materials

MDPI

Article

Optimizing Thermoplastic Starch Film with Heteroscedastic Gaussian Processes in Bayesian Experimental Design Framework

Gracie M. White [1], Amanda P. Siegel [2] and Andres Tovar [3,*]

[1] Luddy School of Informatics, Computing, and Engineering, Integrative Nanosystems Development Institute (INDI), Indiana University Indianapolis, Indianapolis, IN 46202, USA; whitegra@iu.edu
[2] Department of Chemistry and Chemical Biology, Indiana University Indianapolis, Indianapolis, IN 46202, USA; apsiegel@iu.edu
[3] School of Mechanical Engineering, Purdue University, Indianapolis, IN 46202, USA
* Correspondence: tovara@purdue.edu; Tel.: +1-765-495-7765

Abstract: The development of thermoplastic starch (TPS) films is crucial for fabricating sustainable and compostable plastics with desirable mechanical properties. However, traditional design of experiments (DOE) methods used in TPS development are often inefficient. They require extensive time and resources while frequently failing to identify optimal material formulations. As an alternative, adaptive experimental design methods based on Bayesian optimization (BO) principles have been recently proposed to streamline material development by iteratively refining experiments based on prior results. However, most implementations are not suited to manage the heteroscedastic noise inherently present in physical experiments. This work introduces a heteroscedastic Gaussian process (HGP) model within the BO framework to account for varying levels of uncertainty in the data, improve the accuracy of the predictions, and increase the overall experimental efficiency. The aim is to find the optimal TPS film composition that maximizes its elongation at break and tensile strength. To demonstrate the effectiveness of this approach, TPS films were prepared by mixing potato starch, distilled water, glycerol as a plasticizer, and acetic acid as a catalyst. After gelation, the mixture was degassed via centrifugation and molded into films, which were dried at room temperature. Tensile tests were conducted according to ASTM D638 standards. After five iterations and 30 experiments, the films containing 4.5 wt% plasticizer and 2.0 wt% starch exhibited the highest elongation at break (M = 96.7%, SD = 5.6%), while the films with 0.5 wt% plasticizer and 7.0 wt% starch demonstrated the highest tensile strength (M = 2.77 MPa, SD = 1.54 MPa). These results demonstrate the potential of the HGP model within a BO framework to improve material development efficiency and performance in TPS film and other potential material formulations.

Keywords: thermoplastic starch; potato starch; compostable plastic; Bayesian optimization; heteroscedastic noise; material design; physical experiments; mechanical properties

check for updates

Citation: White, G.M.; Siegel, A.P.; Tovar, A. Optimizing Thermoplastic Starch Film with Heteroscedastic Gaussian Processes in Bayesian Experimental Design Framework. *Materials* **2024**, *17*, 5345. https://doi.org/10.3390/ma17215345

Academic Editors: Jacek Mateusz Bajkowski, Czesław I. Bajer, Jerzy Bajkowski and Hanna Piotrzkowska-Wróblewska

Received: 11 October 2024
Revised: 26 October 2024
Accepted: 27 October 2024
Published: 31 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The growing need to reduce plastic environmental pollution [1] and mitigate microplastic health risks [2] demands sustainable alternatives to traditional fossil-based plastics. One alternative is thermoplastic starch (TPS), a biodegradable material derived from natural polymers [3]. TPS has been used in applications like food packaging [4,5] and as a substitute for single-use plastics [6]. However, its mechanical properties are considerably inferior to those of fossil-based plastics [7]. For example, film-grade high-density polyethylene (HDPE), a fossil-based plastic commonly used in packaging, exhibits an elongation at break (EB) between 350% and 1700% and tensile strength (TS) ranging from 23.0 MPa to 89.2 MPa [8]. In contrast, typical TPS film shows EB values around 11% to 31%, and TS values around 1.03 MPa to 2.04 MPa [9,10].

Although the performance of TPS can be enhanced with additives such as Aloe vera [9], chitin nanocrystals [11], zinc oxide [12], nanoclay [13], agar [14], psyllium husk [15], and

different plasticizers [10,16–18], the development and optimization of new compositions is a lengthy and expensive process that typically involves trial-and-error and design of experiments (DOE) methods. While DOE methods such as grid searching and response surface can help us understand the effect of material parameters on its properties, they are often resource-intensive when used for optimization due to the large number of design evaluations required [19]. Experimental noise further complicates the process, especially when time and resource constraints prevent accurate uncertainty quantification. To overcome these challenges, adaptive experimental methods supported by BO are becoming increasingly prevalent as they offer more efficient solutions than traditional DOE methods.

1.1. Bayesian Experimental Design

Bayesian optimization (BO) is a gradient-free, global optimization method suitable for expensive-to-evaluate functions. It uses a probabilistic model, typically a Gaussian process (GP), to predict the function mean and variance (Section 2). It balances the exploration and exploitation of the design space via an acquisition function that selects the next design while minimizing the number of costly and time-consuming experiments (Section 3). In material design, BO offers a probabilistic framework that iteratively plans experiments based on prior results. This enables a more informed and efficient experimental process that guides decision-makers toward a globally optimal design [20–22].

In recent years, BO has gained attention in optimal material design and fabrication for its ability to efficiently navigate high-dimensional search spaces and identify optimal formulations and process parameters with fewer experimental evaluations [23,24]. Packwood [25] optimized the processing parameters of metal alloys and polymer blends, focusing on tuning processing conditions like temperature, pressure, and material composition. Yamashita et al. [26] optimized material atomic configurations, such as crystal, surface, and interface structures, and predicted the most stable structure for given chemical compositions. Talapatra et al. [27] introduced Bayesian objective under model uncertainty (BOMU) to explore the materials design space, accounting for resource constraints and model uncertainty. Liu et al. [28] introduced a BO approach for categorical and non-categorical variables for multimaterial vehicle structures. Xiong et al. [29] improved the mechanical properties of metals for additive manufacturing. Zhang et al. [30] accommodated mixed quantitative and qualitative variables and efficiently identified materials with desired dielectric characteristics. Zhang et al. [31] enhanced the precision and material properties in additive manufacturing, focusing on parameters like layer thickness, build speed, and material feed rate. Gao et al. [32] identified monomers and fabrication conditions of membranes for water purification. Iwama et al. [33] applied a Bayesian adaptive experimental design to optimize the operating conditions and reaction parameters in an ethylene oxide production plant. Valladares et al. [34] introduced a goal-based acquisition function that enables the parallel Bayesian optimization of lithium-ion battery cathode composition. Hickman et al. [35] fine-tuned reaction parameters such as temperature, pressure, and catalyst concentration to maximize reaction yield and selectivity. Guo et al. [36] improved high-throughput reaction screening by optimizing multiple reaction parameters simultaneously, allowing for faster identification of optimal CO_2 and green hydrogen production conditions. Qian et al. [37] leverage BO with the mean objective cost of uncertainty (MOCU) to optimize functional materials. Sattari et al. [38] incorporated physical constraints and process knowledge to improve process precision and mechanical performance of additively manufactured parts.

Although Bayesian optimization is emerging as a valuable tool for material design and has been successfully applied to optimize various material systems, its application in physical experiments is still in the early stages of development, with several challenges yet to be addressed. The challenges addressed in this work include the following: (1) Modeling heteroscedastic noise in physical experiments; (2) The adaptive exploration of the design space; (3) The application toward locating optimal TPS film formulations.

1.2. Heteroscedastic Noise

In most adaptive material design applications, the GP regression model supporting BO captures the underlying relationships between experimental parameters (design variables) and the material properties (objective function). However, traditional GP models assume that noise levels are uniformly distributed across the design space [39]. Since noise levels are not uniform across the design space in many real-world situations, the homeostatic noise assumption presents a challenge in experimental research, specifically relating to the field of materials science [40]. Heteroscedastic noise, which refers to variable levels of uncertainty across the design space, can significantly affect experimental outcomes and data interpretation. This can lead to inaccurate uncertainty quantification and statistical outcomes. Uncertainty in observations may even follow a random distribution, making heteroscedastic inferences analytically intractable [41].

Previous work has yielded methods to approximate the posterior noise variance using heteroscedastic Gaussian process (HGP) interfaces, with the gold standard being the Markov chain Monte Carlo (MCMC) method [42]; however, methods such as MCMC come at a high cost since they require many additional experimental design evaluations and time to evaluate [43]. Therefore, a practical HGP model is needed for the adaptive experimental design of materials within a BO framework.

The first contribution of this work is introducing a practical HGP model suitable for Bayesian experimental design (Section 2.3). The proposed HGP model uses two uncoupled GPs: one to model the mean of the material properties and one to model the variance (noise). Since the heteroscedastic noise function is modeled independently from the material properties, the mean GP model does not require prior noise information, making its implementation practical and efficient. The four statistical outcomes (two per GP) provide a simple and accurate way to estimate the expected value of the material properties and quantify their uncertainty.

1.3. Adaptive Exploration

Most BO implementations for physical experiments follow the same general steps as the ones used for computational experiments. In this configuration, an acquisition function identifies one promising design to be evaluated at a time based on the outputs of the GP (mean and variance). In a physical setting, evaluating one design per iteration causes experiments to advance unnecessarily slowly, especially considering situations where individual designs are timely and expensive to evaluate. Since many physical experiments inherently require time-consuming setups, considering more than one design candidate per iteration is manageable and more efficient. Therefore, there is a need for an approach to identify multiple candidate designs systematically.

Additionally, most implementations use fixed design space boundaries; however, too large boundaries can lead to unfeasible material designs, while boundaries that are too small may leave important regions unexplored. Thus, the boundaries must be adapted as the experiment progresses, allowing the exploration region to be refined based on prior experimental results.

