SAND

SAND2021-4579PE

Accelerating data-driven discovery of materials for hydrogen storage and generation

PRESENTED BY

Matthew Witman

Sandia National Laboratories, Livermore, CA USA

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Collaborators & Acknowledgements

Experimental chemistry/materials science

Mark Allendorf
Vitalie Stavila
Anthony McDaniel
Andrea Ambrosini
Jeffery Chames
David Grant (Nottingham University)
Gavin Walker (Nottingham University)
Gustav Ek (Uppsala University)
Martin Sahlberg (Uppsala University)
Claudia Zlotea (Paris Est)

ML Theory

Sapan Agarwal
Justin Wong

Computational Chemistry

Sanliang Ling (Nottingham University)
Nathan Mahynski (NIST)
Harold Hatch (NIST)
Nick Wunder (NREL)
Nalinrat Guba (NREL)

The authors gratefully acknowledge research support from the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Hydrogen and Fuel Cell Technologies Office through the Hydrogen Storage Materials Advanced Research Consortium (HyMARC). This work was supported by the Laboratory Directed Research and Development (LDRD) program at Sandia National Laboratories.

HyMARC Hydrogen Materials Advanced Research Consortium

LDRD LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Motivation: Discovering improved H_2 storage and generation techniques will facilitate progress towards a hydrogen economy

Personal transportation (Toyota/Hyundai)

Air Travel (Airbus)

Home energy storage (Lavo)

Utility-scale generation (Shell)

Motivation: Do undiscovered materials exist that could improve upon conventional H_2 storage and generation technologies?

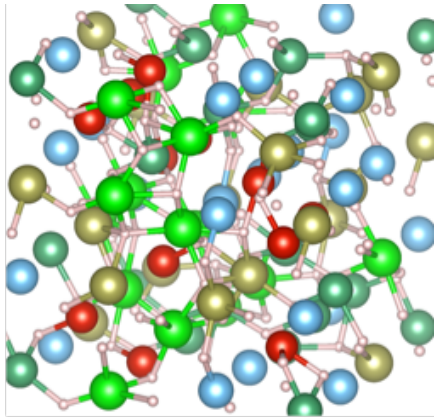
Storage objectives:

- High H_2 gravimetric and/or volumetric density
- Fast, reversible release near ambient T
- Practical/cost-effective

700 bar

"Conventional" (Compressed gas)

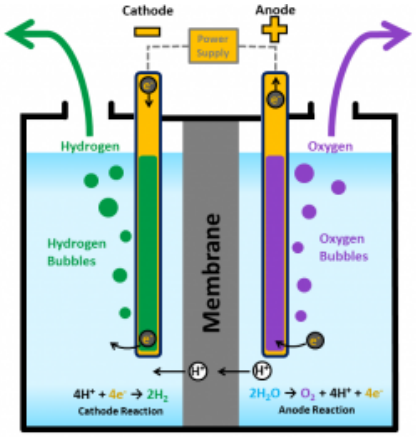
vs.



Material X ??

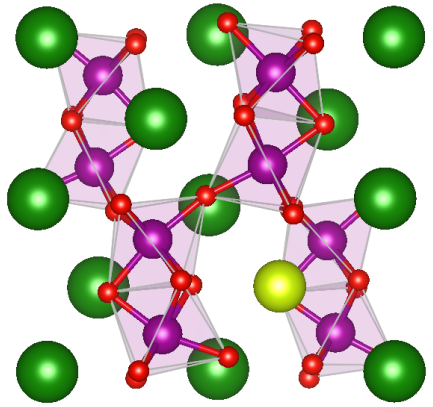
Generation objectives:

- Water-splitting using only renewable energy
- Lower overpotentials than electrolysis
- Practical/cost-effective



"Conventional"
1.2 V in theory
1.8 V in practice

vs.



Material Y ??

Approach: How are new materials discovered?

