

# Aggregated Outputs を活用した回帰モデルによるベイズ最適化

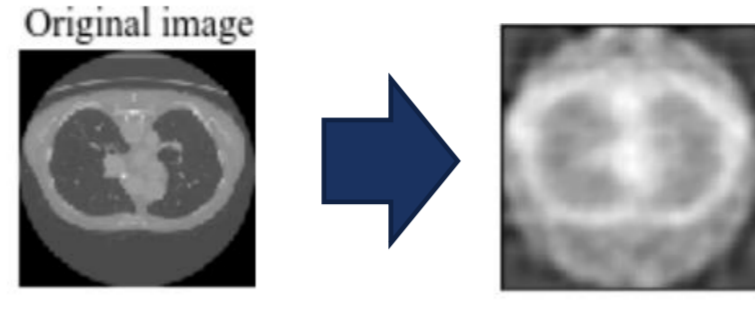
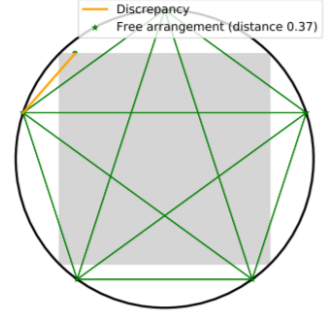
松尾 裕樹 総合研究大学院大学統計科学専攻 博士課程5年

## Aggregated Outputs (集約的観測) とは

入力の集合に対する出力 ( $\{x_i^n\}, y^n$ ) からモデルを構築し、1つの入力( $x$ )に対する出力を予測する方法

例 1. 積分観測値によるCT画像の構築

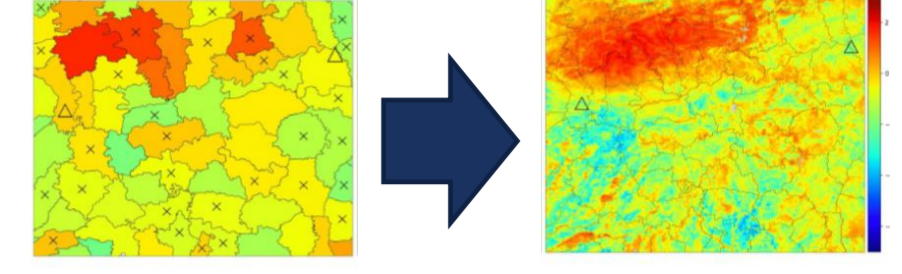
入力: 2つのセンサー間の積分値  
→予測: CT画像



Krause, A., Singh, A., Guestrin, C., 2008. Near-optimal sensor placements in Gaussian processes: theory, efficient algorithms and empirical studies. J. Mach. Learn. Res. 9,235–284

例 2. 地域感染率に基づく各地点の感染率予測

入力: 地域毎に集約して感染率  
→予測: 各地点の感染率



Leon Ho Chung Law, Dino Sejdinovic, Ewan Cameron, Tim C.D. Lucas, Seth Flaxman, Katherine Battle, and Kenji Fukumizu. Variational learning on aggregate outputs with Gaussian processes. In Advances in Neural Information Processing Systems, 2018.

### 【集約的観測の定式化】

$N^i$ 個の入力を含む集合が $M$ セット存在するとき

$$\{ \{x_j^i | j = 1, \dots, N^i\} | i = 1, \dots, M \}$$

集約的観測  $y^i$ は  $\{x_j^i\}_{j=1}^{N^i}$ 、重み $W^i$ から目的関数  $f(x)$  に対し次式となる

$$y^i \sim p(y|g^i), \quad g^i = \sum_j f(x_j^i) W_j^i$$

### 集約的観測の利点

⇒ 一括した観測による評価コスト低減

集約的観測を最適化に活用した先行研究例はなく、最適化への応用を検討

#### 適用対象の要件

集約する条件の範囲( $\{x_j\}, W$ )を指定可能  
集約的観測が可能  
観測のコストが入力を用意するコストより大きい

#### 適用候補の系

- 相溶系のポリマーブレンド
- 混合物のスペクトル吸光度
- 複合材料の機械的物性
- MDシミュレーション

最も単純な集約的観測である  
加成性が成り立つことが期待

## 集約的観測によるベイズ最適化

### 【ベイズ最適化の獲得関数】

#### Max-value Entropy Search (MES)

単一の観測に対するベイズ最適化の獲得関数

$x$ における  $\hat{y}$  の予測分布と  $\hat{y}$  の最大値  $y_*$  の相互情報量

$$\alpha_{MES}(x) = I(\hat{y}(x); y_*)$$

Zi Wang and Stefanie Jegelka. Max-value entropy search for efficient bayesian optimization. In Proceedings of International Conference on Machine Learning (ICML), 2017.

