Learning-based Diagnosis and Control for Non-Equilibrium Plasmas

Ali Mesbah

Department of Chemical and Biomolecular Engineering
University of California Berkeley

Berkeley
UNIVERSITY OF CALIFORNIA

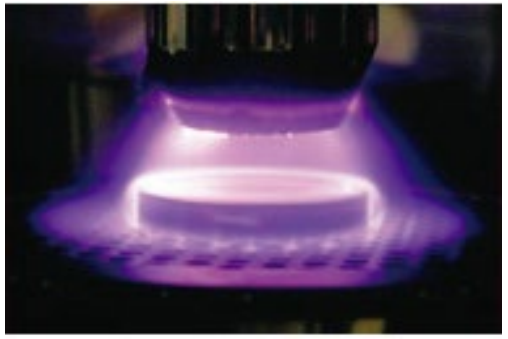
Non-Equilibrium Plasmas

Treatment and processing of complex surfaces

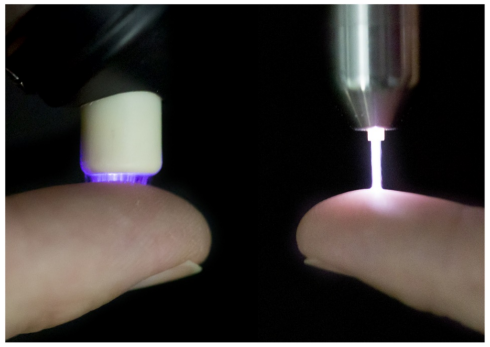
Plasma Etching Process

Surface Coating

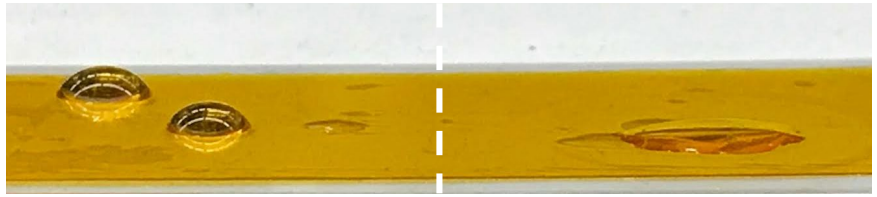
Surface Functionalization



Biomaterials Processing



Water droplets on polyimide film



Not treated Plasma treated

Complex plasma-surface interactions:

- Hard-to-model dynamics across multiple time- and length-scales
- Time-varying surface characteristics
- Lack of real-time diagnostics of surface properties

Grand challenge: Reproducible and precise control of plasma-surface interaction mechanisms

https://www.asbindustries.com/thermal-spray-coatings
http://www.igs.titech.ac.jp/iper/english/iper2/6/detail_44.html
K-D Weltmann and Th von Woedtke, 2016.

Machine Learning for Non-Equilibrium Plasmas

Characterization and control of complex plasma-surface interactions

Unique cross-disciplinary opportunities to exploit advances in machine learning and artificial intelligence in the field of non-equilibrium plasma:

- Modeling and simulation
- Real-time diagnostics
- Learning-based control

Letters

IOP Publishing

Journal of Physics D: Applied Physics

J. Phys. D: Appl. Phys. **52** (2019) 30LT02 (9pp)

https://doi.org/10.1088/1361-6463/ab1f3f

Letter

Machine learning for modeling, diagnostics, and control of non-equilibrium plasmas

Ali Mesbah and David B Graves

Machine Learning for Non-Equilibrium Plasmas

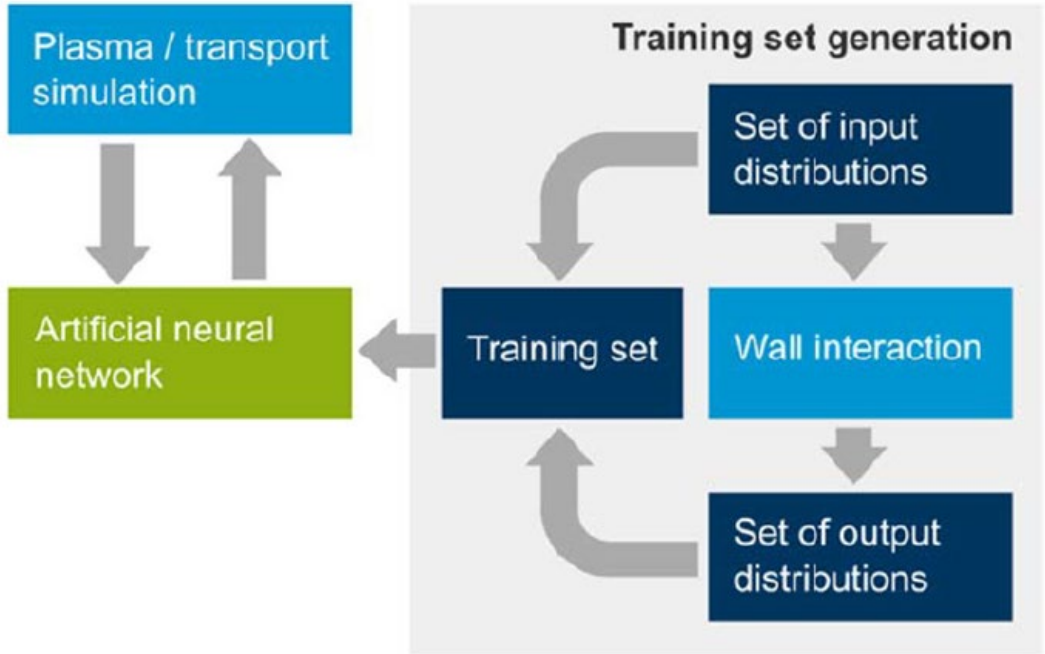
Modeling and simulation – Surrogate modeling

Learn cheap-to-evaluate surrogate models for physics-based predictive models:

- Computational complexity of plasma simulations increases when incorporating surface effects of plasma that occur across multiple length- and time-scales
- Supervised learning for constructing surrogate models from high-fidelity simulations
- Useful for design of experiments, design space visualization, sensitivity analysis, etc.

Multiscale plasma-surface models:

- Deep learning for modelling interactions of energetic particles with a surface and the subsequent particle transport in thin film formation via plasma sputtering
- Significantly more efficient than combining theoretical models



Kruger et al., *PSST*, 2019.

Machine Learning for Non-Equilibrium Plasmas

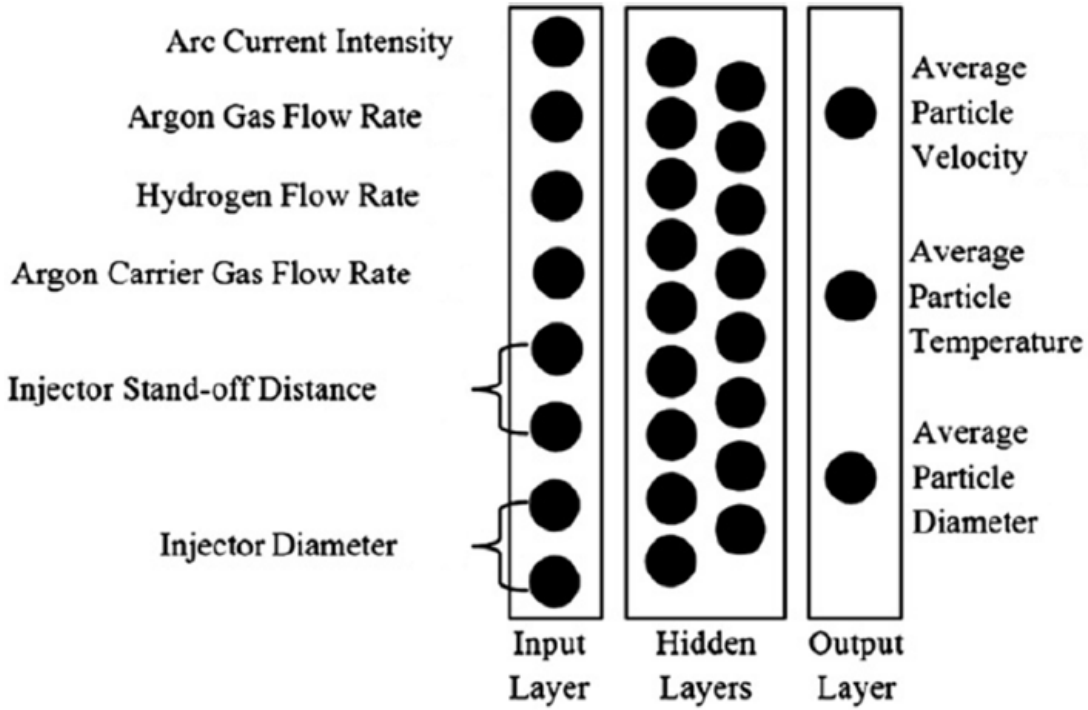
Modeling and simulation – Data-driven modeling

Learn models for plasma-surface interactions from experiments:

- Lack of comprehensive theoretical models for plasma-surface mechanisms
- Supervised learning for constructing nonlinear, multidimensional functional relationships from input-output data, without prior assumptions

Useful for:

- Elucidating the effects of plasma parameters and/or surface characteristics on plasma-surface interactions
- Learning inverse relationships for systematic exploration of the process design space (when process parameter is enormous)



Choudhury et al., *Surf. Coat. Technol.*, 2011.

