2019 Planning Meeting

MATERIALS DATA SCIENCE
RELY

Research Proposal Presentations

September 11-12, 2019
Tinkham Veale University Center
Case Western Reserve University

MDS-Rely Project Proposals

1. Small Data & Early Indicators for Reliability Studies, Satish Iyengar
2. Image Machine Learning for Reliability, Roger French
3. Network Algorithms for Predictive Modeling of Photo-Voltaic Systems, Mehmet Koyuturk
4. Optimization for Better Decision Making and Design, Oleg Prokopyev
5. Reliability of Metallic Parts, John Lewandowski
6. Time Series Analysis of Power Generating Systems in the Field, Mark DeGuire
7. Image Machine Learning for Process Control, Jim McGuffin-Cawley
8. Reliability of Binder Jet 3D Printed Inconel 625: Effect of Binder Burnout, Markus Chmielus
9. Scratch Corrosion Study Protocol and Supervised Image Machine Learning, Jennifer Braid
10. Service Lifetime of Polymers, Laura Bruckman
11. Understanding Mechanical Abrasion Reliability of Multifunctional Coatings and Surfaces, Paul Leu
12. Data Analytics for Benchmarking of Degradable Biopolymer Films and Tapes, Mostafa Bedewy

Project Proposal
(September 2019)

Project Title: Small Data & Early Indicators for Reliability Studiess

Principal Investigator(s): Satish Iyengar

New Project: X

Thrust Area: 3. Reliability Studies

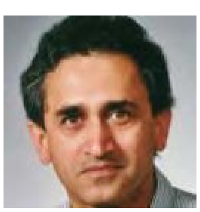
Abstract: I describe the role of statistics and data science for detecting early indicators of process degradation and learning patterns in a sparse-data environment.

I discuss how to identify markers, such as trends in important indexes, of impending problems. In some of my previous with colleagues in Psychiatry we seek indicators of subtypes of disorders that may require differential treatments. Certain aspects of this problem have a long history, but there is renewed interest in them because of recent advances in computing that allows us to handle more complicated situations without having to make simplifying assumptions whose validity may be hard to verify.

Next, I discuss the application of statistical methods in a small data environment. Big data is a current buzzword, but often the available data are rather modest in size. The problem of making sound inferences based on small data sets requires careful modeling. The fact that there is underlying physics and materials science governing the problems also makes the modeling much more interesting than standard techniques like regression. For example, there may be good reasons to believe that a nonlinear function that we are trying to estimate is monotone increasing in some region, in which case, a shape-constrained model fitting procedure would be more efficient than an unconstrained one.

I also give a brief account of previous work and my department.

Tasks: We will work closely with industry partners to identify classes of problems; develop a general framework that addresses the problems; develop statistical procedures and related software; illustrate methods using both actual and simulated data.



Short Bio: Professor and Chair, Statistics Department at University Pittsburgh. Education: s AB in Mathematics from Harvard 1978; PhD in Statistics from Stanford 1982. Research collaborations primarily with biomedical researchers; have collaborated with colleagues at Westinghouse and EPRI, among others.

Statistics and Data Science

Satish Iyengar

Department of Statistics

University of Pittsburgh

September 11, 2019

Industrial Relevance and Novelty

- Develop procedures to identify early indicators of impending problems. Combine modern time series methods with machine learning techniques that are considerably more flexible than classical methods.
- Making decision making in a sparse-data environment. Requires careful modeling of known science and engineering, for example, to obtain or approximate optimal process solutions under constraints.
- Recent advances in statistics and data science aim to improve reproducibility. We also take advantage of computational advances.

Proposal Objectives

- Carefully model physical processes of interest across a range of applications. Ascertain sources of variation (factors within process, uncontrolled noise).
- Develop general procedures that will allow for their application to a wide range of problems.
- Mentor post-doctoral fellow and/or graduate student in the course of this work. Contribute to other educational efforts.
- Make all (R, Python) code that we develop available for public use.
- Introduce wider statistics and data science community to problems of interest.

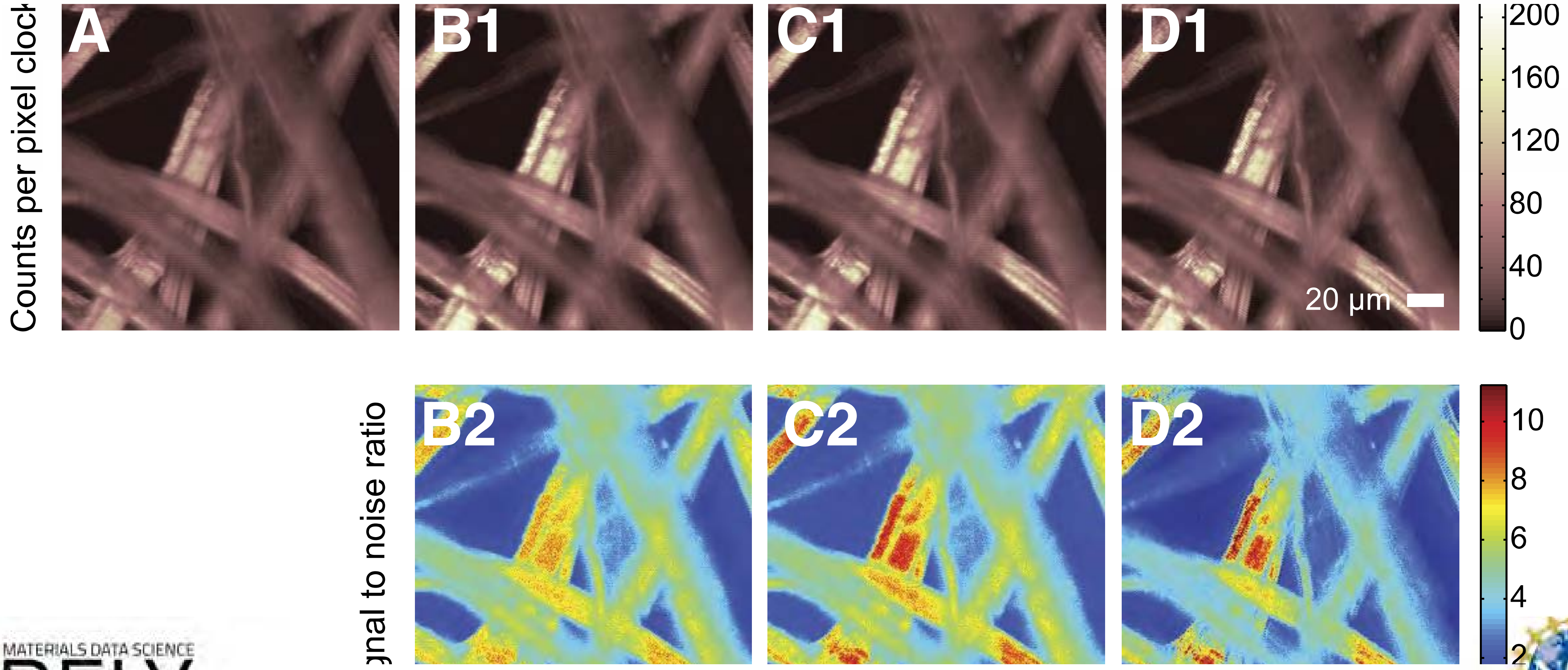
Simplified example

- Seek indicators of performance degradation of outdoor panel array
- Have many measures: weather, images of array, emergence of cracks
- Must first extract quantitative and qualitative summaries; often easy, but may need a model for, say, time series or spatial data
- Which measures are associated with degradation ? At first blush, easy: just check each measure against performance outcome
- Harder: rank the measures with respect to effect size, and derive confidence statements associated with ranking
- Typically start with some sort of regression (GLM); modify as needed
- Be careful about inference after selection; recent (debiasing) approaches allow for more honest estimates and confidence procedures

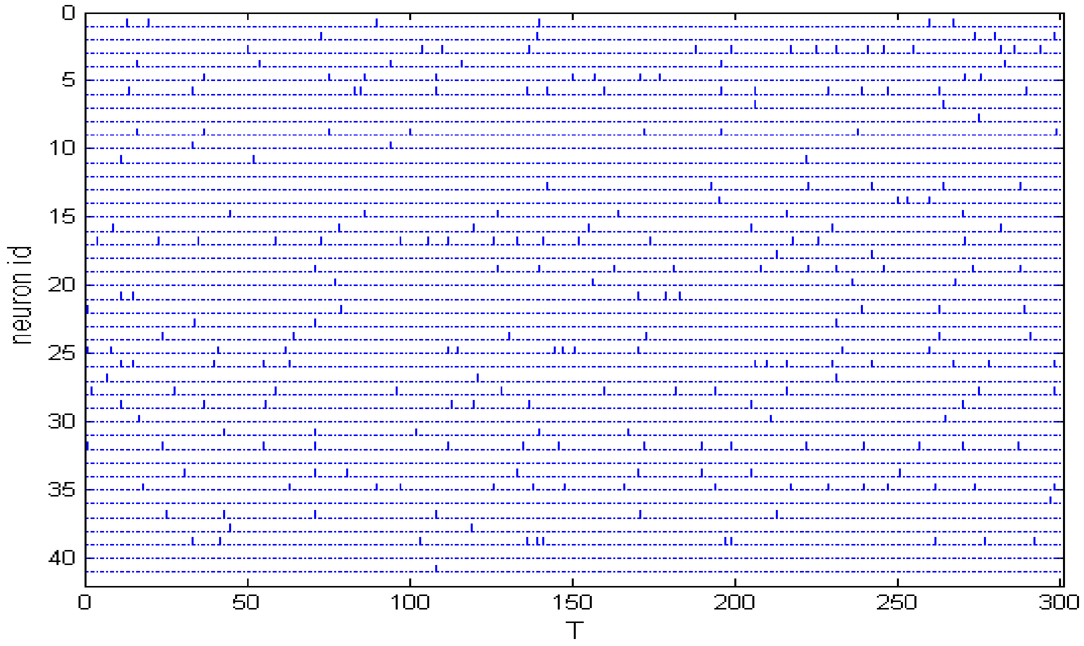
Two earlier examples of our work

- Improve SNR in two-photon laser scanning microscopy (TPLSM): work with David Kleinfeld of UCSD and my PhD student Burcin Simsek. Problem: account for the detector’s dead time, which results in reduced photon counts, to get improved estimate of count. Careful modeling led to considerable improvement in signal-to-noise ratio.
- Detect interactions in pre-motor cortex of macaque monkey in anticipation of task: work with Krishna Shenoy of Stanford, Aaron Batista of Pitt, and my PhD student Mengyuan Zhao.

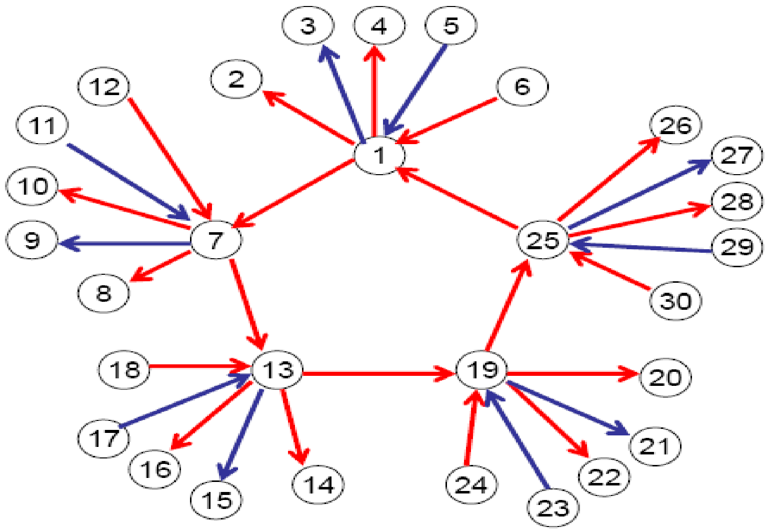
TPLSM


Counts per pixel clock
A
B1
C1
D1
200
160
120
80
40
0
20 μm
gnal to noise ratio
B2
C2
D2
10
8
6
4
2

Spike Train to Interactions (Monkey study)

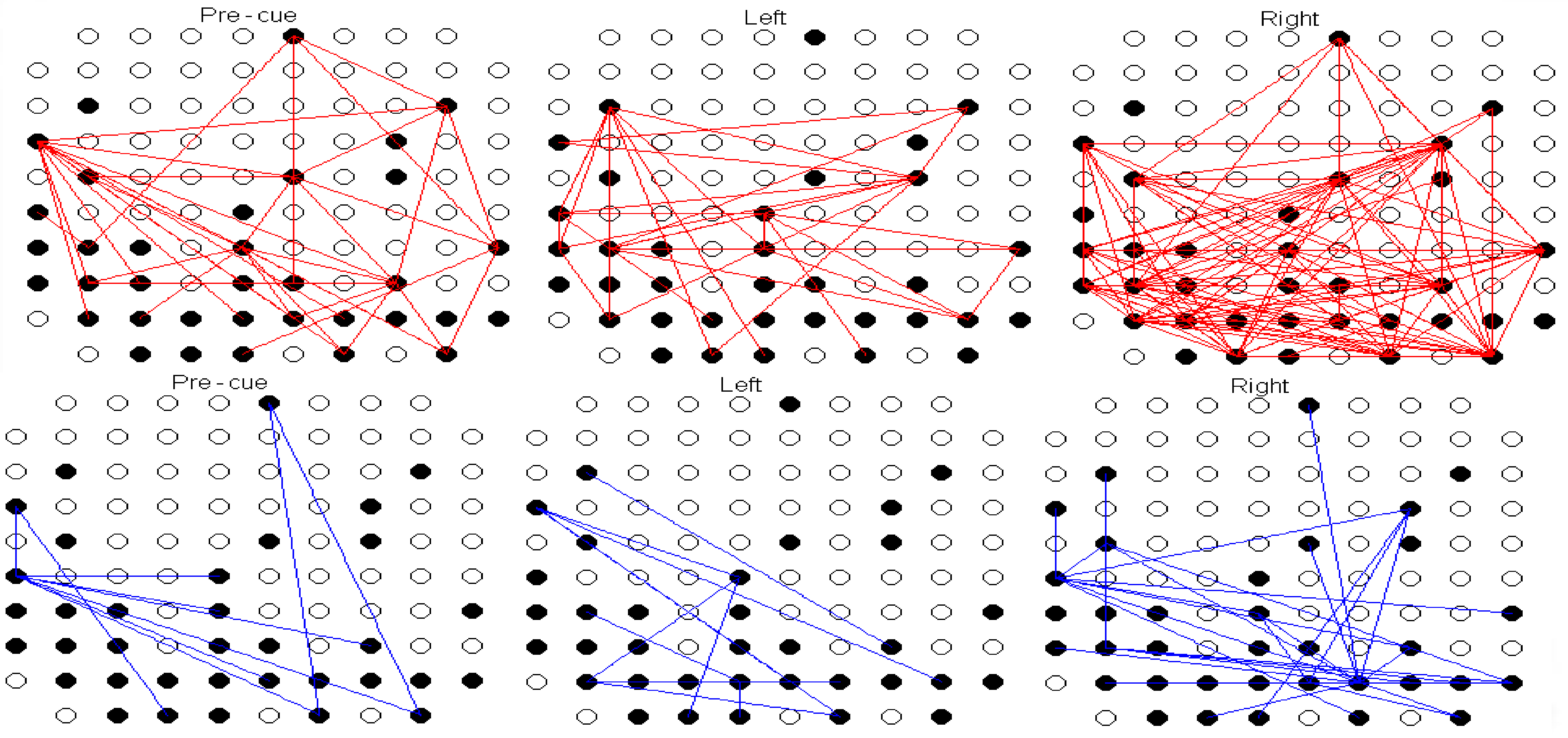


?



MATERIALS DATA SCIENCE
RELY

Interactions in channel map:



Deliverables

- All technical reports, presentations, and publications written by IUCRC-funded personnel
- All software developed by IUCRC-funded personnel

Our Department

- 10 tenure track faculty, all research-active
- Will recruit 2 more next year
- Graduate 3 to 4 PhDs per year
- Mentored many postdoctoral fellows
- Expanding Master's program
- New Data Science undergraduate major in the works

Our Faculty

- Joshua Cape, 2019, Johns Hopkins
- Kehui Chen, 2012, UC Davis
- Yu Cheng, 2006, Wisconsin
- Satish Iyengar, 1982, Stanford
- Christopher McKennan, 2019, U Chicago
- Lucas Mentch, 2015, Cornell
- Zhao Ren, 2014, Yale
- Allan Sampson, 1970, Stanford
- David Stoffer, 1982, UC Davis

Areas of Research

- Functional data analysis (Chen)
- Networks, graphs (Cape, Chen, Ren)
- Machine learning (Cape, Mentch)
- Multivariate analysis (Iyengar, Sampson)
- High-dimensional data (Cape, Ren)
- Dynamic treatment assignment (Cheng)
- Survival analysis (Cheng)
- Stochastic modeling (Iyengar)
- Bayesian inference (McKennan)
- Applications to psychiatry, neuroscience (Chen, Cheng, Iyengar, Sampson)
- Time Series (Stoffer)

Grant support

- NSF (Chen, Mentch, Ren, Stoffer)
- NIMH (Chen, Cheng, Iyengar, Sampson)
- NIH (Ren)
- Stanley Foundation (Iyengar)
- Several others pending

Collaborations and National Committees

- Biomedical engineering, Psychiatry, Neuroscience
- Engineering: EPRI, Westinghouse, Duquesne Light
- Expert witness in legal cases

- FDA panels (psychopharmacology, cardiovascular and renal)
- NSA advisory committee
- NSF and NIH review panels

Facilities / Leveraged Technology / Other Funding Sources

- University of Pittsburgh's Center for Research Computing (CRC).
- If necessary, have access to Pittsburgh Supercomputing Center.
- Consulting Centers at both Department of Statistics and Biostatistics.
- All research-active faculty in our department are well funded by the NSF, NIH, and foundations.
- Vibrant statistics, machine learning, and engineering communities in Pittsburgh.

Thank you !

Please submit questions on life form or save for networking breaks.

MDS-Rely Project Proposal
(A - T2 - 2)
(September 2019)

Project Title: Image Machine Learning for Reliability: Developing Unsupervised Machine Learning Pipelines Using Gray Level Co-occurrence Matrices for Texture Based Degradation and Performance Studies	
Principal Investigator(s): Roger French	**Researcher: Ahmad Karimi**
New Project: XX Renewal: **Term: 1 year, X 2 years**	**Start Date: June 2020**

Thrust Area: 2. Matls. Data Science

Objective: The goal of this project is to develop unsupervised machine learning (ML) approaches, so as to develop an image based ML pipelines for different reliability, degradation and microstructural studies of Materials. The major advantage of this approach is to remove the human intervention and to automate the process for large scale systems and big data sets which are usually impossible for humans to analyze. The idea is to extract features from the images based on their exhibited patterns by using the method of gray level co-occurrence matrix (GLCM) 14 Haralick Features[1]. GLCM is well adapted to textural image problems, where the reliability or performance can be tracked by the evolution of small features or textures at different scales in the acquired images.

Standards Used: Standard exposures of materials, components or systems, such as Damp Heat (IEC 61215), Thermal Cycling (IEC 61215), UV exposures (ASTM G154) or Full spectrum exposures (ASTM G155).

Code Developed: Expanding beyond the SDLE PVimage Python package, we will develop GLCM based unsupervised ML features to track texture-based feature evolution through time or across materials composition or processing.

Datasets Produced: The Unsupervised GLCM ML will be developed using the SunEd EL c-Si PV cell corrosion dataset. In addition we will draw upon image datasets from other MDS-Rely projects as inputs.

---

[1] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural Features for Image Classification, IEEE Transactions on Systems, Man and Cybernetics. SMC-3 (1973) 610–621. doi:10.1109/TSMC.1973.4309314.

Background: ML applied to large image datasets provides an impressive opportunity to extract quantitative information from reliability studies of materials and devices. The initial step of acquiring images in a reproducible manner can be achieved by using a simple photobooth and a dedicated mirrorless camera configured to disable white-balance and other typical settings used to automatically adjust photographs for "best" appearance of people. Once a set of images is acquired, then image processing (**Fig. 1A)**, typically done using Python3[2] with the Sci-kit-image[3] and OpenCV[4] image processing packages, is needed to center, adjust and cut the images to produce images of a given pixel resolution and orientation for use in ML.

In the study of exposure based reliability studies, using either accelerated exposure conditions such as Damp Heat, Thermal Cycling, or Weathering exposures[5], or real world exposure conditions, one of the most powerful methods to transform reliability tests into data and scientific results, is to take photographs over time, typically for 6 to 9 exposure steps. Due to the potential information density, and diverse image content of photographs, ML of step-wise acquired images represent a tremendous opportunity to advance the state of lifetime and degradation science[6].

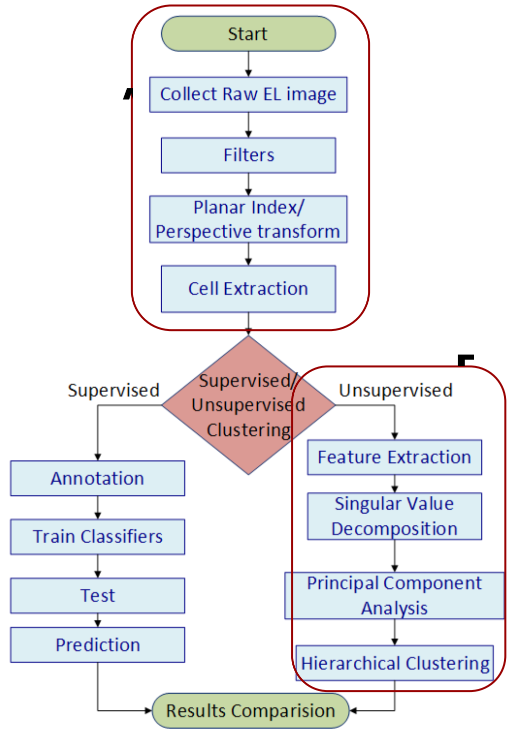


Figure 1. The three parts of an

Supervised or unsupervised ML[7] can be applied to the acquired step-wise images so as to identify significant degradation features and to quantify their occurrence with respect to the predictor variables of the study. We have demonstrated in our earlier works[8,9] that ML methods can classify the degraded cell images with high accuracy of 99%. Depending on the problem,

---

[2] Python Software Foundation: Python 3.6.8 documentation, n.d. https://docs.python.org/3.6/contents.html (accessed June 6, 2019).

[3] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, N. Yager, E. Gouillart, T. Yu, scikit-image: Image Processing in Python, PeerJ. 2 (2014) e453. doi:10.7717/peerj.453.

[4] *opencv: Open Source Computer Vision Library*. OpenCV, 2018.

[5] E44 Committee, Test Methods for Photovoltaic Modules in Cyclic Temperature and Humidity Environments, ASTM International, n.d. doi:10.1520/E1171-15R19.

[6] R.H. French, R. Podgornik, T.J. Peshek, L.S. Bruckman, Y. Xu, N.R. Wheeler, A. Gok, Y. Hu, M.A. Hossain, D.A. Gordon, P. Zhao, J. Sun, G.-Q. Zhang, Degradation science: Mesoscopic evolution and temporal analytics of photovoltaic energy materials, Current Opinion in Solid State and Materials Science. 19 (2015) 212–226. doi:10.1016/j.cossms.2014.12.008.

[7] G. James, D. Witten, T. Hastie, R. Tibshirani, An Introduction to Statistical Learning: with Applications in R, 1st ed. 2013, Corr. 5th printing 2015 edition, Springer, New York, 2013. http://www-bcf.usc.edu/~gareth/ISL/index.html.

[8] A. M. Karimi *et al.*, "Automated Pipeline for Photovoltaic Module Electroluminescence Image Processing and Degradation Feature Classification," *IEEE Journal of Photovoltaics*, pp. 1–12, 2019.

[9] A. M. Karimi *et al.*, "Feature Extraction, Supervised and Unsupervised Machine Learning Classification of PV Cell Electroluminescence Images," in 45th IEEE PVSC, 2018, pp. 0418–0424.

domain predictors could be material, supplier, configuration and even more complex features developed using feature engineering. In addition, image features developed for materials microstructure studies can also play a critical role in defining relevant microstructural features that impact performance and reliability.

Project Tasks: The major components of an image ML pipeline (**Fig. 1)** consist of an initial pipeline segment focusing on image processing **(Fig. 1A)** of the input images to prepare them for the input to ML algorithms, which is either a supervised or unsupervised ML pipeline segment. Supervised ML method, requires an initial phase of human assessment, to sort images into categories (or classes) that represent the desired degradation groups we want to identify. Due to the required human "supervision" of the initial data, supervised ML, while very useful is not robust to changes in the image quality, nor to identifying new degradation modes that may appear later among images being acquired.

In this project, our focus will be on the unsupervised ML segment (**Fig. 1B**) of the pipeline.

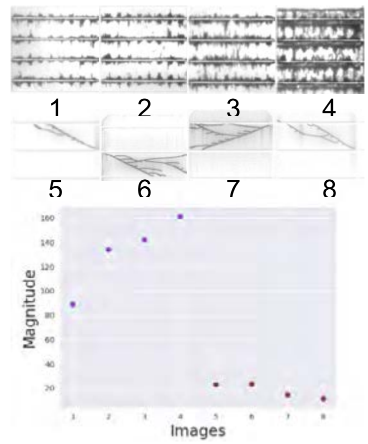


Figure 2: The figure shows the values of a Haralick Feature (Information Measure) for 8

The major tasks in this segment are:

- *Feature Extraction:* Images in themselves are a huge set of data points containing hundreds of thousands of pixel values. So, to convert images into computationally feasible objects and to analyse them, it is imperative to extract meaningful features from them and Haralick feature extraction is one such method. Figure 2 is a demonstration of an Haralick feature called Information Measure (IM) for eight images. We can observe that these images represent two major sets of textural patterns, images 1-4 have high values for the IM features, whereas the images 5-8 have low values. Similarly, remaining Haralick features also helps us distinguish images with different patterns.
- *Identification of critical features:* Usually the nature of the Haralick features are that they are correlated so it is beneficial to identify the set of features which are independent or orthogonal to each other. There are various statistical methods to achieve this such as SVD, NMF, etc. In this step, we would explore these techniques to convert highly correlated feature sets into independent ones.

- *Unsupervised Machine Learning :* This method is also called a clustering technique. In this task we would use the set of features selected above where each feature vector represents an image to cluster them into various categories depending on the data. There are various ways of unsupervised ML methods. We would use a Hierarchical clustering technique because we can leverage it for having multiple sets of clusters depending on the cutoff values.

Benefits to Members: The automated pipeline of image processing and ML pipeline has diverse application in material data science fields. In the manufacturing sector, this will be useful to have in the production assembly line. It can be useful in providing the quantitative analyses to the production team and can also be a part of testing setup. From the perspective of operations management, , the main idea is to automate and speed up the study of the response of material to the external exposures, stresses, stimuli, etc. These kinds of automated tools can be useful in the recycling industry to objectively estimate the worth of the products at the end of their commercial lifecycle and thus better evaluate their products.

As an example, in one of our current projects, we are studying the exposure of Al-Ni alloy in inert levitated environment to find out the different pathways of crystallization that the alloy follows under varying temperature and percentage composition of both metals. The complex model shown in Figure 3(c) can only be possible using image analysis methods.

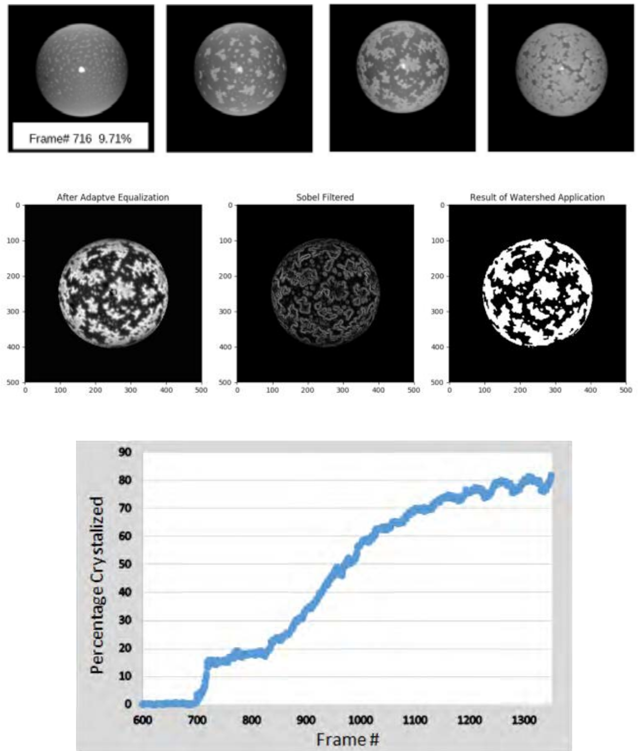


Figure 3: (a) is a collection of ~500,000 input images of levitated alloy taken at various time during exposure. (b) shows different kind of filters applied on the images to find out the regions of crystallization. (c) is the rate of growth of crystallization under one exposure condition

Timeline :

Tasks	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Feature Extraction	15%	20%	20%	25%	20%			
Identification of critical features				10%	20%	40%	30%	
Application of Unsupervised Image Machine Learning							40%	60%

Developing Unsupervised Machine Learning Pipelines: Using Gray Level Co-occurrence Matrices for Texture Based Degradation and Performance Studies

Faculty Member(s): Roger H. French

University / Department: CWRU / MS&E

Proposed Project Duration: 2 years

Industrial Relevance and Novelty

Images have high information density

Machine Learning applied to images

- Can extract useful information very effectively
- They are robust for automated analysis

Taking photographs

- Is effective approach to increase information capture
- From standardized reliability tests

A robust image processing pipeline has broad applicability

- Can be used for supervised (human classified) images
- Or unsupervised images

Proposal Objectives

Develop unsupervised machine learning (ML) pipelines for Materials Research

- **Reliability**
- **Degradation**
- **Microstructural studies**

Automation

- **Large scale systems**
- **Big data sets**
- **Usually impossible for humans to analyze.**

Apply Gray Level Co-occurrence Matrix (GLCM) for Image Feature Extraction

- **Extract 14 Haralick Features[1]**
- **GLCM is well adapted to textural image problems**

1. R.M. Haralick, K. Shanmugam, I. Dinstein, Textural Features for Image Classification, IEEE Transactions on Systems, Man and Cybernetics. SMC-3 (1973) 610–621. doi:10.1109/TSMC.1973.4309314.

Proposal Deliverables

An automated image machine learning pipeline

- Based on our PVimage Python Package[1]

Focused on unsupervised image machine learning

- Using GLCM features, such as the Haralick Features

Applicable to textural features

- Arising in corrosion type studies[2]
- And microstructural studies[3]

1. A. M. Karimi, J. S. Fada, M. A. Hossain, S. Yang, T. J. Peshek, J. L. Braid, R. H. French, **Automated Pipeline for Photovoltaic Module Electroluminescence Image Processing and Degradation Feature Classification**, IEEE Journal of Photovoltaics. (2019) 1–12. doi:10.1109/JPHOTOV.2019.2920732.
2. A. M. Karimi, J. S. Fada, J. Liu, J. L. Braid, M. Koyutürk, R. H. French, **Feature Extraction, Supervised and Unsupervised Machine Learning Classification of PV Cell Electroluminescence Images**, in: 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC), 2018: pp. 0418–0424. doi:10.1109/PVSC.2018.8547739.
3. B.L. DeCost, E.A. Holm, **A computer vision approach for automated analysis and classification of microstructural image data**, Computational Materials Science. 110 (2015) 126–133. doi:10.1016/j.commatsci.2015.08.011.

Leveraged Technology

This builds on our PVImage Python Package[1]

- Developed under DOE SETO funding

Which demonstrated

- A) Image Processing
  - Of PV cell images
- Supervised Image ML
  - For c-Si PV cell corrosion
  - Under Damp Heat exposures

Here we will develop

- B) Unsupervised Image ML
- Using the existing SunEd EL dataset

1. A. M. Karimi, J. S. Fada, M. A. Hossain, S. Yang, T. J. Peshek, J. L. Braid, R. H. French, **Automated Pipeline for Photovoltaic Module Electroluminescence Image Processing and Degradation Feature Classification**, IEEE Journal of Photovoltaics. (2019) 1–12. doi:10.1109/JPHOTOV.2019.2920732.

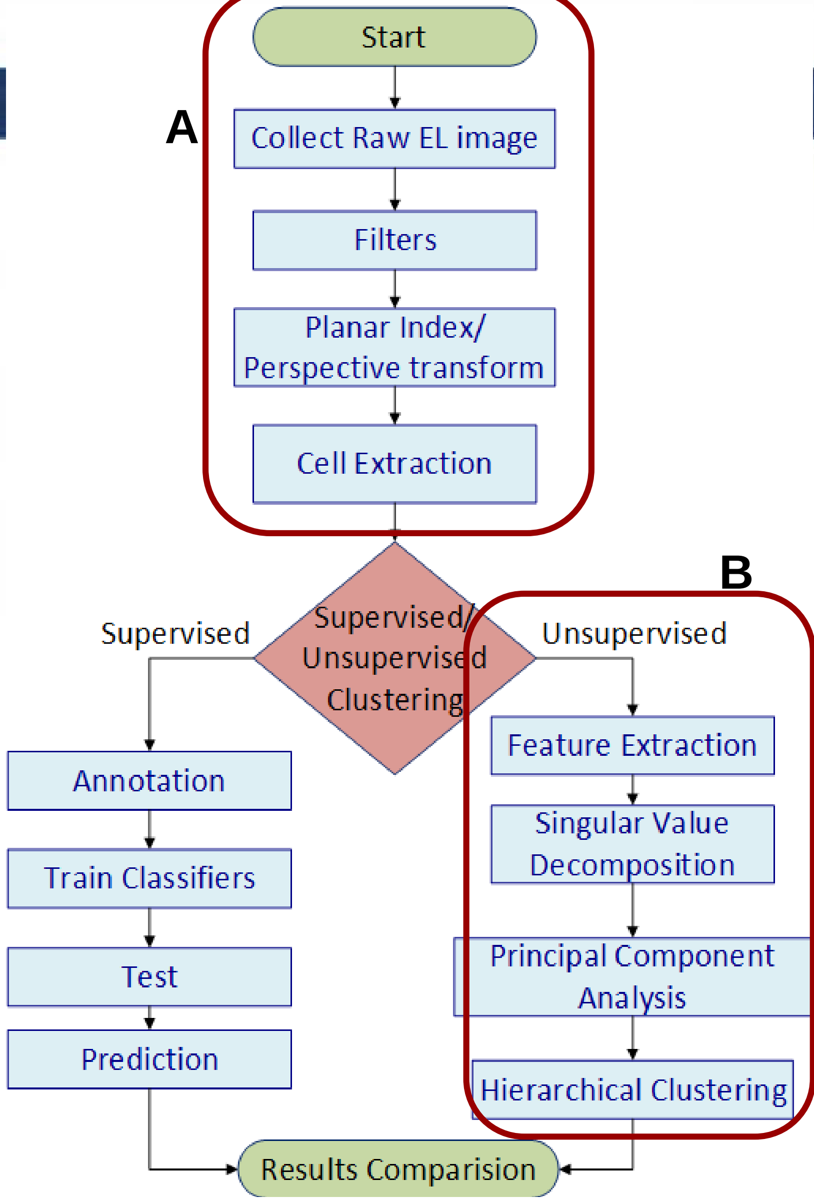


Figure 1. The three parts of an integrated image machine learning pipeline, including A) image pre-processing, and either supervised or B) unsupervised machine learning.

MATERIALS DATA SCIENCE RELY

“SunEd” Dataset Description

5 degradation modes well represented

- Good [Not degraded]
- Cracked
- Busbar corroded
- Edge Darkening
- Between the busbar darkening

3550 sorted cell images from 5400 extracted

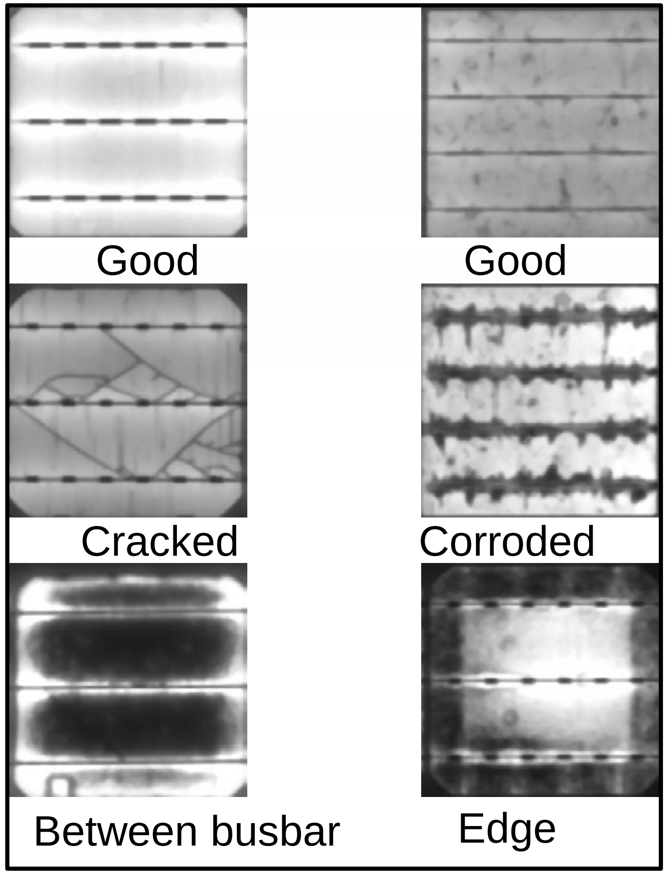
- For supervised learning

Data augmentation to increase training set

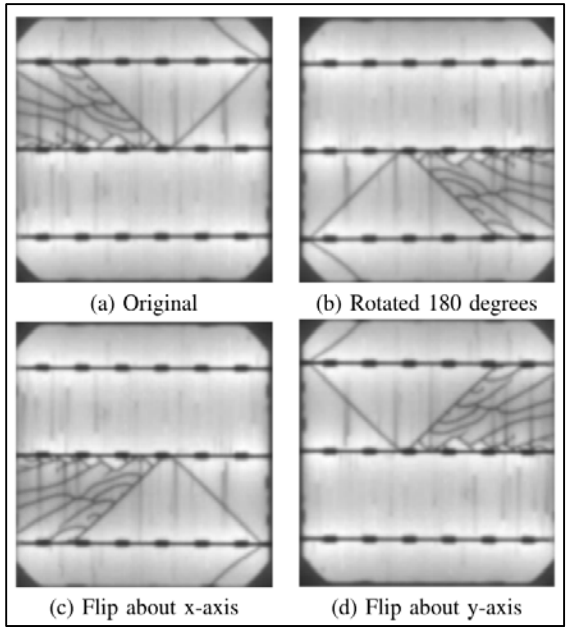
- Rotates 180 degrees, flip about x-axis, flip about y-axis
- Increases data 4 times original size

Stratified sampling

- Samples in each category
- 80:20 training to testing set ratio



EL Image types



Data Augmentation

MATERIALS DATA SCIENCE RELY

[1] J. S. Fada, M. A. Hossain, J. L. Braid, S. Yang, T. J. Peshek, and R. H. French, “Electroluminescent Image Processing and Cell Degradation Type Classification via Computer Vision and Statistical Learning Methodologies,” Washington, D.C., 2017.
[2] J. Friedman, T. Hastie, and R. Tibshirani, The elements of statistical learning, vol. 1. Springer series in statistics New York, 2001.