Experiments

Chemical intuition

↓

Synthesize material

↓

Characterize material

↓

Test Performance

Physics-based simulations

First-principles theory

↓

Run simulation &
predict performance

Machine/statistical learning (ML)

Training data labels/values:
y_i (from experiments or computation)

↓

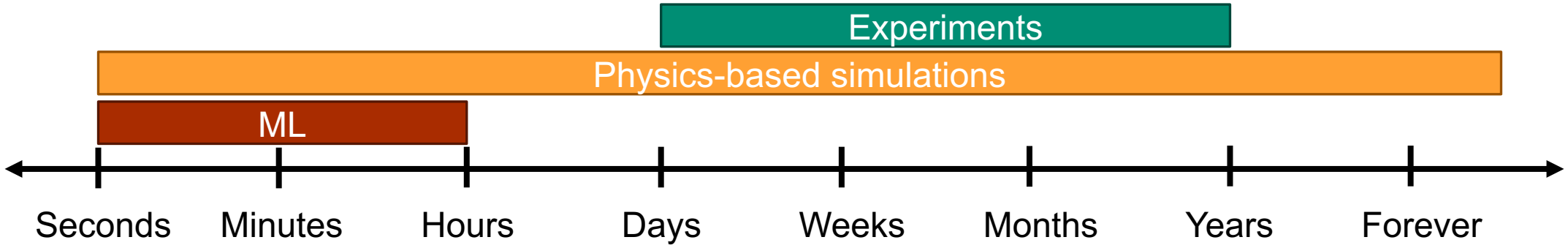
Featurize materials:
$LaNi_5 \rightarrow \boldsymbol{x} = \{\bar{v}_{pa}, \bar{r}_{cov}, \bar{\chi}, \ldots\}$

↓

Train a model:
$\hat{y}_i = f(\boldsymbol{x}_i; \theta)$ and $\mathcal{L} = \sum(\hat{y}_i - y_i)^2$

↓

Run model &
predict performance



Approach: Use of data science and machine learning techniques can…

1) Model material properties *and* provide interpretable predictions to improve fundamental understanding [1]

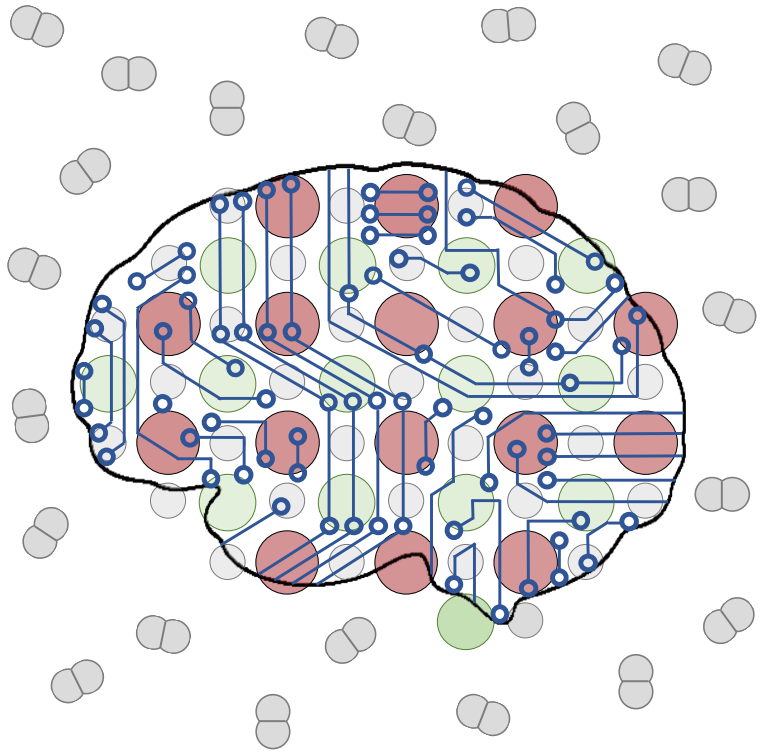
2) Elucidate the design rules (structure-property relationships) for optimal materials [1,2]

3) High-throughput screen large material space to identify top candidates orders of magnitude faster than brute-force experiments or physics-based simulations [3,4,5,6]

4) Accelerate physics-based simulations (critical when experimental training data is lacking) [7]

[1] Witman, Ling, Grant, Walker, Agarwal, Stavila, Allendorf. *J. Phys. Chem. Lett.*, 11 (1), **2020**
[2] Witman, Ling, Stavila, Wijeratne, Furukawa, Allendorf. *Mol. Sys. Des. & Eng.*, 5, **2020**
[3] Ek, Nygard, Pavan, Montero, Henry, Sorby, Witman, et al. *Inorg. Chem.*, 60 (2), **2021**
[4] Witman, Ek, Ling, Chames, Agarwal, Wong, Allendorf, Sahlberg, Stavila. *Chem. Mater.,* **2021***, Accepted*
[5] *In preparation*
[6] Ambrosini, Witman, McDaniel. *Provisional Patent*, **2020**.
[7] *In preparation*

