The University of Utah

Application of a Bayesian Framework for Plasticity Model Selection

Nolan Strauss [1] Karl Garbrecht [1] Geoffrey Bomarito [2] Patrick Leser [2] Jacob Hochhalter [1]

[1]University of Utah [2]National Space and Aeronautics Administration

Abstract

Interpretable Machine Learning (IML) has performed well when tasked with deriving constitutive material models. However, IML has been shown to prefer models that overfit noise in data, which tends to lead to bloat and a decrease in interpretability. Due to these issues, the ability of IML to reliably derive models that fit the data and are both interpretable and generalizable is limited. A method developed recently has shown promise to improve upon traditional IML by using a Bayesian fitness definition for the evolution of free-form models with non-deterministic parameters. This framework was developed for genetic-programming-based symbolic regression (GPSR) and involves model parameter estimation using Sequential Monte Carlo sampling (SMC). The method has demonstrated a reduction in bloat when dealing with noisy data in comparison to conventional GPSR. The results of this framework applied to stress-strain data for copper show models that more effectively predict the experimental data better than was previously shown with GPSR.

Genetic Programming by Symbolic Regression with Uncertainty Quantification

Genetic Programming by Symbolic Regression:

GPSR [4] can be simplified into the following five steps as depicted in Figure 1:

1. Initialize population of individuals
2. Generate new models
3. Evaluate the entire population
4. Select the best performing individuals and use those to evolve a new generation

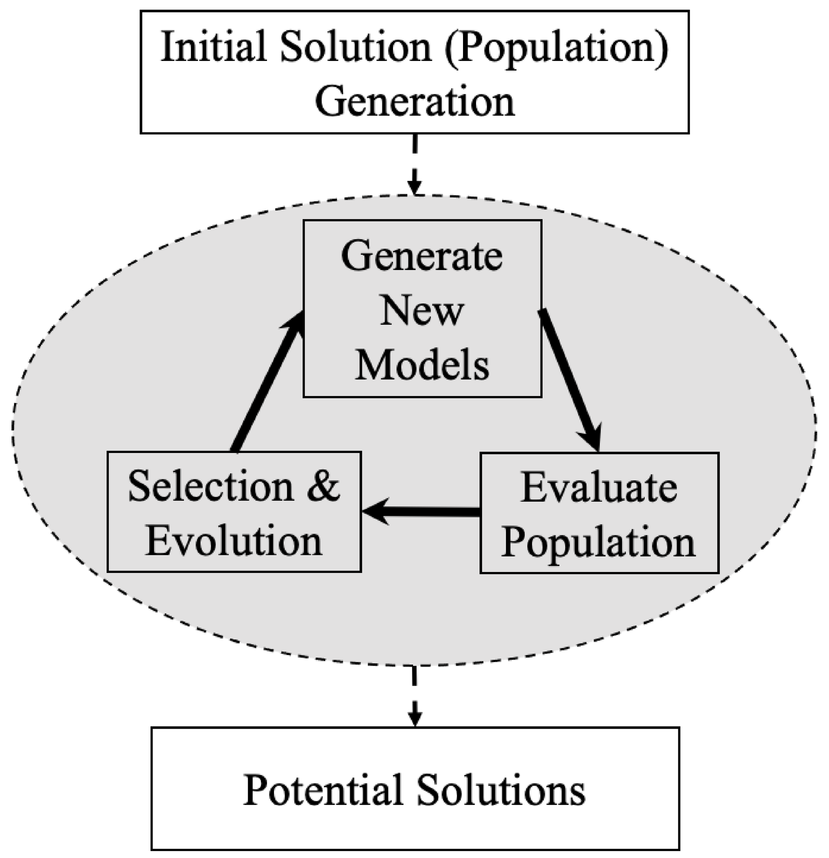

Figure 1. GPSR diagram displaying the initialization following by the iterative process of generation, evaluation, and selection and finally the result of a set of potential solutions.