The second contribution of this work is the application of two acquisition functions while managing the size of the exploration within a BO framework. The two acquisition functions are the lower confidence bound (LCB) and expected improvement (EI). The LCB function allows for increased diversity of new design candidates by promoting more exploration or exploitation than the traditional EI (Section 3.1). In addition, an exploration region management algorithm is implemented to dynamically adjust the size of the exploration region based on the inclusion of new designs in the sampling plan to ensure a more comprehensive modeling approach.

1.4. Optimal Formulation of TPS Film

BO has been applied successfully to optimize various material systems, including metal alloys, crystal structures, battery cathodes, and polymer blends. This work presents

the first application of BO for formulating TPS films. The specific objective of this work is to determine the optimal starch and plasticizer weight percentages (wt%) in the composition of TPS films to maximize elongation at break and tensile strength (Section 4). This marks the first use of BO in the design of TPS films.

In summary, the three main contributions of this work are as follows: (1) The introduction of a practical HGP model suitable for Bayesian experimental design; (2) An adaptive exploration of the design space based on acquisition functions and exploration region management; (3) The optimal TPS film formulation for the maximization of elongation at break and tensile strength. This work shows the potential of Bayesian experimental design in advancing the development of sustainable materials, with a particular focus on TPS films. By leveraging HGP modeling and adaptive exploration, this work contributes to the broader field of material science and offers a blueprint for future research in optimizing renewable materials.

2. Gaussian Process Regression

A Gaussian process (GP) is a collection of random variables such that every finite set of those variables has a multivariate Gaussian distribution. In the context of regression, a GP regression model can be defined as a distribution over functions with inference taking place directly in the space of functions [44]. A GP regression model defines a distribution of functions $f(\mathbf{x})$ such that, for any set of input points $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, the function values $\mathbf{f} = f(\mathbf{X}) = \{f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)\}$ follow a multivariate normal distribution [45]. Unlike traditional regression techniques that assume a specific functional form for the relationship, a GP regression model defines a distribution over possible functions that fit the data, allowing for a more adaptable and probabilistic modeling approach.

2.1. Gaussian Process Regression of Noiseless Data

A Gaussian process (GP) is fully specified by its mean function and covariance function (kernel). A random function $f(\mathbf{x})$ that follows a GP is expressed as

$$f(\mathbf{x}) \sim \mathcal{GP}\left(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\right), \tag{1}$$

where $m(\mathbf{x})$ is the mean function and $k(\mathbf{x}, \mathbf{x}')$ is the covariance function defined as

$$m(\mathbf{x}) = \mathbb{E}[f(x)], \tag{2}$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))], \tag{3}$$

respectively. The joint prior distribution of the observations $\mathbf{f} = \{f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)\}$ and the predictions $\mathbf{f}_*$ at the test locations $\mathbf{X}_*$ is

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{K}(\mathbf{X}, \mathbf{X}) & \mathbf{K}(\mathbf{X}, \mathbf{X}_*) \\ \mathbf{K}(\mathbf{X}_*, \mathbf{X}) & \mathbf{K}(\mathbf{X}_*, \mathbf{X}_*) \end{bmatrix}\right), \tag{4}$$

where $\mathbf{K}(\mathbf{X}, \mathbf{X})$ is the covariance matrix of the observed data, $\mathbf{K}(\mathbf{X}, \mathbf{X}_*)$ is the covariance matrix between the observed data and the predicted data, and $\mathbf{K}(\mathbf{X}_*, \mathbf{X}_*)$ is the covariance matrix of the predictions. The components of the covariance matrices are generated by the evaluation of the covariance function $k(\mathbf{x}, \mathbf{x}')$. Valid covariance functions produce positive semi-definite matrices regardless of the chosen pair of points $(\mathbf{x}, \mathbf{x}')$ [46].

2.2. Gaussian Process Regression of Noisy Data

In modeling problems involving experimental observations, the data do not correspond to observations of $f(\mathbf{x})$, but noisy values, $y = f(\mathbf{x}) + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$. Given a set of noisy observations, $\mathbf{y} = \{y^{(1)}, \ldots, y^{(n)}\}$, the joint prior distribution of $\mathbf{y}$ and the predictions $\mathbf{f}_*$ at the test locations $\mathbf{X}_*$ is

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{K}(\mathbf{X},\mathbf{X}) + \sigma_n^2\mathbf{I} & \mathbf{K}(\mathbf{X},\mathbf{X}_*) \\ \mathbf{K}(\mathbf{X}_*,\mathbf{X}) & \mathbf{K}(\mathbf{X}_*,\mathbf{X}_*) \end{bmatrix}\right), \tag{5}$$

where σ_n^2 is the noise in the data (under the assumption of independent Gaussian noise), and $\mathbf{I}$ is the identity matrix.

In most cases, the covariance function $k(\mathbf{x},\mathbf{x}')$ is flexible enough to fit the data, and the mean function is often assumed to be zero, $m(\mathbf{x}) = 0$ [47]. Then, the predictive equations of a zero-mean GP regression model are obtained by conditioning Equation (5) on the observed data by setting the following condition:

$$\mathbf{f}_* \mid \mathbf{X}, \mathbf{y}, \mathbf{X}_* \sim \mathcal{N}\big(\bar{\mathbf{f}}_*, \mathrm{cov}(\mathbf{f}_*)\big). \tag{6}$$

where $\bar{\mathbf{f}}_*$ and $\mathrm{cov}(\mathbf{f}_*)$ are the mean and the covariance of the prediction, which are given by the posterior distribution mean and variance:

$$\bar{\mathbf{f}}_* = \mathbf{K}(\mathbf{X}_*,\mathbf{X})[\mathbf{K}(\mathbf{X},\mathbf{X}) + \sigma_n^2\mathbf{I}]^{-1}\mathbf{y}, \tag{7}$$

$$\mathrm{cov}(\mathbf{f}_*) = \mathbf{K}(\mathbf{X}_*,\mathbf{X}_*) - \mathbf{K}(\mathbf{X}_*,\mathbf{X})[\mathbf{K}(\mathbf{X},\mathbf{X}) + \sigma_n^2\mathbf{I}]^{-1}\mathbf{K}(\mathbf{X},\mathbf{X}_*). \tag{8}$$

where the diagonal of $\mathrm{cov}(\mathbf{f}_*)$ is the predictive variance [44]. Since the mean function becomes irrelevant, the covariance function (kernel) defines the characteristics of the prediction. Notably, the covariance prediction does not require the observations $\mathbf{y}$ because it is a property of the GP model and the distribution of the inputs only, regardless of the observed values.

2.3. Heteroscedastic Gaussian Process Regression

Heteroscedastic Gaussian Process (HGP) regression is an extension of traditional GP regression that deals with situations where the noise $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$ is not constant across the input space. While standard GP regression assumes a constant σ_n^2 across all inputs, HGP regression allows noise variance to change depending on the input, making it more flexible and accurate for physical world scenarios where uncertainty is not uniform.

The model is structured in HGP regression with two main components: a mean function and a variance function drawn from different Gaussian processes. This is shown below:

$$f_1(\mathbf{x}) \sim \mathcal{GP}\big(0, k_1(\mathbf{x},\mathbf{x}')\big) \tag{9}$$

$$f_2(\mathbf{x}) \sim \mathcal{GP}\big(0, k_2(\mathbf{x},\mathbf{x}')\big)\,. \tag{10}$$

where f_1 follows a mean function GP that captures the underlying relationship between inputs $\mathbf{x}$ and outputs $\mathbf{y}$, while f_2 follows the noise variance function GP that captures the relationship between inputs $\mathbf{x}$ and the noise variance σ_n^2. In this work, the HGP model outputs are given by the following:

$$\bar{\mathbf{f}}_{1*} = \mathbf{K}(\mathbf{X}_*,\mathbf{X})\mathbf{K}(\mathbf{X},\mathbf{X})^{-1}\mathbf{y} \tag{11}$$

$$\mathrm{cov}(\mathbf{f}_{1*}) = \mathbf{K}(\mathbf{X}_*,\mathbf{X}_*) - \mathbf{K}(\mathbf{X}_*,\mathbf{X})\mathbf{K}(\mathbf{X},\mathbf{X})^{-1}\mathbf{K}(\mathbf{X},\mathbf{X}_*) + \mathrm{diag}(\mathbf{f}_{2*}) + \mathrm{cov}(\mathbf{f}_{2*}) \tag{12}$$

$$\bar{\mathbf{f}}_{2*} = \mathbf{K}(\mathbf{X}_*,\mathbf{X})[\mathbf{K}(\mathbf{X},\mathbf{X}) + \lambda\mathbf{I}]^{-1}\boldsymbol{\sigma}_n^2 \tag{13}$$

$$\mathrm{cov}(\mathbf{f}_{2*}) = \mathbf{K}(\mathbf{X}_*,\mathbf{X}_*) - \mathbf{K}(\mathbf{X}_*,\mathbf{X})[\mathbf{K}(\mathbf{X},\mathbf{X}) + \lambda\mathbf{I}]^{-1}\mathbf{K}(\mathbf{X},\mathbf{X}_*)\,. \tag{14}$$

where $\bar{\mathbf{f}}_{1*}$ and $\mathrm{cov}(\mathbf{f}_{1*})$ are the mean and the covariance of the function prediction, $\bar{\mathbf{f}}_{2*}$ and $\mathrm{cov}(\mathbf{f}_{2*})$ are the mean and the covariance of the noise variance, and $\mathrm{diag}(\mathbf{f})$ is the diagonal matrix constructed from the vector $\mathbf{f}$. Here, σ_n^2 is the noise variance, and λ is a trainable hyperparameter representing its variance—the variance of the noise variance.

