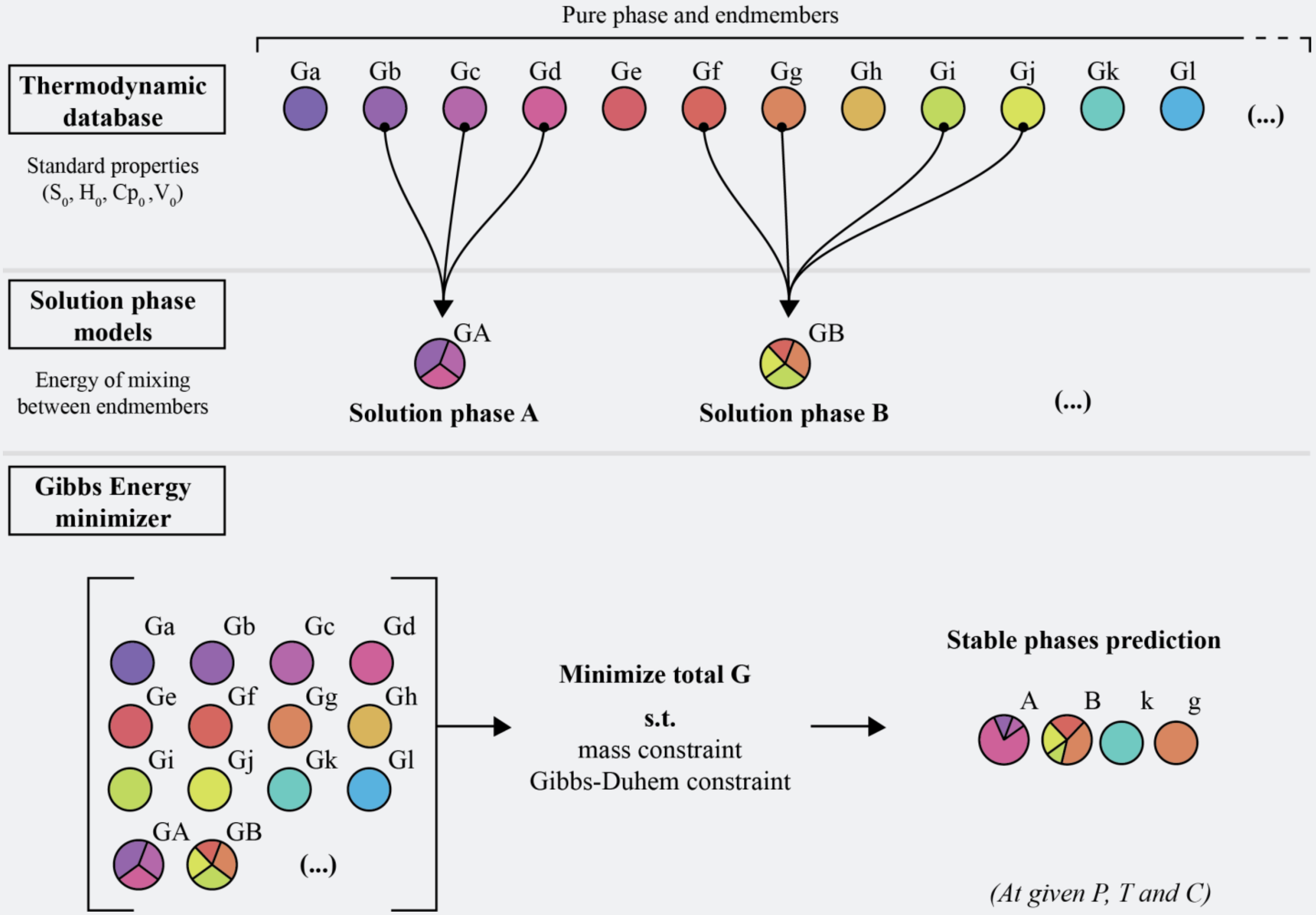
RocMLMs: Predicting Rock Properties through Machine Learning Models

Buchanan Kerswell et al.

Géosciences Montpellier

January 23, 2024

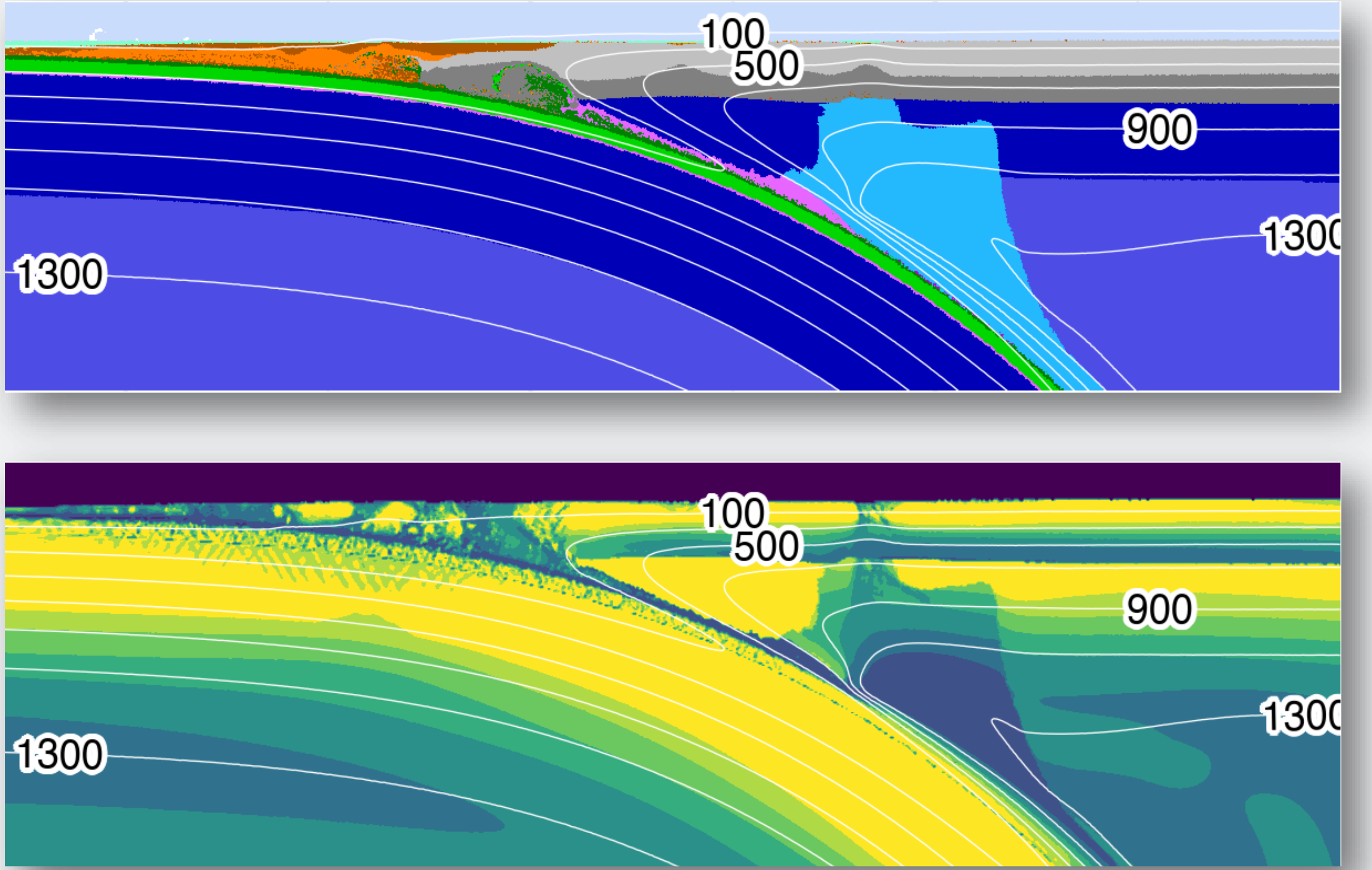
Problem definition



Problem: *predicting mineral assemblages is a tedious minimization problem that is computationally expensive !*