Outline

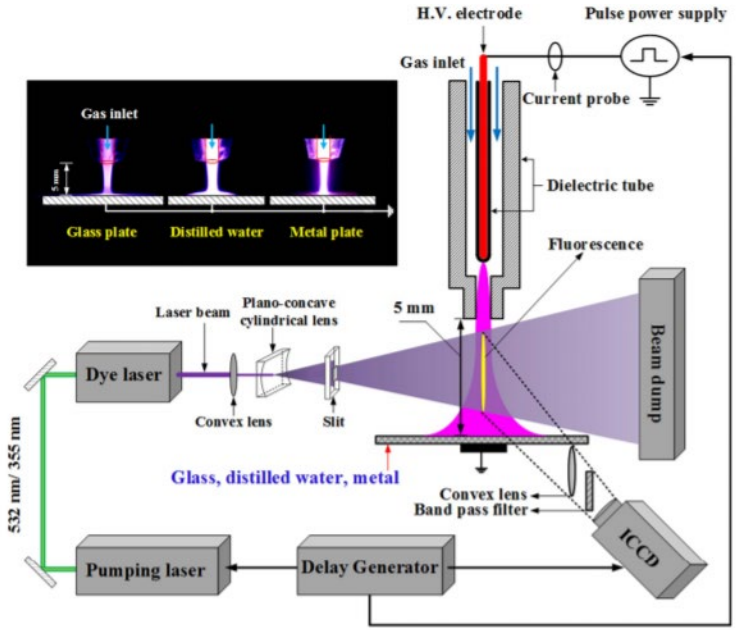
Machine learning for diagnosis and control

- **Learning-based diagnostics**
  - Real-time diagnosis of plasma and surface properties
  - Process disturbance detection
- **Learning-based control**
  - Optimization-based control of plasma-surface interactions
  - Sim-to-real transfer reinforcement learning

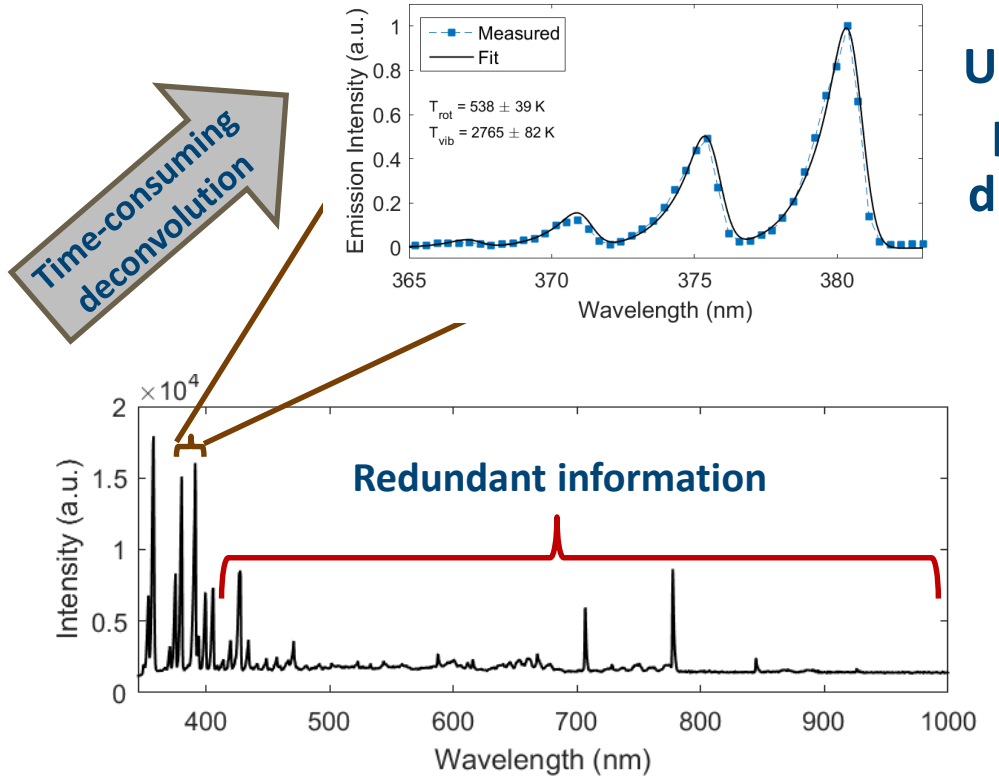
Challenges in Plasma Diagnosis

Real-time plasma and surface diagnosis essential for process control

- Direct measurements require sophisticated instrumentation and specialized experimental configurations that can limit operational flexibility
- Inexpensive and easy-to-implement diagnostics contain implicit information about plasma and substrate, which is computationally expensive to analyze in real-time



Laser-induced fluorescence for direct concentration measurement



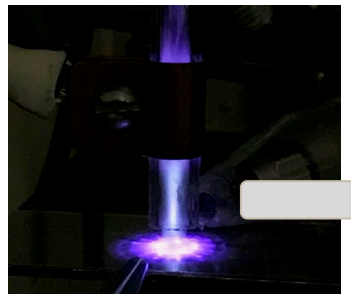
Useful for plasma diagnosis

Dobele et al., *PSST*, 2005.

Real-time Diagnosis of Plasma

Extract process information from spectral data

- Learning-based soft sensors enable inference of plasma characteristics from spectral data in real-time
- Inference of rotational and vibrational temperatures from optical emission spectra

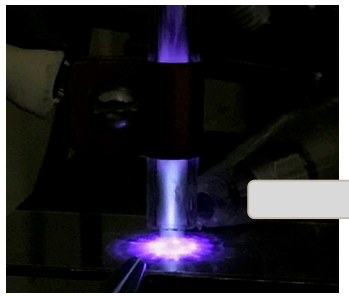


Physics-based Model

Generate Synthetic Spectrum and Fit

Estimated rotational and vibrational temperatures

Slow and iterative procedure, performed offline



Model-based Soft Sensor

Estimated rotational and vibrational temperatures

Real-time diagnosis

Model trained using OES spectra and corresponding rotational and vibrational temperatures

Gidon, Pei, Bonzanini, Graves, and Mesbah, *IEEE TRPMS*, 2019.

Real-time Diagnosis of Plasma

Extract process information from spectral data

Supervised learning: A sparse polynomial regression model trained to predict rotational and vibrational temperatures from OES data:

$$\hat{Y} = \mathbf{w}^\top \mathbf{\Psi}(\mathbf{X})$$

Least-absolute-shrinkage-and-selection operator (LASSO) linear regression

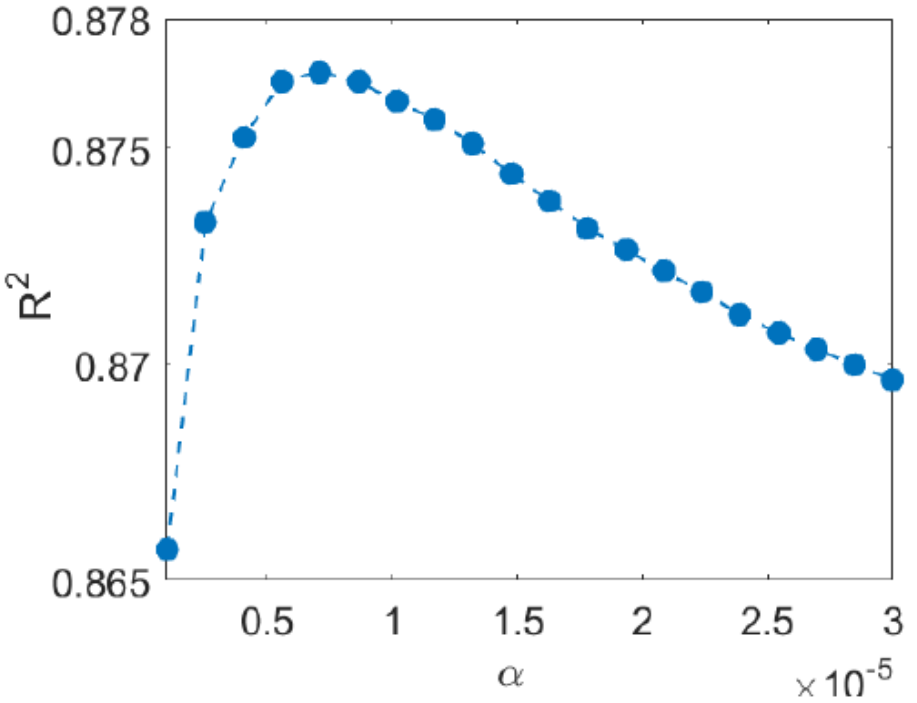
$$\min_{\mathbf{w}} \; ||Y - \mathbf{w}^\top \mathbf{\Psi}(\mathbf{X})||_2^2 + \alpha||\mathbf{w}||_1$$

Polynomial basis defined in terms of features, which are OES intensity at each measured wave length

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{51} \\ x_1^2 \\ x_2^2 \\ \vdots \\ x_{51}^2 \end{bmatrix}$$

Regularization parameter prevents overfitting as order of the model grows

Hyperparameter selection and cross-validation

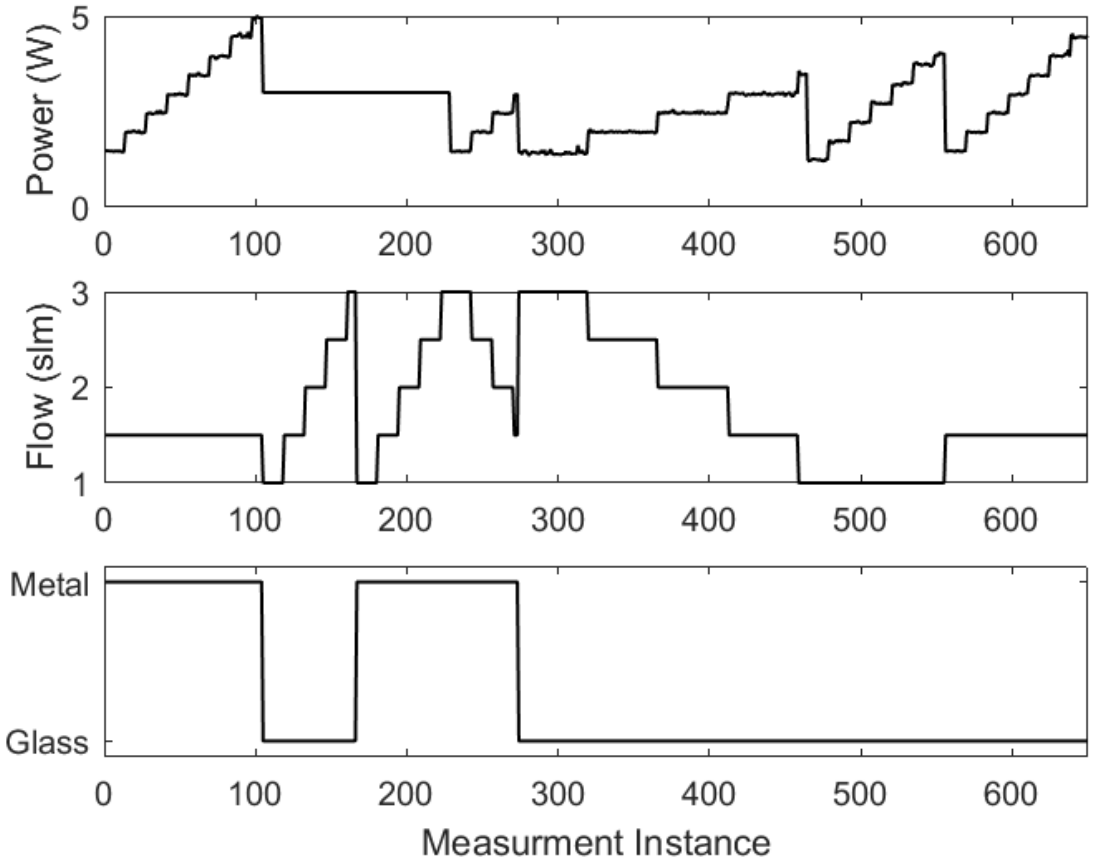


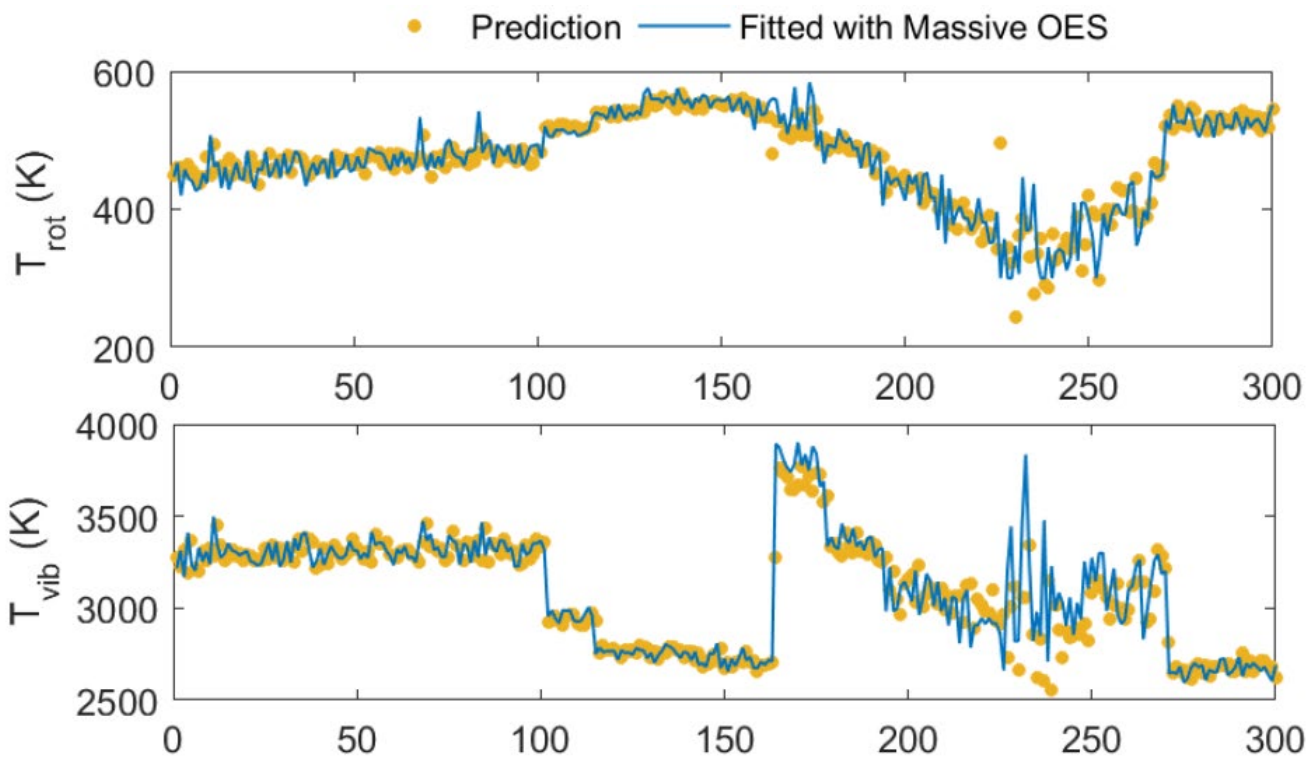
Gidon, Pei, Bonzanini, Graves, and Mesbah, *IEEE TRPMS*, 2019.

Real-time Diagnosis of Plasma

Extract process information from spectral data

Real-time inference of rotational and vibrational temperatures from optical emission spectra as plasma process parameters change in time





Gidon, Pei, Bonzanini, Graves, and Mesbah, *IEEE TRPMS*, 2019.