H_2 **Storage**: A data driven roadmap to discover optimal hydrogen storage alloys



Milestone #1:
Develop an explainable ML framework to predict metal hydride thermodynamics and understand design rules [1]

Milestone #2:
Demonstrate our ability to identify and synthesize new hydrides exhibiting targeted thermodynamic properties [4]

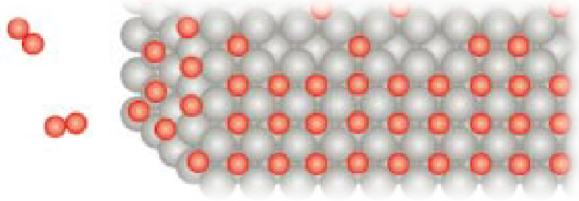
Milestone #3:
Discover the highest performing hydrides across multiple objectives (i.e., Pareto optimal) from an intractably large alloy space [5]

[1] **Witman**, Ling, Grant, Walker, Agarwal, Stavila, Allendorf. *J. Phys. Chem. Lett.*, 11 (1), **2020**
[4] **Witman**, Ek, Ling, Chames, Agarwal, Wong, Allendorf, Sahlberg, Stavila. *Chem. Mater.,* **2021***, Accepted*
[5] *In preparation*

Milestone #1: Explainable ML models predict metal hydride thermodynamics and elucidate simple, first-order design rules

$\ln(P^o_{eq}/P_o)$ target property

H_2 @ P_{eq}, T Metal Hydride



*Intractable for rigorous, atomistic simulation in high-throughput

Training data

HydPARK metal-hydride database

$$\ln(P^o_{eq}/P_o) = -\frac{\Delta H}{R(25^oC)} + \frac{\Delta S}{R}$$

400 out of 2500 are ML-ready…

Featurizing materials

Features derived **only from composition**

e.g. $Zr_{0.04}Ti_{0.96}Nb_{0.04}Fe_{0.95}$ →

$$\boldsymbol{x} = \{\bar{\nu}_{pa}, \bar{r}_{cov}, \bar{\chi}, \ldots\} \in \mathbb{R}^{145}$$

$$\bar{\boldsymbol{\nu}}_{\boldsymbol{pa}} = \sum_i f_i \nu_i$$

$f_i \equiv$ composition fraction

$\nu_i \equiv$ ground state vol. per atom

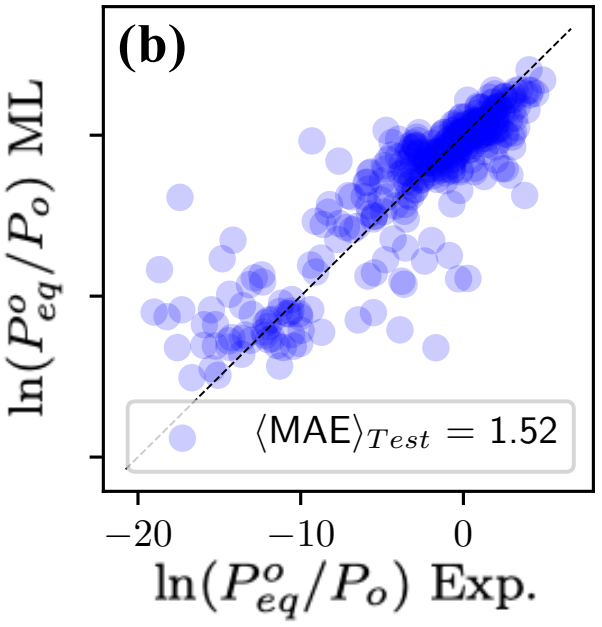
Milestone #1: Explainable ML models predict metal hydride thermodynamics and elucidate simple, first-order design rules

