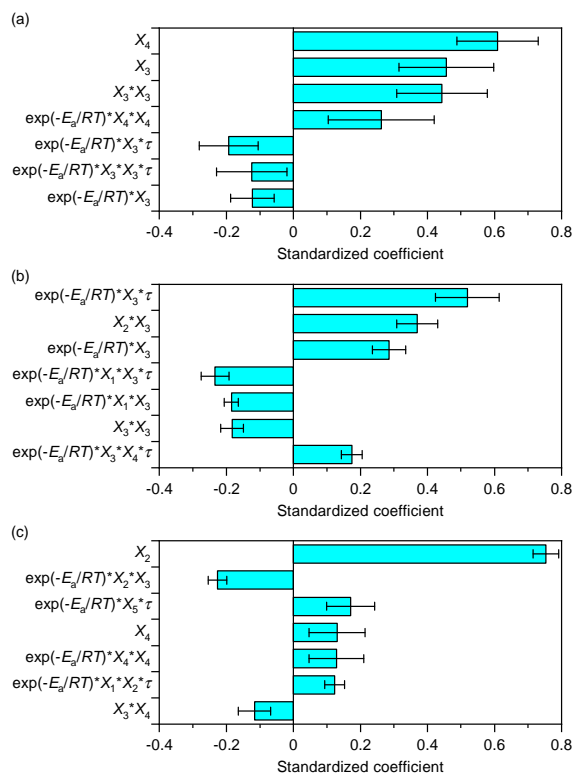


図 4 物理化学に基づく特徴量エンジニアリングを施した LASSO 回帰モデルにおける標準回帰係数。ただし、各モデルにおいて絶対値の大きい 7 個の特徴量の標準回帰係数のみを示した。(a)MAH ; (b)c2PAH ; (c)3PAH/4PAH/coke。

#### 4.4. 様々な回帰手法の比較

物理化学に基づく特徴量エンジニアリングにより LASSO 回帰の予測精度が向上することが確かめられた。そこで、その予測精度を様々な回帰モデルと比較した(表 2)。ここでは目的変数を MAH 生成量とした。基本特徴量のみを用いた場合、線形回帰モデル (OLS、PLS、RR) は予測精度が低いのにに対し、非線形回帰モデル (KRR、SVR、RF、ERT、GBRT) は比較的精度が高かった。これは、生成物生成量と基本特徴量の間非線形の関係があるためである。一方、物理化学に基づく特徴量エンジニアリングと LASSO 回帰を組み合わせたモデルが、線形回帰モデルであるにもかかわらず全ての非線形回帰モデルよりも高い予測精度と小さなばらつきを示した。以上の結果より、線形回帰の 1 種である LASSO 回帰と物理化学に基づく特徴量エンジニアリングを組み合わせることで高い予測精度と解釈可能性の両立ができる可能性を示すことができた。

表 2 異なる回帰モデルを用いた MAH 生成量予測における RMSE と  $R^2$  (平均 ± 標準偏差)

Model	RMSE	$R^2$
基本特徴量を用いた線形回帰		
OLS	0.795 ± 0.302	0.844 ± 0.565
PLS	0.930 ± 1.785	0.427 ± 5.546
RR	0.789 ± 0.316	0.853 ± 0.474
LASSO	0.796 ± 0.316	0.849 ± 0.497
基本特徴量を用いた非線形回帰		
KRR	0.426 ± 0.202	0.964 ± 0.071
SVR	0.477 ± 0.283	0.956 ± 0.064
RF	0.581 ± 0.260	0.930 ± 0.106
ERT	0.465 ± 0.247	0.955 ± 0.059
GBRT	0.470 ± 0.258	0.951 ± 0.079
特徴量エンジニアリング (FE) を施した LASSO 回帰		
Exhaustive FE	0.448 ± 0.260	0.955 ± 0.183
Physics-based FE	0.421 ± 0.217	0.966 ± 0.048

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計19件（うち招待講演 3件 / うち国際学会 2件）

1. 発表者名 安池峻, 長田光正, 福長博, 古山通久, 嶋田五百里
2. 発表標題 バイオオイルモデル物質の接触分解における機械学習を用いた生成物収率予測
3. 学会等名 日本コンピュータ化学会2021年春季年会
4. 発表年 2021年

1. 発表者名 嶋田五百里, 安池峻, 長田光正, 福長博, 古山通久
2. 発表標題 機械学習を用いた重質油及びバイオオイルモデル物質の接触分解反応の生成物組成予測
3. 学会等名 第128回触媒討論会
4. 発表年 2021年

1. 発表者名 嶋田五百里
2. 発表標題 触媒反応工学分野における機械学習と特徴量エンジニアリングの活用
3. 学会等名 令和3年度（2021年度）日本材料科学会若手研究者講演会
4. 発表年 2021年