Riel et al. (2022; G3)

Problem definition



Kerswell et al. (2021; G3)

Numerical Implication: *Cannot change rock properties dynamically in large scale geodynamic simulations*

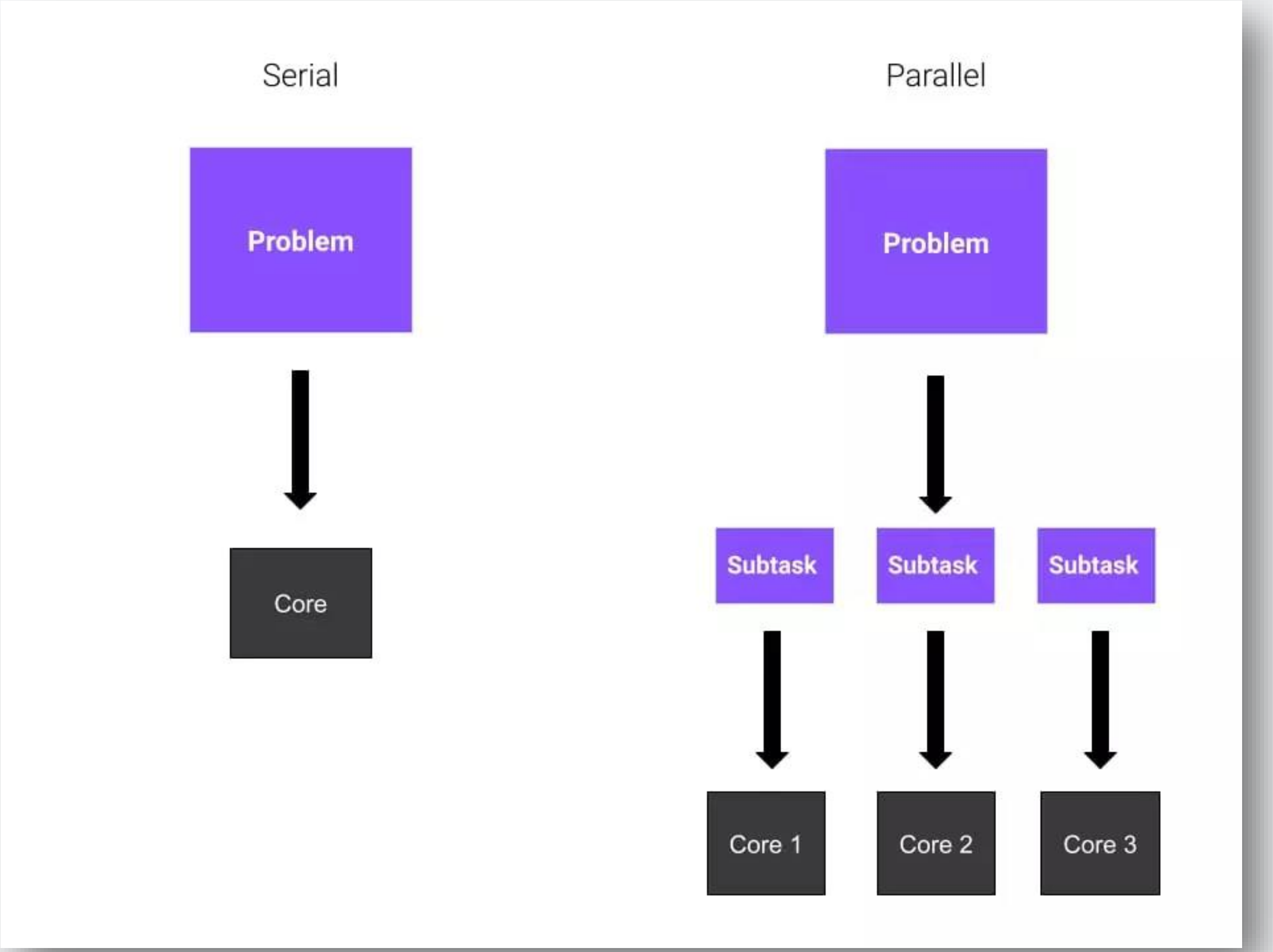
Physical Implication: *Density-driven mantle convection is not self-consistent*

921 x 301 nodes * **1.0 s** to compute stable assemblage = 3d 5h
921 x 301 nodes * **1.0 ms** to compute stable assemblage = 4m 30 s
921 x 301 nodes * **0.1 ms** to compute stable assemblage = 28s

Possible solutions

1. *Execute thermodynamic calculations in parallel (Riel et al., 2022 G3)*

Challenge: need 10^3 cores for 10^3 efficiency improvement

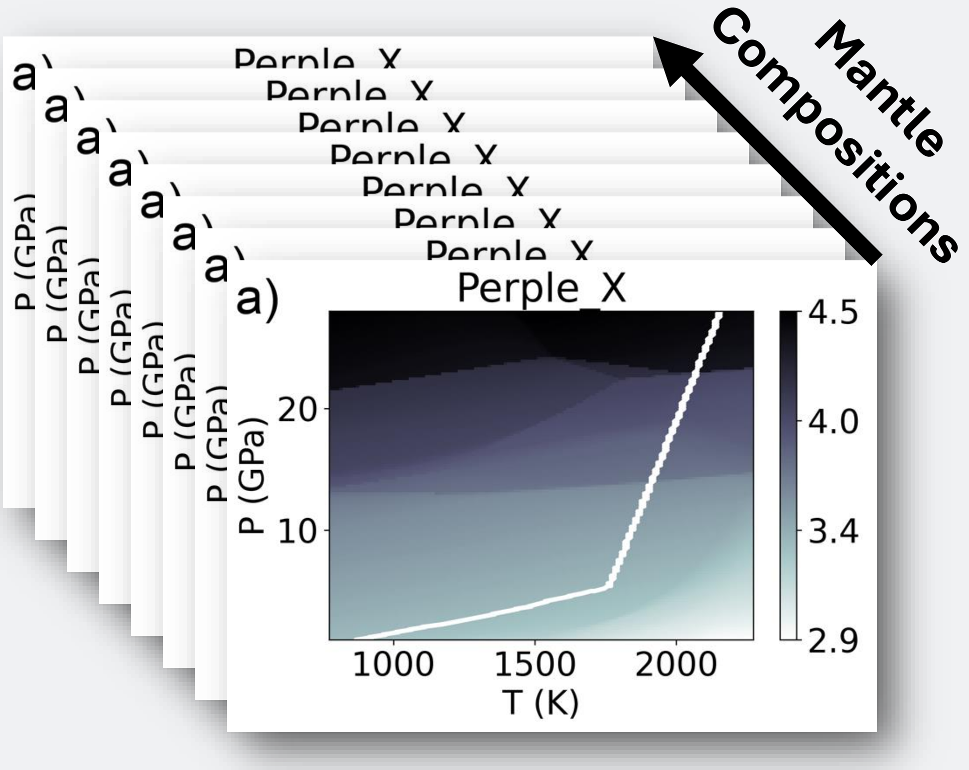


Meso-LR Cluster @ UM: 308 Nodes w/ 28 cores per node
10^3 cores ≈ 36 Nodes ≈ 12% usage of Meso-LR

