Discovering the Underlying Equations Governing Perovskite Solar-Cell Degradation Using Scientific Machine Learning

Richa Ramesh Naik, *Armi Tiihonen, Janak Thapa, Clio Batali, Shijing Sun, Zhe Liu, Tonio Buonassisi*

Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139

Photovoltaic Research Laboratory

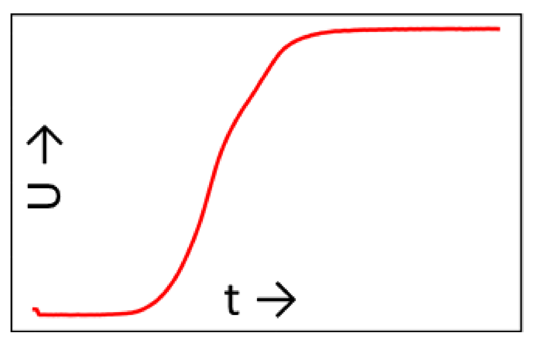
Motivation

- Discovering laws of physics as differential equations *directly from data* would accelerate science across fields
- Recent developments: Scientific ML - automatically learn models from data
- Halide perovskites are potential cost-effective solar energy materials, but they degrade at elevated temperatures and humidity
- We apply scientific ML to discover the laws of perovskite degradation

Scientific ML

- PDE-FIND [2] - sparse regression approach for discovery of physical laws describing dynamical systems
- Library of potential candidate functions consisting of polynomials and other non-linear functions of U built
- Sequential threshold ridge regression algorithm applied to obtain differential equation (DE) from the data.

Collect data



Build non-linear library of data

$$U_t = \begin{bmatrix} 1 & U & U^2 & \dots \end{bmatrix} \beta$$

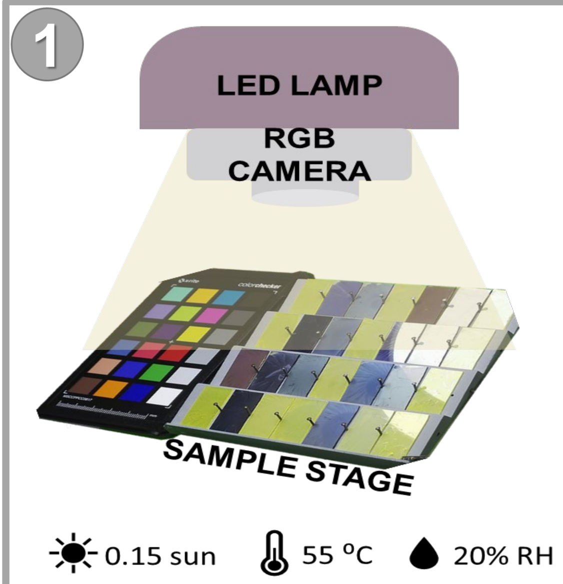
Identify dynamics through sparse regression

$$\arg\min_{\beta} \left(\|\Theta(U)\beta - U_t\|_2 + \lambda_2\|\beta\|_2 + \lambda_0\|\beta\|_0 \right)$$

