



Multi-objective Bayesian alloy design using multi-task Gaussian processes

Danial Khatamsaz^{*}, Brent Vela, Raymundo Arróyave

Materials Science and Engineering Department, Texas A&M University, College Station, TX, USA

ARTICLE INFO

Keywords:

Multi-task Gaussian process
Bayesian optimization
Multi-objective optimization
Multi-task kernels
Correlation

ABSTRACT

In design applications, correlations among material properties (such as the tendency for stronger materials to be less ductile) are often neglected. This approach is echoed in multi-objective optimization techniques which treat each performance characteristic as an independent objective, aiming to optimize scalar functions and find optimal Pareto fronts. However, this overlooks the statistical relationships between performance characteristics inherent in a material system. To address this, we propose the use of Bayesian optimization, a highly efficient black-box optimization algorithm known for constructing Gaussian processes (GPs) – uncorrelated surrogates – to model objective functions. Rather than evaluating multiple GPs for each objective function separately, we argue for a shift towards jointly modeling these objective functions, considering their statistical correlations. This integrated approach utilizes naturally occurring relationships among material properties, providing additional information to enhance the performance of the design framework. This requires the replacement of multiple independent GPs with a single multi-task GP, employing a correlation matrix to construct a multi-task kernel function, wherein each task corresponds to a single objective function. We anticipate this refined methodology will better leverage material correlations, improving design optimization results.

1. Introduction

In alloy design, multiple performance characteristics, or properties, are often optimized for specific operational or environmental conditions. Yet, there is frequently a trade-off between these characteristics, with the improvement of one inadvertently leading to the sacrifice of another. This dynamic is evident in strength versus ductility [1] and density versus creep resistance [2] trade-offs in alloy design. To address this, multi-objective optimization techniques have emerged to identify Pareto-optimal solution sets [3,4]. A solution is considered Pareto-optimal when no other solution better optimizes all objectives simultaneously.

Among these techniques, Bayesian optimization (BO) algorithms represent a class of highly efficient design frameworks [5]. These algorithms iteratively conduct experiments to actively learn the Pareto-optimal designs. Their efficiency derives from modeling objectives with surrogates, typically Gaussian processes (GPs) [6]. This enables computationally economical representations of objective functions, alongside a heuristic-based search to discover non-dominated designs with an optimal number of experiments. Instances of implementing multi-objective BO in the design and discovery of materials with improved properties illustrate the importance and capability of Bayesian

approaches in the field [7,8].

However, a crucial gap in this approach is the neglect of statistical relationships between objectives, or properties. The strength-ductility tradeoff, for example, is a consequence of the fact that mechanisms or features that make a material strong tend to contribute to its brittleness. Traditional multi-objective BO algorithms construct independent GPs for each objective function, disregarding potential interactions between the functions. Such correlations naturally exist when modeling and designing a material system and can enhance our understanding of material properties. In the context of BO, explicitly accounting for these correlations essentially promotes transfer learning among the different property models (GPs), enhancing the predictive ability of the GPs across the entire property space. In this study, using the design of ductile high-strength refractory alloys as a benchmark, we demonstrate how acknowledging these correlations can improve the accuracy of surrogate models and the efficiency of multi-objective BO schemes.

2. Methods

In traditional multi-objective BO algorithms, objective functions are modeled as independent Gaussian Processes (GPs), relying solely on their respective observations, that implies an assumption of no

^{*} Corresponding author.

E-mail address: Khatamsaz@tamu.edu (D. Khatamsaz).

<https://doi.org/10.1016/j.matlet.2023.135067>

Received 29 May 2023; Received in revised form 22 July 2023; Accepted 16 August 2023

Available online 19 August 2023

0167-577X/© 2023 Elsevier B.V. All rights reserved.

correlation between them. However, the multi-task GP approach proposes a single surrogate to jointly model all objective functions [9], thereby acknowledging potential interrelationships. Within this design scheme, the term “task” refers to an objective function to be optimized. In this study, we build upon the design framework out-lined in [8], with a notable alteration: we replace multiple independent GPs with a single multi-task GP. This adaptation allows us to consider the correlation between objective functions within our probabilistic modeling. Thus, a Multi-task GP models the experiment and is searched to find maximizer of an acquisition function to select the next experiment to run.