CRADLE: Distributed Computing & NoSQL database environment

Computation & Performance

- Spark
- Slurm
- Hadoop MapReduce

Data management

Integrate Spark

- **for distributed processing**

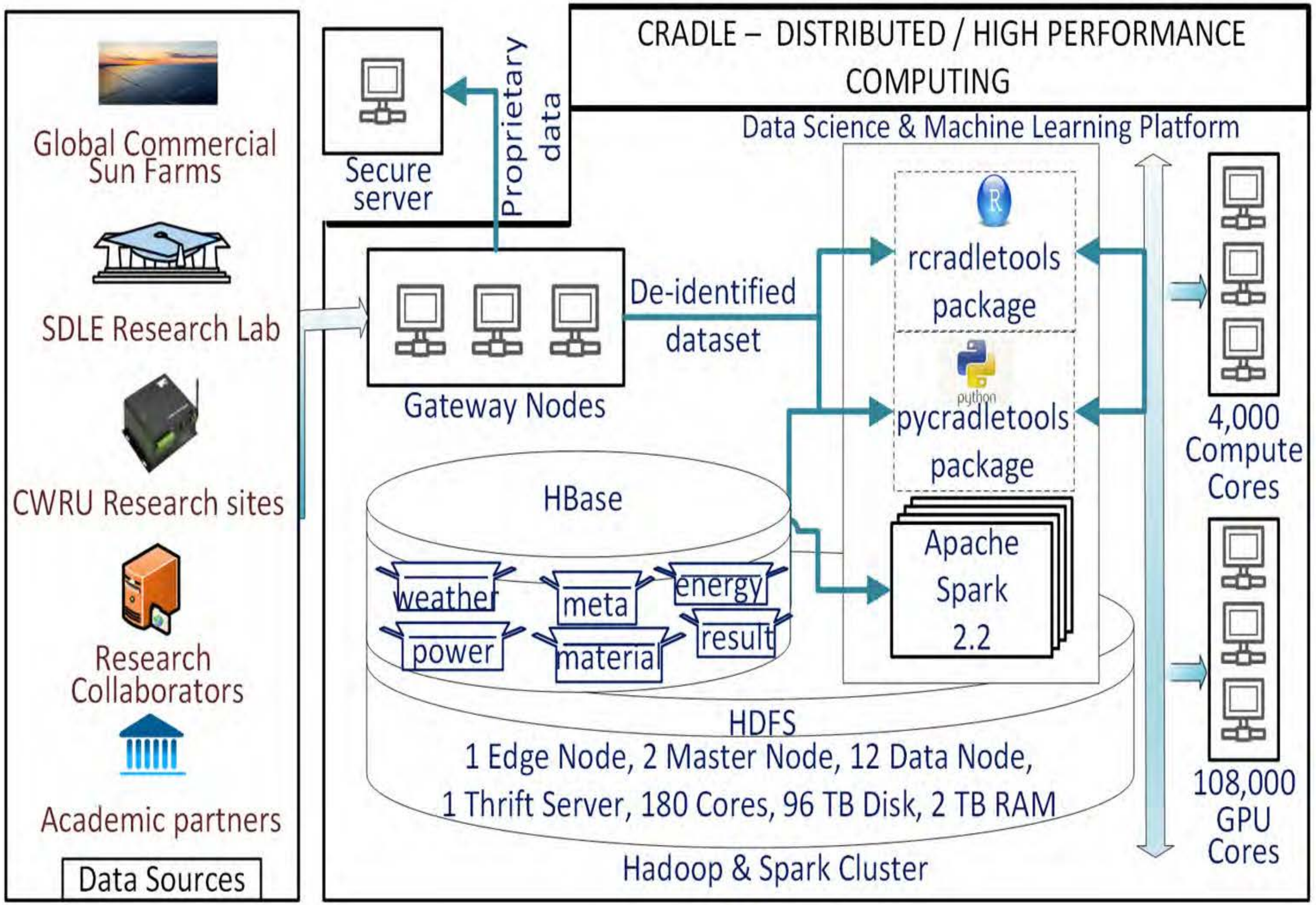
Upgraded functions

- **In cradletools**

Added functionality

- **Similar to cradlesgis**
- **In pytcradletools**

MATERIALS DATA SCIENCE
RELY



Proposed workflow

Image Collection

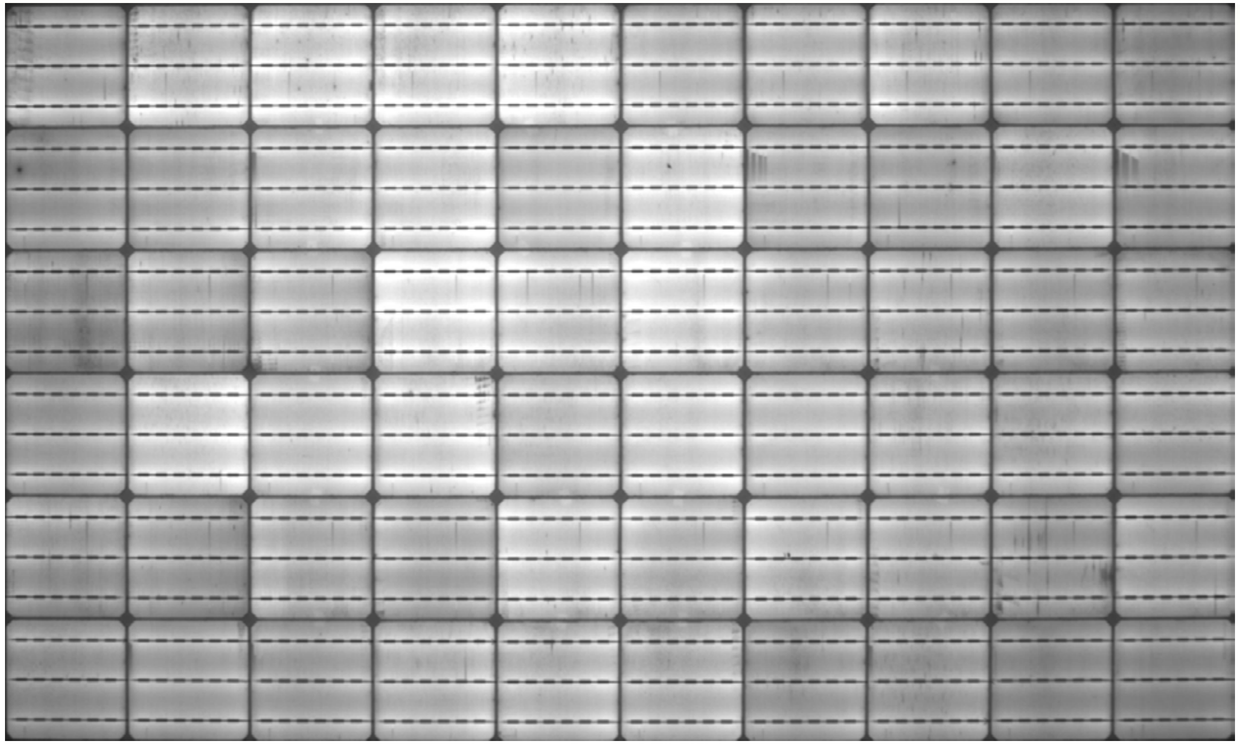
Image processing

- Planar Indexing
- Cell Extraction

Unsupervised Machine Learning

- Development of GLCM
- Assignment of GLCM features
  - To meaningful characteristics

Results



Electroluminescent images

- Through exposure time in Damp Heat
- Of a c-Si PV module

5000 images in the “SunEd” data set

MATERIALS DATA SCIENCE RELY

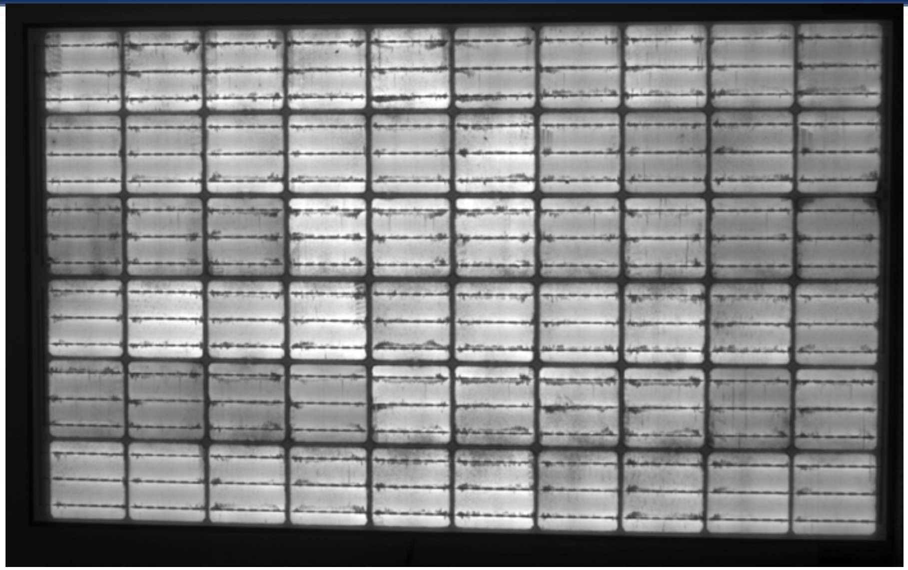
EL: Planar Indexing

Solution: Planar indexing

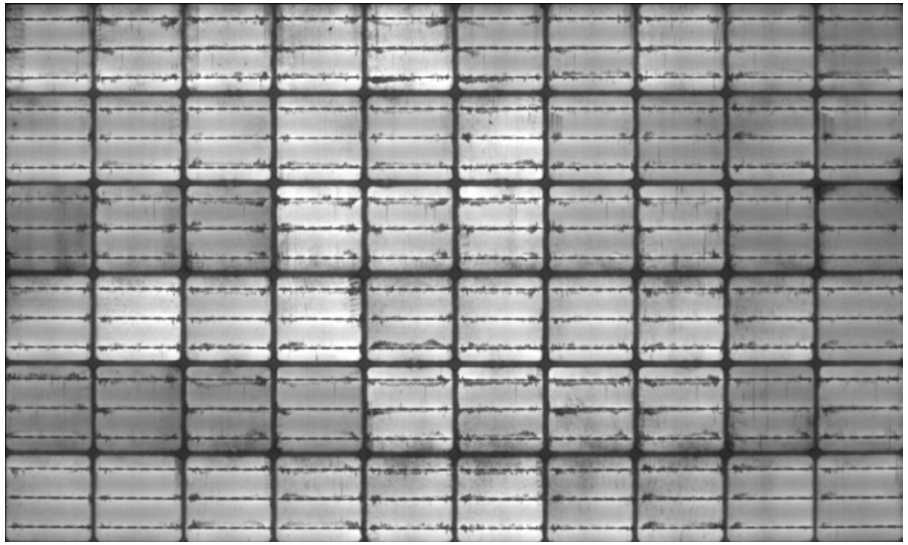
- **Find active area**
  - Filtering
  - Thresholding
  - Convex hull
- **Identify the corners**
  - Edge detection with linear regression
  - Intersection of module edges
- **Perspective transform the images**
  - Transformed module into image plane

Results: ~3% of images rejected

- **Due to severe degradation**



Original Unoriented EL Image

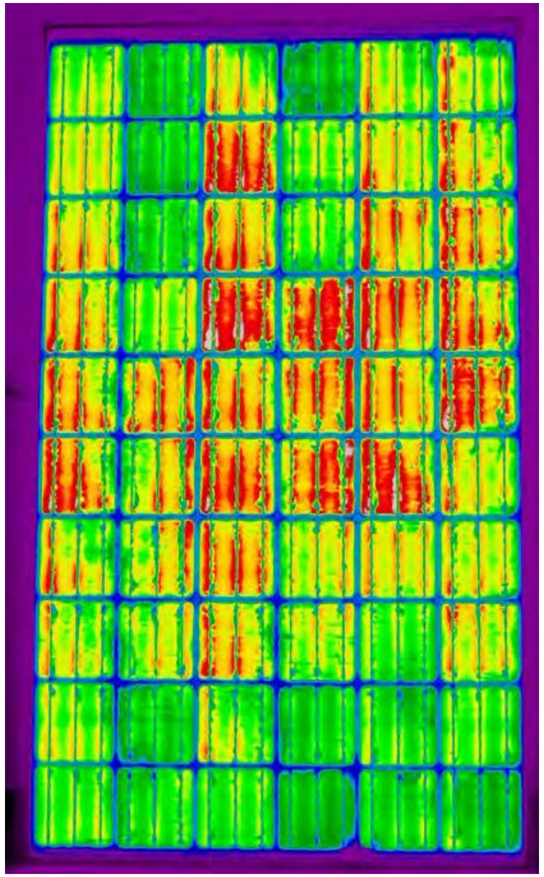


Planar Indexed EL Image

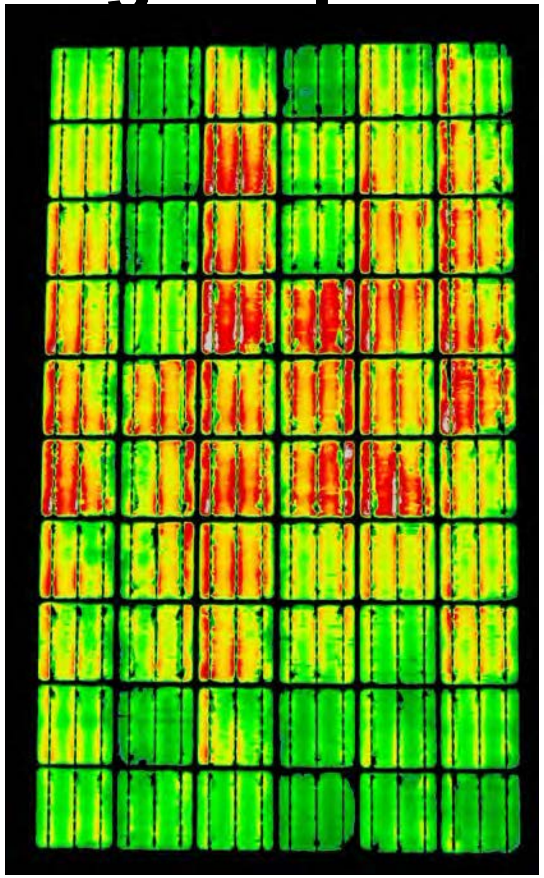
MATERIALS DATA SCIENCE RELY

[1] J. S. Fada, M. A. Hossain, J. L. Braid, S. Yang, T. J. Peshek, and R. H. French, "Electroluminescent Image Processing and Cell Degradation Type Classification via Computer Vision and Statistical Learning Methodologies," Washington, D.C., 2017.
[2] Ulrike Jahn et al., "IEA-PVPS Task 13, SubTask 3.3: Review on Infrared and Electroluminescence Imaging for PV Field Applications," International Energy Agency, 2017.

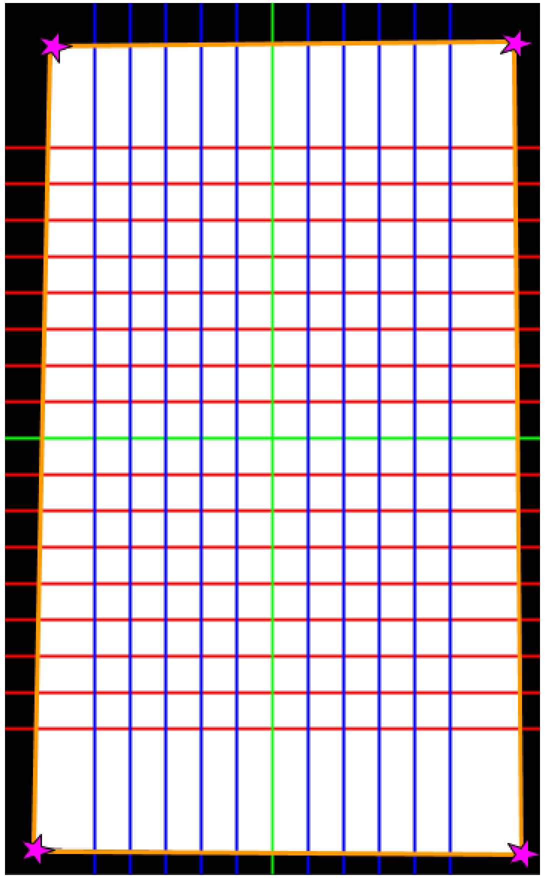
EL: Planar Indexing Steps



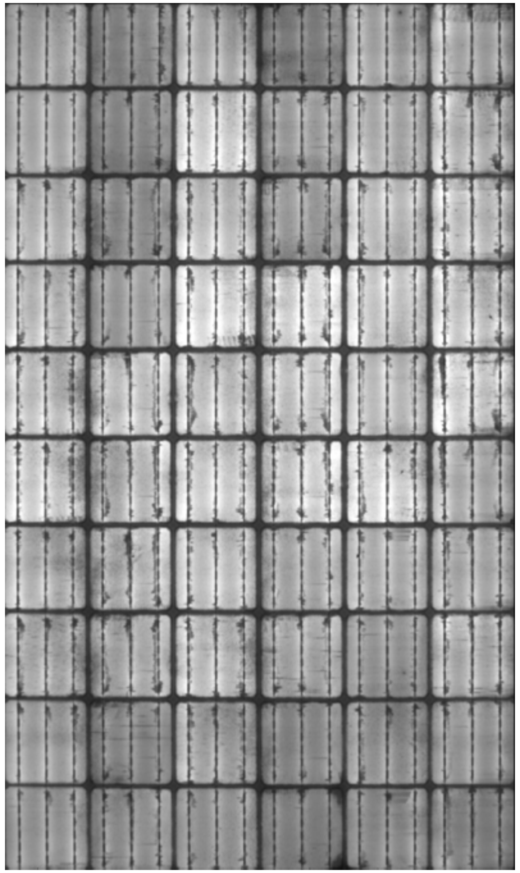
Median Filter
(Spectral Color Mapping)



Background Threshold
(Spectral Color Mapping)



Convex Hull/
Corner Finding



Planar Indexed EL Image
(Perspective Transformed)

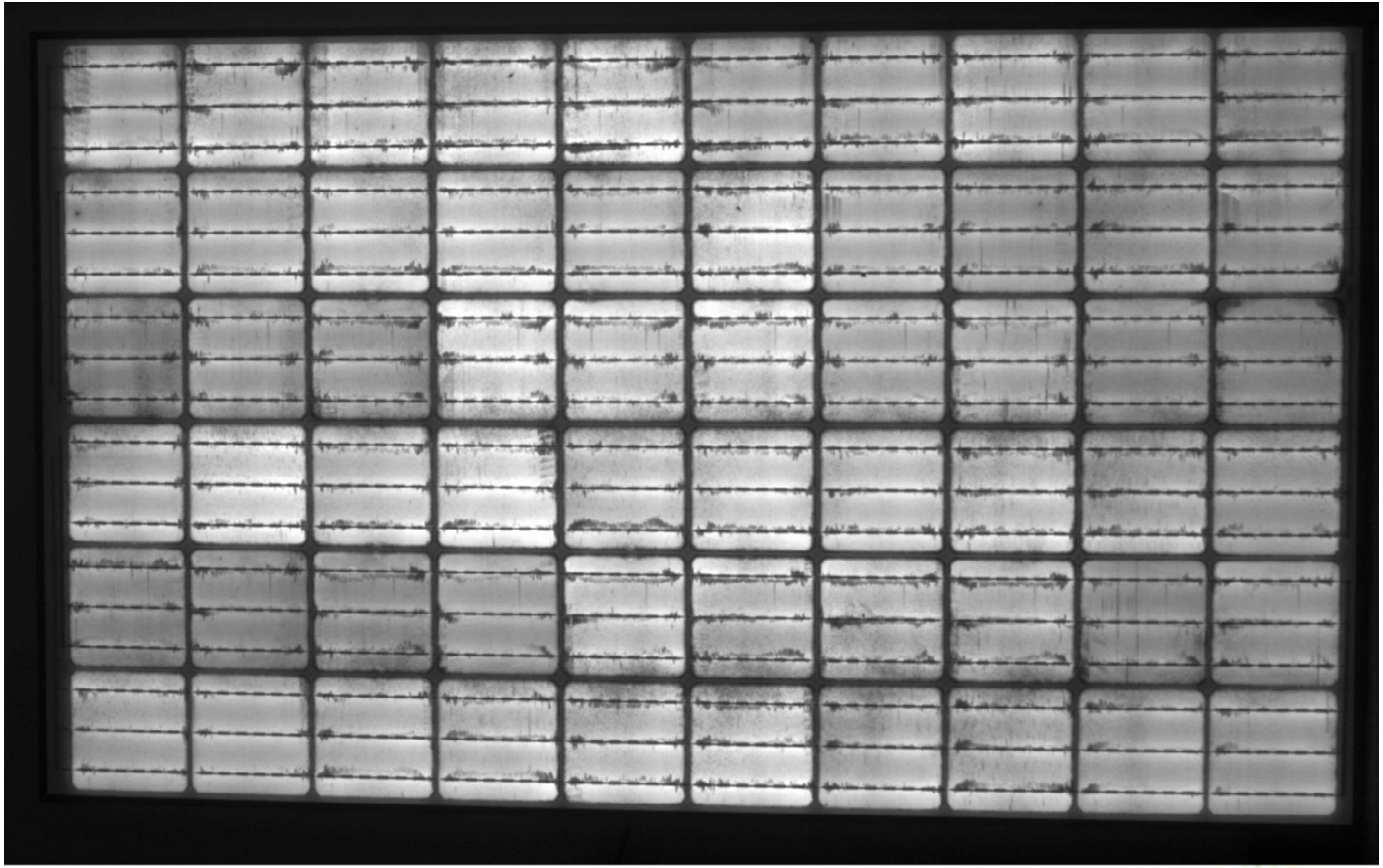
MATERIALS DATA SCIENCE
RELY

[1] J. S. Fada, M. A. Hossain, J. L. Braid, S. Yang, T. J. Peshek, and R. H. French, "Electroluminescent Image Processing and Cell Degradation Type Classification via Computer Vision and Statistical Learning Methodologies," Washington, D.C., 2017.
[2] Ulrike Jahn et al., "IEA-PVPS Task 13, SubTask 3.3: Review on Infrared and Electroluminescence Imaging for PV Field Applications," International Energy Agency, 2017.

Image Processing for planar Indexing

Done in Python 3.6.8

- Using scikit-image
- And OpenCV

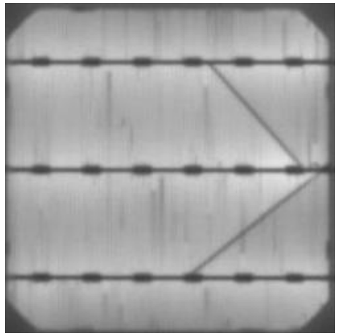


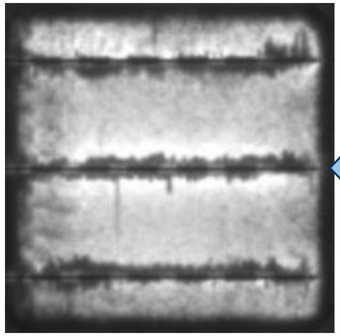
MATERIALS DATA SCIENCE RELY

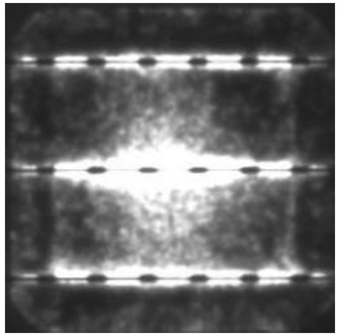
EL: Cell Extraction

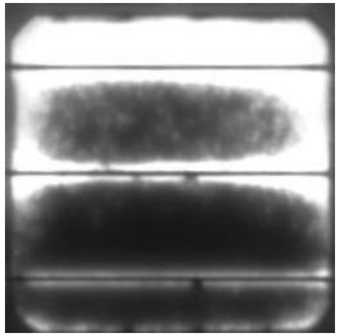
Cell Extraction

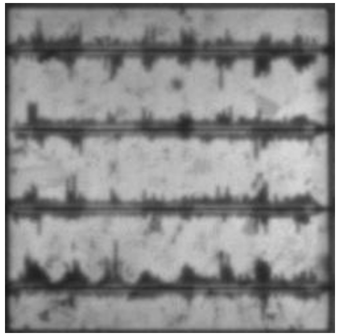
- **Start with planar index module**
- **Simple matrix slicing used to extract cells**
  - Further refined image processing would result in lost information

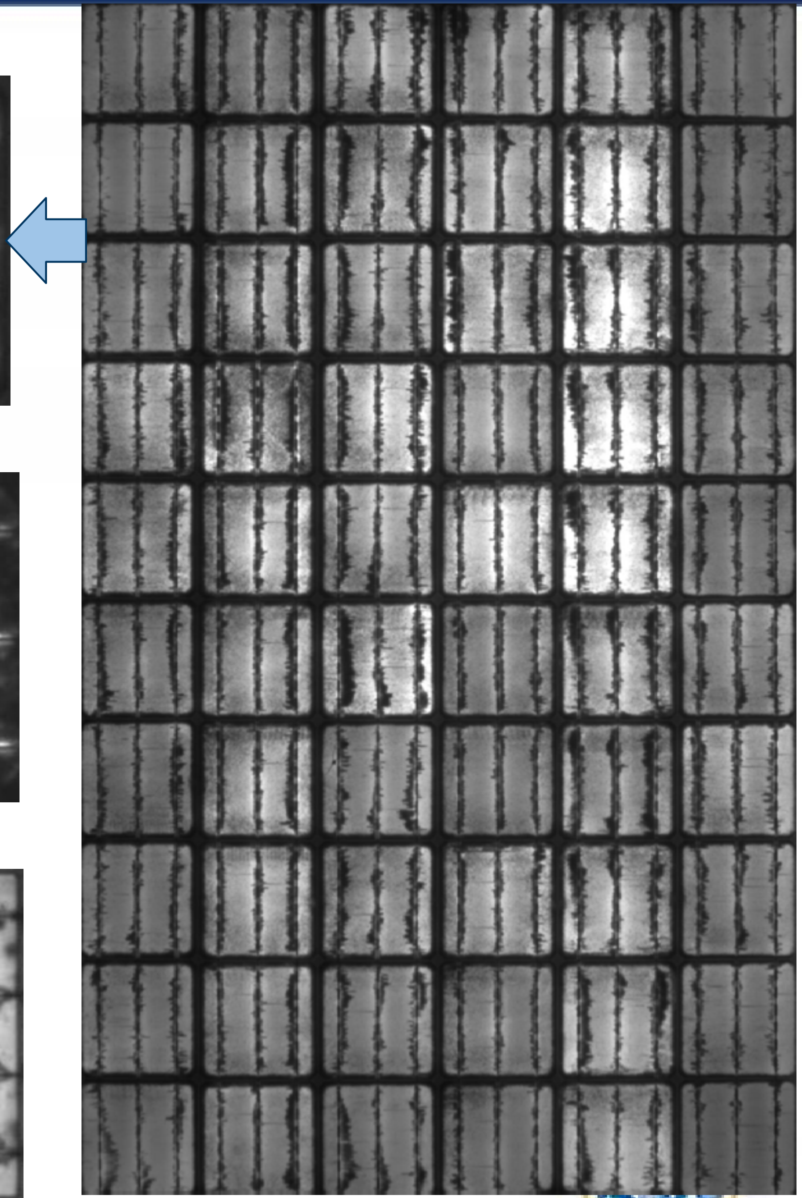












- **Results in single cell images**
  - Resembles face recognition problem

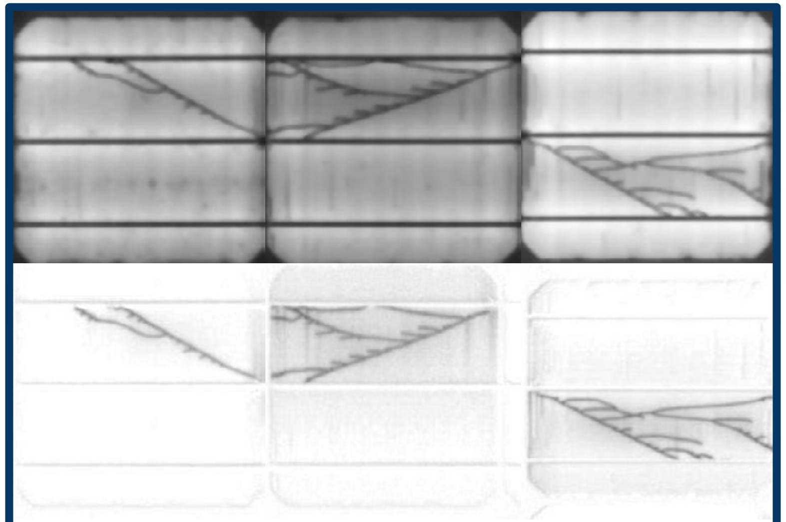
[1] J. S. Fada, M. A. Hossain, J. L. Braid, S. Yang, T. J. Peshek, and R. H. French, "Electroluminescent Image Processing and Cell Degradation Type Classification via Computer Vision and Statistical Learning Methodologies," Washington, D.C., 2017.
[2] Ulrike Jahn et al., "IEA-PVPS Task 13, SubTask 3.3: Review on Infrared and Electroluminescence Imaging for PV Field Applications," International Energy Agency, 2017.

Baseline Image Subtraction

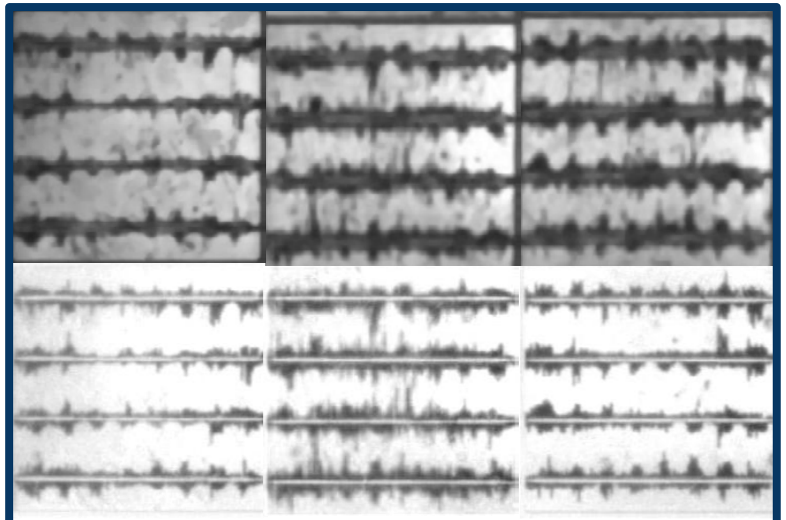
Subtract the degraded image
- from baseline(initial) image

Helps in tracing the defects
- in each c-Si PV cell

Improves the performance
- of clustering algorithm



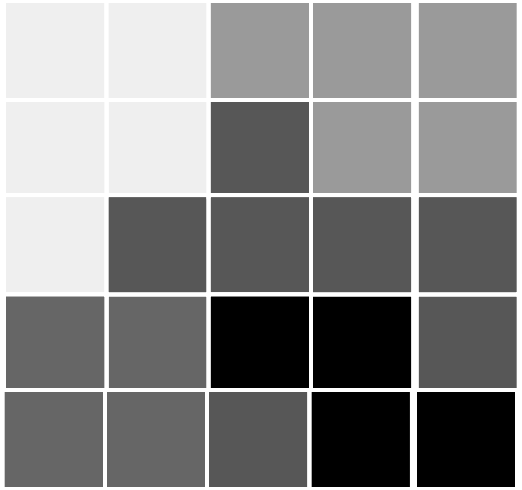
Cracked Cells



Corroded Cells

Feature Extraction

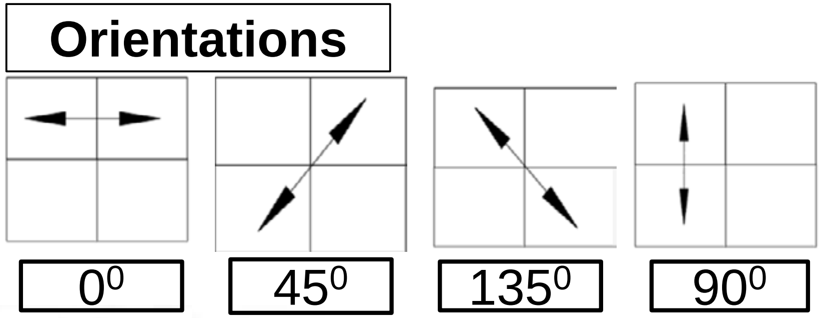
- **Calculation Method**



2-bit Gray Level Image

3	3	2	2	2
3	3	1	2	2
3	1	1	1	1
1	1	0	0	1
1	1	1	0	0

Matrix Representation



[1]Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, *3*(6), 610-621.