Methods

Experimental Process

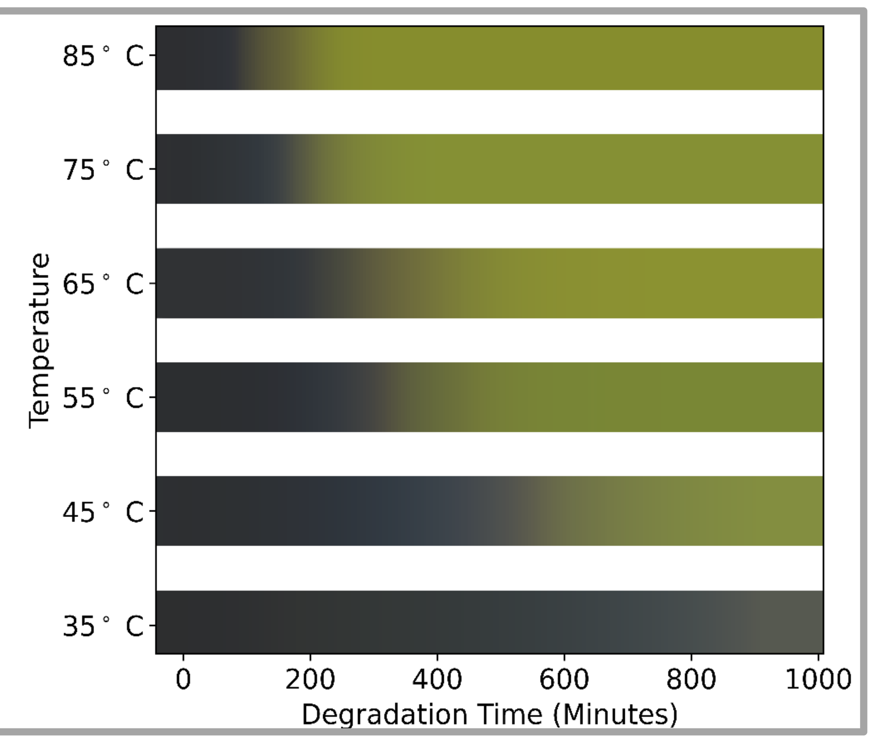
1



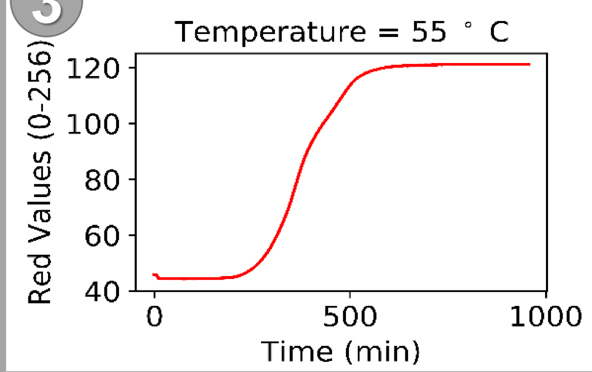
Methylammonium lead iodide (MAPI) thin-film samples degraded at 0.15 Sun, 20% RH, and 35 to 85°C temperature in our environmental chamber

2

Average color of degrading films obtained as a function of time



3



Red color vs time captures the temporal perovskite decomposition

Simulated Data

- Experimental data fit to the Verhulst logistic equation[1] and Arrhenius equation
- Simulated data created with & without Gaussian noise

DOMAIN KNOWLEDGE

Logistic Function

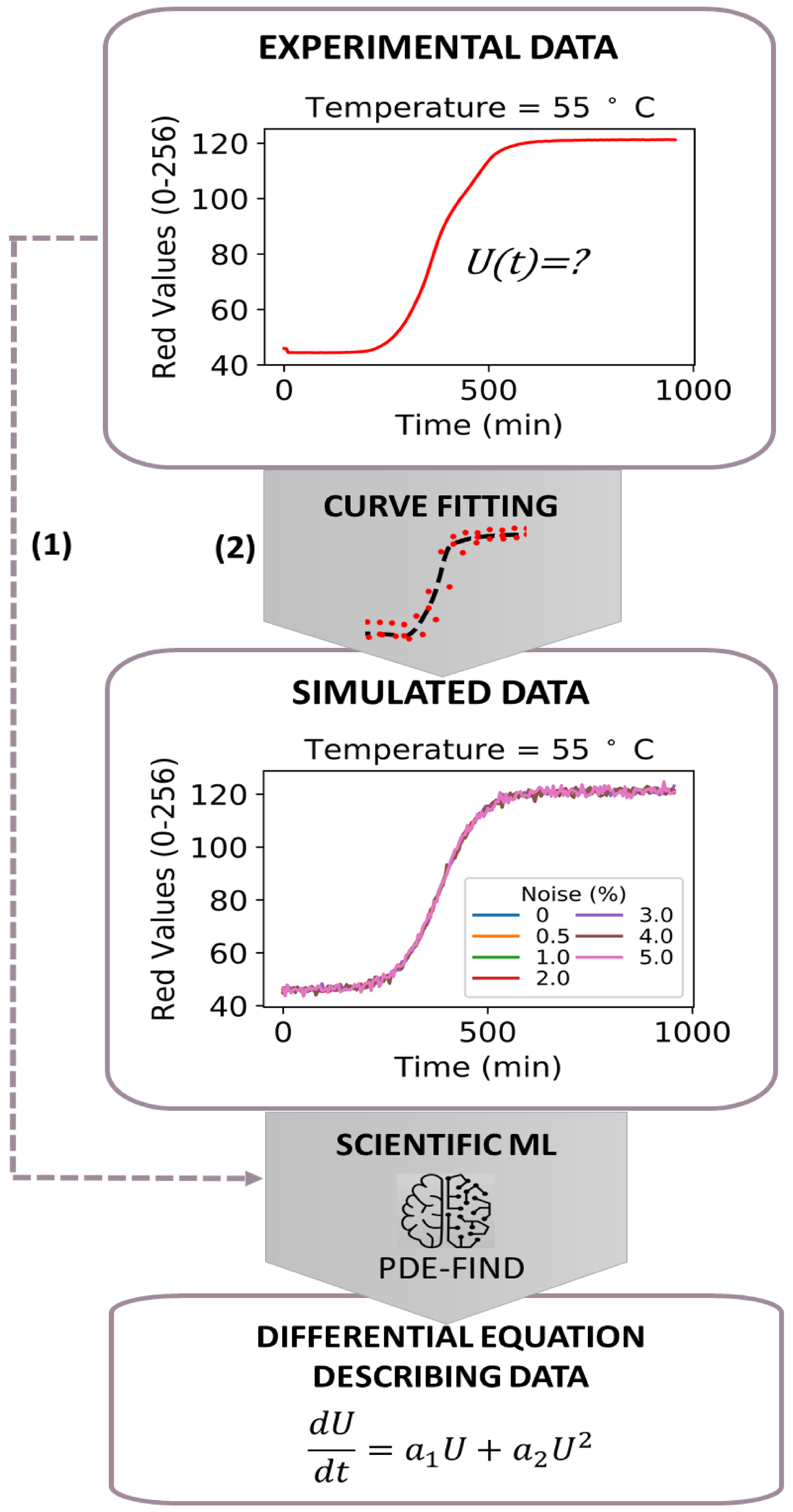
$$U = M + \frac{U_o K e^{kt}}{(K - U_o) + U_o e^{kt}}$$

