

多出力ガウス過程回帰によるベイズ最適化を用いた粉体製造における 粒子径分布の適応的実験計画

北村 智浩^{1*}, 今井 貴史², 河本 薫²

Bayesian Optimization-based Experimental Design with Multi-output Gaussian Process Regression to Obtain Specific Powder Particle Size Distributions

Tomohiro Kitamura^{1*}, Takashi Imai² and Kaoru Kawamoto²

Received 14 May 2025; Accepted 28 July 2025

This paper proposes a Bayesian optimization-based method for optimizing operating parameters in powder processing to achieve desired particle size distributions with a few experimental trials. Such optimization has become increasingly important in recent years due to the trend toward strict quality requirements. A difficulty in applying Bayesian optimization to this task is that the objective variables (i.e., the particle size distributions) are probability distributions, which are required to satisfy the normalization condition. We overcome this difficulty by predicting percentiles using multi-output Gaussian process surrogate models. Experiments on actual equipment showed that the proposed method can provide sufficiently optimized parameter values without either expert knowledge or detailed process modeling. In addition, this paper examines the potential of grey-box modeling for overcoming the limitations of our machine learning approach in extrapolation.

Keywords: Particle size distribution, Powder processing, Design of experiments, Gaussian process, Bayesian optimization.

1. 緒 言

粉体製造における重要な品質指標の一つは、粉体の粒子径分布である。一般に、製造された粉体の粒子径は均一ではなく幅広い分布を持つ。これまでは、品質管理指標には、たとえば粒子径の平均値や中央値 (d_{50}) などの統計量が用いられてきた[1–3]。しかしながら、近年は、粉体製造の品質要求が厳しくなり、製造する粉体の粒子径を目標の粒子径分布 (particle size distribution, PSD) そのものに合わせる事が求められている[4–6]。本論文では、目標の粒子径分布を目標分布と呼ぶ。製造現場では、このような品質要求を満たす製造条件の探索が必要である。

粉体製造における粉体挙動には未知な部分が多い。計算流体力学[7]や離散要素法[8,9]などの粉体シミュレー

ションでは、現実の現象をすべて再現できる技術レベルには至っていない。また、個々の製造装置に対応したエンジニアリングモデルを作ることは、大量の実験を必要とする。たとえば、population balance modeling (PBM) [10,11]をベースにしたエンジニアリングモデルの開発には、製造装置ごとに正確な定式化が必要であり、そのための大量の基礎実験が必要になり現実的ではない。

それゆえに、製造現場では、エンジニアの経験と知識に頼りながら、目標分布を満たす製造条件を探索してきた。しかし、その探索には多くの実験回数が必要となり、その結果、開発期間が長くなり開発コストが増える。また、このような製造条件の探索は熟練したエンジニアに依存しており、その技術伝承も課題である。これらの課題を解決するため、エンジニアの経験に頼らずに、少ない実験回数で目標分布を満たす製造条件を見出す方法が期待される。

条件探索の実験回数を減らす方法として、適応的実験計画法 (adaptive design of experiments, ADoE) が注目されている。従来の実験計画ではあらかじめ計画した実験スケジュールに従って実験するのに対し、ADoEでは実験で得られたデータを活かして次の実験条件を探索する。ADoEを粉体製造に応用した先行研究[12]では、PBMの予測残差をpartial least squares regression (PLS)で予測するモデルを構築し、ADoEによって目標分布を満たす製

粉体工学会 2024 年度春期研究発表会, 2024 年 5 月 14 日, 兵庫県姫路市, BP 賞受賞

1 ホソカワミクロン株式会社

(〒 573-1132 大阪府枚方市招堤田近 1-9)

Hosokawa Micron Corporation

(1-9 Shodaitajika, Hirakata, Osaka 573-1132, Japan)

2 滋賀大学データサイエンス学部

(〒 522-8522 滋賀県彦根市馬場 1-1-1)

Department of Data Science, Shiga University

(1-1-1 Bamba, Hikone, Shiga 522-8522, Japan)

* Corresponding Author tkitamura@hmc.hosokawa.com

造条件を探索している。しかしながら、この方法では、製造装置をPBMとして定式化するための事前の実験と検証が必要であり、さらにPLSを用いているため非線形な現象に対する予測精度が悪くなる課題がある。

これらの解決に有望なADoEとしてベイズ最適化がある[13]。ベイズ最適化とは、目的変数と説明変数の関数関係を実験で得られたデータを用いて統計モデルで推定し、その推定した関数と最適化のための関数（獲得関数と呼ばれる）を用いて次の実験条件を探索する方法である。代表的な統計モデルとしてガウス過程が用いられる。ガウス過程に基づくガウス過程回帰（Gaussian process regression, GPR）はノンパラメトリックかつ非線形な回帰モデルであり、目的変数の事後分布（予測分布）を導出することができる。ベイズ最適化では、ガウス過程回帰によって導出される予測分布を使って、希望する目的変数値が得られると期待できる説明変数値を提案する。ガウス過程回帰によるベイズ最適化は、マテリアル科学[14,15]、システム制御[16]、数値シミュレーション[17]、ロボット工学[18]など幅広い分野で応用されている。しかしながら、粒子径分布を目的変数とした粉体製造条件の最適化に対して、ガウス過程回帰によるベイズ最適化を応用した研究例は見られない。

本研究では、所与の粒子径分布を目標とする粉体製造において、目標達成できる製造条件を探索する際に、ガウス過程回帰に基づくベイズ最適化を応用して効率的に探索する手法を提案する。一般的に用いられているガウス過程回帰に基づくベイズ最適化をそのまま用いることは、以下の3点で困難である。1つ目は、粒子径分布という確率分布をガウス過程回帰の目的変数として扱わなければならない点である。2つ目は、ガウス過程回帰は外挿領域において予測精度が低くなる点である。3つ目は、製造装置の運転制約の中には、事前にはわからず、実験して初めてわかる運転制約がある点である。本研究では、これらの3つの困難を克服して、実用的な手法を提案する。

そして、提案手法について、目標粒子径分布を得る製造条件に辿りつけるのか、また、エンジニアによる経験知に基づいた製造条件の探索と比べて、目標粒子径分布を得るまでの実験回数を低減できるのかについて、数値実験および実機の実験により実証を行う。

なお、粉体製造装置には、さまざまな種類がある。本研究では、製造できる粒子径範囲が広く、また粉碎と同時に分級も可能であり、幅広い分野での粉体製造に使われている air classifier mill（以下、ACM）を対象とした。ACMは、製造条件を探索するための実験において多くの粉体原料を消費し、また実験時間も長時間を要するため、実験回数の低減は、製造業者にとって重要な課題となっている。

2. 回帰モデルと最適化モデルの設計

2.1 ガウス過程回帰によるベイズ最適化

まず、一般的なガウス過程回帰によるベイズ最適化を

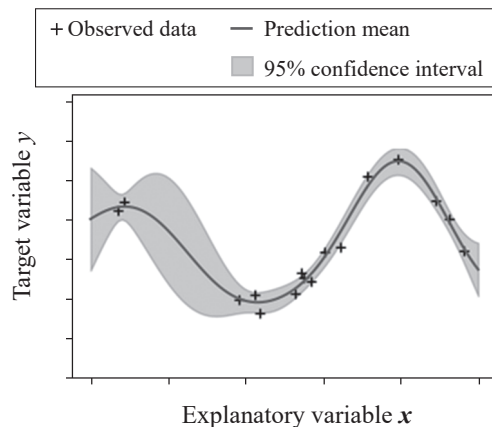


Fig. 1 The posterior predictive distribution of the target variable y obtained via Gaussian process regression. The observed data $\{x_i, y_i | i = 1, \dots, 16\}$ indicated by the crosses are given. The RBF kernel was used as the kernel function for nonlinear regression. Line represents the expected value of y , and gray area represents the 95% confidence interval calculated from the posterior mean and the posterior variance

説明する。ガウス過程回帰は、説明変数 $x \in \mathbb{R}^D$ と目的変数 $y \in \mathbb{R}$ の関係を関数 $f_{\text{GPR}}: \mathbb{R}^D \rightarrow \mathbb{R}$ と観測ノイズ $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ を用いて、

$$y = f_{\text{GPR}}(x) + \varepsilon \quad (1)$$

と表す回帰モデルである[13,19]。ここで、関数 f_{GPR} は

$$f_{\text{GPR}} \sim \mathcal{GP}(0, k(\cdot, \cdot)) \quad (2)$$

のように、平均関数を0とし、共分散関数をカーネル関数 $k(\cdot, \cdot)$ としたガウス過程に従う関数である。カーネル関数 $k(\cdot, \cdot)$ に放射基底関数（radial basis function, RBF）カーネルを用いると、ガウス過程回帰は非線形かつノンパラメトリックな回帰を行うことが可能になる[20]。RBFカーネルを用いたガウス過程回帰の例を Fig. 1 に示す。この例では、観測データ $\{x_i, y_i | i = 1, \dots, 16\}$ が与えられると、ガウス過程回帰が任意の説明変数 x に対応する目的変数 y の事後分布を導出している。図中の灰色の帯が、事後分布の95%信用区間である。事後分布の分散は、観測データに近いほど小さくなり、観測データが疎な領域では大きくなる。このように、ガウス過程回帰から得られる事後分布は、任意の説明変数 x に対する目的変数 y の不確実性を数値化している。