2.4. Covariance Function

The covariance function $k(\mathbf{x}, \mathbf{x}')$ encodes assumptions about the smoothness, periodicity, and other properties of the function to be modeled. The choice of covariance function is crucial in GP regression as it defines the shape and properties of the functions that the GP can model. Commonly used covariance functions include the squared exponential kernel, the Matérn kernel, and the periodic kernel.

The squared exponential kernel, also referred to as the radial basis function (RBF), is defined as

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\ell^2}\right). \tag{15}$$

where σ_f^2 is the signal variance, and ℓ is the length scale, controlling the smoothness of the function. Similarly, the Matérn kernel is given by

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\|\mathbf{x} - \mathbf{x}'\|}{\ell}\right)^{\nu} K_\nu \left(\frac{\sqrt{2\nu}\|\mathbf{x} - \mathbf{x}'\|}{\ell}\right). \tag{16}$$

where $K_\nu()$ is a modified Bessel function and $\Gamma(\nu)$ is the Gamma function [44]: for a positive integer ν, $\Gamma(\nu) = (\nu - 1)!$. Here, ν controls the smoothness, and the Gamma function $\Gamma(\nu)$ normalizes the kernel function and ensures that it correctly describes the covariance structure of the process. The Matérn covariance function becomes simpler when ν is a half-integer, such as $1/2, 3/2, 5/2, 7/2$, and so on. For $\nu = 1/2$, the process becomes very rough, while $\nu \geq 7/2$ may be indistinguishable from $\nu \to \infty$ [44]. The most common cases in GP regression are $\nu = 3/2$ and $\nu = 5/2$, resulting in

$$k_{v=3/2}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \left(1 + \frac{\sqrt{3}\|\mathbf{x} - \mathbf{x}'\|}{\ell}\right) \exp\left(-\frac{\sqrt{3}\|\mathbf{x} - \mathbf{x}'\|}{\ell}\right), \tag{17}$$

$$k_{v=5/2}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \left(1 + \frac{\sqrt{5}\|\mathbf{x} - \mathbf{x}'\|}{\ell} + \frac{5\|\mathbf{x} - \mathbf{x}'\|^2}{3\ell^2}\right) \exp\left(-\frac{\sqrt{5}\|\mathbf{x} - \mathbf{x}'\|}{\ell}\right), \tag{18}$$

respectively. Finally, a periodic kernel can be defined as

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{2\sin^2(\pi\|\mathbf{x} - \mathbf{x}'\|/p)}{\ell^2}\right). \tag{19}$$

where p is the period of the function.

2.5. Training

Training involves finding the optimal values of the unknown kernel parameters, such as σ_f^2, ℓ, and p, which is crucial for achieving good model performance. This task can be accomplished by maximizing the likelihood of the observed data,

$$L(\theta \mid \mathbf{y}) = p(\mathbf{y} \mid \mathbf{X}, \theta) = \mathcal{N}(\mathbf{y} \mid \mathbf{0}, \mathbf{K}(\mathbf{X}, \mathbf{X}; \theta) + \sigma_n^2 \mathbf{I}), \tag{20}$$

where $\mathbf{K}(\mathbf{X}, \mathbf{X}; \theta)$ is the covariance matrix defined by the kernel function with hyperparameters θ. To facilitate the optimization process, the log-likelihood is often used instead of the likelihood, Equation (20) [44]. This log-likelihood is given by

$$\log p(\mathbf{y} \mid \mathbf{X}, \theta) = -\frac{1}{2}\mathbf{y}^T \left(\mathbf{K}(\mathbf{X}, \mathbf{X}; \theta) + \sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{y} - \frac{1}{2}\log\left|\mathbf{K}(\mathbf{X}, \mathbf{X}; \theta) + \sigma_n^2 \mathbf{I}\right| - \frac{n}{2}\log 2\pi. \tag{21}$$

This function can be maximized using a gradient-based algorithm such as L-BFGS-B (limited-memory Broyden–Fletcher–Goldfarb–Shanno with box constraints). The implementation in this work uses the Python library `gpflow` version 2.9.1 (https://www.gpflow.org/, accessed on 10 September 2024).

2.6. Heteroscedastic Gaussian Process Regression Algorithm

The steps involved in generating the posterior mean and variance of a given set of observed data from a noisy (unknown) function are the following:

Step 1: Input points. Define the input points $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ using a sampling plan. This set of points can be selected from a predefined grid, a traditional DOE, or quasi-randomly.

Step 2: Observations. Evaluate the noisy function at every $\mathbf{x}_i$ a given number of times to obtain the corresponding observations $\mathbf{y}_i$. For example, one can evaluate the noisy function five times at each input point.

Step 3: HGP Regression. At each input point, obtain the observed data's mean and variance. Use these values to obtain the posterior mean and variance as defined by (11) and (12).

2.7. Numerical Example

Let us consider a noisy version of Forrester's test function defined by

$$f(x) = (6x - 2)^2 \sin(12x - 4) + \epsilon(x)\,, \tag{22}$$

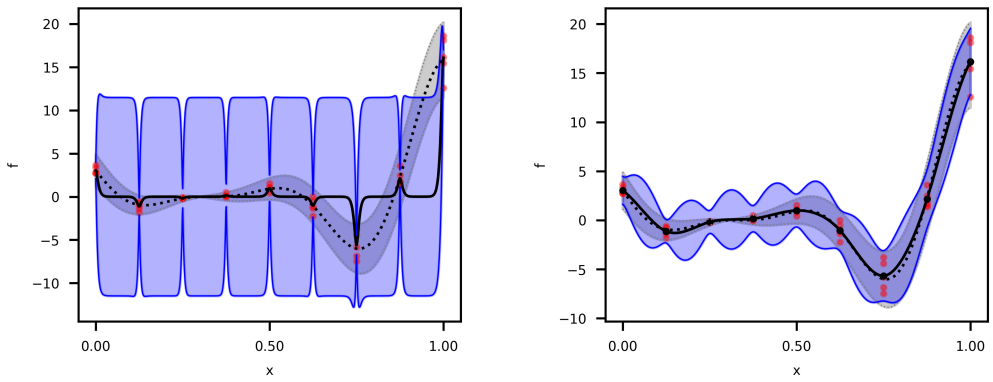
defined in the interval $x \in [0, 1]$ with heteroscedastic Gaussian noise $\epsilon(x) \sim \mathcal{N}(0, \sigma^2(x))$ with variance

$$\sigma^2(x) = 10(x - 0.3)^2\,. \tag{23}$$

Let us consider the sampling plan

$$\mathbf{X} = [0, 1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8, 1]\,,$$

with five random observations per point. Figure 1 shows the comparison of the vanilla GP regression from (7) and (8), and the HGP regression from (11) and (12). Both models utilized the Matérn 3/2 kernel (17). Since the vanilla GP cannot capture the noise variance, it fits a constant value across the design space, increasing the prediction's uncertainty. The predicted mean significantly differs from the true mean of the function. On the other hand, the HGP better fits both the variance and the mean of the true function.



(**a**) Vanilla GP (**b**) Heteroscedastic GP

Figure 1. Gaussian process regressions of the noisy Forrester's function. The red dots represent the random samples. The dotted line and gray-shaded area correspond to the 95% confidence interval. The black line and blue-shaded area correspond to the prediction based on the posterior distributions for the following: (**a**) The vanilla GP; (**b**) The HGP.

3. Bayesian Optimization

Bayesian optimization (BO) is an efficient method for optimizing expensive black-box functions [21,22]. Unlike traditional optimization techniques, which often rely on gradient information or extensive sampling of the parameter space, BO leverages a probabilistic model to guide the search for optimal solutions. This approach is particularly well-suited for applications where each function evaluation is costly in terms of time or resources [48].

In this work, BO begins by modeling the noisy objective function $f(\mathbf{x})$ using a HGP regression (Section 2.6). The HGP predicts the posterior mean (11) and variance (12) of $f(\mathbf{x})$. Then, an acquisition function is defined to determine the next points to sample based on the HGP outcomes (Section 3.1). The function is evaluated in the new points, the HGP model is updated, and this process continues until no improvement is possible. The next sections explain in more detail the acquisition functions implemented in this work, the exploration region management, and the convergence criteria of the BO algorithm.

3.1. Acquisition Functions

This work implements two acquisition functions: lower confidence bound (LCB) and the expected improvement (EI). Both functions are designed to balance exploration and exploitation by considering the GP model's predicted mean function $\mu(\mathbf{x})$ defined from (11) and the predicted standard deviation $\sigma(\mathbf{x})$ obtained from (12).

Lower Confidence Bound. The Lower Confidence Bound (LCB) is an acquisition function defined for minimizing a function $f(\mathbf{x})$. The LCB for a given input $\mathbf{x}$ is defined by:

$$\mathrm{LCB}(\mathbf{x}) = \mu(\mathbf{x}) - \kappa\,\sigma(\mathbf{x})\,, \tag{24}$$

where κ is a positive parameter that controls the trade-off between exploration ($\kappa \to \infty$) and exploitation ($\kappa = 0$). If the objective is to minimize $f(\mathbf{x})$, then the new point to be added to the sampling plan is

$$\mathbf{x}_{\mathrm{new}} = \arg\min \mathrm{LCB}(\mathbf{x})\,. \tag{25}$$

In the proposed algorithm, the value of κ is incrementally increased until $\mathbf{x}_{\mathrm{new}}$ is different than the one from maximizing the expected improvement.

