CMP User Group Winter 2024

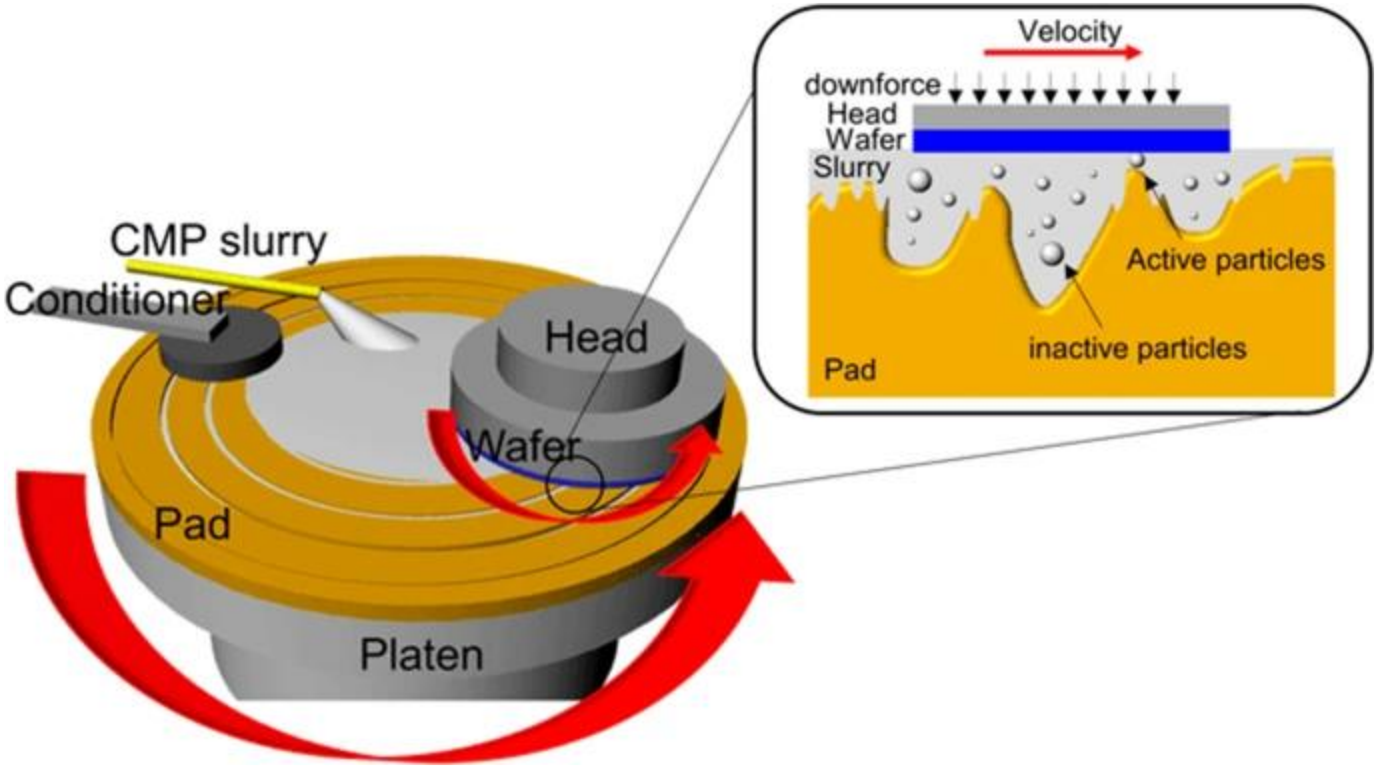
Spectroscopic Analysis and Machine Learning to Expedite Oxide Planarization Development (SAMPLE-OPD)

Shanmukh Kutagulla, Richard F Vreeland, Glenn Whitener, Ken Delbridge, Thomas Hoff, Eric Jorgenson, Jeff Armstrong, William Brezinski, Herbert Barnett, Eric Mattson, Shouvik Banerjee, Rupak Thapa, Christopher Pell, Cindy Taylor, Nam Nguyen, Camille McGill, Abraham Trejo

Intel Technology Research

Agenda

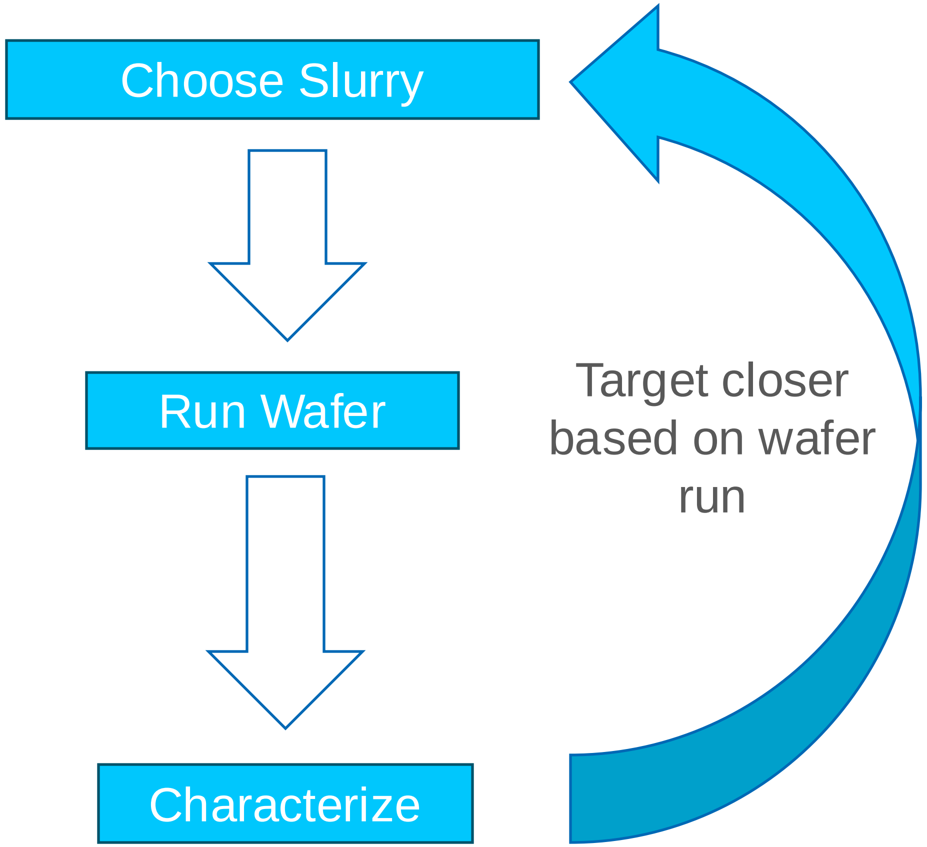
The CMP Process & Current Challenges



Many parameters that all interact, making 1:1 modeling difficult. Development is primarily empirical.

Seo, J. A review on chemical and mechanical phenomena at the wafer interface during chemical mechanical planarization. *Journal of Materials Research* **36**, 235–257 (2021). https://doi.org/10.1557/s43578-020-00060-x

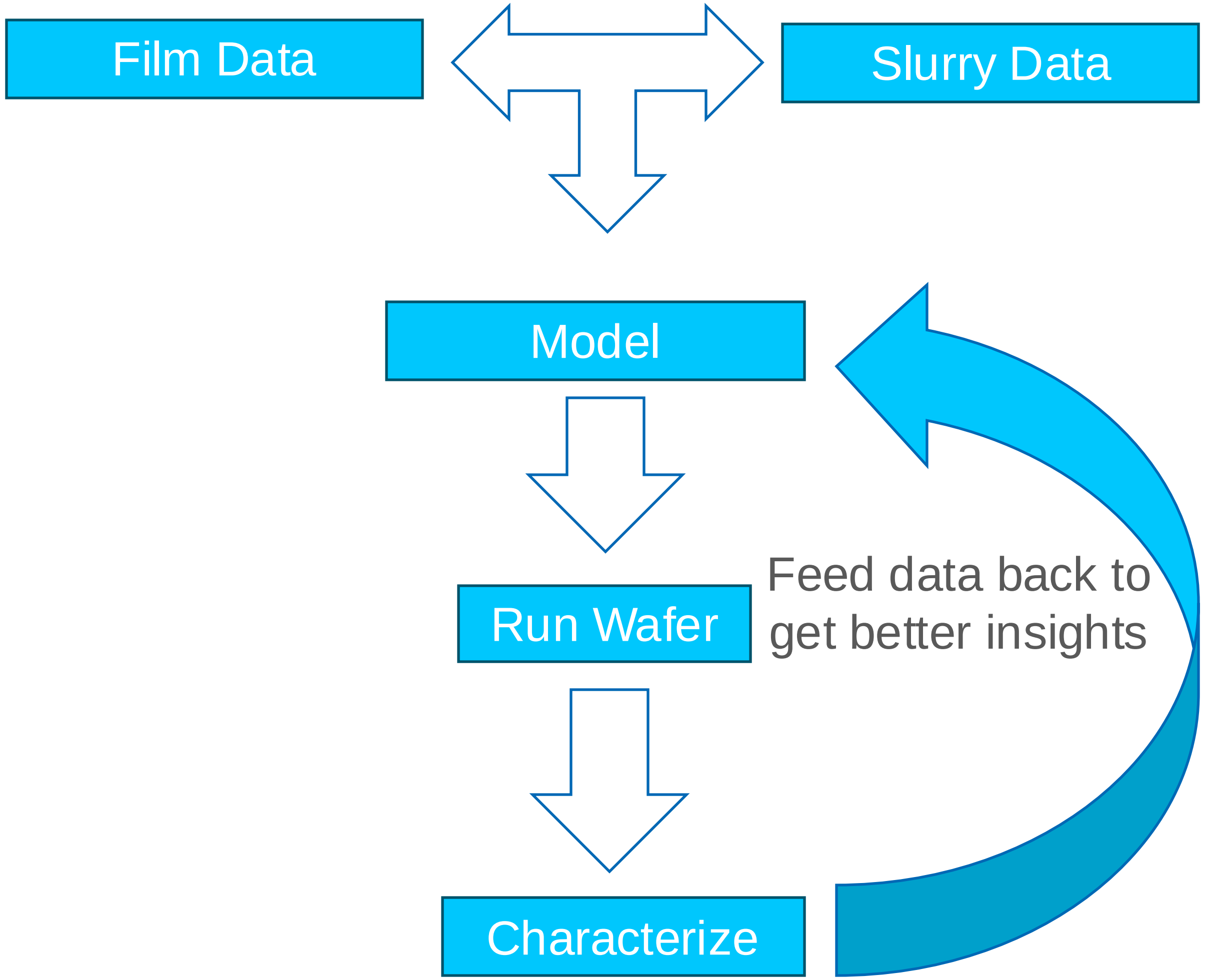
Current CMP Development- High Throughput Experimentation



Make process changes based on previous/institutional knowledge, empirical data, gut instinct

*sometimes, you get it right the first time using onboard metrology

The Vision for CMP Development



Make changes based on model-based insight-agnostic of previous processing

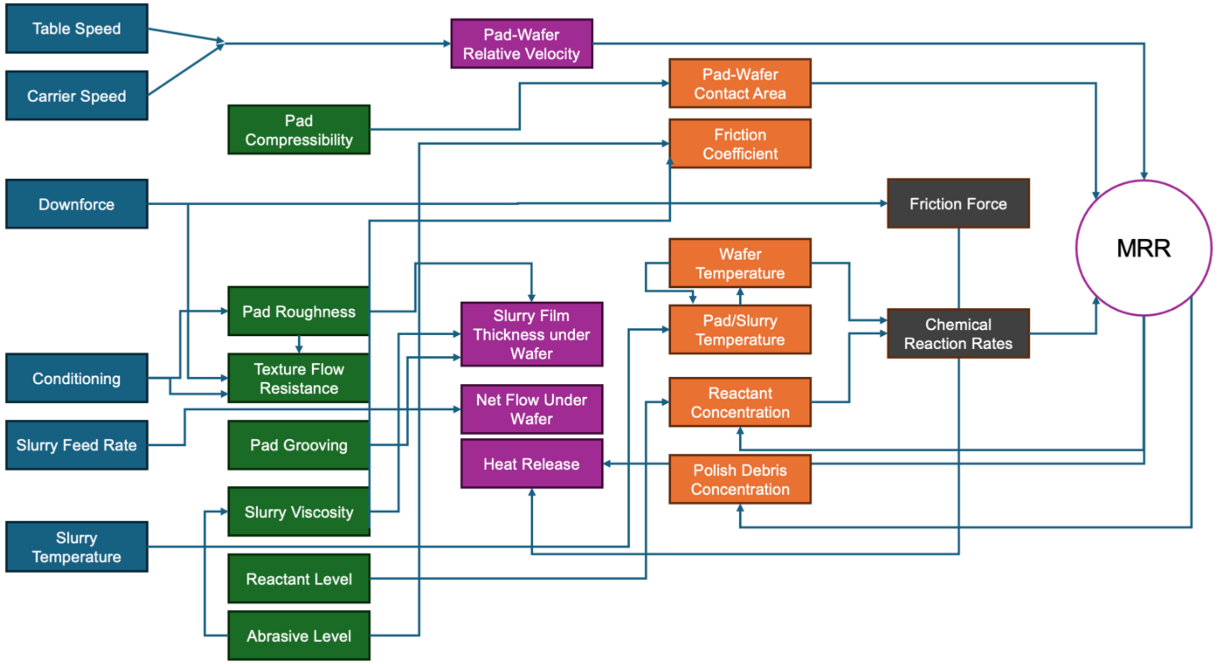
Do not need to start from zero when working on new films

Can we develop such a model with AI?