#### Aggregated Outputs Max-value Entropy Search (AOMES)

集約的観測に対応出来るように拡張した獲得関数

$X = (\{x_j | j = 1, \dots, N\}, W)$  における集約的予測 ( $\hat{y}_{agg}$ ) の予測分布と  $\hat{y}$  の最大値  $y_*$  の相互情報量

$$\alpha_{AOMES}(X) = I(\hat{y}_{agg}(X); y_*)$$

$$= \iint p(\hat{y}_{agg}(X), y_*) \log \frac{p(\hat{y}_{agg}(X), y_*)}{p(\hat{y}_{agg}(X))p(y_*)} d\hat{y}_{agg} dy_*$$

### 【ベンチマーク関数による数値実験】

#### 予測モデル

ガウス過程回帰を使用

事前分布

$$\hat{f}(x) \sim GP(m(x), k_{rbf}(x, x') | \sigma_s, l)$$

尤度関数

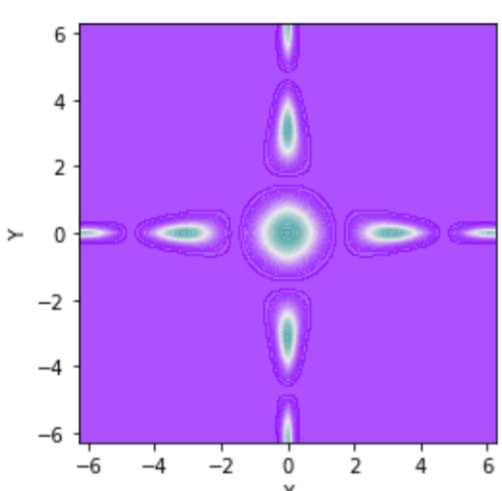
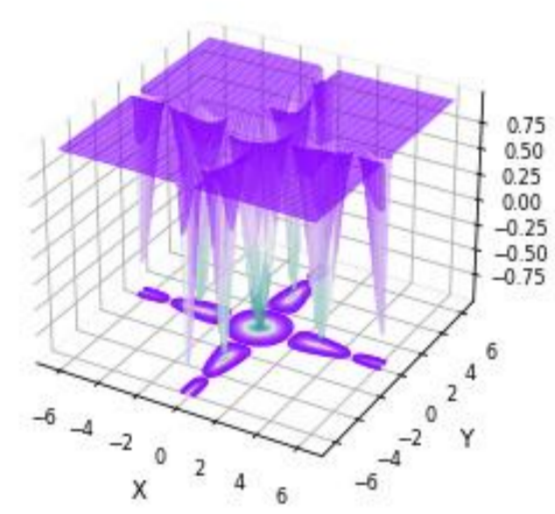
$$y^i \sim N(\sum_j \hat{f}(x_j^i) W_j^i, \sigma_n)$$