Assuming constructing a GP for objective function $f(\mathbf{x})$, the prediction at an unobserved location $\mathbf{x}$ in existence of N previously observed data denoted by $\{\mathbf{X}_N, \mathbf{y}_N\}$, where $\mathbf{X}_N = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ and $\mathbf{y}_N = (f(\mathbf{x}_1), \dots, f(\mathbf{x}_N))$, is given as:

$$f_{GP}(\mathbf{x})|\mathbf{X}_N, \mathbf{y}_N \sim \mathcal{N}(\mu(\mathbf{x}), \sigma^2(\mathbf{x})) \quad (1)$$

where

$$\begin{aligned} \mu(\mathbf{x}) &= K(\mathbf{X}_N, \mathbf{x})^T [K(\mathbf{X}_N, \mathbf{X}_N) + \sigma^2 \mathbf{I}]^{-1} \mathbf{y}_N \\ \sigma^2(\mathbf{x}) &= k(\mathbf{x}, \mathbf{x}) - K(\mathbf{X}_N, \mathbf{x})^T [K(\mathbf{X}_N, \mathbf{X}_N) + \sigma^2 \mathbf{I}]^{-1} K(\mathbf{X}_N, \mathbf{x}) \end{aligned} \quad (2)$$

with k as a real-valued kernel function, $K(\mathbf{X}_N, \mathbf{X}_N)$ as a N by N matrix with m, n entry as $k(\mathbf{x}_m, \mathbf{x}_n)$, and $K(\mathbf{X}_N, \mathbf{x})$ is a N by 1 vector with m^{th} entry as $k(\mathbf{x}_m, \mathbf{x})$. The term σ^2 represents experimental error. The kernel function uses the notion of relative distance to obtain the correlation between observations. A common kernel function is squared exponential:

$$k^x(\mathbf{x}, \mathbf{x}') = \exp\left(-\sum_{h=1}^d \frac{(\mathbf{x}_h - \mathbf{x}'_h)^2}{2l_h^2}\right)$$

k^x indicates distance-based correlation, d is dimensionality of the input space and l_h is the characteristic length-scale to determine correlation strength between observations within the dimension h .

To better leverage the correlation between tasks and facilitate joint modeling of all objective functions, we first define the correlation matrix K^f , a covariance function over the tasks to capture the statistical relationships [9]. Then, we construct the kernel functions and covariance matrices in Eq. (2) as proposed in Ref. [9]:

$$K = K^f \otimes K^x, k = K^f \otimes k^x \quad (4)$$

where $\otimes$ is the Kronecker product operation. Note that the variable $\mathbf{y}_N$ in Eq. (2) is a column vector, thus, each task data should be ordered in columns, stacked vertically and the predictions are similarly stacked vertically. One can use GPyTorch Python package to build multi-task GPs and tune model hyperparameters [11].

3. Results

Here, we benchmark multi-objective BO frameworks, integrated with multi-task GPs (MTGPs), against a previously investigated multi-objective alloy design problem [7,8] in which the GPs for each property were considered to be independent (IGPs). Our objective is to maximize two indicators of ductility (Pugh ratio, Cauchy pressure) along with the yield strength, as predicted by the Maresca-Curtin model [10], within the Mo-Nb-Ti-V-W high-entropy alloy system. These alloys are relevant to next-generation gas turbine engine blades [8].

In this study, we replace the multiple IGPs with a single MTGP, jointly modeling all objectives by capitalizing on the existing

correlations. Data from our previous work [7,8] shows the correlation coefficients between the three objective functions, as indicated in Table 1. It is worth noting that these correlation coefficients should be dynamically updated as additional data is acquired, to maintain the most accurate statistical relationships between the objective functions. The calculated correlation coefficients indicate a discernible degree of correlation among the objectives. Notably, a strong inverse correlation is observed between the Pugh ratio and yield strength, offering the potential for exploitation of this correlation in the probabilistic modeling of these properties.

To assess the impact of integrating correlations into our design framework, we use both MTGP and IGPs to approximate the problem's optimal Pareto front, aiming to maximize the hypervolume, an indicator of proximity to the Pareto front. The BO over a design space consisting of 10,000 alloys is carried out 80 times in each case to assess the robustness of each approach, as shown in Fig. 1.

Utilizing a MTGP enhances design efficiency, achieving larger hypervolumes at each iteration than when using IGPs. While both methods converge to the same Pareto front, leveraging objective correlations improves modeling precision and design performance. Crucially, exploiting these correlations accelerates the discovery of the Pareto set.

The Uniform Manifold Approximation and Projection (UMAP) [12] in Fig. 2 visualizes the design space and Pareto points discovered after 300 iterations. Notably, the MTGP and IGP Pareto Sets agree with each other, but the MTGP Pareto set arrived at much faster. The trade-offs between objective functions and the strong inverse correlation between Pugh ratio and Yield strength are evident by comparing the colored spaces.