ガウス過程をベースにしたベイズ最適化では、ガウス過程によって得られる事後分布を用いて、探索と活用をうまく両立させる獲得関数を定義する。そして、獲得関数を最小化（あるいは最大化）する説明変数 x を求め、その説明変数 x を次の製造条件にする。

2.2 多出力ガウス過程回帰モデル

本研究では、粒子径分布の最適化を目標とするため、確率分布を目的変数とするガウス過程回帰を設計しなければならない。最初に考えられる方法は、確率分布を離散確率分布として近似し、粒子径ごとの確率質量を目的変数にする方法である。しかしながら、この方法は、全

確率質量の総和は1であるという制約の元で回帰問題を解く困難を伴う。この規格化条件を回避するため、本研究では、粒子径分布のパーセンタイル値を目的変数とする。具体的には、0パーセンタイルから10パーセンタイル刻みで100パーセンタイルまで、計11個のパーセンタイルを目的変数とした。 $m\%$ パーセンタイルに対応する粒子径を d_m と表記する。目的変数 \mathbf{y} はEq. (3)に示す通り、11個のパーセンタイルを成分とするベクトルになる。本論文では、このベクトル \mathbf{y} をパーセンタイルベクトルと呼ぶ。

$$\mathbf{y} = (d_0, d_{10}, d_{20}, \dots, d_{80}, d_{90}, d_{100})^\top \quad (3)$$

なお、各パーセンタイル d_m は正の値でなければならないため、計算時には各 d_m を対数化して用いる。

パーセンタイルベクトル \mathbf{y} を目的変数とするガウス過程回帰を実現する単純な方法は、 \mathbf{y} の成分ごとに(すなわちパーセンタイルごとに)ガウス過程回帰モデルを設けることである。しかし、この方法は、パーセンタイル間の相関関係を考慮できないため、表現力の面で不十分である。そこで本研究では、ベクトル成分間の相関関係も取り込める多出力ガウス過程回帰(multi-output Gaussian process regression, MOGPR) [21]の一種であるintrinsic coregionalization model (ICM) [22]を用いた。MOGPRでは、説明変数 $\mathbf{x} \in \mathbb{R}^D$ と目的変数 $\mathbf{y} \in \mathbb{R}^L$ の関係を関数 $f_{\text{MOGPR}}: \mathbb{R}^D \rightarrow \mathbb{R}^L$ と観測ノイズ $\boldsymbol{\varepsilon} \in \mathbb{R}^L$ ($\varepsilon \sim \text{i.i.d. } \mathcal{N}(0, \sigma^2)$)を用いて

$$\mathbf{y} = f_{\text{MOGPR}}(\mathbf{x}) + \boldsymbol{\varepsilon} \quad (4)$$

として表す。ここで、関数 f_{MOGPR} は

$$f_{\text{MOGPR}} \sim \mathcal{GP}(\mathbf{0}, \mathbf{B} \otimes k(\cdot, \cdot)) \quad (5)$$

のように、平均関数を $\mathbf{0}$ (各要素が0のベクトル)とし、共分散関数を行列 $\mathbf{B} \in \mathbb{R}^{L \times L}$ とカーネル関数 $k(\cdot, \cdot)$ のテンソル積

$$\mathbf{B} \otimes k(\cdot, \cdot) = \begin{pmatrix} b_{1,1}k(\cdot, \cdot) & \dots & b_{1,L}k(\cdot, \cdot) \\ \vdots & \ddots & \vdots \\ b_{L,1}k(\cdot, \cdot) & \dots & b_{L,L}k(\cdot, \cdot) \end{pmatrix} \quad (6)$$

とするガウス過程に従う関数である。行列 \mathbf{B} は目的変数 \mathbf{y} の分散共分散行列であり目的変数 \mathbf{y} の各要素間の相関情報を持つ。この行列 \mathbf{B} のすべての要素を学習データで推定しなければならないが、パラメータ数は L^2 個と多い。そこで、ICMでは、行列 $\mathbf{B} \in \mathbb{R}^{L \times L}$ をベクトル $\mathbf{w} \in \mathbb{R}^L$ 同士の直積 $\mathbf{w}\mathbf{w}^\top \in \mathbb{R}^{L \times L}$ とベクトル $\boldsymbol{\kappa} \in \mathbb{R}^L$ を対角成分に持つ対角行列の和

$$\mathbf{B} = \mathbf{w}\mathbf{w}^\top + \text{diag}_{\text{vector} \rightarrow \text{matrix}}(\boldsymbol{\kappa}) \quad (7)$$

として近似的に表現する。この表現により、行列 \mathbf{B} 内のパラメータ数は L^2 個から $2L$ 個へ減少する。

カーネル関数 $k(\cdot, \cdot)$ には、RBFカーネルとLinearカーネルを組み合わせた

$$k(\mathbf{x}, \mathbf{x}') = \theta_{\text{Constant}}^2 + \theta_{\text{Linear}}^2 \mathbf{x}^\top \mathbf{x}' + \theta_{\text{RBF}}^2 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|_2^2}{l_{\text{RBF}}^2}\right) \quad (8)$$

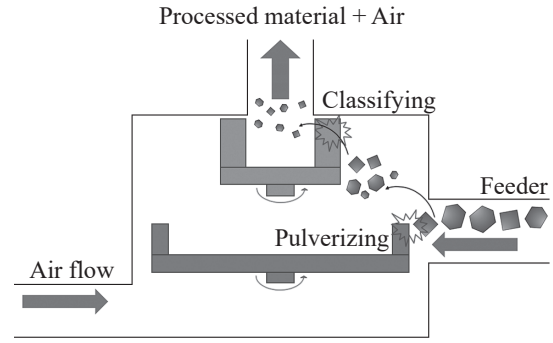


Fig. 2 An overview diagram of ACM. The pulverizing rotor pulverizes the feed material, and the fine particles pass through the classifier rotor

を用いる。このカーネル関数により、線形と非線形を組み合わせたモデル化を可能にする。以下では、Eq. (8)のカーネル関数を用いたICMのことをMOGPRモデルと呼ぶことにする。

以上により、MOGPRモデルのモデル学習時に推定すべきパラメータは、

$$\boldsymbol{\theta} = \{\mathbf{w}, \boldsymbol{\kappa}, \sigma^2, \theta_{\text{Constant}}^2, \theta_{\text{Linear}}^2, \theta_{\text{RBF}}^2, l_{\text{RBF}}^2\} \quad (9)$$

となる。これらのパラメータは、実験で得られた n セットの観測データ $\{\mathbf{X}, \mathbf{Y}\}$ 、ここで

$$\mathbf{X} = \{\mathbf{x}_i | i = 1, \dots, n\} \quad (10)$$

$$\mathbf{Y} = \{\mathbf{y}_i | i = 1, \dots, n\} \quad (11)$$

を用いて最尤推定[23]にて求める。そして、推定したパラメータを持つMOGPRモデルに基づき、観測データ $\{\mathbf{X}, \mathbf{Y}\}$ が得られたもとの、任意の \mathbf{x} に対する \mathbf{y} (本研究ではパーセンタイルベクトル)の予測分布を出力する。

2.3 ガウス過程回帰に物理モデルを組み込む

機械学習の外挿領域における予測性能の向上のために、機械学習に物理モデルによるガイドを組み込む手法が提案されている[24]。本研究においても、MOGPRモデルに粉体製造の物理モデルを組み込んだモデルを設計する。

本研究の対象とするACMの機械構造をFig. 2に示す。ACMでは、原料粉が粉砕部に供給されて粉砕され、その後、分級部で微粉と粗粉に分けられる。気流により、微粉は製品として回収され、粗粉は粉砕部で再度粉砕される。これらは、連続的かつ定期的処理される。ACMの可変な製造条件は、原料粉の供給速度 $F[\text{kg/h}]$ 、粉砕ロータの回転速度 $R[\text{rpm}]$ 、分級ロータの回転速度 $N[\text{rpm}]$ 、および空気の通過速度 $Q[\text{m}^3/\text{min}]$ の4つである。すなわち、説明変数 \mathbf{x} は

$$\mathbf{x} = (F, R, N, Q)^\top \quad (12)$$

となる。なお、これらの製造条件パラメータは正の値でなければならないため、計算時には各製造条件パラメータを対数化して用いる。また、MOGPRモデルでは、RBFカーネルを利用しており製造条件パラメータ間でスケールを合わせる必要があるため、対数化を行う前に各パラ