Gradient boosting trees performance

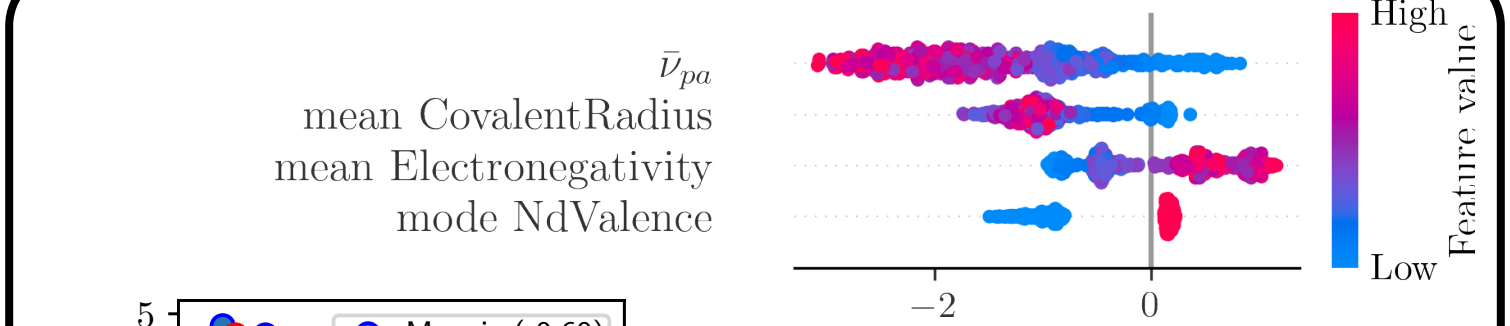
$y_i = \ln(P^o_{eq}/P_o)$ Exp.

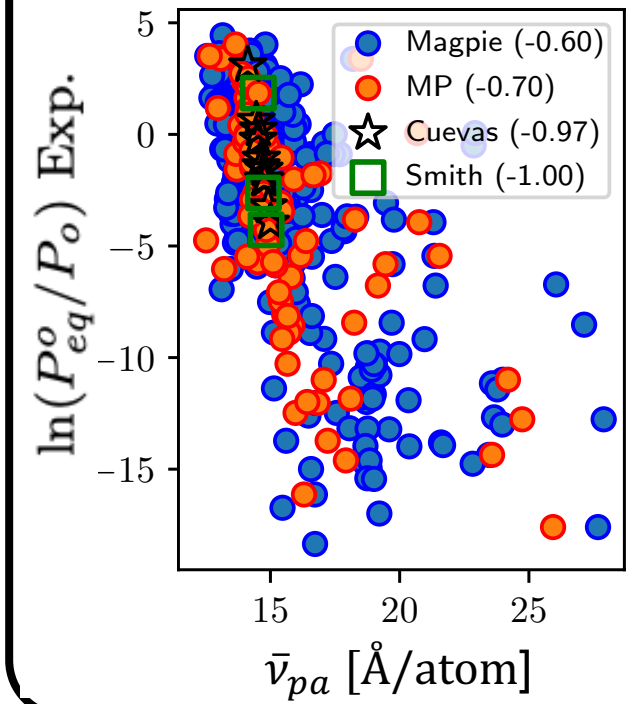
$\hat{y}_i = \sum_{k=1}^{K} f_k(\boldsymbol{x}_i)$ ML

$\mathcal{L} = \frac{1}{n}\sum_i (\hat{y}_i - y_i)^2$ Loss



Model explainability



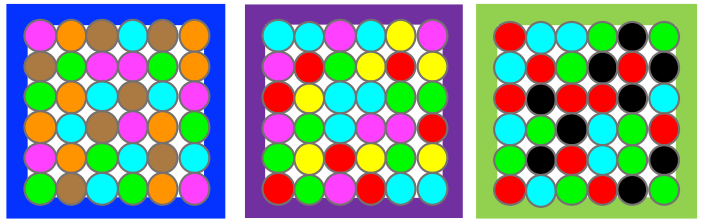


First-order design rule applicable across structural classes & diverse composition space!

$$\ln(P^o_{eq}/P_o) \approx C_1\,\bar{\nu}_{pa} + C_2$$

Milestone #2: ML model rapidly screens for high entropy alloy (HEA) hydrides with a desired thermodynamic stability

HEA overview:



- > 4 elements, ~ equimolar
- Solid solution character necessitates a compositional ML model

Enumerating refractory HEA (rHEA) space:

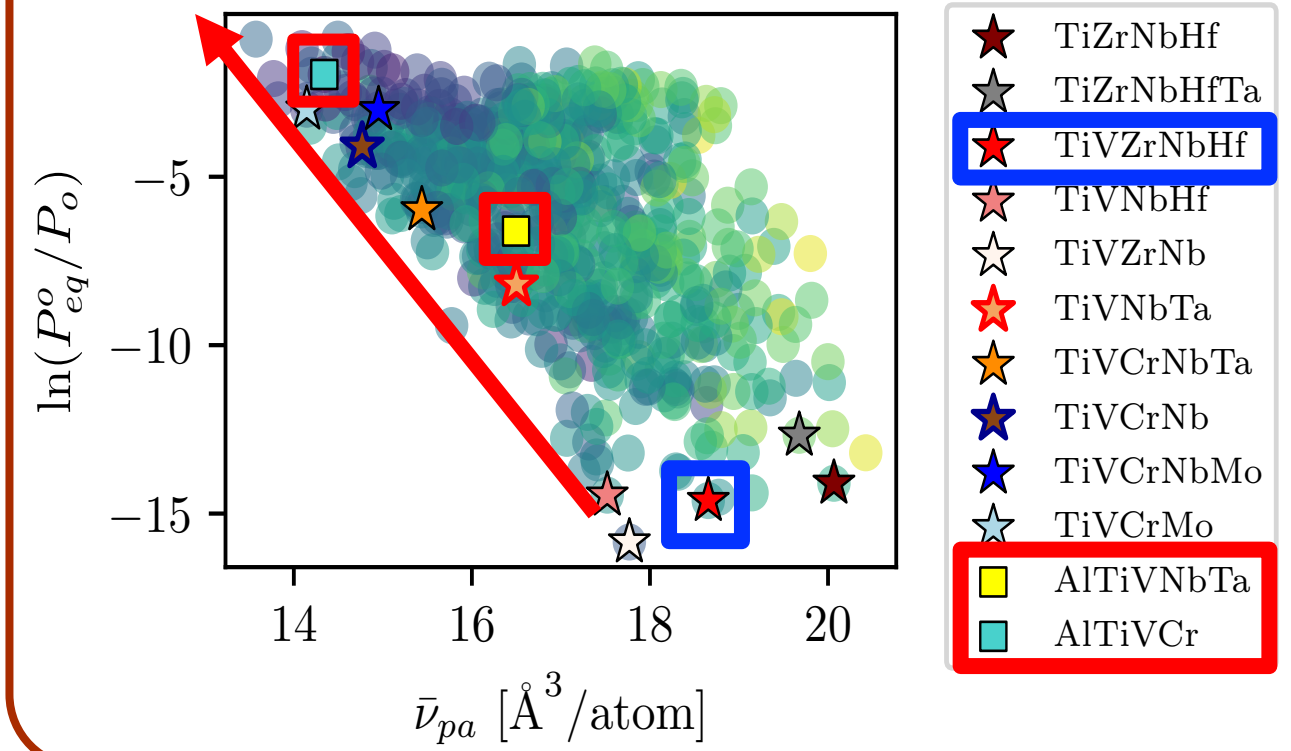
E = {Al, Ti, V, Cr, Zr, Nb, Mo, Pd, Hf, Ta}

$$\binom{E}{4} + \binom{E}{5} + \binom{E}{6} \rightarrow 672 \text{ compositions}$$

Far too many for experiments…

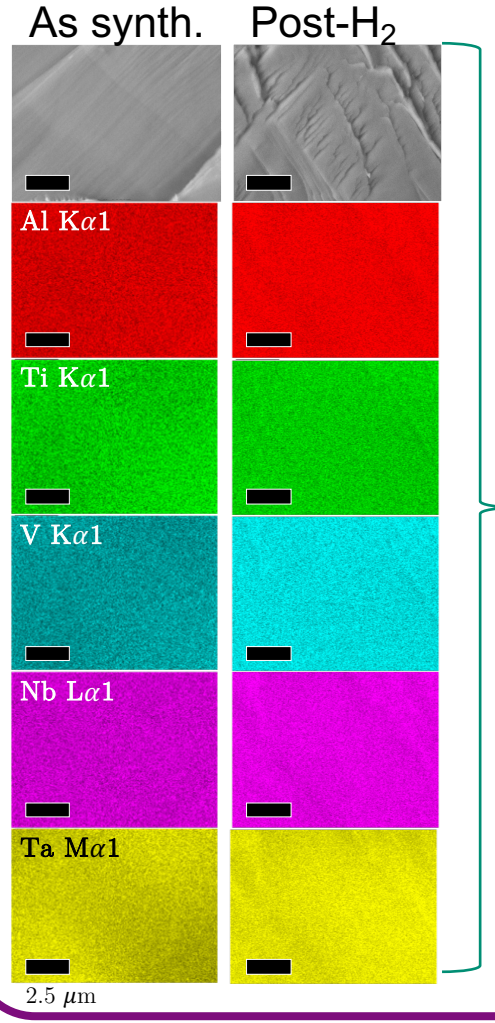
Screening rHEA space:

Novel, destabilized HEA hydrides



Milestone #2: ML model rapidly screens for high entropy alloy (HEA) hydrides with a desired thermodynamic stability