- It has never been easier to study interrelated variables using AI techniques

- Previous attempts at modeling utilized 20 or more descriptors and thousands of data points to predict polish rates- infeasible for R&D

- Select features of importance based on surface interactions and basic spectroscopy data



Can we develop a model in the low data regime?

Can We Use Noncontact Metrology to Predict Polish Rates?

Custom Blends from Fujimi

Slurry	Particle Size	Material
X	30	SiO_2
Y	30	"Low Silanol" SiO_2
2X	75	SiO_2

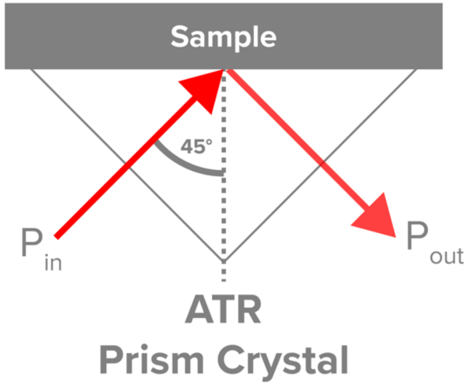
16 unique films

Films	Deposition Information
fCVD 1	fCVD- Thermal Cure
fCVD 2	fCVD- UV Cure
PECVD	PECVD (SiH_4+N_2O)
HDP	High Density Plasma (SiH_4+O_2)

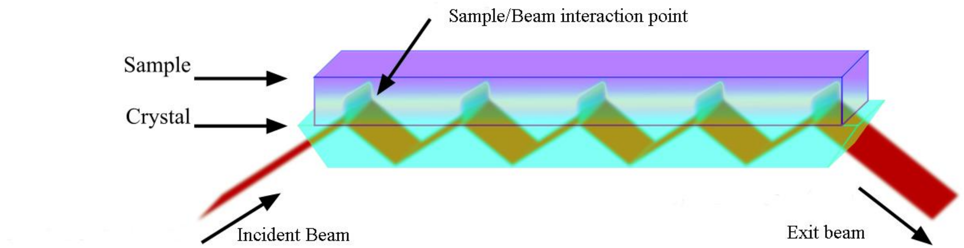
Steam Treatment Conditions
No Process - NP
200 C (1 hr) + 500 C (1 hr)- 200/500
500 C (2 hr) - 500
700 C (2 hr) - 700

3 slurries x 4 films x 4 steam treatments= **48 polish rates**

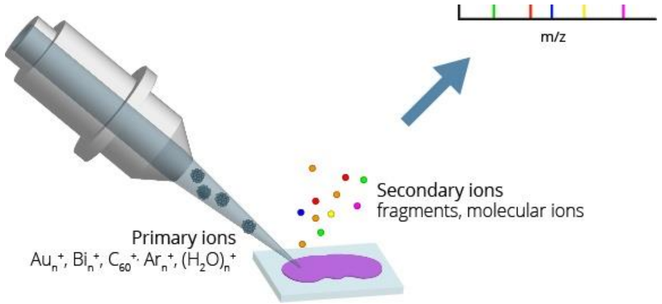
Our Toolkit



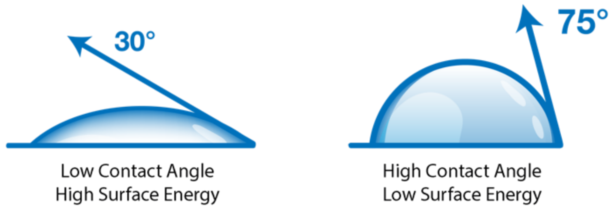
ATR-FTIR



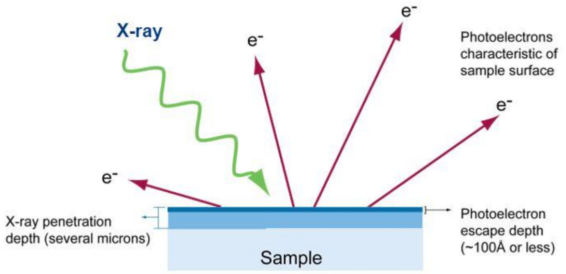
MIR-FTIR



ToF-SIMS



Water Contact Angles (WCA)

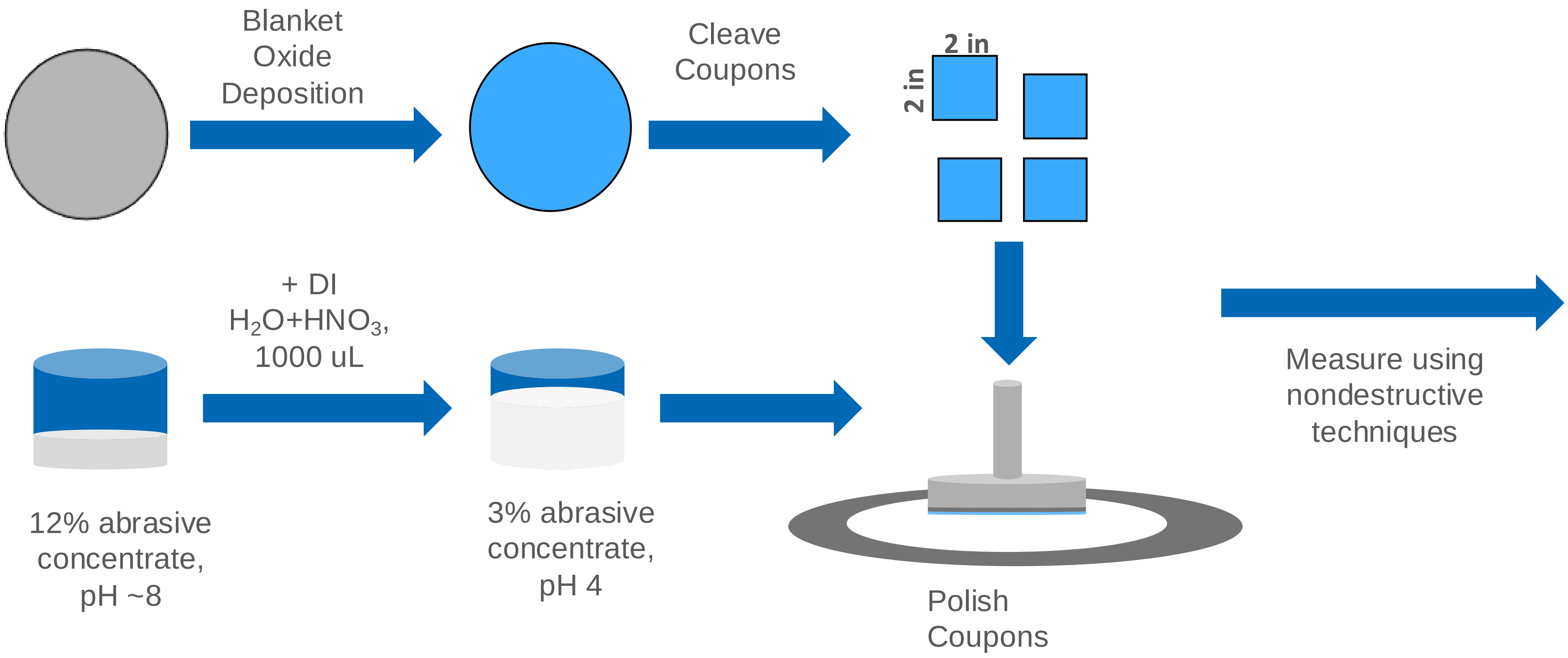


X-ray Photoelectron Spectroscopy (XPS)

Density Measurements

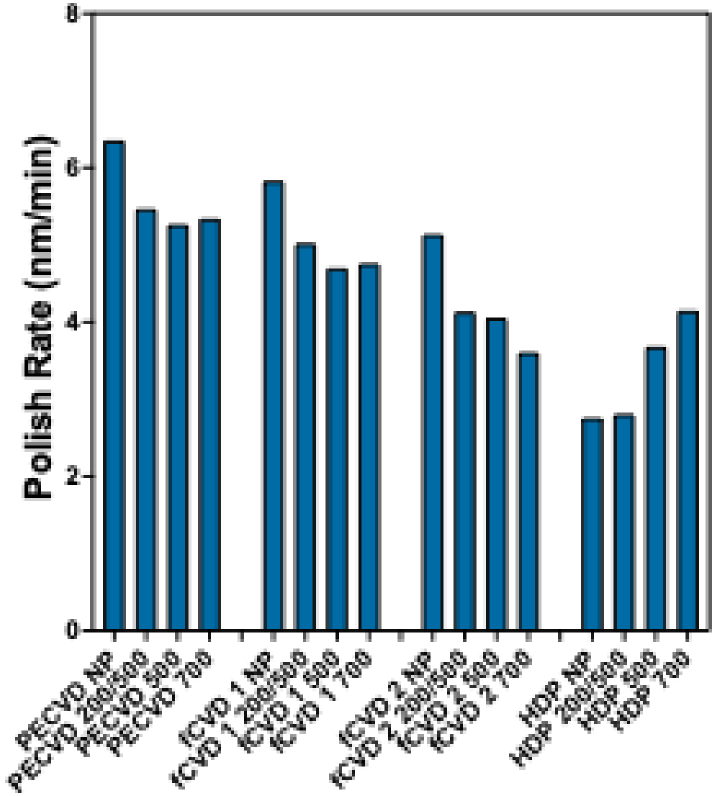
Characterization battery was completed on all wafers post growth, and on selected "interesting" coupons

Experimental Design

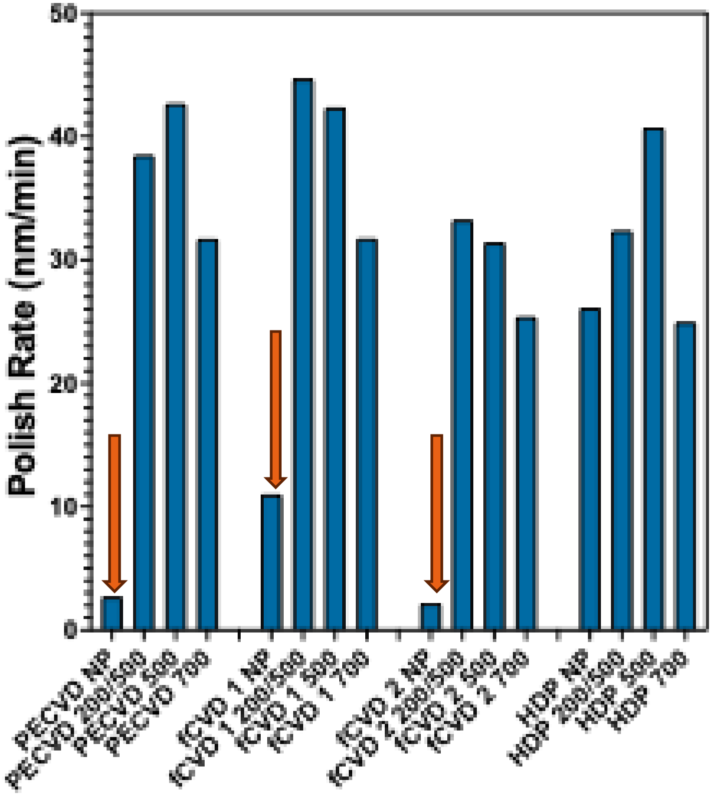


Based on the collected data, we can identify some immediate trends on factors that affect polish rates

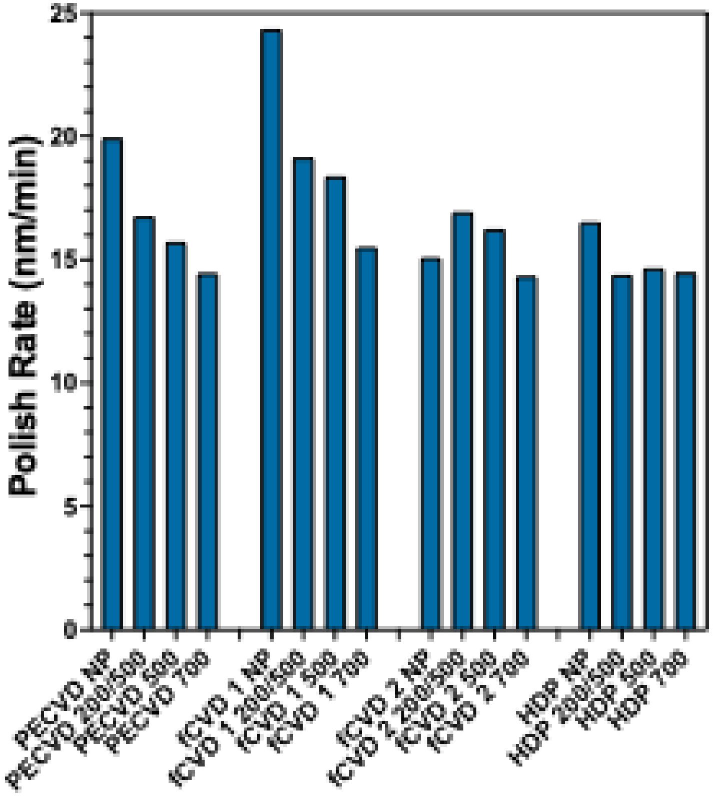
Steam Annealing Decreases Polish Rates – With Notable Exceptions