メータを可変範囲で正規化する。

ACMの分級工程は、ストークスの式に基づいた

$$\mathbf{y}_{\text{Physical}} = cQ^{0.5}N^{-1.0} \quad (13)$$

で表すことができる[25]。ここで、 $\mathbf{y}_{\text{Physical}}$ はパーセンタイルベクトルを表しており、Eq. (3)における \mathbf{y} と同義である。 c は、原料粉の物理特性や粉体製造装置の機械的特徴により決まるパラメータベクトルである。本論文では、Eq. (13)をPhysicalモデルと呼ぶ。

本研究では、このPhysicalモデルとMOGPRモデルを並列に組み合わせたモデルを採用する。すなわち、Physicalモデル（関数 f_{Physical} ）は説明変数 \mathbf{x} が与えられると目的変数の予測 $\mathbf{y}_{\text{Physical}}$ を出力し、MOGPRモデル（関数 f_{MOGPR} ）は実データ \mathbf{y} と予測 $\mathbf{y}_{\text{Physical}}$ の残差 $\mathbf{y}_{\text{Residual}} = \mathbf{y} - \mathbf{y}_{\text{Physical}}$ を予測する。以下では、このモデルをPhysical+MOGPRモデルと呼ぶ。

2.4 評価指標の定義

目標とする粒子径分布とPhysical+MOGPRモデルで予測した粒子径分布の差異を評価する指標を定義しよう。単純な指標としては、パーセンタイルベクトルの差のノルムが考えられる。しかし、この評価指標は、変動幅の大きい d_{90} や d_{100} といった上位のパーセンタイルに大きく影響されてしまう。この問題を避けるために、2つの確率分布の差異を測るカルバック・ライブラー情報量（以下、KL情報量と呼ぶ）[26]を評価指標として用いる。KL情報量は、Eq. (14)のとおり定義される。

$$D_{\text{KL}}(p||q) = \int_{z \in \mathbb{R}_{>0}} p(z) \log \frac{p(z)}{q(z)} dz \quad (14)$$

ここで z は任意の正の実数で、 p と q は連続確率分布であり、 $p(\cdot)$ と $q(\cdot)$ は p と q の確率密度関数である。KL情報量は、統計学の分野でよく用いられる評価指標である。KL情報量は、非負であり2つの確率分布が一致すれば0になる。

それぞれの確率分布について、共通の離散値を用いて離散確率分布で近似表現するならば、KL情報量はEq. (15)のように表現される。

$$D_{\text{KL}}(t||s) = \sum_{j=1}^J (t_j (\log t_j - \log s_j)) \quad (15)$$

ここで $t \in \mathbb{R}^J$ は目標分布であり、 $s \in \mathbb{R}^J$ は前項で提案したモデルが予測する粒子径分布である。 J は粒子径分布を離散確率分布で近似する場合の離散値の数であり、 t_j と s_j はそれぞれ目標分布と粒子径分布の j 番目の離散値の確率質量を表す。

モデルの出力はパーセンタイルベクトルであるため、Eq. (15)に従ってKL情報量を計算するには、パーセンタイルベクトルから粒子径分布を導出する必要がある。そこで、まずパーセンタイルベクトルを線形補間することで累積確率分布を導出し、次にそれを離散確率分布に変換した。

なお、前項で提案したモデルはパーセンタイルベクトルの確率分布を出力するものであるから、それに基づき計算するKL情報量もまた確率分布として得られることになる。

2.5 獲得関数の設計

ここまで述べてきた提案手法を整理すると、2.3項で設計したPhysical+MOGPRモデルを用いて、任意の製造条件 (\mathbf{x}) におけるパーセンタイルベクトル (\mathbf{y}) の予測分布を求め、次にそれを2.4項で定義した評価指標(KL情報量)の確率分布に変換し、最後にその確率分布をもとに次に実験する製造条件 (\mathbf{x}) を決定する、という流れである。

ベイズ最適化では、次に実験する条件を比較選択するために、獲得関数と呼ばれる関数を導入する。獲得関数は、製造条件 (\mathbf{x}) を引数とし、その製造条件の「不適合」を返す関数である。

獲得関数は評価指標の確率分布をもとに定義され、これまでの実験で得られた最小値を更新する確率を返す関数(probability improvement)や最小値を更新する幅の期待値を返す関数(expected improvement)など、さまざまな関数が提案されている。本研究では、評価指標の確率分布の期待値と標準偏差を用いてEq. (16)のとおり定義されるlower confidence bound (LCB) [27]を採用する。

$$LCB(\mathbf{x}) = \mu(\mathbf{x}) - \beta \cdot \sigma(\mathbf{x}) \quad (16)$$

ここで、 $\mu(\mathbf{x})$ と $\sigma(\mathbf{x})$ はそれぞれ \mathbf{x} における評価指標（本研究ではKL情報量）の確率分布の期待値と標準偏差である。

Eq. (16)において β を大きくすれば、評価指標の不確実性が大きい \mathbf{x} 、すなわち、これまで実験した条件から遠く離れた製造条件を提案するようになる。一方、 β を小さくすれば、評価指標の期待値が大きい \mathbf{x} 、すなわち、これまで実験した条件のうち評価指標を最小にする条件に近い製造条件を提案するようになる。

評価指標の確率分布が正規分布などの場合は、期待値と標準偏差を解析的に計算できる。しかしながら、本研究では評価指標にKL情報量を用いており、また、KL情報量の算出に際してパーセンタイルベクトルから粒子径分布への変換処理を行っているため、評価指標の期待値や標準偏差を解析的に導くことはできない。そこで、モンテカルロ法[28]を用いて、Physical+MOGPRモデルから得られるパーセンタイルベクトルの事後分布から複数個ランダムサンプリングし、サンプリングされたパーセンタイルベクトルから2.4項で示した手順に従ってKL情報量を計算し、KL情報量の経験分布を生成する。サンプリング回数は、経験分布の近似精度と計算負荷のバランスから1000回に設定した。経験分布上でのLCBは、正規分布の累積分布関数からEq. (16)の値の累積確率を求め、それと同じ累積確率をもつ経験分布上の点の値を求めることで算出する。すなわち、(期待値 $-\beta$ 標準偏差)が正規分布において p パーセンタイルに相当すれば、生成した経験分布における p パーセンタイルをLCBとする。たとえば、Eq. (16)の β が1の場合は、正規分布の15.9パーセンタイルに相当するので、1000回のランダムサンプリングからなるKL情報量の経験分布の小さい方から159番目の値をLCBとする。

さらに、実験回数 (n) が増えるに従って β を小さくするように

$$\beta = \sqrt{\frac{\log n}{n}} \quad (17)$$

と設定した。この設定により、実験回数が少ない段階では分散に重みを置く獲得関数になり、次項の最適化手法は観測データが疎な領域を探索する。実験回数が増えるに伴い期待値に重みを置く獲得関数になり、最適化手法はこれまでの中で最善の観測データの付近を探索するようになる。これにより、効率的に製造条件を探索することが期待される。

なお、サンプリングしたパーセントイルの大小関係に、たとえば $d_{10} < d_{30} < d_{20} < d_{40}$ のような逆転を生じる可能性がある。その場合、逆転が発生したパーセントイルの前後のパーセントイルで線形補間した値で代替した。前述の例では、 d_{20} と d_{30} のペアの前後にある d_{10} と d_{40} の線形補間値で d_{20} と d_{30} を代替した。

2.6 最適化手法

前項で定義した獲得関数を最小化する説明変数 \mathbf{x} を見つけたい。獲得関数が説明変数 \mathbf{x} で微分できるならば勾配法などの手法を用いることができる。しかし、本研究での獲得関数は KL 情報量の経験分布に基づいており、説明変数 \mathbf{x} での微分は不可能である。そこで、メタヒューリスティクス[29]の一種であり、微分不可能な関数であっても大域最適化が可能な、差分進化 (Differential Evolution, DE) [30,31]を採用した。差分進化法による最適化は次のフローにて行われる。

- i. ラテン超方格サンプリング[32]に従って説明変数 \mathbf{x} の候補空間から候補点 $\mathbf{x}_{\text{Candidate}}$ を所定の個数選ぶ。
- ii. 差分進化の戦略 (best/1/bin) に従って、その複数の候補点 $\mathbf{x}_{\text{Candidate}}$ から一部 (親) を抽出し、その親から新たな候補点 (子) を生成する。
- iii. 親および子について獲得関数を評価し、関数値が小さい候補点を残す。
- iv. iii で残った候補点と ii で子を生成しなかった候補点を次の世代の候補点とする。
- v. ii~iv を終了条件に達するまで繰り返す。
- vi. 残った複数の候補点のうち獲得関数の値がもっとも小さい候補点を最適解 \mathbf{x}_{Next} とする。

