

最少実験データ(≦15)による機械学習→大幅なコスト削減

1. 基礎研究: ガウス過程回帰によるフロー反応条件(2次元パラメータ)の迅速最適化

- 予測収率の可視化 -, Chem. Commun. 2020, 56, 1259

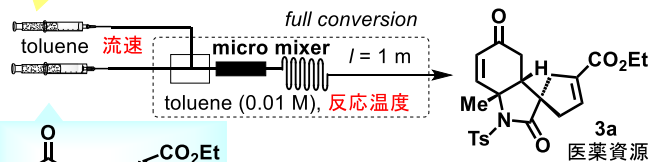
新規化学反応の開発

目標: 目的物を高収率にて合成する
⇒ 多種多様な反応パラメータの探索・最適化が必要
N種類のパラメータをM通りずつ検討する場合、網羅的な検討ではM^N個の実験が必要となることもある(時間・資源の浪費)

本研究: 最少の実験と機械学習による効率的な反応条件最適化

フロー反応

- 高い反応効率・再現性・安全性
- スケールアップが容易
- 煩雑なパラメータ設定が必要(流速・管長等)



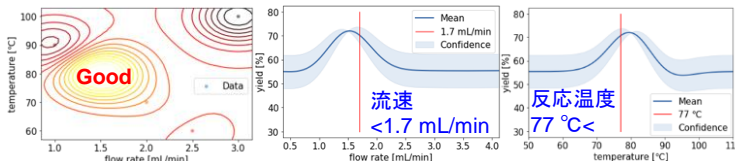
収率とeeに影響を与える3因子
①流速 ②反応温度 ③試薬2aの当量
3因子を5通りずつ網羅的に検討
= 125通りの実験が必要

最少の実験 × 機械学習 (ガウス過程回帰)

- 2つのパラメータを同時に探索
- 収率をイメージング
- ⇒ 実験数の大幅な削減

反応温度と流速の検討

Entry	流速 (mL/min)	温度 (°C)	2a (eq)	収率 (%)	Ee (%)
1	1.0	90	2.0	49	92
2	1.5	80	2.0	72	90
3	2.0	70	2.0	58	90
4	2.5	60	2.0	55	93
5	3.0	100	2.0	43	89



2aの当量・反応温度についても検討し、最終的に10例の学習データから最適反応条件を予測: 流速 = 1.7 mL/min, 反応温度 = 80 °C, 試薬2aの当量 = 2.0 eq.
⇒ 機械学習から予測した条件を適用すると、76%収率, 94% eeで3aが得られた。

⇒ 本条件最適化法をシアノヒドリンのスケールアップフロー合成に適用することにも成功 (岡大・菅教授との共同研究)

Application of an Electrochemical Microflow Reactor for Cyanosilylation: Machine Learning-assisted Exploration of Suitable Reaction Conditions for Semi-Large-Scale Synthesis

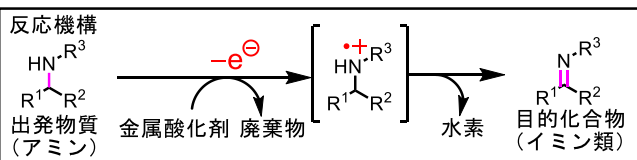
J. Org. Chem. 2021, 86, 16035

Flow Chemistry | Electro-organic Synthesis | Machine Learning

2. 基礎研究: ベイズ最適化による電解ケチミン合成反応条件(5次元パラメータ)の迅速最適化

Green Chem. 2021, 23, 5825

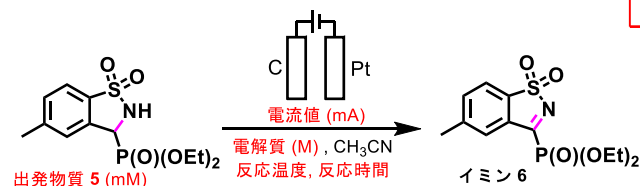
全合成や不斉合成に頻用される有用な合成中間体
イミン類
従来の合成法は金属酸化剤を用いるため、大量の金属廃棄物を副生
⇒ 金属フリーかつ高効率な合成法が求められている



金属酸化剤を用いず、クリーンな手法でアミンから電子を奪うために有機電解反応を活用

有機電解反応

- 金属・廃棄物フリー
- 高効率
- 煩雑なパラメータ設定が必要(電流値・電解質の量等)