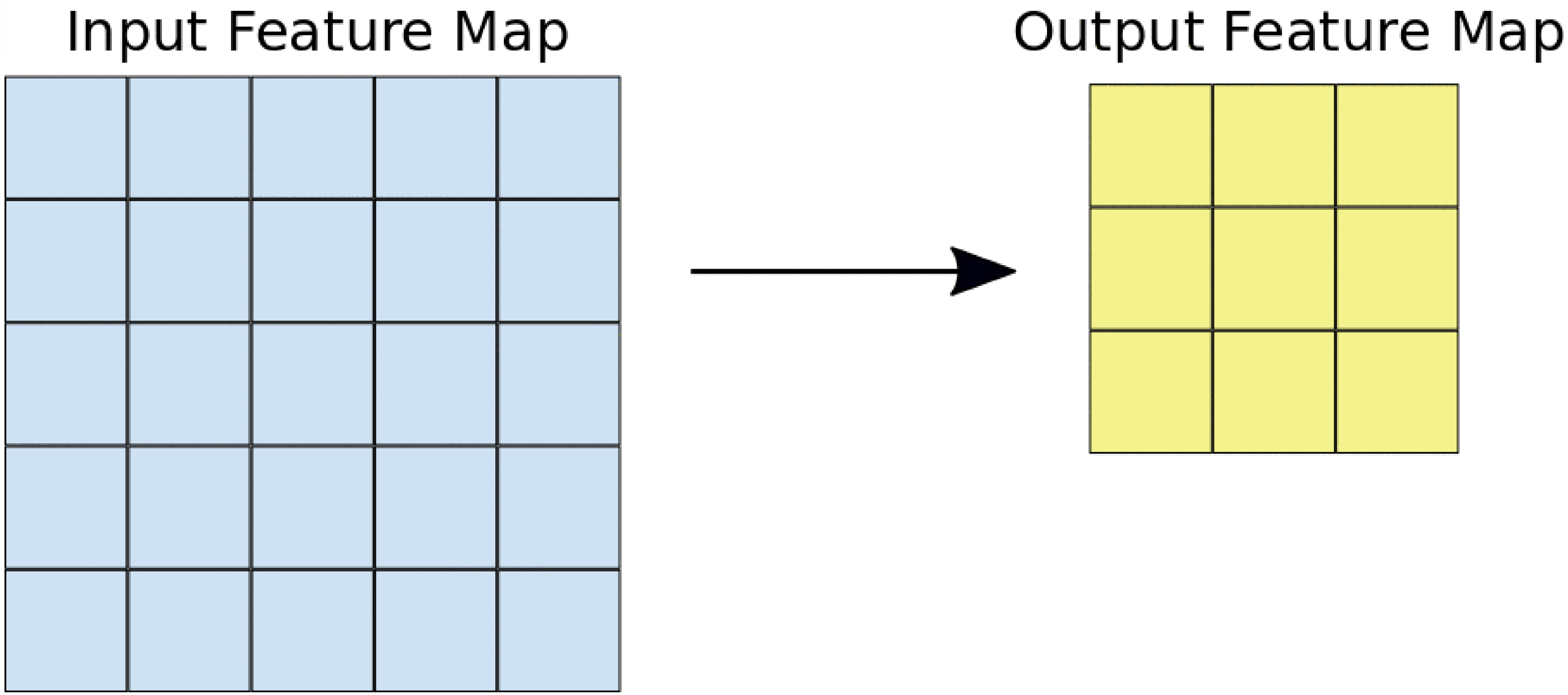
Features	Description
Energy/ Angular Second Moment (ASM), $\sum_{i,j}^{N_g} p(i,j)^2$ N_g is the no. of gray scale levels	ASM is also called energy, it's value is 1 for constant image & range [0,1]. $p(i,j)$ is $(i,j)^{th}$ value in a normalized gray-tone
Contrast, $\sum_{i,j}^{N_g} (i-j)^2 p(i,j)$	Measure of contrast in intensity between adjacent pixels
Correlation, $\sum_{i,j} \frac{(i.j)p(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y}$ μ and σ are mean and std. deviation of p_x p_y	Measure of how correlated a pixel to it's neighbor in an image. p_x, p_y are marginal probabilities
Sum of square, $\sum_{i,j}(i-\mu)^2 p(i,j)$	squared distance from the mean pixel intensity
Inverse Difference Moment $\sum_{i,j} \frac{p(i,j)}{1+(i-j)^2}$	Also called Homogeneity, measures values by the inverse of the contrast weight
Sum Average, $\sum_{k=2}^{2N_g} kp_{(i+j)}(k)$	sum of the average values of whole image
Sum Average, $\sum_{k=2}^{2N_g} kp_{(i+j)}(k)$	sum of the average values of whole image
Sum Variance, $\sum_{k=2}^{2N}(k-\mu_{x+y})^2 p_{(x+y)}(k)$	sum of the variance values of whole image. N is number of distinct gray level
Entropy, $-\sum_{i=1}^{N}\sum_{j=1}^{N} P(i,j) log P(i,j)$	measure of randomness that can be used to characterize the texture of the input image.
Sum Entropy, $-\sum_{k=2}^{2n} p_{(x+y)}(k) log(p_{x+y}(k))$	sum of the entropy.
Difference Variance, $\sum_{k=2}^{2n}(k-\mu_{x-y})^2 p_{(x-y)}(k)$	Difference between the variance.
Difference Entropy, $-\sum_{k=2}^{2n} p_{(x-y)}(k) log(p_{x-y}(k))$	Difference between the entropy.
Information Measure of Correlation I, $\frac{HXY-HXY1}{max(HX,HY)}$	HX and HY are entropies of $P_x and P_y$ $HXY = -\sum_i \sum_j p(i,j) log(p(i,j))$ $HXY1 = -\sum_i \sum_j p(i,j) log(p_x(i)p_y(j))$
Information Measure of Correlation II, $(1-exp[-2.0(HXY2-HXY)])^{\frac{1}{2}}$	$HXY2 = -\sum_i \sum_j p_x(i)p_y(j) log(p_x(i)p_y(j))$
Maximum Correlation coefficient $Q(i,j) = \sum \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}$	Method for identifying shape or numerical parameter for probability distribution

Calculating the GLCM

Gray Level Co-occurrence Matrices (GLCM)

1	0	-1
1	0	-1
1	0	-1

Filter image

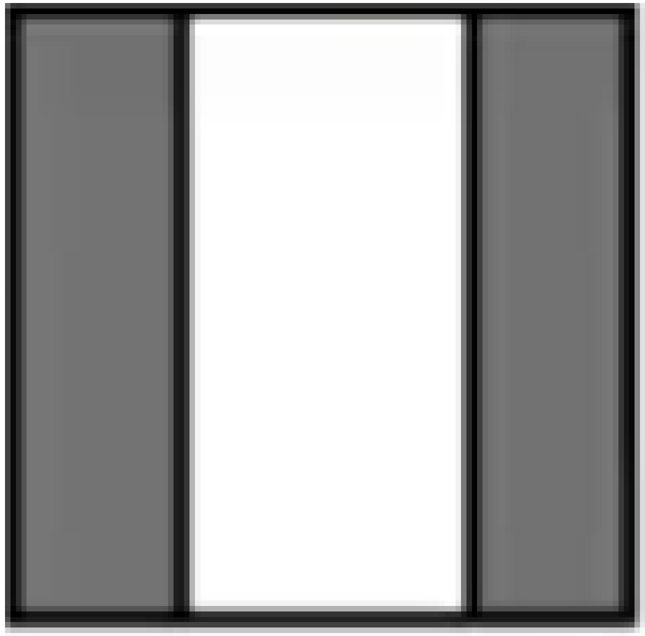


Convolve the filter image with the input image

MATERIALS DATA SCIENCE RELY

Result of applying image filter, and Result

0	30	30	0
0	30	30	0
0	30	30	0
0	30	30	0



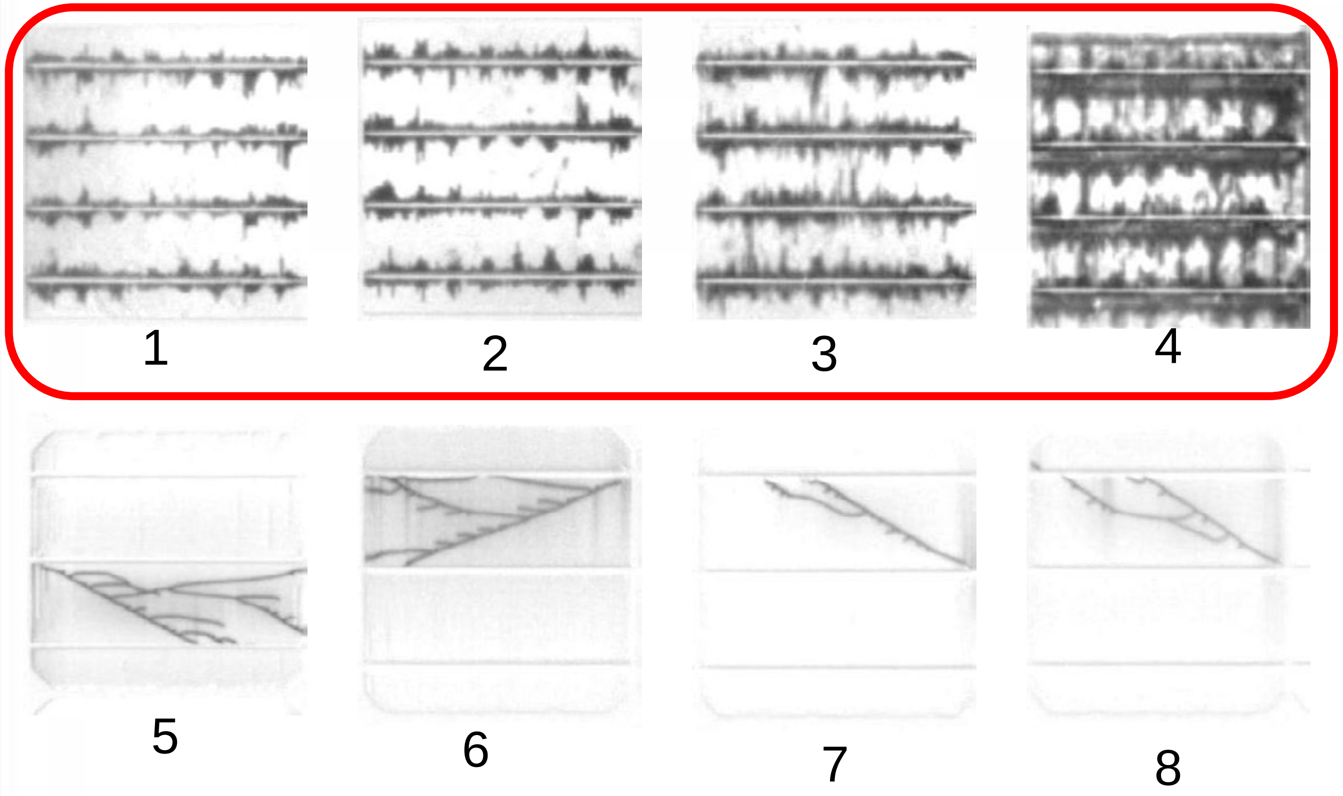
This filter, locates a vertical bar in the image

- A set of GLCM features identifies diverse patterns

MATERIALS DATA SCIENCE RELY

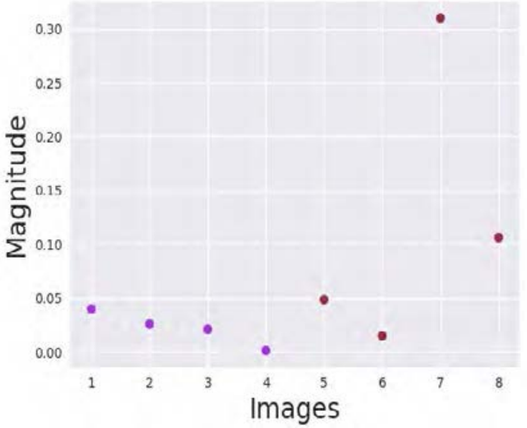
Feature calculations on sample images

8 Sample Images

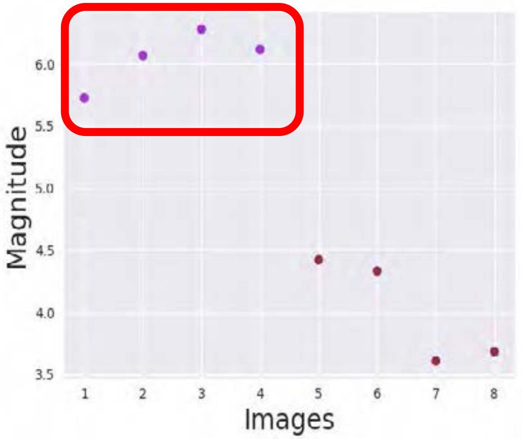


4 Haralick Features of those 8 images

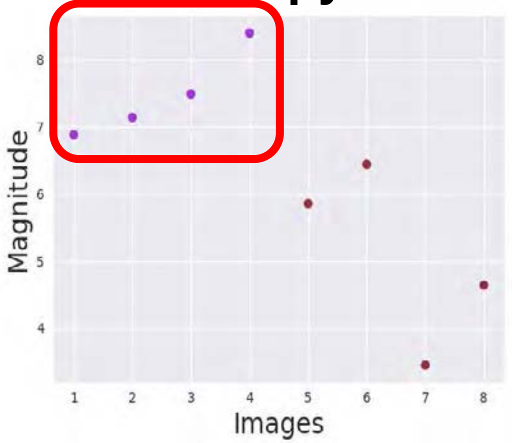
Angular 2nd Moment



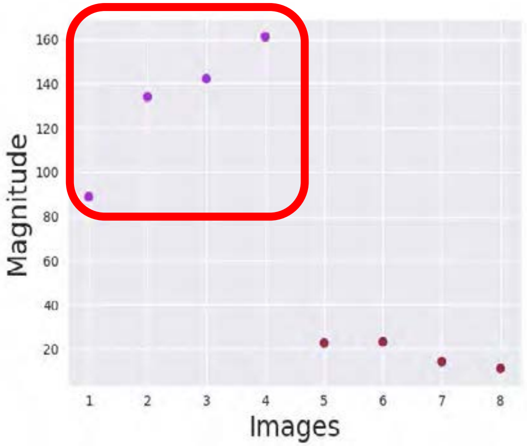
Contrast



Entropy



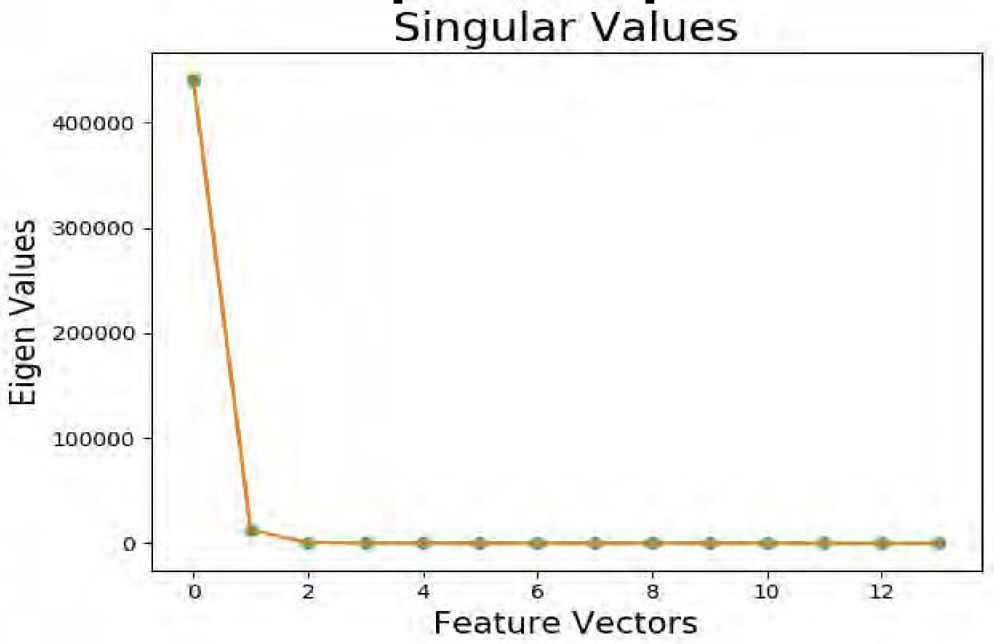
Information Measure

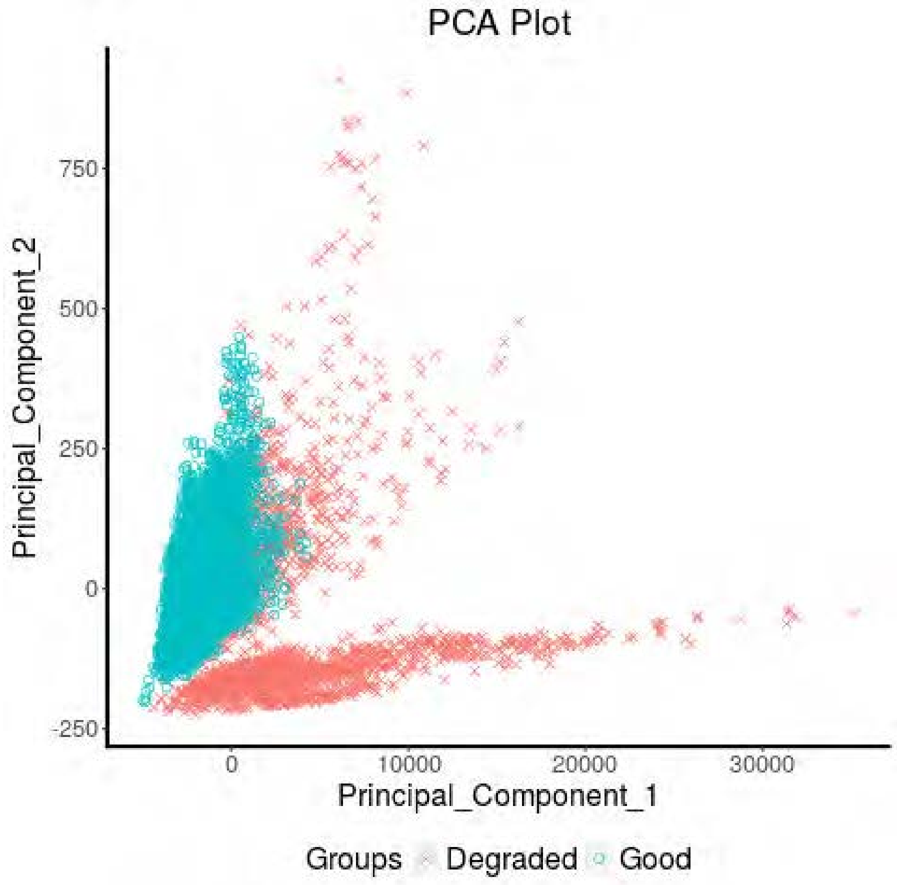


MATERIALS DATA SCIENCE
RELY

Unsupervised Learning Algorithms

- **Singular Value Decomposition on Extracted Features**
- **Principal Component Analysis**





- **Agglomerative Hierarchical Clustering**
  - Two principal component
  - Euclidean distance as dissimilarity measure

MATERIALS DATA SCIENCE RELY

Utility for MDS-Rely

New Product Development

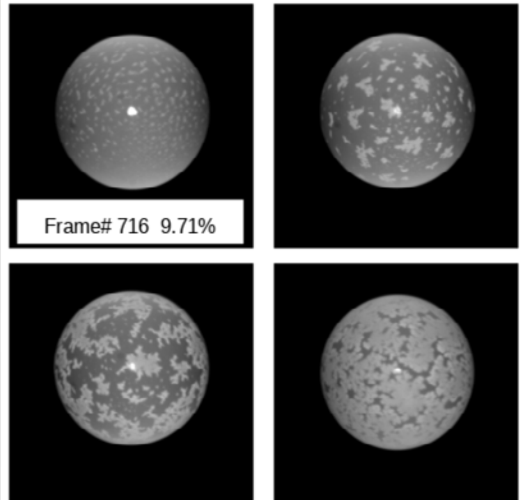
- Accelerate cycles of learning

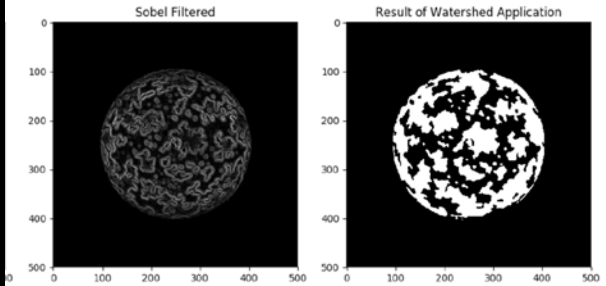
Manufacturing Industry:

- Quantitative analyses
- Production assembly line

Operations & Management:

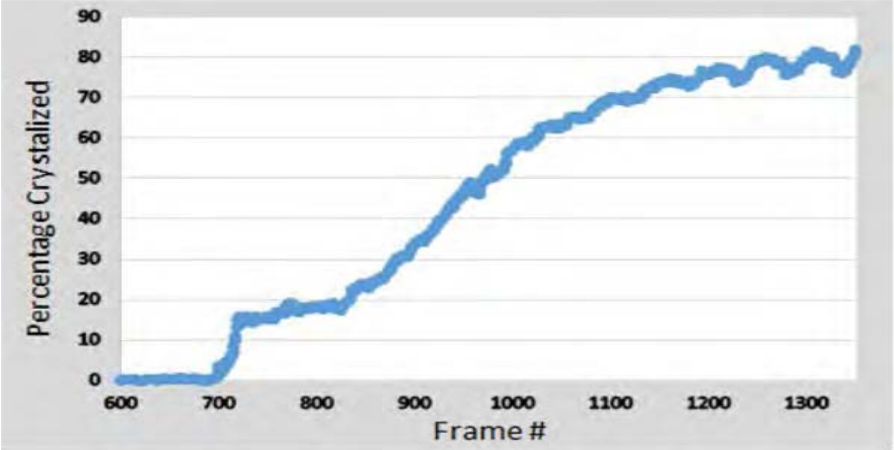
- Automate materials analysis





Characterization of AlN Growth

- **From Al-Ni alloy**
- **Using 400,000 video images**



MATERIALS DATA SCIENCE RELY

Thanks

MATERIALS DATA SCIENCE RELY

Proposal Timeline

Tasks	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Feature Extraction	15%	20%	20%	25%	20%			
Identification of critical features				10%	20%	40%	30%	
Application of Unsupervised Image Machine Learning							40%	60%

MDS-Rely Project Proposal
(RS Area A - Thrust Area 2 - Proposal Number 3)
(September 2019)

Project Title: Network Algorithms for Predictive Modeling of Photo-Voltaic Systems	
Principal Investigator(s): Mehmet Koyuturk	**Researcher: Ahmad M. Karimi**
New Project: XX Renewal: **Term: X1 year, 2 years**	**Start Date: June 2020**

Thrust Area: 2. Matls. Data Science

Objective:

Energy and sustainability are two challenges facing the United States, its citizens and the world. Photovoltaic (PV) power plants supply power for utilities by converting light directly into electricity. PV plants offer a promising source of energy in terms of its sustainability. Since large-scale deployment of PV systems has been realized very recently, the knowledge on their behavior (e.g., fluctuations in performance due to changes in solar irradiation, degradation as a function of time, location, climate, materials, and combinations of these) is based on lab experiments, which are limited in scope and scale. The Solar Durability and Lifetime Extension (SDLE) Research Center at Case Western Reserve University, has ingested into its Hadoop/Hbase/Spark infrastructure (Energy-CRADLE) 20 years worth of large-scale and high-resolution "real world" data from thousands of PV plant sites and systems distributed across the globe. This data consists of minute interval time-series data-streams of power, weather and solar irradiance, encompassing 800 sites, 6000 systems, 60 PV module brands, 35 DC/AC inverters constituting 3.4 GW, or 2%, of the global PV power generation assets.

This project will effectively mine these datas using a spatio-temporal graph-based network modeling approach to inform increasing PV penetration. Analyzing the spatio-temporal electricity production of thousands of PV systems we will develop simplified probabilistic predictive capacity and generation models. Retrospective analysis across the full time range of the PV system/inverter datasets will determine the probability density function (pdf) of the mean PV power (Y) and its variance, and identify the predictors (variables) that most strongly influence the answer and its uncertainty. We then use these detailed probabilistic (net conditional expectation) results for each specific question, and fit probabilistic non-time series models for the most likely (mean) answer.

This approach renders the data analysis tractable and robust to noise, uncertainty, and the curse of dimensionality. These new models, provided as open-source algorithms, will have higher accuracy than existing models and will enable higher PV penetration levels with stable and reliable power system operation and performance.

Standards Used:

Code Developed: We will use Python scripting to generate an open source code. We will port to MATLAB to integrate the developed algorithms into OpenDSS.

Datasets Produced:

Spatio-temporal and correlation networks representing the relationships between 600 PV systems in 800 sites.

Background:

A limiting factor in optimizing the cost and yield of PV plants is their degradation over time. While producers continue to offer 25-year warranties, degradation-induced failures are commonly observed for new and promising PV cells. On the other hand, observations on real-world applications suggest that the efficiency of most PV cells does not drop below 75% for up to 50 years [1]. These observations suggest that the dynamic range for the degradation of PV systems is quite wide, and with improved understanding of the factors and stressors that underlie degradation, this dynamic range can be exploited to significantly improve the lifetime and reliability of PV plants up to 50 year lifetimes.

Degradation of energy materials distinguishes itself in that it evolves over long time-frames due to a multitude of distinct, complex, and interacting mechanisms that can lead to a variety of slow and/or rare events that eventually cause failure. There are severe knowledge gaps in identifying, modeling, and reliably predicting the mesoscopic evolution that produce degradation, and in establishing an effective monitoring system of the evolving process of degradation over the relevant timescales to prevent failures (especially catastrophic failures) [2]. It is essential to connect the mechanistic degradation pathways and their temporal evolution at the mesoscale so as to enable the identification of improved and longer-lived energy materials in real life. Hence, degradation science examines degradation of a material or system, guided by real-world or realistic outcomes, whose fundamentals include modeling, monitoring, and prediction of a degradation process, as well as intervention, feature selection, and optimization aimed at improvement of materials and reduction of system failures. A new interdisciplinary approach to degradation science calls for the involvement of materials science, physics, chemistry, statistics, computer science, engineering, and energy industries in the investigation of real-world degradation of energy material over their lifetimes.

There have been many studies of the PV power generation impact on the performance of power systems. For example the recent work by the VADER project has taken a data-driven modeling approach to real time monitoring of distributed energy resources, but do not consider long term PV degradation or the needs for long term planning models [3]. One advantageous aspect of PVs in the power system, from a power generation and its temporal variability perspective, is the effect of spatio-temporal averaging of distributed PV systems in a specific location area. As the area increases, and the number of PV systems in that area increases, the variability of the power generation, and ramp rates of power generation are reduced due to averaging, or smoothing. The impact of solar variability is at its most extreme during an eclipse, and for example in the European 2016 and the U. S. 2017 total eclipse had no negative impact on the power system [4]. Using a probabilistic approach to PV power capacity and generation modeling, means that understanding the uncertainties of the models and their forecasts plays an important role, and the major contributors to uncertainty propagation have been studied for long term yield prediction and planning [5] and on the financial modeling [6] of PV systems and these uncertainty analyses will inform the proposed work.

Project Tasks:

Task 1: Calculate expected power from retrospective spatio-temporal PV system datasets.

We will construct spatio-temporal graph networks capturing different controlling factors for PV capacity and generation models that are intuitive, interpretable, and tractable by humans. These spatio-temporal graph-based network models are **spatially aggregative**, i.e., they can study PV systems in any location area defined by a GIS shapefile, defined for a state, a utility's distribution service area, substation, an individual feeder or a service transformer to determine "net" power output of the PV inverters of interest. These graph models must **forecast power production for multiple different time instances**: (1) how much PV power will be generated in the next 15 minutes, hour, 4 hours, or next day (for power system operation), and (2) how much power will PV be providing maximally (efficient capacity) and at the peak load for every year during a15 years period (for power systems planning).

The retrospective cohort study of the PV systems involve three types of graphical network models: spatial graph (s-graph) network $G^s(V, E)$, spatial-temporal graph (st-graph) network $G^{st}(V, E)$ (st-graph, and correlational graph (c-graph) networks $G^c(V, E)$, where V denotes the set of vertices (or nodes) representing PV systems and E represents either euclidean distance (for s-graphs or st-graphs) or correlations (for c-graphs).

A simple additive model is applied to calculate the power ($\boldsymbol{Y_{i,j}}$) of a given PV system (i) at a time instance (t_j) by power generation under ideal irradiance ($\boldsymbol{Y_{ij}^{ideal}}$), power loss caused by cloud cover ($\boldsymbol{Y_{ij}^{cloud}}$), power loss caused by degradation ($\boldsymbol{Y_{ij}^{degr.}}$) and power loss caused by inverter saturation ($\boldsymbol{Y_{ij}^{sat.}}$), where these four additive terms are assumed to be independently distributed:

$$\begin{aligned} \boldsymbol{Y_{ij}} &= f(\boldsymbol{Y}_{ij}^{ideal}, \boldsymbol{Y}_{ij}^{cloud}, \boldsymbol{Y}_{ij}^{degr.}, \boldsymbol{Y}_{ij}^{sat.} | P_o = p_o, L = l, T = t_j, \ldots) \\ &= (\boldsymbol{Y}_{ij}^{ideal} - \boldsymbol{Y}_{ij}^{cloud} - \boldsymbol{Y}_{ij}^{degr.} - \boldsymbol{Y}_{ij}^{sat.} | P_o = p_o, L = l, T = t_j, \ldots) \end{aligned} \quad \textbf{Eqn. 1}$$

Each independent component of the additive model is associated with an st-graph which is treated as a layer of the st-graph model for $\boldsymbol{Y}_{i,j}$. The simple additive model for the output of each graph layer gives us the conditional probability density function of power $Y_{i,j}$ (i $\in$){1,..., n} and j T)
calculated in **Eqn. 1** for a set of PV systems under given values of different parameters such as time instance t_j , location *l, and* nameplate power p_o. Here, n is the number of PV systems in the region under study and $\boldsymbol{T}$ is the set of timestamps for each power time-series. Timestamps $\boldsymbol{T}$ start from t_s and end at t_e with granularity □$_t$.

We represent the network graph for power output from a PV system under ideal solar conditions as $\boldsymbol{Y_{ij}^{ideal}}$. $\boldsymbol{Y}_{ij}^{ideal} = \eta * p_o$ where $\eta = I_i/I_{sun}$, I_i is the ideal geometric solar irradiance, I_{sun} solar irradiance at 1 sun, and p_o is the nameplate power such that $0 \leq \eta \lesssim 1$. Similarly, we represent the network graph for power loss due to the cloud cover as $\boldsymbol{Y_{ij}^{cloud}}$. $Y_{ij}^{cloud} = \zeta * p_o$ where $\zeta = (I_i - I_c)/I_i$ and I_c is the value of solar irradiance under the cloud cover such that $0 \leq$ □ ≤ 1. We represent the network graph for the power loss in PV due to the degradation of the modules, $\boldsymbol{Y_{ij}^{degr.}}$. $Y_{ij}^{degr.} = \nu * p_o$ where, □ is the accumulated change in the PV system due to degradation and can be calculated using either a constant or a piecewise constant degradation rate model. Constant degradation is modeled as □ =□$_o$ N_s , where □$_o$ is a ROC calculated using historical data, and N_s is the PV system's age. Piecewise constant degradation is modeled as follows:

$$\nu = \begin{cases} \beta_0 * N_s; & (N_s \leq \tau) \\ \beta_0 * \tau + \beta_1 * (N_s - \tau); & (N_s > \tau) \end{cases}; \quad ROC = \begin{cases} \beta_0 & ; (N_s \leq \tau) \\ \beta_1 & ; (N_s > \tau) \end{cases}$$

where τ is a changepoint determined by modeling using the historical data of when the PV system starts to degrade at a different rate. The values of β_0 and β_1 are influenced by the technology of the PV modules such as thin-film, mono-Si and multi-Si. The piecewise constant degradation rate model is a necessary modeling option since ROC is influenced by the technology and system age and changepoints are often observed over system lifetime. **We** represent the network graph for the loss in power due to inverter saturation, $\boldsymbol{Y_{ij}^{sat.}}$. $Y_{ij}^{sat.} = (I_c - I_{sat.}) * p_o / I_{sun}$ when $\boldsymbol{Y_{ij}^{ideal}}$-$\boldsymbol{Y_{ij}^{cloud}}$-$\boldsymbol{Y_{ij}^{degr.}}$>$\boldsymbol{Y_{ij}^{sat.}}$ or *0 otherwise*, where $I_{sat.}$ is the solar irradiance at which inverter saturates.

We estimate the net generated power by all the PV systems at retrospective time (t_j) and location area (*l*) using the additive function $Y_{.j} = \sum_{i=1}^{n} Y_{i.j}$, where *n* is the number of PV systems at the given time and location. Note that we define *n* as a function of time to account for the growth in the number of PV systems and ensure that inactive systems do not contribute to the total generated power. Similarly, we can calculate $\boldsymbol{Y_{.j}^{ideal}}$, $\boldsymbol{Y_{.j}^{cloud}}$, $\boldsymbol{Y_{.j}^{degr.}}$, $\boldsymbol{Y_{.j}^{sat.}}$ and create the PV power probability density function of each component of the additive model and the generated net power at each location area (*l*) and retrospective time (t_j). Next, median, mean, and higher order moments of the generated pdf, such as the *variance*, *skewness* and *kurtosis*, are calculated and stored in their associated s- and c-graphs. For example, the expected net generated power at Δt_f hours ahead ($t_f = t_p + \Delta t_f$) for location area *l*, and a current time t_p is $E(Y_{.t_f}) = \Sigma_{j=1}^{m} Y_{.j}/m$, where m is the number of the time stamps equal to t_f over $\boldsymbol{T}$.

Task 2: Develop a long-term probabilistic spatial PV capacity model.

PV efficient capacity is defined as power generated under an ideal solar source (for a given location and time), while still considering the long term PV system degradation (for different PV technologies and ages). We we will develop the spatial graph network model $G^s(V,E)$ for describing the distribution of expected efficient capacity for each PV system and obtain the expected net efficient capacity and its uncertainty in the long term such as for Δt_f = 5 years ahead,10 or 15 years ahead. In $G^s(V,E)$, the nodes (*V*) represent the expected efficient capacity of each single PV system, and the edges (*E*) represent the nodes located in the location area we want to study as defined by the GIS shapefile.

We define $\boldsymbol{Y_{ij}}$ as the efficient capacity of a given PV system (*i*) at a time instance (t_j) and calculate it by **Eqn. 1** only considering the $\boldsymbol{Y_{ij}^{ideal}}$ and $\boldsymbol{Y_{ij}^{degr.}}$ terms. Using the probability density functions calculated for the first task, we can obtain the expected net efficient capacity of a region at a given future time instance t_f using **Eqn. 2** and the spatial graph model of expected efficient capacity of each PV systems $E(Y_{i,t_f})$ **(Fig. 6)**

$$E(Y_{.t_f}) = \Sigma_{j=1}^{m}(Y_{.j}^{ideal} - Y_{.j}^{degr.})/m = E(Y_{.t_f}^{ideal}) - E(Y_{.t_f}^{degr.}) \qquad \textbf{Eqn. 2.}$$

The model can be expanded to calculate expected efficient capacity at peak load by adding power loss of PVs due to irradiance reduction because the peak load may not happen at the same time as the maximum PV production.

Task 3: Develop a short-term probabilistic spatial PV power generation model

The goal of the short-term power generation model is to estimate the expected power output $E(\boldsymbol{Y}_{.t_f})$ for a future time instance t_f in a location area covered by the shapefile l. We use the correlational graph to include similar PV systems in the study because adding more data improves the accuracy of expected value calculations. Nodes in the network are pdfs of the PV system calculated as part of Task 1 and if the dissimilarity measure among the nodes are less than a threshold value, then edges are drawn between them showing association among those PVs. . For the correlation graph of cloud cover, we consider the cluster with the largest number of the PV systems in the location l and define n_c as the total number of the PVs in that cluster. . For n_c, the net power loss due to cloudiness is $E(\boldsymbol{Y}^{c\_cloud}_{.t_f})$, power under ideal solar source is $E(\boldsymbol{Y}^{c\_ideal}_{.t_f})$ and λ is the ratio of net nameplate power (p_n) for PV systems in region l over net nameplate power (p_c) for n_c ($\lambda = p_n/p_c$). Likewise, for the correlation graph of degradation loss, the set of all the PV systems that belong to the cluster with majority of PV systems from region l be n_d. For n_d, $E(\boldsymbol{Y}^{c\_degr.}_{.t_f})$ is the net expected power loss due to degradation and □ is the ratio of net nameplate power p_n over net nameplate power (p_d) for n_d PVs (□ $= p_n/p_d$). Similarly, n_s is the number of PV systems that will lie in the cluster having majority of PV systems from region l, p_s is the net nameplate power for n_s systems, $E(\boldsymbol{Y}^{c\_sat.}_{.t_f})$ is the net expected power loss due to inverter saturation and $\omega = p_n/p_s$. The equation to calculate net expected power for region l is:

$$E(\boldsymbol{Y}_{.t_f}) = E(\boldsymbol{Y}^{ideal}_{.t_f}) - \lambda(E(\boldsymbol{Y}^{c\_cloud}_{.t_f})) - \gamma(E(\boldsymbol{Y}^{c\_degr.}_{.t_f})) - \omega E(\boldsymbol{Y}^{c\_sat.}_{.t_f}) \quad \textbf{Eqn. 3}.$$

In **Eqn. 3**, all the expected value terms on the RHS of the equation can be calculated from the pdfs computed as part of Task 1.

Benefits to Members:

Significant investments have been made to develop power and weather time-series data sets, to perform analysis of PV system degradation, to develop ROC analysis, and to analyze impact of PV systems. The major impact of this investment will be to bring together these efforts to develop more accurate PV models for any spatio-temporal scale that will reduce uncertainty and eliminate need for conservative interconnection assumptions. The new technology required to do this is graph-based network models. The innovation in our concept lies in network modeling using spatio-temporal graphs, to identify rank-ordered predictors that capture the controlling variables in PV generation. Modeling PV system performance with these dominant predictors and conditioning the models on these predictors enable us to represent BTM PV systems probabilistically in areas with limited PV observability. The proposed proposed PV model will decrease interconnection costs, provide a faster and more accurate interconnection process, increase observability, increase reliability and resilience and allow for increased penetrations of PV. Since the model will be integrated with tools utilities are already using, it will be no-cost option that can significantly reduce burden to the utilities during planning process, interconnection studies or operation.

Timeline :

Tasks	Q1	Q2	Q3	Q4
Task 1				
Task 2				
Task 3				

References

[1] Jordan D, Kurtz S, VanSant K, Newmiller J, "Compendium of photovoltaic degradation rates," *Progress in Photovoltaics*, Vol 24, Issue 7, 2016.

[2] Hemminger, John, et al., "From Quanta to the Continuum: Opportunities for Mesoscale Science," U.S. Department of Energy, 2012.

[3] E.C. Kara, C.M. Roberts, M. Tabone, L. Alvarez, D.S. Callaway, E.M. Stewart, Disaggregating solar generation from feeder-level measurements, Sustainable Energy, Grids and Networks. 13 (2018) 112–121.

[4] Y. Saint-Drenan, R. Fritz, D. Jost, How an energy supply system with a high PV share handled a solar eclipse, 2016, IEA-PVPS Task 13, 2017.

[5] C. Reise, B. Müller, D. Moser, G. Belluardo, P. Ingenhoven, A. Driesse, G. Razongles, M. Richter, Uncertainties in PV System Yield Predictions and Assessments, 2018.

[6] M. Richter, et al., Technical Assumptions Used in PV Financial Models - Review of Current Practices and Recommendations, IEA-PVPS Task 13, 2017.

