Physics-Informed ML for Nanocomposite Hardness

Predicting Vickers hardness of metal–ceramic composites, to screen candidates before the furnace

Ronnie W. · materials science

The problem: discovery is slow and data is messy

- **Cost.** Every candidate you fabricate and indent is days of work. The win is spending lab time on the *right* candidates.
- **Messy data.** Hardness is reported as HV, GPa, or kgf/mm². Reinforcement loading as weight *or* volume fraction. It all has to be made consistent first.
- **A subtle trap.** One paper gives 5–10 near-duplicate samples. If a model is graded on samples from a paper it also trained on, the nums look accurate but has just memorized that paper.

Easy to fool yourself into a nice looking number that won't hold up on a new material.

What I built, and how I made it trustworthy

A pipeline from composition + processing → predicted hardness, with three main rules -

1. **Tested the hard way** — the model is never graded on samples from a study it trained on, so the accuracy is real and not memorized.
2. **Validated on real public data** — the same method run on ~11,000 measured materials, not just my small set.
3. **Calibrated error bars** — every prediction comes with an honest uncertainty, and the model flags when it's extrapolating.

The hardness model is the application. Trustworthiness is the point of this project.

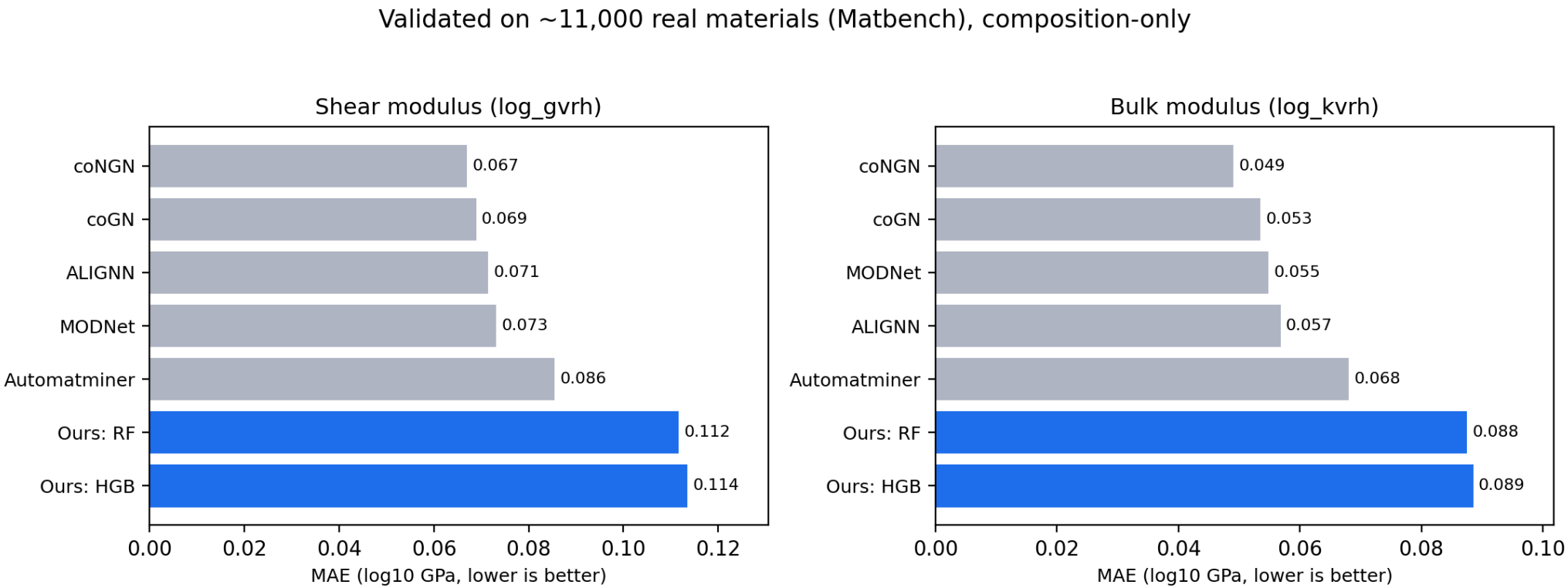
How it works

```
composition + processing  →  features  →  model  →  ranked candidates + uncertainty
```

- **Normalize** units and convert weight↔️volume fraction
- **Physics-informed features**: Hall-Petch (grain size), Orowan (particle strengthening), thermal-expansion mismatch, known metallurgy, its encoded for the model
- **Composition descriptors** (matminer/pymatgen) alongside the physics
- Gradient-boosted trees (interpretable, good for small data)