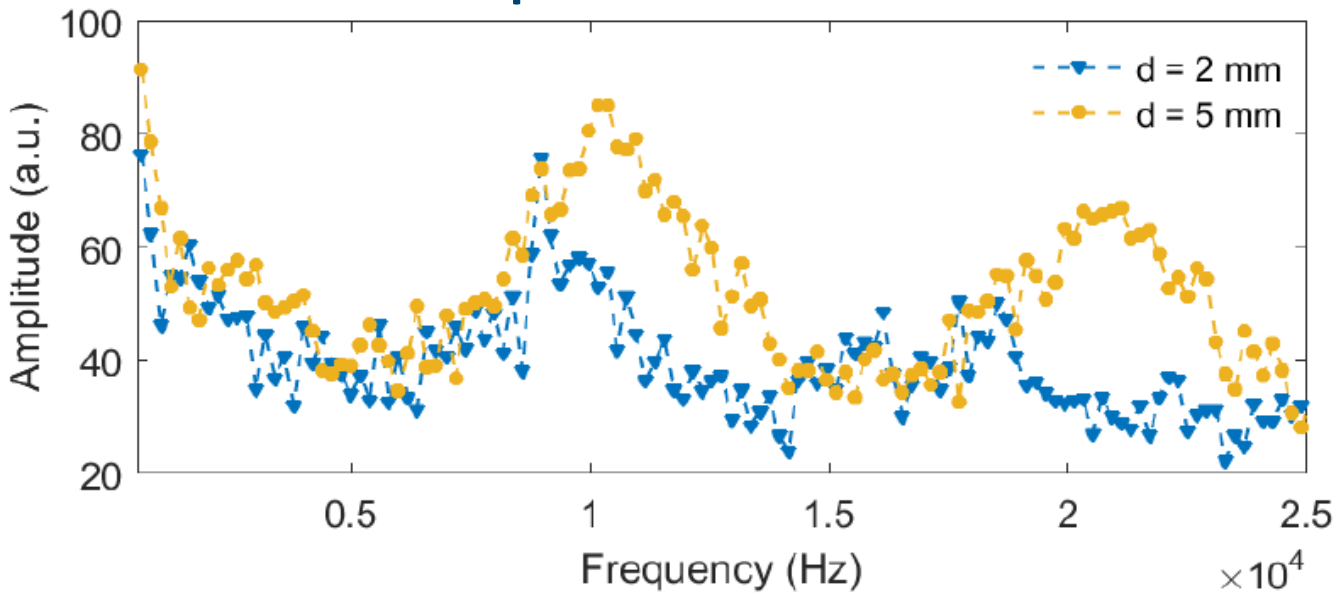
Process Disturbance Detection

Extract process information from spectral data

A learning-based soft sensor for real-time detection of variations in device tip-to-surface separation distance (disturbance) from electro-acoustic emissions

Variations in the device tip-to-surface separation distance

Electro-acoustic emission at two separation distances



Process Disturbance Detection

Extract process information from spectral data

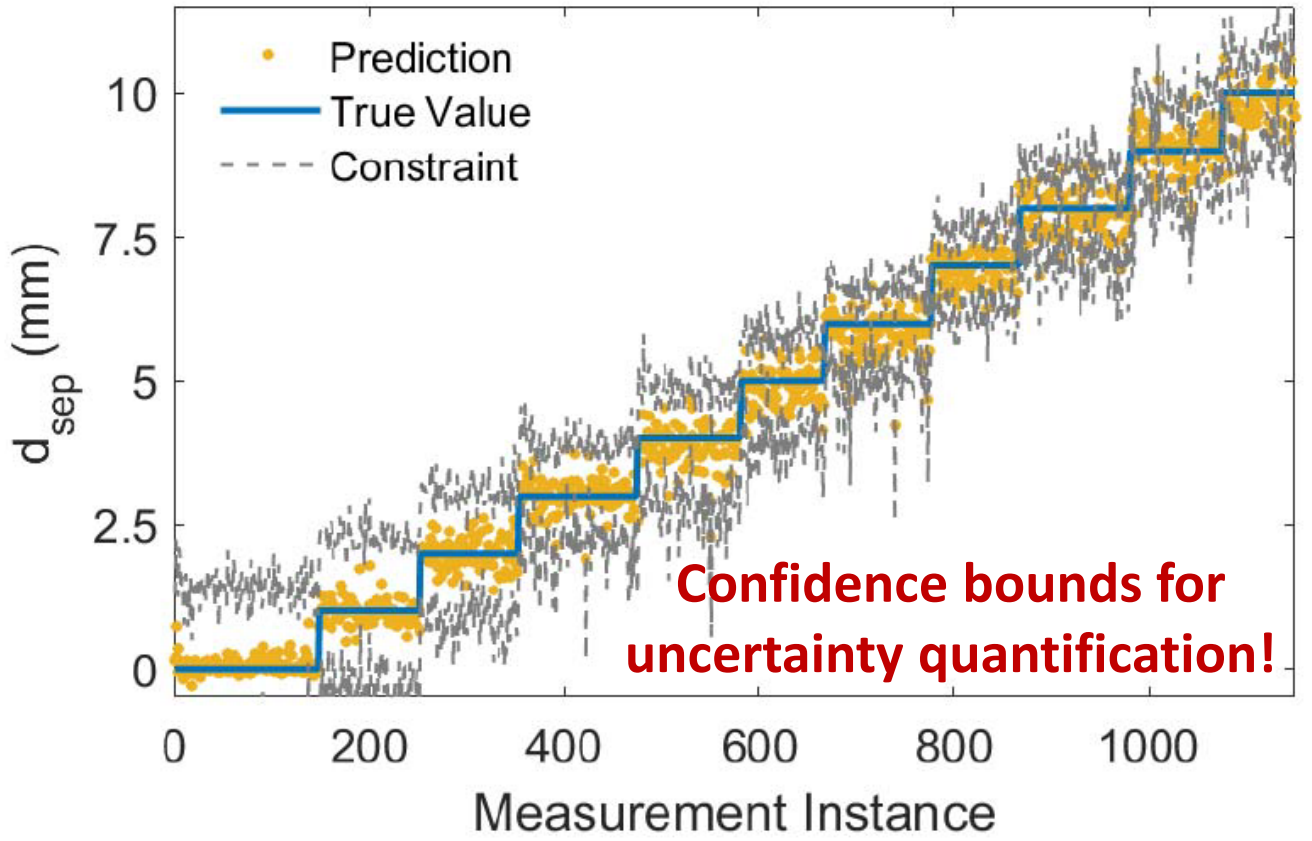
Supervised learning: A Gaussian process (GP) regression model predicts changes in the device tip-to-surface separation distance in real-time

GP predicts probability distribution of process variables conditioned on training data:

$$\begin{pmatrix} Y \\ \hat{Y} \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \Sigma\right)$$