Expected Improvement. Given the mean best-observed value of the function so far $f(\mathbf{x}_{\mathrm{best}})$ of the function $f(\mathbf{x})$ to be minimized, the Expected Improvement (EI) at a given point $\mathbf{x}$ is defined as:

$$\mathrm{EI}(\mathbf{x}) = \mathbb{E}[\max(0, f(\mathbf{x}) - f(\mathbf{x}_{\mathrm{best}}))]\,. \tag{26}$$

When the function $f(\mathbf{x})$ is modeled as a GP, the EI function can be expressed as:

$$\mathrm{EI}(\mathbf{x}) = (\mu(\mathbf{x}) - f(\mathbf{x}_{\mathrm{best}}))\Phi\left(\frac{\mu(\mathbf{x}) - f(\mathbf{x}_{\mathrm{best}})}{\sigma(\mathbf{x})}\right) + \sigma(\mathbf{x})\phi\left(\frac{\mu(\mathbf{x}) - f(\mathbf{x}_{\mathrm{best}})}{\sigma(\mathbf{x})}\right), \tag{27}$$

where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution, and $\phi(\cdot)$ is the probability density function (PDF) of the standard normal distribution. Then, the new point to be added to the sampling plan is

$$\mathbf{x}_{\mathrm{new}} = \arg\max \mathrm{EI}(\mathbf{x})\,. \tag{28}$$

3.2. Exploration Region Management

The exploration region is defined as

$$\Gamma = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}_{\mathrm{LB}} \leq \mathbf{x} \leq \mathbf{x}_{\mathrm{UB}}\}\,, \tag{29}$$

where $\mathbf{x}_{\text{LB}}$ and $\mathbf{x}_{\text{UB}}$ are the lower and upper boundaries, respectively. The boundaries are defined as a function of the distance between the designs $\mathbf{x}$ in Γ along every dimension $k = 1, \ldots, n$. This is

$$x_k^{\text{UB}} - x_k^{\text{LB}} = r\left(x_k^{\max} - x_k^{\min}\right), \tag{30}$$

where $(x_k^{\max}$ and $x_k^{\min}$ are the maximum and minimum values along the dimension k and $r > 1$ is the expansion ratio. In this work, $r = 1.5$ by default. The corners of the exploration region Γ are also added to the sampling plan. Hence, the exploration grows by a factor r in every iteration until it reaches absolute limits, such as a fraction $x_k = 0$ or $x_k = 1$ of a given ingredient x_k.

3.3. Convergence Criterion

Soft convergence is achieved when there is no significant change in the value of $\max \text{EI}(\mathbf{x})$ and $\mathbf{x}_{\text{best}}$ in three consecutive iterations. This is shown below:

$$\Delta^3 \max \text{EI}(\mathbf{x}) \leq \epsilon_e \tag{31}$$

$$\Delta^3 \mathbf{x}_{\text{best}} \leq \epsilon_x\,, \tag{32}$$

where ϵ_e and ϵ_x are small quantities. Hard convergence occurs when the allotted number of experiments has been exhausted.

3.4. Bayesian Optimization Algorithm

The steps involved in the proposed Bayesian optimization (BO) approach are summarized in (Figure 2) and include the following steps:

Step 1: Objective Function Regression. BO begins by modeling the objective function $f(\mathbf{x})$ using a HGP regression. The HGP predicts the posterior mean (11) and variance (12) of the noisy objective function $f(\mathbf{x})$ as a distribution over possible functions that fit the observed noisy data $\mathbf{y}$ from the sampling plan $\mathbf{X}$.

Step 2: Acquisition Function Optimization. An acquisition function is defined to determine the next point to sample based on the HGP posterior mean and variance. In this work, the new points to be added to the sampling plan correspond to the solutions of (25) and (28).

Step 3: Check Convergence. If the convergence conditions (31) and (32) are satisfied, and the exploration region Γ has reached its maximum size, then the $\arg\min \mu(\mathbf{x})$ is the optimal design $\mathbf{x}^*$. Otherwise, the process continues by adding $\mathbf{x}_{\text{new}}$ to the sampling plan.

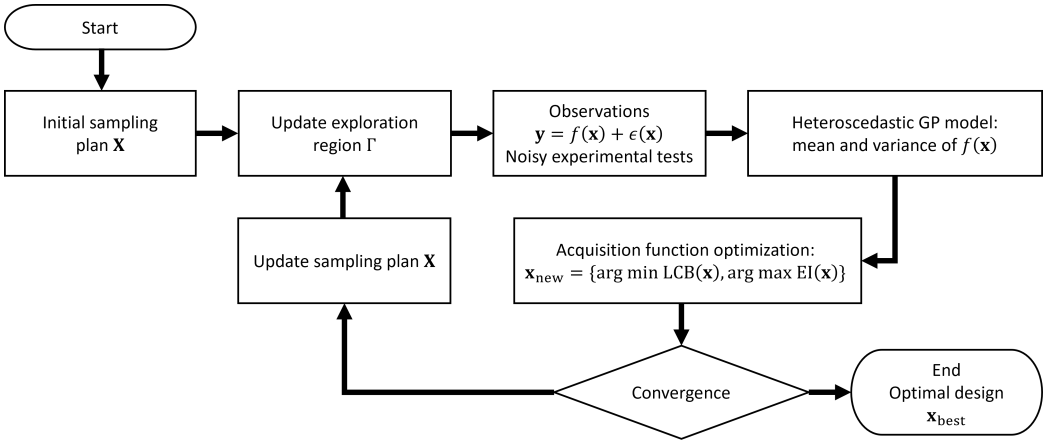


Figure 2. BO adaptive experimental design algorithm implementing an HGP surrogate model and an exploration region.

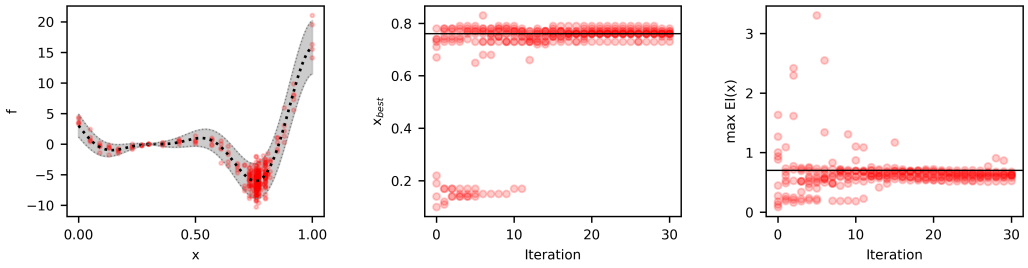
3.5. Numerical Results

To illustrate the implementation of the BO algorithm, let us consider two benchmark problems: Forrester's function (1D) and the three-hump camel function (2D).

Forrester's function (1D). The first problem is the minimization of the noisy Forrester's function defined in (22). The global minimizer of this function is $x^* \approx 0.7572$. For this study, let us consider random initial sampling plans with five random designs. To replicate typical physical experiments, the data are defined as

$$\mathbf{X} = \{x_r - 2\delta, x_r - \delta, x_r, x_r + \delta, x_r + 2\delta\},$$

where x_r and δ are random numbers such that all x in $\mathbf{X}$ is in the interval $[0, 1]$. The minimum distance between adjacent designs is set to 0.01. The results are summarized in Figure 3. These results show that, in most cases, convergence can be achieved in about five iterations.



(**a**) Function evaluations (**b**) Best evaluated design (**c**) Expected improvement

Figure 3. Bayesian optimization of the noisy Forrester's function. Results summarize ten optimization algorithm runs. Each optimization algorithm was initialized with five random initial designs and ran for 30 iterations. The red dots represent the designs evaluated during all the optimization runs. (**a**) The black dotted line is the mean of the true function, and the gray shaded area is the 95% confidence interval. Most of the function evaluations were around the minimizer. (**b**) The algorithm usually finds the minimizer or a close design in less than five iterations. The black solid line represents the minimizer $x^* \approx 0.7572$. (**c**) Accordingly, the expected improvement also remains at a constant value after the fifth iteration.

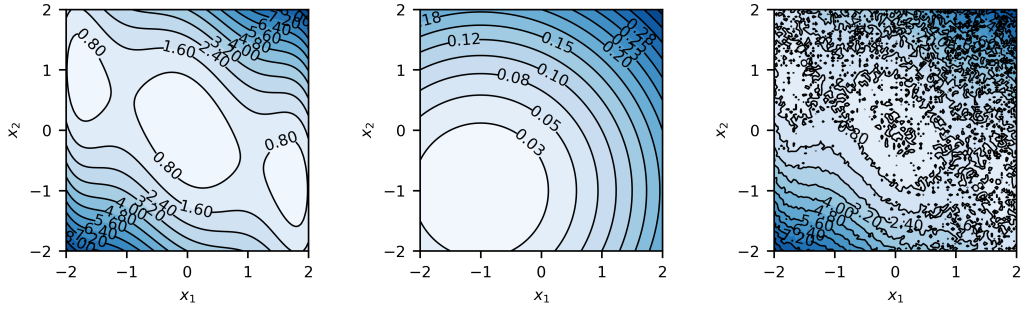
Three-hump camel function (2D). The second problem is the minimization of the noisy three-hump camel function:

$$f(x_1, x_2) = 2x_1^2 - 1.05x_1^4 + \frac{x_1^6}{6} + x_1x_2 + x_2^2 + \epsilon(x_1, x_2), \tag{33}$$

defined in the region $-2 \leq x_1 \leq 2$ and $-2 \leq x_2 \leq 2$, with heteroscedastic Gaussian noise $\epsilon(x_1, x_2) \sim \mathcal{N}(0, \sigma^2(x_1, x_2))$ with variance

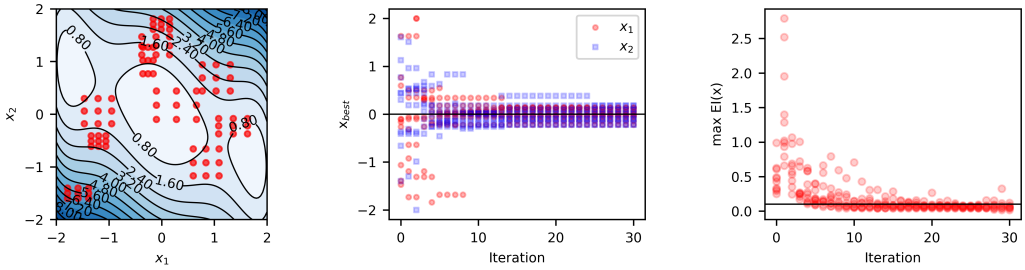
$$\sigma^2(x_1, x_2) = \frac{1}{50}\left[(x_1 + 1)^2 + (x_2 + 1)^2\right]. \tag{34}$$

The global minimizer of this function is $\mathbf{x}^* = [0, 0]$. For this study, let us consider initial sampling plans with nine random designs, $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_9\}$, as shown in Figure 4. Results summarized in Figure 5 show that in the case of 2D, the algorithm usually reaches convergence in about ten iterations.