2.7 事前に想定できない運転制約への対応

本研究の対象である ACM の製造条件は、原料粉の供給速度 F [kg/h]、粉砕ロータの回転速度 R [rpm]、分級ロータの回転速度 N [rpm]、および、空気の通過速度 Q [m³/min] の4つである。各製造条件には運転可能な上下限値が設けられており、その範囲を候補空間として、前項で説明した差分進化法で、評価指標を最適化する製造条件を導出する。

しかしながら、ACM には、候補空間に含まれる製造条件であっても、実際に実験してみると、消費電力や消費電流が上限を超えてしまい運転不可になることがある。消費電力や消費電流は、原料粉の種類や製造装置の機械

的特徴、製造条件により複雑に変化するため、上限を超えるか否かを事前に判定することは困難である。

このような事前に想定できない制約も、ベイズ最適化に組み込みたい。そこで、ベイズ最適化が提案する製造条件 (\mathbf{x}_{Next}) で実際に運転をした結果、消費電力ないしは消費電流が上限を超えると判明した場合は、製造条件のうち原料粉の供給速度 F [kg/h] だけを下方に調整することで消費電力および消費電流を上限以下におさめた。ここで原料粉の供給速度だけを調整したのは、粉砕機のメカニズムから、消費電力および消費電流を低減したい場合には原料粉の供給速度を低減することが効果的とわかっているからである。修正後の製造条件 ($\mathbf{x}_{\text{Next,Fix}}$) により ACM で製造した粒子径分布のパーセントイルベクトル ($\mathbf{y}_{\text{Next,Fix}}$) を得る。そして、 $\{\mathbf{x}_{\text{Next,Fix}}, \mathbf{y}_{\text{Next,Fix}}\}$ ではなく、 $\{\mathbf{x}_{\text{Next}}, \mathbf{y}_{\text{Next,Fix}}\}$ を観測データセット \mathcal{D} に加えることで、事前には想定できない制約も考慮したベイズ最適化を可能にした。

2.8 提案手法の全体フロー

本研究で提案するベイズ最適化の概要を Fig. 3 に示す。まず、Physical+MOGPR モデルに製造条件 \mathbf{x} を入力しパーセントイルベクトルの予測分布を得る。次に、予測分布からパーセントイルベクトルを 1000 回ランダムサンプリングし、それぞれを粒子径分布に変換して目標分布との KL 情報量を計算することで、KL 情報量の経験分布を得る。最後に、KL 情報量の LCB を獲得関数として算出する。この流れで算出される LCB を最小にするような製造条件 \mathbf{x} を差分進化法により導出する。

提案手法の全体アルゴリズムを Fig. 4 に示す。まず、回帰モデルの学習フェーズ (パラメータの推定) では、すでに実験で得られた観測データセット $\mathcal{D} = \{\mathbf{X}, \mathbf{Y}\}$ を基に、Physical モデル (関数 f_{Physical}) のパラメータ \mathbf{c} を最尤法で推定する。そして、Physical モデルの残差を目的変数として、MOGPR モデル (関数 f_{MOGPR}) のパラメータ θ を最尤法で推定する。

次に、差分進化法のラテン超方格サンプリングにより実験条件の複数候補 $\mathbf{x}_{\text{Candidate}}$ を抽出し、各候補について、学習済みの Physical+MOGPR モデルからパーセントイルベクトルの予測分布を導出し、それを KL 情報量の確率分布に変換し、その LCB を算出する。差分進化法に従い、複数候補 $\mathbf{x}_{\text{Candidate}}$ の中で最小の LCB が得られた $\mathbf{x}_{\text{Candidate}}$ を次の実験条件 \mathbf{x}_{Next} とする。

そして、ベイズ最適化が提案した \mathbf{x}_{Next} にて粉体製造装置を運転し、新たな粒子径分布のパーセントイルベクトル (\mathbf{y}_{Next}) を得た後、 $\{\mathbf{x}_{\text{Next}}, \mathbf{y}_{\text{Next}}\}$ を観測データセット \mathcal{D} に追加する。ただし、消費電力ないし消費電流が上限を超えた場合には、製造条件のうち原料粉の供給速度 F [kg/h] だけを下方に調整することで消費電力および消費電流を上限以下におさめ、修正後の製造条件 ($\mathbf{x}_{\text{Next,Fix}}$) で製造した粒子径分布のパーセントイルベクトル ($\mathbf{y}_{\text{Next,Fix}}$) を得る。そして、 $\{\mathbf{x}_{\text{Next}}, \mathbf{y}_{\text{Next,Fix}}\}$ を観測データセット \mathcal{D} に加える。

以上のプロセスを、目標分布 t との差異が許容範囲に

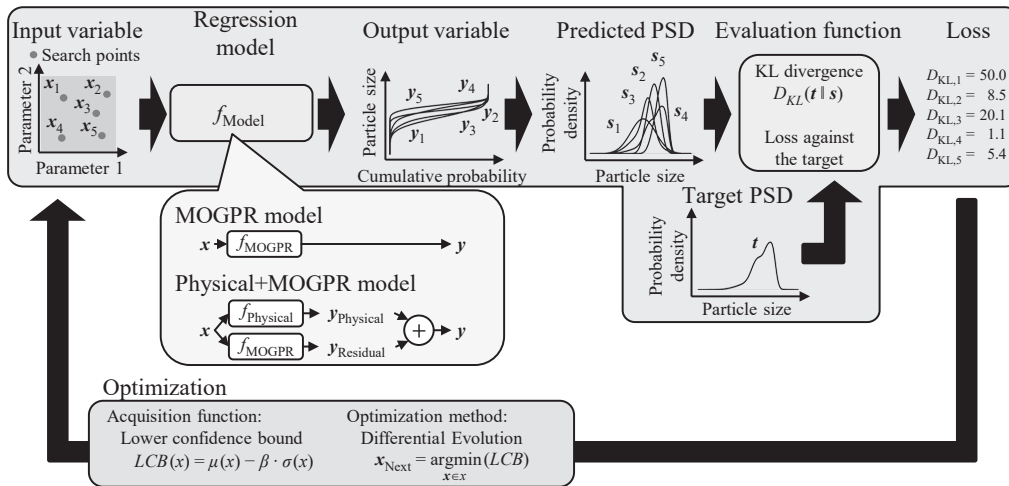


Fig. 3 A schematic diagram of our Bayesian optimization to obtain the desired particle size distribution. The diagram shows the structure of regression, evaluation function, acquisition function, and optimization

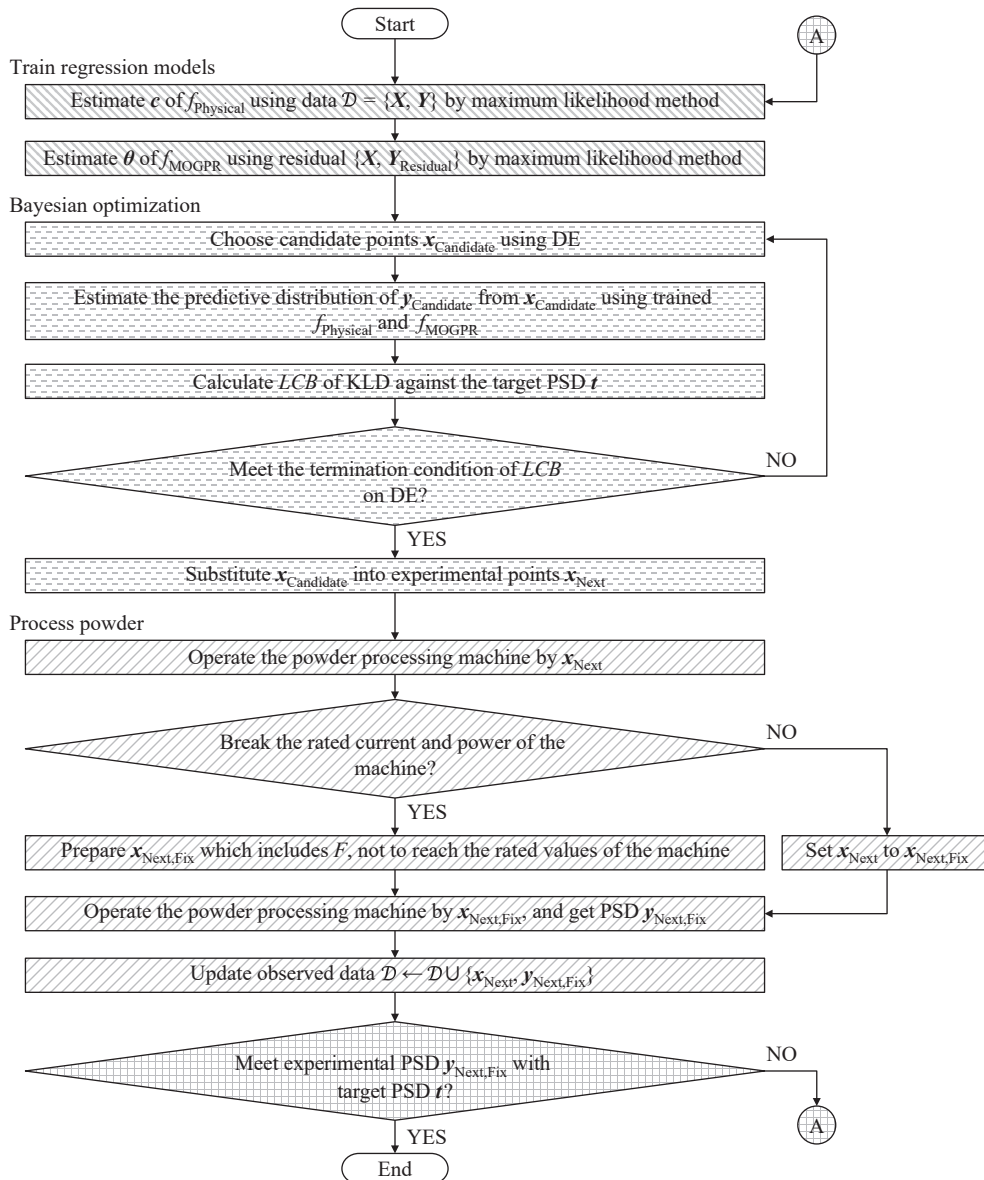


Fig. 4 Flowchart of our Bayesian optimization to obtain the desired particle size distribution. It is divided into three phases: training regression model, Bayesian optimization, and operating powder processing equipment