Possible solutions

2. *Use precomputed lookup tables*

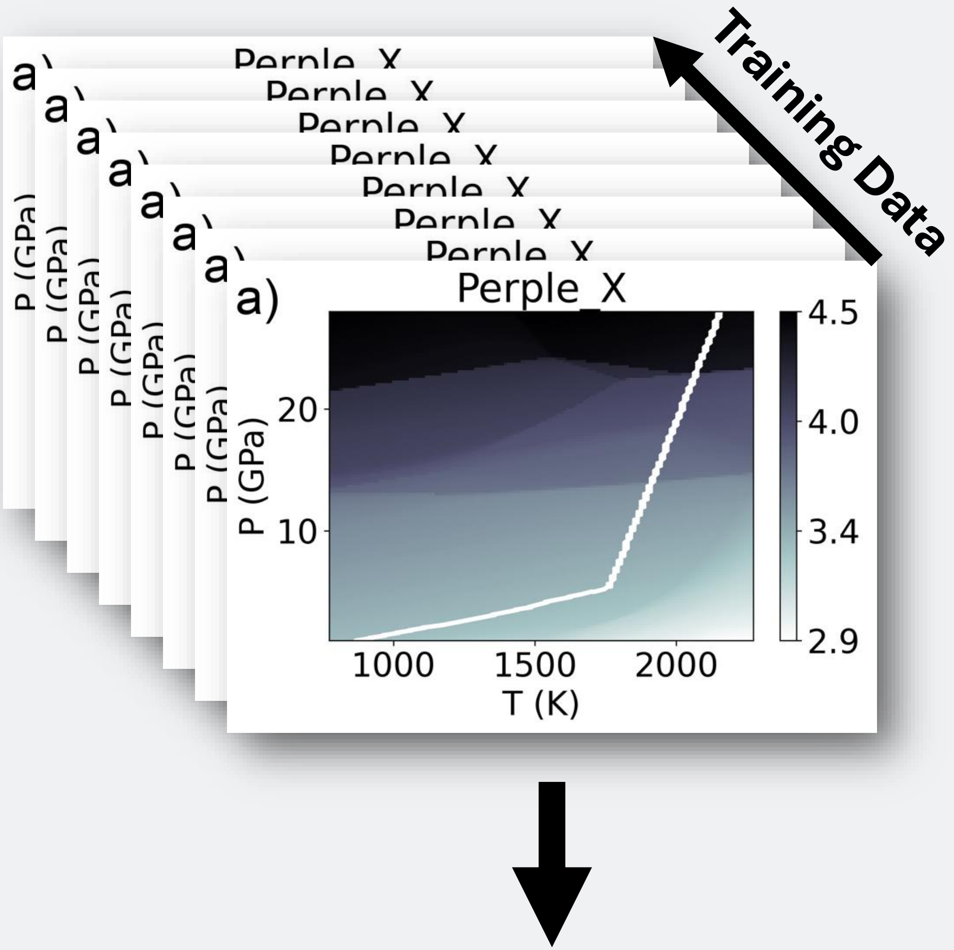
Challenge: need to store independent lookup tables for each rock type and target rock property



Possible solutions

3. *Use pretrained machine learning (ML) models*

Challenge: compress many thermodynamic calculations into a small efficient function of P, T, and X

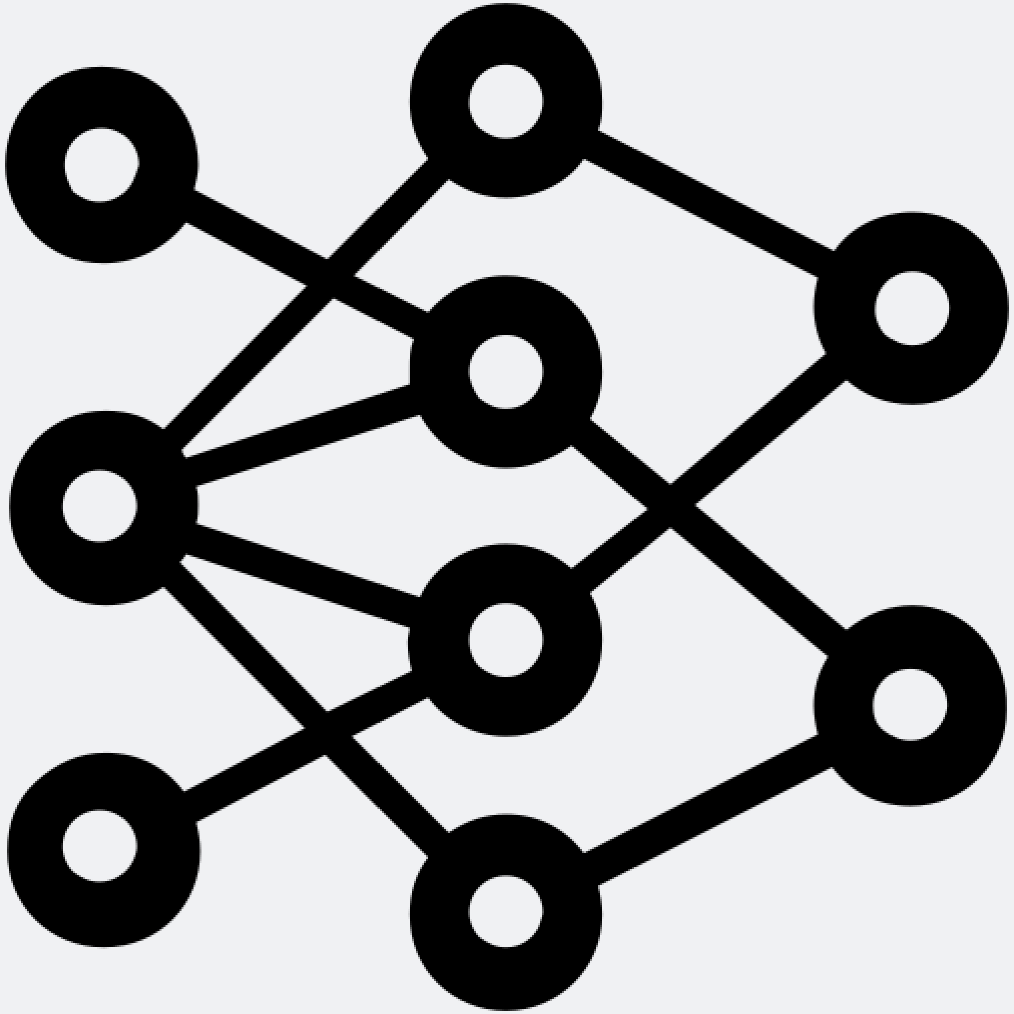


ƒ(P, T, X)

