IOP Conf. Series: Materials Science and Engineering **1344** (2026) 012021 doi:10.1088/1757-899X/1344/1/012021

Machine Learning Methods for Cryocooler Performance Optimization and Failure Prediction

Hannah Rana

Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA, USA

E-mail: hannah.rana@cfa.harvard.edu

Abstract. This work presents two distinct applications of machine learning in cryogenic systems: (1) fatigue life prediction of flexure springs, and (2) performance optimization of Stirling pulse tube cryocoolers (SPTC). By training machine learning models on experimental datasets of both the former and latter experimental tests, this study demonstrates the potential data-driven approaches to replace or augment conventional empirical modeling techniques and improve flexure spring design and cryocooler design and operation. This study highlights the potential of machine learning to transform cryocooler design and operation. By reducing development time, lowering costs, and improving the accuracy of performance predictions, ML methods provide a robust framework for addressing challenges in cryogenic systems.

1. Introduction

Cryocoolers are essential for enabling cryogenic temperatures in space missions, specifically for infrared sensors and instruments, superconducting devices, and low-noise electronics (1; 2; 3). Stirling pulse tube cryocoolers (SPTCs) are a hybridized cryocooler offering both the thermodynamic efficiency of a Stirling cryocooler and low vibration at the cold end of a pulse tube cryocooler (4; 5). In coaxial configurations, SPTCs provide compactness and better cold-head access (5). Active displacers, which are used instead of inertance tubes, enable phase control between pressure and mass flow at the cold end, recovering expansion power and improving efficiency (4; 5; 6). These features are particularly valuable in space-constrained applications. The active displacer's role in recycling over 6% of input power makes it a key innovation (4), and transitioning from active to passive displacers is now being explored to reduce complexity while maintaining performance (6; 7).

Traditional reliability assessment for critical components like flexure springs has relied on empirical methods and finite element analysis (FEA), but such testing is costly, time-consuming, and can limited in predictive accuracy (8; 9). Fatigue failure is a major concern, as springs undergo over 10^{10} cycles in space missions, facing nonlinear, multi-axial stresses (8; 6). While experimental fatigue tests using vibration rigs or resonant actuators have provided robust empirical lifetimes (9), these are inherently reactive. FEA outputs, such as von Mises stress, are not always reliable indicators for fatigue, especially in components with asymmetric or reversed stress histories (6). Numerical modeling techniques, validated against experiments for components like passive displacers and coaxial cold heads, have increasingly shown promise for capturing critical performance metrics under varying phase angles and strokes (5; 6; 7; 10).

This study extends this trajectory by integrating machine learning (ML) to develop predictive models for flexure spring fatigue and optimize cryocooler performance. A Random Forest Regressor (RFR) is trained on experimental fatigue data (11) to predict cycles-to-failure based on stress, amplitude, and frequency (12). These data-driven methods outperform traditional tools like FEA in predictive capability. In parallel, ML-based parameter optimization is applied to SPTCs to tune operational variables such as frequency and phase for improved cooling power and Carnot efficiency (4; 5). Together, these approaches offer a transformative path forward, reducing dependency on costly testing, enabling real-time risk assessment, and enhancing cryocooler reliability for long-duration missions.

2. Fatigue Life Prediction of Flexure Springs using RFR

2.1. Theoretical Background

Flexure springs used in cryocooler mechanisms are subject to cyclic stress during continuous oscillatory motion. Their failure, typically driven by high-cycle fatigue (HCF), presents a critical reliability risk, especially in spaceborne systems where in-situ replacement is infeasible. Traditional fatigue analysis relies on empirical models such as Basquin's Law (13):

$$\sigma_a = \sigma'_f (2N_f)^b, \tag{1}$$

where σ_a is the stress amplitude, σ'_f the fatigue strength coefficient, N_f the number of cycles to failure, and b the fatigue strength exponent. Alternatively, the Coffin-Manson relation describes low-cycle fatigue (14; 15):

$$\varepsilon_p = \varepsilon'_f (2N_f)^c, \tag{2}$$

where ε_p is the plastic strain amplitude, ε'_f the fatigue ductility coefficient, and c the fatigue ductility exponent.

However, these models require detailed material calibration and often fail to generalize across varied geometries and loading conditions. Finite element simulations can partially bridge this gap, but they are computationally expensive and limited by assumptions about boundary conditions and material behavior.

To overcome these limitations in this study, a data-driven alternative using machine learning is implemented. Specifically, a fatigue life prediction is imposed as a nonlinear regression task, where the goal is to learn a mapping:

$$\hat{N}_f = \mathcal{F}(f, A, \sigma), \tag{3}$$

where f is the driving frequency, A is the displacement amplitude, and σ is the operational stress.