To quantify the improvement in modeling accuracy afforded by the MTGP implementation, we create training and testing sets using an 80/20 split of data, and compute the mean squared error for objective functions modeled using both independent and MTGP. The average mean squared errors, obtained via 5-fold cross-validation, are displayed in Fig. 3 for the three objectives under consideration. As shown in Fig. 3, for Pugh ratio and Cauchy pressure, the mean squared errors are about 100 times smaller when they are jointly modeled, however, the modeling error is smaller by around 10 times for independent modeling of the yield strength. Overall, MTGP models the objective functions

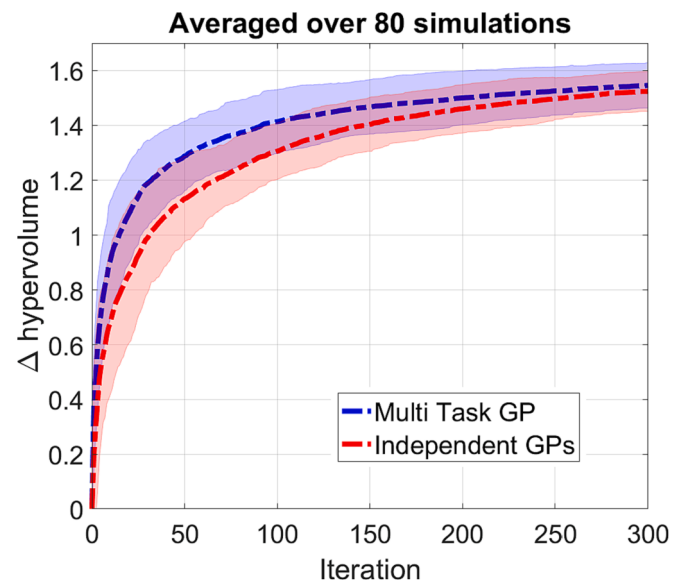


Fig. 1. Mean Hypervolume as a Function of Iteration Number. The graph reveals an enhancement in design performance when objective functions are jointly modeled. This improvement is indicated by larger hypervolumes, which suggest the discovery of higher-quality Pareto fronts.

Table 1
Pearson correlation coefficients between the 3 objective functions.

	Pugh ratio	Cauchy pressure	Yield strength
Pugh ratio	1	0.466	-0.917
Cauchy pressure	0.466	1	-0.235
Yield strength	-0.917	-0.235	1

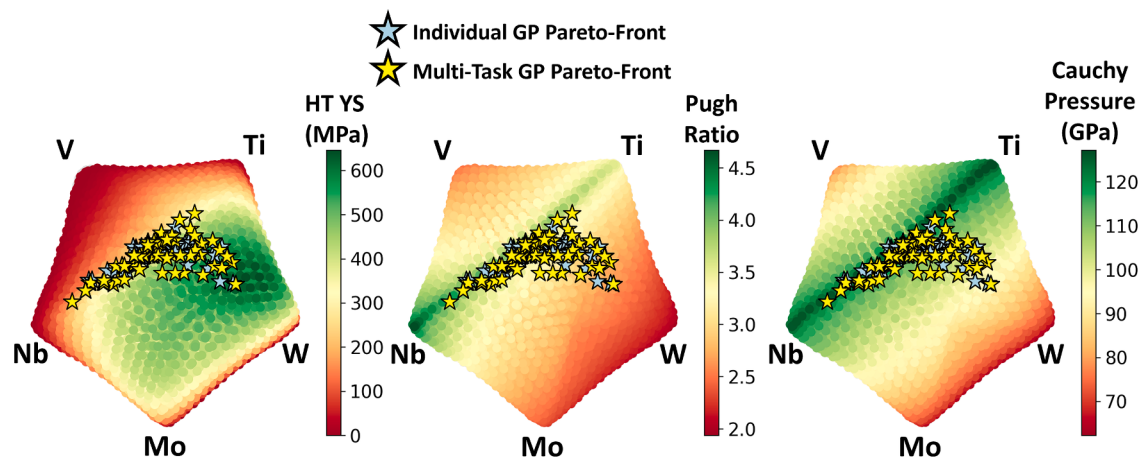


Fig. 2. Color coded UMAPs including discovered non-dominated designs. The non-dominated region discovered via using multi-task GP is slightly wider.