Covariance defined in terms of a distance measure between training and test data

GP provides (probabilistic) uncertainty bounds

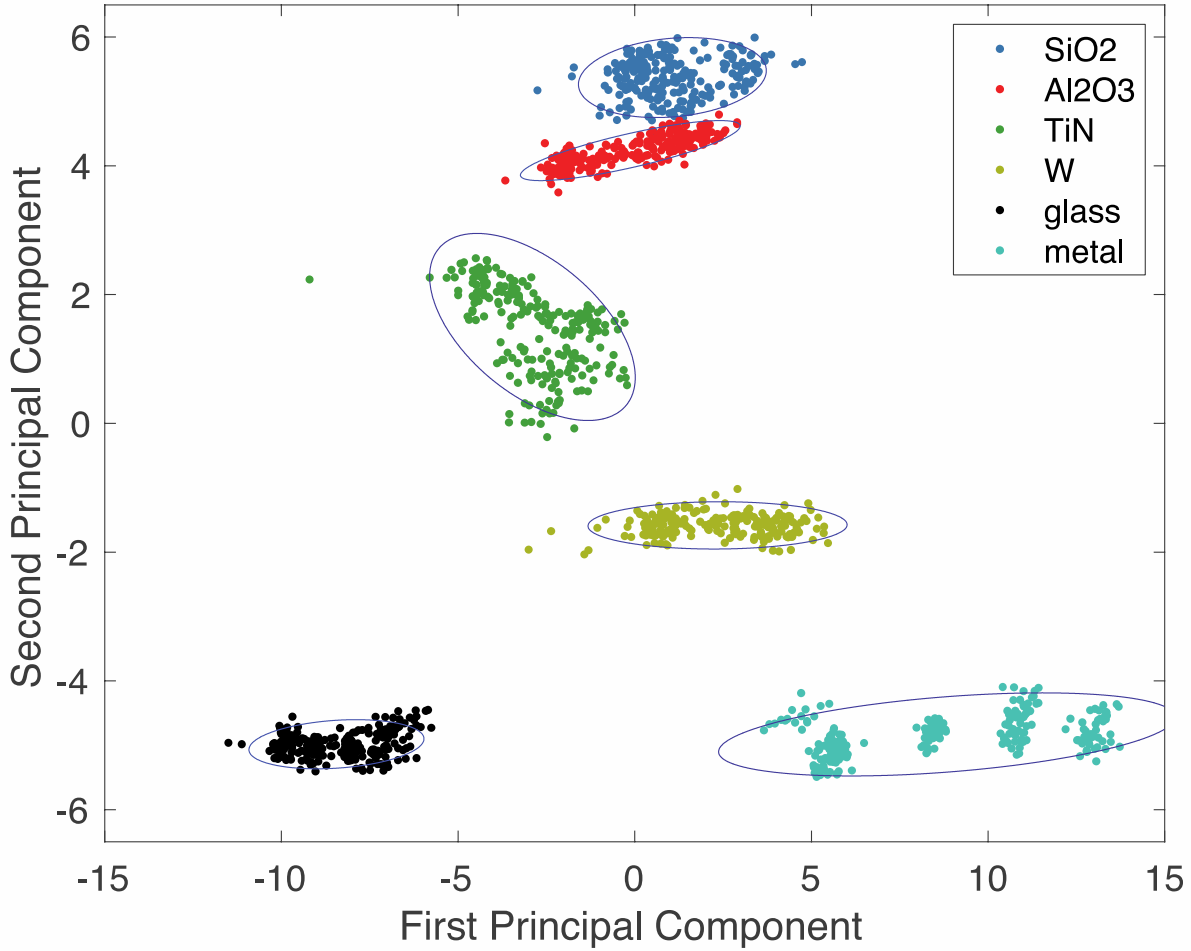


Gidon, Pei, Bonzanini, Graves, and Mesbah, *IEEE TRPMS*, 2019.

Real-time Diagnosis of Surface Properties

Extract process information from spectral data

Supervised and unsupervised learning: Real-time detection of six different surfaces with vastly different thermal and electrical properties



- Leverage the interplay between plasma and surface to detect and monitor changes in the chemical, electrical, or mechanical state of a surface over time
- Dimensionality reduction to extract latent information from OES data
- Pattern recognition for surface classification based on latent information of OES
- Probabilistic quantification of our confidence in predictions of learning-based models

Outline

Machine learning for diagnosis and control

- **Process diagnosis**
  - Real-time diagnosis of plasma properties
  - Process disturbance detection
- **Learning-based control**
  - Optimization-based control of plasma-surface interactions
  - Sim-to-real transfer reinforcement learning

Data-Driven Models for Optimization-based Process Control

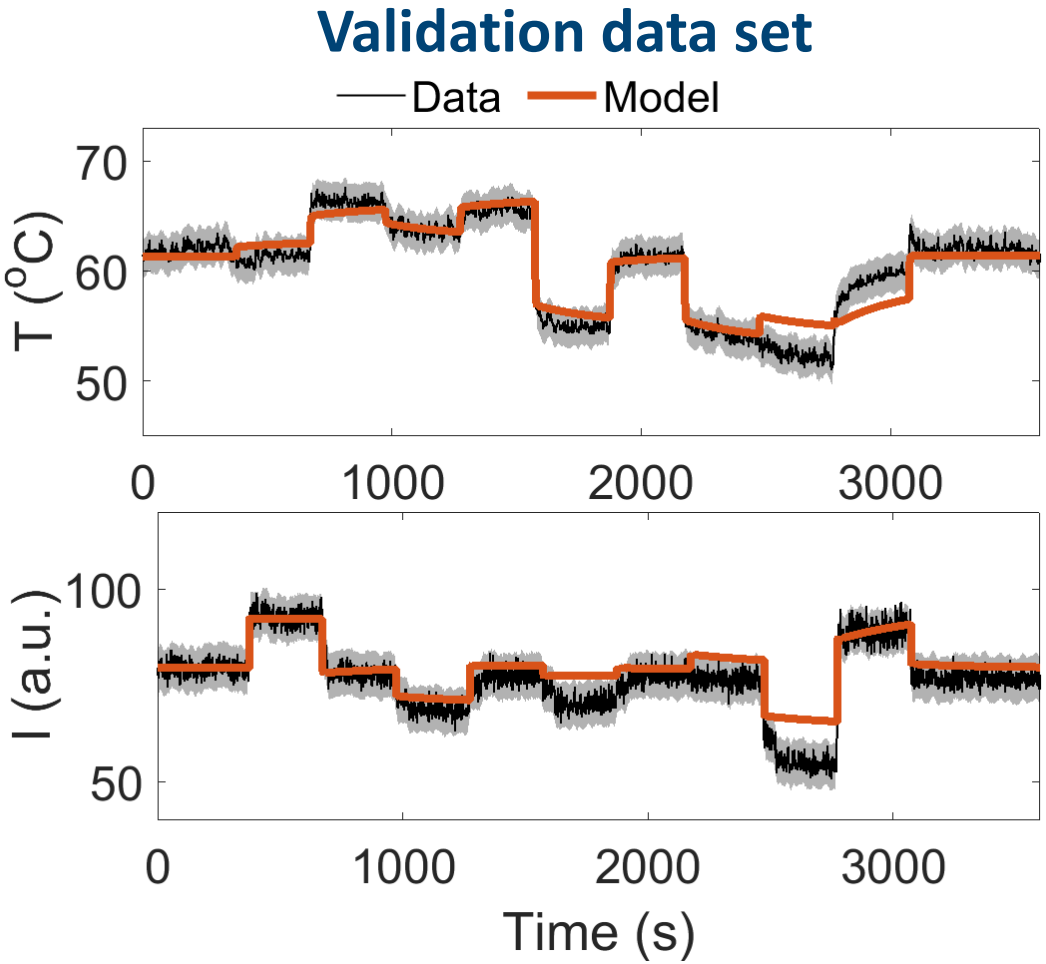
Reduced-order models of multivariable plasma-surface dynamics

Multivariable modeling: The multi-input multi-output nature of dynamics can be described by data (no physics-based knowledge); amenable to online optimization

$$\frac{dx}{dt} = Ax + B\begin{bmatrix} V_{p2p} \\ f \\ q \end{bmatrix}$$

$$\begin{bmatrix} T_s \\ I \end{bmatrix} = Cx$$

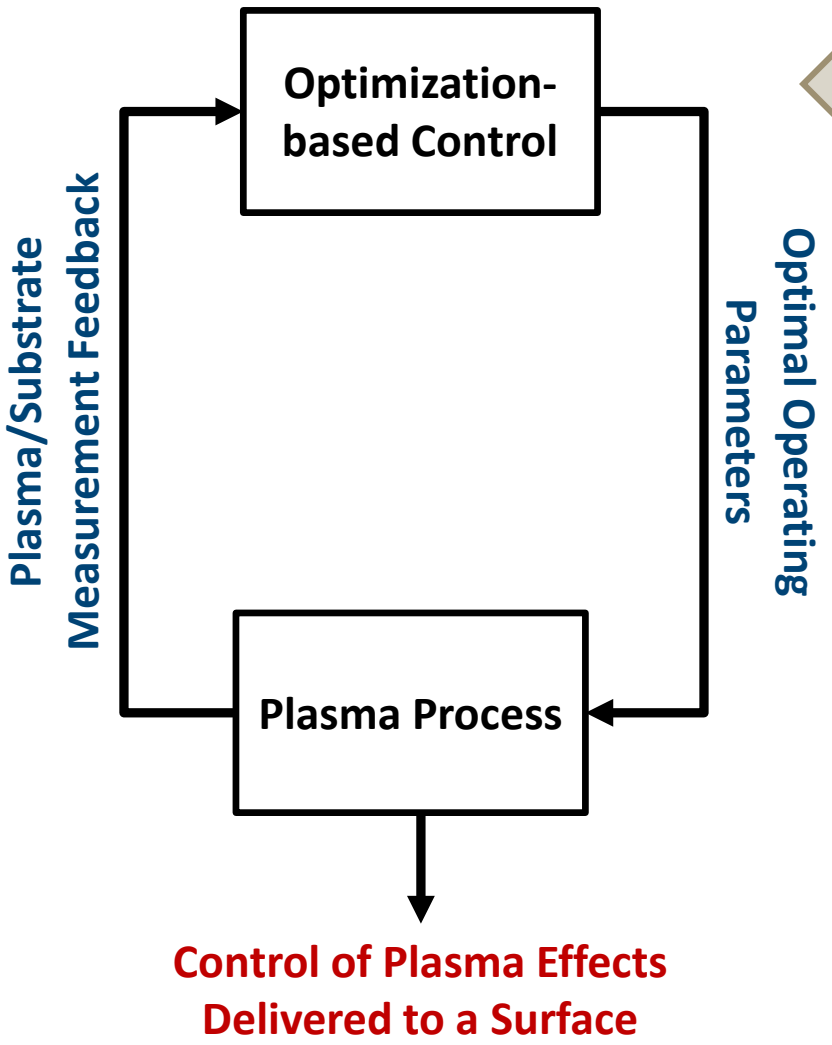
Subspace identification model to describe the input-output behavior of the system while state variables have no physical meaning



Gidon, Curtis, Paulson, Graves, and Mesbah, *IEEE TRPMS*, 2018.