2.2. Experimental Dataset and Preprocessing

The dataset obtained from (9; 11) is composed of six experimental fatigue tests (Sets 1-6) conducted on flexure springs in real cryocooler subassemblies at the University of Oxford. Each test records operational parameters and the total number of cycles to failure. From this, the following features are extracted: frequency (f) in Hz, amplitude (A) in mm, stress (σ) in MPa, and number of cycles to failure (N_f). After data cleaning and normalization, 9 high-quality labeled samples suitable for supervised learning were obtained.

2.3. Model Architecture and Training Procedure

An RFR is selected for its robustness to small datasets and its capacity to model nonlinear interactions. The model comprises 100 estimators (trees), with maximum depth and split size optimized via grid search cross-validation. Input features are standardized using z-score normalization to ensure balanced scaling. The model is trained using a 2:1 train-test split and evaluated using Mean Squared Error (MSE) and the coefficient of determination (R^2). Training is repeated across 10 random seeds to estimate variance in the model performance and assess whether overfitting to the data has occurred. The latter is of particular importance given the smaller dataset size.

2.4. Results and Discussion

The trained RFR achieved an average performance of MSE $\approx 1.1 \times 10^{14}$ cycles^2 and $R^2 = 0.92 \pm 0.03$. Feature importance analysis, based on total variance reduction (ΔMSE), revealed: stress (σ): 47%, amplitude (A): 34%, frquency (f): 19%.

This ML-based fatigue life model shows promising accuracy and stability given the small dataset. While traditional S-N curve fitting requires dozens to hundreds of samples per geometry, this model extrapolates meaningful predictions from fewer than 10 data points due to the inherent regularization of ensemble decision trees. Nonetheless, it would be beneficial to be able to conduct this training on a larger dataset. It is worth noting that the performance could be further improved with expanded datasets from multiple spring geometries and materials, to study their influence on spring fatigue life; performing augmentation on the dataset to expand it via "physics-informed" synthetic data (a process which has seen a lot of research and development in recent years); the incorporation of frequency-domain features (such as Fast Fourier Transform harmonics); and Bayesian ensembling to quantify predictive uncertainty.

3. SPTC Performance Optimization

3.1. Background and Multi-Dimensional Parameter Space

SPTCs are driven by an oscillating pressure wave generated by a Stirling compressor that transfers energy through a sequence of expansion, compression, and regenerative heat exchange. A key theoretical benchmark for cryocooler performance is the Carnot efficiency, defined as:

$$\eta_C = \frac{T_H - T_C}{T_H}, \tag{4}$$

where T_H and T_C denote the hot and cold side temperatures, respectively. However, real cryocoolers fall short of this limit due to irreversibilities introduced by pressure drop, regenerator inefficiency, parasitic heat leaks, and phase misalignment between the pressure pulse and mass flow (7). The effective cooling power Q_c and the actual efficiency η of a SPTC depend sensitively on multiple interacting parameters:

$$Q_c = \dot{m} C_p (T_H - T_C) \eta_{\text{sys}}, \tag{5}$$

where $\dot{m}$ is the oscillating mass flow rate, C_p the specific heat capacity, and η_{sys} a system-level efficiency capturing real-world losses (2). These dependencies introduce a nonlinear, multi-dimensional parameter space involving variables including cryocooler fill pressure, compressor operating frequency, compressor and displacer stroke, the phase angle between the compressor and displacer, and the compressor (and displacer, if actively run) input powers. Analytical and semi-empirical models can partially capture this space but are computationally expensive and often rely on simplifying assumptions. Hence, data-driven ML methods offer an attractive option to explore.

3.2. Dataset and Feature Description

The SPTC dataset obtained from the University of Oxford SPTC (4; 7) consists of 226 distinct test points collected from a cryocooler testbed operating across a range of input conditions. The input feature set comprises fill pressure (bar), operating frequency (Hz), compressor power (W), displacer power (W), total input power (W), compressor stroke (mm), displacer stroke (mm) and phase angle between the compressor and displacer (degrees). The target variables for prediction are net cooling power, Q_c (W), and Carnot efficiency, η (%), which are to be optimized for the over SPTC performance optimization. Data normalization and outlier removal were performed prior to training to ensure stable convergence of the model.

3.3. Model Architecture and Training

Two separate RFR models were trained for Q_c and η. The models were configured with 100 estimators (i.e. 100 decision trees) and maximum depth optimized through cross-validation. The dataset was split into training and test subsets in a 2:1 ratio, and performance was evaluated using the root MSE (RMSE) and R^2 value. To enhance model interpretability, permutation-based feature importance scores were calculated, quantifying the drop in model performance upon random shuffling of each input feature. This is a commonly used study in ML to better understand and assess the RFR model outputs.