Particle X



Particle Y (Surface Engineered)

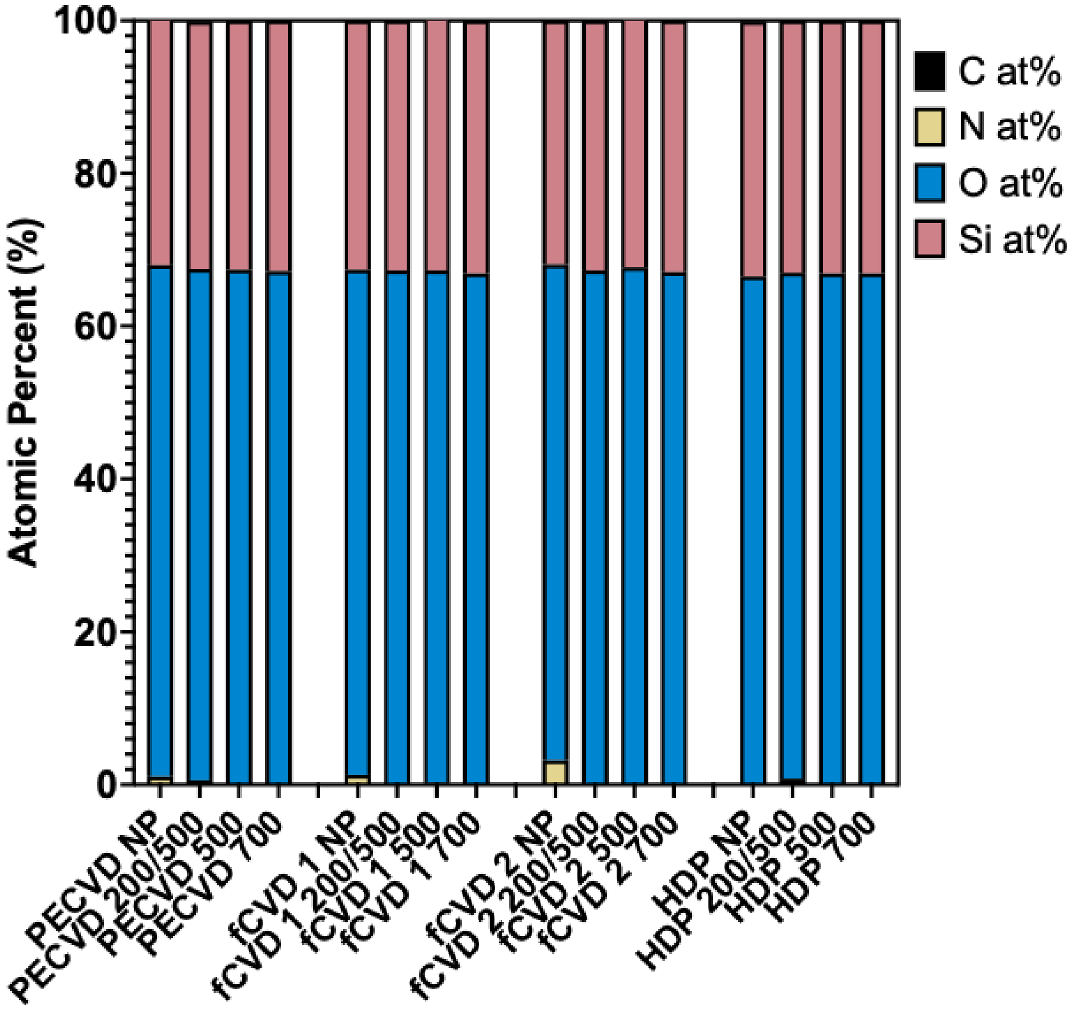


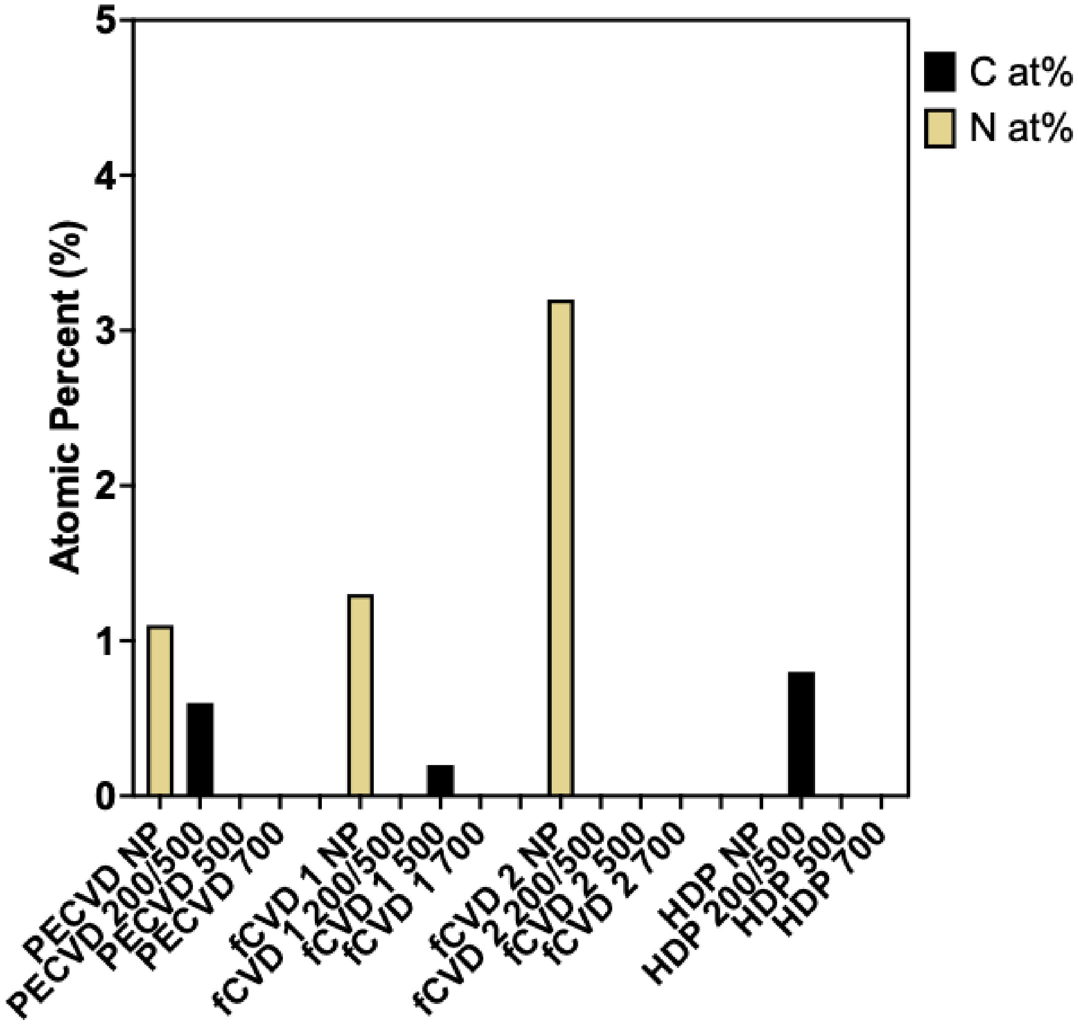
Particle 2X

No Process films for fCVD 1, fCVD 2, & PECVD show reduced polish rates with Particle Y while the rest show enhanced rates

The Role of Different Chemical Species in Films

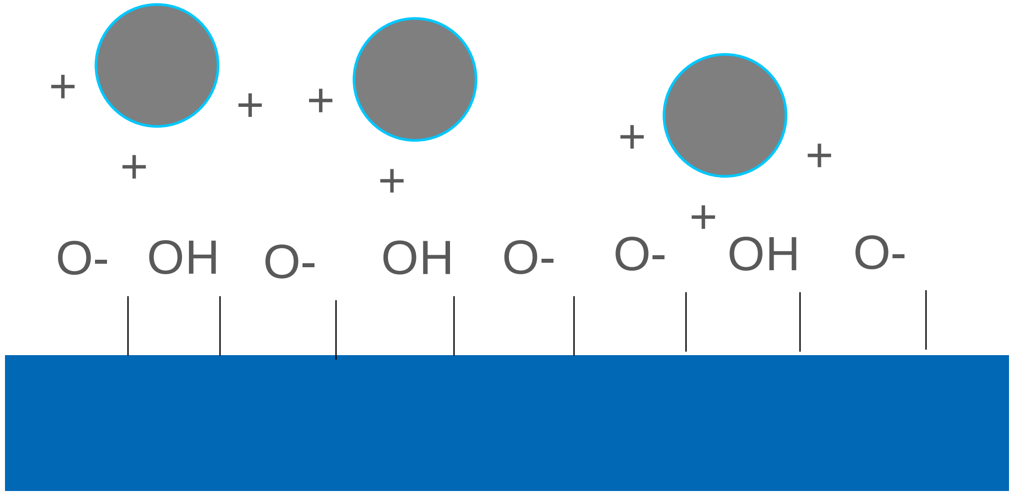
All 3 Particle Y Nonpolishing Films Show >1 at% N



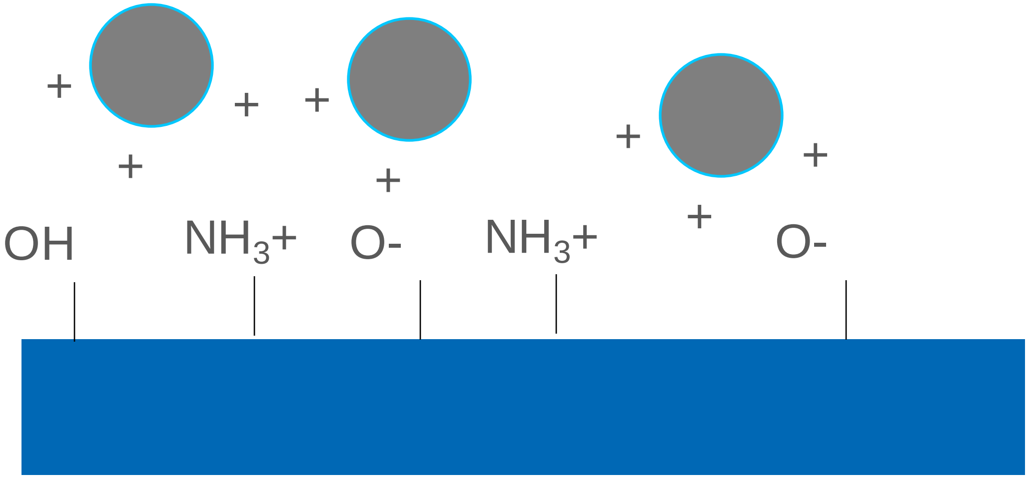


A small amount of nitrogen causes large scale effects!

Why Does 1-3% Nitrogen Reduce Polish Rate With Particle Y?