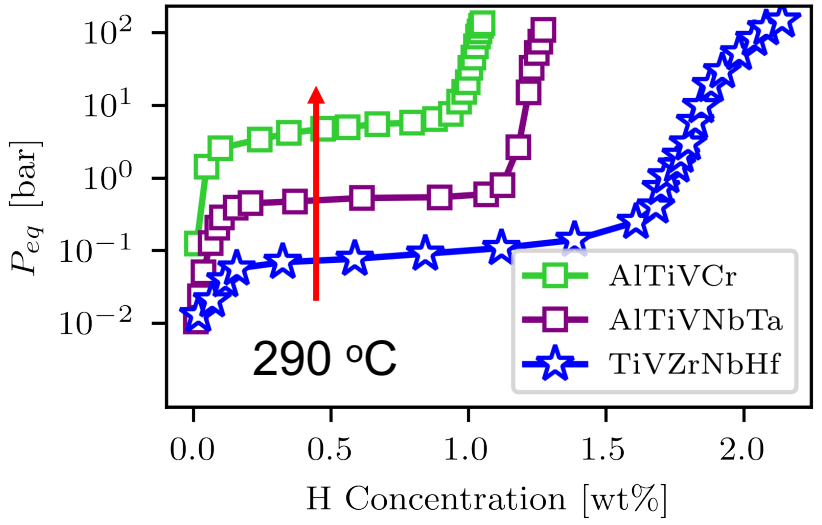
AlTiVNbTa & AlTiVCr synthesis

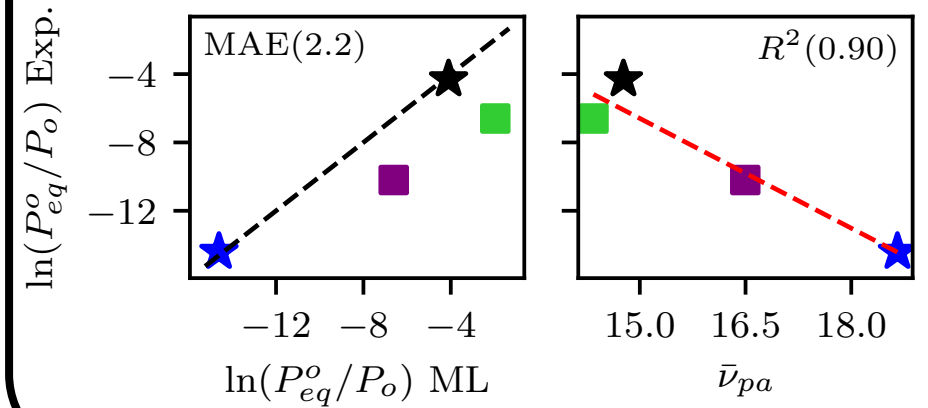


No elemental segregation

ML model & design rule confirmed by experiments

Successfully targeted destabilized hydrides (increase in P_{eq})





Validated ML model & design rule

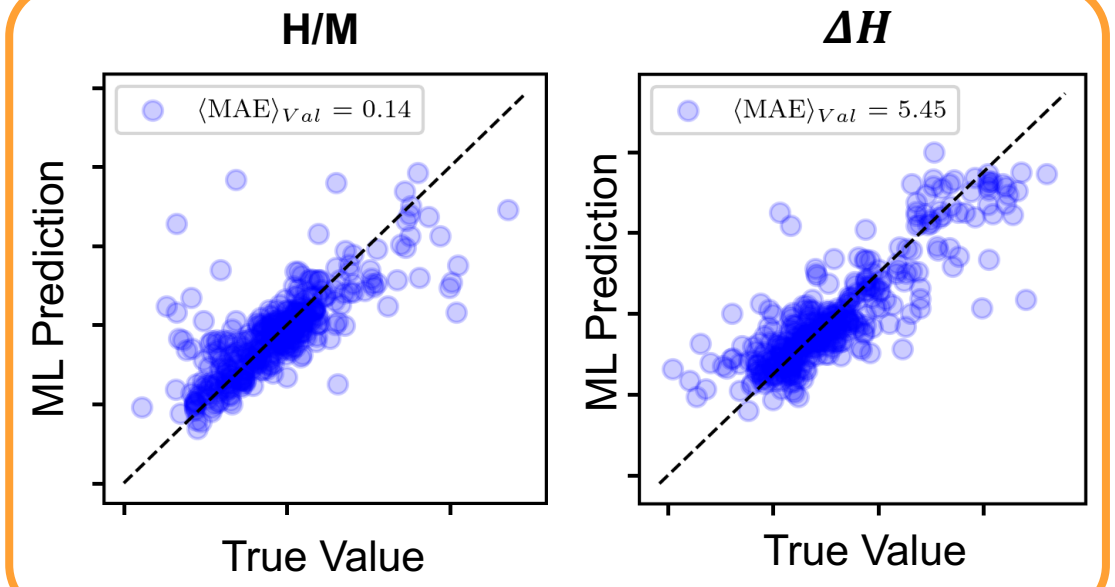
Milestone #3: Identify the Pareto optimal hydrides from an expansive space of ~21,000 candidate HEAs [3]