Mehmet Koyuturk

Department of Computer and Data Sciences

Case Western Reserve University

Proposed Project Duration: 1 year

Industrial Relevance and Novelty

- Significant investments have been made to develop power and weather time-series data sets to analyze the impact of Photovoltaic (PV) systems
  - Network-based models and algorithms can robustly extract hidden relationships in these datasets
- We will develop more accurate PV models for any spatio-temporal scale
  - To reduce uncertainty
  - To eliminate need for conservative interconnection assumptions
- Predictive models will represent PV systems probabilistically in areas with limited PV observability
  - Reduced interconnection costs
  - Faster and more accurate interconnection process
  - Increased observability, reliability, and resilience

Proposal Objectives

Objective 1: Calculate expected power from retrospective spatio-temporal PV system datasets

- Key idea: Use data from neighboring and/or similar plants through network modeling

Objective 2: Develop a long-term probabilistic spatial PV capacity model

- Degradation

Objective 3: Develop a short-term probabilistic spatial PV power generation model

- Weather, cloudiness, weather, time of day

MATERIALS DATA SCIENCE RELY

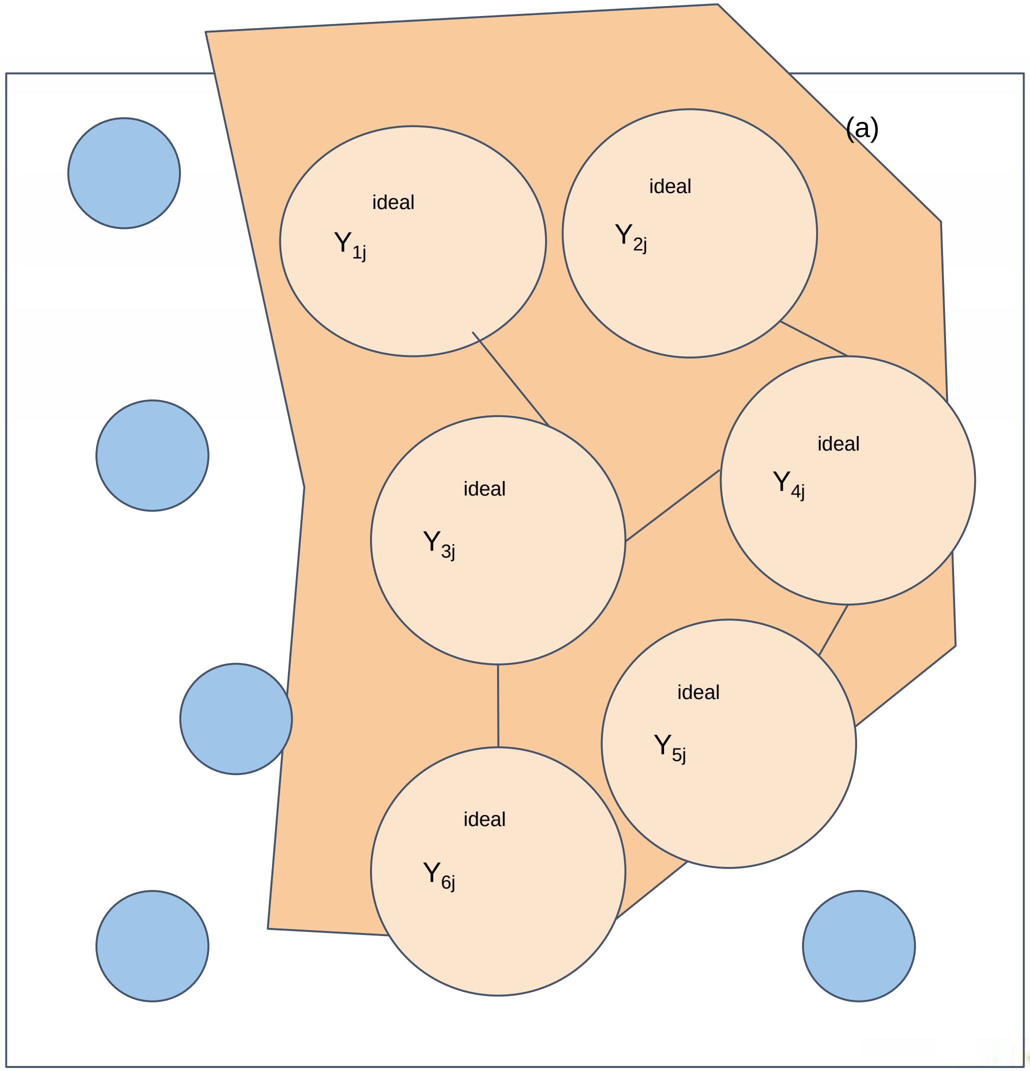
Proposal Deliverables

- A suite of algorithms and tools for predictive modeling of PV systems
  - Multiple levels of user control
  - Can use existing database or add data and retrain
  - Can query for the output of a (set of) PV system(s) at a given interval
  - Intuitive user interface
  - Real-time query performance

Proposal Timeline

Task / Milestone	Year 1			
	Q1	Q2	Q3	Q4
Objective 1				
Objective 2				
Objective 3				

Proposal Work Description



This is a Map

- holding the time series data

We define our region of interest

- with a GIS shapefile

This can be different areas and regions

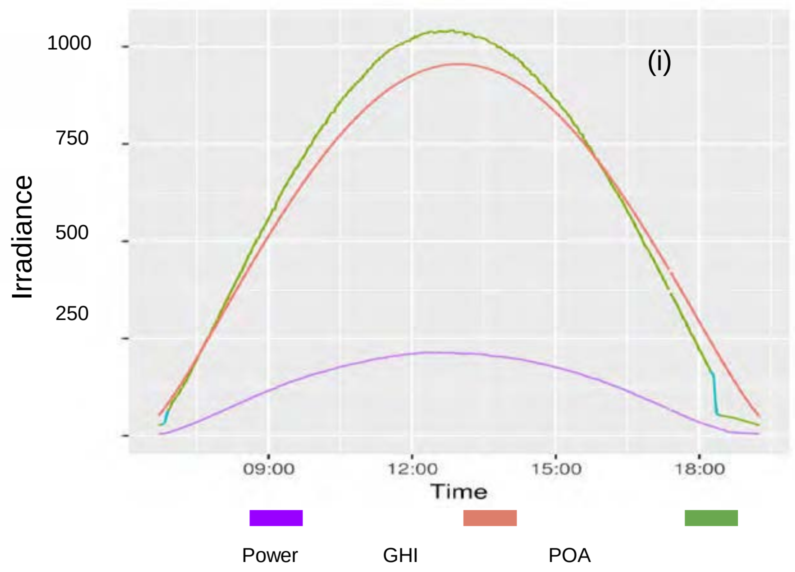
- A Feeder in the Distribution Network
- A Substation
- A Service Area
- A State or Country

We want to be able to rapidly Forecast

- For any region of interest

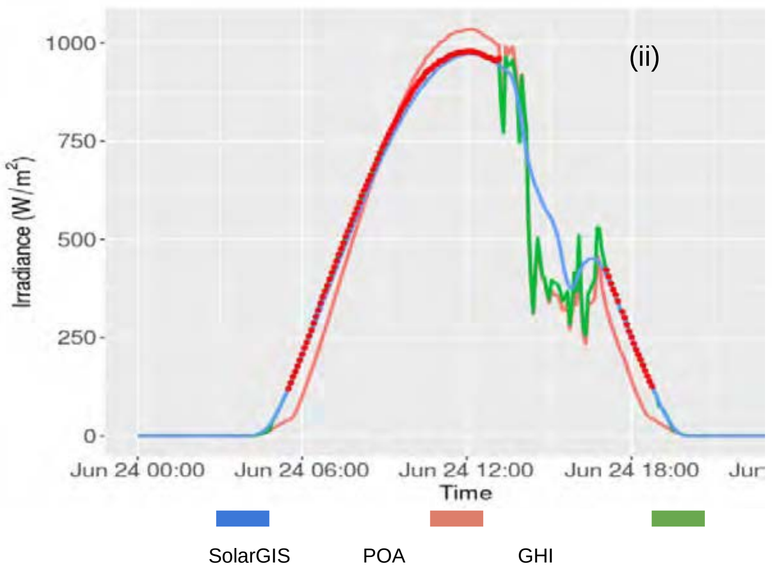
MATERIALS DATA SCIENCE RELY

Typical Temporal Data Stream for a Single Node



Sample time-series for a typical datastream

- output power
- Global Horizontal Irradiance (GHI)
- Plane of Array Irradiance (POA)



Data processing for clear sky points

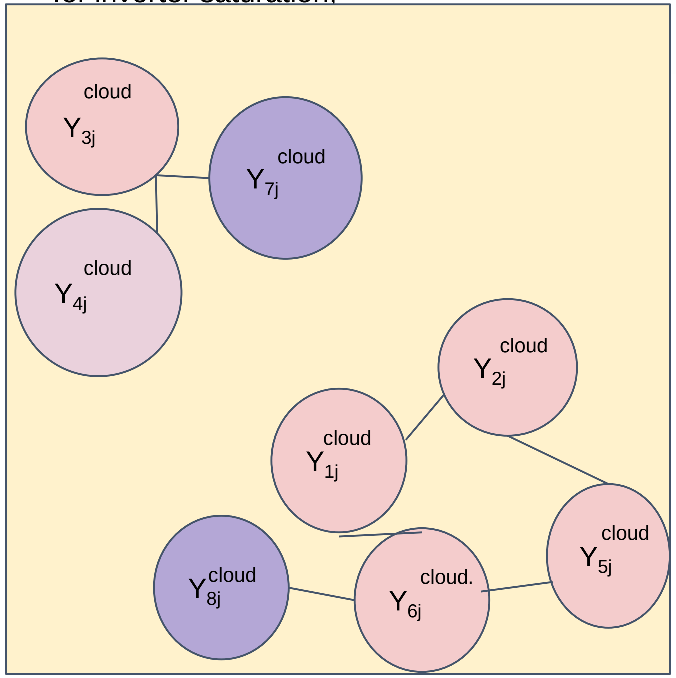
- SolarGIS
- POA
- GHI
- Red points are clearsky part of day

MATERIALS DATA SCIENCE RELY

Spatial and Correlational Networks

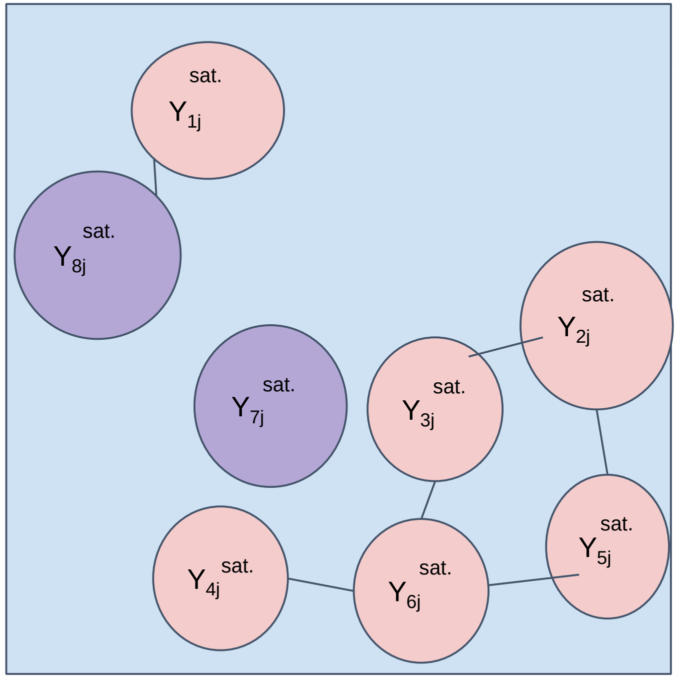
Correlational network graph for

- power loss due to cloud cover and
- for inverter saturation.



Similarly correlational networks

- Can be constructed for other variables



MATERIALS DATA SCIENCE RELY

General Formulation

Using the network graphs (a,b,c & d),

- Calculate the power output
  - at time instance j
  - for PV system i using equation.

$$Y_{i,j} = Y_{i,j}^{ideal} - Y_{i,j}^{cloud} - Y_{i,j}^{degr.} - Y_{i,j}^{sat.}$$

Net output power for a set of systems

- at instance j
- is calculated using graph e

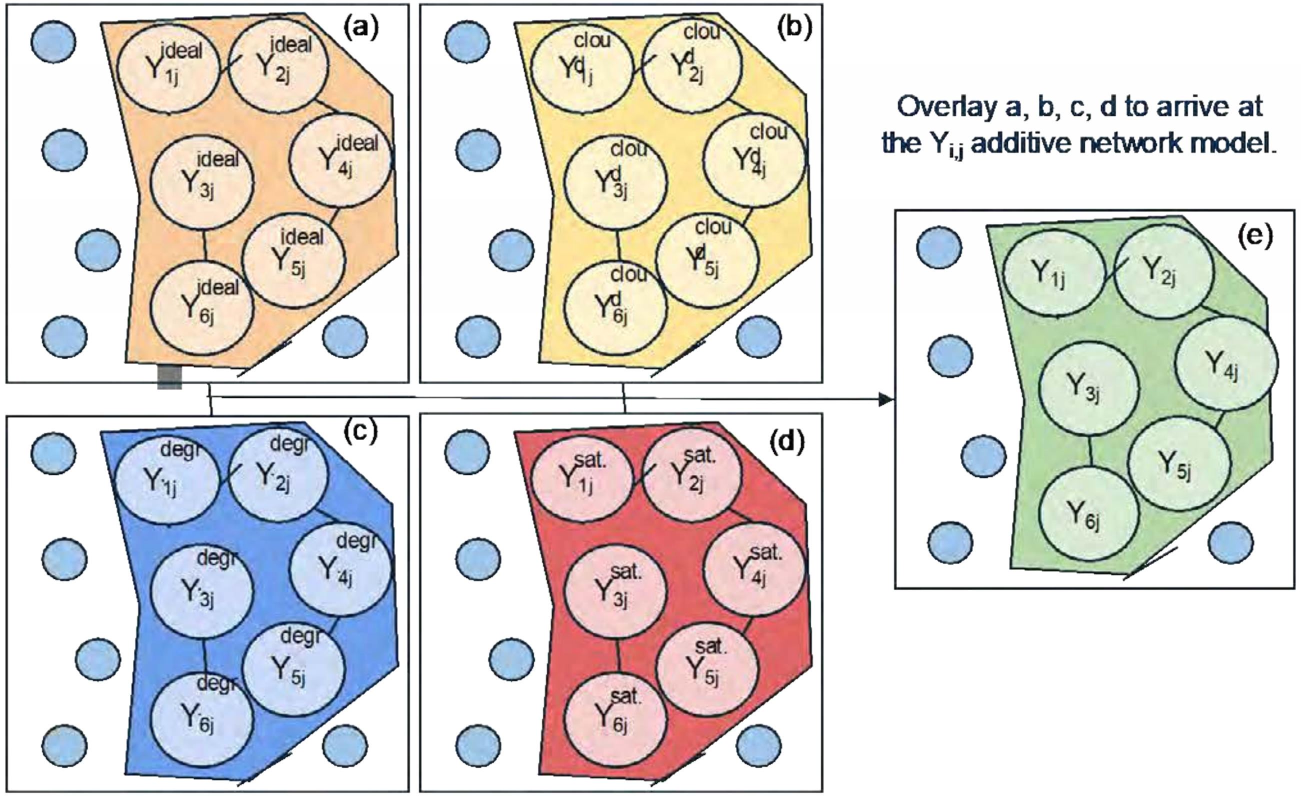
Using Figure (e) and (f) we find

- similar systems outside the given region

Study additional PVs

- to improve the prediction
- of expected output power:

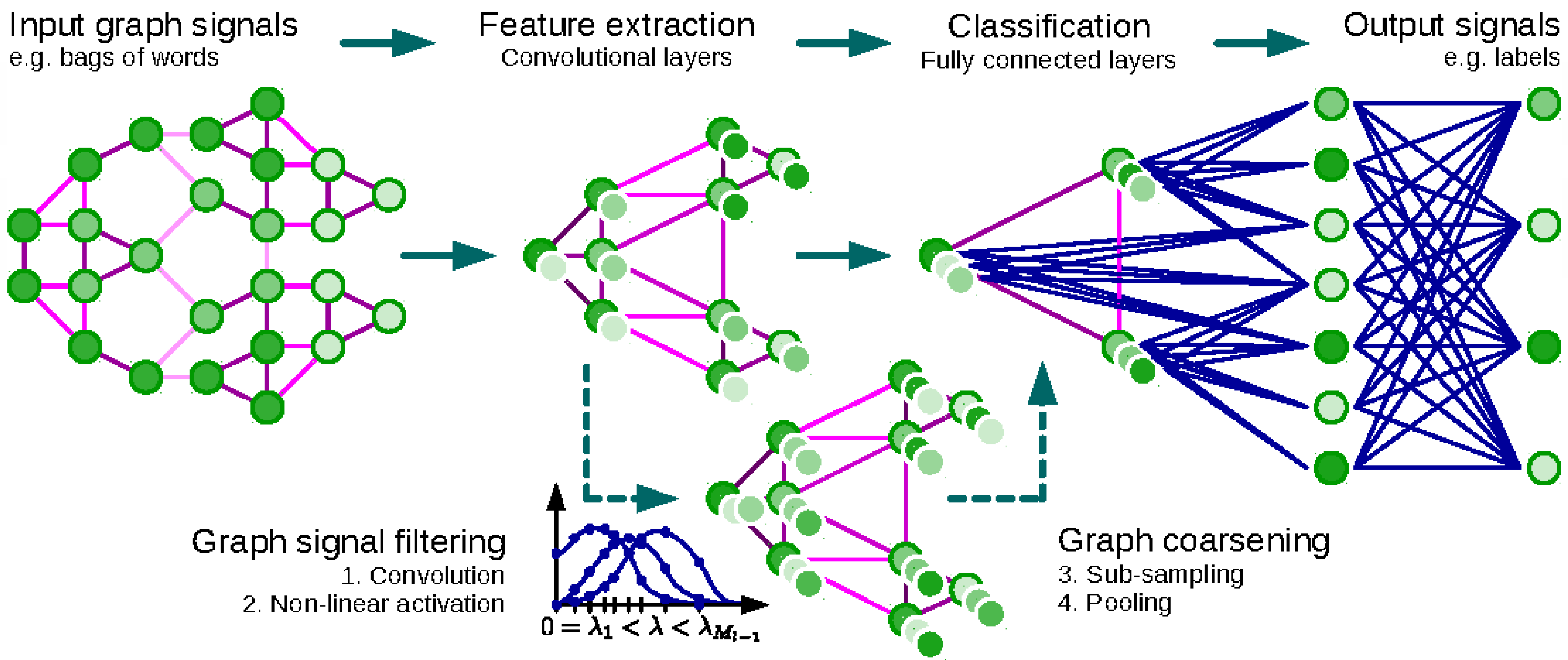
$$E(Y_{i,j}) = E(Y_{i,j}^{ideal}) - \lambda(E(Y_{i,j}^{a\_cloud})) - \gamma(E(Y_{i,j}^{a\_degr.})) - \omega E(Y_{i,j}^{a\_sat.})$$



Overlay a, b, c, d to arrive at the $Y_{i,j}$ additive network model.

Photovoltaic power($Y_{i,j}$) shown in spatio-temporal graphs for a defined location at time instance *j* under different condition in figure a,b,c & d. The figure shows $Y_{i,j}$ values from Eqn-1.

Graph Convolutional Networks



Idea: Learn how to smoothen the features across the networks using deep learning

MATERIALS DATA SCIENCE RELY

Facilities / Leveraged Technology / Other Funding Sources

- CRADLE
  - Distributed computing cluster embedded in the CWRU High Performance Computing cluster
  - Can ingest, explore and model petabyte-scale data volumes

- Graph algorithms
  - Representation of the relationship between various factors for large, heterogenous, multi-dimensional, and incomplete datasets

Risks / Countermeasures

Risk	Countermeasure(s)
Overfitting of predictive models	Hide part of data that is available and do not include in the analyses until final model is identified
Models not sufficiently accurate	Investigate the potential causes of inaccurate predictions via data mining

Thank You!

(September 2019)

Project Title: Optimization for Better Decision Making and Design

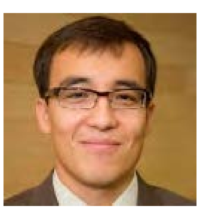
Principal Investigator(s): Oleg A. Prokopyev

New Project: X

Thrust Area: 2. Materials Data Science

Abstract: Optimization involves maximization or minimization of one or more objectives (often, economic in nature) subject to some application-specific constraints (e.g., budget, space, time limitations). Constructing an appropriate optimization model heavily depends on: the trade-offs (e.g., cost vs. benefit) and the underlying decision-making process in the application of interest (e.g., single-level vs. multi-level optimization, one-time decisions vs. decisions over multiple time periods) as well as the data availability (deterministic vs. optimization under uncertainty). In this talk, we briefly describe several of our projects that have been funded by the National Science Foundation, Air Force Office of Scientific Research, Office of Naval Research, Department of Veteran Affairs and Defense Threat Reduction Agency. In particular, we highlight the decision-making problems arising in these projects, the involved trade-offs and challenges as well as the methodological advances. Finally, we briefly discuss possible research directions for collaboration with industry or government labs.

Code Developed: This project will develop new optimization codes.



Short Bio: Dr. Oleg A. Prokopyev is a Professor in the Department of Industrial Engineering at the University of Pittsburgh. He received MS and PhD degrees in industrial and systems engineering from the University of Florida and BS and MS degrees in applied mathematics and physics from Moscow Institute of Physics and Technology (Moscow, Russia). Dr. Prokopyevs research interests are in the areas of combinatorial optimization, optimization under uncertainty, bilevel optimization and applications of Operations Research in health care, bioinformatics, network analysis and military problems. His research has been supported by the National Science Foundation, Air Force Office of Scientific Research (AFOSR), Department of Veteran Affairs, Office of Naval Research and the Defense Threat Reduction Agency. Dr. Prokopyev is a recipient of the AFOSR Young Investigator Program Award. He is the Co-Editor-in-Chief of Optimization Letters and serves on the editorial boards of IISE Transactions, Journal of Global Optimization and Omega.

Optimization for Better Decision Making and Design

Oleg A. Prokopyev

Department of Industrial Engineering

University of Pittsburgh

September 11, 2019

Optimization: Data and Models

- Mathematical model: objective, constraints, types of decisions
  - linear programs, linear and nonlinear mixed-integer programs
  - fractional programs
  - bilevel programs and multi-level programs (when multiple decision-makers)
- Data:
  - parameters are known: deterministic optimization
  - distributions are known or can be well estimated: stochastic optimization
  - worst-case over some uncertainty sets: robust optimization
  - incomplete information, learning, exploit vs. explore data: online optimization
- Solution methods and outcome:
  - global vs. local optimality
  - Pareto optimality (for multi-objective models)
  - exact, approximation and heuristic methods

Recent projects funded by NSF

- "A Novel Approach to Multistage Decision Making under Uncertainty," PI: Schaefer, Co-PI: Prokopyev
  - solution methodology for hard stochastic optimization problems
- "Integrating Proactive and Reactive Operating Room Management," PI: Prokopyev, Co-PI: Schaefer
  - scheduling under uncertainty (e.g., uncertain surgery durations)
- "Optimizing Implanted Cardiac Device Follow-Up Care," PI: Maillart, Co-PI Prokopyev
  - maintenance optimization models for patients with cardiac devices
- "Advancing Fractional Combinatorial Optimization: Computation and Applications," PI: Gomez, Co-PI: Prokopyev
  - solution methodology for hard fractional optimization models
- "Bilevel Optimization with Learning," PI: Prokopyev
  - sequential decision making with uncertainty and learning in hierarchical settings (multiple-decision makers)

Recent projects funded by AFOSR

- Two major themes:
  - detection of critical/key elements of the network
  - finding subgraphs with specific structure (e.g., “cohesive” or “tightly nit” node clusters)
- Additional complexities:
  - the problems arising in the contexts of these themes are not necessarily disjoint
  - the underlying networks may be multi-scale and/or dynamic
  - the information about network structure may be incomplete
  - the problem may involve dynamic processes in the network, e.g., information or failure cascades

Networks and Data

- Why networks?
  - Many data sets can be naturally represented by network abstractions
  - Network-based concepts are often useful for analysis of complex data sets
  - Using networks in representing data may allow intuitive interpretations of the results
  - ...

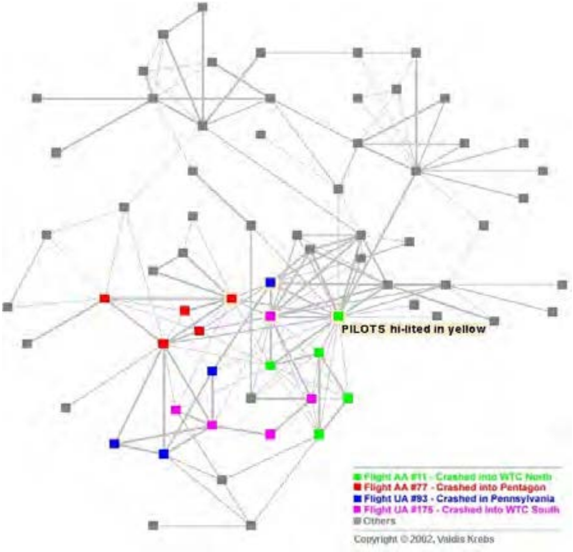
- Examples:
  - Phone call networks, World Wide Web, social networks, power grids, transportation networks, food, energy and water (FEW) networks, ...

"*Networks are present everywhere. All we need is an eye for them.*"
Barabási, "Linked: The New Science of Networks"

Greg's Cable map: undersea communications

Source: http://www.cablemap.info

9/11 hijackers (Krebs, 2002)

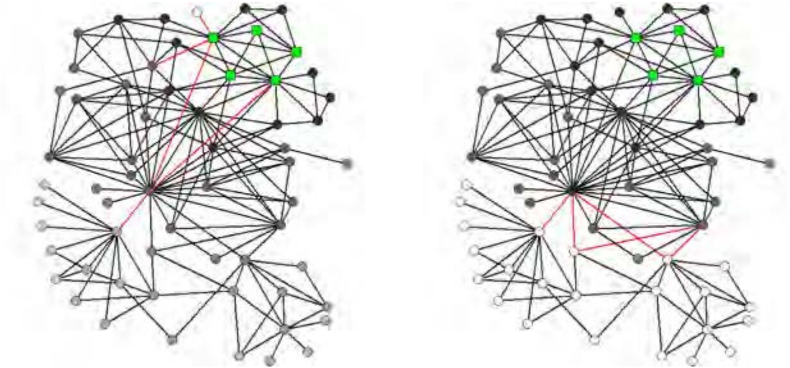


Source: http://www.orgnet.com/hijackers.html

Key Players & Links in a Network

- Given graph $G = (N, E)$, what are the most **"important" nodes and edges** of G?
- Two possible ways to address this question:
  - **"Negative" view**: removal of a set of graph elements *maximally degrades connectivity/cohesion* of G according to a pre-defined metric
    - such elements are referred to as **critical nodes and/or edges**
    - closely related to **network interdiction** problems
  - **"Positive" view**: a set of graph elements that are *maximally "connected"* to all other nodes according to a pre-defined metric
    - for example, where to locate monitoring devices?
- Various aspects and generalizations can be explored: global vs. local properties, dynamics, structures, etc.
- Useful for characterization of various important properties of networked systems
  - Application domains: telecommunication, transportation, social network analysis, homeland security and defense areas

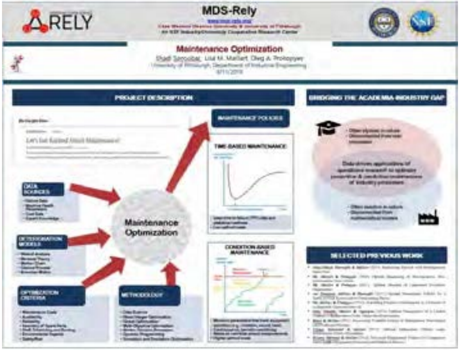
Example: Critical Links



- Five **most** critical links with respect to the decay centrality measure: local (for a group of nodes) vs. global connectivity

Maintenance Optimization

- Strike an optimal balance between the costs of
  - performing preventive and/or predictive actions on a system
  - entering an undesirable state, e.g., failure
- Please see poster (joint with Shadi Sanoubar and Lisa Maillart):



Stochastic Modeling, Analysis and Control (SMAC) Laboratory at Pitt IE

- Primary research mission: mathematical modeling, analysis and control of engineering, service and other systems that have inherently stochastic elements
- Research in the lab emphasizes
  - analytical and computer-based modeling of such systems, e.g.,
    - reliability, maintenance, production, inventory, medical decision making, healthcare operations and policy,...
  - and their optimization
    - e.g., by exploiting applied probability, stochastic processes, stochastic optimal control techniques,...
- Director: Dr. Lisa Maillart
  - Email: maillart@pitt.edu

Summary and concluding remarks

- What type of collaborators/projects are we interested?
  - Broadly speaking, we are interested in applications that can be stated as (or involve) complex decision-making problems, possibly, sequential, hierarchical and under uncertainty
- What do we have to offer?
  - Mathematical modeling of decision-making problems/processes
  - Methodologies:
    - Combinatorial optimization, global optimization: optimization on graphs/networks is one example
    - Stochastic (robust) optimization: data (e.g., some of the problem parameters) are uncertain
    - Hierarchical (bilevel) optimization: the decision-making process involves multiple decision-makers, possibly acting in a hierarchical and decentralized manner
    - Maintenance optimization: optimization of procedures to achieve better maintainability, reliability, and availability

MDS-Rely Project Proposal
(RS Area C - Thrust Area 1 - Proposal Number TBD)
(September 2019)

Project Title: Reliability of Metallic Parts	
Principal Investigator(s): John Lewandowski	**Researcher: Nat Tomczak/Dr. Janet Gbur**
New Project: XX Renewal: **Term: 1 year, X 2 years**	**Start Date: June 2020**

Thrust Area: 1. Study Protocols

Objective: The goal of this project is to develop accelerated protocols for mechanical testing and analysis to address the reliability of metallic parts for biomedical (e.g. wires, implantable devices) and/or structural (e.g. additively manufactured) applications. The Advanced Manufacturing and Mechanical Reliability Center (AMMRC: https://ammrc.case.edu) at CWRU houses extensive equipment for accelerated mechanical testing along with remote access and/or control of certain machines. Continuing interactions with various ASTM committees on testing of biomedical wires and additively manufactured materials provides close contact with industrially relevant concerns in these areas. Rapid assessment of processing effects on strength, fatigue resistance, and toughness is possible on both standardized samples as well as samples excised from parts. One of the major advantages of these approaches is the possibility to include performance assessment of manufactured parts.

Standards Used: Mechanical testing of materials, components or systems, such as Tension Testing (ASTM E8, F2516), High Cycle Fatigue (ASTM E466, E796, E2948), Fatigue Crack Growth (ASTM E647), Toughness (ASTM E399, E1820, E23).

Code Developed: Expanding upon our ability to provide remote access/control of certain experiments, we will continue to develop techniques to share data as it evolves, enabling more rapid assessment as well as adjustments of test protocols as necessary.

Datasets Produced: The digital data files as well as digested results will be provided. Processing effects on material behavior can be correlated with microstructural changes and characteristic fracture surface features.

Background: The PI and collaborators have recently written a number of review articles on reliability of wires for biomedical applications [1], reliability of additively manufactured materials [2], and environmental effects on stress corrosion cracking of commercial Al-Mg alloys [3]. There remains a need for more rapid assessment of properties of interest to the industrial community. Figure 1 shows the effects of strain level on the high cycle fatigue behavior of Nitinol implantable wires, a shape memory/superelastic material used widely in the biomedical industry. Considerable variability of performance is shown due to differences in the processing/cleanliness. The AMMRC at CWRU has developed techniques to rapidly assess the fatigue performance of such wires, followed by detailed microstructure/fractographic investigations to determine the source(s) of such variability. Similar techniques can be utilized to assess the behavior of wires used in a variety of sectors, including automotive, aerospace, and biomedical. Facilities also exist, or can be designed, to examine the reliability of components containing wires. Dr. Janet Gbur at CWRU, working on various teams with the PI has been directly involved with evolving ASTM standards [4] and the associated Interlaboratory Studies Program (ILS), as well as with FDA approvals for implantable biomedical devices containing such wires. Similar, and/or customizable approaches can be developed for other applications such as flexible electronics and other industry sectors that utilize flexible cables and thin films.

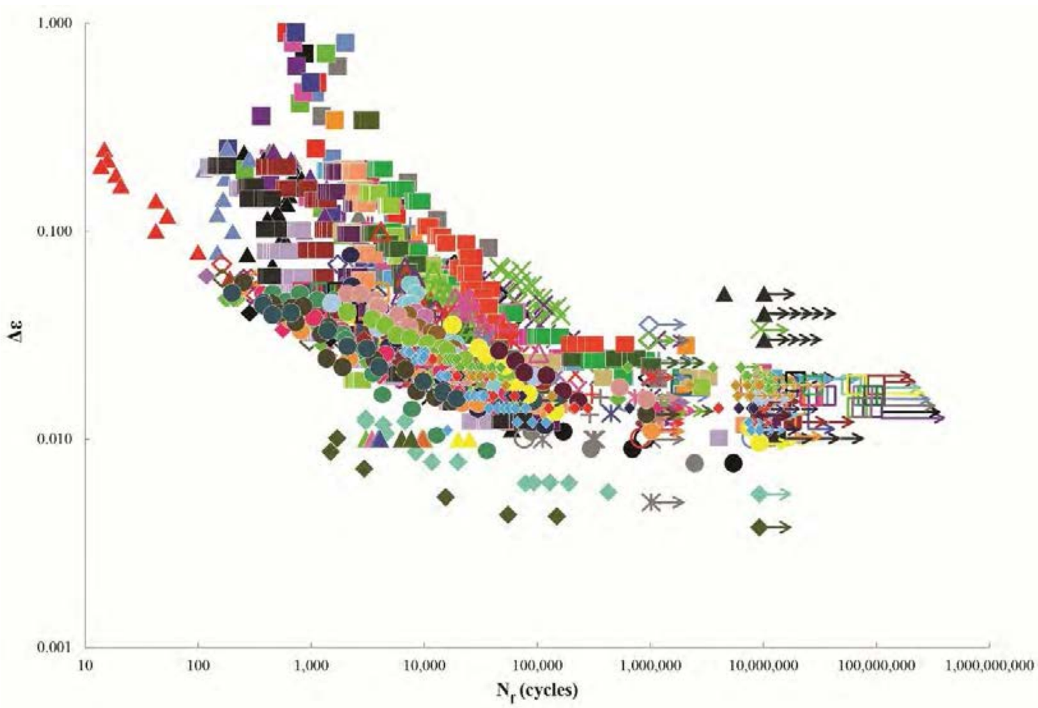
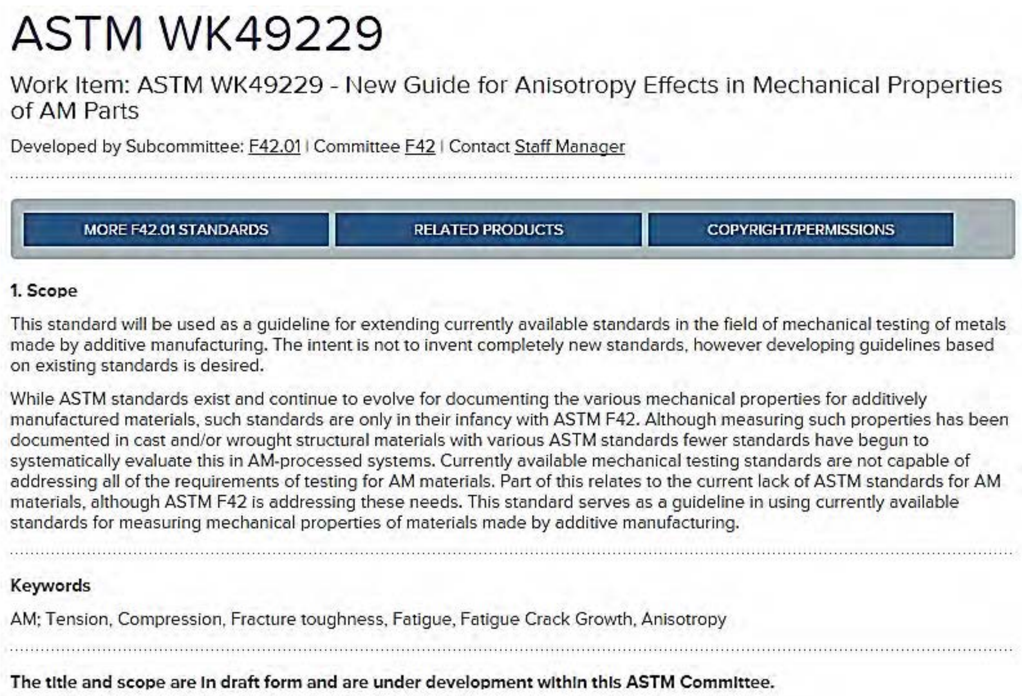


Figure 1 – Fatigue behavior of commercially available Nitinol shape memory/superelastic materials showing wide variation in fatigue lifetimes [1].

In the area of Additive Manufacturing, the PI has been directly involved with the ASTM Director of Global Additive Manufacturing, Dr. Mohsen Seifi, on the development of an ASTM work item WK49229 [5] for testing of additively manufactured samples, with extension to components. Figure 2 summarizes this effort and builds directly on the extensive ongoing work at CWRU on mechanical reliability of AM samples/parts. The PI currently has five (5) separately funded grants on mechanical reliability of AM samples/parts. Accelerated testing techniques are being explored to examine the fatigue performance of samples processed using different parameters (e.g. power, velocity, etc.). This approach also is relevant for additively manufactured flexible electronics.


ASTM WK49229

Work Item: ASTM WK49229 - New Guide for Anisotropy Effects in Mechanical Properties of AM Parts

Developed by Subcommittee: F42.01 | Committee F42 | Contact Staff Manager

MORE F42.01 STANDARDS | RELATED PRODUCTS | COPYRIGHT/PERMISSIONS

1. Scope

This standard will be used as a guideline for extending currently available standards in the field of mechanical testing of metals made by additive manufacturing. The intent is not to invent completely new standards, however developing guidelines based on existing standards is desired.

While ASTM standards exist and continue to evolve for documenting the various mechanical properties for additively manufactured materials, such standards are only in their infancy with ASTM F42. Although measuring such properties has been documented in cast and/or wrought structural materials with various ASTM standards fewer standards have begun to systematically evaluate this in AM-processed systems. Currently available mechanical testing standards are not capable of addressing all of the requirements of testing for AM materials. Part of this relates to the current lack of ASTM standards for AM materials, although ASTM F42 is addressing these needs. This standard serves as a guideline in using currently available standards for measuring mechanical properties of materials made by additive manufacturing.

Keywords

AM; Tension, Compression, Fracture toughness, Fatigue, Fatigue Crack Growth, Anisotropy

The title and scope are in draft form and are under development within this ASTM Committee.

Figure 2 – PI and collaborators are directly involved with ASTM work item WK49229 [5] on assessing anisotropy effects in mechanical properties of AM samples/parts.

In the area of environmental effects on fracture of commercial Al-Mg alloys used in a variety of naval and aquatic applications, the effects of various time/temperature thermal exposures on the susceptibility to stress corrosion cracking are being determined [3]. In addition to conducting experiments on commercial alloy plates given prescribed time/temperature combinations and subsequently evaluating the stress corrosion cracking (SCC) behavior in a variety of environments, material removed from a service application is also being tested. Accelerated test techniques are being explored to quickly determine the susceptibility to SCC. These results are being correlated with existing ASTM standards to indicate regimes of susceptibility. Three (3) current projects are being supported on this type of work.