Arrhenius Equation

$$k = Ae^{\left(-\frac{E_a}{RT}\right)}$$

Analysis Workflow

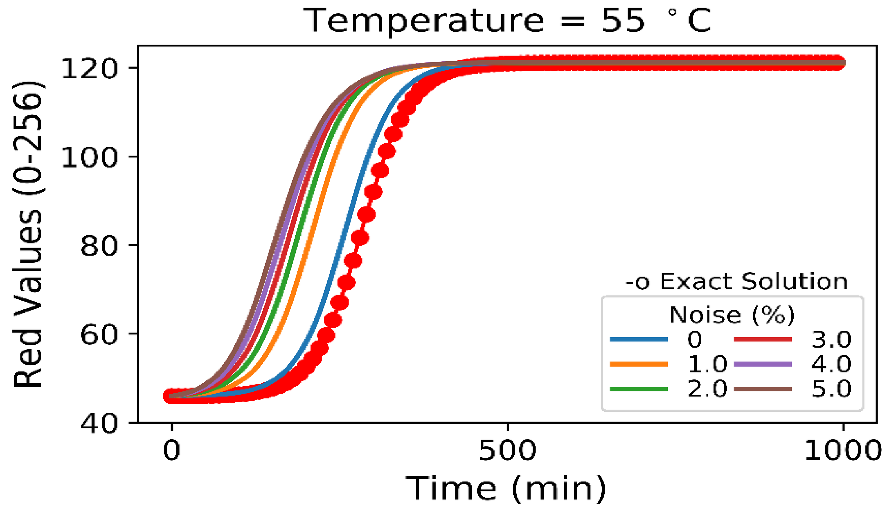
We apply PDE-FIND [2] to Experimental data (Workflow (1)) and Simulated data with and without Gaussian noise (Workflow (2))



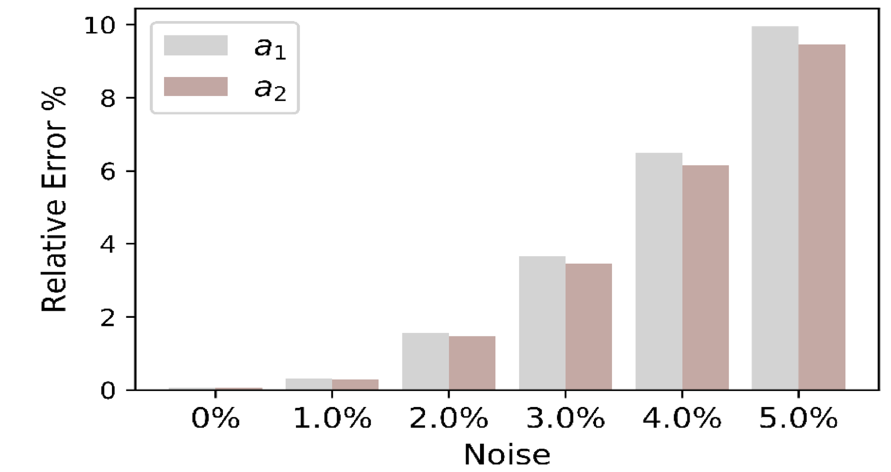
Results

Equation Identified	$\frac{dU}{dt} = a_0 + a_1 U + a_2 U^2$		
T = 55 °C	**Parameter Estimates**		
Noise	$\mathbf{a_0}$	$\mathbf{a_1}$ $(\times 10^{-2})$	$\mathbf{a_2}$ $(\times 10^{-4})$
0 %	0.000	2.078	-2.763
1.0 %	0.000	2.070	-2.753
2.0 %	0.002	2.044	-2.721
3.0 %	0.004	2.001	-2.666
4.0 %	0.007	1.942	-2.591
5.0 %	0.010	1.870	-2.500
Exact Solution	**0**	**2.076**	**-2.761**

- Underlying DE describing the simulated data identified when up to 5% Gaussian noise is added



- Relative error in the parameters accompanying the function terms is 10% at 5% Gaussian Noise.



- PDE-FIND on experimental data yields DE that does not to fit data

Conclusions

- We showcase scientific ML (PDE-FIND) on perovskite degradation data
- Scientific ML can aide with understanding the underlying scientific phenomena, make simulations faster and extrapolate beyond our dataset.
- In its current state, is well-suited to be applied to domains where obtaining large quantities of low-noise data is possible, and will find more applications with methods that are robust to noise

Acknowledgements

We acknowledge support from DARPA, TOTAL, and DOE.

References

1. Tsoularis, A. & Wallace, J. Analysis of logistic growth models. Mathematical Biosciences 179, 21–55 (2002).
2. Rudy *et al.* Data-driven discovery of partial differential equations. Science Advances 3, e1602614 (2017).