Genetic Programming by Symbolic Regression with Uncertainty Quantification:

GPSR-UQ [2] is a recent improvement to GPSR via the inclusion of a modified model evaluation and selection. GPSR-UQ was chosen for the following reason:

1. Increased generalizability of discovered models
2. Reduced tendency towards bloat
3. Significant improvement in convergence rate w.r.t generations

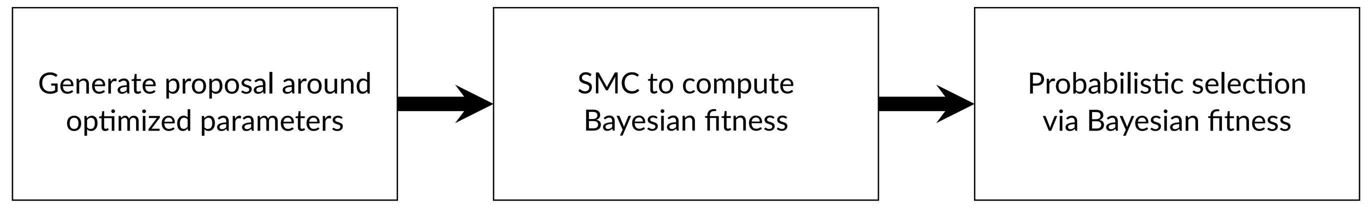

Figure 2. Bayesian reformulation of evaluation stage shown in Figure 1.

Bingo [1], an open-source package for performing symbolic regression that provides the Bayesian reformulation of the evaluation stage shown in Figure 1, and SMCPy [5], an open source package for uncertainty quantification, were used to perform GPSR-UQ.

Previous Works

1. "Data driven modeling of plastic deformation", Versino et. al. proposed the use of Symbolic Regression (SR) to develop free-form models to compare to traditionally used plasticity models
2. Versino et. al. discovered the following equation, referred to as SR-1, using datasets from several different resources in addition to synthetic data representing "expert knowledge"

$$\tau = 6.025 + 0.0095\mathbf{T} + \varepsilon^p(3217.611 + 2314.333\varepsilon^p - 4514.091\sqrt{\varepsilon^p} - 317.766(0.5835\varepsilon^p)^{\dot{\varepsilon}^p} - \mathbf{T}) \quad (1)$$

Where τ is stress [MPa], $\mathbf{T}$ represents temperature [K], ε^p represents the plastic strain, and $\dot{\varepsilon}^p$ represents the plastic strain rate $[s^{-1}]$

.

Current Work and Results

1. Using the datasets described in the previous section, GPSR-UQ was used to discover new plasticity models
2. From a run on a combination of datasets, the following equation, representing the maximum a posteriori (MAP), was found

$$\tau = 2123.11 e^{-\varepsilon^p} \left(\frac{e^{(\varepsilon^p)^2}}{\varepsilon^p} \right)^{\varepsilon^p} \sqrt{\frac{\varepsilon^p}{\sqrt{\mathbf{T}}}} \quad (2)$$