#### 検証手続き

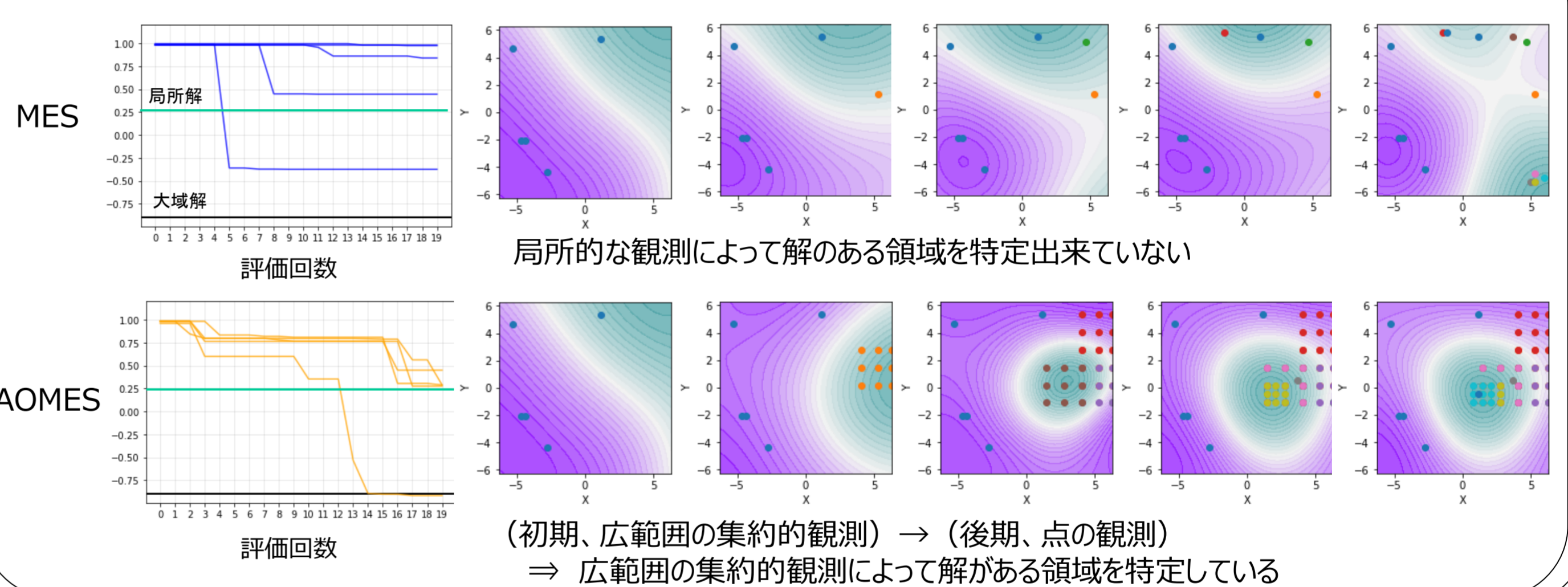
MES, AOMES で下記手続きを実施し比較

- ランダムに初期点を5点生成
- ハイパーパラメータ ( $l, \sigma_s, \sigma_n$ ) を最適化
- 獲得関数を最大化する候補条件を選択
- 選択条件を評価
- 2~4を既定回数繰り返す