Data-Driven Models for Optimization-based Process Control

Reduced-order models for real-time control of plasma-surface interactions



Online process optimization:

$$\min_{u} J(x,u)$$
$$\text{s.t.: } \frac{dx}{dt} = f(x,u,\theta)$$
$$x \in \mathcal{X}, \quad u \in \mathcal{U}$$

Regulate nonlinear plasma effects on surface

Process constraints:

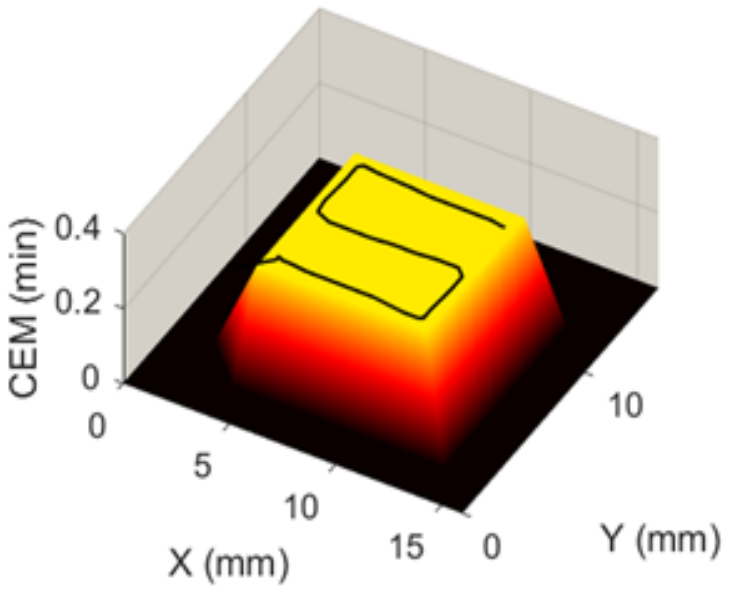
- Device limitations
- Maintain discharge and/or surface properties in a desired operating range
- Avoid mode transitions, etc.

Gidon, Graves, and Mesbah, *PSST*, 2019.

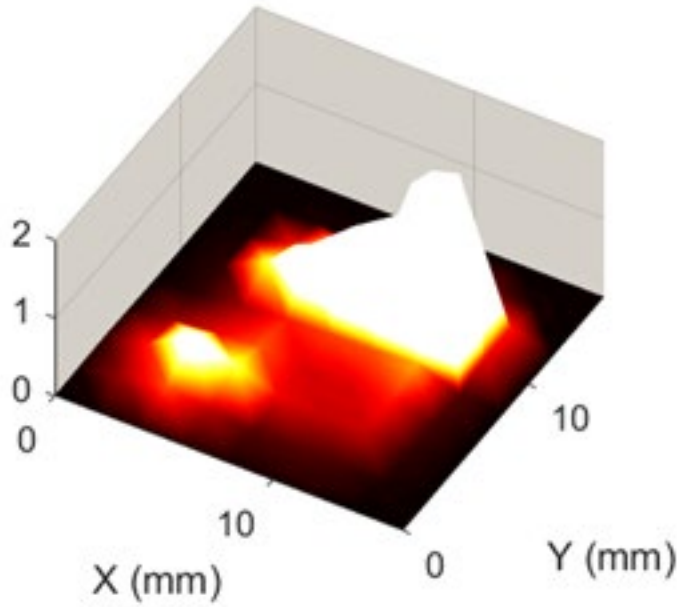
Data-Driven Models for Optimization-based Process Control

Real-time control of nonlinear plasma effects on surface (in the presence of disturbances)

Reference spatial thermal effects (CEM_{43}=0.25 min)



No feedback control (Constant applied voltage)



Real-time control

Nonlinear thermal effects of plasma on surface:

$$\mathrm{CEM}_{43} = \int_0^t K^{(43-T_s(\tau))}\, d\tau$$

Gidon, Graves, and Mesbah, *PSST*, 2019.

Data-Driven Models for Optimization-based Process Control

Real-time control with online model adaptation

- A key challenge in plasma modeling is to account for nonlinear and time-varying nature of dynamics while maintaining computational tractability of a model

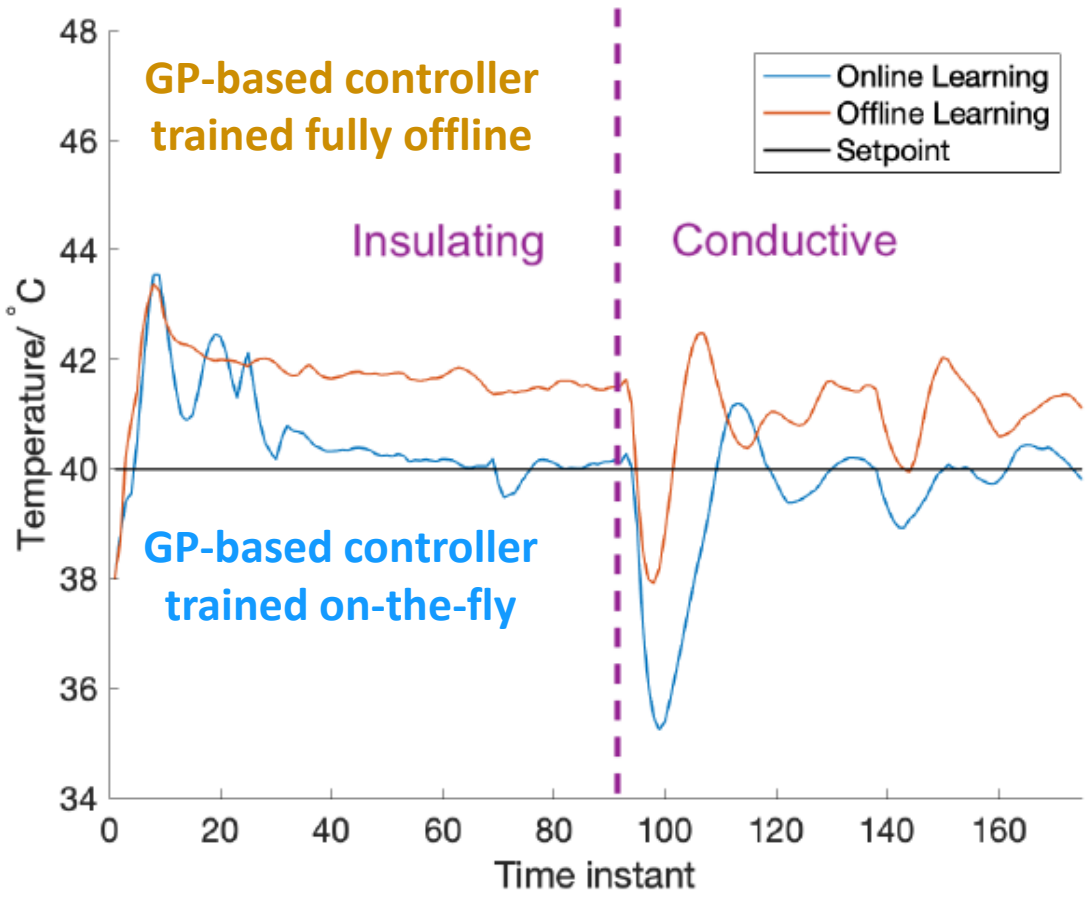
e.g., physics-based model or data-driven model (derived offline)

$$\frac{dx}{dt} = f(x,u) + w(x)$$

state-dependent "model uncertainty" term learned online

Nonlinear Gaussian process regression

$$w(x) = \begin{pmatrix} Y \\ \hat{Y} \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \Sigma\right)$$



Bonzanini, Graves, and Mesbah, *IEEE TCST*, Submitted.