Screening a novel, expansive HEA space

E = {Mg, Al, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Zr, Nb, Mo, Pd, Hf, Ta}

$$\binom{E}{4} + \binom{E}{5} + \binom{E}{6} \rightarrow 20{,}944 \text{ compositions}$$

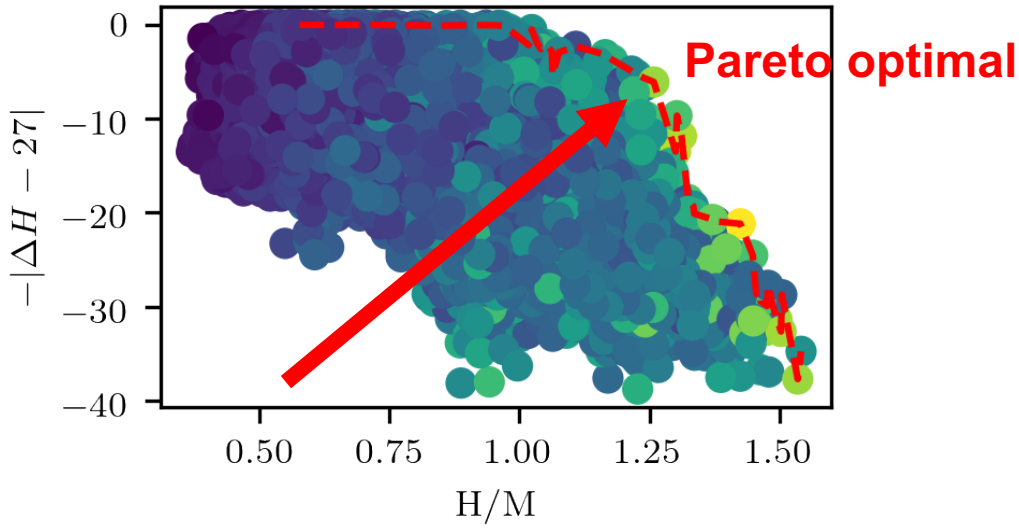
Multiple ML property predictions



Identification of ~100 Pareto optimal materials

Objectives / Quantity to maximize:

- Optimal thermodynamics -> $-|\Delta H - 27|$
- High volumetric capacity -> H/M
- High gravimetric capacity -> Hwt%
- Low Cost -> -cost [USD/kg]



ML: ***Seconds*** to identify all Pareto optimal candidates
Experiment: ***Years*** to identify one by trial-and-error

Ongoing Milestones: Address the lack of robust, standardized, and centralized hydrogen storage data

NREL Data Hub app for standardized experimental data reporting

- Better experimental reproducibility
- More standardized data reporting
- More efficient data collection
- Better ML models!

Experimental parameters/metadata (string fields):
1. Composition = TiVNbCr
2. Mechanical preparation = Arc Melting
3. Branch = Desorption

...TBD

Specialty parameters/metadata (string fields):
1. If nanoscale hydride: N/A
2. If nanoscale hydride: N/A
3. If nanoscale hydride: N/A

...TBD

Thermodynamic data (variable row data table):

dH	dS	van't Hoff Pts	T	P	H/M	Step
47	122	(des1_348K.csv, 16)	348 K	0.202	1.42	1
		(des2_373K.csv, 9)	373 K	0.600	1.39	
		(des3_388K.csv,10)	383 K	1.08	1.43	

$$\Delta H = -m * \frac{8.3145}{1000} * \left(\frac{1}{T}\right)$$
$$\Delta S = b * 8.3145$$