収まるような粒子径分布が観測されるまで繰り返す。

3. 提案手法の有効性の評価

3.1 シミュレータを用いた数値実験による評価

提案手法の有効性を評価するために、ホソカワミクロン製の ACM-15H を用いた実験を行う。

まず、ACM-15H のシミュレータを用いた数値実験により、提案手法の挙動を確認する。シミュレータは、ACM-15H を用いて重質炭酸カルシウムの粉体を製造したデータを用いて開発したものである。具体的には、41 通りの製造条件で重質炭酸カルシウムを製造し、それぞれにおける粒子径分布、消費電力、および消費電流を計測し、これら計測データにフィットするような現象論的シミュレーションモデルを事前に開発した。開発したシミュレーションモデルに 41 通りの製造条件を入力したときに出力されるパーセンタイル粒子径 (d_{10} , d_{50} , および d_{100}) について、観測値との比較を Fig. 5 に示す。一般的に、シミュレータは、ACM-15H により製造される重質炭酸カルシウムの粒子径分布をよく再現できていることがわかる。

このシミュレータ上で目標とする確率分布を得るための製造条件を得るべく、2 節で開発した手法に従って製造実験を 10 回重ねた。目標分布は、公平性を確保するために、4 つの製造条件それぞれについて上下限範囲内でランダムな値を選択し、シミュレータ上でその製造条件を与えて得られた粒子径分布とした。また、ベイズ最適化のアルゴリズムには乱数を用いており、生成した乱数により結果は変わるため、この 10 回の製造実験を 5 ラウンド繰り返すことで乱数によるバラつきも評価した。さらに、比較のために、予測モデルに MOGPR モデルだけを用いるケース、および Physical モデルだけを用いるケースでも同様の実験を行った。

Fig. 6 は、製造実験の回数（何回目の製造実験か）とそのときの KL 情報量の推移を示したものである。図中の帯は、5 ラウンド繰り返した結果から計算される 95% 信頼区間である。予測モデルに Physical+MOGPR モデルを用いた場合および MOGPR モデルを用いた場合は、実験を重ねるに従い KL 情報量は減少し、また、両者の間では信頼区間が重なり有意な差は見られなかった。一方、Physical モデルは、実験を重ねても KL 情報量は改善されなかった。このことから、MOGPR は KL 情報量を小さくする製造条件の探索に有効に働いていること、および Physical モデルによる外挿領域のガイド効果は本実験では小さかったことが推察される。

Fig. 7 は、予測モデルに Physical+MOGPR モデルを用いた場合と MOGPR モデルを用いた場合における、製造条件の推移を示したものである。前者においては、1 回目の実験から、製造条件の候補空間における内部の候補点を提案している。一方、後者においては、2 回目の実験まで、製造条件の候補空間の境界にある候補点を提案している。このことから、Physical モデルによる製造条件探索のガイドは、機能していると思われる。

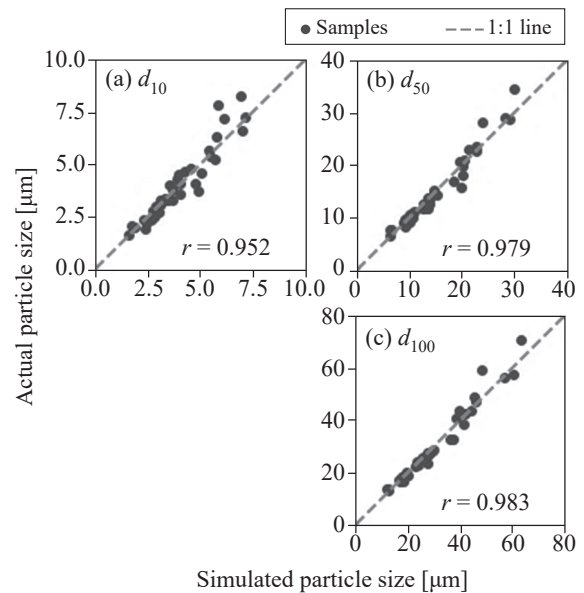


Fig. 5 Representative particle sizes from actual data and simulator output. (a) is d_{10} , (b) is d_{50} , and (c) is d_{90} . The dots represent the 41 data points. The dashed line indicates the position of the case where the actual data and the simulator output agree

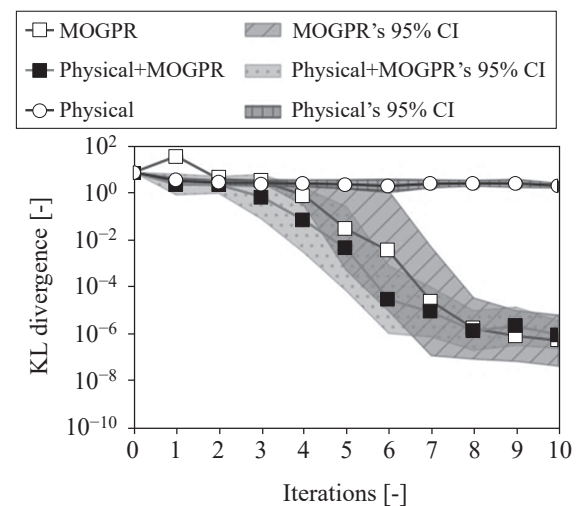


Fig. 6 The changes of the KL divergence for increasing iterations on the numerical experiment. The confidence intervals (CI) were calculated from the results of five times experiments

Fig. 8 (a) は、9 回目の実験後に得られた Physical モデルについて、パーセンタイルごとの予測誤差の相関行列（観測されたパーセンタイルベクトルと Physical モデルが出力するパーセンタイルとの差分ベクトルから計算した相関行列）を表したものである。Fig. 8 (b) は、Physical+MOGPR モデルにおける共分散行列 B から換算した相関行列を示す。Fig. 8 (c) は、Fig. 8 (a) と Fig. 8 (b) の相関行列の差分を示す。2 つの相関行列の傾向は類似しており、Physical モデルが予測するパーセンタイルベクトルの誤差について、MOGPR モデルはパーセンタイル間相関を捉えられていると考えられる。一方、共分散行列 B から換算した相関係数の値は、Physical モデルの予測誤差の相関行列に一致していない。これは、本研究では MOGPR

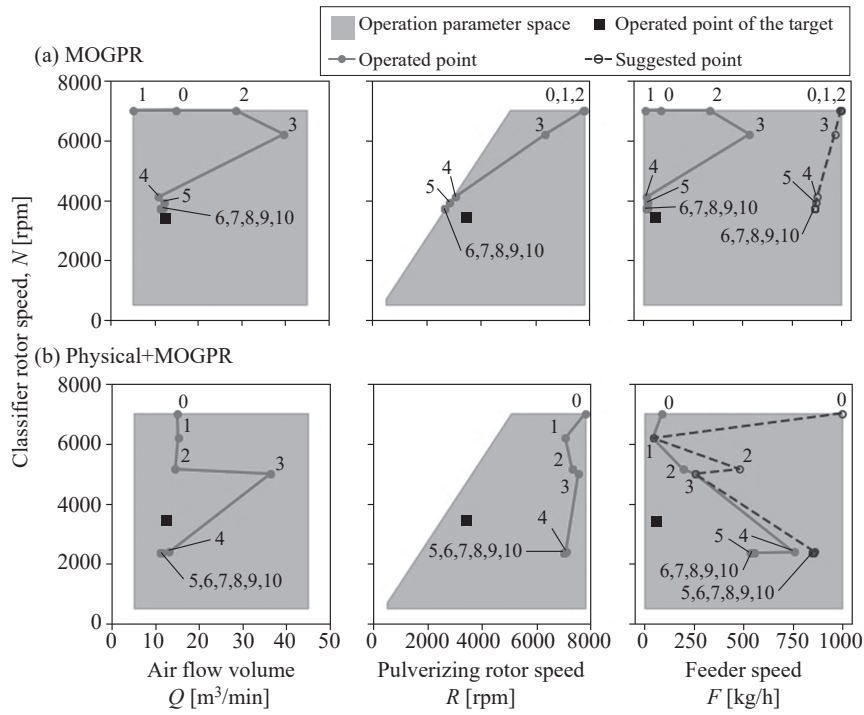


Fig. 7 The operating parameters of MOGPR and Physical+MOGPR model on numerical experiments. (a) is MOGPR model, and (b) is the result of Physical+MOGPR model. Each axis represents the operating parameters. “Operating parameter space” indicates the upper and lower bounds of the ACM operating parameters. “Operated point of the target” is the operating parameter that produced the target PSD. “Operated point” is adjusted downward parameters $x_{Next,Fix}$ by the rated electrical current and power. “Suggested point” is the suggested operating parameter x_{Next} from Bayesian optimization. “Suggested point”s of N , Q , and R are not shown in the figures because they are the same as “Operated point”