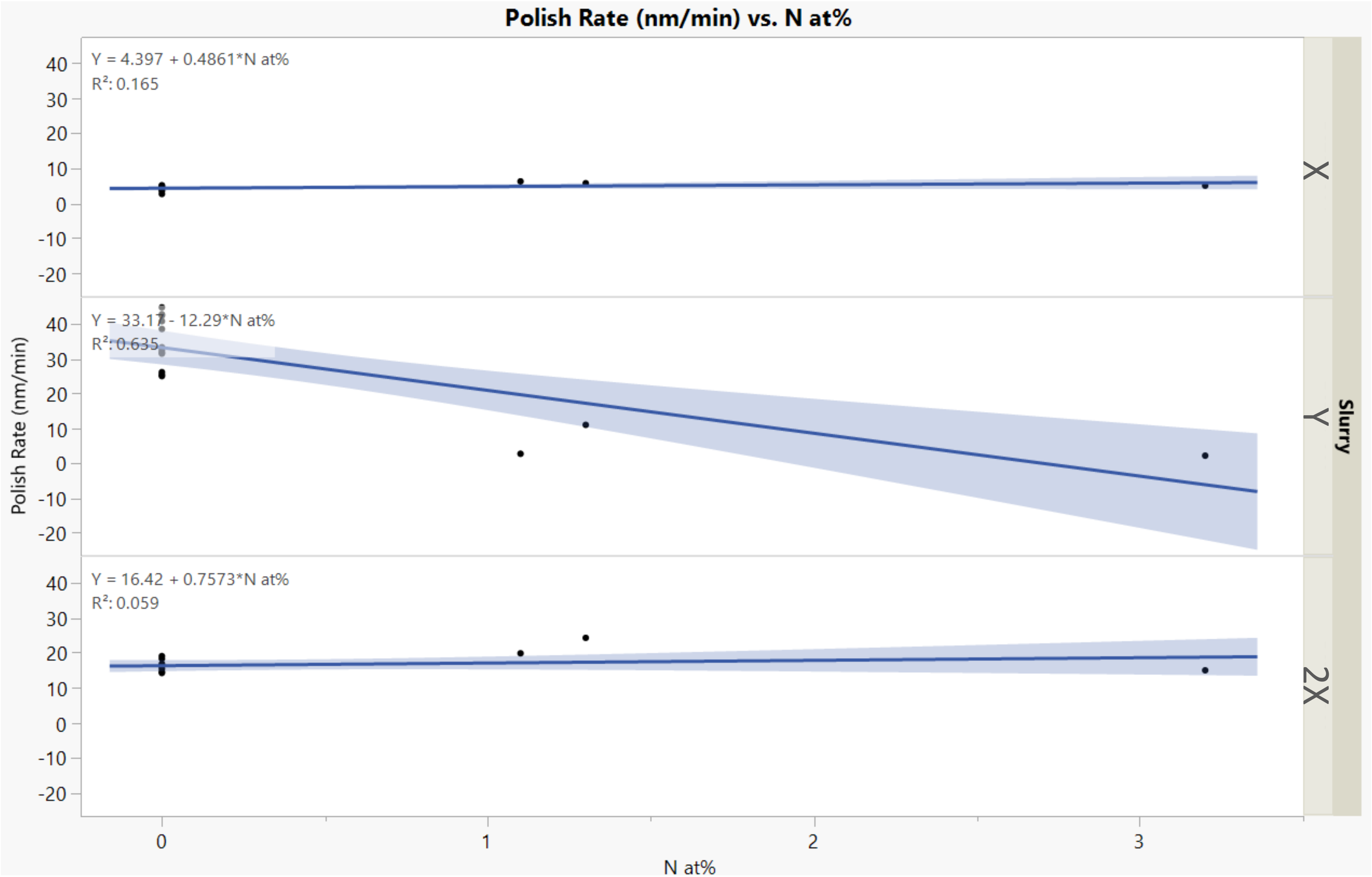


Attractive force between particle and surface, MRR increase

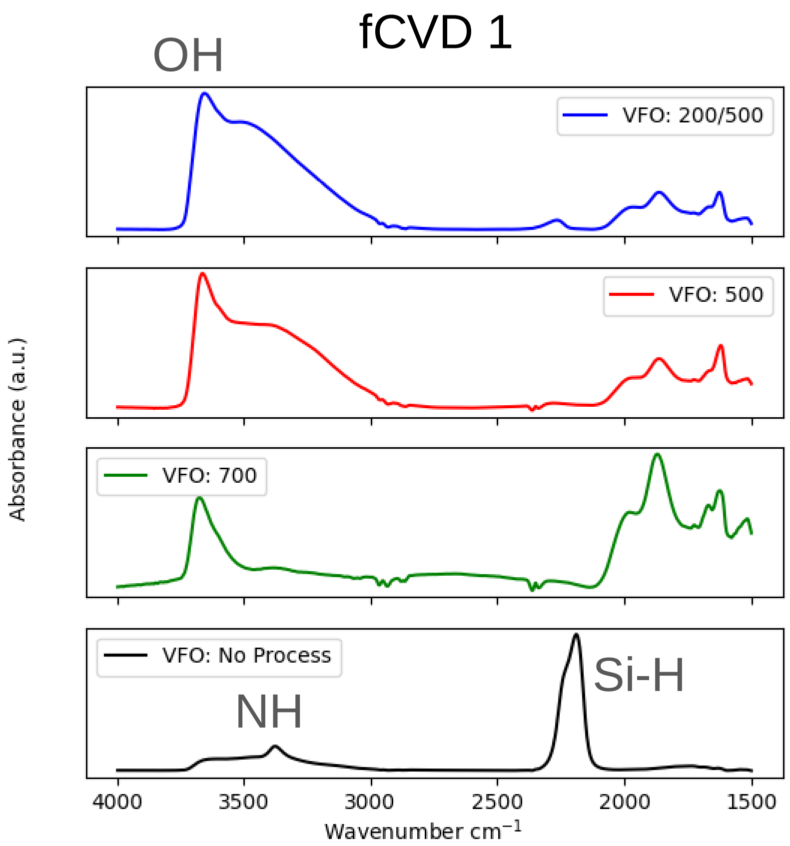


More generally abrasive action due to minimal electrostatic interactions on surface- repulsion and charge neutralization from NH_3+

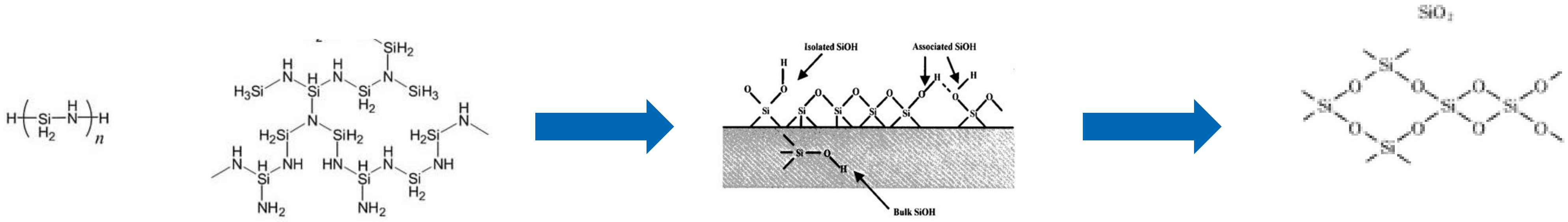
Nitrogen Content is Not Correlative to Polish Rates



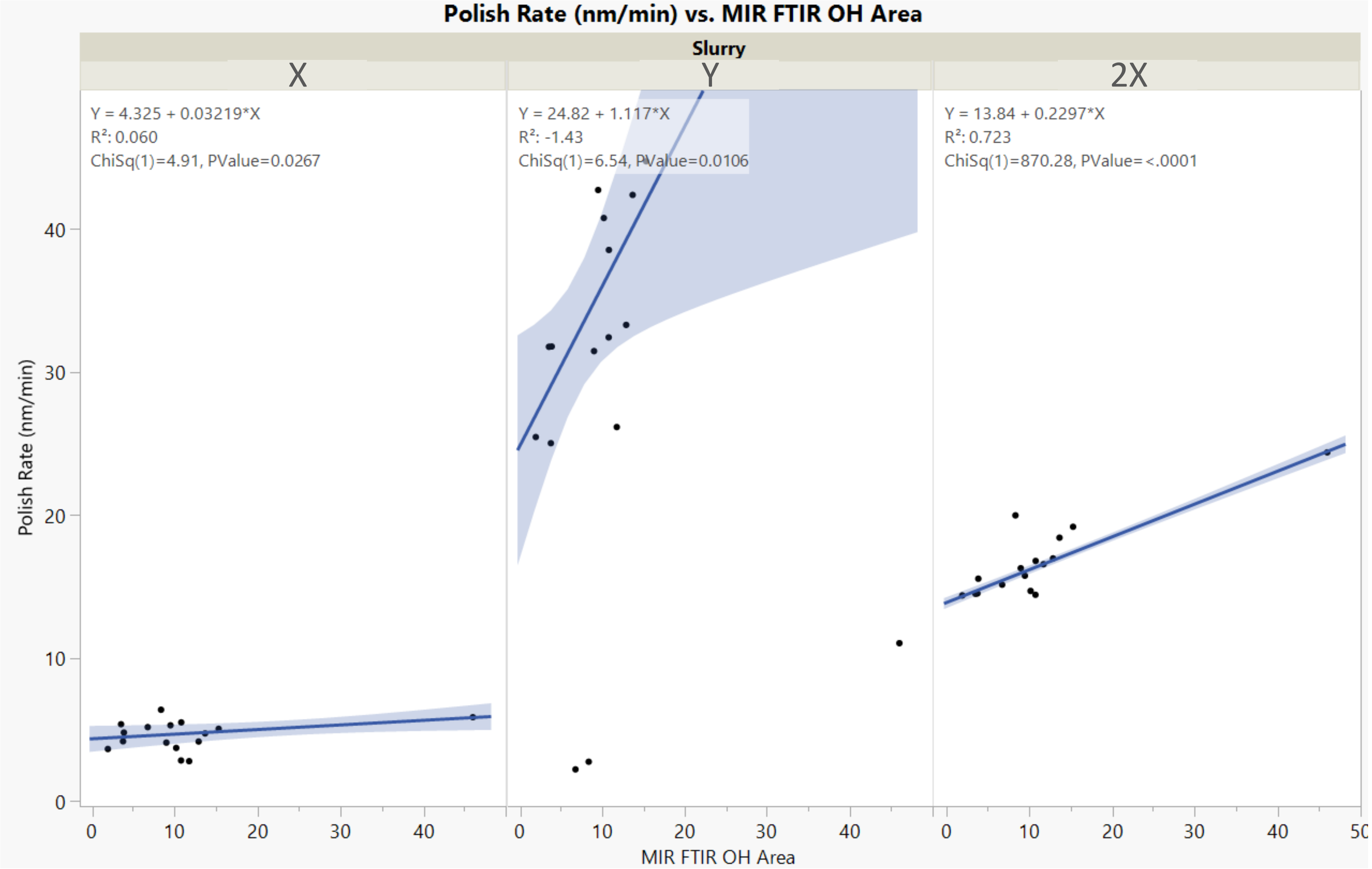
FTIR Provides A More Direct Understanding of Anneal Effects



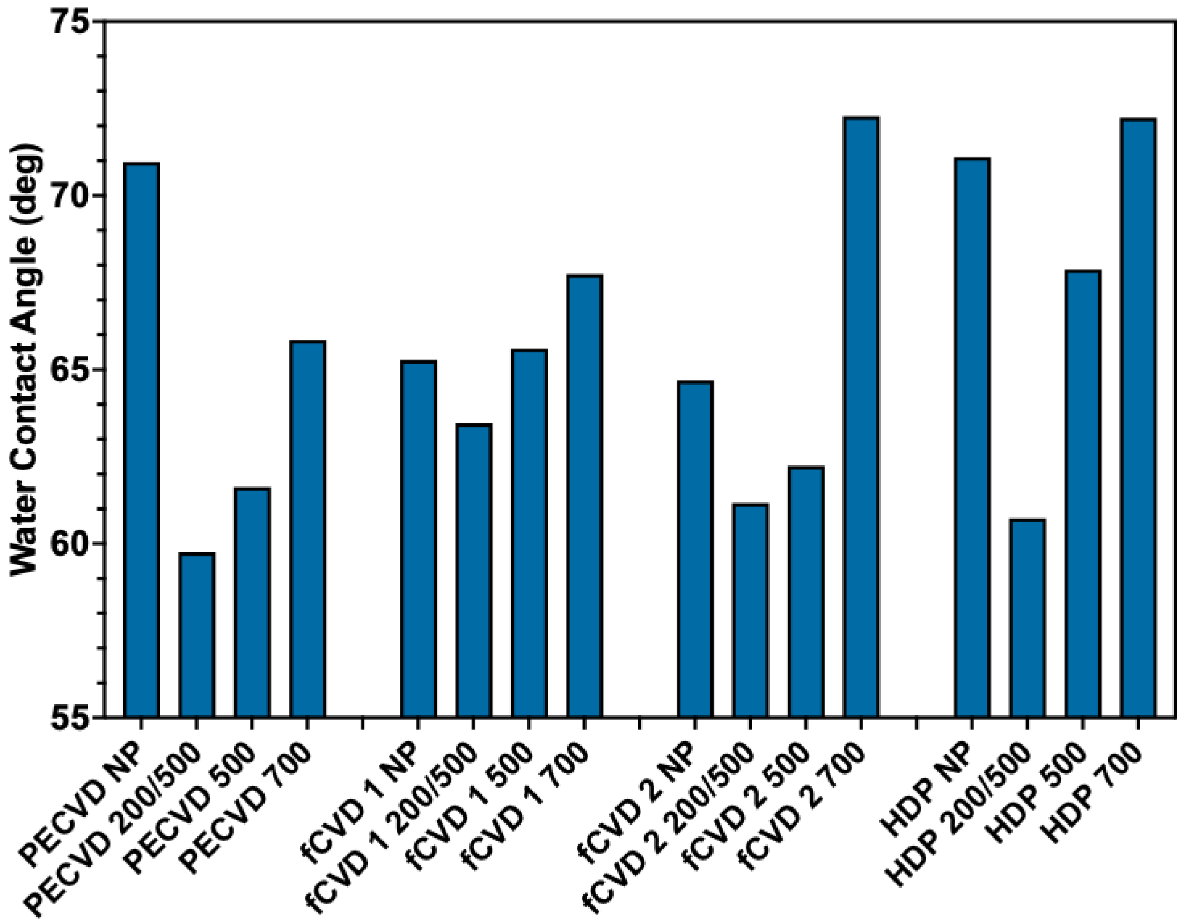
- In both fCVD films, the N and Si-H peak disappears after a steam anneal
- As the steam anneal gets more aggressive, the broad OH peak narrows and decreases