Research question

Can a pretrained ML models infer changes to rock properties ***accurately*** *and* ***more efficiently*** *than thermodynamic programs?*

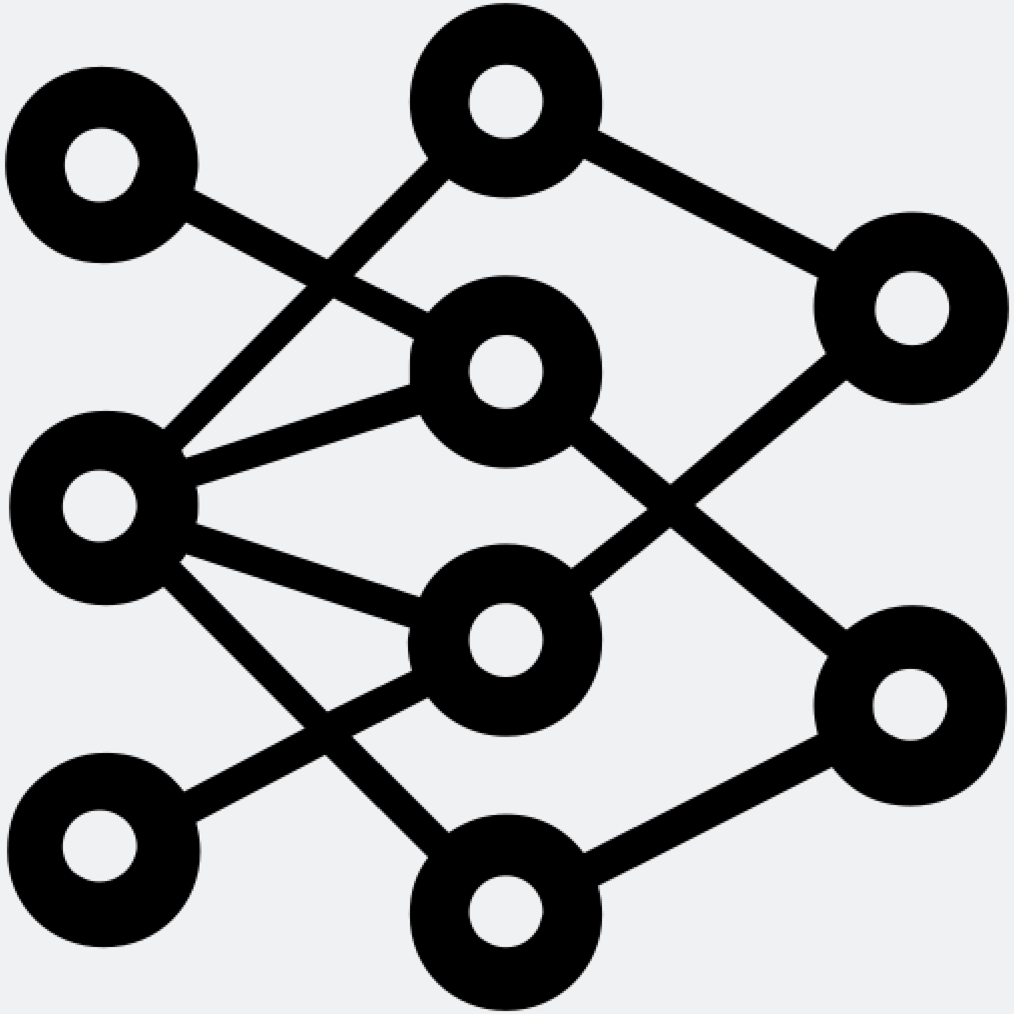
Implication: *pretrained ML models can replace thermodynamic programs for generalized tasks*



Hypothesis

ML models trained on an array of precomputed rock properties can improve the speed of predicting rock properties by 10^3–10^4 times

***Implication**: it is now feasible to simulate rock property changes self-consistently in large-scale geodynamic simulations*