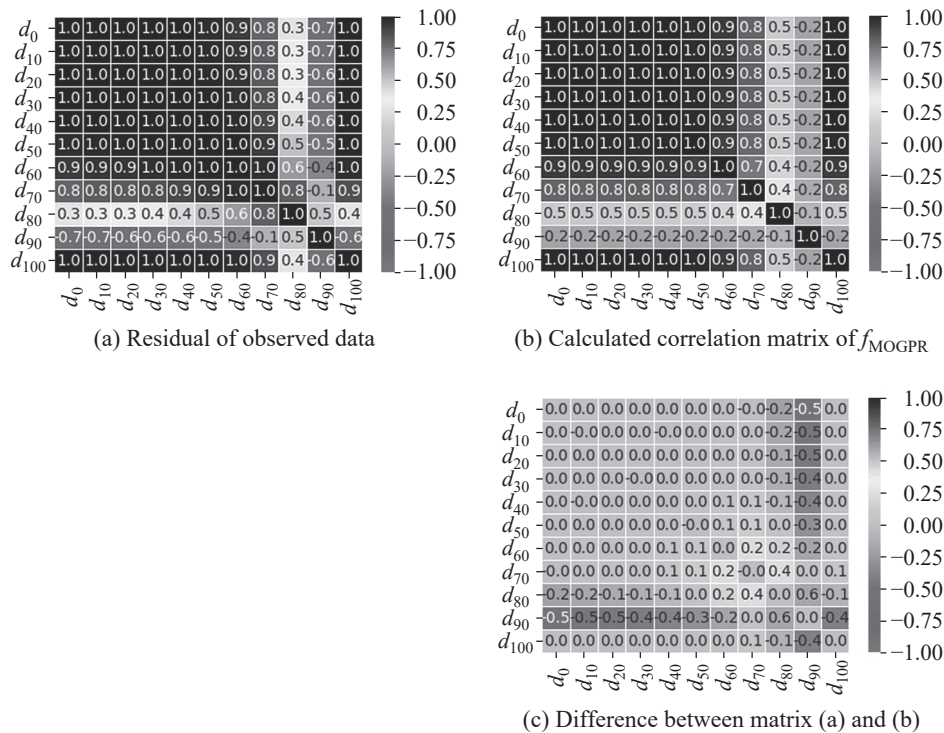


Fig. 8 Correlation matrix of particle size per quantile at the 9th iteration on the numerical experiment. (a) is the correlation matrix of the predicted residuals of $f_{Physical}$ from Physical+MOGPR model, (b) is the correlation matrix converted from the covariance matrix B of f_{MOGPR} from Physical+MOGPR model, and (c) is the difference between (a) and (b)

モデルとして Eq. (7) で定義した ICM を用いており、共分散行列を表現する自由度が小さいためと考えられる。

Fig. 9 は、Physical+MOGPR モデルで用いているカーネル関数のパラメータについて、実験の回数を重ねるに伴った推移を示したものである。5 回目の実験まではいずれのパラメータも変動は大きく、6 回目の実験以降は収束に向かった。これは、実験回数が少ないときは少数の観測データでパラメータ推定するため、観測データが追加されるたびに推定値は大きく変動したと考えられる。一方、RBF カーネルのパラメータ θ_{RBF}^2 は、線形カーネルのパラメータ θ_{Linear}^2 を一貫して上回っており、本実験で観測データから学習したベイズ過程帰帰は非線形性が強いと考えられる。

3.2 実機を用いた実験による評価

次に、実機による実験を行った。製造装置には ACM-2EC (ホソカワミクロン) を使用し、原料粉には寒水石を用いた。

目標分布は、前項同様に、製造条件の候補空間からランダムに製造条件を選択し、その条件で実際に製造して得られた粉体の粒子径分布とする。しかしながら、観測される粒子径分布は測定誤差を伴う。そこで、粒子径分布の計測を 3 回繰り返し、その平均を目標分布とした。また、3 回の計測のそれぞれで得られた粒子径分布について目標分布との間の KL 情報量を算出し、そのバラつきを測定誤差に起因する KL 情報量の変動とみなした。そして、このバラつきの標準偏差を算出し、標準偏差の 1 倍に相当する値を、製造実験で観測された粒子径分布と目標分布との間の KL 情報量の合格基準とした。

まず、本研究で提案する Physical+MOGPR モデルを用いた手法に従って、製造実験を 8 回重ねた。次に、比較のために、2 名のエンジニア (エンジニア経験 2 年とエンジニア経験 13 年) それぞれに、目標分布を得る製造条件を見出すべく、製造実験を 8 回行ってもらった。

Fig. 10 は、それぞれのケースについて、製造実験の回数 (何回目の製造実験か) に対する KL 情報量の推移を示したものである。図中の破線は、前述の合格基準を表す。Fig. 10 から、本研究で提案する Physical+MOGPR モデルを用いた手法も、2 名のエンジニアも、8 回目の実験までに目標分布との KL 情報量が合格基準を満たす製造条件を見出せた。Fig. 11 は、Physical+MOGPR モデルを用いた手法に従って実験を重ねた場合の、実験回数ごとに得られた粒子径分布を表す。製造実験が 6 回目までは、実験回数が増えるに従い粒子径分布は目標分布に近づいていることが見てとれる。一方、7 回目と 8 回目は、目標分布より分布の裾が狭い粒子径分布になり、Fig. 10 では KL 情報量が大きくなった。

しかしながら、Fig. 10 からわかるように、Physical+MOGPR モデルを用いた手法は、KL 情報量が合格基準に達するまでに、エンジニアより多く実験回数を要した。Fig. 12 に、各ケースにおける製造条件の推移を示す。図中の四角印は、目標分布を生成した製造条件を表す。目標分布を生成した製造条件は、粉碎ロータの回転速度 R

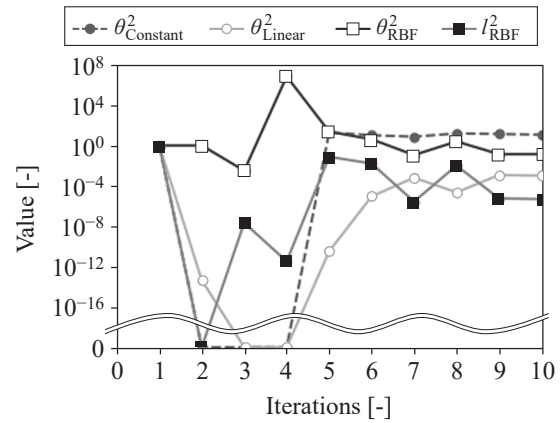


Fig. 9 Estimated parameter value of kernel function, Eq. (8), on numerical experiment

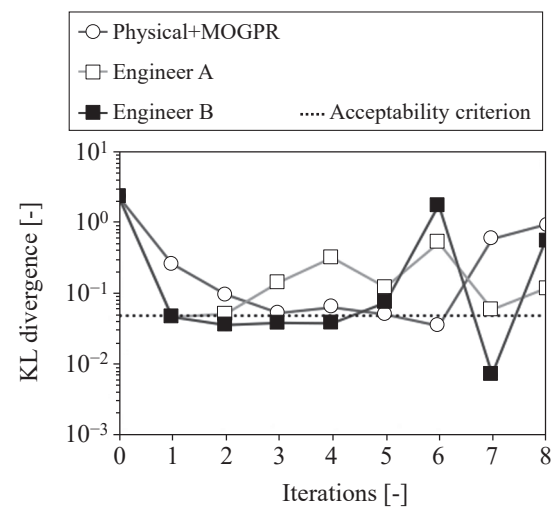


Fig. 10 The changes of the KL divergence for increasing iterations in the actual experiment. “Acceptability criterion” in the legend is a standard deviation of KL divergence due to measurement errors in the three PSD measurements for the target