Reinforcement Learning (RL)

Sim-to-real transfer learning

RL for learning-based control of plasma processes involving treatment of complex surfaces with time-varying or non-uniform characteristics:

- In the absence of first principles, RL agents can be trained directly via experimental data in a model-free fashion
- Using simulation data for training RL can address challenges with safe training of RL agents as well as excessive sample complexity
- There is a reality gap between simulation data and the real-world environment

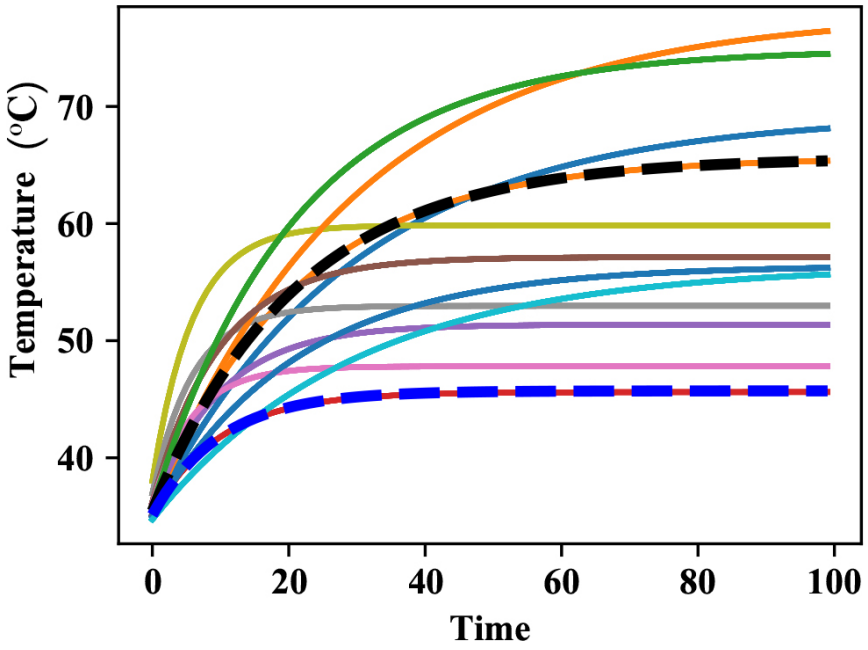
How to enrich simulation data for effective sim-to-real transfer learning?

Deep Reinforcement Learning

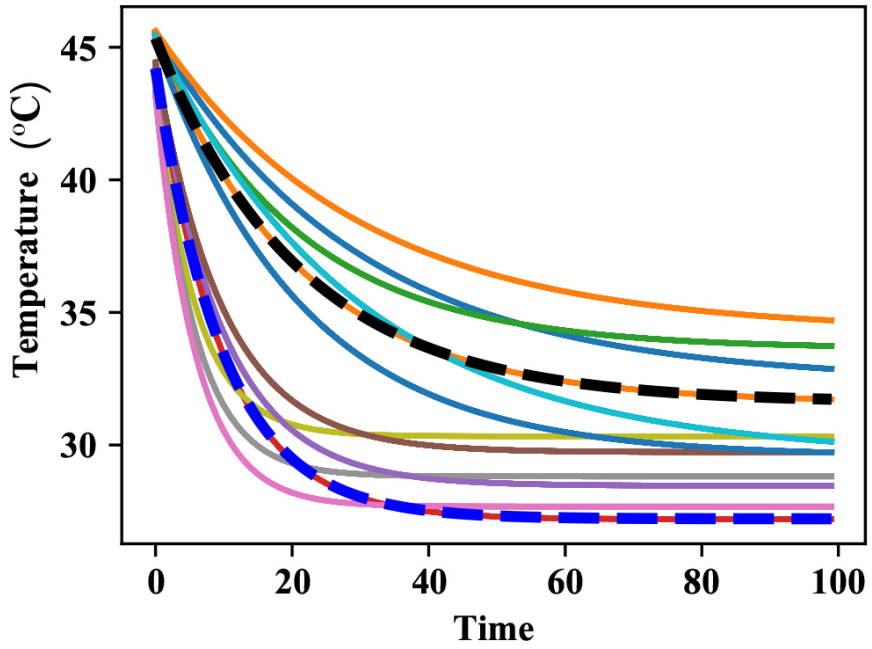
Sim-to-real transfer learning using dynamics randomization

RL agents trained using a physics-based model to control temperature of substrates with different thermal and electrical properties (glass, aluminum, and polyimide)

Heating dynamics



Cooling dynamics



Aluminum

Glass

Model parameters randomized to enrich simulation training data by accounting for different substrate properties

Peng et al., *CoRR*, 2017.

Sim-to-Real Transfer of Deep Reinforcement Learning

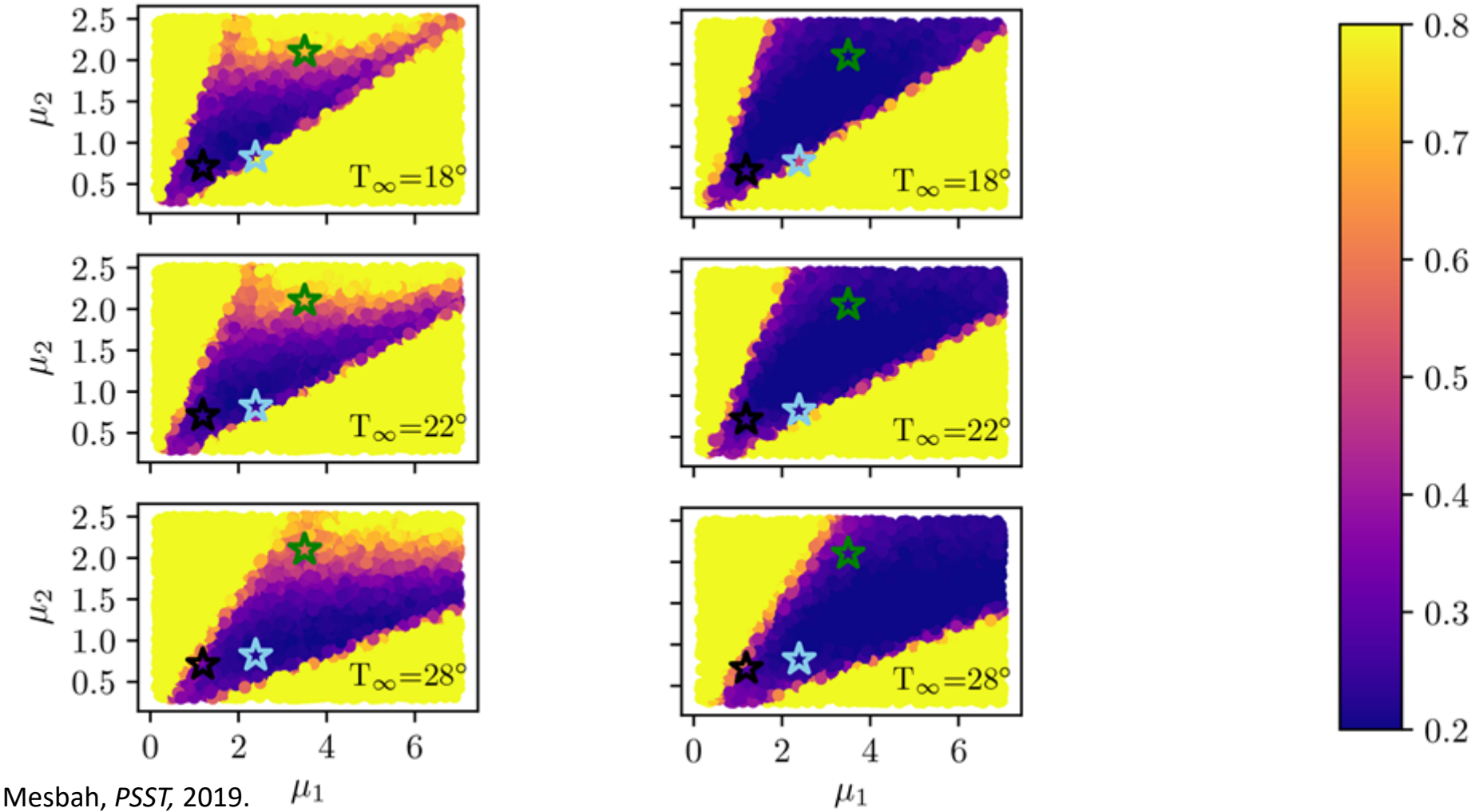
In silico performance evaluation

- T_∞: Ambient temperature
- μ_1, μ_2: Model parameters
- Black star: Aluminum
- Blue star: Glass
- Green star: Polyimide