FTIR –OH Area Is Not Independently Predictive of Polish Rates

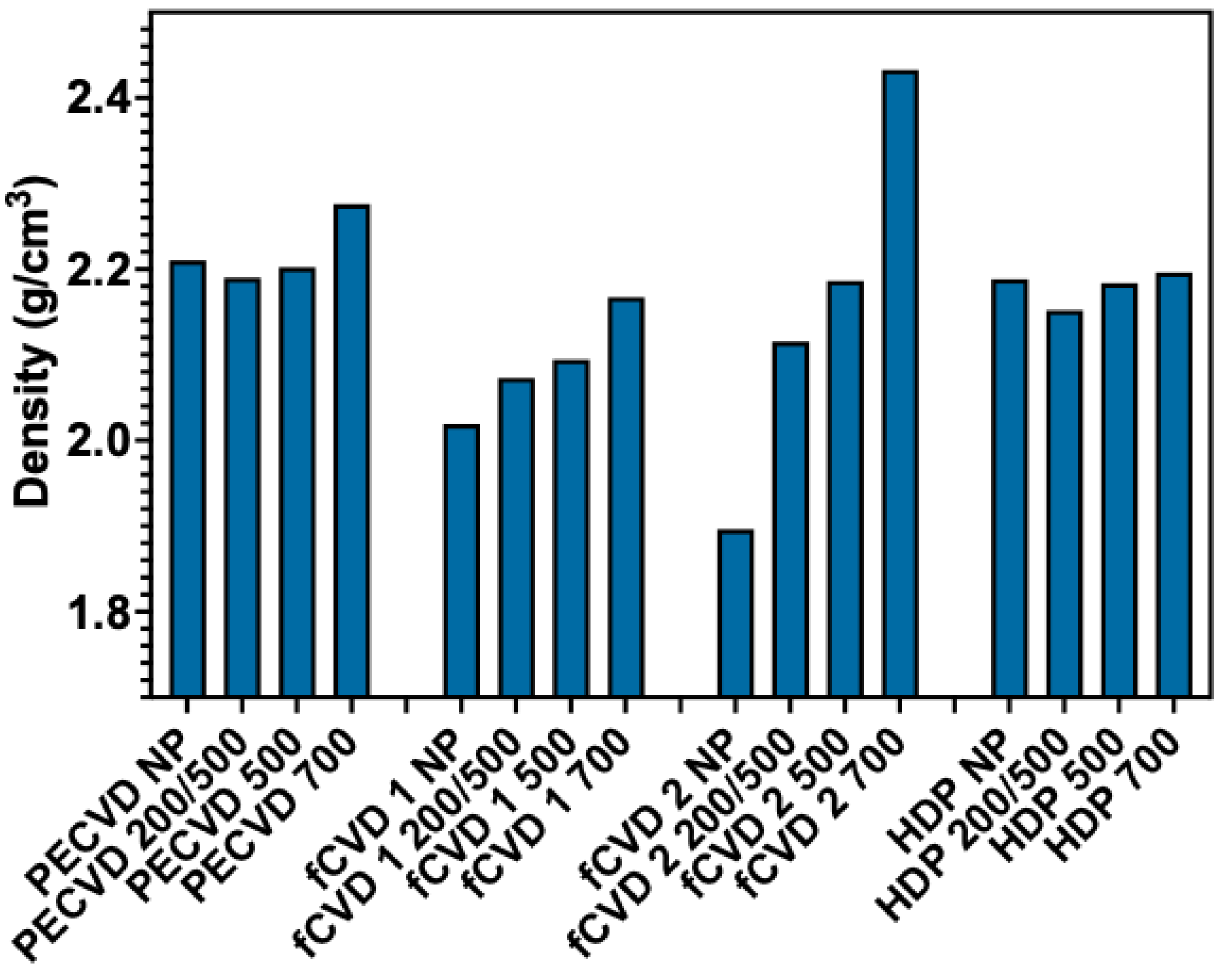


Why does a higher anneal condition lead to a lower polish rate?



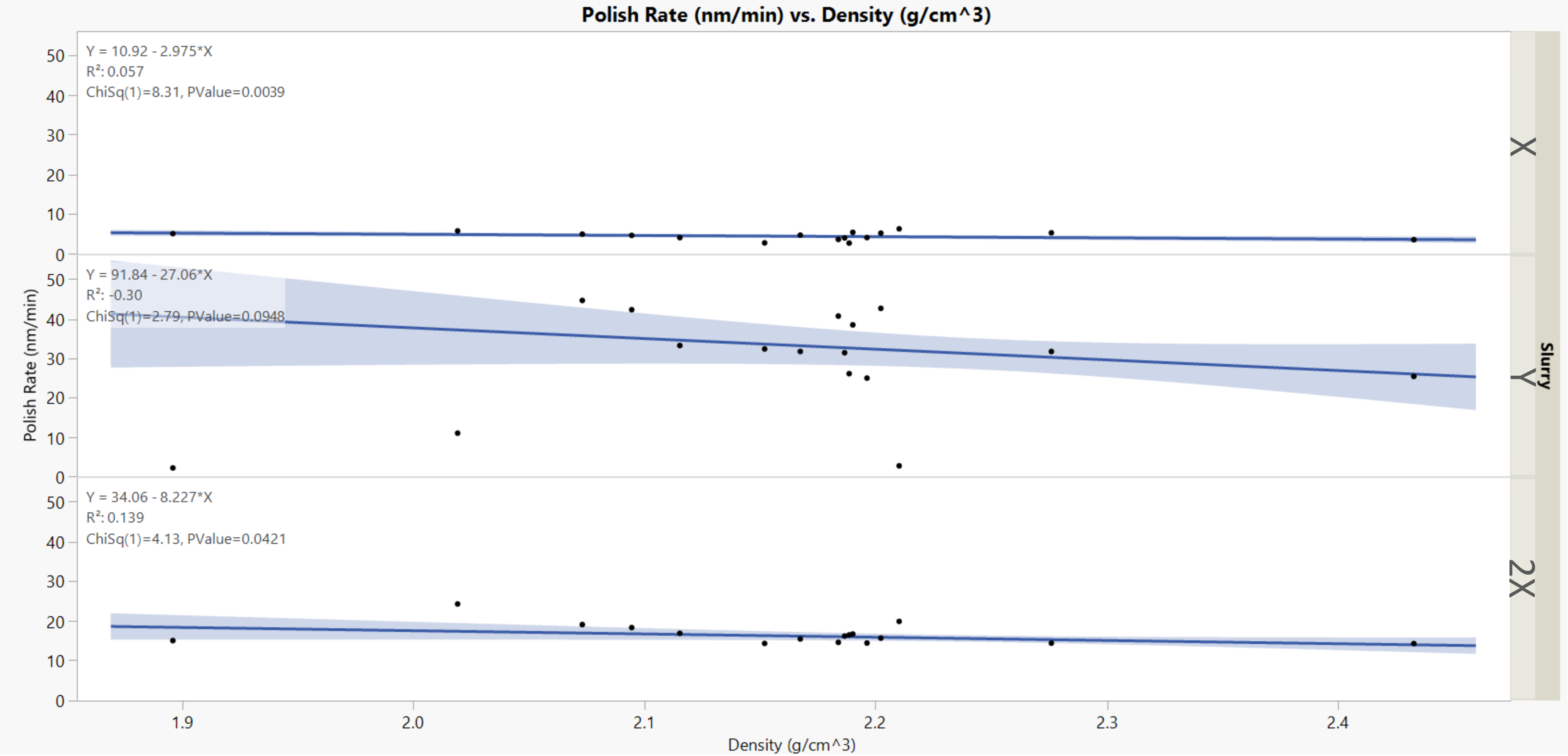
Contact angles show a clear trend- as the anneal gets more aggressive, hydrophobicity increases

Density Data Further Reinforces WCA Findings



As the films are annealed, they also become denser- a result of hydrolysis into Si-O-Si

Neither is Density Alone



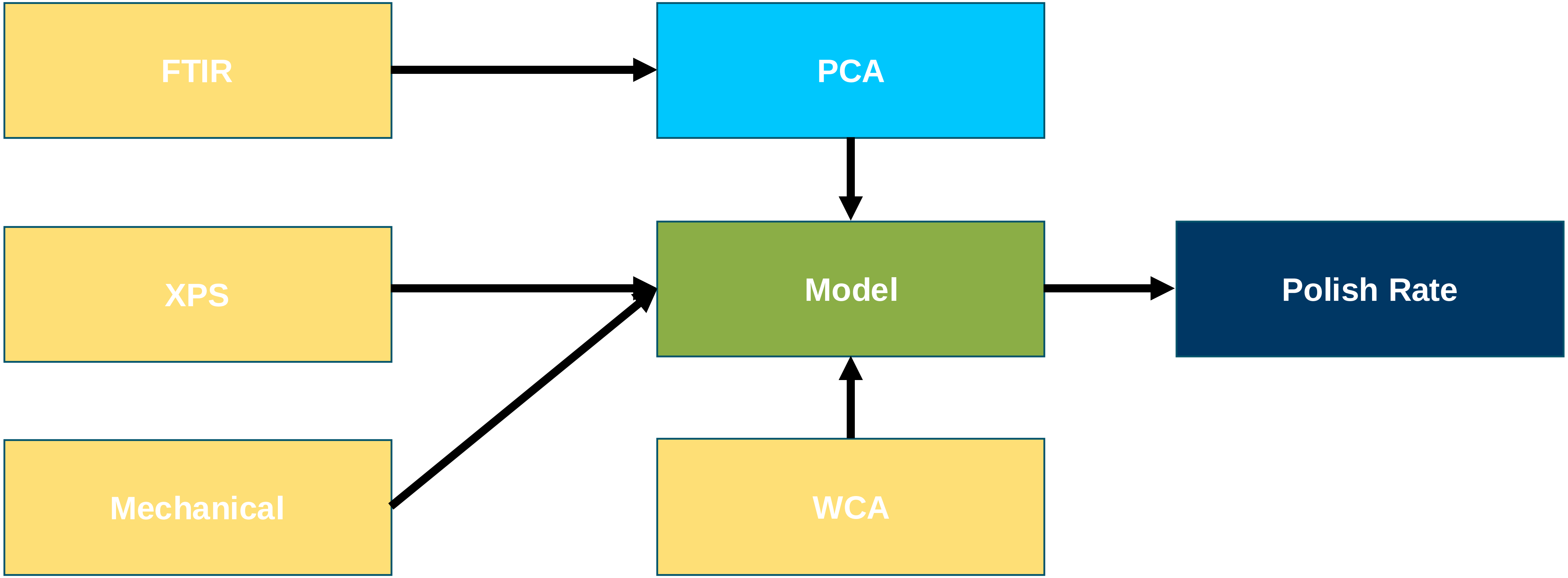
Surprisingly, not that strongly correlated, but F-test suggests strong significance of the density in Slurry X (P<0.05) and Slurry 2X (P<0.05)

Takeaways

- Particle Y enhances the polish rate for pure oxides
- Particle Y shows a negligible polish rate if films contain nitrogen (>1 at%)
- Nitrogen, oxygen and density are all related to polish rate (P<0.05) but are not 1:1 correlative

ML Modeling

If Surface Interactions Are Key, Can We Use Surface Spectra As Predictors of Polish Rate?