Project Tasks:

The major components of this work require input from industry on mechanical properties of interest. Specific examples have been provided regarding wires, additively manufactured materials, and environmental cracking resistance.

The major tasks outlined in this proposal are:

1. Obtain industrial input on materials/properties of interest given existing/evolving processing techniques.
2. Identify appropriate test techniques to conduct testing in the AMMRC in addition to examining possible accelerated test techniques. Conduct preliminary testing to evolve the data base and determine appropriate changes to the procedures, if needed.
3. Continue to conduct mechanical characterization of property(ies) of interest and correlate these to existing/evolving processing techniques.

Benefits to Members: The extensive facilities in the AMMRC (https://ammrc.case.edu), experience of the PI and staff in various mechanical characterization techniques, in addition to interactions with evolving ASTM standards for both wire testing as well as mechanical characterization of additively manufactured materials/components, provides a unique opportunity to address evolving reliability issues. In the manufacturing sector, the ability to rapidly characterize some of the key mechanical properties and correlate them with changes in processing details will be useful to various members along the supply chain. Such approaches can be useful in providing the quantitative analyses to the production team and can also be a part of the original and/or evolving testing setup. These kinds of accelerated techniques can be useful across the supply chain for new/evolving materials and products, in addition to the evaluation of properties of materials/systems at the end of their design lifetime.

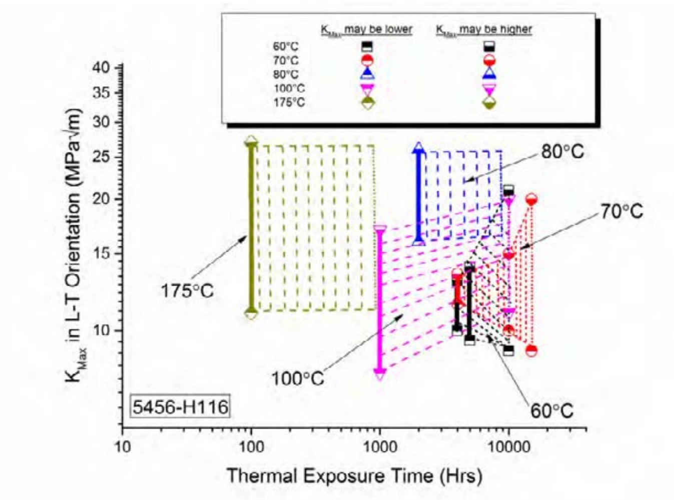


Figure 3 – Regimes of susceptibility to Stress Corrosion Cracking after different thermal exposure times/temperatures for a commercial Al-Mg alloy, AA 5456-H116.

As an example, in one of our current projects on commercial solid solution strengthened Al-Mg alloys used in naval and aquatic applications, we are studying the effects of controlled laboratory thermal exposures at moderate temperatures on the subsequent resistance to stress corrosion cracking in both humid air and simulated seawater. This is being compared to the behavior of similar ship plate material removed after 42 years in service. In all cases, accelerated testing is being used to establish regimes of susceptibility. Figure 3 shows regimes of cracking that occur for different thermal exposure time/temperature combinations on commercial plate and illustrate the ability to accelerate the cracking response of this commercial Al-Mg alloy. Exposures of 175C/100 hours are sufficient to accelerate the cracking response compared to exposures in excess of 1,000's of hours at 60C. These observations are being correlated with ASTM standard G67 Nitric Acid Mass Loss Test (NAMLT) values on both commercial plate that has been thermally exposed in addition to ship plate removed from service.

Timeline:

Tasks	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**
Determine Properties of Interest	**15%**	**20%**	**20%**	**25%**	**20%**			
Identify/Conduct Accelerated Testing		**5%**	**5%**	**20%**	**20%**	**25%**	**25%**	
Correlate Processing Changes with Performance				**10%**	**10%**	**20%**	**30%**	**30%**

References:

[1] "Fracture and Fatigue of Wires and Cables Used in Biomedical Applications", J.L. Gbur and J.J. Lewandowski, International Materials Reviews, 61(4), pp. 231-314, 2016.

[2] "Additive Manufactured Metals - A Review of Mechanical Properties, J.J. Lewandowski and M. Seifi, Annual Review of Materials Research, pp. 151-186, 2016.

[3] Seifi, S.M., Ghamarian, I., Samimi, P., Collins, P.C., Holroyd, N.J.H., and Lewandowski, J.J. (2018). "Sensitization and Remediation Effects on Environmentally Assisted Cracking of Al-Mg Naval Alloys", Corrosion Sci., 138, pp. 219-241.

[4] "Standard Test Method for Conducting Rotating Bending Fatigue Tests of Solid Round Fine Wire", ASTM International, ASTM E2948, West Conshohocken, PA, pp. 1-10, 2016.

[5] https://www.astm.org/DATABASE.CART/WORKITEMS/WK49229.htm.

Reliability of Metallic Parts

John J Lewandowski

Case Western Reserve University

Department of Materials Science and Engineering

Proposed Project Duration: 2 years

Industrial Relevance and Novelty

1. Rapid and novel assessment of mechanical reliability of samples and/or manufactured parts in accordance with existing/evolving ASTM standards.
   PI, staff, and former student directly involved with evolving ASTM standards.
2. Access to unique and extensive facilities and expertise within the CWRU Advanced Manufacturing and Mechanical Reliability Center (AMMRC) (https://ammrc.case.edu)
   Remote access/control of some computer controlled equipment possible.
3. Project(s): Wires/cables/films for implantable biomedical/other applications;
   Additively manufactured metallic samples/parts;
   Environmental effects on cracking in structural materials.
4. PI and students have written recent major review articles in each topic that highlight industrially relevant and ASTM compliant research on mechanical reliability regarding:
   a) test techniques for fatigue of wires/cables/films
   b) effects of surface roughness and defects in AM samples/parts
   c) environmental effects on cracking in structural materials
5. The accumulated knowledge over many decades with a focus on key metallurgical/processing effects affecting reliability using advanced characterization tools and techniques.

MATERIALS DATA SCIENCE RELY

Proposal Objectives

Long term objectives:

- Develop/extend accelerated protocols for mechanical reliability
- Impact existing/evolving ASTM standards for accelerated characterization (wires/cables/films; fatigue testing for AM; environmental cracking)
- Expand the industrial use and facilities in the AMMRC
- Extend sample testing approaches to parts/components
- Increase the use of graduate internships/co-ops with interested parties
- Integrate data science/machine learning/vision (microstructure, fracture surfaces)

Short term objectives:

- Establish/extend linkages with relevant and interested parties
- Focus on materials system(s) and properties of most current interest
- Lay the groundwork for additional focused projects

Proposal Deliverables

1. Access to student(s) fully trained in mechanical reliability concepts/techniques
2. Access to student(s) with extensive metallurgical expertise-ongoing
3. Willingness to share student(s) for internships/co-ops-ongoing
4. Access to unique equipment/staff and expertise–ongoing
5. Interaction with ongoing/evolving ASTM standards in areas supported–ongoing
   (Wires/cables; AM samples/parts; Environmental Cracking)
6. Detailed analyses conducted on mechanically characterized samples/parts
   (Fracture Surfaces, Microstructure, Surface Analyses, Microscopy)
7. Possibility of part-time pursuit of graduate degree for current employees
8. Access to specialty course(s) on CWRU Mediavision
   (Failure Analysis, Fracture of Materials)
9. Access to JJ Lewandowski network (>graduated 100 MS/PhD, >50 post-docs)

Proposal Timeline

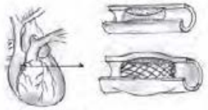
Task / Milestone	Year 1				Year 2				Year 3			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Determine Matl's System(s) and Properties of Interest (Wires/Cables/Films; Fatigue of AM; Environment Effects)	15%	20%	20%	25%	20%							
Identify and Conduct Accelerated Testing (Maintain/establish contact with existing/evolving ASTM)		5%	5%	20%	20%	25%	25%					
Correlate Processing Changes with Reliability				10%	10%	20%	30%	30%				

MATERIALS DATA SCIENCE RELY

Proposal Work Description

Task	Work to be Done	Outcomes / Deliverables
Reliability of Wires/Cables (Biomedical, Automotive, Aerospace, etc.)	Establish/extend accelerated test protocols; Determine processing/cleanliness effects on fatigue performance	Understanding of factor(s) controlling fatigue durability/reliability
		Continue interactions with existing/evolving ASTM standards

Medical device performance and reliability
- Incidence of failures from 18.1[1]-65%[2]
- Nitinol wires prevalent in literature



Cardiac stent
Photo courtesy:
Gopal: HCL White Paper, 2009

FDA Approved Medical Devices with Nitinol[3]

Device Category	Sample Devices	Number of Devices	Usage Time	Recalls for Failure
Cardiology	Stents, filters	234	< 24 hours or > 30 days	Yes
Orthopaedics	Staples, clips	128	< 24 hours or > 30 days	Yes
Surgical Instruments	Catheter, needle	49	< 24 hours	Yes
Urology	Temporary stent	2	< 24 hours or > 30 days	No
Neurology	Lead anchor band	1	> 30 days	No
Totals		414		

Literature reference:
[1] Allie et al., Endovascular Today, 2004
[2] Diehm et al., Journal of Vascular and Interventional Radiology, 2009
[3] FDA approved medical devices containing Nitinol, www.fda.gov

8

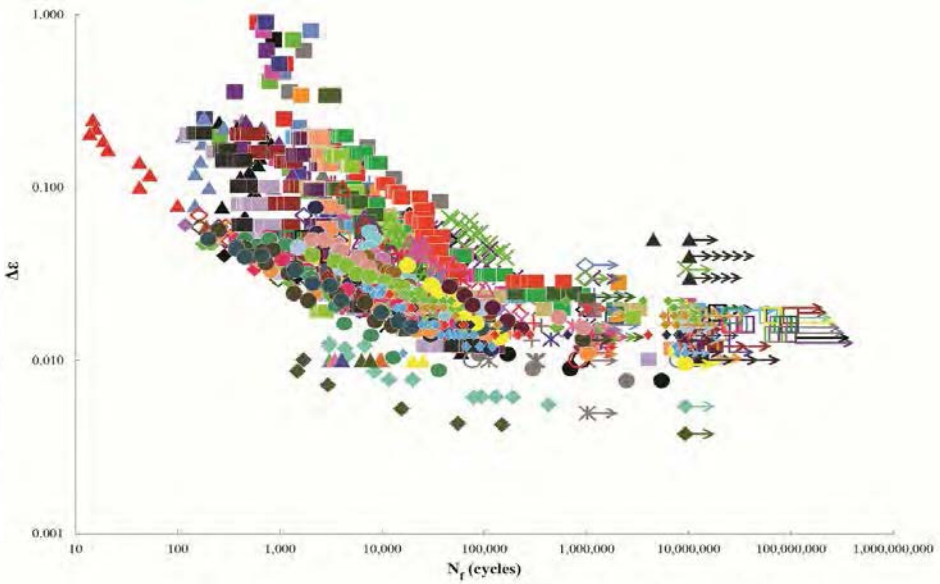
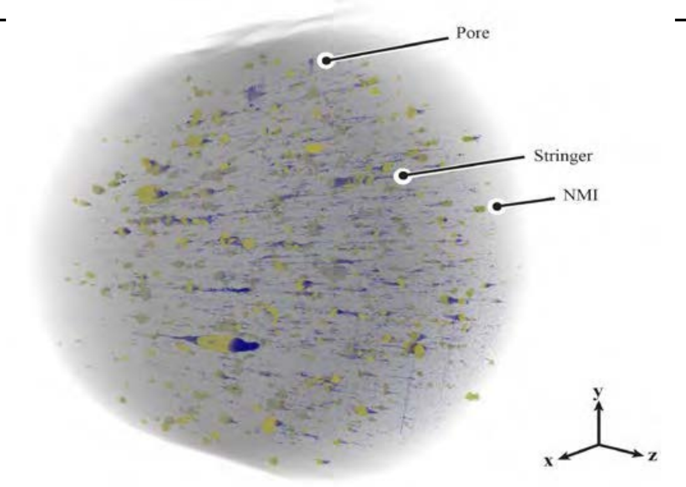


Figure 1: Fatigue behavior of commercially available Nitinol superelastic materials showing wide variation in fatigue lifetimes [1].

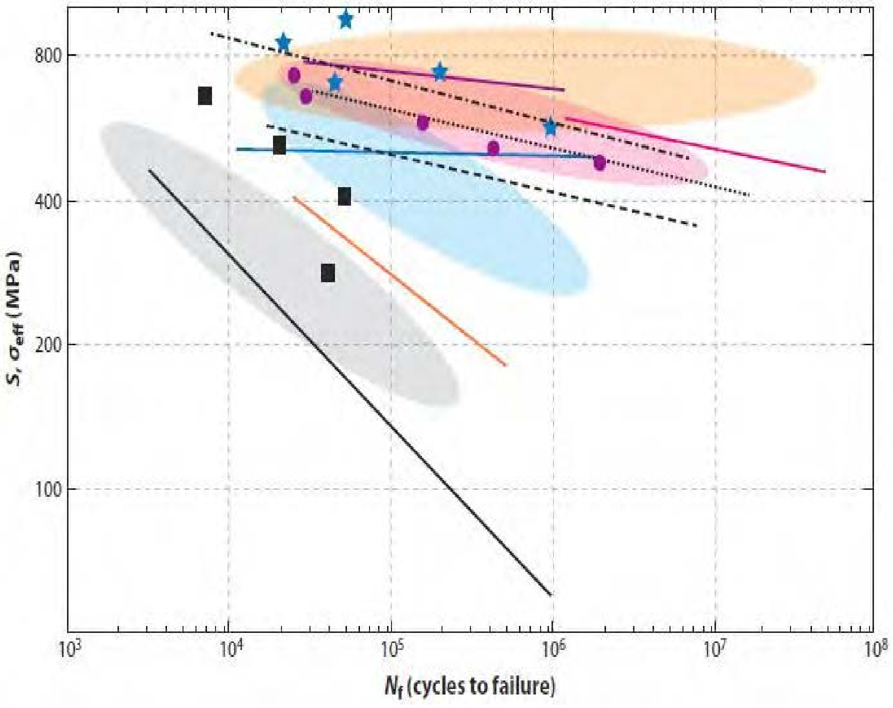


μ-CT scan showing non-metallic inclusions (NMI)

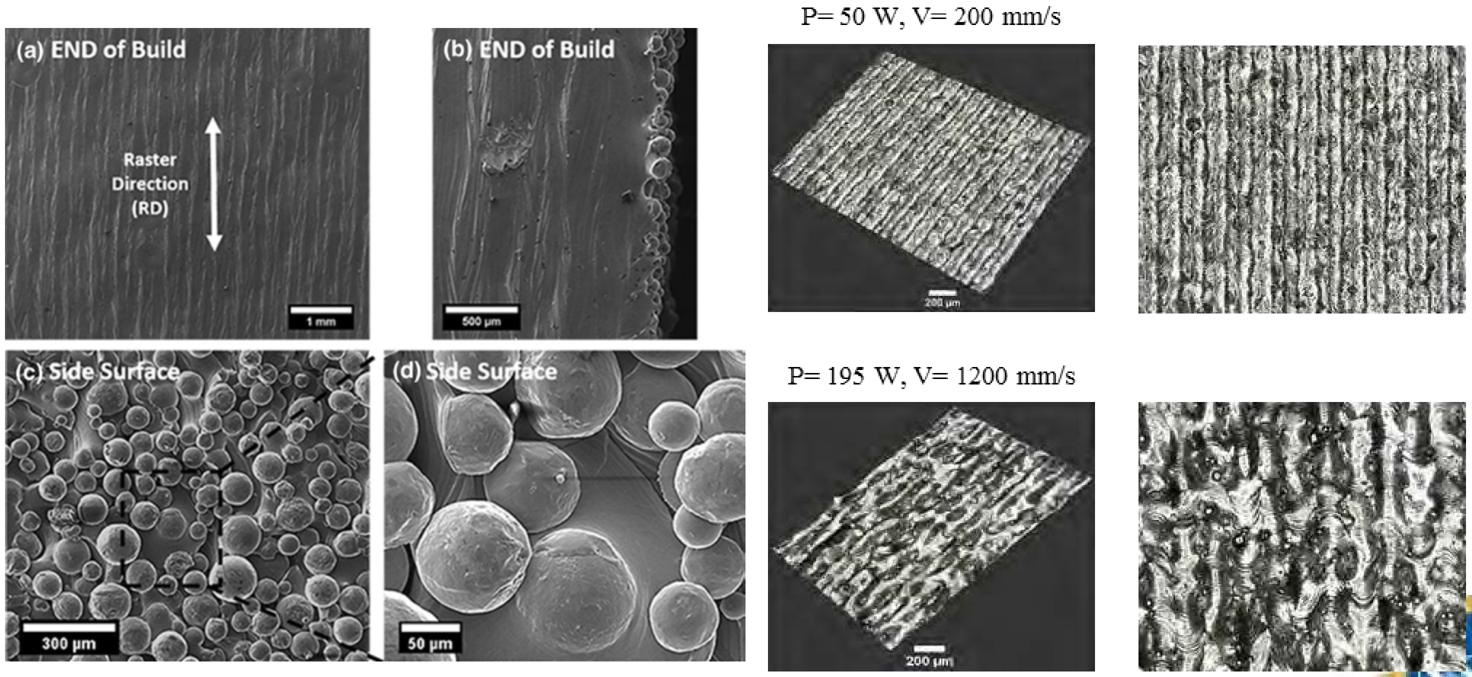
MATERIALS DATA SCIENCE RELY

Proposal Work Description

Task	Work to be Done	Outcomes / Deliverables
Fatigue of Additively Manufactured Materials/Parts	Determine effects of surface roughness, defects on fatigue performance;	Quantify effects of processing on surface roughness/defects and fatigue performance
		Continue interactions with existing/evolving ASTM standards



Fatigue Variability in AM – PBF [2]



Examples of Surface Roughness on AM Samples/Parts [2]

Proposal Work Description

Task	Work to be Done	Outcomes / Deliverables
Environmentally Assisted Cracking (EAC) of Structural Materials	Develop accelerated test protocols to determine EAC susceptibility, understand EAC mechanisms operating	Understanding of mechanisms controlling EAC and how to accelerate testing without changing mechanisms; relation to service history
	Relate relevance of laboratory exposures to service exposure	Continue interactions with existing/evolving ASTM standards

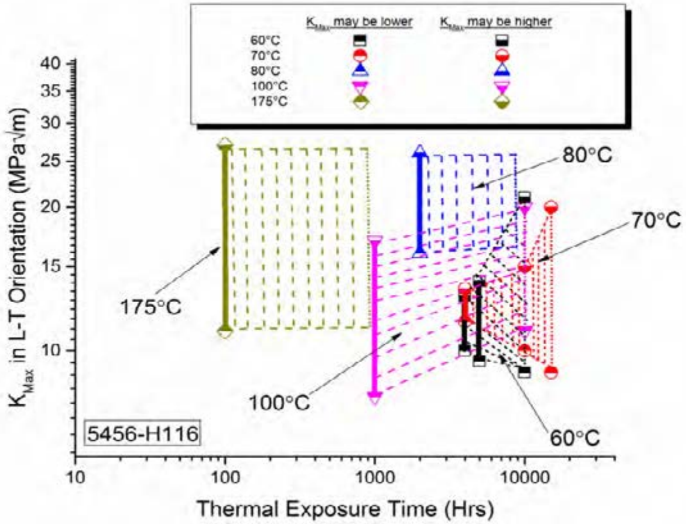
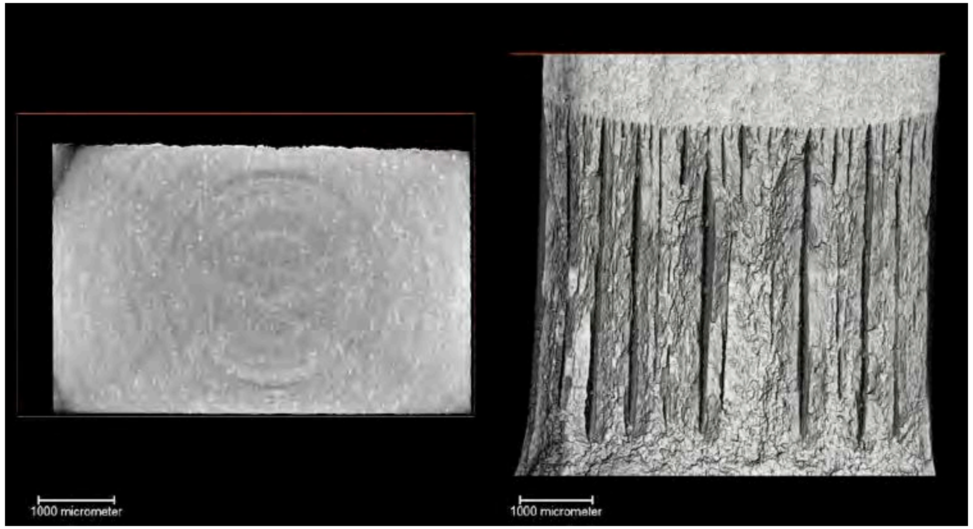


Figure 3 – Regimes of EAC susceptibility in Al-Mg alloy [3]. Shaded regions show thermal exposure time/temperatures that will cause EAC in HUMID air!



High-resolution tomography of EAC crack path – Diamond Light Source Synchrotron (UK)

MATERIALS DATA SCIENCE RELY

Facilities/Leveraged Technology/Other Funding Sources

1. Facilities: Advanced Manufacturing and Mechanical Reliability Center (AMMRC) (https://ammrc.case.edu)
   Equipment Value = $5.5M, Sub-scale and Large-scale testing equipment, Computer controlled, Web access
2. Leveraged Technology – existing and evolving interactions with ASTM standards (wires/cables, AM materials)
   - PI and staff/students have written three major review articles outlining industrially relevant research
   - Wires/Cables; Mechanical Characterization of AM Materials/Samples/Parts; Environmental Cracking
   - Accelerated Test Protocols
3. Other Knowledge/Funding Sources Leveraged
   - Wires/Cables for Biomedical Applications-Next Generation Artificial Limbs: DARPA/NIH
   - Processing and Reliability of Additive Manufactured Materials/Parts:
     - NAVAIR/ONR: Uncertainty Qualification for Fatigue of LPBF Ti-6Al-4V
     - ONR: High Deposition Rate AM of CP-Ti Grade 2
     - NASA: In-situ Alloying of Dispersion Strengthened GR-Cop 42
     - NASA ULI: Development of an Ecosystem for Qualification of AM Processes and Matl's in Aviation
     - America Makes: Maturation of Advanced Manufacturing for Low-cost Sustainment (MAMLS)
   - Environmentally Assisted Cracking (EAC) Resistance of Structural Materials
     - ONR: Thermal Exposure Effects on EAC and Remediation for Commercial Al-Mg Alloys
     - ONR: Fracture/Fatigue Behavior of Al-Mg Structures on HMCS Iroquois (Retired Navy Ship-42 yrs.)
     - Diamond Light Source Synchrotron Facilities (UK) – In-situ testing/tomography beamtime
       (Discussions/collaborations with US Synchrotron Facilities Evolving)

Risks / Countermeasures

Risk	Countermeasure(s)
Inadequate Access to Pedigree Materials/Parts	Wires/Cables: Link with existing suppliers/end users Additively Manufactured Samples/Parts: Link with supply chain
Test Machine Access Delays Due to Other Supported Research	Additional test machines can be purchased if enough funds exist Surplus machines are often available at reduced costs
Timing of Funding in Relation to Next Student Recruitment Cycle	Need to continue to recruit graduate students, Sufficient time is needed regarding Potential award(s)/project(s) to recruit appropriate students into the program

Thank You!

MDS-Rely Project Proposal
(RS Area D - Thrust Area 3 - Proposal Number TBD)
(September 2019)

Project Title: Time Series Analysis of Power Generating Systems in the Field	
Principal Investigator(s): Mark De Guire	**Researcher: TBD**
New Project: XX Renewal: **Term: 1 year, XX 2 years**	**Start Date: June 2020**

Thrust Areas: 1. Study Protocols, 2. Matls. Data Science, 3. Reliability Studies

Objective: The goal of this project is two-fold: 1) to assemble field data on performance of electrical power generating or storage systems in the field, such as fuel cells, solar cells, wind turbines, and electrochemical energy storage systems (batteries, flow batteries); and 2) correlate system performance with relevant parameters related to e.g. service conditions (current, voltage, demand, fluctuations thereof, etc.) and ambient external conditions (e.g. weather/climate, malfunctions, disruptive events). For power generating and storage systems, long-range degradation often limits their lifetime and — more importantly — their cost-effective implementation. Therefore time-series data will be especially useful, as they afford great opportunity for improved understanding of degradation mechanisms in these systems.

Standards Used: These will depend on the systems for which datasets become available. All relevant existing standards for performance, operational stresses, and safety will be applied to the system(s) under study.

Code Developed: Again, this will depend on the systems for which datasets become available.

Datasets Produced: When data analysis detects significant correlations between operating conditions or external circumstances and performance changes, the PI will work with industry partners to design and (time permitting) conduct lab simulations of the critical conditions and monitor relevant performance parameters, to replicate and confirm the effects observed in the field.

Background: For power generating or storage systems, monitoring of performance, especially deviations from normal performance, and correlating it with operating conditions, or with events not usually encountered in lab settings but to be expected in service, can provide understanding into what conditions are most conducive to successful long-term operations. In some cases, such correlations serve as clues to the underlying mechanisms of performance changes. These can yield insights that lead to improved designs or better protocols for system operation, and ultimately to longer and more reliable system life.

For example, ceramic solid oxide fuel cells (SOFCs) directly convert chemical energy, released during oxidation of a fuel, into DC electrical energy. Certain operating conditions are known to affect long-term performance adversely: thermal cycling, load cycling, redox cycling, and operating at or beyond the peak power of the cell's current-voltage characteristic.[1–4] Isolated events in the field, which can occur especially to remote or distributed power systems, can also lead to performance degradation. The PI's group has documented such effects on a lab scale: the consequences of e.g. interruption of the fuel stream, leakage of seals, and inadequate exhaust of oxidation products.[5] The last of these situations, though easily avoided under lab conditions, can occur in the field if a random event obstructs the system's exhaust line. Figure 1 illustrates the distinctive signal of such an occurrence: a sudden drop in output voltage (and corresponding increase in cell area specific resistance, ASR) after 30 hours of operation at fairly high current density (571 mA cm^{-2}) and fuel utilization (44.5%). The decreased output voltage, and resulting drop in power, is bad enough. If uncorrected (as in the case of Figure 1), the back-pressure of steam causes irreversible chemical and microstructural damage on the fuel-side electrode of the SOFC,[6] which in this case led to irrecoverable failure of the cell in about 24 hours. Once the connection was recognized between exhaust blockage and the resulting characteristic drop in cell voltage, the operator could recognize the circumstance that was causing the change in cell performance and effectively eliminate the problem.

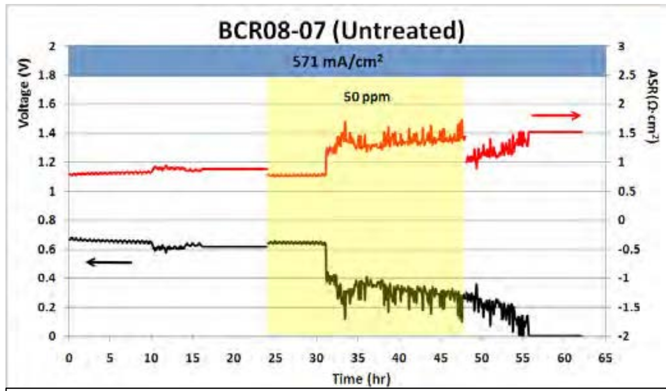


Figure 1. Typical effect of moisture accumulation in the exhaust line after 30 h of testing: an abrupt drop in voltage (bottom curve) and concurrent rise in area specific resistance (top curve). Current density: 0.571 A/cm^2; fuel utilization = 44.5%. Cell was exposed to 50 ppm H_2S from hours 24 to 48.

[1] A. Weber, J. Szasz, S. Dierickx, C. Endler-Schuck, and E. Ivers-Tiffee. "Accelerated Lifetime Tests for SOFCs," *ECS Transactions* **68** [1] (2015): 1953-1960.

[2] Yeon-Hyuk Heo, Jong-Won Lee, Seung-Bok Lee, Tak-Hyoung Lim, Seok-Joo Park, Rak- Hyun Song, Chong-Ook Park, and Dong-Ryul Shin, "Redox-induced Performance Degradation of Anode-supported Tubular Solid Oxide Fuel Cells." *International Journal of Hydrogen Energy* 36.1 (2011): 797-804.

[3] G. Almutairi, K. Kendall, and W. Bujalski, "Cycling Durability Studies of IP- SOFC." *International Journal of Low-Carbon Technologies* 7 [1] (2011): 63-68.

[4] M. Näslund and H. Iskov, *Accelerated Lifetime Testing and Standardization of SOFC Systems*. Rep. no. 734-90. Danish Gas Technology Centre, 2012.

[5] C. C. Wu and Mark R. De Guire, "Performance of Solid Oxide Fuel Cells under Fuel-Side Operational Stresses," Proceedings of *EnergyTech 2012,* Cleveland, Ohio, 30 May 2012, published online (DOI: 10.1109/EnergyTech.2012.6304695).

[6] K. Du, F. Ernst, M. Garrels, and J. Payer, "Formation of Nickel Nanoparticles in Nickel-Ceramic Anodes during Operation of Solid-Oxide Fuel Cells," *Internat. J. Mater. Res.*, **98** (2008): 548-552.

The bright side of such an occurrence is that continuous remote monitoring of the cell output, combined with well-designed data analysis algorithms operating in real time, could detect the “signature” voltage drop, diagnose the specific problem, and alert system operators to initiate remedial action. This is one example of the kind of correlation between conditions and performance that the proposed project aims to discover.

Another phenomenon in SOFC performance that lends itself well to data analytics is low-amplitude cycling of cell output at 24-hour intervals (Figure 2).[7] This phenomenon is known to system researchers and developers, but it is rarely reported in the open literature and is not well understood. When detected, the voltage fluctuations are very small (~5–10 mV) and regular, with a distinctive “plateau” shape; but in some single 500-hour tests (not shown here) they have appeared and disappeared over periods of several days. The diurnal nature of the effect suggests that ambient operating conditions (such as temperature, relative humidity, or barometric pressure) could be at play, though it is unclear whether the effect is in the cell itself or in the electronics monitoring the cell performance. Although the instantaneous effects of these voltage fluctuations on cell power are miniscule, their long-term effects on cell performance (if any) are unknown. As mentioned above, the load cycling that was also being applied during this test (at each vertical red line in Figure 1) is recognized[2,4] as a cause of degradation in cell performance. A challenge to address using data analytics would be to sort out the long-term effect on cell performance of short-term perturbations in operating conditions whose immediate effect is small.

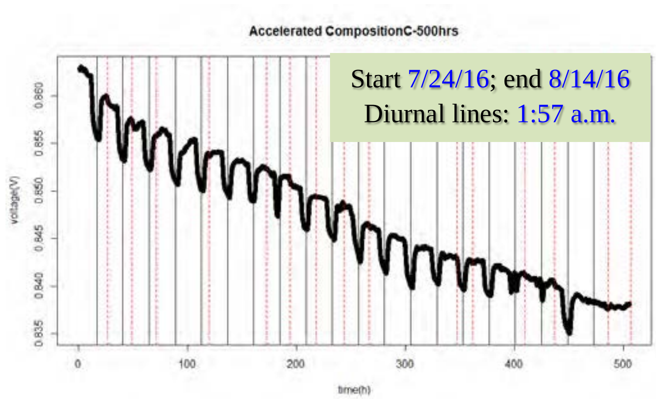


Figure 2. Diurnal fluctuations in the output of an SOFC “button” cell during accelerated testing of 500 h. Current density: 0.760 A/cm^2; 1,000 °C. The vertical black lines are spaced daily at 1:57 a.m.; nearly every one coincides with a drop in output voltage of ~5 mV. The red lines are spaced whenever a loading cycle (0–800 A/cm^2) was applied; many (but not all) of these can be seen to coincide with subsequent smaller perturbations in the cell voltage.

In the context of energy storage, a related type of correlation that could emerge from the project proposed here would be to analyze the field performance of battery packs in electric vehicles (EVs) with respect to weather and use conditions. Such an analysis would complement a modeling study that was recently carried out by Yang et al.[8] They used mathematical models to predict how ambient temperature, frequency of recharging, and use cycles would affect battery degradation in EVs. Inputs to the models included average consumer driving patterns and annual weather data in all 50 of the United States. Their results predicted “battery life ranging between 5.2 years in Florida and 13.3 years in Alaska.” Correlating field data on the performance of EV

---

7 Hanke Gu, Celeste Cooper, Roger H. French, and Mark R. De Guire, “Data Analytics Applied to SOFC Durability Time-series Datasets,” poster presented at the 18th Annual Solid Oxide Fuel Cell Project Review Meeting, Pittsburgh, Pennsylvania, 12–14 June 2017. Available online at https://netl.doe.gov/sites/default/files/event-proceedings/2017/sofc%20proceedings/poster/De-Guire.pdf

8 Fan Yang, Yuanyuan Xie, Yelin Deng & Chris Yuan, Predictive modeling of battery degradation and greenhouse gas emissions from U.S. state-level electric vehicle operation, *Nature Comm.* (2018) **9** 2429–2437. doi: 10.1038/s41467-018-04826-0.

batteries to actual use and ambient temperature data would not only provide a valuable cross-check to model predictions; it would afford refinement of the models being used, and would enhance physical understanding of battery degradation mechanisms.

Project Tasks: As discussed in the Background section within the context of fuel cells, certain performance "signatures" that correlate to operating or external conditions are known in many of the types of systems relevant to this project. Those systems for which partners are willing to share time-series performance data from the field (company-provided and de-identified) will determine which systems will be studied, especially if operating or external conditions can be mapped onto the performance data versus time. **Task 1** will consist of collecting such time-series datasets. **Task 2** will entail development of a robust data frame (e.g. in RStudio) that will accommodate the relevant operational and external parameters for the particular system, and importing the available time-series datasets into data frame. **Task 3** will apply the tools of time-series data analysis, such as cross-correlation analysis, and fitting the temporal evolution of the data to current models of performance changes for the chosen system. The real value of data science is in its ability to detect signatures from large datasets that are not as obvious as those illustrated in Figures 1 and 2.

Benefits to Members: A long-term goal of such research would be to be able to use advance warning of critical external conditions, or of distinctive changes in non-environment-driven operational parameters, to forestall a system failure. Knowing what patterns of operating conditions degrade the performance of a power generating or storage system would enable designers to build protection against such conditions into the system from the start. Ideally, since critical external circumstances cannot be completely foreseen and prevented, knowing the signature responses of the system to such conditions could allow remote detection of impending problems and initiation of preventive measures, ultimately prolonging the service life of the system.