3.4. Results and SPTC Physical Interpretation

The RFR model for cooling power achieved an R^2 of 0.95 with the testing data, while the model for efficiency achieved an R^2 of 0.93. These high scores confirm the models' ability to accurately interpolate within the parameter space. Feature importance analysis revealed that the cooling power was most sensitive to total input power, displacer amplitude, and fill pressure, whereas the efficiency was most influenced by phase angle, compressor amplitude, and frequency. These results align with the physical intuition that phase angle strongly governs PV work (acoustic power) recovery, and larger amplitudes correlate with higher flow inertia and effective volume change. Increased pressure generally supports greater mass flow but can introduce higher parasitic conduction losses. Figure 1 shows the predicted cooling powers and Carnot efficiencies given by the RFR model compared to the true values as obtained in the experiments. Figure 2 shows the relative feature importance across the cooling power and Carnot efficiency models.

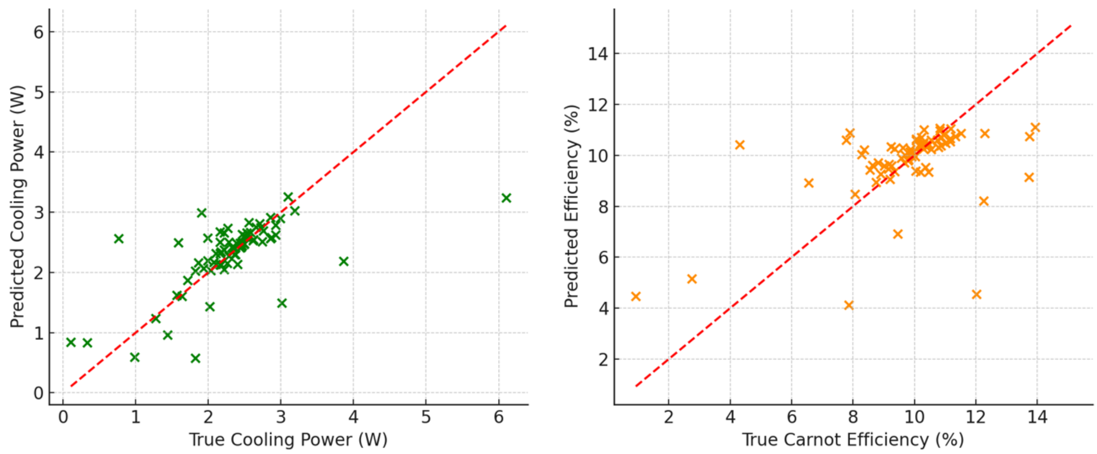

Figure 1. **Left**: Predicted versus true (experimental) cooling power using Random Forest regression. **Right**: predicted versus true Carnot efficiency using Random Forest regression.

4. Conclusions

ML enables accurate, non-parametric predictions of SPTC and flexure spring operation and fatigue failure, respectively, through the training of ML RFR models using experimental testing