Steps for using ML models for predicting rock properties

01

Build database of rock properties for a defined range of (P, T, X) ✅

02

Train ML models to predict rock properties ✅

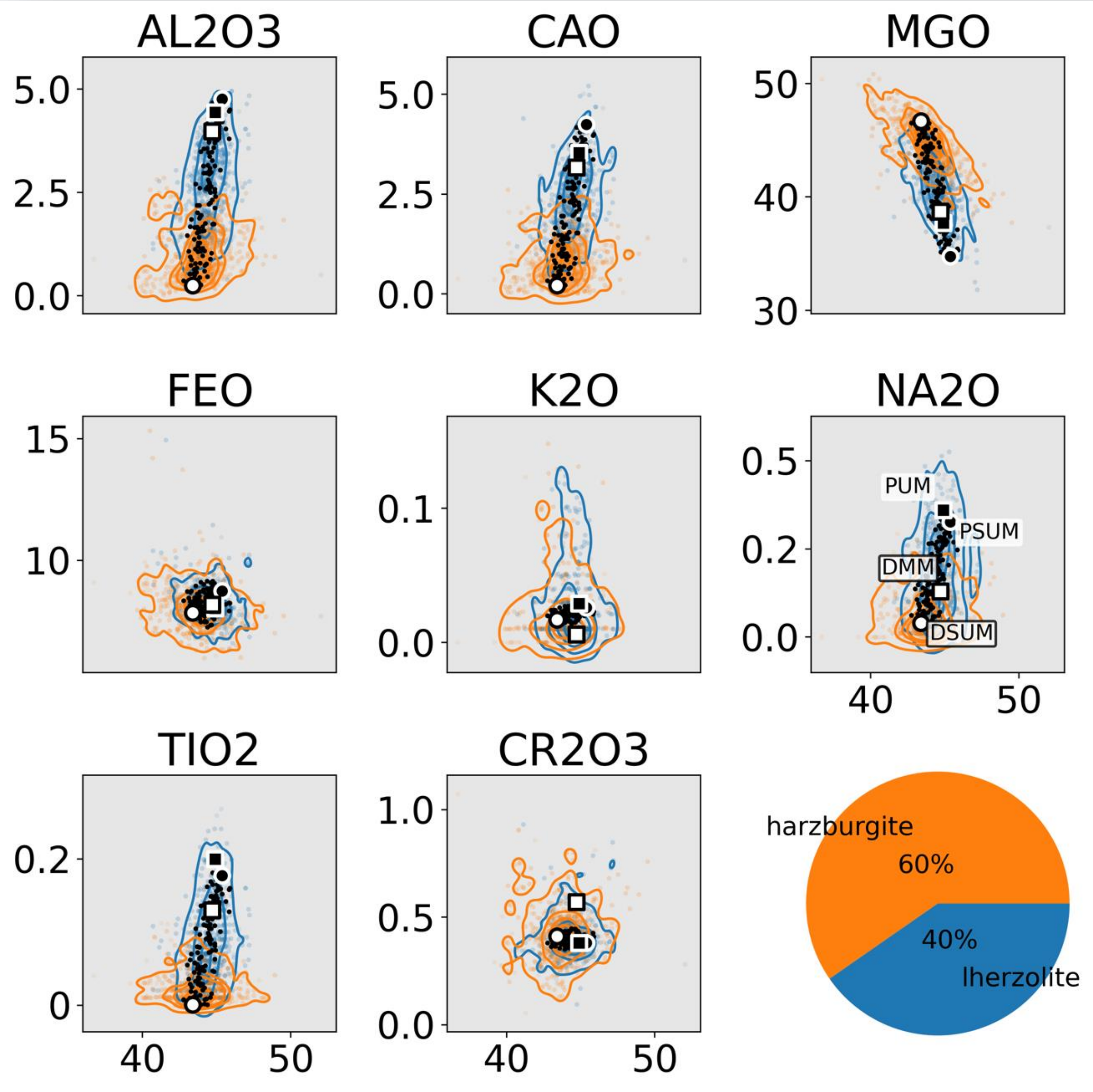
03

Benchmark ML models against incumbent thermodynamic programs ✅

04

Implement ML models into large-scale geodynamic simulations ⏳

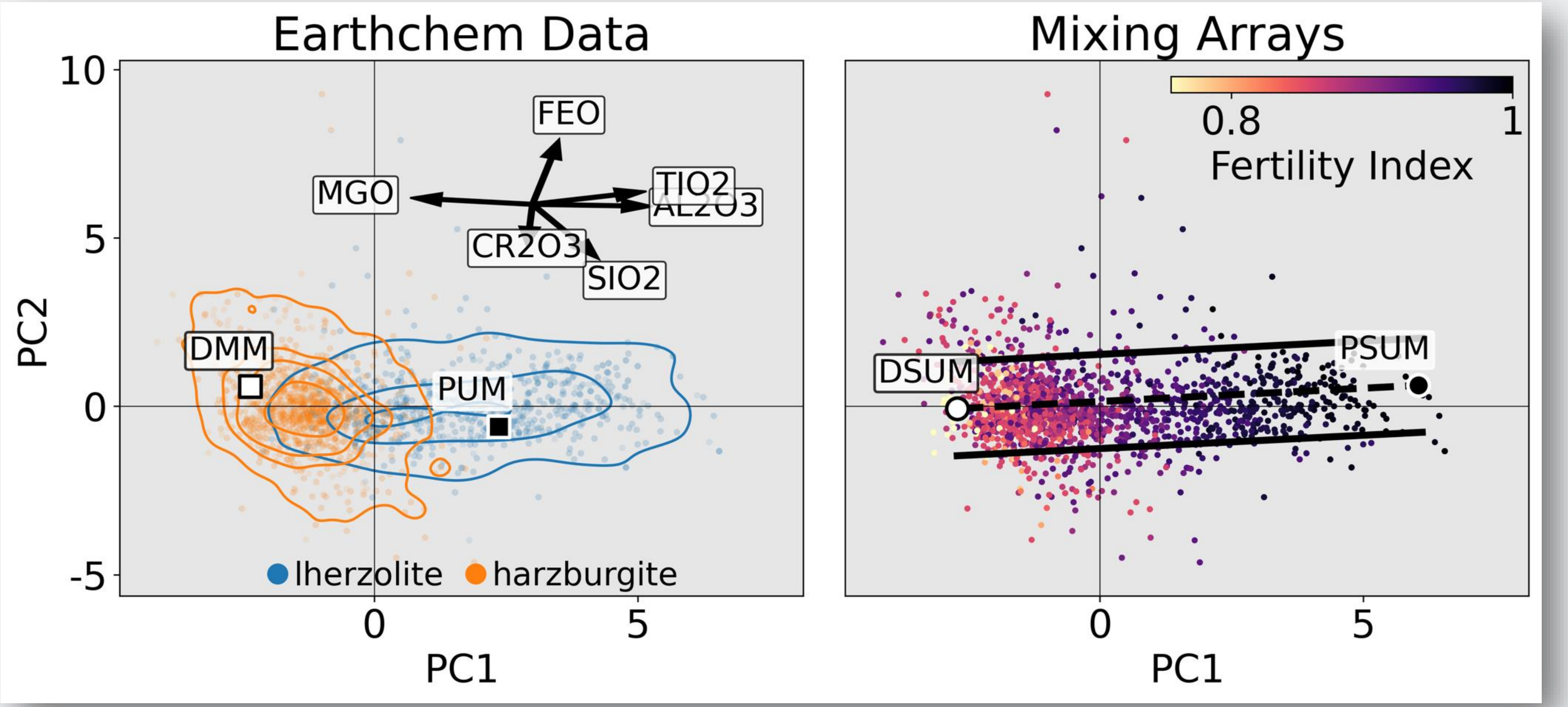
Step 1a: compile peridotite data



Typically, rocks are modelled with up to 11 chemical components (e.g., Al_2O_3, CaO, MgO, FeO, K_2O, NaO, TiO_2, Cr_2O_3)

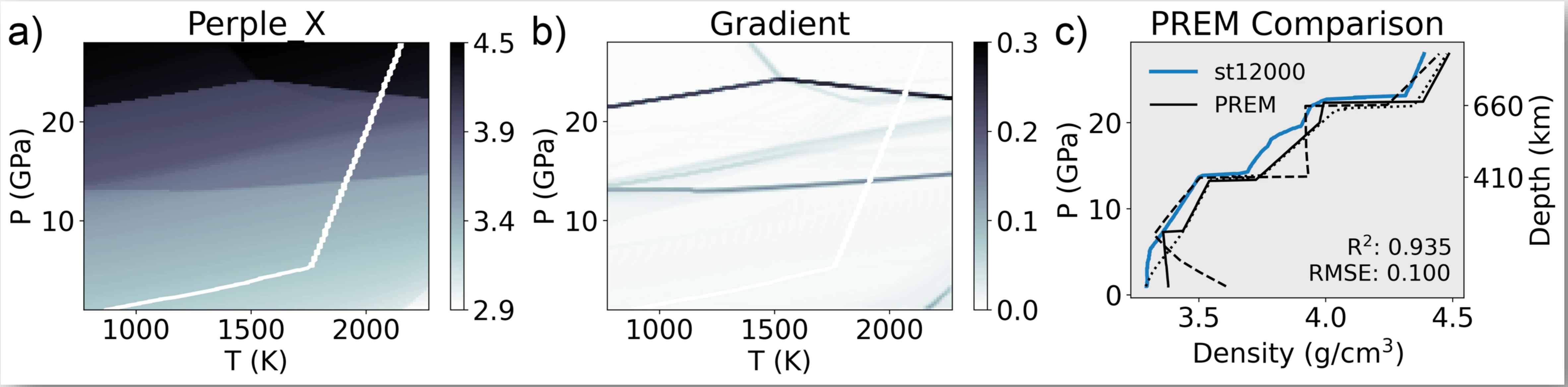
Implication: *pretrained ML models need 13 inputs (PT + 11 oxides) to predict rock properties*

Step 1b: reduce dimensionality



Implication: pretrained ML models need 3 inputs (PT + Fertility Index) to predict rock properties