Timeline:

Tasks	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**
1) Obtaining datasets	■	■						
2) Develop data frame		■	■	■				
3) Application of data analytical tools				■	■	■	■	■

Time Series Analysis of Power-Generating Systems in the Field

Mark R. De Guire

Case Western Reserve University

Department of Materials Science and Engineering

Proposed Project Duration: 2 years

Industrial Relevance and Novelty

- ***Reliability limits implementation*** of renewable/advanced power and storage technologies
- Lab conditions don’t capture conditions of in-field service
- Need to know:
  - What conditions matter most?
  - What degradation can be reduced by design?
  - What critical conditions can be ***diagnosed*** and ***averted*** in real time?
- Goals:
  - ***Improve reliability***
  - ***Reduce cost***
  - ***Increase deployment***

MATERIALS DATA SCIENCE RELY

Proposal Objectives

1. ***Assemble in-field time-series data*** from electrical systems
   - Power — fuel cells, photovoltaics, wind …
   - Storage — batteries, electrolyzers, …
2. ***Correlate*** performance with
   - Service conditions — current, voltage, demand
   - External conditions — weather, malfunctions, disruptions

- Short-term:
  - ***Recognize drivers*** of degradation
  - ***Understand loss processes***
- Long-term: ***Diagnose*** and proactively ***avert*** degradation

Proposal Deliverables

For the selected systems:

- Methodologies for analyzing large, multi-variable, time series, and other types of datasets
- Design rules for improved reliability
- Strategies for diagnosis and remediation in real time

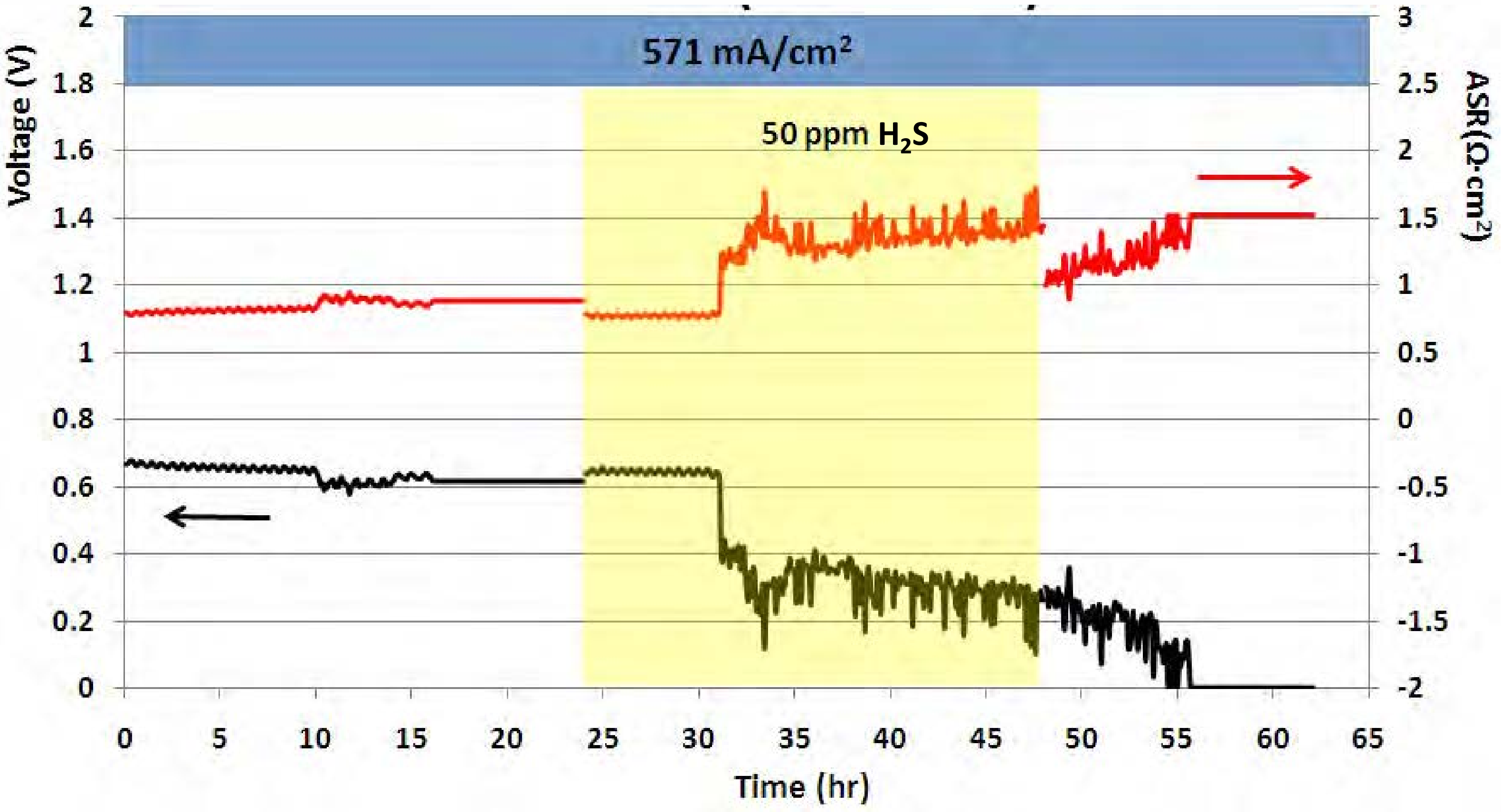
Proposal Timeline

Task / Milestone	Year 1				Year 2			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
1) Access in-field datasets								
2) Develop data frame(s)								
3) Apply data analytical tools								
4) Replicate key disruptions in lab simulations								

Timeline can be expanded
as more systems or datasets become available

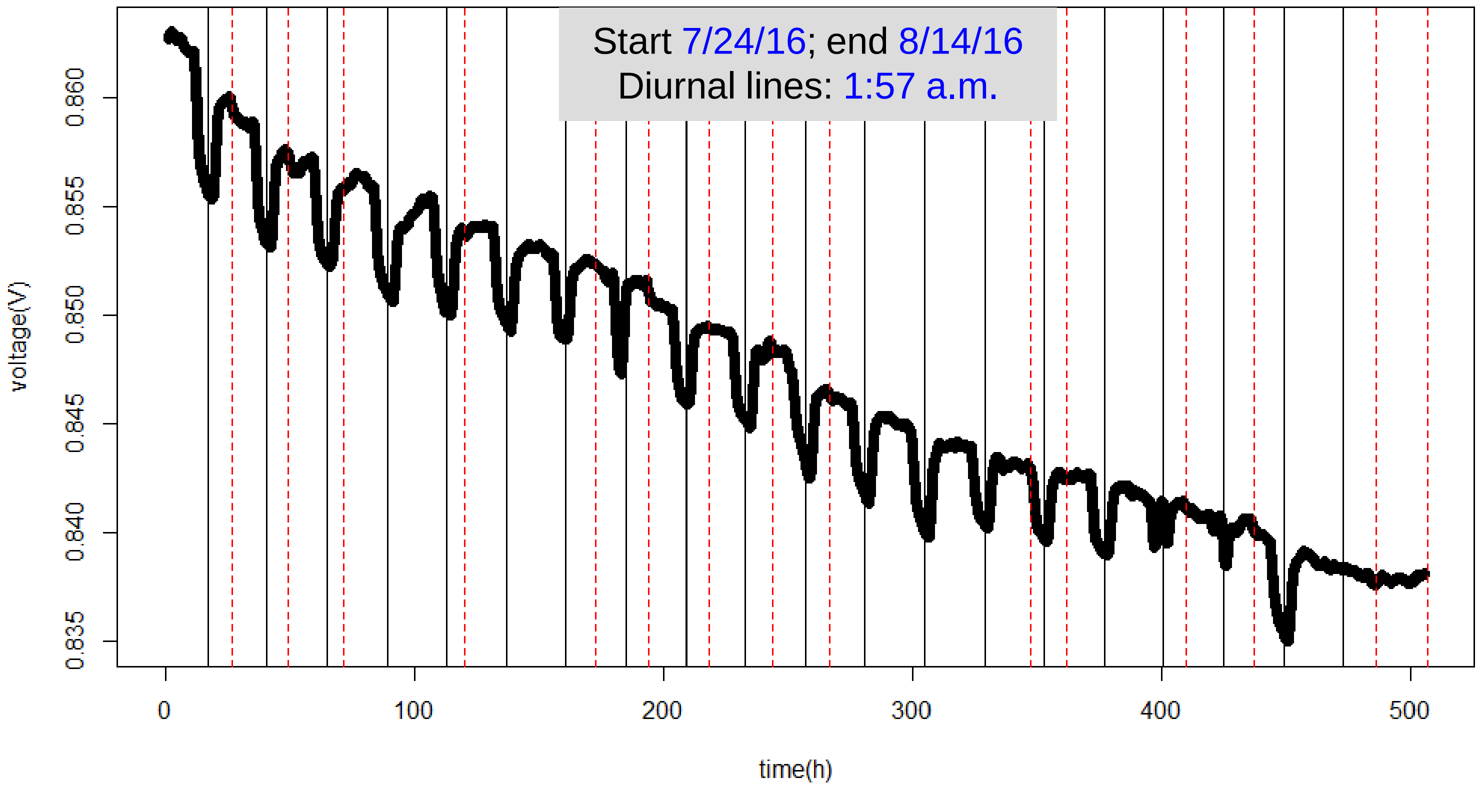
Example: solid oxide fuel cells

- ***Monitor*** symptom of impending failure: erratic ***voltage drops***
- ***Diagnose*** cause: ***inadequate exhaust*** of H_2O
- ***Remediate*** (not achieved here)
  - Remove obstruction
  - Hold cells at open-circuit voltage, 24 h



Example: solid oxide fuel cells

- ***Small daily drops*** in output voltage (2 a.m.)
- Working to ***diagnose cause***: monitor ambient T, humidity, barometric pressure
- ***Is it a degradation mechanism?***
  - Data analytics being applied: cross-correlations; filters

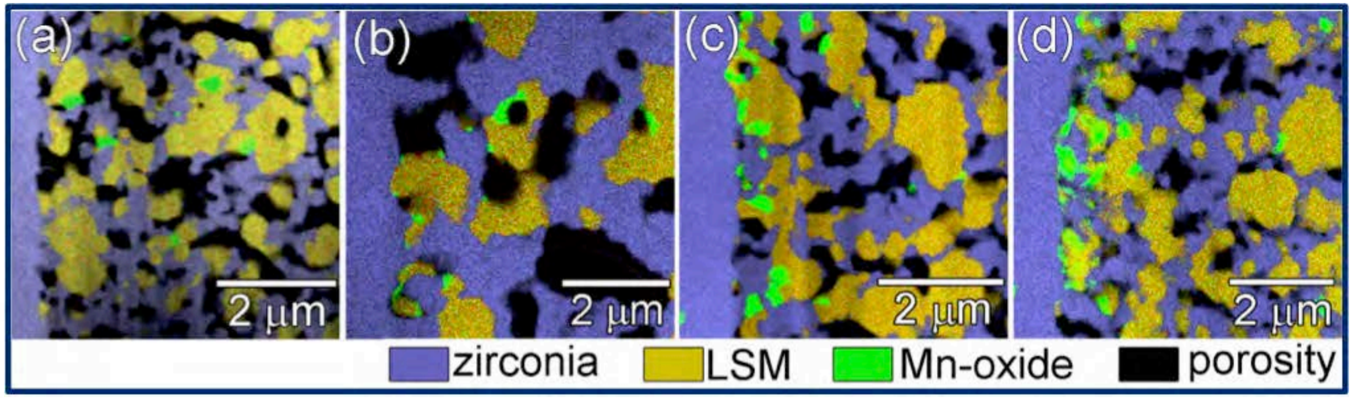


MATERIALS DATA SCIENCE RELY

Facilities / Leveraged Technology / Other Funding Sources

For solid oxide fuel cells:

- Ongoing analysis of lab-scale tests up to 624 h
- Extensive time-series datasets (NASA, DoE funding)
- Fuel cell testing lab
- Microstructural analysis facilities and expertise



Thank You!

MDS-Rely Project Proposal
(RS Area C - Thrust Area 2 - Proposal Number TBD)
(September 2019)

Project Title: Image Machine Learning for Process Control	
Principal Investigator(s): James McGuffin-Cawley	**Researcher:** TBA
New Project: XX Renewal: **Term: 1 year, 2 years**	**Start Date: Jun 2020**

Thrust Area: 1. Study Protocols, **2. Matls. Data Science**, 3. Reliability Studies (choose one)

Objective: The goal of this project will be to apply existing strategies for image processing to two classes of images, thermograms and optical micrographs, for purposes of process control. This will impose tight timeliness constraints on processing. Identified opportunities exist in in primary production operations for metals and alloys as well as in both casting and deformation processing. In addition, we anticipate evolving a well-established industry standard for qualifying tool steels to use this work to develop a new standard for, for example, additively manufactured tooling casting and molding.

Standards Used: For optical imaging the following ATSM standards will be consulted and the relevant subset employed: ASTM-E930: Test Methods for Estimating the Largest Grain Observed in a Metallographic Section (ALA Grain Size), 2007; ASTM-E1245: Practice for Determining the Inclusion or Second-Phase Constituent Content of Metals by Automatic Image Analysis, 2008; ASTM-E1382: Test Methods for Determining Average Grain Size Using Semiautomatic and Automatic Image Analysis, 2010; ASM-E3: Guide for Preparation of Metallographic Specimens, 2011. For thermography the following will be used: ASTM E2582-19 Standard Practice for Infrared Flash Thermography of Composite Panels and Repair Patches Used in Aerospace Applications; MIL MIL-HDBK-731 Military Standardization Handbook; Nondestructive Testing Methods of Composite Materials – Thermography; ASTM E1934-99a(2018) Standard Guide for Examining Electrical and Mechanical Equipment with Infrared Thermography; ASTM E1311-14(2018) Standard Practice for Minimum Detectable Temperature Difference for Thermal Imaging Systems.

Code Developed: Anticipated codes will integrate with unique experimental arrangements to process image generated in real-time and execute commands to alter process variables to control and improve process yield.

Datasets Produced: Set of images that characterize the range of process output from successful to unsuccessful in multiple dimensions will be generated. Initially, focus will be on optical micrographs from production runs of primary metals producers and thermograms generated in-house on-campus in the Case Metal Processing Laboratory. Tabulated results of analysis will be generated that define the precision required of the image analysis, the uniqueness of the interpretation of results, and the rate of computation correlated to the demands of the process stream.

Background: The manufacture of metal components continues to evolve in fundamental ways. For example, over the last two decades there has been a fundamental shift from mild steel to advanced high strength steels [1]. This offers advantages in weight reduction, improved safety, and power delivery redesign. However, the processing window becomes tighter due to the need to reliably produce metastable phases within the metal microstructure, reduce inclusion count, and achieve tight dimensional specifications. The consequences are substantial – for example, the yield point of steels can be altered by as much as a factor of three during press hardening [2]. These development impose requirements on sheet metal producers to achieve "clean room" level standards for steel cleanliness (reduction in nonmetallic inclusions). Processes exist for control through chemistry (equilibration with given slag chemistry), fluid flow (limitation of exposure to atmosphere, turbulence in contact with refractories or slag), residence time, etc. [4,5] and truly impressive gains have been achieved. Improvements have been so dramatic that new approaches to characterization have had to be developed. For example, the performance of the steel is frequently limited by the behavior at the extreme values of the distribution rather than the details of the distribution near the mean or mode [6,7]. Methodologies for characterizing steel cleanliness have become sophisticated, but remain typically offline. A personal communication from an engineer in steel production at a large global firm comments on the state-of-the-art "online/in-process steel cleanliness analysis is something that the industry dearly desires… [Existing] automated sample analysis… that is offline… has proven a useful tool for new advanced high strength steel grade development… [allowed development of] ultra clean, very high strength steels (up to 1180 MPa tensile commercially, and up to 2000 MPa in development)… [However, the "quicker" feedback at about a week turnaround time if rushed.. [and t]he ultimate answer… usually take 6-8 weeks...[Yet w]e are making a heat about every 40 minutes." There is on-going work directed at automating image analysis to document inclusions [8.9]. Furthermore, extraction of in-process samples for metallography is standard practice in producers. The challenge and opportunity is uniting these.

An example of successful application of machine learning in metal processing is machine learning applied to anomaly detect in thermal images of diecasting tooling in a production setting [10]. Die casting involve forcing molten metal into a permanent die under high-pressure. Control variable include melt temperature, velocity of injection, cavity geometry, peak load, dwell time, and ejection temperature. It was recognized that the thermal information that is used currently is indirect (e.g., cooling water temperature) and it was hypothesized that the 2D thermal map of the die face must reflect the local conditions that give rise to at least some of the conditions that lead to unusable parts (i.e., scrap). This was shown to be the case. Mounted thermal cameras collected full field 2D thermograms each process cycle (typically under two minutes). Machine learning was used to define the oscillatory steady state, including the range of noise that is seen under "stable" processing conditions, including, for example, variations due to water draws in other parts of the plant that would affect the cooling water in a given machine. A three-step algorithm was developed that including both anomalies associated with local excursions (hot or cold spots) and drift of baseline, which was validated. On-going work is being directed at converting this to production tool. Applications include simple documenting the thermal images associated with the fabrication of a given serialized part for later engineering-epidemiology, displaying thermal image with color-coded ID's for particular anomalies to alert operators who would otherwise be unaware of changes, and an alert/auto-stop for a well-defined set of anomalies. These workers have already employed their algorithm to analysis of transmission x-rays of die cast parts to analyze the distribution of porosity and search for informative anomalies in these results. Extension to inclusion analysis is logical and straightforward. One key variable they identified was precision in fixturing. A key advantage of this work is that "a subject matter expert can invest a small amount of time understanding and classifying clusters…" and that "…selected thresholds can be highlighted…[such that they are] easily visualized by humans."

In the initial phases of this project, it is proposed to adapt an existing national standard test for thermomechanical fatigue testing to thermal cycling of additive manufacturing tooling. This represents an emerging technology offering a qualitatively new dimension of control to molders in general and metal casters in particular. This will serve as a well-controlled laboratory test to validate codes and image handing protocols.

Project Tasks:

- Consultation with industrial partners to identify points of opportunity, processes of high-value, and areas of persistent problems. Development of tables of 2D images that typify the results of a typical range of process variables.
- Algorithm development and adaption, making use of others work to the extent possible.
- Test of the ability to define the distribution of registered anomalies, including, especially, their extremes.
- Laboratory tests, including the thermomechanical fatigue to validate codes – intentionally probing unusual conditions to ensure robustness.
- Results from in-plant case studies.

Benefits to Members:

- Identification of new applications of well-established algorithm to effect process.
- Development of image analysis into a practical process control tool.
- Development and characterization of lab test appropriate for process control development qualification testing.
- Protocols for adapting approaches and algorithms to an array of 2D image fields.

Timeline :

Tasks	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**
Industrial consultation								
Algorithm literature review								
Algorithm development and validation								
Laboratory testing								
In-plant trials								

References

1. Frahm, Bill. “Mastering fundamentals opens opportunities for advancement, *The Fabricator*, August 15, 2016 https://www.thefabricator.com/article/metalsmaterials/mastering-fundamentals-opens-opportunities-for-advancement (accessed Sept 8, 2019)
2. Löbbe, Christian, Oliver Hering, Lars Hiegemann, and A. Tekkaya. "Setting mechanical properties of high strength steels for rapid hot forming processes." *Materials* 9, no. 4 (2016): 229.
3. Zhang, Lifeng, and Brian G. Thomas. "State of the art in evaluation and control of steel cleanliness." *ISIJ international* 43, no. 3 (2003): 271-291.
4. Yoon, B-H., K-H. Heo, J-S. Kim, and H-S. Sohn. "Improvement of steel cleanliness by controlling slag composition." *Ironmaking & steelmaking* 29, no. 3 (2002): 214-217.
5. Zhang, Lifeng, Subo Yang, Kaike Cai, Jiying Li, Xiaoguang Wan, and Brian G. Thomas. "Investigation of fluid flow and steel cleanliness in the continuous casting strand." *Metallurgical and materials transactions b* 38, no. 1 (2007): 63-83.
6. Beretta, Stefano, and Yukitaka Murakami. "Largest-extreme-value distribution analysis of multiple inclusion types in determining steel cleanliness." Metallurgical and Materials Transactions B 32, no. 3 (2001): 517-523.
7. Ekengren, Jens, and Jens Bergström. "Extreme value distributions of inclusions in six steels." Extremes 15, no. 2 (2012): 257-265.
8. Harris, Marc, Obinna Adaba, Simon Lekakh, Ron O’Malley, and Von Richards. "Improved methodology for automated SEM/EDS non-metallic inclusion analysis of mini-mill and foundry steels." In *AISTech Proceedings*, pp. 3315-3325. 2015.
9. Slattery, G. F., P. O'Riordan, M. E. Lambert, and S. M. Green. "An integrated electron and optical metallographic procedure for the identification of precipitate phases in type 316 stainless steel." *Journal of Microscopy* 122, no. 2 (1981): 109-130.
10. Blondheim, D. Jr. “Unsupervised machine learning and statistical anomaly detection applied to thermal images,” Proceedings of the North American Die Casting Association Die Casting Congress & Exposition, 2018.

Image Analysis for Process Control

Faculty Member(s): J. McGuffin-Cawley

University / Department: Materials Science & Engineering

Proposed Project Duration: 2 years

MATERIALS DATA SCIENCE RELY

Industrial Relevance and Novelty

The technical view of manufacturing is a sequence of unit operations that control the geometry (shape) of a part, and/or the material properties, and/or the condition of the surface.

The economic view is a series of steps that incrementally add value to the part, component, or assembly.

Yield (i.e., not producing scrap (unusable or unsalable parts) is a key component of profitability.

Bohn, Roger E., and Christian Terwiesch. "The economics of yield-driven processes." *Journal of Operations Management* 18, no. 1 (1999): 41-59.

Industrial Relevance and Novelty

Image analysis in medicine is being revolutionized by machine learning and AI in image analysis. Two examples are:

- Thermal imaging using state-of-the art IR camera technology

"…present efforts are focused on automatic analysis of temperature distribution of regions of interest and their statistical analysis for detection of abnormalities…"

Lahiri, B. B., S. Bagavathiappan, T. Jayakumar, and John Philip. "Medical applications of infrared thermography: a review." *Infrared Physics & Technology* 55, no. 4 (2012): 221-235.

Industrial Relevance and Novelty

Image analysis in medicine is being revolutionized by machine learning and AI in image analysis. Two examples are:

- Optical micrographs of prepared sections of tissue

  …algorithms have begun to be developed for disease detection, diagnosis, and prognosis prediction to complement to the opinion of the pathologist…

Gurcan, Metin N., Laura Boucheron, Ali Can, Anant Madabhushi, Nasir Rajpoot, and Bulent Yener. "Histopathological image analysis: A review." IEEE reviews in biomedical engineering 2 (2009): 147.

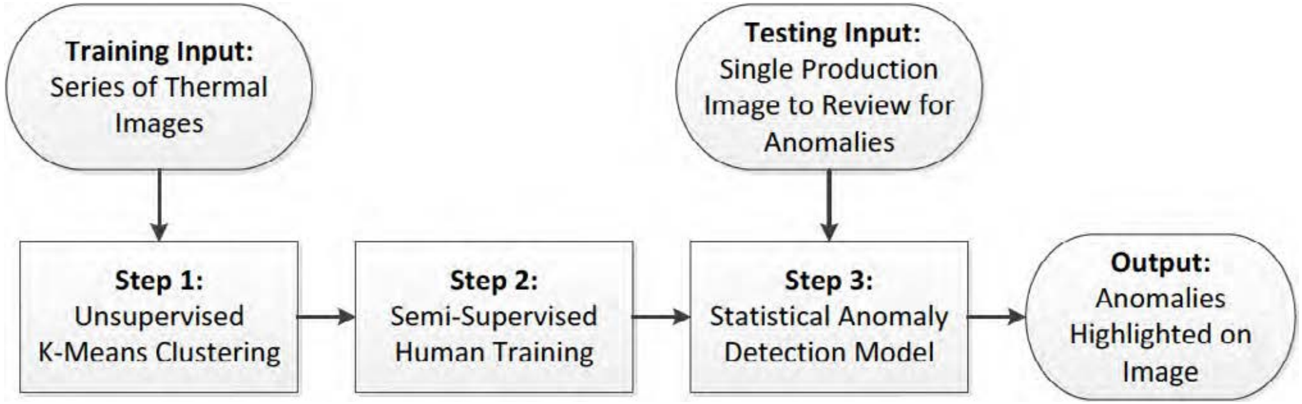
Janowczyk, Andrew, and Anant Madabhushi. "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases." Journal of pathology informatics 7 (2016).

Industrial Relevance and Novelty

Emerging applications in metal production and forming

- thermal imaging using state-of-the art IR camera technology

Figure 3 – Thermal Image Camera System Installed on Die Cast Machine



Blondheim, D. Jr. “Unsupervised machine learning and statistical anomaly detection applied to thermal images,” Proceedings of the North American Die Casting Association Die Casting Congress & Exposition, 2018.

Industrial Relevance and Novelty

Emerging applications in metal production and forming

- thermal imaging using state-of-the art IR camera technology

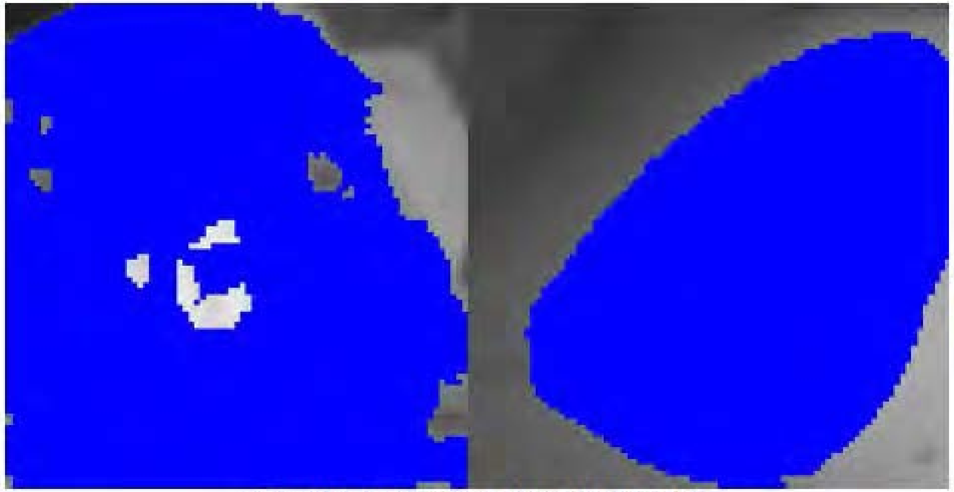


Figure 12: Anomaly Detection on Warmup Shot

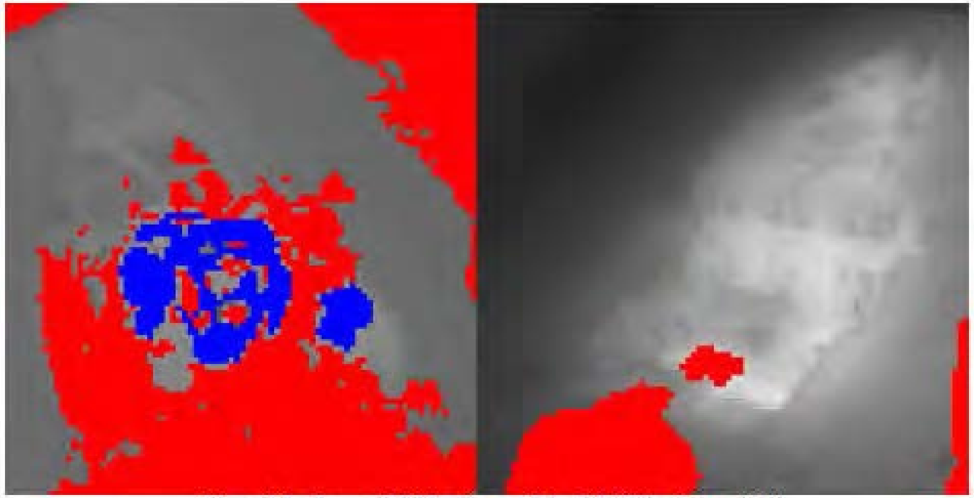


Figure 13: Anomaly Detection on Die with Water Turned Off

Blondheim, D. Jr. “Unsupervised machine learning and statistical anomaly detection applied to thermal images,” Proceedings of the North American Die Casting Association Die Casting Congress & Exposition, 2018.

Industrial Relevance and Novelty

Emerging applications in metal production and forming
- Documenting the thermal images associated with the fabrication of a given serialized part for later engineering-epidemiology
- Displaying thermal image with color-coded ID's for particular anomalies to alert operators who would otherwise be unaware of changes
- An alert/auto-stop for a well-defined set of anomalies.

Key advantage: "a subject matter expert can invest a small amount of time understanding and classifying clusters…" and that "…selected thresholds can be highlighted…[such that they are] easily visualized by humans."

Blondheim, D. Jr. "Unsupervised machine learning and statistical anomaly detection applied to thermal images," Proceedings of the North American Die Casting Association Die Casting Congress & Exposition, 2018.

Industrial Relevance and Novelty

Logical extensions for thermal imaging:

- Metal forging
- High-speed sheet metal forming
- Deformation processing of tailor-welded blank
- Performance of conformally cooled dies & tools
- Application of adhesives
- Curing of coatings

Blondheim, D. Jr. "Unsupervised machine learning and statistical anomaly detection applied to thermal images," Proceedings of the North American Die Casting Association Die Casting Congress & Exposition, 2018.

Industrial Relevance and Novelty

Assessing steel cleanliness through on-line optical microscopy

“online/in-process steel cleanliness analysis is something that the industry dearly desires...

[Existing offline] automated sample analysis allowed development of] ultra clean, very high strength steels (up to 1180 MPa tensile commercially, and up to 2000 MPa in development)...but

Feedback... usually take 6-8 weeks...but, heat about every 40 minutes.”

Extraction of in-process samples for metallography is standard practice in producers. The challenge and opportunity is uniting these.

Personal communication, Manager - Steel Producing Quality

MATERIALS DATA SCIENCE RELY

Proposal Objectives

- Identification of new applications of well-established algorithms to effect process control.
- Development of image analysis into a practical process control tool.
- Development and characterization of lab test appropriate for process control development qualification testing.
- Protocols for adapting approaches and algorithms to an array of 2D image fields.

Proposal Deliverables

- Consultation with industrial partners to identify points of opportunity, processes of high-value, and areas of persistent problems. Development of tables of 2D images that typify the results of a typical range of process variables.
- Algorithm development and adaption, making use of others' work to the extent possible.
- Result of laboratory tests, including the thermomechanical fatigue to validate codes – intentionally probing unusual conditions to ensure robustness.
- Results from in-plant case studies.

MATERIALS DATA SCIENCE RELY

Facilities / Leveraged Technology

- Case Metal Processing Lab
  Thermomechanical test facility
- North American Die Casting Assoc.
- America Makes

Aluminum bath: 380 Al Casting Alloy

- (~125 lbs)
- Maintained at 732 °C (1350 °F)

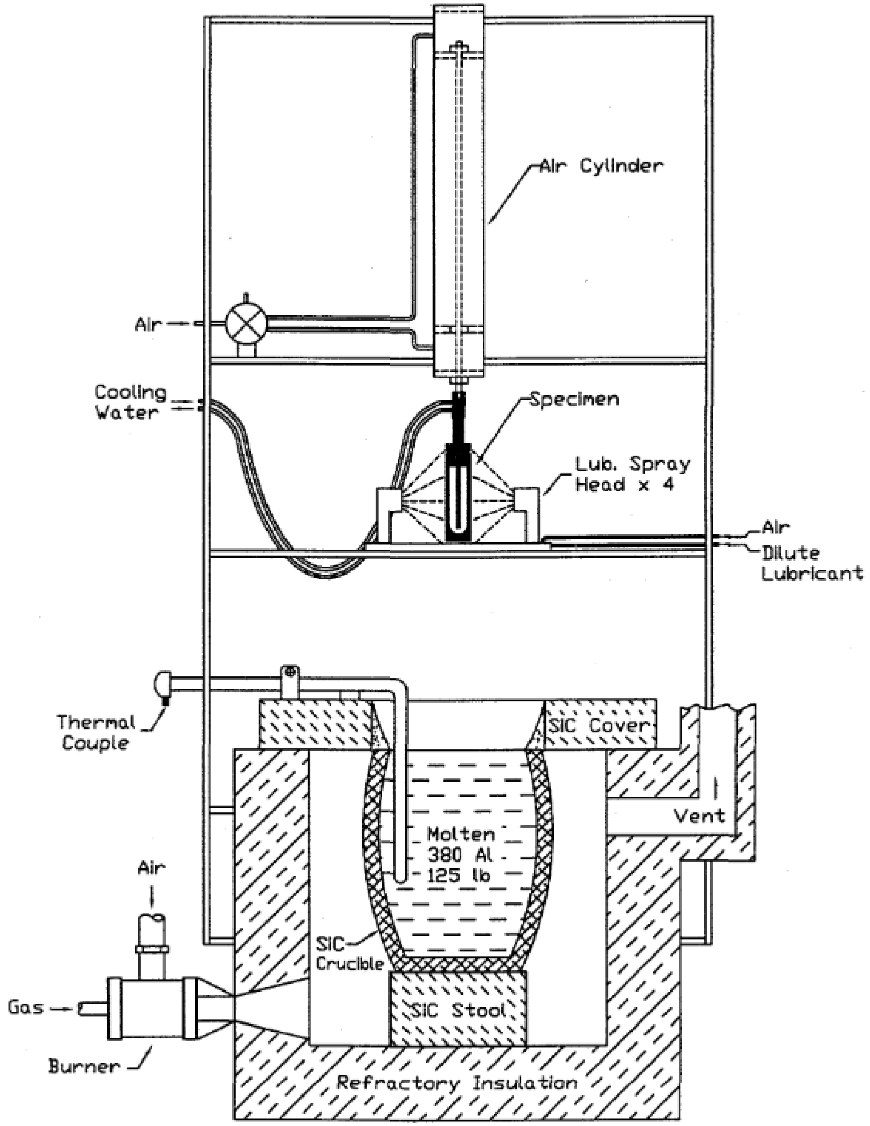


Figure 4.5. Thermal Fatigue Test Equipment

MATERIALS DATA SCIENCE RELY

Proposal Timeline

Tasks	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Industrial consultation								
Algorithm literature review								
Algorithm development and validation								
Laboratory testing								
In-plant trials								

MATERIALS DATA SCIENCE RELY

Risks / Countermeasures

Risk	Countermeasure(s)
Data acquisition requires hardware in rough service environment	Work with established vendors for partner plants, characterize environment
High cost	Engineering design in-house, migrate to vendors
Noise in data	Algorithm improvement, signal conditioning

Thank You!

Project Proposal
(September 2019)

Project Title: Reliability of Binder Jet 3D Printed Inconel 625: Effect of Binder Burnout

Principal Investigator(s): Markus Chmielus

New Project: X

Thrust Area: 3. Reliability Studies

Abstract: As additive manufacturing (AM) technologies become more widely studied and adapter, it is important to understand the implications of these methods on the material reliability. Binder jet 3D printing (BJ3DP) is one AM method where layer by layer, powder is loosely bound together with a polymer binder to create a fragile green part. After printing, the green part is then sintered at high temperatures to densify, achieving higher than 99% densities. During the sintering process, the polymer binder burns out and the metal powder particles fuse together, eliminating porosity. Compared to other AM methods, the final parts have homogeneous microstructures and no thermal stresses. However, little attention has been given to the effects of the binder burnout on the final microstructure and part reliability. When the binder burns off, excess carbon is left behind, slightly altering the composition of the material that could lead to new phases forming in the material. These phases have the potential of altering the performance of the final parts. Therefore, understanding the burnout process can help to mitigate potential challenges and inconsistencies in BJ3DP materials. Inconel 625 is especially of interest due to its wide range of application and use in the AM industry.

Project proposal Tasks: Thermogravimetric analysis on various AM binders, 3D printing and sintering of parts, surface treatment of parts, microstructural analysis and mechanical testing of final parts. This project will develop standards for sintering with optimized binder burnout steps and provide implications of binder burnout on final part performance and reliability.

Benefits to Members: Companies would benefit from this project through the implementation of binder jet printing for large scale, fast additive manufacturing and improved reliability of final parts through the implementation of sintering best practices.

Companies: PPG, NASA Glenn, NETL, Lockheed Martin, US Army Armaments Center, ANSYS, Solar Turbines, 3M Company

Short Bio: Dr. Markus Chmielus is an Associate Professor in the Mechanical Engineering and Materials Science Department with a background in Materials Science and Engineering and Aerospace Engineering. Dr. Chmieluss areas of research include the influence of production and processing parameters on the properties and microstructure of metals, carbides and functional magnetic materials produced via additive manufacturing (also known as 3D printing) and ultra-high purity thin film deposition. His work spans a wide range of materials including Alloy 625, 718 and Ti64, WC-Co, magnetic shape memory alloys and magnetocaloric materials. The combining umbrella of all research areas is quantitative characterization of microstructure, defects, mechanical, electrical, magnetic and thermal properties on different length scales using local, national, and international facilities including synchrotron and neutron diffraction.

Reliability of binder jet printed Inconel 625: Effect of binder burnout

Katerina Kimes and Erica Stevens

Proposed Project Details

PI: Dr. Markus Chmielus

chmielus@pitt.edu

University of Pittsburgh

Department of Mechanical Engineering and Materials Science

Proposed Project Duration: 1 year

Inconel 625

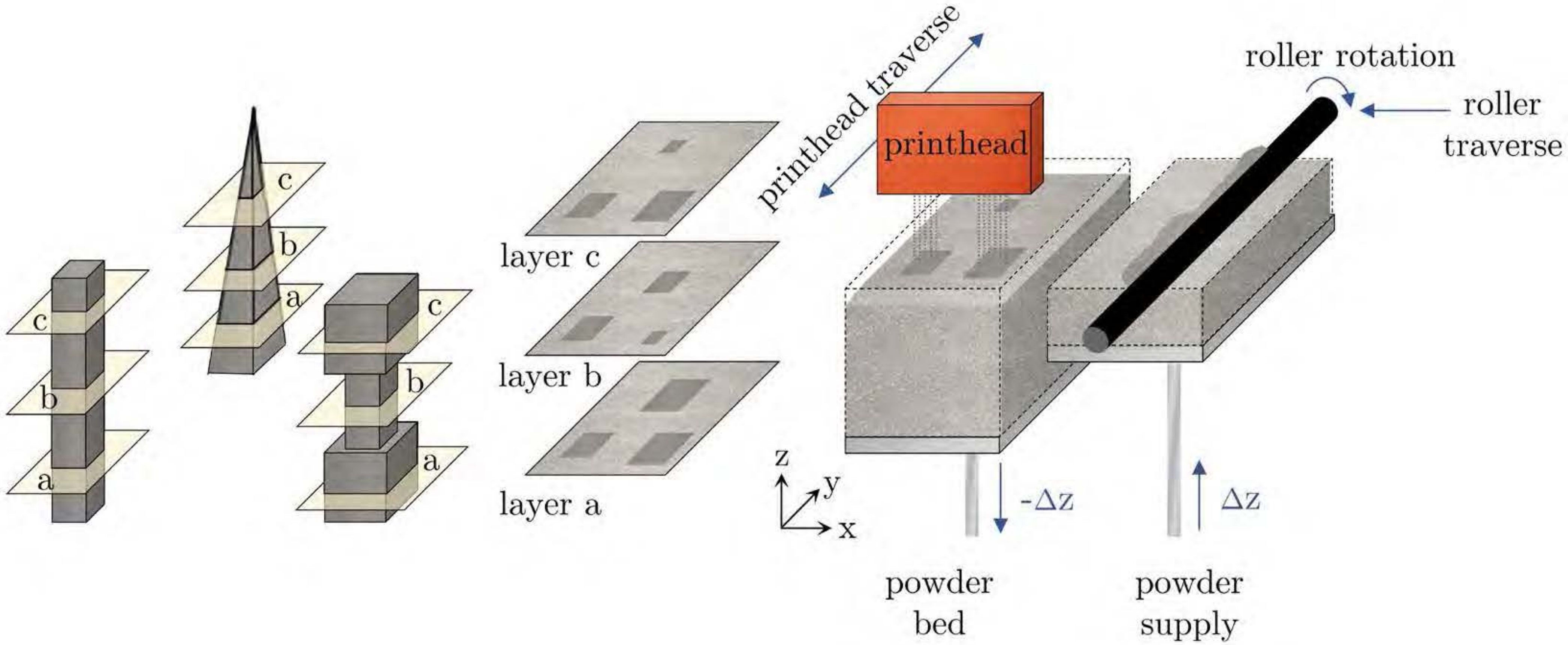
G.P. Dinda, et al., *Mat. Sci. and Eng. A.* **509** (2009) 98-104.