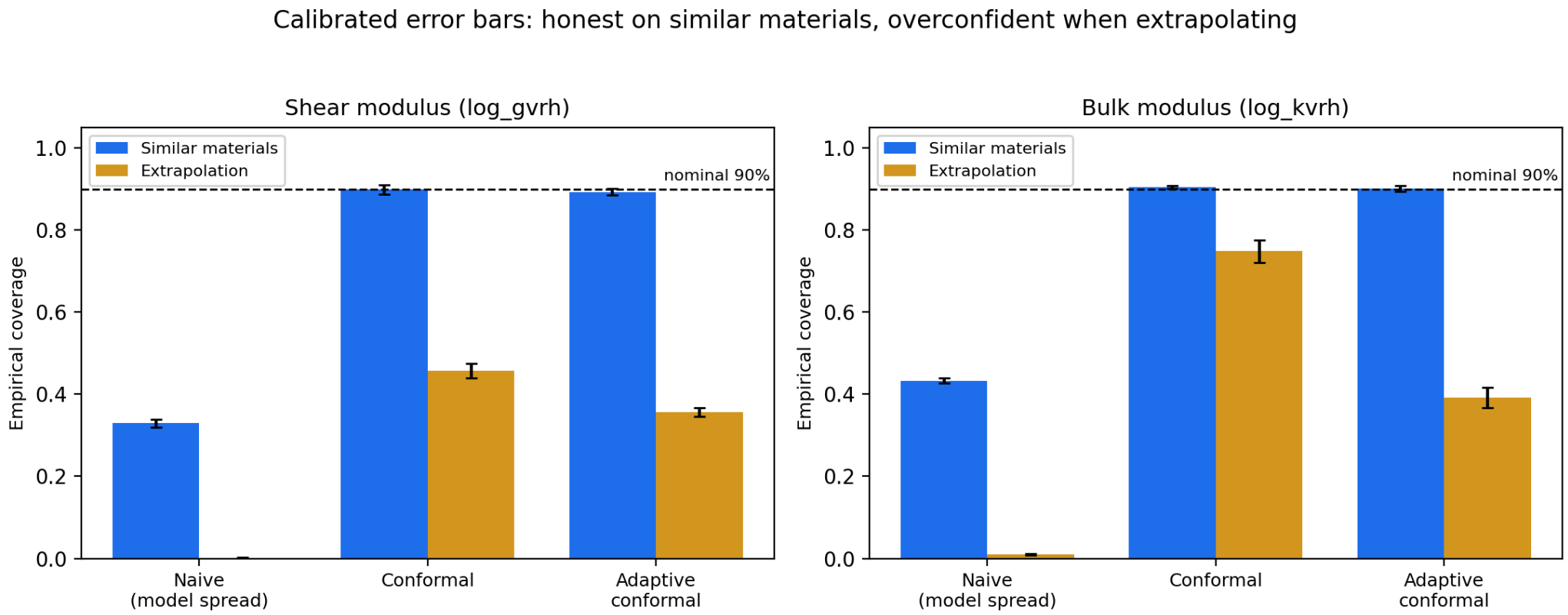
Physics priors give the model structure when data is thin.

Validated on real materials



Run on ~11,000 measured materials using **only composition**, the realistic case for a brand-new composite. Specialized models that use the full crystal structure do better. That gap is expected because we won't have a structure for a material that doesn't exist yet.

The model knows when it's guessing



- The model's **raw confidence is overconfident** (claims 90%, right ~33–43% of the time)
- **Calibration fixes it tho**: "90% confident" actually means 90% — on materials like the training set

Demo

live: predict from a composition · the honesty check · a ranked candidate shortlist

(switch to screen recording)

It agrees with the metallurgy

- Feature attribution (SHAP) shows **reinforcement fraction** and the **physics terms** drive the predictions
- The model leans on the same levers metallurgy says matter and not a inauthentic correlation
- When the physics and the learned model agree, you can start to trust it

Limitations & where this goes

Honest limitations

- Hardness data here is a small synthetic set. Real literature extraction is the next step

Already proven & ready to scale

- The method is validated on ~11,000 real materials · runs on a cluster
- The bottleneck to a real hardness result is **gathering the data, not the code**

Next: real literature data into the pipeline · use the uncertainty to choose which experiments to run next (active learning)

The point

Encode the physics · test honestly · give every prediction an error bar ·
know when the model is out of its depth

A tool you could maybe use before committing lab time to a composite.

Ronnie W. · wronnie@seas.upenn.edu · github.com/RonTonPeso/nanocomposite-hv-pipeline