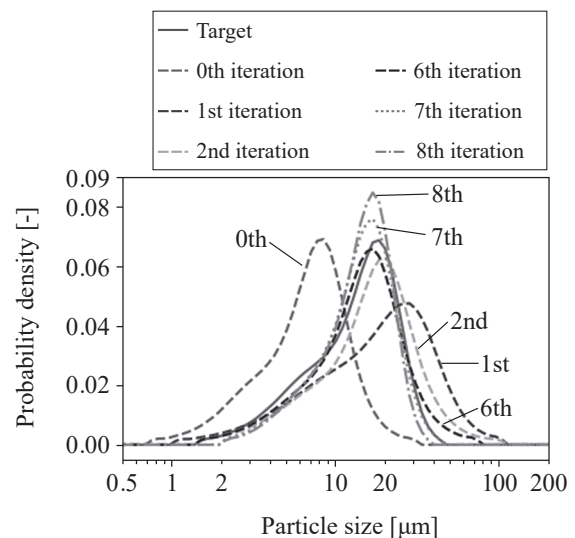


Fig. 11 Particle size distribution at each iteration on the actual experiment

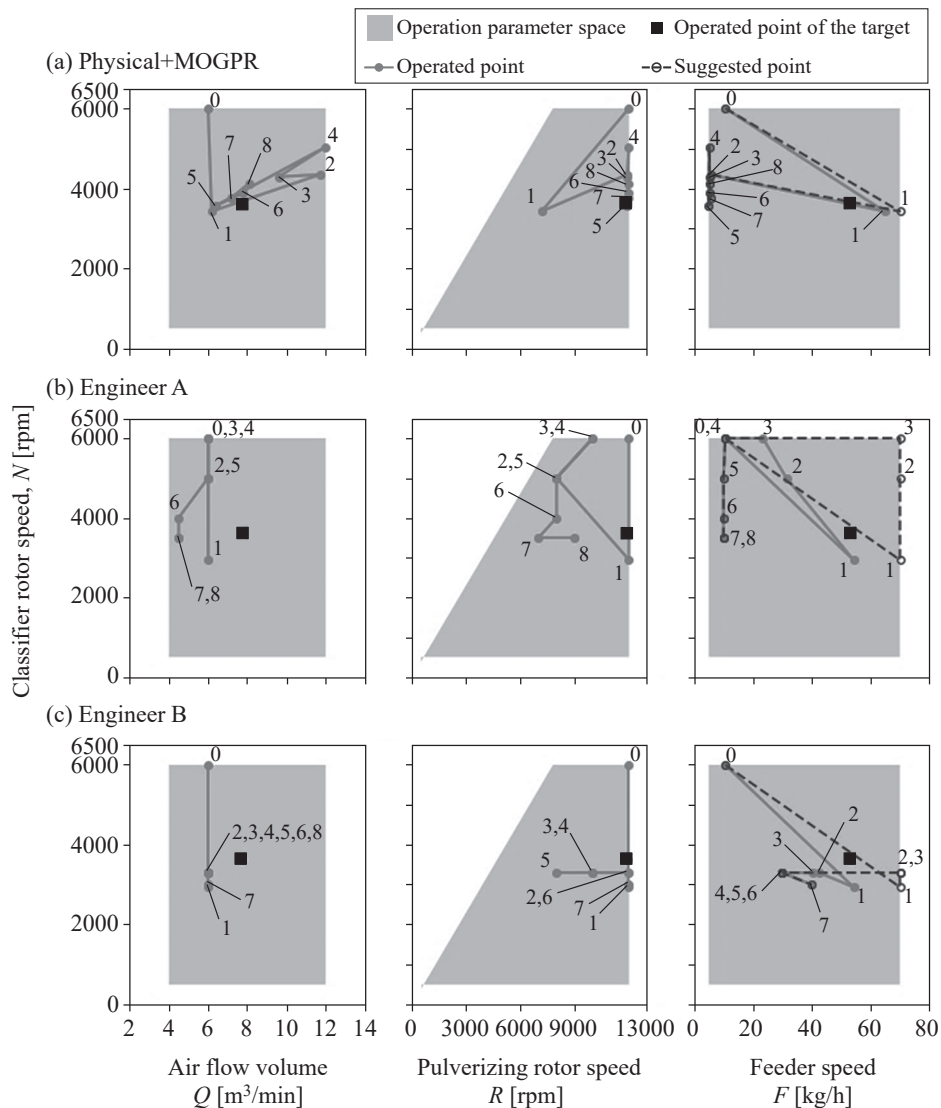


Fig. 12 The operating parameters of Physical+MOGPR model and two Engineers on the actual experiment. (a) shows the results for Physical+MOGPR model, (b) is for Engineer A, and (c) is for Engineer B. The legend is the same as in Fig. 7

が運転上限値に近かった。一方、エンジニアは、経験則により1回目の条件は R を運転上限値で提案していた。偶然にも、この両者がほぼ一致したため、エンジニアはPhysical+MOGPRモデルを用いた手法よりも少ない実験回数で合格基準を満たす製造条件に達した可能性が考えられる。

4. 結 言

本研究では、製造する粉体の粒子径分布を目標分布に合わせることを目標とする、多出力ガウス過程に基づくベイズ最適化を用いた適応的実験計画法を開発した。対象とした粉体製造装置は、ACM (air classifier mill) である。開発した手法の特徴は、以下の5点である。

1. 粒子径分布を目的変数とする多出力ガウス過程回帰を設計するにあたり、粒子径の離散分布の確率質量ベクトルを目的変数にすると、総和は1である制約のもとで回帰問題を解かなければならない。この制

約を回避するため、確率分布のパーセントイルベクトルを目的変数とした。また、通常の高出力ガウス過程回帰はパラメータ数が多く学習を困難にするため、ICM (intrinsic coregionalization model) を用いることでパラメータ数を抑制した。

2. 製造する粉体の粒子径分布と目標分布の差異を評価する指標について、パーセントイルベクトル間のノルムでは大粒径の影響が大きくなるため、パーセントイルベクトルを確率質量ベクトルに戻し、確率質量ベクトル間のKL情報量を評価指標とした。
3. ベイズ最適化にあたっては、獲得関数をLCB (lower confidence bound) とし、さらに実験当初は分散の重みを大きく、実験を重ねるに従い分散の重みを小さくした。また、最適化アルゴリズムには差分進化法を用いた。
4. ガウス過程回帰の外挿領域における予測精度を向上するため、物理モデルとガウス過程回帰のハイブリッ

ドモデルとした。物理モデルには、分級工程を表すストークスの式を用いた。

5. 粉体製造装置は、事前にわかっている運転可能範囲で実験しても、消費電力や消費電流が上限を超過してしまうことがあり、その場合は上限を超えないよう製造条件を修正する。この場合、修正前の製造条件と修正後の運転で製造した粉子粒分布を対応付けて観測データセットに加えることで、ベイズ最適化の枠組みに組み込んだ。

提案手法の有効性評価のために、ホソカワミクロン製の ACM-15H および ACM-2EC を使用した数値実験と実機による実験を行った。数値実験では、多出力ガウス過程回帰と物理モデルのいずれも期待通りに機能し、また、実験を重ねるに従い目標分布に近づくことを確認した。実機による実験では、目標分布との差異は計測誤差範囲内とみなされる粒子径分布を製造する製造条件に、実験 8 回以内に到達した。一方、エンジニアと比べると、目標分布に到達するまでの実験回数は劣化した。ただし、

実機実験で目標とした分布を製造する製造条件は、エンジニアの経験則に近かったため、本実験をもって、提案手法はエンジニアの実験計画よりも劣るとは断定できない。

われわれが提案した手法は、粉体製造において、エンジニアの経験知を必要とせずに、目標分布と同じ粒子径分布を製造する実験条件を見出だすことを可能にする。この手法は、任意の粒子径分布を対象とできるため、たとえば多峰性のある粒子径分布が目標分布であっても、本手法の適用ができる。ただし、製造条件のパラメータが非常に多い場合でも、目標分布が得られるか否かについては、さらなる検証が必要である。

今後の展望として、われわれの提案手法とエンジニアの間で、目標分布に至るまでの実験回数の比較を、さまざまな目標分布に対して行いたい。また、観測データが少ないときには、MOGPR の推定パラメータが局所解に陥っている可能性があるため、パラメータの数をさらに抑えた回帰モデルを開発していきたい。