(**a**) Noiseless function (**b**) Noise variance (**c**) Noisy function

Figure 4. Three-hump camel function prior to optimization. Heteroscedastic noise and variance are to be added to the black box function to simulate physical experimental conditions. (**a**) Noiseless three-hump camel function. (**b**) Heteroscedastic noise variance of the three-hump camel function. (**c**) Noisy three-hump camel function to be optimized.



(**a**) Random initial designs (**b**) Best evaluated design (**c**) Expected improvement

Figure 5. Bayesian optimization of the noisy three-hump camel function. Results summarize ten optimization algorithm runs utilizing an exploration region for each model iteration. Each algorithm was initialized with nine random design points and ran for 30 iterations. The red dots represent the random initial designs. (**a**) The initial designs prior to noisy observations. Most of the initial function evaluations were around local minima, with only one near the minimizer. (**b**) With noisy observations in two dimensions, the algorithm usually finds the minimizer or a close design in less than ten iterations. The black solid line indicates the minimizer values, $x_1^* = 0$ and $x_2^* = 0$. (**c**) The maximum expected improvement remains about constant after the tenth iteration.

4. Thermoplastic Starch Film Optimization

The proposed Bayesian experimental design is applied to the formulation of TPS film, aiming to optimize the plasticizer (x_1) and potato starch (x_2) weight percentages to maximize the film's elongation at break (y_1) and tensile strength (y_2). Given the conflicting effects of plasticizer and starch content on the mechanical properties of TPS films and the difference in their noise effect, this application provides a template for tackling similar material optimization challenges.

4.1. Materials

The materials used in this study were sourced as follows: potato starch was obtained from Gefen Foods (Bayonne, NJ, USA); distilled white vinegar, 5 (m/v)% acetic acid, from Walmart's Great Value (Bentonville, AR, USA); distilled water from Sigma-Aldrich (Burlington, MA, USA); and vegetable glycerol from Florida Laboratories (Fort Lauderdale, FL, USA).

Polymers are sourced from potato starch, and glycerol:vinegar mixture (in a 2:1 ratio) is utilized as the plasticizer. The acetic acid in the vinegar acts as an acid catalyst for facilitating the plasticization process [10]. Distilled water is used as a solvent to maintain a homogeneous mixture. It also acts as an additional plasticizing agent in the plasticization of TPS [49].

All ingredients utilized in this study adhere to the Food Chemical Codex (FCC) grade standards, ensuring their compostability. The TPS formation process outlined in Section 4.1 is designed to preserve the compostability characteristics of these ingredients and the TPS films.

4.2. Synthesis Protocol

The synthesis protocol for all experimental compositions is as follows:

Mixing. Wet ingredients (glycerol, vinegar, distilled water) are first combined in a 600 mL beaker. Potato starch is sifted using a 60-mesh sieve to ensure uniform granules, and is added while stirring continuously to achieve a homogeneous mixture. For all experimental compositions, the total initial mass of the mix is maintained at 120 g.

Gelatinization. The beaker containing the solution is placed on an electric hot plate and brought to the gelatinization temperature at around 80 °C at a heating rate of about 10 °C/min while stirring consistently. At the gelatinization point, the TPS color changes from milky to colorless and becomes more viscous. At this point, the beaker is taken off heat and cooled to about 60 °C.

Degassing. The TPS mixture is then centrifuged for two minutes at 2000 rpm to remove bubbles (Figure 6). The remaining mass of the TPS solution usually ranges between 72 and 78 g at this point.

Drying. The TPS mixture is poured into Petri dishes (12 g per Petri dish) and left to air dry in controlled laboratory ventilated air at room temperature (20 to 23 °C) for 72 h.

Film preparation. Finally, the dried film is detached from the Petri dish, and testing specimens are prepared for mechanical testing (Section 4.3).

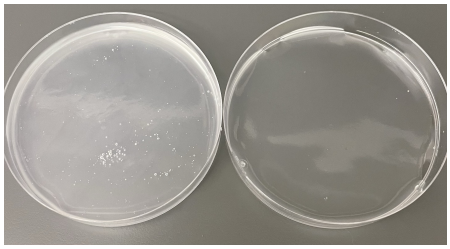


Figure 6. Centrifuging of TPS slurry effectively degases solution before air drying, leading to higher quality film. **Left**: 12 g of TPS slurry in a Petri dish without centrifuging. The solution is cloudy with many air bubbles. **Right**: 12 g of TPS slurry in a Petri dish after being centrifuged for two minutes at 2000 rpm. The solution is degassed and translucent.

4.3. Tensile Testing

TPS film samples were prepared to fit coupon-sized rectangles with a nominal dimension of 88 mm long by 13 mm wide, according to ASTM standard D638 [50]. Film thickness was measured with a Mitutoyo Digimatic digital micrometer with a 0.5 µm resolution. The specimen's width was measured using a Mitutoyo Digimatic digital caliper with a 0.01 mm resolution. Three measurements were taken for each film sample and then averaged to estimate the film's cross-sectional area. Tensile strength and elongation at break were determined by tensile testing on a universal testing machine (UTM) from Jinan Focus Test Instrument Co., Ltd. (Jinan City, Shandong, China). Five samples per formulation were tested, EB and TS were recorded, and the mean and variance of each property were determined.

4.4. Optimization

The iterative steps of the BO process, summarized in Figure 2, are the following:

Initial sampling plan. The initial sampling plan used in this work follows a full-factorial design of experiments (DOE) with two factors, x_1 and x_2, each at three levels, resulting in a sampling plan of nine points.

Update exploration region. The initial size of the exploration is 1.5 times larger than the initial DOE grid. The exploration region is updated dynamically, and when new experimental data are added, the region boundaries are adjusted based on the new sampling plan (Section 3.2).

Observations. One observation corresponds to the evaluation of a mechanical property (Section 4.3). After each new experiment, the observed data for the EB and TS properties, y_1 and y_2, and corresponding compositions, x_1 and x_2, are recorded. These observations provide the data to train the HGP model.

HGP model. Two GP models are trained for each property (y_1 and y_2) for a total of four GP models: mean(Y_1), var(Y_1), mean(Y_2), and var(Y_2). Each GP model has two outputs: mean and variance; therefore, the total number of outputs of the HGP model is eight: mean(mean(y_1)), var(mean(y_1)), mean(var(y_1)), var(var(y_1)), mean(mean(y_2)), var(mean(y_2)), mean(var(y_2)), and var(var(y_2)). The HGP model updates continuously with new experimental results, refining the predictions and uncertainty quantification predictions and providing data to the acquisition functions.

Acquisition function. The acquisition functions, LCB (25) and EI (28), guide the selection of new sampling points based on the HGP model's outputs. In each iteration, four new points are located: one from EI and one from LCB point for each targeted mechanical property. The four corners of the exploration region Γ are evaluated to promote broader exploration. While up to eight distinct points can be added per iteration, in practice, some points may overlap or already be included in the sampling plan, resulting in fewer new additions.

Convergence. The convergence criteria are satisfied when there is minimal improvement in material properties over successive iterations or when a predefined number of experiments is met (Section 3.3). At the point of convergence, an optimal or near-optimal solution within the explored region has been found. If the convergence is reached, the optimal design corresponds to $\mathbf{x}_{\text{best}}$. This application will have one optimal design for each material property y_1 and y_2.

Update sampling plan. If the convergence criteria are not satisfied, the new designs are tested, and the sampling plan is updated. The BO process iterates until a convergence criterion is met.

5. Results and Discussion

Five experimental iterations were evaluated. Each iteration takes roughly one week to complete. The TPS elongation at break (EB) and tensile strength (TS) are affected by the plasticizer (glycerol:vinegar) (x_1) and starch content (x_2) in a nonlinear manner (Figure 7). Specifically, higher content of plasticizer (up to about 4.5 wt%) and low starch content (down to about 2.0 wt%) maximize EB. In contrast, low content of plasticizer (down to about 0.5 wt%) and medium-level content of starch (about 7.0 wt%) maximize TS. Notably, an excessive amount of plasticizer (above 6.0 wt%) or excessive amount of starch (above 12.0 wt%) has a negative effect on the overall structural stability of the film and its mechanical properties, leading to increased film brittleness and decreased EB. Figure 8 displays the variations in film quality with varying amounts of starch and plasticizer content. Furthermore, our approach allows quantifying the uncertainty of the response as a function of the material content (Figure 9). For EB, higher variance is observed in specimens with higher content of both plasticizer and starch. The variance function is more flat for TS but increases with the starch content.

As summarized in Table 1, the optimal formulation for the highest EB was found in the last iteration to be 4.5 wt% plasticizer, 2.0 wt% starch, and 93.5 wt% water (Figure 9). The corresponding optimal TPS contains 3.6 g glycerol, 1.8 g vinegar, 2.4 g starch, and 112.2 g water. Compared to the EB value of about 11% to 31% reported in previous studies [9,10], this composition leads to a higher EB value (M = 96.7%, SD = 5.6%), which represents an increase of about 212%.