Step 1c: compute phase diagrams

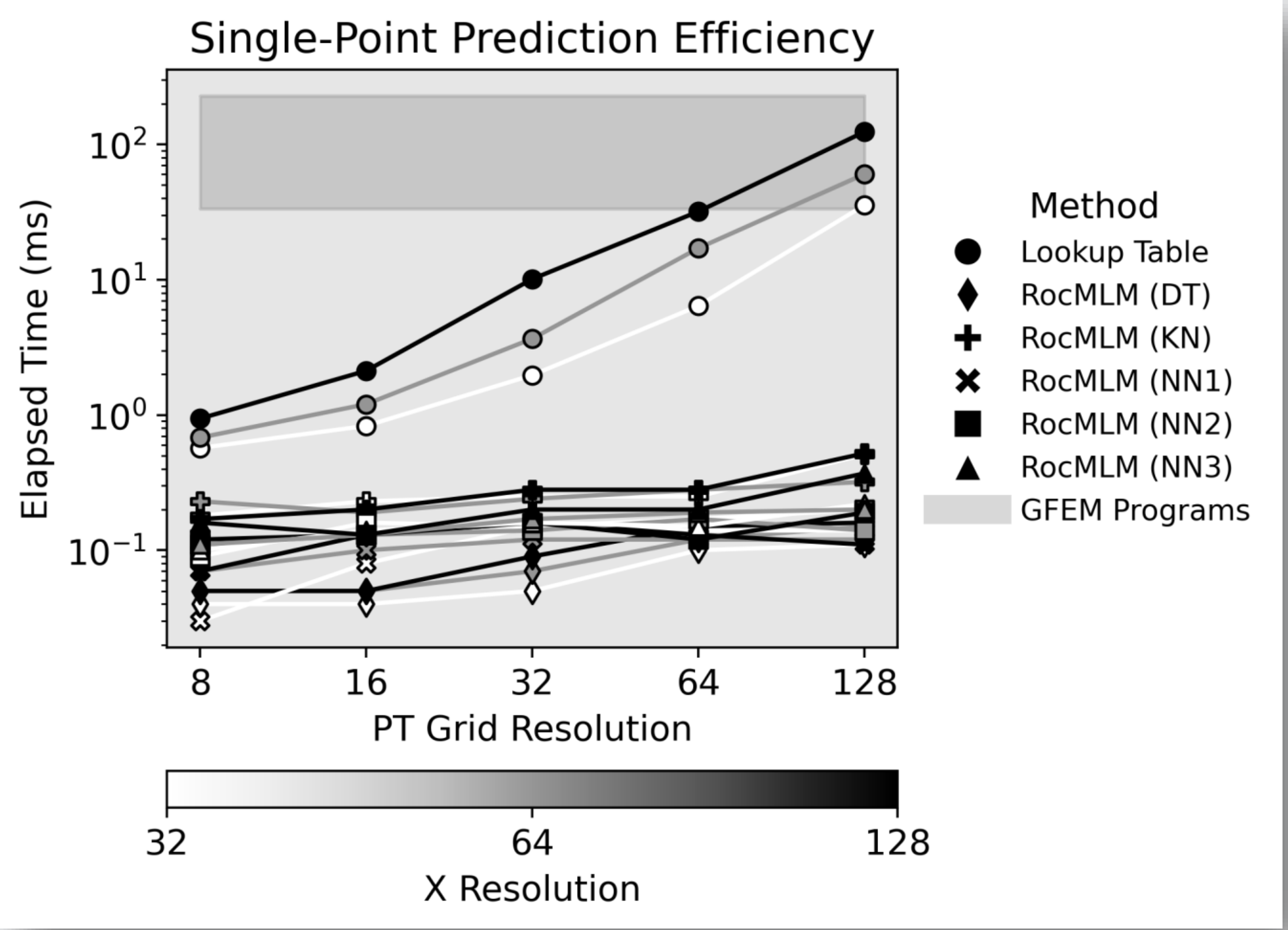


PTX resolution: 128 x 128 x 128 = 128^3 training examples (~2.1M)

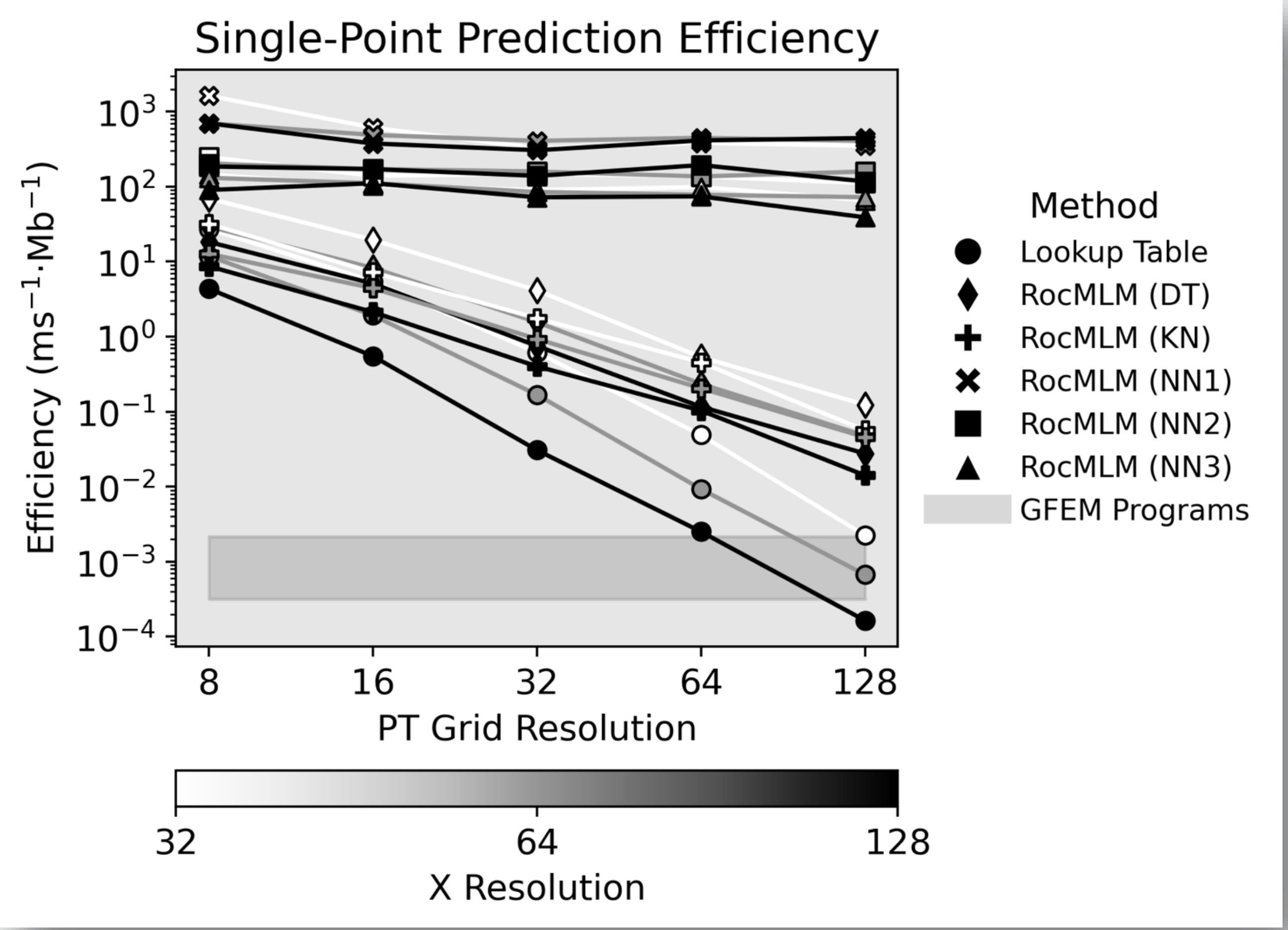
Step 2 & 3: train and benchmark ML models