1. 発表者名 嶋田五百里, 長田光正, 福長博, 古山通久
2. 発表標題 多環芳香族の接触分解反応における機械学習を用いた生成物組成予測
3. 学会等名 第51回石油・石油化学討論会
4. 発表年 2021年

1. 発表者名 嶋田五百里, 安池峻, 長田光正, 福長博, 古山通久
2. 発表標題 機械学習と物理に基づく特徴量エンジニアリングによる触媒反応の生成物組成予測
3. 学会等名 第31回日本MRS年次大会
4. 発表年 2021年

1. 発表者名 Iori Shimada
2. 発表標題 Prediction of product composition using machine learning in catalytic cracking of polycyclic aromatic hydrocarbons
3. 学会等名 The 20th Kuwait/Japan Joint Symposium (招待講演) (国際学会)
4. 発表年 2022年

1. 発表者名 嶋田五百里, 長田光正, 福長博, 古山通久
2. 発表標題 重質油接触分解反応の生成物組成予測における機械学習と特徴量エンジニアリングの利用
3. 学会等名 化学工学会第87年会
4. 発表年 2022年

1. 発表者名 嶋田五百里, 長田光正, 福長博, 古山通久
2. 発表標題 機械学習と特徴量エンジニアリングを用いた多環芳香族接触分解反応の生成物組成予測
3. 学会等名 第129回触媒討論会
4. 発表年 2022年

1. 発表者名 嶋田五百里, 長田光正, 福長博, 古山通久
2. 発表標題 接触分解反応の生成物組成予測における機械学習の利用
3. 学会等名 第51回化学工学会秋季大会
4. 発表年 2020年

1. 発表者名 嶋田五百里, 長田光正, 福長博, 古山通久
2. 発表標題 接触分解反応の生成物組成予測における機械学習と特徴量エンジニアリングの利用
3. 学会等名 第127回触媒討論会
4. 発表年 2021年

1. 発表者名 嶋田五百里, 長田光正, 福長博, 古山通久
2. 発表標題 機械学習と特徴量エンジニアリングを用いた接触分解反応の生成物組成予測
3. 学会等名 化学工学会第86年会
4. 発表年 2021年

1. 発表者名 安池峻, 長田光正, 福長博, 古山通久, 嶋田五百里
2. 発表標題 機械学習を用いたバイオオイルと重質油の混合接触分解の生成物収率予測
3. 学会等名 化学工学会第86年会
4. 発表年 2021年



1. 発表者名 安池峻, 嶋田五百里
2. 発表標題 重質油およびバイオオイルモデル物質の混合接触分解における機械学習を利用した生成物収率予測
3. 学会等名 第26回JPIJSポスターセッション
4. 発表年 2022年

1. 発表者名 嶋田五百里
2. 発表標題 重質油やバイオマスの接触分解反応の解析における機械学習の利用
3. 学会等名 化学工学会反応工学部会若手会（反好会）第23回講演会（招待講演）
4. 発表年 2022年

1. 発表者名 安池峻, 長田光正, 嶋田五百里
2. 発表標題 接触分解反応の生成物収率予測における物理化学に基づく特徴量エンジニアリングの効果
3. 学会等名 化学工学会第53回秋季大会
4. 発表年 2022年

1. 発表者名 嶋田五百里
2. 発表標題 接触分解プロセスにおけるバイオマス原料のコプロセッシング 実験及び機械学習を用いた反応解析
3. 学会等名 第52回石油・石油化学討論会（招待講演）
4. 発表年 2022年

1. 発表者名 Iori Shimada, Mitsumasa Osada, Hiroshi Fukunaga, Michihisa Koyama
2. 発表標題 Prediction of catalytic reaction product composition using machine learning and physics-based feature engineering
3. 学会等名 The 9th Tokyo Conference on Advanced Catalytic Science and Technology (TOCAT9) (国際学会)
4. 発表年 2022年

1. 発表者名 安池峻, 嶋田五百里
2. 発表標題 機械学習を用いた接触分解反応の生成物収率予測モデルの構築
3. 学会等名 日本コンピュータ化学会2022年秋季年会
4. 発表年 2022年

1. 発表者名 安池峻, 長田光正, 福長博, 高橋伸英, 嶋田五百里
2. 発表標題 バイオオイル/重質油混合接触分解における機械学習を用いた生成物組成予測
3. 学会等名 化学工学会第88年会
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<p>嶋田研究室ウェブサイト  <a href="http://www.iorishimada-lab.com/index.html">http://www.iorishimada-lab.com/index.html</a></p>
---

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------