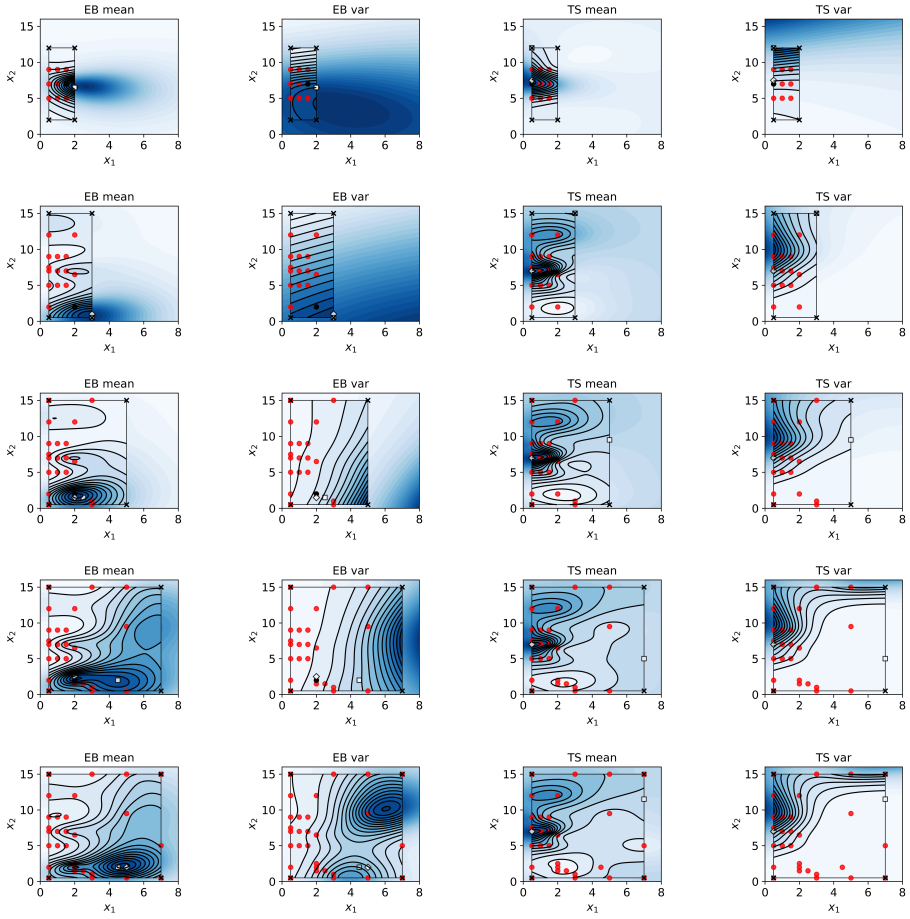


Figure 7. Results from HGP predictions from TPS DOEs for five experimental iterations. As in the numerical experiments, an initial design was evaluated, and an exploration region was initialized. BO guided the subsequent experiments until formulations yielding optimal mechanical properties were achieved. The red dots represent the design points evaluated in each iteration. HGP models were used to predict the mean and variance of EB and TS for varying compositions of plasticizer (x_1) and starch content (x_2). The far left column displays HGP predictions for EB, with the second column showing the predicted variance for EB. The next column shows the mean TS predictions, with the final column showing the predicted variance for TS.

On the other hand, the optimal formulation for the highest TS was found in the first experimental iteration to be 0.5 wt% plasticizer, 7.0 wt% starch, and 93.5 wt% water (Figure 9 and Table 2). The corresponding optimal TPS contains 0.4 g glycerol, 0.2 g vinegar, 8.4 g starch, and 111.0 g water. Compared to the TS value of about 1.03 MPa to 2.04 MPa from previous studies [9,10], this formulation leads to around a higher TTS value (M = 2.77 MPa, SD = 1.54 MPa), which represents a 38% increase.

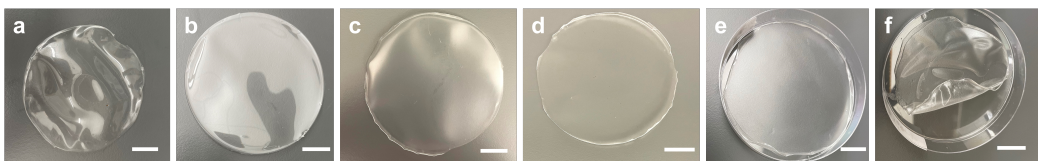
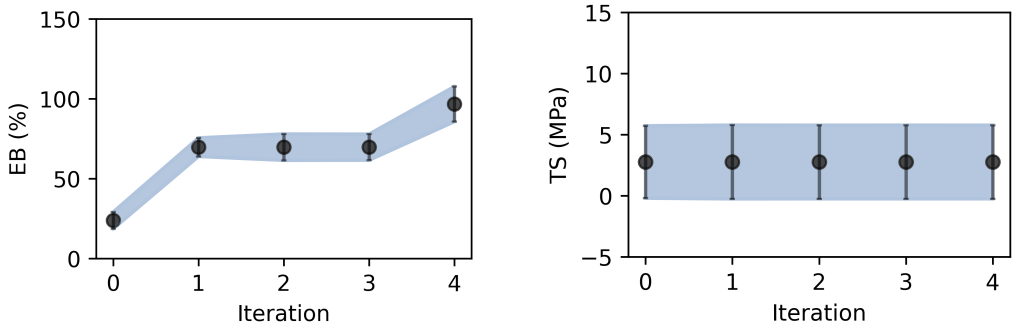


Figure 8. TPS film: Increased starch content with decreased plasticizer content to decreased starch content with increased plasticizer content. TPS film with increased and decreased starch content led

to brittleness and warping. Similarly, decreased starch and increased plasticizer content also caused film warping and moisture. (**a**) The film specimen with the highest starch content and least plasticizer content. The film is warped and extremely brittle, with low mechanical properties. (**b**) The film specimen with increased plasticizer content. The film is less brittle and exhibits optimal TS. (**c**) The film specimen has roughly equal parts of plasticizer and starch content. Mechanical properties are suitable for both EB and TS. (**d**) The film specimen with slightly more plasticizer than starch. EB is increased, but subsequently, the TS begins to diminish. (**e**) Increased plasticizer leads to slight warping, films retain more moisture and exhibit optimal EB, and diminished TS. (**f**) Film specimen with the highest plasticizer concentration and the lowest starch concentration. Films are extremely deformed and moist. TS is diminished, and EB begins to decrease with worsening film quality.



(**a**) Best observed EB per iteration.

(**b**) Best observed TS per iteration.

Figure 9. Best observed properties. The black dots represent the mean best-observed property per experimental observation. The vertical bar and blue-shaded region indicate the 95% confidence interval. (**a**) The optimal formulation for maximizing EB was found after four iterations, with increased plasticizer concentrations leading to greater EB variability. (**b**)The optimal formulation for TS was found after the first iteration, with relatively constant variability.

Table 1. Summary of iterative results for elongation at break (EB). Results include the plasticizer and starch composition of the best design, the EB mean and standard deviation, the expected improvement at the end of the iteration, and the cumulative number of experiments.

Iter	Plast (wt%)	Starch (wt%)	EB (%)	EI	Exp
1	1.5	7.0	23.9 ± 2.7	6.06	9
2	2.0	2.0	69.8 ± 3.0	32.5	15
3	2.0	2.0	69.8 ± 4.2	2.97	20
4	2.0	2.0	69.8 ± 4.2	2.23	25
5	4.5	2.0	96.7 ± 5.6	3.16	30

Table 2. Summary of iterative results for tensile strength (TS). Results include the plasticizer and starch composition of the best design, the TS mean and standard deviation, the expected improvement at the end of the iteration, and the cumulative number of experiments.

Iter	Plast (wt%)	Starch (wt%)	TS (MPa)	EI	Exp
1	0.5	7.0	2.77 ± 1.51	0.615	9
2	0.5	7.0	2.77 ± 1.54	0.609	15
3	0.5	7.0	2.77 ± 1.54	0.605	20
4	0.5	7.0	2.77 ± 1.54	0.604	25
5	0.5	7.0	2.77 ± 1.54	0.604	30

Plasticizing agents, extracted starch, and acid catalysts significantly affect the mechanical, structural, and physical properties of TPS films [18]. Additional factors, such as

the time and temperature of film preparation, may also affect the material's mechanical properties. Previous studies have investigated the effects of various plasticizing agents such as choline chloride and urea [10], xylitol and sorbitol [51], and glycerol-sorbitol [16]. The most commonly utilized plasticizer in the fabrication of TPS films is glycerol/water [49]. The presence of low-weight amide plasticizing structures in the addition of water, such as glycerol, has been shown to perform significantly better at suppressing starch retrogradation as opposed to other plasticizing agents [17]. Additionally, other low-weight organic acids such as acetic acid or citric acid [52] further contribute to suppressing starch reintegration, effectively reducing the material's degradation rate while also acting as a catalyst for plasticization [53].

Increased amounts of glycerol in relation to starch tend to amplify variations in the observed mechanical properties of TPS film [54]. This is most likely explained by changes in the starch's crystalline structure post-retrogradation. Studies such as those by Paluch et al. [55] and Surendren et al. [6] highlight that the addition of plasticizers disrupts the natural crystalline structure of starch, leading to a higher proportion of amorphous regions. This structural transformation results in increased chain mobility and free volume within the polymer matrix, which promotes more efficient starch gelatinization and flexibility. However, this increase in flexibility is accompanied by a reduction in material strength since the crystalline regions responsible for stiffness are diminished by the plasticizer-induced amorphous structures.