PCA Crash Course

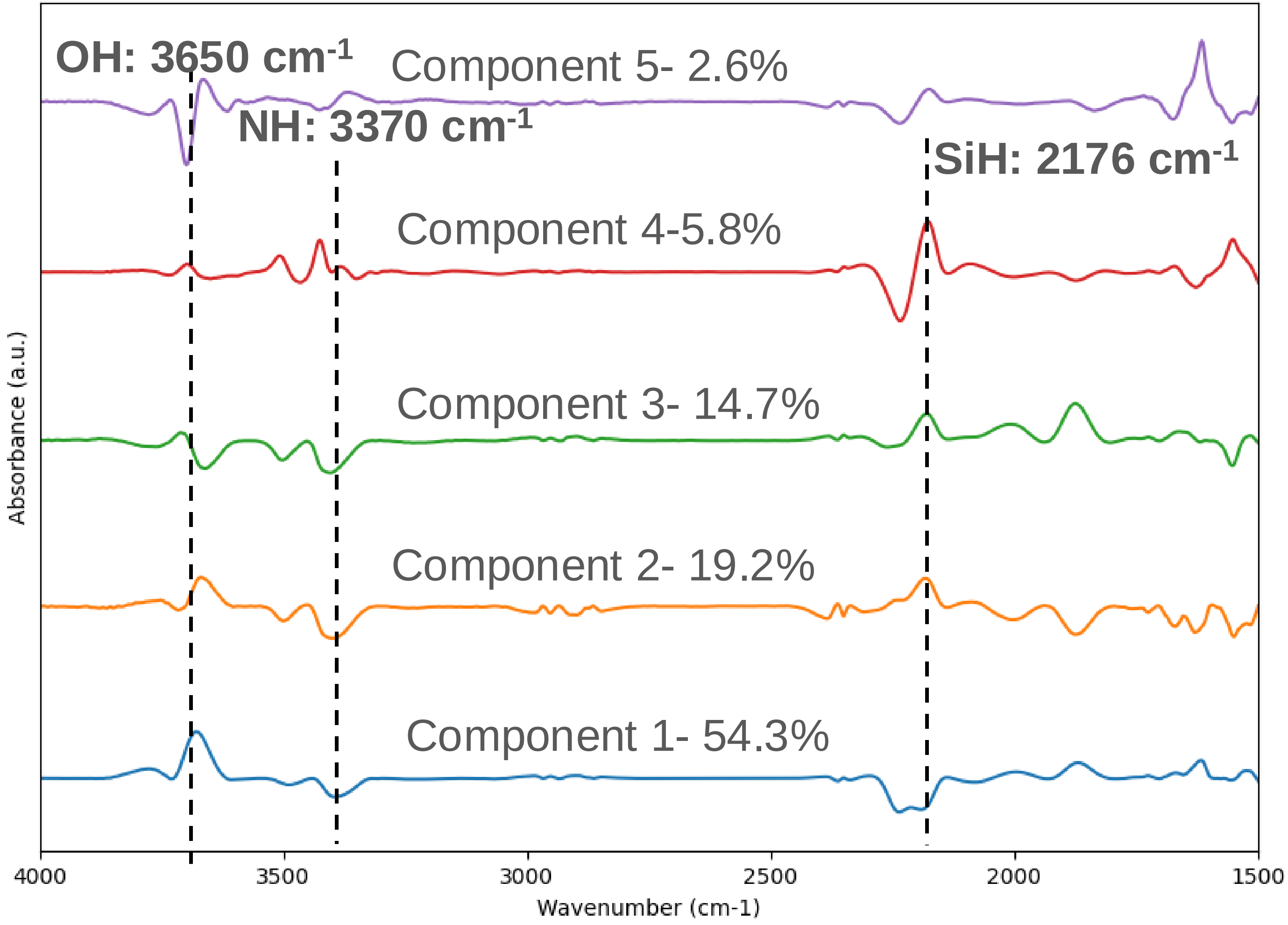
- In non math terms, a PCA ("Principal Component Analysis") finds the features in a dataset that maximize variance.

- "What features in this data make one different than another?"
  - Component 1- maximizes variance, Component 2- 2nd most variance etc.

- Putting together the components in some unique weight will rebuild all the data (almost). Each weight is known as a score.

- This results in a lower-dimensional representation of data that still retains the most important patterns or structures.

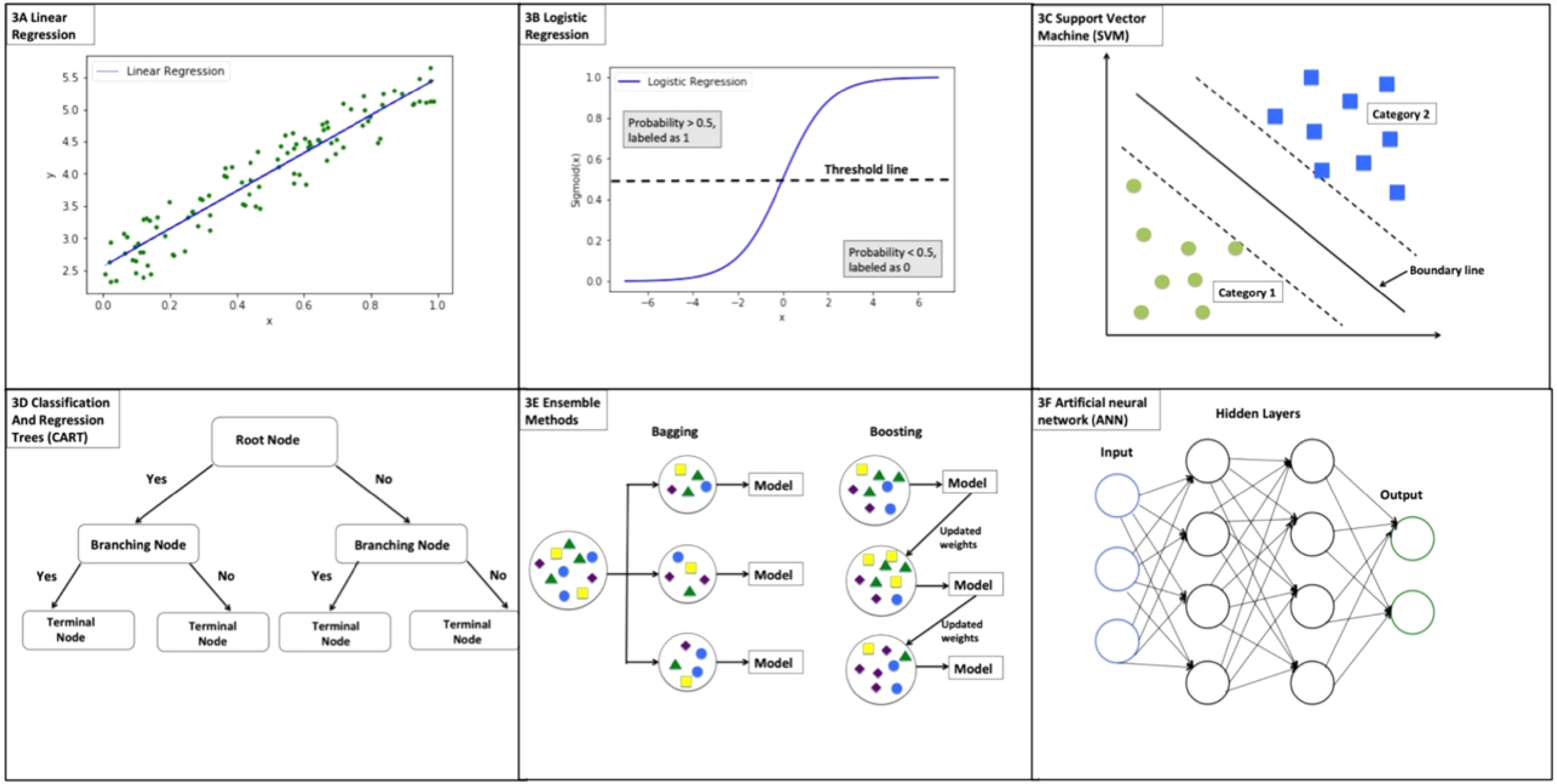
Modeling- FTIR MIR PCA



PCA analysis found impactful wavelengths, correctly identifying them as OH, NH and SiH peaks **without deconvolution**

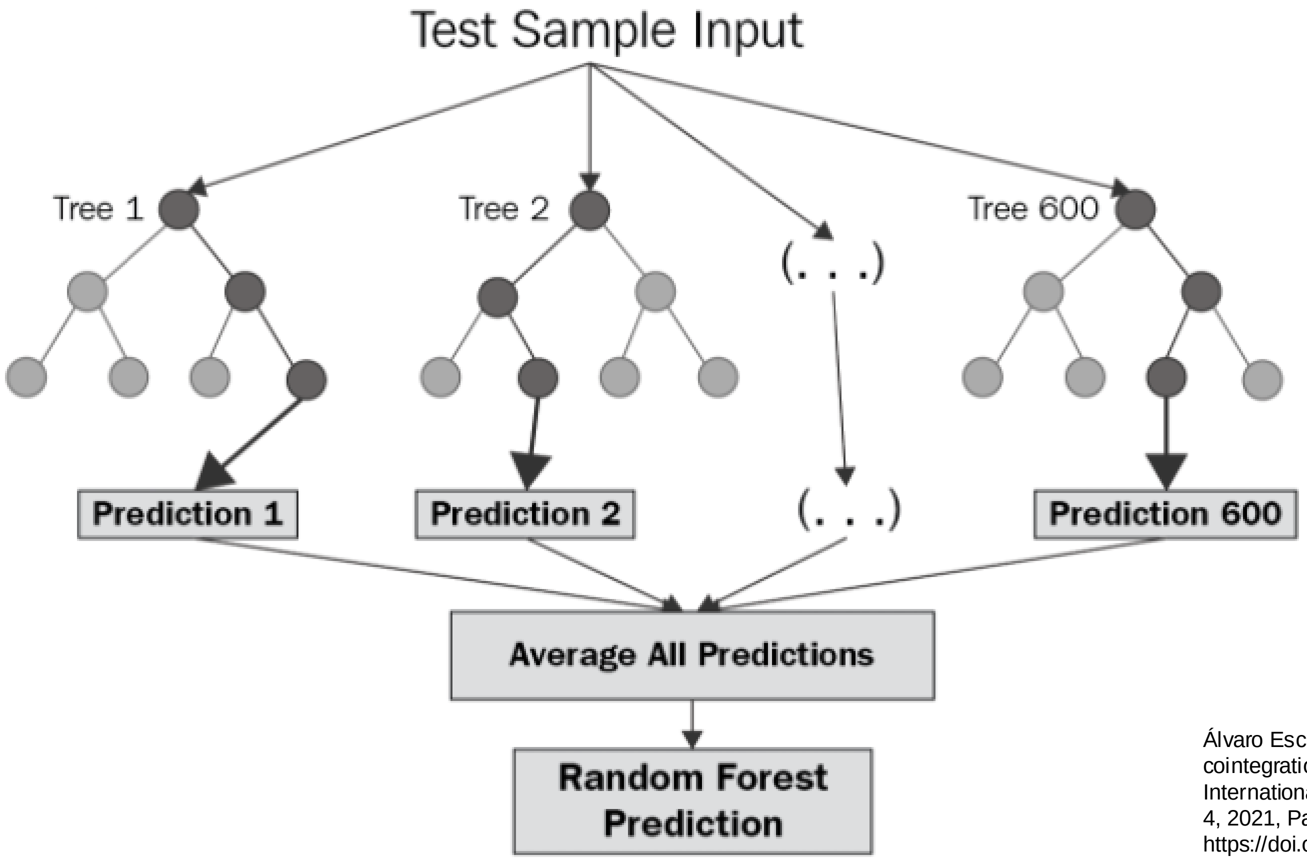
Machine Learning Methods Crash Course

Lin, W., Chen, J.S., Chiang, M.F., & Hribar, M.R. (2020). Applications of Artificial Intelligence to Electronic Health Record Data in Ophthalmology. *Translational Vision Science & Technology, 9*.



General Guidance: Stay as simple as you can (start at linear regression and work higher)