3. Credible and predicted outputs of equation 2 can be seen in the following plots

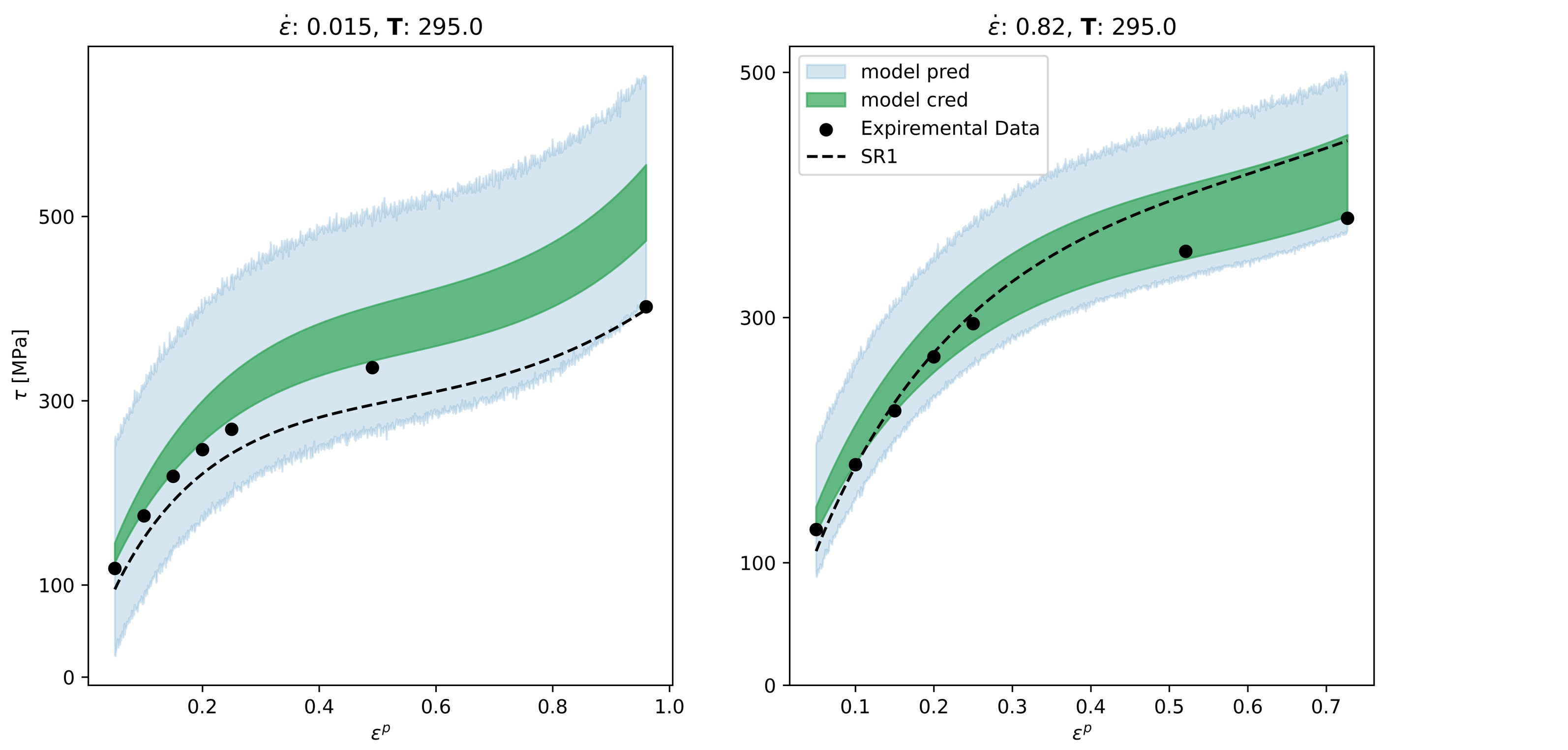

Figure 3. Plot of 95% credible and predicted interval of equation 2, plotted in green and blue, vs. SR-1, plotted as a black dashed line, along with the data from the two data subsets of varying strain rate ($\dot{\varepsilon}$) and Temperature ($\mathbf{T}$).

Conclusion

While traditional attempts to discover novel plasticity models using IML have shown to be promising, the use of GPSR-UQ has shown to outperform previous attempts by preferring more generalizable models and reducing the tendency to fit to noise, as shown in figure 3. These models are easier to implement than traditional plasticity models such as MTS while supplying similar accuracy.

Future Work

- Test discovered plasticity models by implementing them into ABAQUS as a user material (VUMAT) and compare to traditionally used models.

References

[1] G. Bomarito. Bingo. https://github.com/nasa/bingo, 2019.

[2] G. Bomarito, P. Leser, N. Strauss, J. Hochhalter, and K. Garbrecht. Bayesian model selection for reducing bloat and overfitting in genetic programming for symbolic regression. *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 526–529, 2022.

[3] K. Garbrecht, M. Aguilo, A. Sanderson, A. Rollet, R. Kirby, and J. Hochhalter. Interpretable machine learning for texture-dependent constitutive models with automatic code generation for topological optimization. *Metal Additive Manufacturing Modeling Challenge Series 2020*, 10:373–392, 2021.

[4] J. Koze. Genetic programming as a means for programming computers by natural selection. *Statistics and Computing*, 4:87–112, 1994.

[5] P. Leser. Smcpy. https://github.com/nasa/SMCPy, 2018.

[6] D. Versino, A. Tonda, and C. Bronkhorst. Data driven modeling of plastic deformation. *Computer Methods in Applied Mechanics and Engineering*, 10(1016):981–1004, 2017.