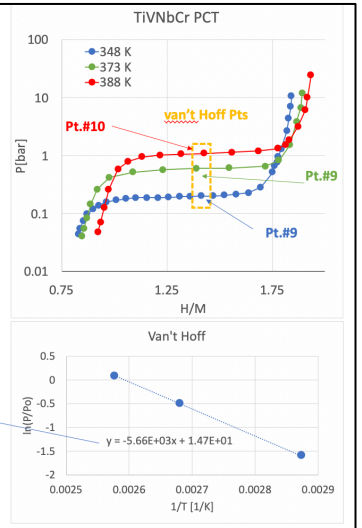
Capacity data:
1. Total capacity [H/M] = 1.9
2. Total reversible capacity [H/M] = 0.9

Step	H/M
1	0.9

# of steps/rows must equal max of Step col above

Kinetic data:

...TBD

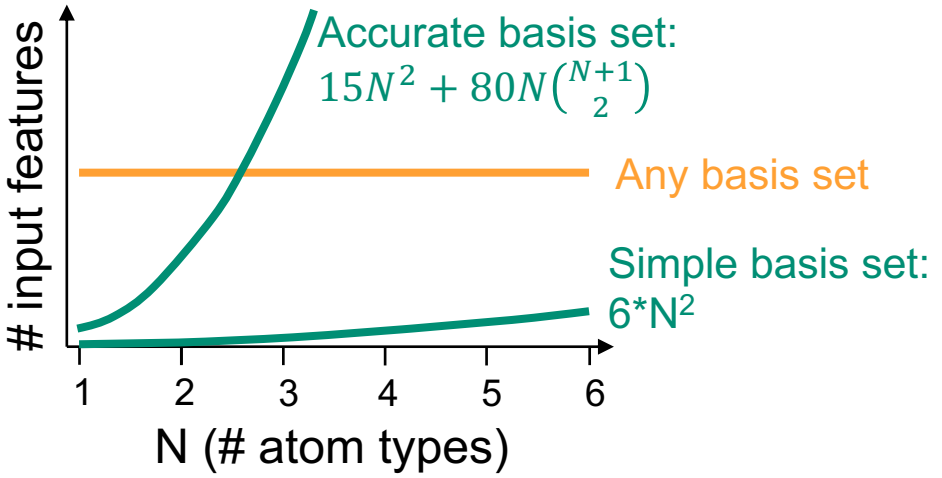


Data that is actually plotted on PCT:
(Each csv has col1=H/M, col2 = P)
des2_373K.csv
des1_348K.csv
pct3_388K.csv

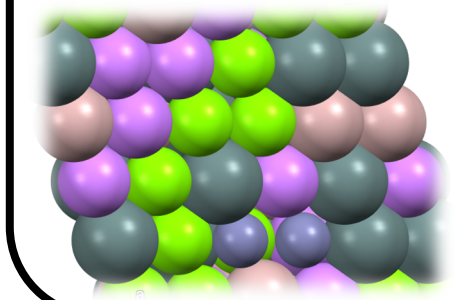
Raw PCT instrumentation file for reproducibility:
(User responsible for creating *.csv above from this file)
des2_373K.dat
des1_348K.dat
des3_388K.dat

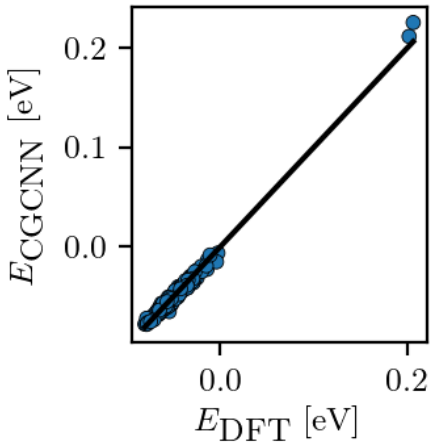
Surrogate models for DFT in many element systems

Conventional vs. *Graph Convolutional* NN



5 element system…
w/accurate potential:





Research Aspirations (in 5 years can we confidently say that we have…)

Storage

- Performed truly comprehensive HEA screening for ***Pareto optimal hydrides?***

Generation

- An ML framework to identify ***optimal metal oxides*** for thermochemical water-splitting?

Data

- Created ***standardized data management tools*** for hydrogen storage experiments?

Simulations

- ML driven simulations that can predict physical/chemical properties of ***many-element hydrogen/material systems***?

More data & better ML methods leads to

- Improved models
- More accurate materials predictions
- Faster experimental success
- *Quicker hydrogen technology adoption*

Thank you for your attention.

Questions?

Contact: mwitman@sandia.gov