Since amorphous regions are more flexible and less rigid than crystalline regions, the overall mechanical properties (such as tensile strength and elongation at break) of the material become more sensitive to small changes in the glycerol concentration. With more glycerol, the proportion of amorphous regions in the starch increases while the amount of crystalline regions decreases. This sensitivity contributes to greater variability in the observed properties of TPS film because the balance between amorphous and crystalline regions can significantly affect the material's properties. Starch content and the presence of amylose have little effect on TS because the overall balance of crystalline and amorphous regions in TPS is more dependent on plasticizer content and is not dramatically altered by varying starch content [56,57]. On the contrary, increased amounts of starch content are found to decrease the EB and, in large concentrations, contribute to severe film brittleness [56,58].

The trade-off between TS and EB must be carefully managed depending on the desired application of the TPS film by carefully identifying optimal starch and plasticizer concentrations. Plasticizers like glycerol disrupt the intermolecular hydrogen bonding between starch molecules, leading to a more amorphous and flexible material, but at the cost of TS. While starch content may not significantly influence the mechanical properties of TPS due to its limited effect on the overall crystalline-amorphous balance, plasticizers weaken these interactions, resulting in lower TS values across most TPS formulations. Increasing plasticizer content leads to softer, more deformable films with reduced tensile strength and increased EB; however, it contributes to more variation in the observed properties. Increased starch content results in a reduction of EB, with little impact on the TS; however, it also leads to increased film brittleness. A balance between plasticizer and starch content is needed to maintain satisfactory film quality free of warping, brittleness, or excess moisture retention that may negatively impact the mechanical properties.

6. Conclusions

This study presents a novel Bayesian optimization approach for maximizing the elongation at break and tensile strength of TPS film by varying the plasticizer and starch content. The proposed framework addresses the challenges of DOE methodologies and traditional Bayesian approaches, including inefficiency, handling of heteroscedastic noise, and exploration region management. The proposed Bayesian approach leverages an HGP model suitable for physical experiments. This model independently predicts the mean and variance of the material properties, allowing for more accurate predictions of non-uniform

noise. The optimization process is guided by two acquisition functions (LCB and EI) and a dynamic exploration strategy to adjust the design space boundaries. Validation via numerical benchmark tests, including the Forrester's and three-hump camel functions, demonstrated the framework's capability to efficiently handle noise in multidimensional optimization problems. When applied to TPS films, the approach iteratively achieved EB and TS improvements, highlighting its potential to accelerate material optimization and enhance sustainable material performance.

This work's main contributions are as follows: First, it introduces a practical HGP model for Bayesian experimental design built from traditional GP models. The HGP implementation allows the use of established modeling tools with minimal modifications. Second, it presents an adaptive exploration scheme that systematically expands the design space by managing the exploration region. The two acquisition functions, along with the samples from the corner of the exploration region, enhance the diversity of candidate designs. In practice, this approach enables the generation of multiple designs per iteration, making it more feasible for physical experiments. Finally, this work pioneers a BO approach to formulate TPS films.

The results of this work are limited by the number and nature of design variables, the number of experiments, and the number of objectives. Only two variables are considered: plasticizer and starch content. Additional variables include additives, process parameters, or the relative molecular weight of TPS polymers through Gel permeation chromatography (GPC), which may positively affect the film's mechanical properties. Additionally, this work is constrained by having only 30 total experiments. More experiments could refine the results and further validate the optimization framework. Finally, this study focused on optimizing two objectives: elongation at break and tensile strength. Expanding the optimization framework to include additional properties such as hydrophobicity or compostability could offer deeper insights. Moreover, further characterization that can be gathered from stress–strain curves such as Young's modulus, yield strength, and strain hardening could provide a more comprehensive understanding of the material's potential to replace traditional plastic films. Despite the limitations of this study, the proposed work serves as a template to optimize material systems more efficiently.

Author Contributions: Conceptualization, A.T., A.P.S., and G.M.W.; methodology, A.T. and G.M.W.; software, G.M.W. and A.T.; validation, A.T. and A.P.S.; formal analysis, G.M.W., and A.P.S.; investigation, G.M.W.; resources, A.T.; data curation, G.M.W.; writing—original draft preparation, G.M.W. and A.T.; writing—review and editing, G.M.W., A.T., and A.P.S.; visualization, G.M.W.; supervision, A.T. and A.P.S.; project administration, A.T. and A.P.S.; funding acquisition, A.T. and A.P.S. All authors have read and agreed to the published version of the manuscript.

Funding: The authors gratefully acknowledge the support of the Center for Research and Learning (CRL) at Indiana University in Indianapolis through the Multidisciplinary Undergraduate Research Institute (MURI) Program, the Integrative Nanosystems Development Institute (INDI) at Indiana University Indianapolis, and the Research Experience for Undergraduates (REU) program of the National Science Foundation (NSF), Award #1950672. Any opinions, findings, conclusions, or recommendations expressed in this manuscript are those of the authors and do not necessarily reflect the views of our supporters.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article. The raw data and the data analysis code are openly available on the following GitHub repository: https://github.com/andres-tovar-purdue/materials/, accessed on 26 October 2024. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors express their sincere gratitude to Mangilal Agarwal, Director of the Integrated Nanosystems Development Institute (INDI) at Indiana University in Indianapolis for his support in the analysis and development of TPS film.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ahmed, S. Three ways to solve the plastics pollution crisis. *Nature* **2023**, *616*, 234–237.
2. Jones, N. More than 4000 plastic chemicals are hazardous, report finds. *Nature* **2024** . [CrossRef] [PubMed]
3. Marichelvam, M.; Jawaid, M.; Asim, M. Corn and rice starch-based bio-plastics as alternative packaging materials. *Fibers* **2019**, *7*, 32. [CrossRef]
4. Bangar, S.P.; Whiteside, W.S.; Ashogbon, A.O.; Kumar, M. Recent advances in thermoplastic starches for food packaging: A review. *Food Packag. Shelf Life* **2021**, *30*, 100743. [CrossRef]
5. Khan, B.; Bilal Khan Niazi, M.; Samin, G.; Jahan, Z. Thermoplastic starch: A possible biodegradable food packaging material—A review. *J. Food Process. Eng.* **2017**, *40*, e12447. [CrossRef]
6. Surendren, A.; Mohanty, A.K.; Liu, Q.; Misra, M. A review of biodegradable thermoplastic starches, their blends and composites: Recent developments and opportunities for single-use plastic packaging alternatives. *Green Chem.* **2022**, *24*, 8606–8636. [CrossRef]
7. Mohammadi Nafchi, A.; Moradpour, M.; Saeidi, M.; Alias, A.K. Thermoplastic starches: Properties, challenges, and prospects. *Starch-Stärke* **2013**, *65*, 61–72. [CrossRef]
8. MatWeb, LLC. Material Property Data for High Density Polyethylene (HDPE), Film Grade. Available online: https://www.matweb.com/search/DataSheet.aspx?MatGUID=af3d832d12b64e0ea1c0657208ae5b87&ckck=1 (accessed on 9 October 2024).
9. Abd Karim, S.F.; Idris, J.; Jai, J.; Musa, M.; Ku Hamid, K.H. Production of thermoplastic starch-aloe vera gel film with high tensile strength and improved water solubility. *Polymers* **2022**, *14*, 4213. [CrossRef]
10. Staker, J.; Schott, S.; Singh, R.; Collier, K.; Druschel, G.; Siegel, A.P.; Tovar, A. Influence of Choline Chloride/Urea and Glycerol Plasticizers on the Mechanical Properties of Thermoplastic Starch Plastics. *Polymers* **2024**, *16*, 751. [CrossRef]
11. Salaberria, A.M.; Diaz, R.H.; Labidi, J.; Fernandes, S.C. Role of chitin nanocrystals and nanofibers on physical, mechanical and functional properties in thermoplastic starch films. *Food Hydrocoll.* **2015**, *46*, 93–102. [CrossRef]
12. Vaezi, K.; Asadpour, G.; Sharifi, H. Effect of ZnO nanoparticles on the mechanical, barrier and optical properties of thermoplastic cationic starch/montmorillonite biodegradable films. *Int. J. Biol. Macromol.* **2019**, *124*, 519–529. [CrossRef]
13. Cyras, V.P.; Manfredi, L.B.; Ton-That, M.T.; Vázquez, A. Physical and mechanical properties of thermoplastic starch/montmorillonite nanocomposite films. *Carbohydr. Polym.* **2008**, *73*, 55–63. [CrossRef]
14. Taharuddin, N.H.; Jumaidin, R.; Ilyas, R.A.; Kamaruddin, Z.H.; Mansor, M.R.; Md Yusof, F.A.; Knight, V.F.; Norrrahim, M.N.F. Effect of Agar on the Mechanical, Thermal, and Moisture Absorption Properties of Thermoplastic Sago Starch Composites. *Materials* **2022**, *15*, 8954. [CrossRef] [PubMed]
15. Beer-Lech, K.; Skic, A.; Skic, K.; Stropek, Z.; Arczewska, M. Effect of Psyllium Husk Addition on the Structural and Physical Properties of Biodegradable Thermoplastic Starch Film. *Materials* **2022**, *15*, 4459. [CrossRef]
16. Schmitt, H.; Guidez, A.; Prashantha, K.; Soulestin, J.; Lacrampe, M.; Krawczak, P. Studies on the effect of storage time and plasticizers on the structural variations in thermoplastic starch. *Carbohydr. Polym.* **2015**, *115*, 364–372. [CrossRef] [PubMed]
17. Niazi, M.B.; Zijlstra, M.; Broekhuis, A.A. Influence of plasticizer with different functional groups on thermoplastic starch. *J. Appl. Polym. Sci.* **2015**, *132* . [CrossRef]
18. Mohammed, A.A.; Hasan, Z.; Omran, A.A.B.; Elfaghi, A.M.; Khattak, M.; Ilyas, R.; Sapuan, S. Effect of various plasticizers in different concentrations on physical, thermal, mechanical, and structural properties of wheat starch-based films. *Polymers* **2022**, *15*, 63. [CrossRef]
19. Jankovic, A.; Chaudhary, G.; Goia, F. Designing the design of experiments (DOE)–An investigation on the influence of different factorial designs on the characterization of complex systems. *Energy Build.* **2021**, *250*, 111298. [CrossRef]
20. Frazier, P.I.; Wang, J. Bayesian optimization for materials design. *Information Science for Materials Discovery and Design*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 45–75.
21. Frazier, P.I. Bayesian optimization. In *Recent Advances in Optimization and Modeling of Contemporary Problems*; Informs: Catonsville, MD, USA , 2018; pp. 255–278.
22. Greenhill, S.; Rana, S.; Gupta, S.; Vellanki, P.; Venkatesh, S. Bayesian optimization for adaptive experimental design: A review. *IEEE Access* **2020**, *8*, 13937–13948. [CrossRef]
23. Wang, K.; Dowling, A.W. Bayesian optimization for chemical products and functional materials. *Curr. Opin. Chem. Eng.* **2022**, *36*, 100728. [CrossRef]
24. Wang, X.; Jin, Y.; Schmitt, S.; Olhofer, M. Recent advances in Bayesian optimization. *ACM Comput. Surv.* **2023**, *55*, 287. [CrossRef]
25. Packwood, D. *Bayesian Optimization for Materials Science*; Springer: Berlin/Heidelberg, Germany, 2017; Volume 3.
26. Yamashita, T.; Sato, N.; Kino, H.; Miyake, T.; Tsuda, K.; Oguchi, T. Crystal structure prediction accelerated by Bayesian optimization. *Phys. Rev. Mater.* **2018**, *2*, 013803. [CrossRef]
27. Talapatra, A.; Boluki, S.; Duong, T.; Qian, X.; Dougherty, E.; Arroyave, R. Autonomous efficient experiment design for materials discovery with Bayesian model averaging. *Phys. Rev. Mater.* **2018**, *2*, 113803. [CrossRef]
28. Liu, K.; Wu, T.; Detwiler, D.; Panchal, J.; Tovar, A. Design for Crashworthiness of Categorical Multimaterial Structures Using Cluster Analysis and Bayesian Optimization. *J. Mech. Des.* **2019**, *141*, 121701. [CrossRef]
29. Xiong, Y.; Duong, P.L.T.; Wang, D.; Park, S.I.; Ge, Q.; Raghavan, N.; Rosen, D.W. Data-driven design space exploration and exploitation for design for additive manufacturing. *J. Mech. Des.* **2019**, *141*, 101101. [CrossRef]