RL agent trained using data generated for glass substrate (G-RLC)

RL agent trained using randomized data for all substrates (E-RLC)

$$\mathrm{MAE} = \frac{1}{T}\sum_{t=0}^{T} |T_{s,t} - T_s^{sp}|$$

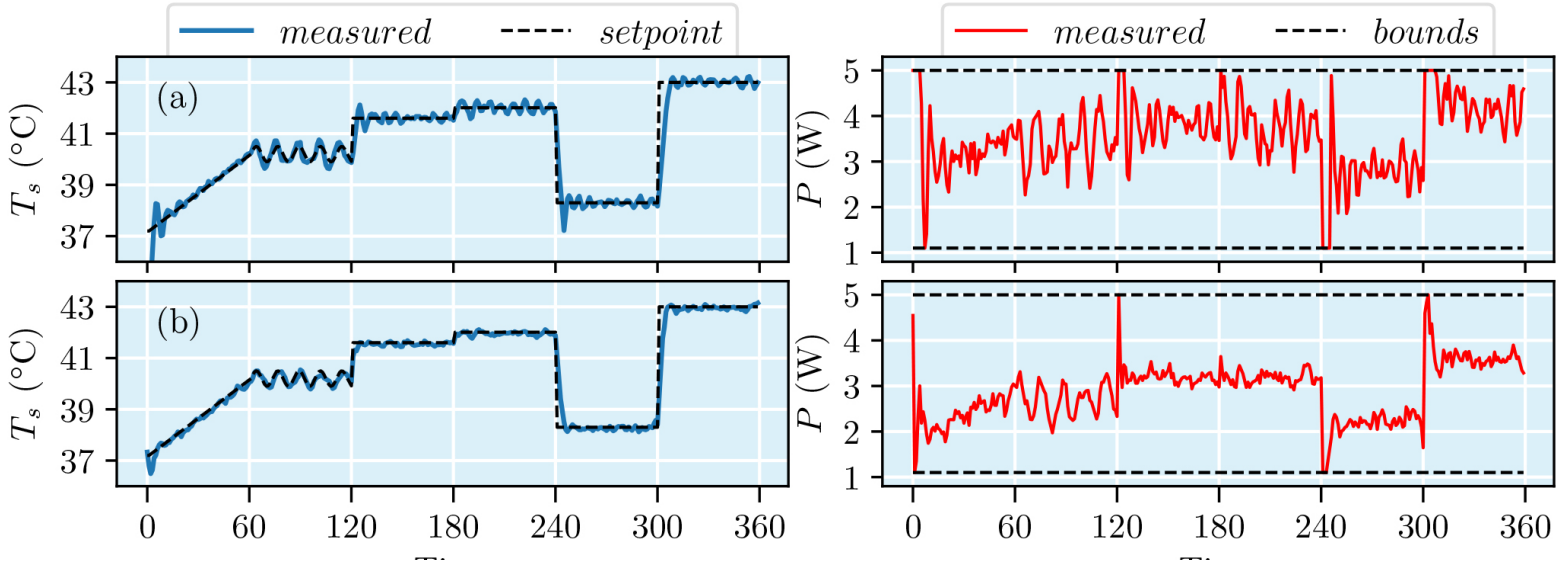


Witman, Gidon, Graves, Smit, and Mesbah, *PSST,* 2019.

Sim-to-Real Transfer of Deep Reinforcement Learning

Experiments

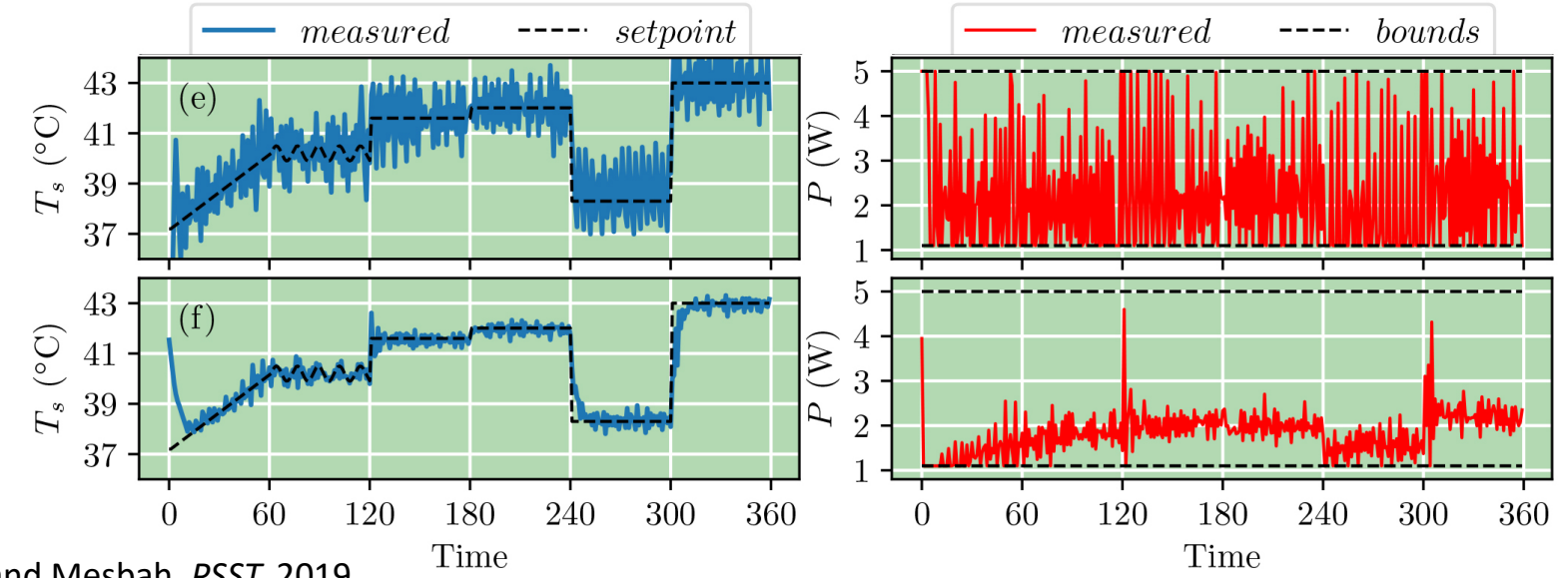
Glass Substrate



G-RLC

E-RLC

Polyimide Substrate



G-RLC

E-RLC

Witman, Gidon, Graves, Smit, and Mesbah, *PSST,* 2019.

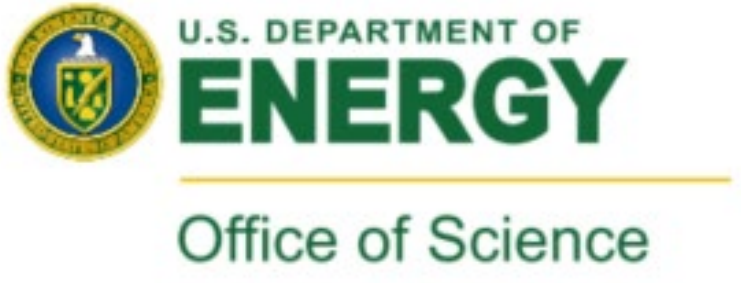
Vision and Future Work

Towards learning-based autonomous and embedded control systems technologies for non-equilibrium plasma processing of complex surfaces:

- Autonomous learning of complex dynamics of plasma-surface interactions
- Embedded process diagnostics and control

Closing thoughts:

- Machine learning may enable development of modeling and simulation frameworks that are self-aware and self-correcting
- There may be need to develop specialized machine learning formalizations for non-equilibrium plasma processes
- Bayesian machine learning can become the leading paradigm

Acknowledgments

Prof. David B. Graves
Dr. Dogan Gidon
Angelo Bonzanini
Ketong Shao
Dr. Joel Paulson
Dr. X. Pei