収率に影響を与える5因子

- ①電流値 ②5の濃度 ③電解質の濃度 ④反応時間 ⑤反応温度

5因子を5通りずつ網羅的に検討
= 3125通りの実験が必要

ガウス過程回帰による5次元パラメータの収率イメージングは困難

新たなアプローチ

最少の実験 × 機械学習 (ベイズ最適化)

- 5つのパラメータを同時に探索
- 局所解に陥りにくい
- ⇒ 実験数の大幅な削減

Entry	電流値 (mA)	5 (mM)	電解質 (M)	時間 (min)	温度 (°C)	収率 (%)
1	1	0.8	0.2	150	60	4
2	2	1.2	0.6	60	20	10
3	3	0.4	0.2	240	40	43
4	3	0.8	0.4	60	60	8
5	4	1.2	0.4	150	40	67
6	5	0.4	0.6	240	20	32
7	4	1.1	0.27	150	40	76
8	5	0.4	0.2	150	35	88

電解質: LiClO₄ (過塩素酸リチウム)

収率: NMR収率

Entries 1-6: 収集した実験データ(学習データ)

Entry 7: Entries 1-6のデータに基づいて提案されたパラメータ

Entry 8: Entries 1-7のデータに基づいて提案されたパラメータ

Single EI (一つずつ実験・評価)

8例の学習データと機械学習によって

88%収率で目的物6が得られる反応条件を見出した。

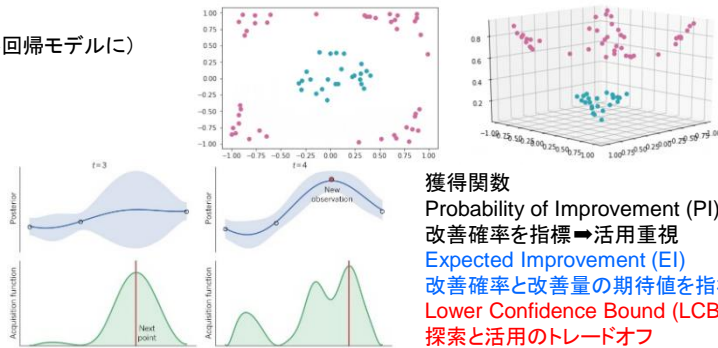
応用分野は、ライフサイエンス・創薬化学・精密有機合成反応プロセスです。 ファインケミカル関連企業様との連携・共同研究を希望しております！

ガウス過程回帰

- ・非線形モデル化でき高汎用性(右上段図:カーネルトリックにて非線形回帰モデルに)
- ・目的変数の推定値だけでなく分散も計算(右中段図:データが存在する場所は正確に、足りない場所は曖昧に)
- ・関数 $y=f(x)$ が曖昧な状態でも最大値・最小値を推測(データ数減)

ベイズ最適化

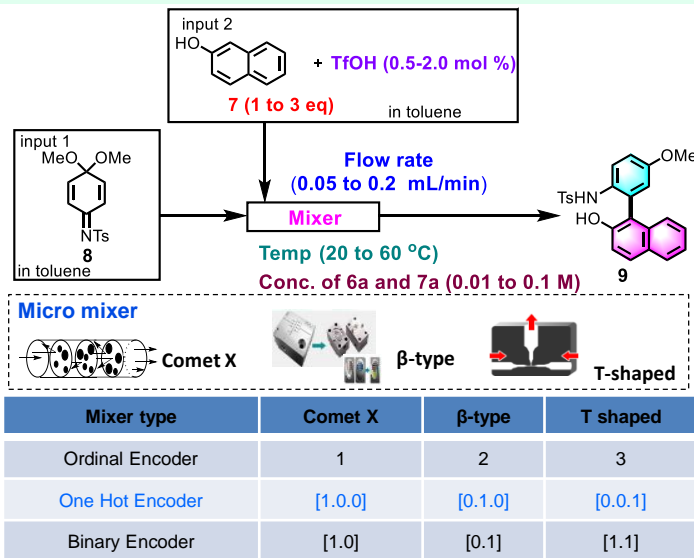
- ・不確かさを用いながら次に探索すべき点を決定するブラックボックス最適化アルゴリズム。
- ・獲得関数が最大になる点を探索(右下段図)
- ・真の関数を近似する獲得関数の構成に確率過程、特にガウス過程を用いる
- ・期待値と分散から計算される獲得関数($f(x)$ 由来の近似関数)を利用(期待値が最大を示す近傍のパラメータ値だけでなく(=活用)、分散が大きく最大値になり得るパラメータ値も探す(=探索))