#### ベンチマーク関数



#### 結果



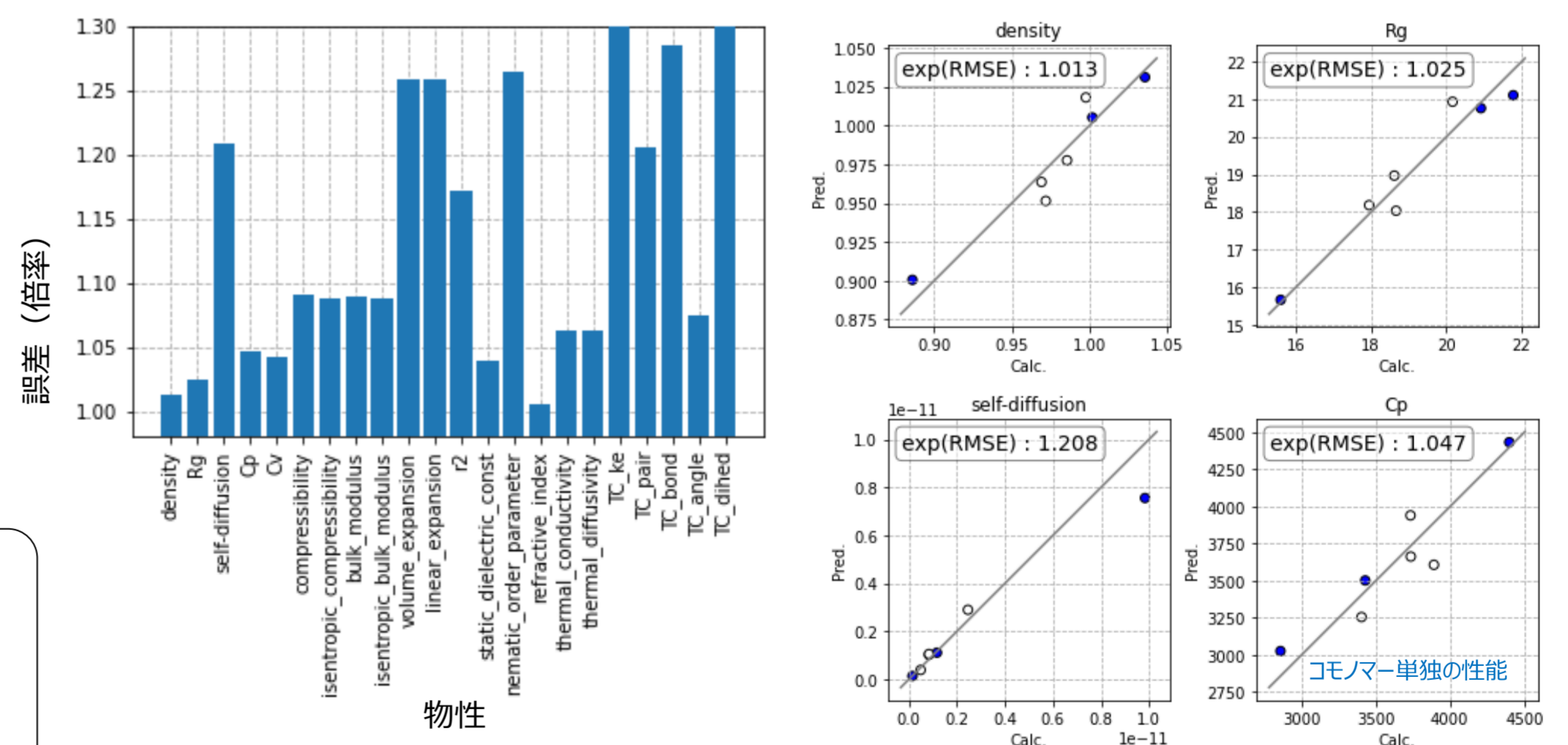
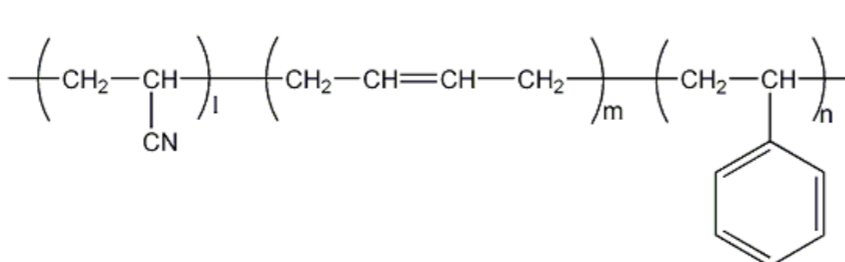
## ポリマーMDシミュレーションへの適用可能性検討

適用対象の要件	MDシミュレーション
集約する条件の範囲を指定可能	ポリマー構造を指定可能
集約的観測が可能	物性で加成性が成り立つ可能性有り
観測のコストが入力を用意するコストより大きい	構造の作製よりも観測 (シミュレーション) のコストが大きい

「MDシミュレーションによるモノマー探索」に対して集約的観測を使った性能最適化は可能か?  
→ ABS樹脂系にて加成性を検証し、複数の物性に対して適用可能性があることを確認 (RadonPyでMDシミュレーションを実施)

Y. Hayashi, J. Shiomi, J. Morikawa, R. Yoshida, "RadonPy: Automated Physical Property Calculation using All-atom Classical Molecular Dynamics Simulations for Polymer Informatics," npj Comput. Mater. 8:222 (2022)

#### ABS樹脂



## まとめ

- 効率的な最適化のために集約的観測を活用したベイズ最適化の指標である Aggregated Outputs Max-value Entropy Search (AOMES) を設計した。
- 数値実験による比較で、AOMES が解の探索に優れ得ることを確認した。
- MDシミュレーションによるモノマー探索において集約的観測が可能なることを確認した。

## 今後の予定

- 多峰性を持つ関数等、複数のベンチマーク関数でも検証を実施する。
- MDシミュレーションによるモノマー探索に対してAOMESによる最適化を試みる。
- AOMES 最適化時、効率的に  $(\{x_j\}, W)$  を求める方法を開発する。