30. Zhang, Y.; Apley, D.W.; Chen, W. Bayesian optimization for materials design with mixed quantitative and qualitative variables. *Sci. Rep.* **2020**, *10*, 4924. [CrossRef]
31. Zhang, M.; Parnell, A.; Brabazon, D.; Benavoli, A. Bayesian optimisation for sequential experimental design with applications in additive manufacturing. *arXiv* **2021**, arXiv:2107.12809.
32. Gao, H.; Zhong, S.; Zhang, W.; Igou, T.; Berger, E.; Reid, E.; Zhao, Y.; Lambeth, D.; Gan, L.; Afolabi, M.A.; et al. Revolutionizing membrane design using machine learning-bayesian optimization. *Environ. Sci. Technol.* **2021**, *56*, 2572–2581. [CrossRef]
33. Iwama, R.; Kaneko, H. Design of ethylene oxide production process based on adaptive design of experiments and Bayesian optimization. *J. Adv. Manuf. Process.* **2021**, *3*, e10085. [CrossRef]
34. Valladares, H.; Li, T.; Zhu, L.; El-Mounayri, H.; Hashem, A.; Abdel-Ghany, A.; Tovar, A. Gaussian process-based prognostics of lithium-ion batteries and design optimization of cathode active materials. *J. Power Sources* **2022**, *528*, 231026. [CrossRef]
35. Hickman, R.J.; Aldeghi, M.; Häse, F.; Aspuru-Guzik, A. Bayesian optimization with known experimental and design constraints for chemistry applications. *Digit. Discov.* **2022**, *1*, 732–744. [CrossRef]
36. Guo, J.; Ranković, B.; Schwaller, P. Bayesian optimization for chemical reactions. *Chimia* **2023**, *77*, 31–38. [CrossRef] [PubMed]
37. Qian, X.; Yoon, B.J.; Arroyave, R.; Qian, X.; Dougherty, E.R. Knowledge-driven learning, optimization, and experimental design under uncertainty for materials discovery. *Patterns* **2023**, *4*, 100863. [CrossRef] [PubMed]
38. Sattari, K.; Wu, Y.; Chen, Z.; Mahjoubnia, A.; Su, C.; Lin, J. Physics-constrained multi-objective bayesian optimization to accelerate 3d printing of thermoplastics. *Addit. Manuf.* **2024**, *86*, 104204. [CrossRef]
39. Williams, C.K. Prediction with Gaussian processes: From linear regression to linear prediction and beyond. In *Learning in Graphical Models*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 599–621.
40. Liang, Q.; Gongora, A.E.; Ren, Z.; Tiihonen, A.; Liu, Z.; Sun, S.; Deneault, J.R.; Bash, D.; Mekki-Berrada, F.; Khan, S.A.; et al. Benchmarking the performance of Bayesian optimization across multiple experimental materials science domains. *npj Comput. Mater.* **2021**, *7*, 188. [CrossRef]
41. Le, Q.V.; Smola, A.J.; Canu, S. Heteroscedastic Gaussian process regression. In Proceedings of the 22nd International Conference on Machine Learning, Bonn, Germany, 7–11 August 2005; pp. 489–496.
42. Goldberg, P.; Williams, C.; Bishop, C. Regression with input-dependent noise: A Gaussian process treatment. In Proceedings of the Advances in Neural Information Processing Systems, Denver, CO, USA, 1–6 December 1997; Volume 10.
43. Robert, C.P.; Elvira, V.; Tawn, N.; Wu, C. Accelerating MCMC algorithms. *Wiley Interdiscip. Rev. Comput. Stat.* **2018**, *10*, e1435. [CrossRef]
44. Edward, C. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006; Volume 211, p. 212.
45. Ebden, M. *Gaussian Processes for Regression: A Quick Introduction*; Robotics Research Group in Department on Engineering Science, University of Oxford: Oxford, UK, 2008; Volume 91, pp. 424–436.
46. Frazier, P.I. A tutorial on Bayesian optimization. *arXiv* **2018**, arXiv:1807.02811.
47. Jones, D.R.; Schonlau, M.; Welch, W.J. Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* **1998**, *13*, 455–492. [CrossRef]
48. Kotthoff, L.; Wahab, H.; Johnson, P. Bayesian optimization in materials science: A survey. *arXiv* **2021**, arXiv:2108.00002.
49. Zdanowicz, M.; Johansson, C. Mechanical and barrier properties of starch-based films plasticized with two-or three component deep eutectic solvents. *Carbohydr. Polym.* **2016**, *151*, 103–112. [CrossRef]
50. *ASTM Designation: D638-14*; Standard Test Method for Tensile Properties of Plastics. ASTM International: West Conshohocken, PA, USA, 2014.
51. Talja, R.A.; Helén, H.; Roos, Y.H.; Jouppila, K. Effect of type and content of binary polyol mixtures on physical and mechanical properties of starch-based edible films. *Carbohydr. Polym.* **2008**, *71*, 269–276. [CrossRef]
52. Qin, Y.; Wang, W.; Zhang, H.; Dai, Y.; Hou, H.; Dong, H. Effects of citric acid on structures and properties of thermoplastic hydroxypropyl amylomaize starch films. *Materials* **2019**, *12*, 1565. [CrossRef] [PubMed]
53. Caicedo, C.; Pulgarin, H.L.C. Study of the physical and mechanical properties of thermoplastic starch/poly (lactic acid) blends modified with acid agents. *Processes* **2021**, *9*, 578. [CrossRef]
54. Zuo, Y.; Gu, J.; Tan, H.; Zhang, Y. Thermoplastic starch prepared with different plasticizers: Relation between degree of plasticization and properties. *J. Wuhan Univ. Technol.-Mater. Sci. Ed.* **2015**, *30*, 423–428. [CrossRef]
55. Paluch, M.; Ostrowska, J.; Tyński, P.; Sadurski, W.; Konkol, M. Structural and thermal properties of starch plasticized with glycerol/urea mixture. *J. Polym. Environ.* **2022**, *30*, 728–740. [CrossRef]
56. Myllärinen, P.; Partanen, R.; Seppälä, J.; Forssell, P. Effect of glycerol on behaviour of amylose and amylopectin films. *Carbohydr. Polym.* **2002**, *50*, 355–361. [CrossRef]
57. Pelissari, F.M.; Yamashita, F.; Garcia, M.A.; Martino, M.N.; Zaritzky, N.E.; Grossmann, M.V.E. Constrained mixture design applied to the development of cassava starch–chitosan blown films. *J. Food Eng.* **2012**, *108*, 262–267. [CrossRef]
58. Mali, S.; Grossmann, M.V.E.; García, M.A.; Martino, M.N.; Zaritzky, N.E. Effects of controlled storage on thermal, mechanical and barrier properties of plasticized films from different starch sources. *J. Food Eng.* **2006**, *75*, 453–460. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.