Nomenclature

\mathbf{B}	: covariance matrix	[-]	\mathcal{X}	: experimental parameters space	[-]
b	: element of \mathbf{B} in Eq. (6)	[-]	\mathbf{x}	: vector of explanatory variable	[-]
\mathbf{c}	: constant vector in Eq. (13)	[-]	$\mathbf{x}_{\text{Candidate}}$: candidate experimental parameter for experiment	[-]
d_m	: particle size corresponding to $m\%$ percentile	[μm]	\mathbf{x}_{Next}	: suggested experimental parameters for next iteration	[-]
$\text{diag}_{\text{vector} \rightarrow \text{matrix}}(\cdot)$: function transform from vector to diagonal matrix	[-]	$\mathbf{x}_{\text{Next,Fix}}$: adjusted downward experimental parameters for next iteration	[-]
D	: number of dimensions of explanatory variable	[-]	\mathbf{y}	: vector of particle sizes at each percentile	[-]
$D_{\text{KL}}(\cdot\ \cdot)$: a function of Kullback–Leibler divergence	[-]	\mathbf{y}_{Next}	: vector of particle size gotten at next iteration	[-]
$f_{\text{GPR}}(\cdot)$: a function of Gaussian process regression	[-]	$\mathbf{y}_{\text{Next,Fix}}$: vector of particle size gotten under $\mathbf{x}_{\text{Next,Fix}}$	[-]
$f_{\text{MOGPR}}(\cdot)$: a function of multi-output Gaussian process regression	[-]	y	: scalar of target variable	[-]
F	: feeder speed	[kg/h]	y_j	: j th element of \mathbf{y}	[μm]
$\mathcal{GP}(\cdot, \cdot)$: Gaussian process	[-]	z	: any positive real number	[-]
$k(\cdot, \cdot)$: kernel function	[-]	β	: variable for lower confidence bound in Eq. (16) and Eq. (17)	[-]
L	: number of dimensions of target variable	[-]	$\boldsymbol{\varepsilon}$: vector of observation noise	[-]
l_{RBF}^2	: a length scale for radial basis function kernel in Eq. (8)	[-]	ε	: scalar of observation noise	[-]
N	: classifier rotor speed	[rpm]	$\boldsymbol{\theta}$: vector of parameter in MOGPR model in Eq. (9)	[-]
\mathbf{p}	: probability distribution	[-]	$\theta_{\text{Constant}}^2$: coefficient for constant kernel in Eq. (8)	[-]
$p(\cdot)$: probability density function of \mathbf{p}	[-]	θ_{Linear}^2	: coefficient for linear kernel in Eq. (8)	[-]
\mathbf{q}	: probability distribution	[-]	θ_{RBF}^2	: coefficient for radial basis function kernel in Eq. (8)	[-]
$q(\cdot)$: probability density function of \mathbf{q}	[-]	$\boldsymbol{\kappa}$: vector of constant in Eq. (7)	[-]
\mathbf{s}	: vector of particle size distribution	[-]	$\mu(\cdot)$: expected value function of evaluation index in Eq. (16)	[-]
s_j	: j th element of \mathbf{s}	[-]	$\sigma(\cdot)$: standard deviation function of evaluation index in Eq. (16)	[-]
Q	: air flow volume	[m^3/min]	σ^2	: variance of observation noise	[-]
R	: pulverizing rotor speed	[rpm]			
\mathbb{R}	: the entire set of real numbers	[-]			
\mathbf{t}	: vector of target particle sizes	[-]			
t_j	: j th element of \mathbf{t}	[μm]			
\mathbf{w}	: vector of constant in Eq. (7)	[-]			

References

- [1] T. Allen, *Powder Sampling and Particle Size Determination*, Elsevier Science Limited (2003).
- [2] K. Leschonski, Representation and evaluation of particle size analysis data, *Part. Part. Syst. Charact.* 1 (1984) 89–95.
- [3] R. Hogg, Issues in particle size analysis, *KONA* 26 (2008) 81–93.
- [4] A. J. R. Rennie, V. L. Martins, R. M. Smith, P. J. Hall, Influence of particle size distribution on the performance of ionic liquid-based electrochemical double layer capacitors, *Sci. Rep.* 6 (2016) 22062.
- [5] S. Wu, B. Yu, Z. Wu, S. Fang, B. Shi, J. Yang, Effect of particle size distribution on the electrochemical performance of micro-sized silicon-based negative materials, *RSC Adv.* 8 (2018) 8544–8551.
- [6] L. Bläubaum, F. Röder, C. Nowak, H. S. Chan, A. Kwade, U. Krewer, Impact of particle size distribution on performance of lithium-ion batteries, *ChemElectroChem* 7 (2020) 4755–4766.
- [7] C. Fushimi, K. Yato, M. Sakai, T. Kawano, T. Kita, Recent progress in efficient gas–solid cyclone separators with a high solids loading for large-scale fluidized beds, *KONA* 38 (2021) 94–109.
- [8] L. M. Tavares, A review of advanced ball mill modelling, *KONA* 34 (2017) 106–124.
- [9] A. Hadi, R. Roepfl, Y. Pang, D. L. Schott, Dem modelling of segregation in granular materials: A review, *KONA* 41 (2024) 78–107.
- [10] M. Hounslow, The population balance as a tool for understanding particle rate processes, *KONA* 16 (1998) 179–193.
- [11] L. Vogel, W. Peukert, Modelling of grinding in an air classifier mill based on a fundamental material function, *KONA* 21 (2003) 109–120.
- [12] F. J. Doyle III, C. A. Harrison, T. J. Crowley, Hybrid model-based approach to batch-to-batch control of particle size distribution in emulsion polymerization, *Comput. Chem. Eng.* 27 (2003) 1153–1163.
- [13] S. Greenhill, S. Rana, S. Gupta, P. Vellanki, S. Venkatesh, Bayesian optimization for adaptive experimental design: A review, *IEEE Access* 8 (2020) 13937–13948.
- [14] T. Ueno, T. D. Rhone, Z. Hou, T. Mizoguchi, K. Tsuda, COMBO: An efficient Bayesian optimization library for materials science, *Mater. Discov.* 4 (2016) 18–21.
- [15] A. Denzel, J. Kästner, Gaussian process regression for geometry optimization, *J. Chem. Phys.* 148 (2018) 094114.
- [16] A. Candelieri, R. Perego, F. Archetti, Bayesian optimization of pump operations in water distribution systems, *J. Glob. Optim.* 71 (2018) 213–235.
- [17] Y. Morita, S. Rezaeiravesh, N. Tabatabaei, R. Vinuesa, K. Fukagata, P. Schlatter, Applying Bayesian optimization with Gaussian process regression to computational fluid dynamics problems, *J. Comput. Phys.* 449 (2022) 110788.
- [18] R. Martinez-Cantin, Bayesian optimization with adaptive kernels for robot control, 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore (2017) pp. 3350–3356.
- [19] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, N. de Freitas, Taking the human out of the loop: A review of Bayesian optimization, *Proceedings of the IEEE* 104 (2016) 148–175.
- [20] J.-P. Vert, K. Tsuda, B. Schölkopf, A primer on kernel methods, *Kernel Methods in Computational Biology*, The MIT Press (2004).
- [21] H. Liu, J. Cai, Y.-S. Ong, Remarks on multi-output Gaussian process regression, *Knowl. Based Syst.* 144 (2018) 102–121.
- [22] E. V. Bonilla, K. Chai, C. Williams, Multi-task Gaussian process prediction, *Adv. Neural Inf. Process. Syst.* 20 (2007) 505–512.
- [23] M. A. Branch, T. F. Coleman, Y. Li, A subspace, interior, and conjugate gradient method for large scale bound-constrained minimization problems, *SIAM J. Sci. Comput.* 21 (1999) 1–23.
- [24] A. Daw, A. Karpatne, W. Watkins, J. Read, V. Kumar, Physics-guided neural networks (PGNN): An application in lake temperature modeling, *Knowledge Guided Machine Learning*, Taylor & Francis (2022).
- [25] C. Spötter, K. Legenhausen, A. P. Weber, Separation characteristics of a deflector wheel classifier in stationary conditions and at high loadings: New insights by flow visualization, *KONA* 35 (2018) 172–185.
- [26] S. Kullback, R. A. Leibler, On information and sufficiency, *Ann. Math. Statist.* 22 (1951) 79–86.
- [27] J. T. Wilson, F. Hutter, M. P. Deisenroth, Maximizing acquisition functions for Bayesian optimization, *Adv. Neural Inf. Process. Syst.* 31 (2018) 9884–9895.
- [28] R. L. Harrison, Introduction to Monte Carlo Simulation, *AIP Conf. Proc.* 1204 (2010) 17–21.
- [29] K. Rajwar, K. Deep, S. Das, An exhaustive review of the metaheuristic algorithms for search and optimization: Taxonomy, applications, and open challenges, *Artif. Intell. Rev.* 56 (2023) 13187–13257.
- [30] R. Storn, K. Price, Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces, *J. Glob. Optim.* 11 (1997) 341–359.
- [31] J. Lampinen, A constraint handling approach for the differential evolution algorithm, *Proceedings of the 2002 Congress on Evolutionary Computation. CEC’02 (Cat. No.02TH8600)*, USA (2002) pp. 1468–1473.
- [32] M. D. McKay, R. J. Beckman, W. J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics* 21 (1979) 239–245.