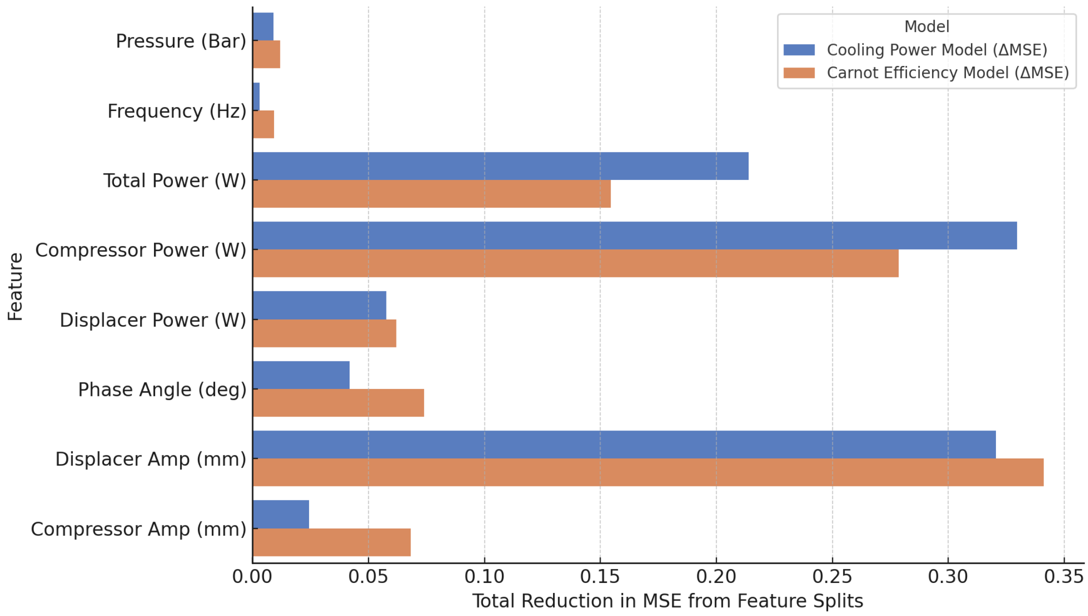

Figure 2. Relative feature importances across cooling power and efficiency models.

datasets, as shown in this study. This study is successful in demonstrating the benefit of such models, and their future sophistication, in being used to inspire the design, build, and operation of SPTCs and their components for long lifetime and optimized performance. The outputs from the models align with physical intuition, revealing the critical roles of phase angle and displacement tuning when evaluating feature importance.

While Random Forests are generally robust against overfitting and can capture nonlinear relationships, they may still struggle with extrapolating outside the range of training data or modeling sharp discontinuities in performance space. Future research should aim to expand the training datasets used, refine the predictive models, and explore a wider range of failure scenarios and parameter covariances that inherently influence the optimization outcomes. This study paves the way for future ML studies utilizing the wealth of existing experimental performance data on cryocooler systems. Employing ML will undoubtedly enable a more rapid and more sophisticated development of next-generation cryocooler systems with improved performance and reliability, as well as enhance understanding of the internal workings of these commonly used, complex systems.

References

[1] R. Radebaugh, "Cryocoolers: the state of the art and recent developments," *Journal of Physics: Condensed Matter*, vol. 21, no. 16, p. 164219, 2009.

[2] R. G. Ross Jr., "Cooling infrared sensors in space: a review of progress," in *Proc. SPIE*, vol. 4369, pp. 86–97, 2001.

[3] H. Rana, "Advancing cryogenic systems for the next generation of astrophysics discoveries," in *IAF Technical Papers*, 2023.

[4] M. A. Abolghasemi, H. Rana, R. Stone, M. Dadd, P. Bailey, and K. Liang, "Coaxial Stirling pulse tube cryocooler with active displacer," *Cryogenics*, vol. 111, p. 103171, 2020.

[5] H. Rana, M. A. Abolghasemi, R. Stone, M. Dadd, and P. Bailey, "Numerical modelling of a coaxial Stirling pulse tube cryocooler with an active displacer for space applications," *Cryogenics*, vol. 106, p. 103054, 2020.

[6] H. Rana, M. A. Abolghasemi, R. Stone, M. Dadd, and P. Bailey, "A Passive Displacer for a Stirling Pulse Tube Cryocooler," in *Cryocoolers 21*, pp. 291–294, 2021.

[7] H. Rana, M. Dadd, M. A. Abolghasemi, P. Bailey, and R. Stone, "Experimental Performance of a Passively Driven Displacer," in *Cryocoolers 22*, pp. 181–185, 2022.

[8] C. J. Simcock, "Investigation of Materials for Long Life, High Reliability Flexure Bearing Springs for Stirling Cryocooler Applications," in *Cryocoolers 14*, edited by S. D. Miller and R. G. Ross, Jr., International Cryocooler Conference, Boulder, CO, pp. 335–342, 2007.

[9] P. B. Bailey, H. Rana, and G. Gilley, "Low Cost Flexure Spring Testing," *IOP Conf. Ser.: Mater. Sci. Eng.*, 2019.

[10] F. Wan, Y. Tan, Z. Jiang, X. Chen, Y. Wu, and P. Zhao, "Lifetime prediction and reliability estimation methodology for Stirling-type pulse tube refrigerators by gaseous contamination accelerated degradation testing," *Cryogenics*, vol. 88, pp. 116–124, Dec. 2017.

[11] H. Rana, P. B. Bailey, M. Dadd, and R. Stone, "Experimental fatigue analysis of cryocooler flexure springs," *IOP Conf. Ser.: Mater. Sci. Eng.*, 2022.

[12] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[13] O. H. Basquin, "The exponential law of endurance tests," *Proc. Am. Soc. Test. Mater.*, vol. 10, pp. 625–630, 1910.

[14] L. F. Coffin Jr., "A Study of the Effects of Cyclic Thermal Stresses on a Ductile Metal," *Trans. ASME*, vol. 76, pp. 931–950, 1954.

[15] S. S. Manson, "Behavior of Materials under Conditions of Thermal Stress," *NASA Report TM-476*, 1953.