Ni-based superalloy

Used in aerospace, chemical, petrochemical, and marine applications

Diverse applications in wide temperature ranges from cryogenic to over 1000 °C

Binder Jet 3D Printing (BJ3DP)

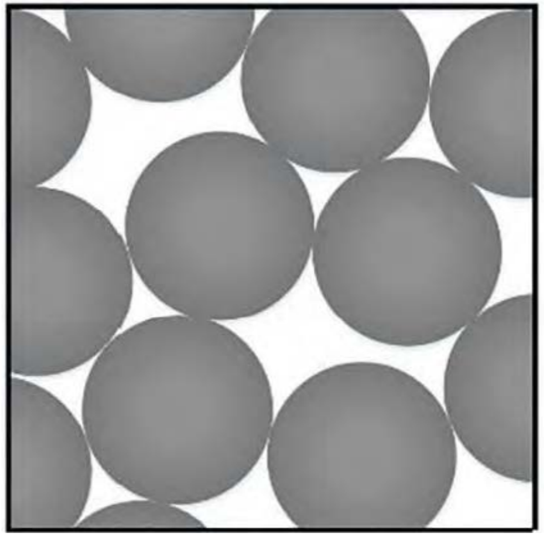
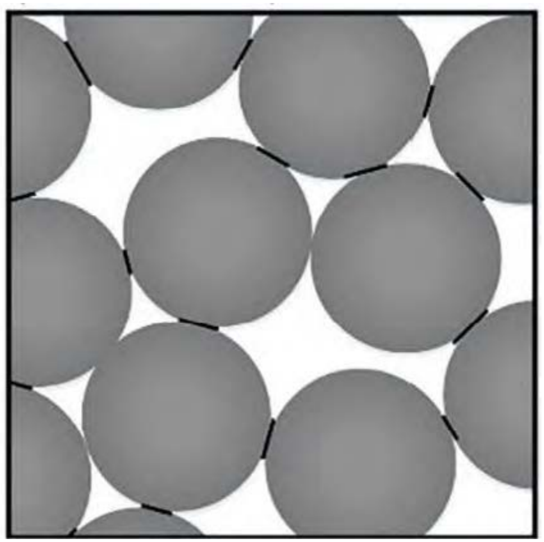
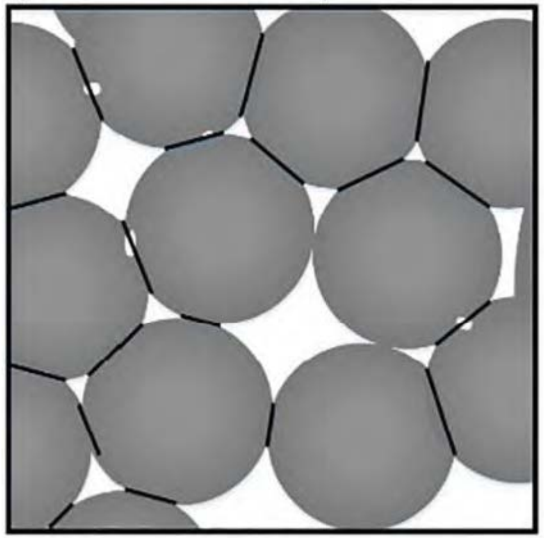
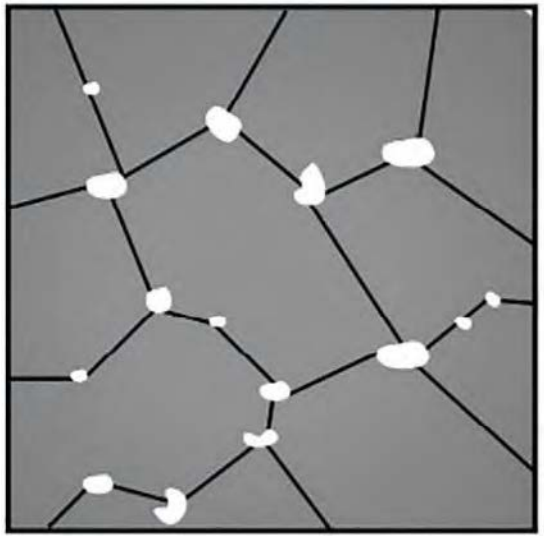


Produces complex parts that are otherwise very expensive to produce due to extensive machining

Binder Jet 3D Printing (BJ3DP)

Sintering

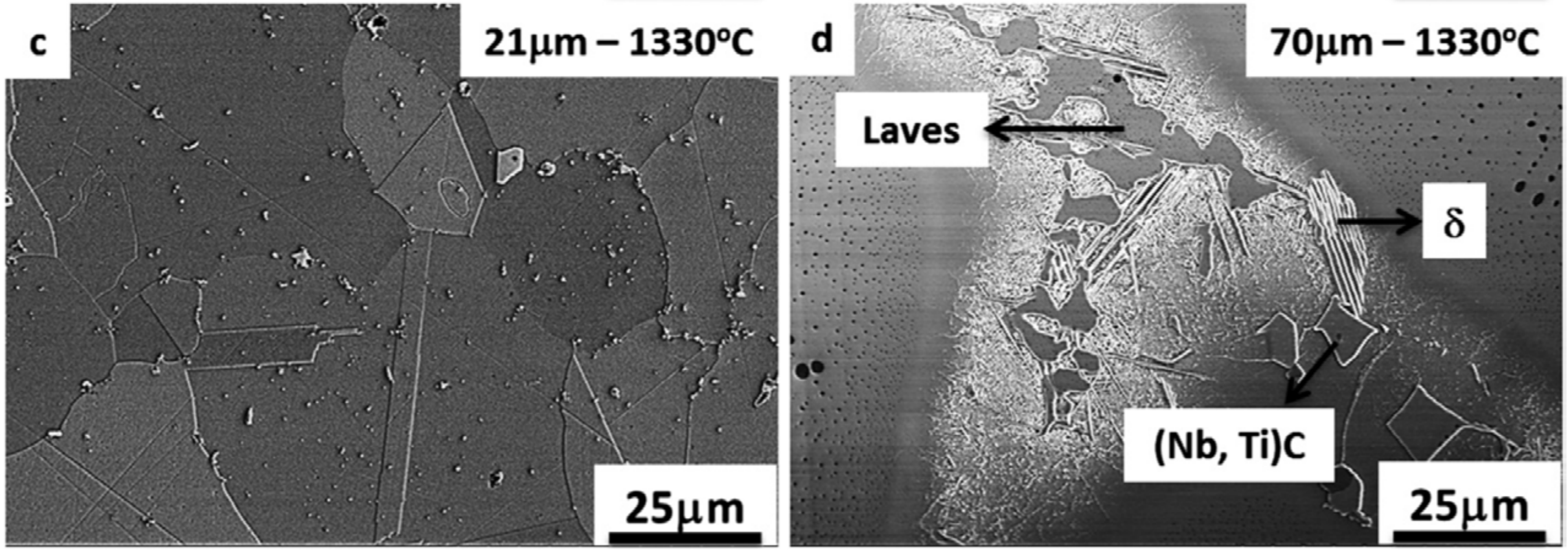
Can achieve final part densities above **99%**

Loose Powder	Initial Stage (Surface Diffusion)	Intermediate Stage (Volumetric Diffusion)	Final Stage
			
40-55% dense	density increases to up to 70%	pores sections split, start closing, density up to 92%	removal of pore sections, density above 99%

Effect of Binder Saturation on Microstructure

Nandwana, et al., Current Opinion in Solid State and Materials Science. 21 (2017) 207-218.

2017 Study on BJ3DP Inconel 718:



High binder saturation

Low binder saturation

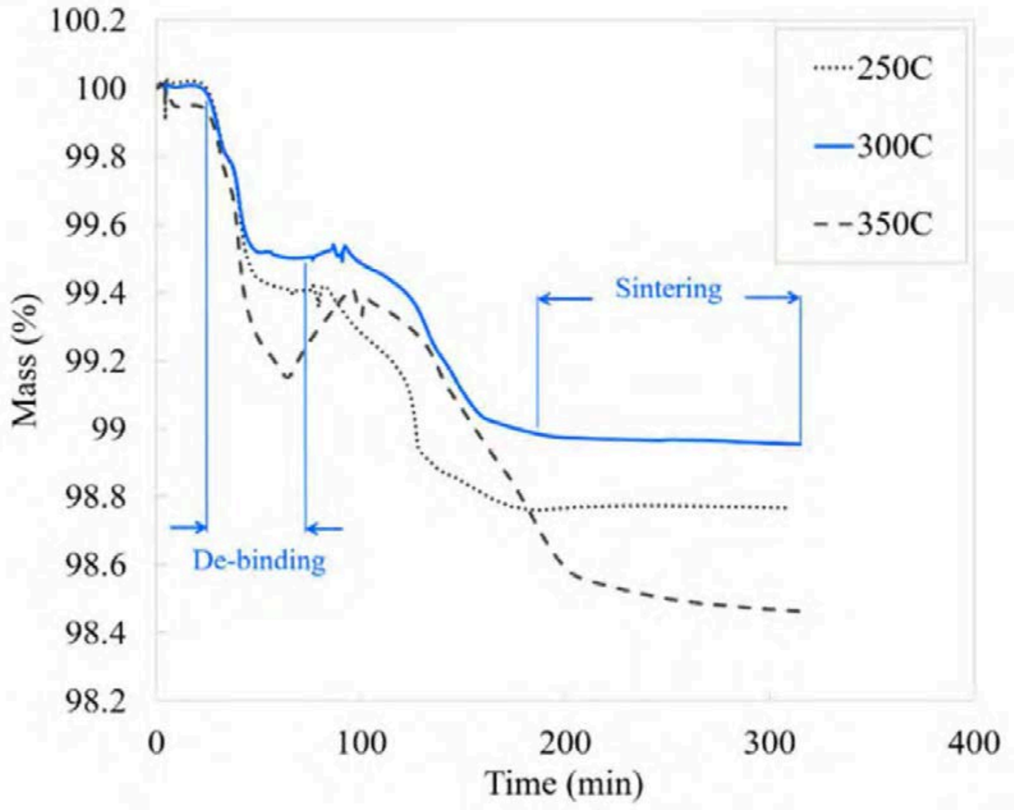
Carbon-rich residue present after burnout

Binder saturation has influence on phases present in microstructure

Effect of Burnout Temperature on Mass Loss

2018 Study on BJ3DP Pure Iron Parts:

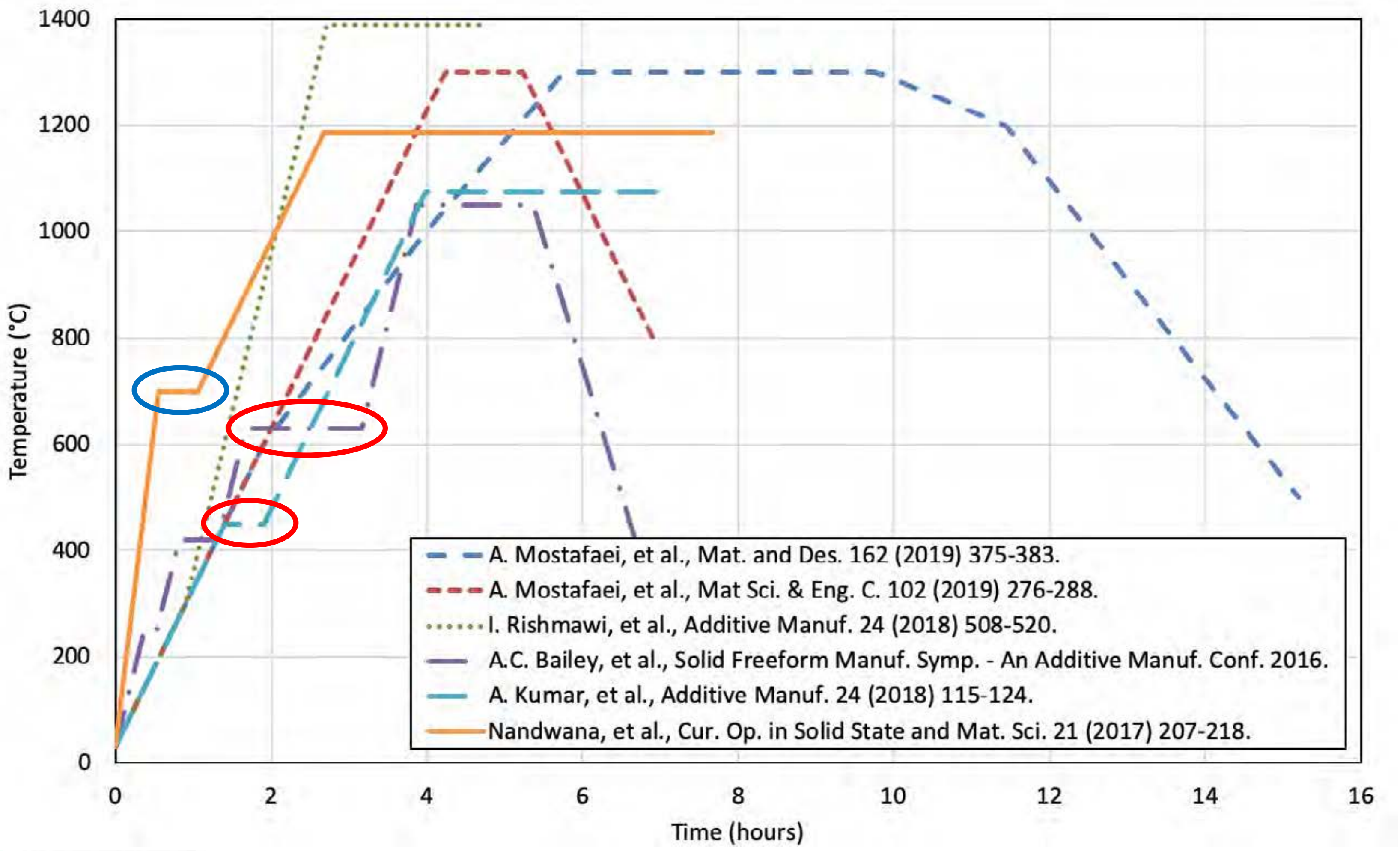
I. Rishmawi, et al., Additive Manuf. 24 (2018) 508-520.



TGA mass loss of green samples for three burnout temperatures

Binder burnout temperature has influence on amount of residual binder

Binder Burnout



Mostafaei: ExOne Solvent
Rishmawi: Z Corporation Aqueous
Bailey: ExOne Solvent
Kumar: ExOne Solvent
Nandwana: ExOne Aqueous

Need for Understanding Binder Burnout

Excess binder in the final parts can influence microstructural phases

↓

Microstructure influences mechanical behavior and reliability of materials

↓

Understanding effects of binder burnout temperature on mechanical properties and reliability

Proposal Objective and Deliverables

Objective: Determine relation between binder burnout parameters and residual impurities on sintered part microstructure and reliability (Inconel 625)

Deliverables:

- Thermogravimetric data on commercially available binder
- Database of processing microstructure-property relations that can be used to extrapolate optimized parameter combinations
- Best practices of binder burnout parameters to maximize reliability (fatigue life) for optimum printing parameters of Inconel 625

Experimental

Evaluate Binder

- Aqueous- and solvent-based binders
- Thermogravimetric analysis

Thermogravimetric data

Printing

- Standard print parameters for Inconel 625
- Fatigue and tensile test specimens and coupons for metallurgical analysis

Sintering

- Various temperatures and hold times for binder burnout

Mostafaei, et al., *Additive Manufacturing.* **24** (2018) 200-209.
Mostafaei, et al., *Acta Materialia.* **124** (2017) 280-289.

Surface Treatment

- Eliminate stress concentrations on surface

Mostafaei, et al., *Additive Manufacturing.* **24** (2018) 200-209.

Characterization

- Microstructural analysis
- Mechanical reliability testing (fatigue)

Microstructure-property relations database

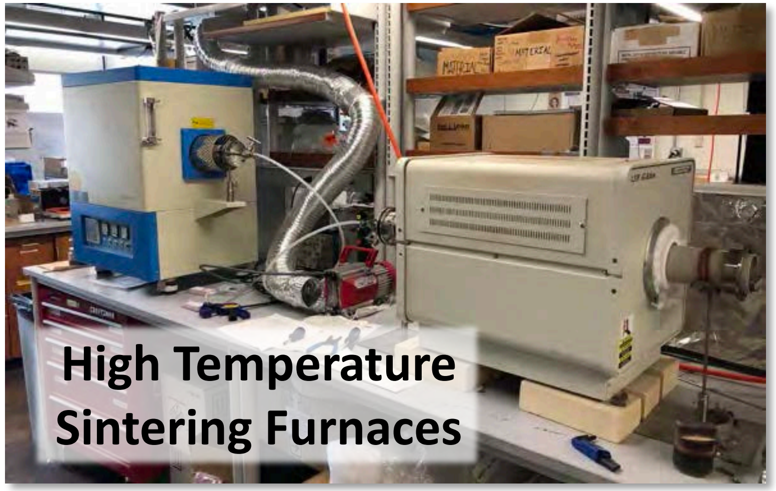
Best practices for optimal reliability

MATERIALS DATA SCIENCE RELY

Proposal Timeline

Task	Expected Duration			
	Q1	*Q2*	*Q3*	*Q4*
Thermogravimetric analysis on binders				
Printing				
Sintering				
Surface Treatment				
Microstructural analysis				
Mechanical testing				

Leverage of Facilities



MATERIALS DATA SCIENCE RELY

NSF

Risks and Countermeasures

Risk	Countermeasure(s)
Difference between thermal properties of binder alone vs. binder in green part	TGA measurements on small samples of green part and comparison with single binders
Inconsistencies of powder batches	Full powder characterization and sieving and remixing of powder to achieve comparable size distributions if needed
Inconsistencies in surface roughness after surface treatment	Test samples with surface profilometry to ensure consistent surface roughness

Outlook

Long Term:

Obtain a comprehensive understanding of binder burnout parameters on the reliability of the most commonly used BJ3DP structural metals.

Additional 2nd and 3rd year objectives:

1. Determine how binder saturation and drying affects burnout time and sintered part microstructure and reliability (Inconel 625)
2. Investigate consistency of results among various structural materials

Thank you

Please submit questions on life form or save for networking breaks

MDS-Rely Project Proposal
(RS Area - Thrust Area - Proposal Number)
(September 2019)

Project Title: Scratch Corrosion Study Protocol and Supervised Image Machine Learning	
Principal Investigator(s): Jennifer Braid	**Researcher: Kunal Rath**
New Project: XX Renewal: **Term: 1 year, X 2 years**	**Start Date: June 2020**

Thrust Area: 3. Reliability Studies

Objective: To extract greater scientific information and insights from the scratch corrosion test, and to make the results more accurate and predictive of the large number of degradation modes observed in the scratch corrosion test.

Develop and validate a photographic-image-based study protocol for the Scratch Corrosion test, and the materials data science codes for image processing and supervised image machine learning.

Standards Used:

1. Test Method for Evaluation of Painted or Coated Specimens Subjected to Corrosive Environments[1]
2. Practice for Operating Salt Spray (Fog) Apparatus[2].
3. Standard Practice for Modified Salt Spray (Fog) Testing[3]
4. Prohesion test for metallic corrosion[4].

Figure 1. A typical panel is

[1] D01 Committee, Test Method for Evaluation of Painted or Coated Specimens Subjected to Corrosive Environments, ASTM International, 2017. doi:10.1520/D1654-08R16E01.

[2] G01 Committee, Practice for Operating Salt Spray (Fog) Apparatus, ASTM International, n.d. doi:10.1520/B0117-18.

[3] G01 Committee, ASTM G85-11 Standard Practice for Modified Salt Spray (Fog) Testing, ASTM International, n.d. doi:10.1520/G0085-11.

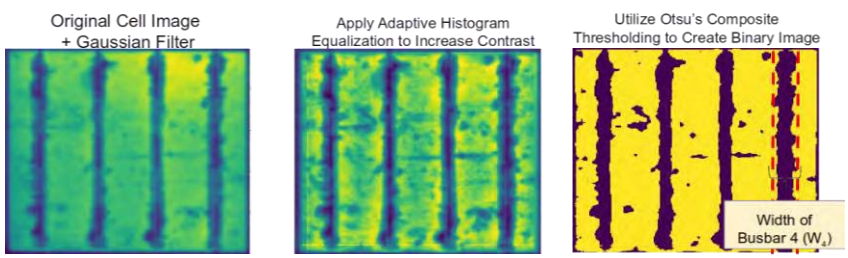
[4] E. Almeida, D. Santos, J. Uruchurtu, Corrosion performance of waterborne coatings for structural steel, Progress in Organic Coatings. 37 (1999) 131–140. doi:10.1016/S0300-9440(99)00064-8.

Code Developed: Image preprocessing and supervised machine learning codes, in a data analysis pipeline, to evaluate the progression of the salt fog, or prohesion, scratch corrosion test. This will build on the supervised image machine learning pipeline we have developed for corrosion of c-Si photovoltaic cells[5].

Datasets Produced: Scratch Corrosion tests will be run on coatings of steel substrates, using salt fog or prohesion spray solutions, and white light camera images taken stepwise through 3000 or 6000 hours. The image dataset of these coatings, taken stepwise through exposure time, and the convolutional neural network model developed and fitted to this image dataset will be made available to member companies.

Background:

Coatings applied to metal structures such as bridges, automobiles or in aquatic environments such as lakes and oceans, are tested routinely with the scratch corrosion test, in which an engineered scratch is scored in the coating, prior to its exposure to either salt or Prohesion[6] mists[7,8,9]. A typical scratch corrosion panel, after six weeks of exposure is shown in Figure 1[10]. These tests are typically rated for 2000 exposure hours, but many tests are run to exposure times of 6000 hours, to check for extreme failures. Yet the standard test is only rated visually, by viewing the corroded scratch and rating it on a scale of 1 to 10. For the amount of time, effort, and cost expended, a highly variable human rating, represents only the ability to identify extreme failures. The goal in this project is to automate the human rating, through photographic image acquisition, and then supervised image machine learning, to rank the coatings on the 1 to 10 scale, while also identifying the many differing degradation modes that are present in the images.



Initially the study protocol for the scratch corrosion exposures will be developed to include image acquisition in a simple photo booth, Figure 2. Images will be

[5] A. M. Karimi, J. S. Fada, M. A. Hossain, S. Yang, T. J. Peshek, J. L. Braid, R. H. French, Automated Pipeline for Photovoltaic Module Electroluminescence Image Processing and Degradation Feature Classification, IEEE Journal of Photovoltaics. (2019) 1–12. doi:10.1109/JPHOTOV.2019.2920732.

[6] What is Prohesion? - Definition from Corrosionpedia, (n.d.). https://www.corrosionpedia.com/definition/1356/prohesion (accessed April 19, 2019).

[7] E. Almeida, D. Santos, J. Uruchurtu, Corrosion performance of waterborne coatings for structural steel, Progress in Organic Coatings. 37 (1999) 131–140. doi:10.1016/S0300-9440(99)00064-8.

[8] D. Li, P. Leroux, Corrosion Resistance of Coating After Scratch Testing, 2016. doi:10.13140/RG.2.1.1566.7606.

[9] T.A. Considine, Standard Operating Procedure for Accelerated Corrosion Testing at ARL, (2017) 14.

[10] Scratching the Surface on Corrosion Testing of Automotive Coatings, (n.d.). https://www.pfonline.com/articles/scratching-the-surface-on-corrosion-testing-of-automotive-coatings (accessed September 5, 2019).

acquired following each 500 hours of exposure, to achieve at least 6 to 9 images across the total exposure time.

Image processing steps for determining scratch width will be developed based on the approach we have used previously for determining the busbar width in corrosion of c-Si photovoltaic cells (Figure 3). The three image processing steps used are shown in in Figure 4[11]. Pixel-based algorithmic image quantification will be used to confirm the responses and sensitivities of the convolutional neural network (CNN) (e.g., scratch width, texture).

Once the images have been planar indexed (to be in comparable perspective and sizes), then they are split in this project into training and testing sets. The initial convolutional neural network will be constructed using Keras[12] and TensorFlow[13,14] from Python 3.6[15] running on a GPU enabled computer. This CNN will be based on the one developed for c-Si cell corrosion shown in Figure 5.

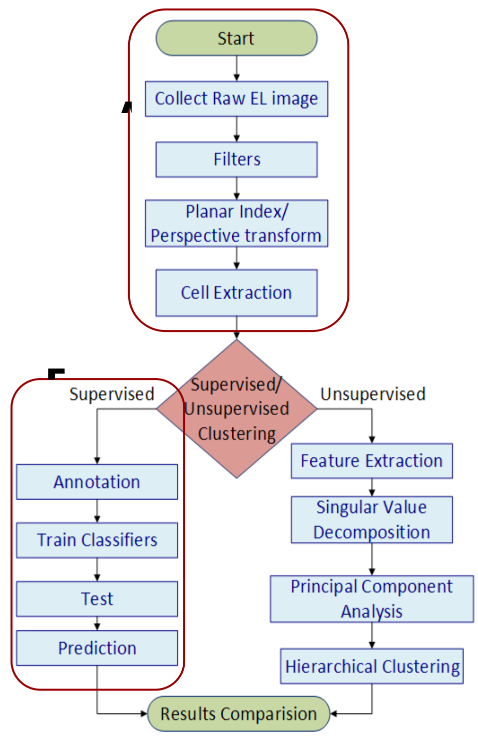


Figure 4. The parts of an integrated image machine learning pipeline, including A)

Project Tasks:

1. Fabricate samples and begin Salt Fog or Prohesion exposures.

2. White light image collection of each scratch before and at 500 hour increments of exposure

3. Image processing, including planar indexing and cropping

4. Training/Testing split and Supervision of images

5. Gage R&R for 3 human test evaluators vs. ImageML rankings of 1-10.

6. Trained Convolutional Neural Net model of Scratch Corrosion, for use by others

Benefits to Members:

By developing an improved study protocol for the scratch corrosion test standard, and combining it with an image processing and machine learning pipeline, we can transform this test into a test that provides improved results, and a greater

[11] Simple photobooth, ~ $60, from Amazon.

[12] F. Chollet, others, Keras, GitHub, 2015. https://github.com/fchollet/keras.

[13] A. M. Karimi, J. S. Fada, J. Liu, J. L. Braid, M. Koyutürk, R. H. French, Feature Extraction, Supervised and Unsupervised Machine Learning Classification of PV Cell Electroluminescence Images, in: 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC), 2018: pp. 0418–0424. doi:10.1109/PVSC.2018.8547739.

[14] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, ArXiv:1603.04467 [Cs]. (2016). http://arxiv.org/abs/1603.04467 (accessed February 24, 2018).

[15] Python Software Foundation: Python 3.6.8 documentation, n.d. https://docs.python.org/3.6/contents.html (accessed June 6, 2019).

amount of information, and degradation mode insights from this common and important test.

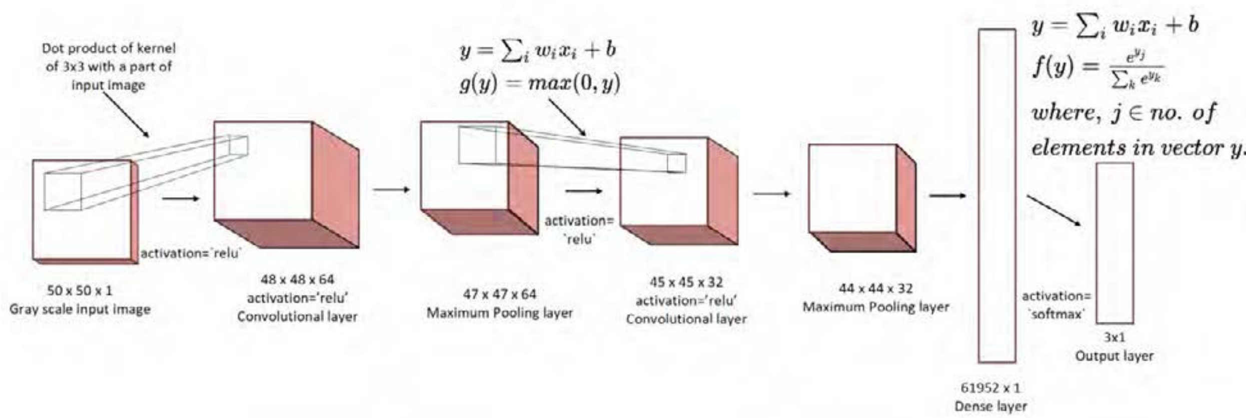


Figure 5. Architecture of the CNN. Input data is an image of dimensions 50 × 50 × 1. The feature maps after the first and second convolutional layer have dimensions of 48 × 48 × 64 and 45 × 45 ×

Timeline :

Tasks	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**
1. Fabricate samples and being Salt Fog or Prohesion exposures.								
2. White light image collection of each scratch at 500 hour increments of exposure								
3. Image processing, including planar indexing and cropping								
4. Training/Testing split and Supervision of images								
5. Gage R&R for 3 human test evaluators vs. ImageML rankings of 1-10.								
6. Trained Convolutional Neural Net model of Scratch Corrosion, for use by others								

Scratch Corrosion Study Protocol and Supervised Image Machine Learning

Faculty Member: Jennifer L. Braid
University / Department: CWRU / MS&E

Proposed Project Duration: 2 years

Industrial Relevance and Novelty

Automated classification algorithms

- to a traditionally manual classification problem

Eliminate operator bias and error in corrosion classification

Mechanism for data input beyond the defined project

- for continuing model improvement

Proposal Objectives

To extract greater scientific information and insights

- From the scratch corrosion test
- Comparing Salt Fog[1] with Prohesion[2]

To make the results more accurate and predictive for

- the large number of degradation modes
- observed in the scratch corrosion test.

Develop and validate a photographic-image-based

- Study protocol for the Scratch Corrosion test

Develop the materials data science codes

- image processing
- supervised image machine learning.

Typical coated panels before (left) following (right) the six-week cycle corrosion test. Some rust run-off was observed from the scribe and stamped panel hole.

1. G01 Committee, ASTM G85-11 Standard Practice for Modified Salt Spray (Fog) Testing, ASTM International, n.d. doi:10.1520/G0085-11.
2. What is Prohesion? - Definition from Corrosionpedia, (n.d.). https://www.corrosionpedia.com/definition/1356/prohesion (accessed April 19, 2019).

MATERIALS DATA SCIENCE RELY

Leveraged Technology

This builds on our PVImage Python Package[1]

- Developed under DOE SETO funding

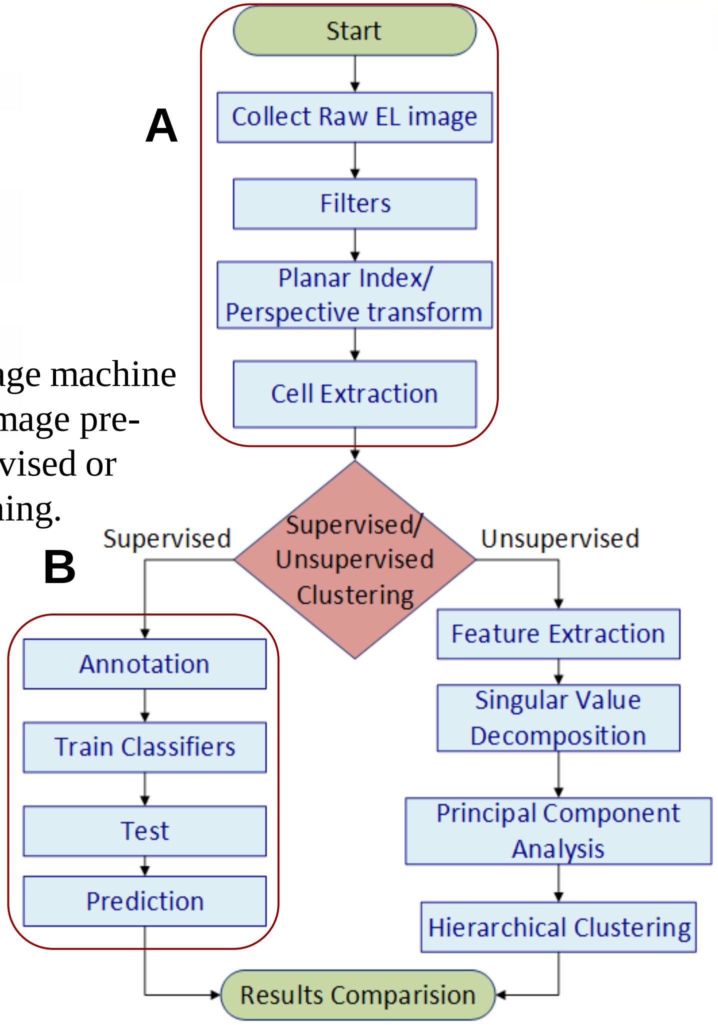
This demonstrated

- A) Image Processing
  - Of PV cell images
- B) Supervised Image ML
  - For c-Si PV cell corrosion
  - Under Damp Heat exposures

Here we will develop

- Improved Study Protocol for Data Acquisition
- Supervised ML for Scratch Corrosion Rating the Test

The three parts of an integrated image machine learning pipeline, including A) image pre-processing, and either B) supervised or unsupervised machine learning.



1. A. M. Karimi, J. S. Fada, M. A. Hossain, S. Yang, T. J. Peshek, J. L. Braid, R. H. French, **Automated Pipeline for Photovoltaic Module Electroluminescence Image Processing and Degradation Feature Classification**, IEEE Journal of Photovoltaics. (2019) 1–12. doi:10.1109/JPHOTOV.2019.2920732.

Time Steps and Exposure Protocol

Exposures using a Salt Fog Chamber with

- Salt Spray (NaCl) or
- Prohesion
  - Ammonium sulphate & sodium chloride
  - More comparable to outdoors for steels

Test is typically done for 2000 hours

- Often extended to 4000, 6000 hours

Employ stepwise exposure and evaluation

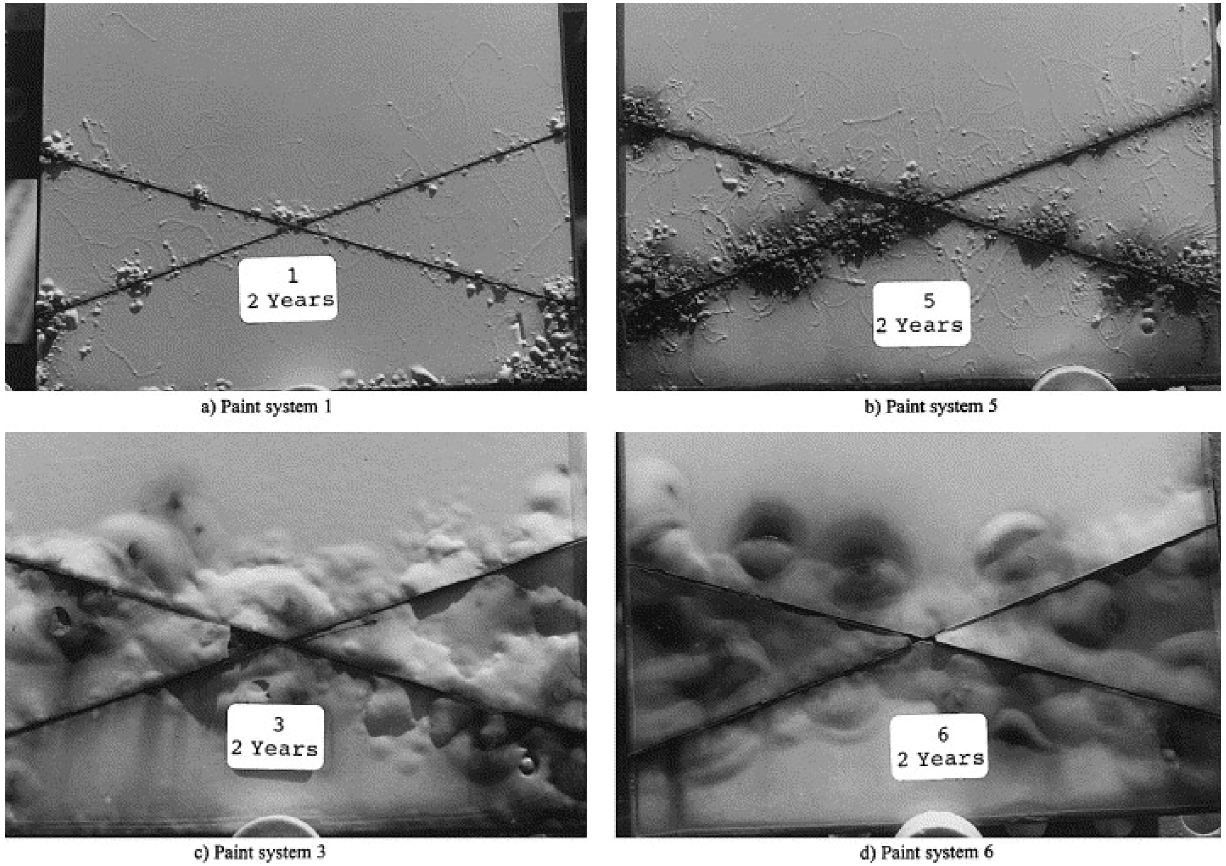
- To observe corrosion trend and rate
- E.g., 500 hour time steps
- Yield 9 step evaluations (images) in 4000 hours

Fig. 2 Examples of a single linear scribe (left) and an X-scribe (right).

MATERIALS DATA SCIENCE RELY

Image processing for scratch corrosion test

1. Planar indexing
   a. to de-tilt and crop image
2. Apply metrics to scratch
   a. Such as scratch width
   b. Scratch length
   c. Texturing
   d. Color changes (rusting)
3. Tag images with exposure time



MATERIALS DATA SCIENCE RELY

Almeida, Elisabete, Dulcínea Santos, and Jorge Uruchurtu. "Corrosion performance of waterborne coatings for structural steel." *Progress in organic coatings* 37.3-4 (1999): 131-140.