On average, ML models make predictions 10^3 times faster than thermodynamic (GFEM) programs and lookup tables

Implication: *pretrained ML models are a major improvement over incumbent methods*



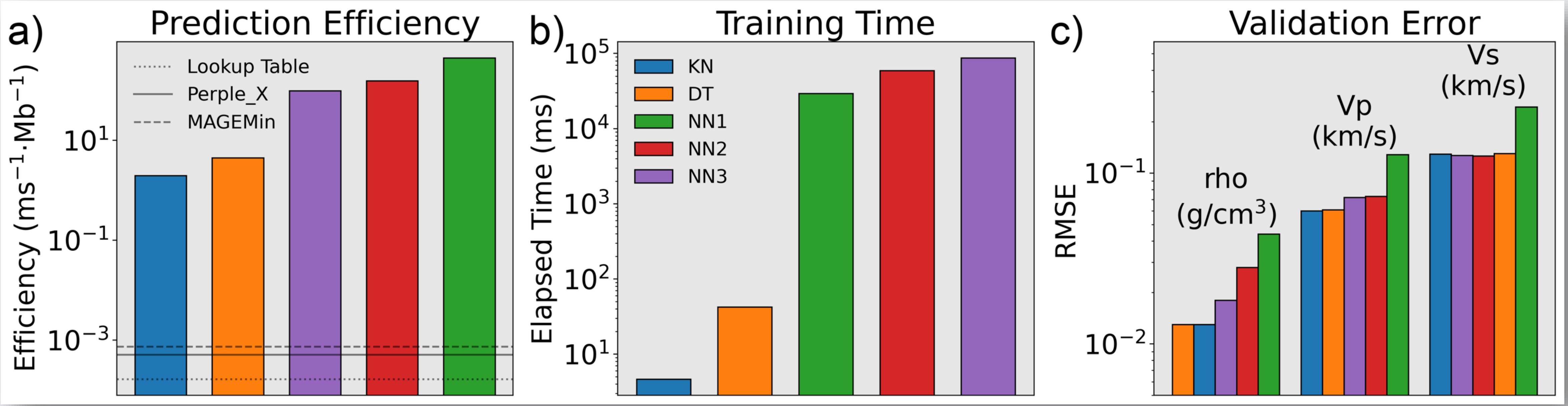
Step 2 & 3: train and benchmark ML models



The efficiency of some ML models scales poorly when considering the memory cost

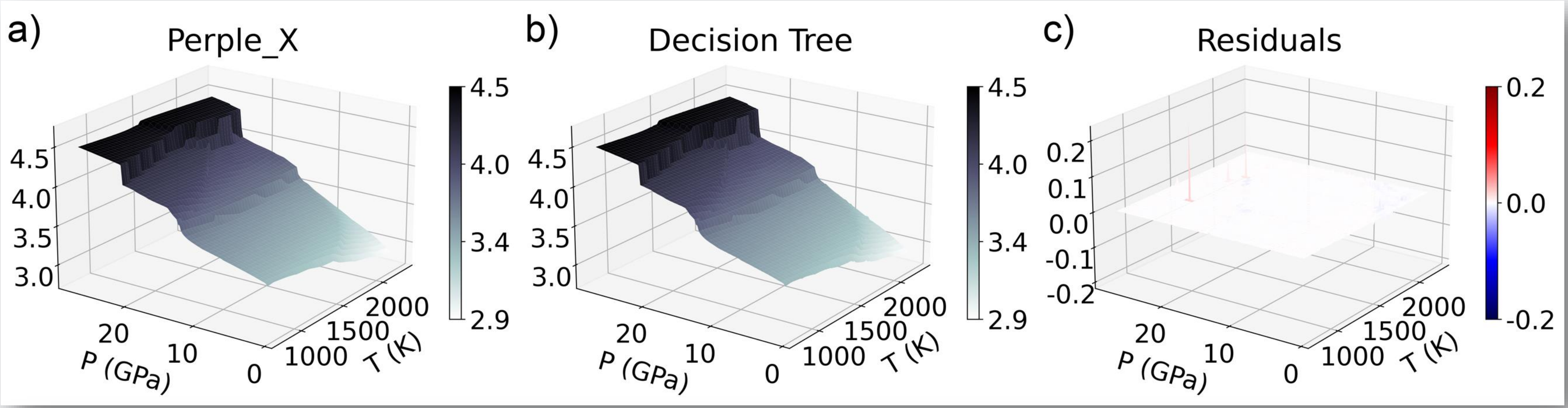
Implication: *NN models compress information better than KN or DT*

Step 2 & 3: train and benchmark ML models

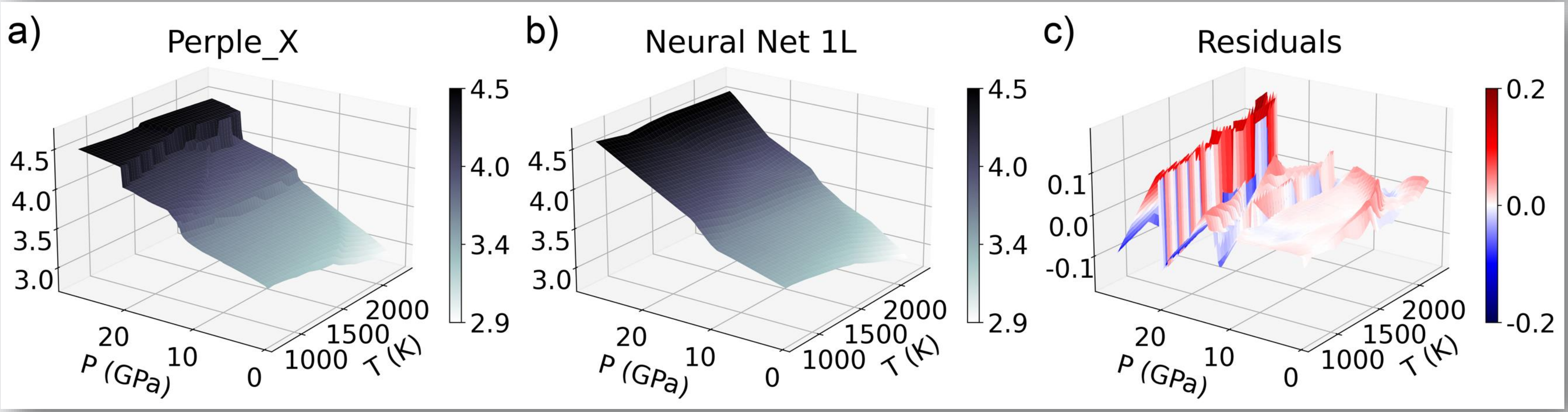


Implication: there are tradeoffs between prediction efficiency, training time, and accuracy