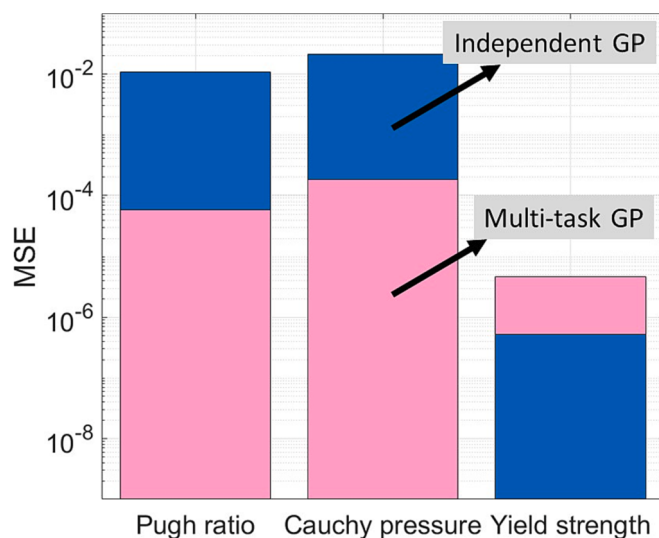


Fig. 3. Mean squared errors in prediction of test data using independent and joint modeling of quantities of interest. Overall, multi-task GP suggests higher accuracy in predictions because of considering existing correlations.

more accurately and improves the design efficiency by lowering the data-dependency of the design framework.

4. Conclusions

- Generally, properties of a material system have some degree of correlation.
- Multi-task GP exploits correlations among tasks in predictive modeling.
- Replacing uncorrelated GPs by a multi-task GP suggests more efficient multi-objective Bayesian alloy design.

CRedit authorship contribution statement

Danial Khatamsaz: Conceptualization, Methodology, Writing – original draft. **Brent Vela:** Methodology, Visualization, Writing – review & editing. **Raymundo Arróyave:** Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We acknowledge the support from the U.S. Department of Energy (DOE) ARPA-E ULTIMATE Program through Project DE-AR0001427 and DEVCOM-ARL under Contract No. W911NF2220106 (HTMDEC). The authors also acknowledge the support of NSF through Grants No. 1746932 (BV), CDSE-2001333 (RA,DK), and NSF-CISE- 1835690 and NSF-DMREF-2119103 (RA).

References

- [1] W. Fu, H. Li, Y. Huang, Z. Ning, J. Sun, A new strategy to overcome the strength-ductility trade off of high entropy alloy, *Scr. Mater.* 214 (2022) 114678.
- [2] R.C. Reed, T. Tao, N. Warnken, Alloys-by-design: application to nickel-based single crystal superalloys, *Acta Mater.* 57 (19) (2009) 5898–5913.
- [3] K. Deb, K. Deb, Multi-objective optimization, in: Search methodologies: Introductory tutorials in optimization and decision support techniques, Springer, 2013, pp. 403–449.
- [4] Tamaki, Hisashi, Hajime Kita, and Shigenobu Kobayashi. "Multi-objective optimization by genetic algorithms: A review." *Proceedings of IEEE international conference on evolutionary computation*. IEEE, 1996.
- [5] B. Shahriari, K. Swersky, Z. Wang, R.P. Adams, N. De Freitas, Taking the human out of the loop: A review of bayesian optimization, *Proc. IEEE* 104 (1) (2015) 148–175.
- [6] C.E. Rasmussen, C.K.I. Williams, *Gaussian Processes for Machine Learning* (Adaptive Computation and Machine Learning), The MIT Press, 2005.
- [7] D. Khatamsaz, B. Vela, P. Singh, D.D. Johnson, D. Allaire, R. Arróyave, Multi-objective materials bayesian optimization with active learning of design constraints: Design of ductile refractory multi-principal-element alloys, *Acta Mater.* 236 (2022) 118133.
- [8] D. Khatamsaz, B. Vela, P. Singh, D.D. Johnson, D. Allaire, R. Arróyave, Bayesian optimization with active learning of design constraints using an entropy-based approach, *npj Comput. Mater.* 9 (1) (2023).
- [9] E.V. Bonilla, K. Chai, C. Williams, Multi-task gaussian process prediction, *Adv. Neural Inf. Proces. Syst.* 20 (2007).
- [10] F. Maresca, W.A. Curtin, Mechanistic origin of high strength in refractory bcc high entropy alloys up to 1900k, *Acta Mater.* 182 (2020) 235–249.
- [11] J. Gardner, et al., Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration. *Advances in Neural Information Processing Systems* 31, 2018.
- [12] L. McInnes, J. Healy, N. Saul, L. Großberger, UMAP: Uniform Manifold Approximation and Projection, *JOSS* 3 (29) (2018) 861.