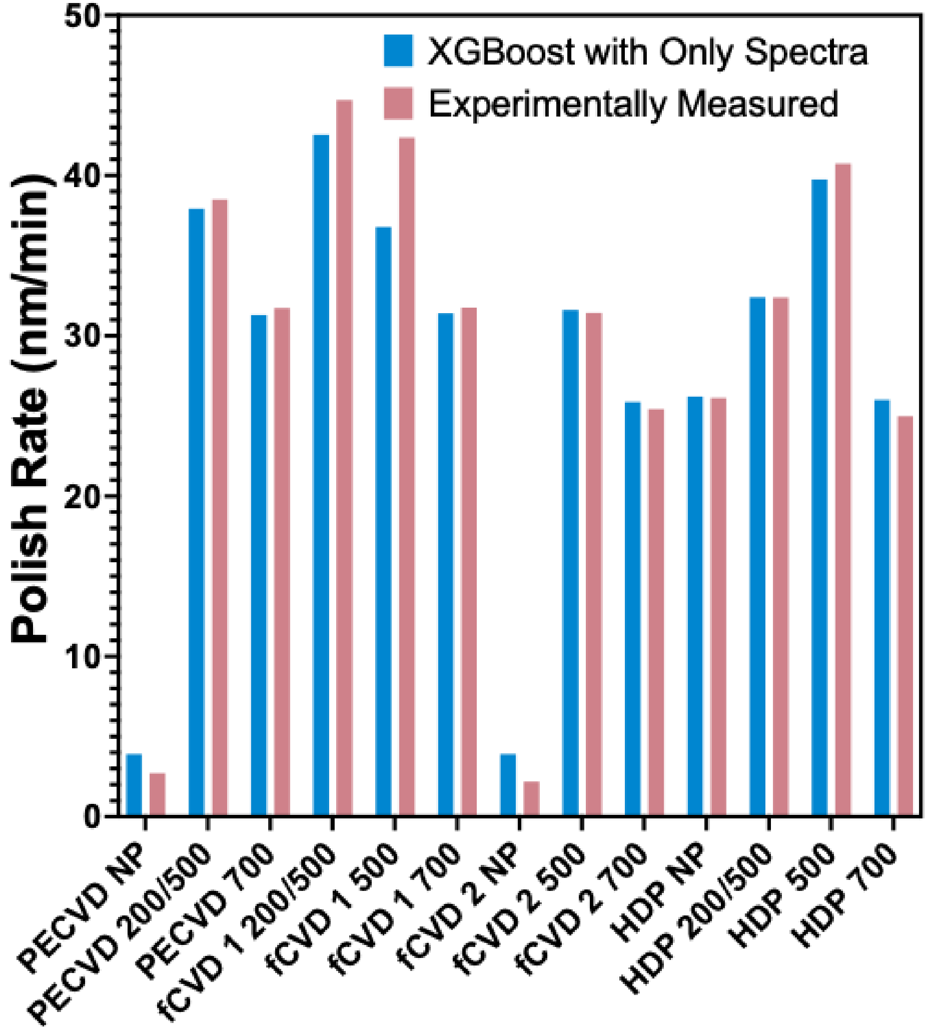
Random Forest Crash Course

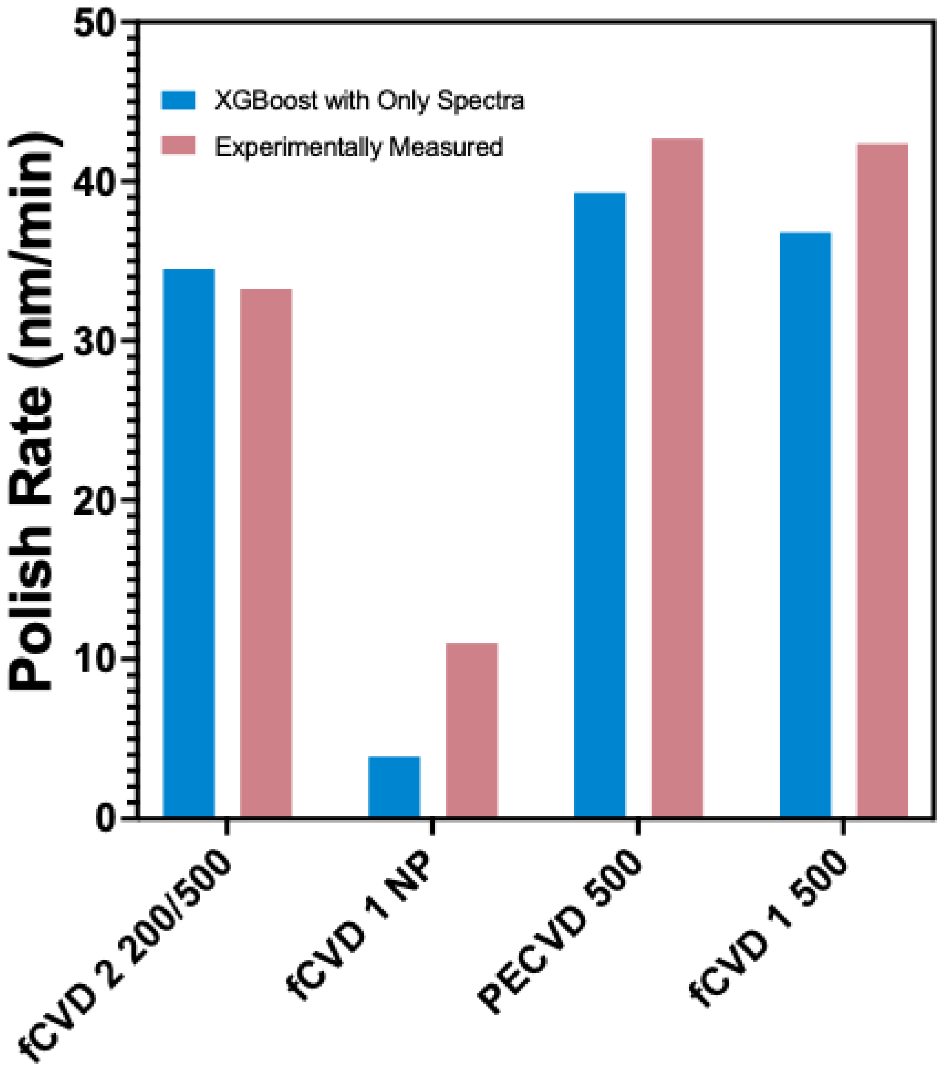


Álvaro Escribano, Dandan Wang, Mixed random forest, cointegration, and forecasting gasoline prices, International Journal of Forecasting, Volume 37, Issue 4, 2021, Pages 1442-1462, ISSN 0169-2070, https://doi.org/10.1016/j.ijforecast.2020.12.008.

Each "tree" independently results in a solution based on different features. The average of these is our final solution.

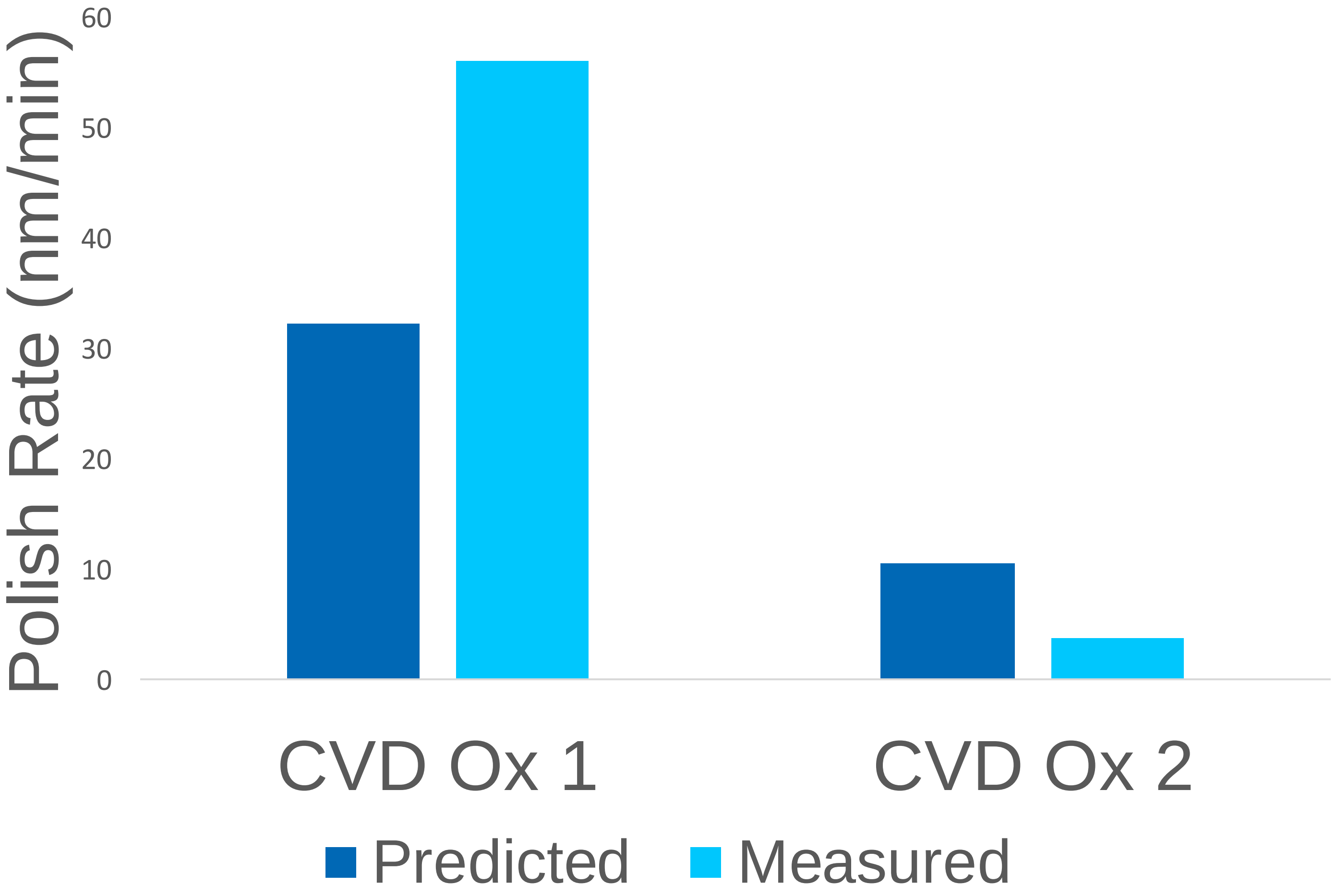
XGBoost Modeling w/Spectra Alone





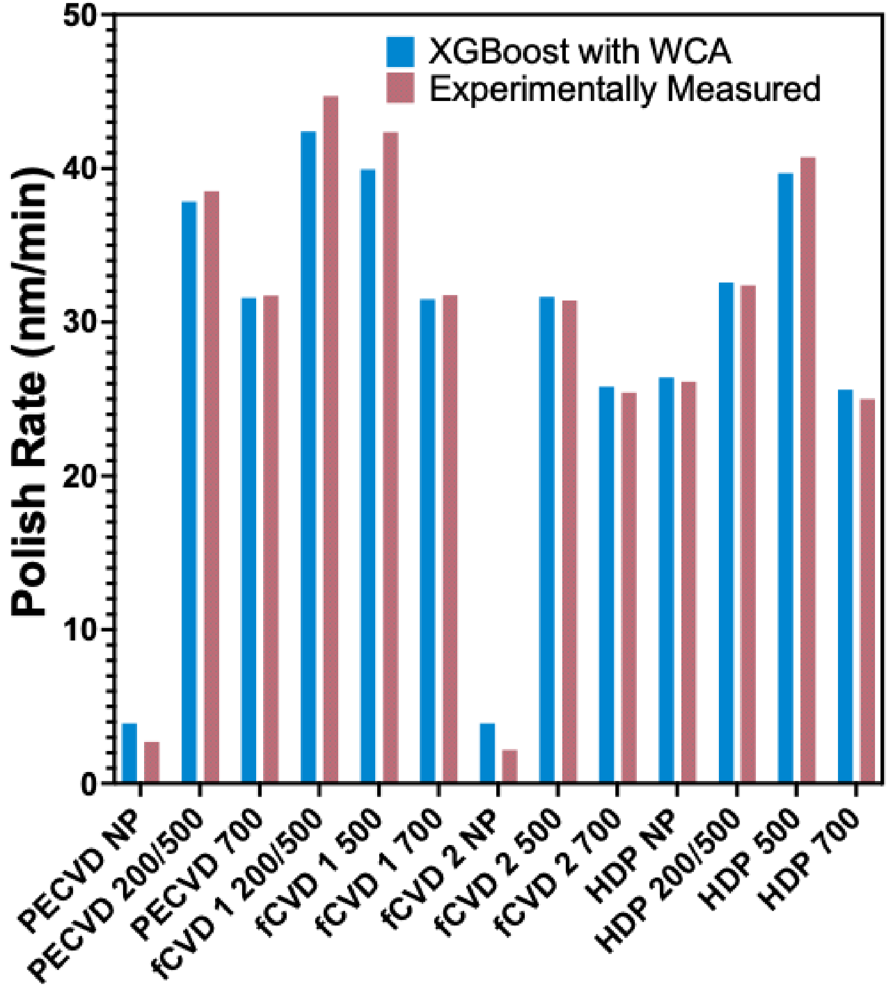
Train Dataset R2: 95.4%
Test Dataset R2: 87.4%

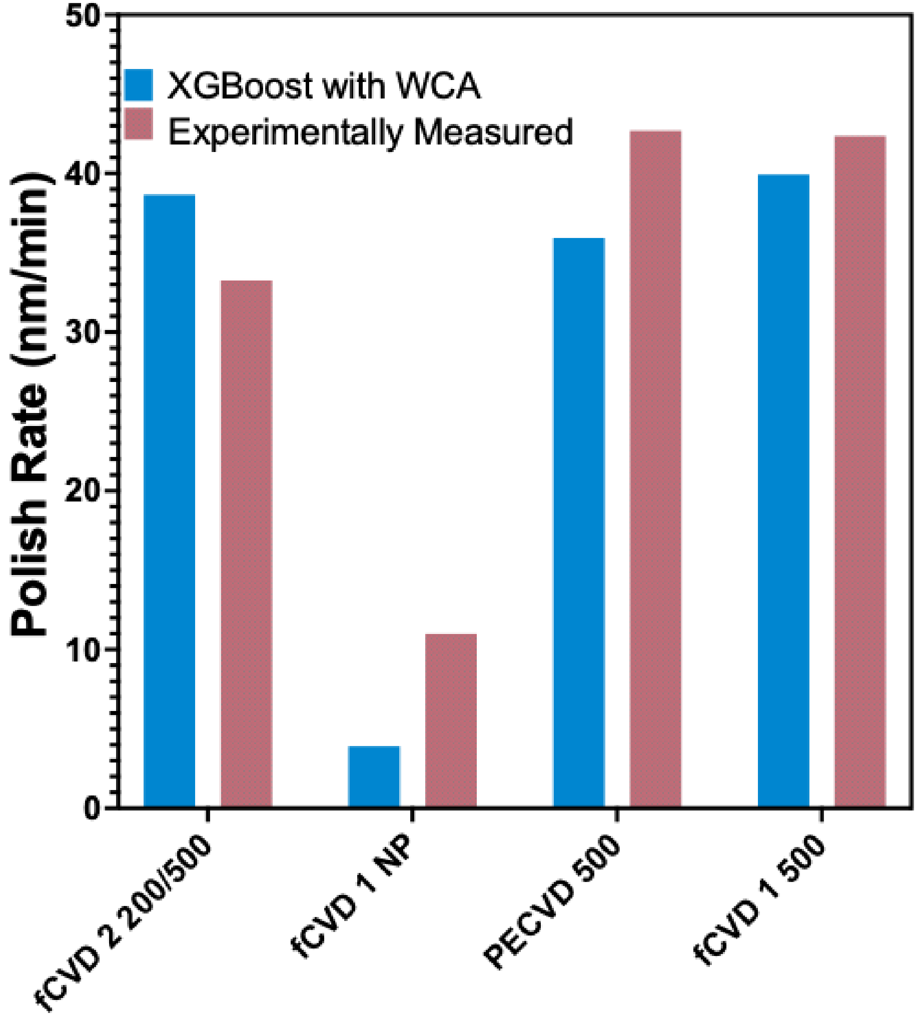
What about films it has never seen?