[redacted]

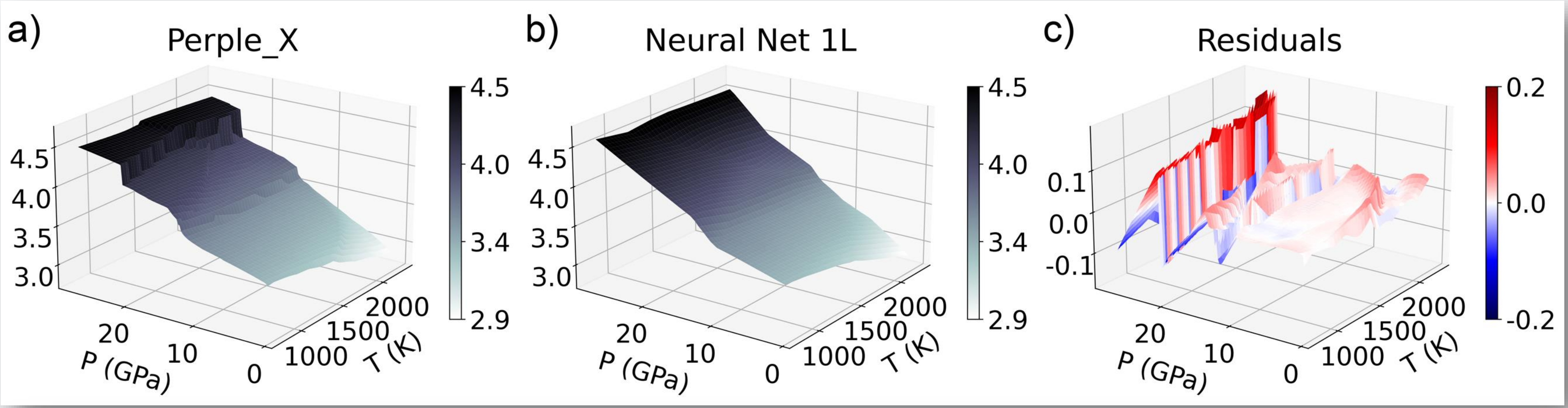


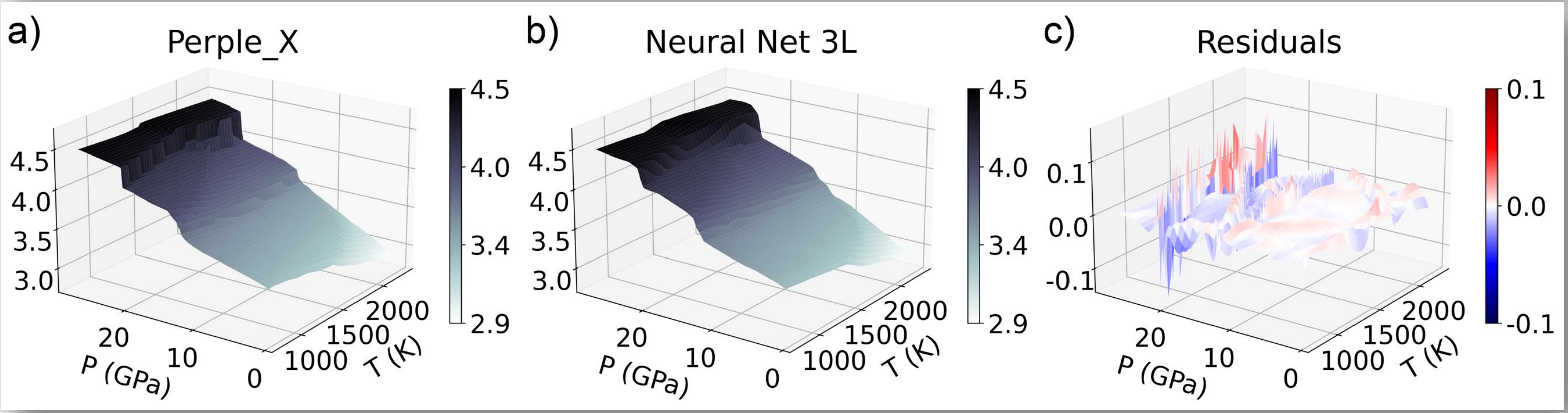
a)
Perple_X
b)
Decision Tree
c)
Residuals
4.5
4.0
3.5
3.0
4.5
4.0
3.4
2.9
20
10
0
P (GPa)
1000
1500
2000
T (K)
0.2
0.1
0.0
-0.1
-0.2



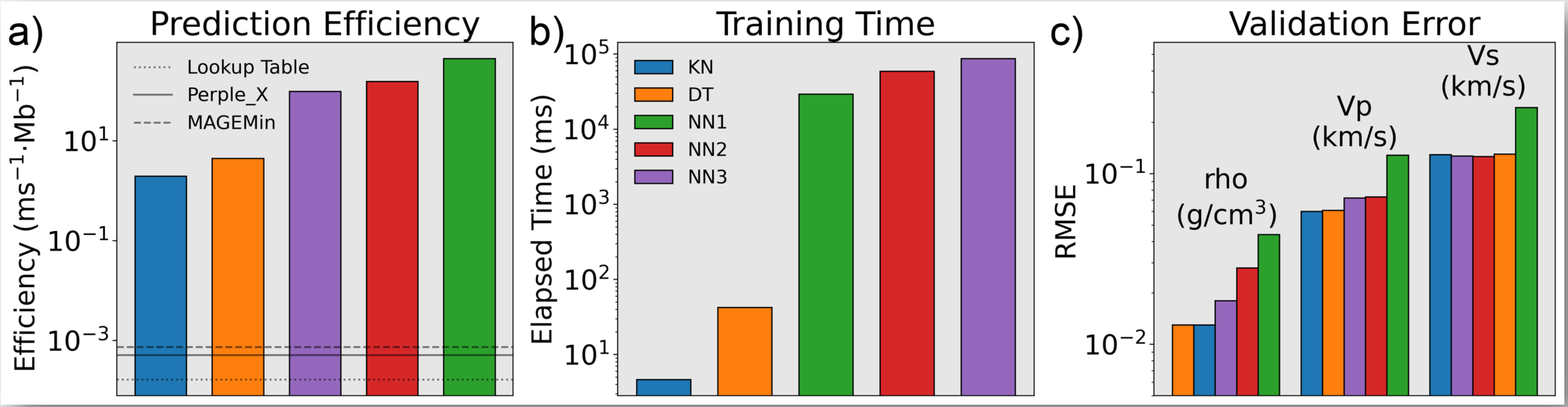
a)
Perple_X
b)
Neural Net 1L
c)
Residuals
4.5
4.0
3.5
3.0
4.5
4.0
3.4
2.9
20
10
0
P (GPa)
1000
1500
2000
T (K)
0.1
0.0
-0.1
0.2
0.0
-0.2

[redacted]


a)
Perple_X
b)
Neural Net 1L
c)
Residuals
P (GPa)
T (K)


a)
Perple_X
b)
Neural Net 3L
c)
Residuals
P (GPa)
T (K)

Which ML model is the "best"?



Fastest: DT or KN
Most accurate: DT or KN
Most efficienct overall (best compression): NN

Questions?

Thanks for the attention