獲得関数
 Probability of Improvement (PI)
 改善確率を指標 → 活用重視
 Expected Improvement (EI)
 改善確率と改善量の期待値を指標
 Lower Confidence Bound (LCB)
 探索と活用のトレードオフ

3. 基礎研究：並列ベイズ最適化によるビアリアル化合物のフロー合成反応条件

(ミキサー選択を含む6次元パラメータ)の迅速最適化 *Commun. Chem.* 2022, 5, 148



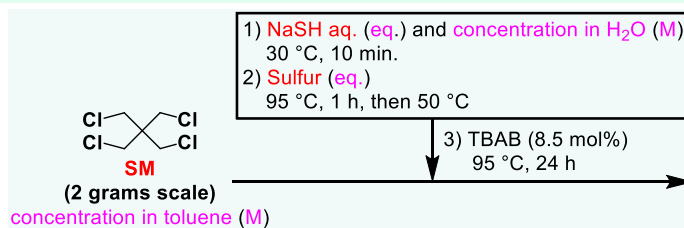
Entry	Micro mixer	7 (equiv.)	Temp. (°C)	8 in toluene [conc. (M)]	Flow rate (mL/min)	TfOH (mol %)	Yield of 9 (NMR %)
1	Comet X	2	60	0.05	0.05	1	68
2	Comet X	3	40	0.01	0.2	0.5	73
3	β-type	1	60	0.01	0.1	2	42
4	β-type	3	20	0.1	0.1	0.5	28
5	T-shaped	1	40	0.05	0.05	2	55
6	T-shaped	2	20	0.1	0.2	1	75
Prediction with ITS (entries 1 to 6)							experimental results
7	T-shaped	2	20	0.1	0.15	1	81
8	Comet X	2.3	55	0.039	0.04	1	77
9	β-type	1.1	85	0.15	0.1	2.4	40
Prediction with TS (entries 1 to 9)							experimental results
10	T-shaped	1.3	15	0.11	0.15	1.2	78
11	T-shaped	2.1	30	0.061	0.15	1.1	76
12	Comet X	2.8	50	0.014	0.11	1	79
Prediction with TS (entries 1 to 12)							experimental results
13	Comet X	3.4	55	0.01	0.032	1.3	88
14	Comet X	2.2	15	0.1	0.14	1.7	44
15	Comet X	3.0	25	0.015	0.08	1.5	96

ITS: Initial Training Set; TfOH: Trifluoromethanesulfonic acid

Parallel LCB (三つずつ実験・評価) ⇒ 時間の節約
 15例の学習データと機械学習によって
 96%収率で目的物8が得られる反応条件を見出した。

4. 社会実装研究：ベイズ最適化の化成品バッチ合成への応用

(旭化学工業株式会社との共同研究) 論文投稿中



Entry	NaSH aq. (eq.)	NaSH aq. concentration in H ₂ O (M)	Sulfur (eq.)	SM concentration in toluene (M)	Yield (GC %)
1	6.5	14.8	2.4	1.19	65
2	4.5	7.4	2.4	2.78	51
3	5.5	14.8	3.0	0.76	39
4	5.5	7.4	1.2	1.19	76
5	6.5	9.8	2.0	2.78	66
6	4.5	9.8	1.2	0.76	52
Prediction with ITS (entries 1 to 6)					experimental results
7	5.8	8.1	1.4	1.22	78
Champion data in company (none-BO screening): 2 grams scale					
	6.0	11.9	3.0	1.71	68

Entry	NaSH aq. (eq.)	NaSH aq. concentration in H ₂ O (M)	Sulfur (eq.)	SM concentration in toluene (M)	Yield (GC %)
1	5.5	5.3	2.1	4.54	67
2	2.7	7.0	3.3	2.16	31
3	6.9	18.4	2.7	1.42	68
4	8.3	4.3	0.9	0.84	48
5	4.1	10.0	1.5	1.06	66
Prediction with ITS (entries 1 to 5)					experimental results
6	6.7	16.2	2.7	1.50	70
Prediction with ITS (entries 1 to 6)					experimental results
7	6.5	29.8	1.3	4.54	89 (87)
ITS prepared by random sampling					20 grams scale
ITS prepared by Latin hypercube sampling (LHS)					