Definitely some overfit, but appears to be a solid classifier

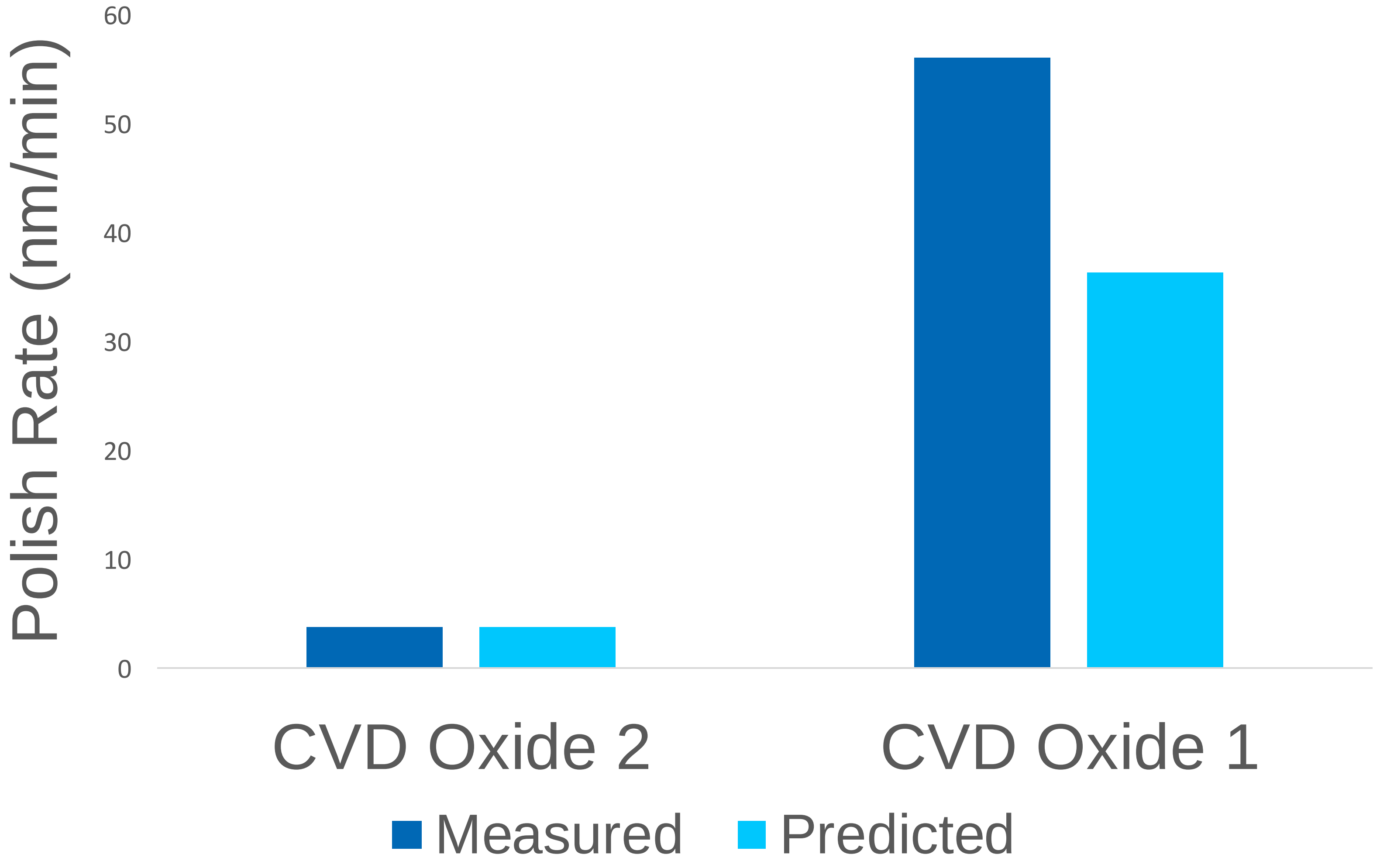
XGBoost Modeling w/WCA Added





Train Dataset R2: 96.0%
Test Dataset R2: 85.7%

XGBoost Modeling w/WCA Added



The Modeling Approach Used Is A Balance

Strengths

- Fully built on nondestructive techniques, allowing for faster iterating
- An extremely successful classifier- will it polish or will it not polish?
- Functions well as a regressor as well, with room for improvement
- The successful predictive use of only surface spectra supports the electrostatic interaction qualitative model

Limitations

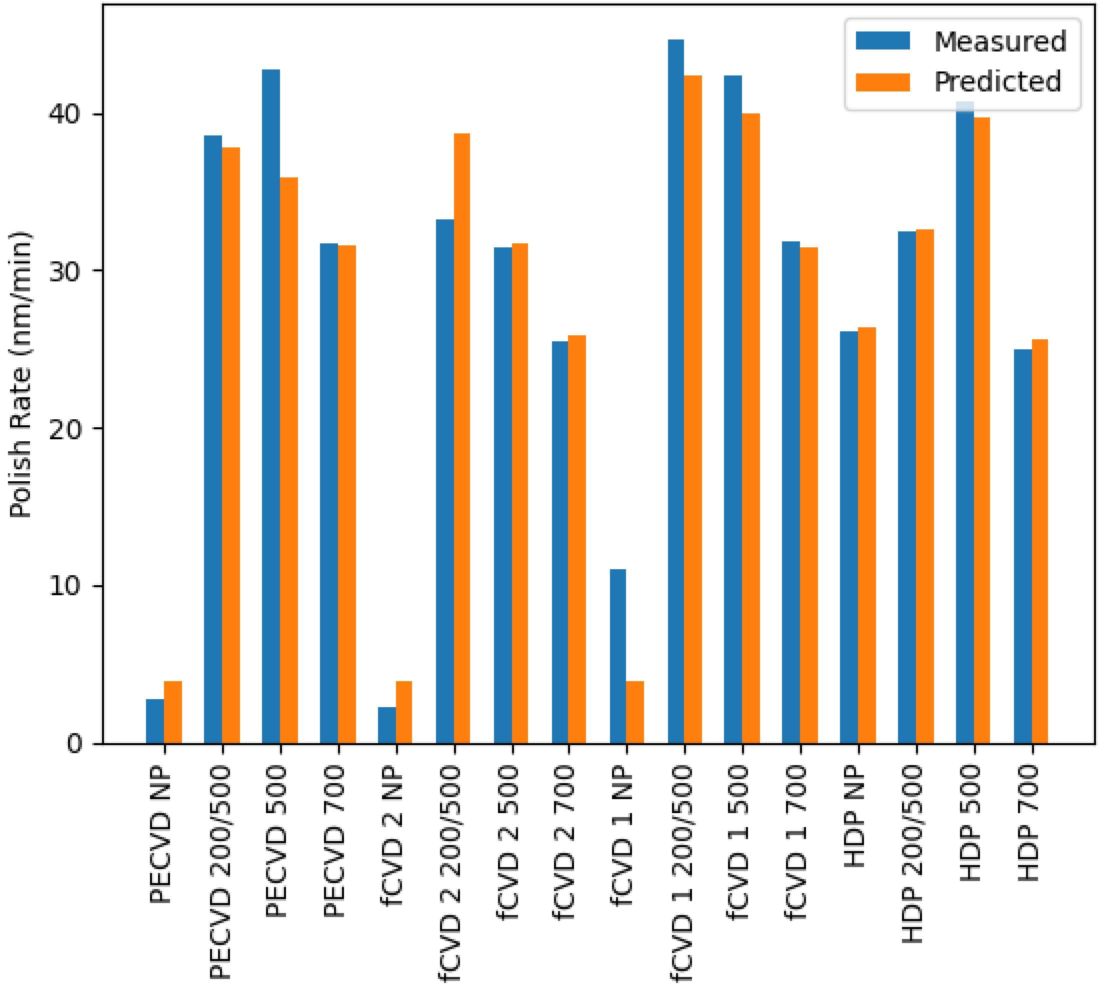
- Cannot capture extremes in polish rate (>40nm/min)- need to add more data points in the extremes
- Need to add more data on other moieties (carbon + nitrogen containing films) to build robustness

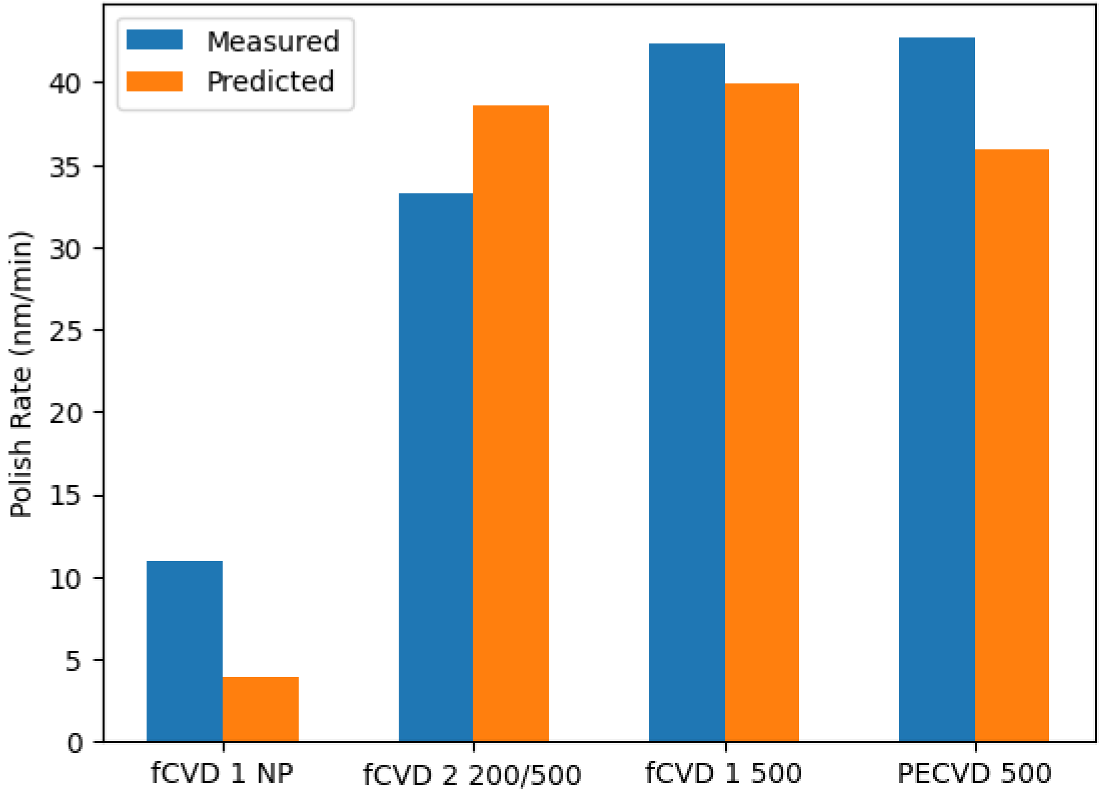
Conclusions

- This is the **<u>first successful modeling using noncontact metrology to predict oxide CMP polish rates</u>** (to our knowledge)
- AI Models are powerful tools to deconvolve interacting variables in CMP with immediate use cases in R&D and pathfinding
- No single spectroscopic variable appears to be predictive of polish rate changes in a vacuum
- A 2-step model is proposed, where the first step is electrostatic interaction with the surface, and the second is chemical binding to the surface followed by removal

Reference to research results, including comparisons to products, services or technology performance are estimates and do notimply availability. May contain information on products, services and technologies in development. Learn more at www.Intel.com/PerformanceIndex and www.Intel.com/ProcessInnovation.

XGBoost Modeling w/WCA & Density Added





Train Dataset R2: 94.6%
Test Dataset R2: 80.1%