NSF

Measuring corrosion line widths

Similar as done for c-Si busbar corrosion

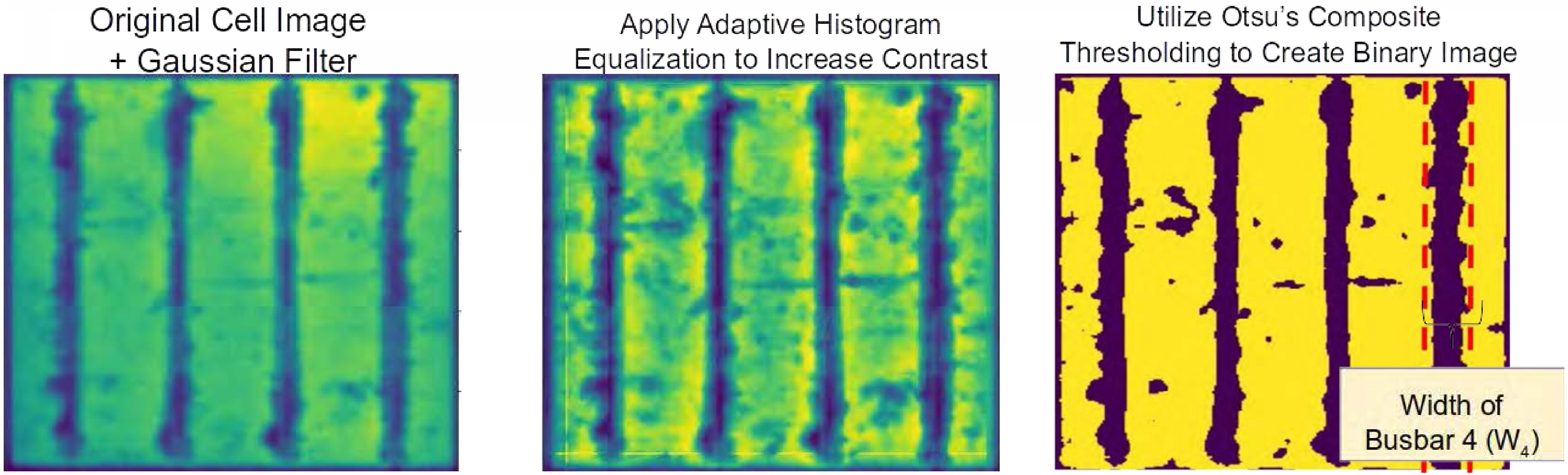


Image processing steps used to determine the width of a corroded busbar in a c-Si photovoltaic cell.

MATERIALS DATA SCIENCE RELY

Supervised Machine Learning Protocol

Supervised classification of images

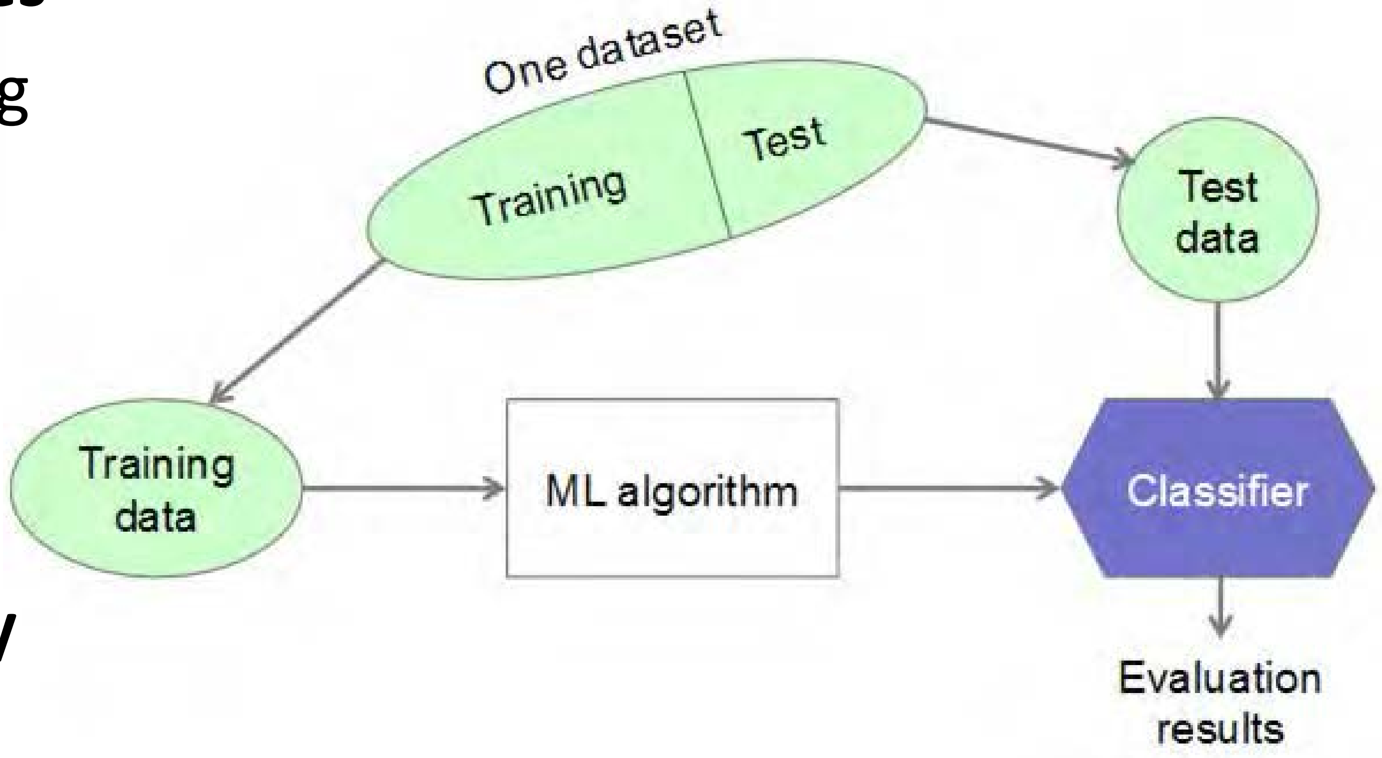
- Manual 1 to 10 corrosion rating

Train the machine

- On the training images

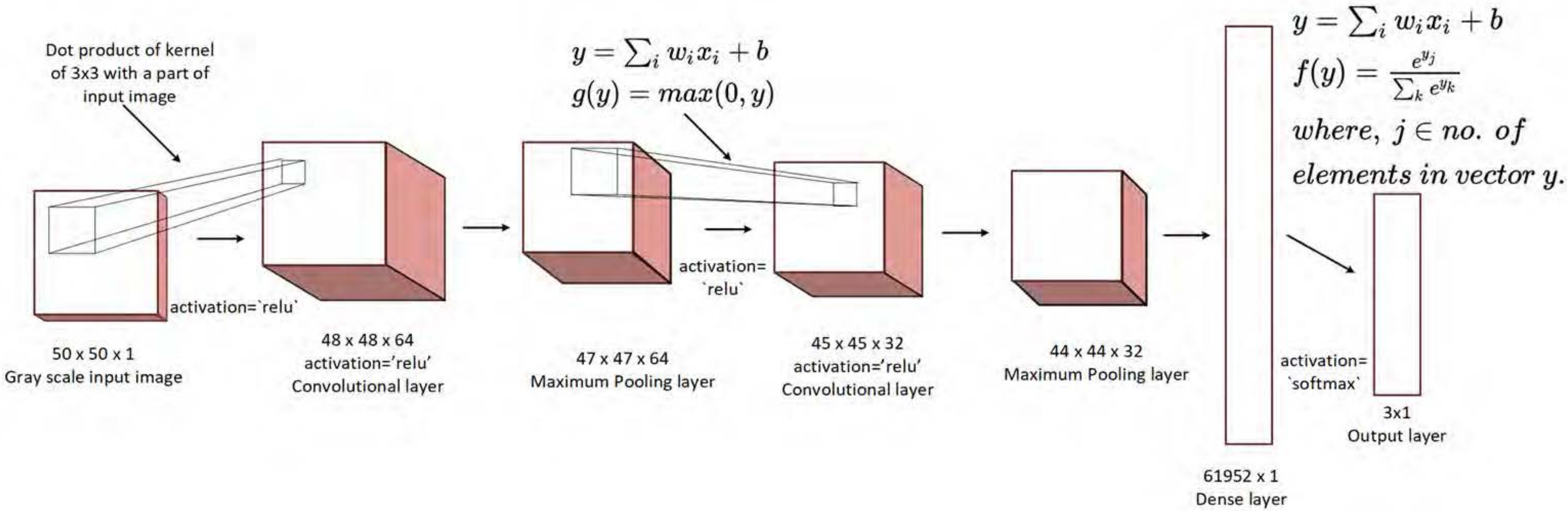
Evaluate performance & accuracy

- On the test images



MATERIALS DATA SCIENCE RELY

Example Convolutional Neural Network Model



Architecture of the CNN. Input data is an image of dimensions 50 × 50 × 1. The feature maps after the first and second convolutional layer have dimensions of 48 × 48 × 64 and 45 × 45 × 32, respectively, formed by convolving sixty four 3 × 3 kernels. The maximum pooling layers are of dimensions 47 × 47 × 64 and 44 × 44 × 32 using a 2 × 2 filter. The dimensions of the dense layer are 61952 × 1, and last layer is the output layer for the three classes. g(y) and f (y) are the ReLU and Softmax activation functions.

An example of the c-Si PV cell corrosion Convolutional Neural Net Model

MATERIALS DATA SCIENCE RELY

Evaluate Various Models for Classification

Convolutional Neural Network

- Using Keras/TensorFlow on GPU

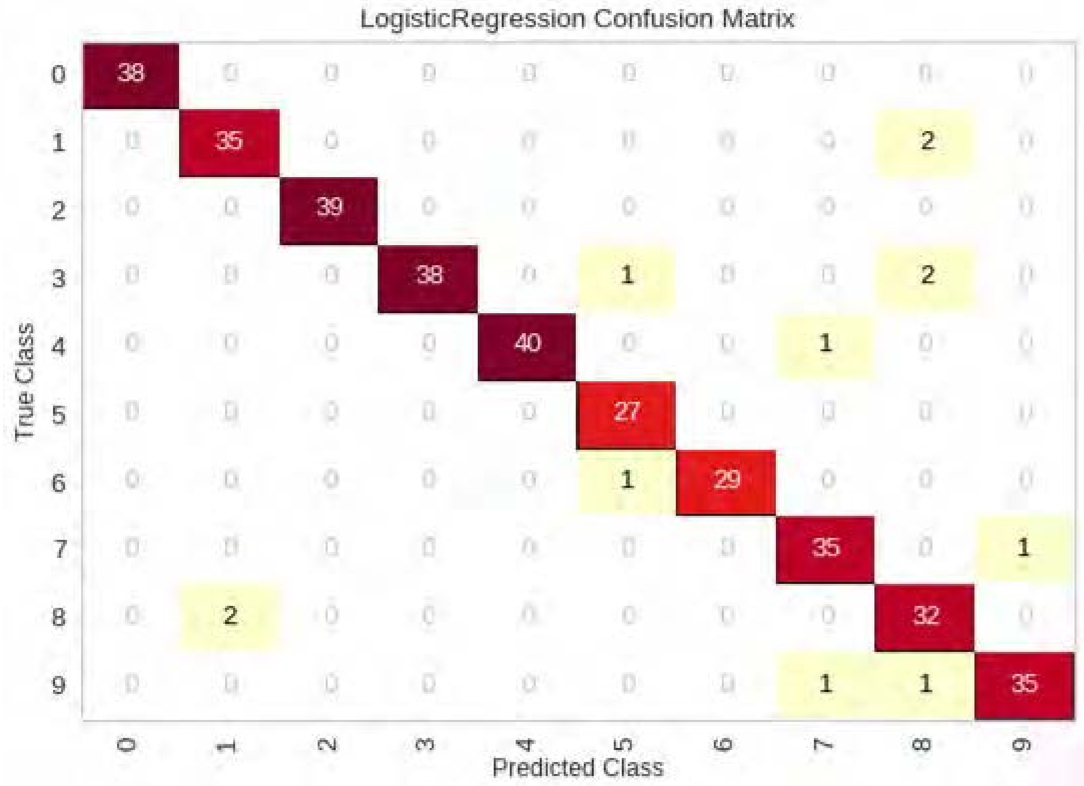
Support Vector Machine

And other models (e.g., clustering)

Compare model results

- with pixel-based quantities
- to validate model responses

Determine Misclassification Error Rates



MATERIALS DATA SCIENCE RELY

Proposal Deliverables

1. Specification of new data acquisition protocol
   a. Photographic images of each sample
   b. Taken at 500 hour steps
   c. Using a photobooth
2. Image Processing Code
   a. To crop and planar index scratch corrosion images
3. Procedures for Supervising a set of test images
4. An initially trained CNN model for scratch corrosion
5. An automated pipeline code package
   a. For scratch corrosion test

Proposal Tasks & Timeline

Tasks	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
1. Fabricate samples and being Salt Fog or Prohesion exposures.								
2. White light image collection of each scratch at 500 hour increments of exposure								
3. Image processing, including planar indexing and cropping								
4. Training/Testing split and Supervision of images								
5. Gage R&R for 3 human test evaluators vs. ImageML rankings of 1-10.								
6. Trained Convolutional Neural Net model of Scratch Corrosion, for use by others								

MATERIALS DATA SCIENCE RELY

Thank You!

MDS-Rely Project Proposal
(RS Area B - Thrust Area 3 - Proposal Number TBD)
(September 2019)

Project Title: Service Lifetime of Polymers	
Principal Investigator(s): Laura S. Bruckman	**Researcher: Menghong Wang**
New Project: XX Renewal: **Term: 1 year, 2 years**	**Start Date:** **June 2020**

Thrust Area: 3. Reliability Studies

Objective: The goal of this project is to develop service lifetime prediction of polymers with data driven models and network degradation models of polymers under various stress conditions and to understand the role of each stressor and stress level. The advantage to this methodology is to develop data driven models of lifetime and to use network models to visualize degradation pathways and to develop a systems of equations that can be used to predict performance within the stressor parameter space. This project would take samples exposed to multiple stressor conditions and stressor levels through time and measure the samples performance and its underlying chemical changes. The project would focus on polymer samples (e.g., urethane, acrylic, silicone, PET). Data driven and netSEM models include a system of equations that can predict a material's lifetime performance in the stressor parameter space.

Standards Used: This project will use ASTM and other standards common in the industry for exposures of samples. Additional exposures will be used to understand the impact of changes in stressors on samples. Samples will be evaluated during exposure at periodic points to gain performance and mechanistic insights into degradation.

Code Developed: A netSEM package has been developed in R to visualize the networks of degradation. Additional different data driven models will be used that currently exist.

Datasets Produced: This project will produce data sets of polymer degradation including performance metrics (e.g., color, mechanical strength, adhesion, cracking) and mechanistic variables (from e.g., FTIR, GC-MS, NMR, spectroscopic techniques). Included will be images of the samples as well as retained sample library of samples from different stages of exposure.

Background: Data driven models can be used to predict the lifetime of materials under various stress conditions. These models go beyond simple linear models or typical acceleration factors. These types of models bring into account the various material grades, stressor and stressor levels as well as the often non-linear behavior of materials degradation where one degradation product needs to build up in a material until a second degradation pathway begins to occur.

Linear regression models are often preferred to quantify relationships between variables of interest. However, in a longitudinal experimental design with multiple coincident observations of variables at multiple points in time, and specifically with multiple number of predictors, model selection is a challenge. Models, in this case, are expected to consider all the variables in a multilevel way and the interactions between them[1]. Considering the vast behavior of material degradation in polymers under multi-factor stressor conditions, multivariate linear regression modeling is a well suited method for lifetime prediction of polymers. Using fixed-effects[2] and mixed-effects modeling[3], in which fixed and random effects are included, can allow for multi-level modeling of variables and determination of contributing factors to degradation[4]. Service lifetime prediction of polymers, performed using results of lab-based standardized tests, is not technically feasible because of the presence of multiple and variable stress conditions in the real world. Owing to their ability to model multilevel structured and longitudinal experimental design, multivariate regression modeling approach, however, can be extended to model real-world scenarios.

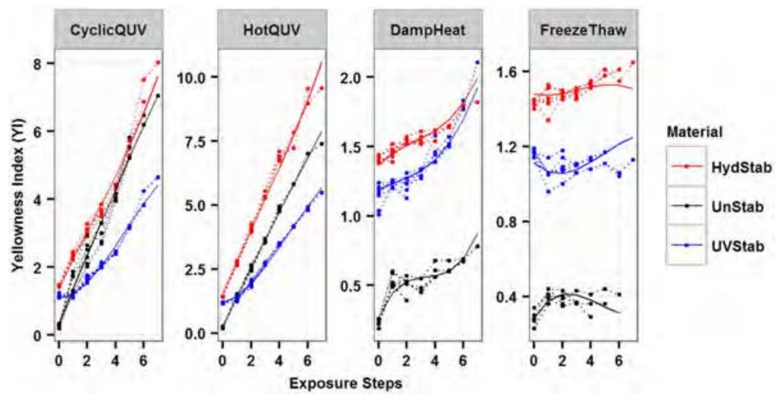


Figure 1: Predictive modeling of change in yellowness index via fixed-effects linear regression modeling. Unstab is unstabilized PET, HydStab is hydrolytically stabilized PET, and UVStab is UV-stabilized PET. PET, polyethylene-terephthalate.

Figure 1 shows a fixed-effects linear regression model of 3 grades of polyethlyene-terephthalate (PET) under 4 different exposure conditions[5]. Equation 1 shows the multi-level model for the change in yellowness index (YI) in PET over 7 exposure steps.

$$\begin{aligned} YI \;&\approx (\beta_0 + \beta_{01}M_1 + \beta_{02}M_2 + \beta_{03}X + \beta_{04}M_1X + \beta_{05}M_2X) \\ &+(\beta_1 + \beta_{11}M_1 + \beta_{12}M_2 + \beta_{13}X)t \\ &+(\beta_2 + \beta_{21}M_1 + \beta_{22}M_2 + \beta_{23}X)t^2 \\ &+(\beta_3 + \beta_{31}M_1 + \beta_{32}M_2 + \beta_{33}X)t^3 \end{aligned} \tag{1}$$

In Equation 1, β's are parameter estimates, t is exposure step, and $M1$, $M2$, and X are as follows:

$$M_1 = \begin{cases} 1 & \text{if Material} = \text{UnStab} \\ 0 & \text{otherwise} \end{cases} \qquad M_2 = \begin{cases} 1 & \text{if Material} = \text{UVStab} \\ 0 & \text{otherwise} \end{cases}$$

$$X = \begin{cases} 1 & \text{if Exposure} = \text{HotQUV or FreezeThaw} \\ 0 & \text{otherwise} \end{cases}$$

SEM[6] is a common technique used in social sciences to derive and map casual relationships between latent (unobserved) variables from observable variables through mathematical systems of equations. However, traditional SEM uses linear models to evaluate these relationships and lacks quadratic, exponential, and logarithmic models that are naturally observed in physical and chemical processes. For this reason, the traditional SEM methodology has been adapted to introduce nonlinear models and exploratory data analysis in order to determine variables related to particular degradation mechanisms. This methodology, referred to as netSEM, is available as an open source code package in R[7,8,9]. netSEM is semi-supervised, whereby domain knowledge is used to supervise the stepwise variable selection and model development, and generalized so as to incorporate nonlinear models in the analysis. This netSEM is highlighted in Figure 2 where the different stages of acrylic degradation occur. The strengths of the relationships between stressors and mechanistic variables are give by ⍰ in a system of equations (REFERENCE COSSMS).

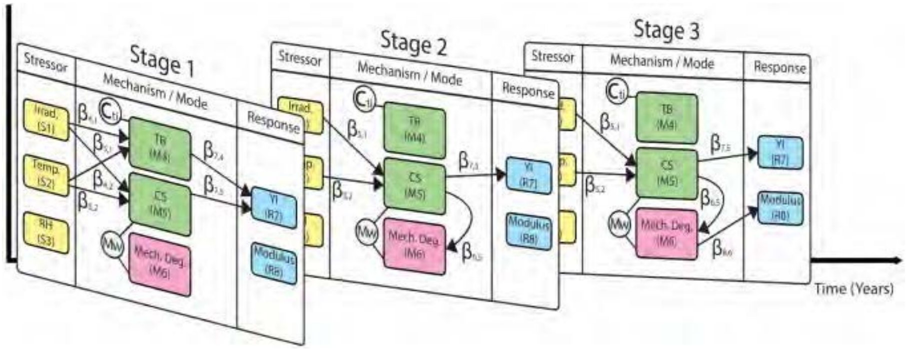


Figure 2: A network of degradation stages of acrylic. Stage 1 includes bleaching of the light stabilizers. Stage 2 shows chain scission of the polymer backbone and Stage 3 shows the mechanical integrity of the polymer changing.

netSEM models of PET degradation under the same temperature and irradiance, but with spray or no spray present in the exposure conditions are shown in Figure 3 where the exposure with water spray (A) shows very little surface degradation because the degraded surface is removed and the exposure without water spray (B) shows significant degradation of the PET[10]. The surface removal of PET was confirmed with excitation-emission fluorescence spectroscopy of PET and with PARAFAC analysis[11].

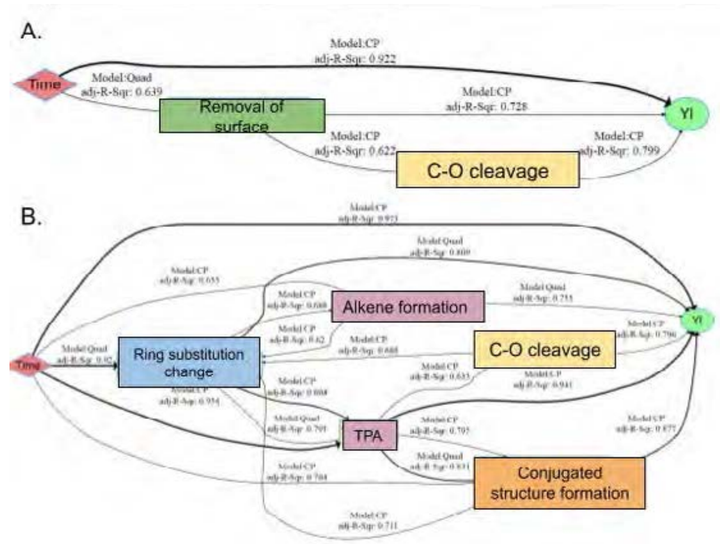


Figure 3. netSEM degradation pathway of PET with water spray (A) and without water spray (B) at 0.8 W/m²/nm at 340 nm and 80℃.

Project Tasks:

1. Develop degradation study protocol
   a. Polymer or set of polymers
   b. Exposure conditions and time steps in exposure
   c. Performance and mechanistic variables to follow
2. Expose polymer sets in exposures conditions and measure performance and mechanistic variables
3. Develop data driven and netSEM models through exposure time with additional time steps improving lifetime prediction

Benefits to Members: The benefit of the members is to further develop the methodology for lifetime prediction of polymer materials so that standard testing can add scientific value on lifetime prediction as well as to give early indicators of new materials that will not have sufficient lifetime early in the tests by identifying key early degradation that occurs in materials.

References

1. A. Gelman, J. Hill, Data Analysis Using Regression and Multilevel/Hierarchical Models, 2006. https://doi.org/10.1017/CBO9780511790942.
2. P.D. Allison, Fixed Effects Regression Models, SAGE Publications, 2009. ISBN 978-1-4833- 8927-1. Google-Books-ID: 3UxaBQAAQBAJ.
3. R.H. Baayen, D.J. Davidson, D.M. Bates, Mixed-effects modeling with crossed random effects for subjects and items, Journal of Memory and Language 59 (4) (2008) 390e412. https://doi.org/10.1016/j.jml.2007. 12.005.
4. J. Fox, Applied Regression Analysis and Generalized Linear Models, third ed., SAGE Publications, Inc, Los Angeles, 2015. ISBN 978-1-4522-0566-3.
5. Gok A, Ngendahimana DK, Fagerholm CL, French RH, Sun J, et al. (2017) Predictive models of poly(ethylene-terephthalate) film degradation under multi-factor accelerated weathering exposures. PLOS ONE 12(5): e0177614. https://doi.org/10.1371/journal.pone.0177614.
6. K.B. Natasha, S. Guo, Structural equation modeling, in: Social Work Research Methods, first ed., Oxford University Press, Oxford;
7. L.S. Bruckman, N.R. Wheeler, J. Ma, E. Wang, C.K. Wang, I. Chou, et al., Statistical and domain analytics applied to PV module lifetime and degradation science, IEEE Access 1 (2013) 384e403. https://doi.org/10.1109/ACCESS. 2013.2267611 .
8. R.H. French, R. Podgornik, T.J. Peshek, L.S. Bruckman, Y. Xu, N.R. Wheeler, et al., Degradation science: mesoscopic evolution and temporal analytics of photovoltaic energy materials, Current Opinion in Solid State & Materials Science 19 (4) (2015) 212e226. https://doi.org/10.1016/j.cossms.2014.12.008.
9. W.H. Huang, N. Wheeler, A. Klinke, Y. Xu, W. Du, A.K. Verma, et al., netSEM: Network Structural Equation Modeling, 2018.
10. Bruckman, Laura, et al. *Backsheets: Correlation of Long-Term Field Reliability with Accelerated Laboratory Testing*. No. DOE-UL-0007143. Renewable Energy, Underwriter's Laboratories, Northbrook, IL 60062, 2019.
11. Gordon, Devin A., Zhonghao Zhan, and Laura S. Bruckman. "Characterizing the weathering induced degradation of Poly (ethylene-terephthalate) using PARAFAC modeling of fluorescence spectra." *Polymer degradation and stability* 161 (2019): 85-94.

Service Lifetime of Polymers

Faculty Member(s): Laura S. Bruckman

University / Department: CWRU / MS&E

Proposed Project Duration: 2 years

Proposed Project Budget:

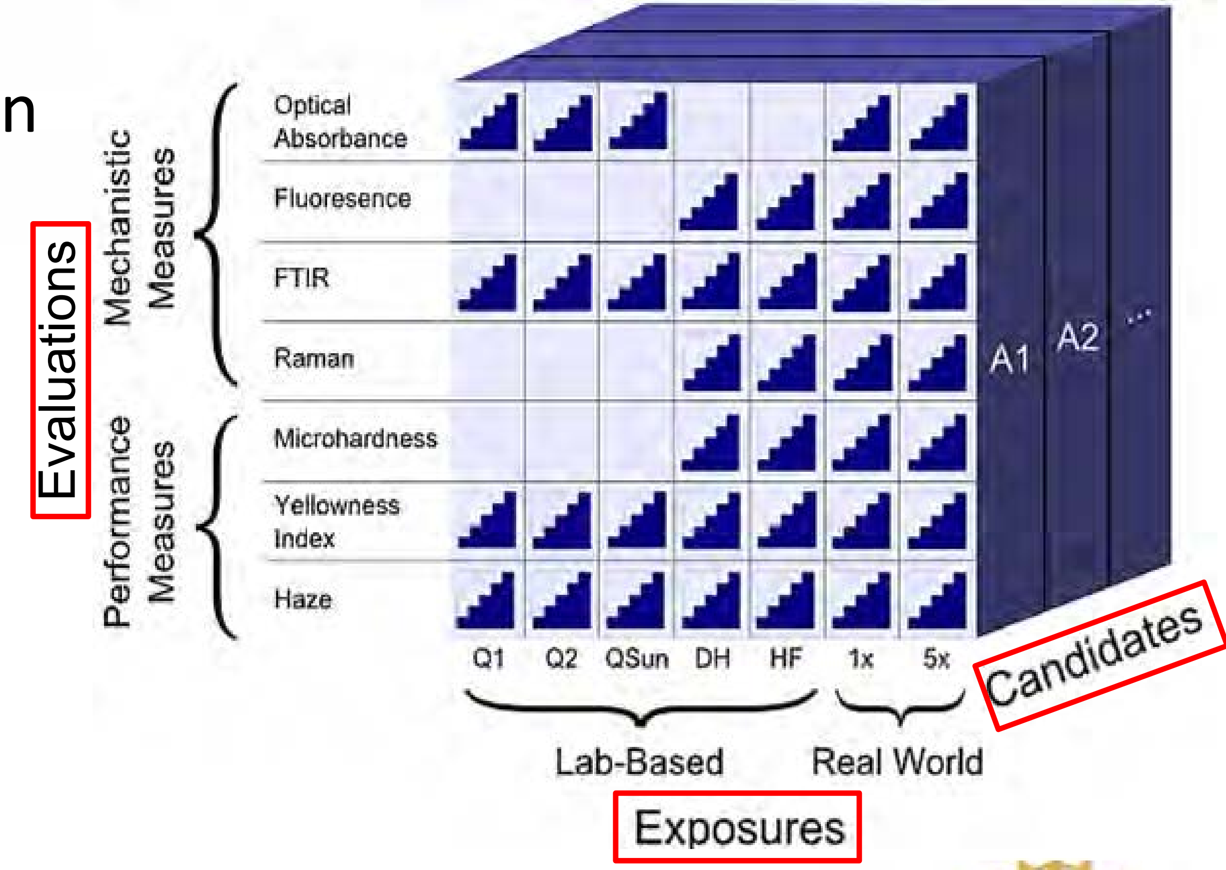
Year 1: $ 60k , Year 2: $ 60k

Total: $ 120k

Industrial Relevance and Novelty

Study Protocols and Materials Data Science

- To determine service lifetime prediction
  - Polymers
- Capture more information from
  - standard testing
  - by monitoring samples periodically
  - mechanistic measurements
- Model behavior using
  - data-driven models
  - network degradation models (netSEM)



MATERIALS DATA SCIENCE RELY

Proposal Objectives

Develop Models for Lifetime Prediction of Polymers

- data-driven models
- network degradation models (netSEM)
- cross-correlation of multiple stressors

Extract greater scientific information and insights

- standard exposure conditions for polymer materials
- identify failing materials early in testing

Improve the current material science data codes

- network degradation models
- cross-correlation

Leveraged Technology

This builds on our netSEM R Package[1]

- Developed under UL funding[2]
- Improved under DOE SETO funding

This demonstrated

- netSEM network models
- system of equations to model

Combined with data driven modeling

- data sets to give insights

Here we will develop

- Improved study protocol for lifetime prediction
- Lifetime prediction models for polymers in various conditions

Stress | Mechanism | Response

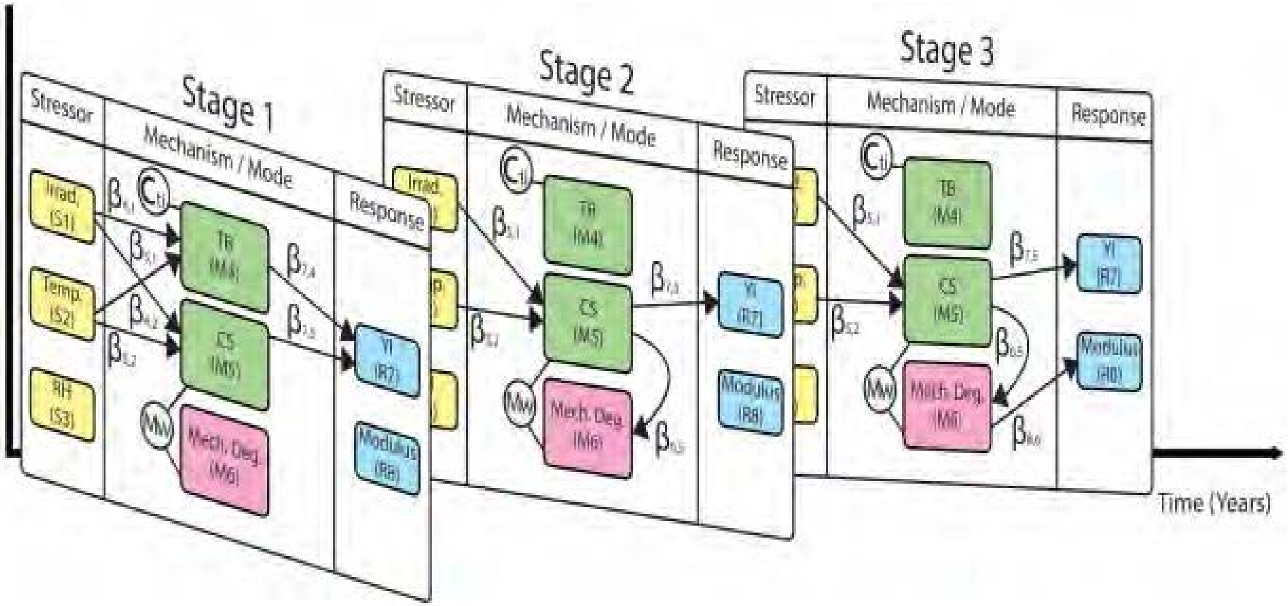


Figure 1. Acrylic degradation occurring in 3 stages as the UV stabilizer is bleached, the acrylic backbone degrades, and the mechanical integrity fails.

1. W.H. Huang, N. Wheeler, A. Klinke, Y. Xu, W. Du, A.K. Verma, et al., netSEM: Network Structural Equation Modeling, 2018.

2. Bruckman, Laura S., et al. "Statistical and domain analytics applied to PV module lifetime and degradation science." *IEEE Access* 1 (2013): 384-403.

Study Protocol

Meta data collection

Exposures using:

- Standard protocols
- Modified protocols
- Real-use conditions

Stressors

- Irradiance (UV and Full Spectrum)
- Heat
- Humidity, Water Spray
- Thermal Cycling

Multiple replicates (population behavior)

Tests need multiple time points

- Extend to 4000, 6000 hours

Evaluations (data acquisition)

- Performance & Mechanistic
  - Color, cracking, adhesion, etc.
- Every 100-500 hours

Yield ~9 step evaluations in 4000 hours

Identify change points (mechanism changes)

MATERIALS DATA SCIENCE RELY

Data Driven Model: PET

Study Protocol

- 3 different grades of PET
- 4 exposures
- Yellowness index: Performance
- Fixed effects models

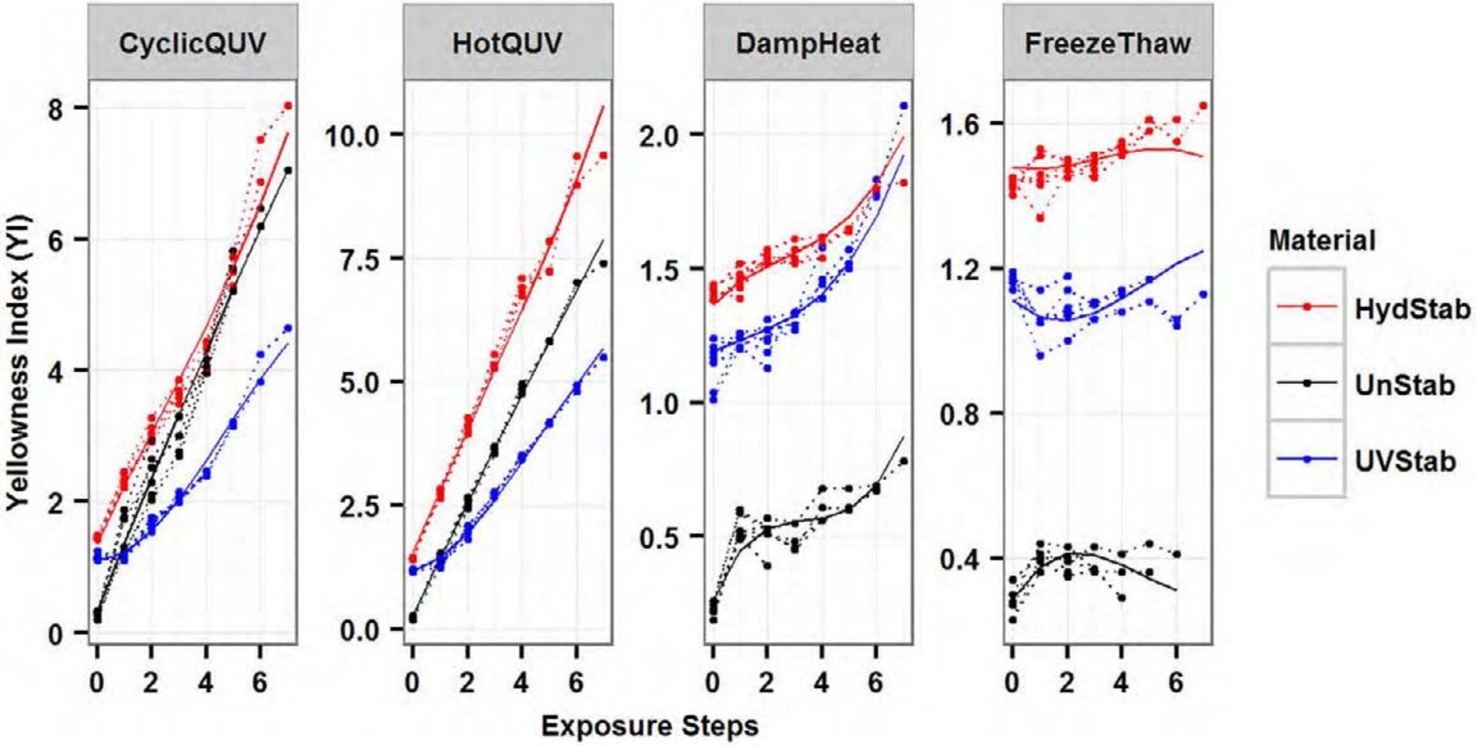


Fig. 2 Fixed effects modeling of PET yellowing. Y-axis are different ranges

$$
\begin{aligned}
YI \approx\ & (\beta_0 + \beta_{01}M_1 + \beta_{02}M_2 + \beta_{03}X + \beta_{04}M_1X + \beta_{05}M_2X) \\
& +(\beta_1 + \beta_{11}M_1 + \beta_{12}M_2 + \beta_{13}X)t \\
& +(\beta_2 + \beta_{21}M_1 + \beta_{22}M_2 + \beta_{23}X)t^2 \\
& +(\beta_3 + \beta_{31}M_1 + \beta_{32}M_2 + \beta_{33}X)t^3
\end{aligned}
$$

$$
M_1 = \begin{cases} 1 & \text{if Material} = \text{UnStab} \\ 0 & \text{otherwise} \end{cases} \qquad M_2 = \begin{cases} 1 & \text{if Material} = \text{UVStab} \\ 0 & \text{otherwise} \end{cases}
$$

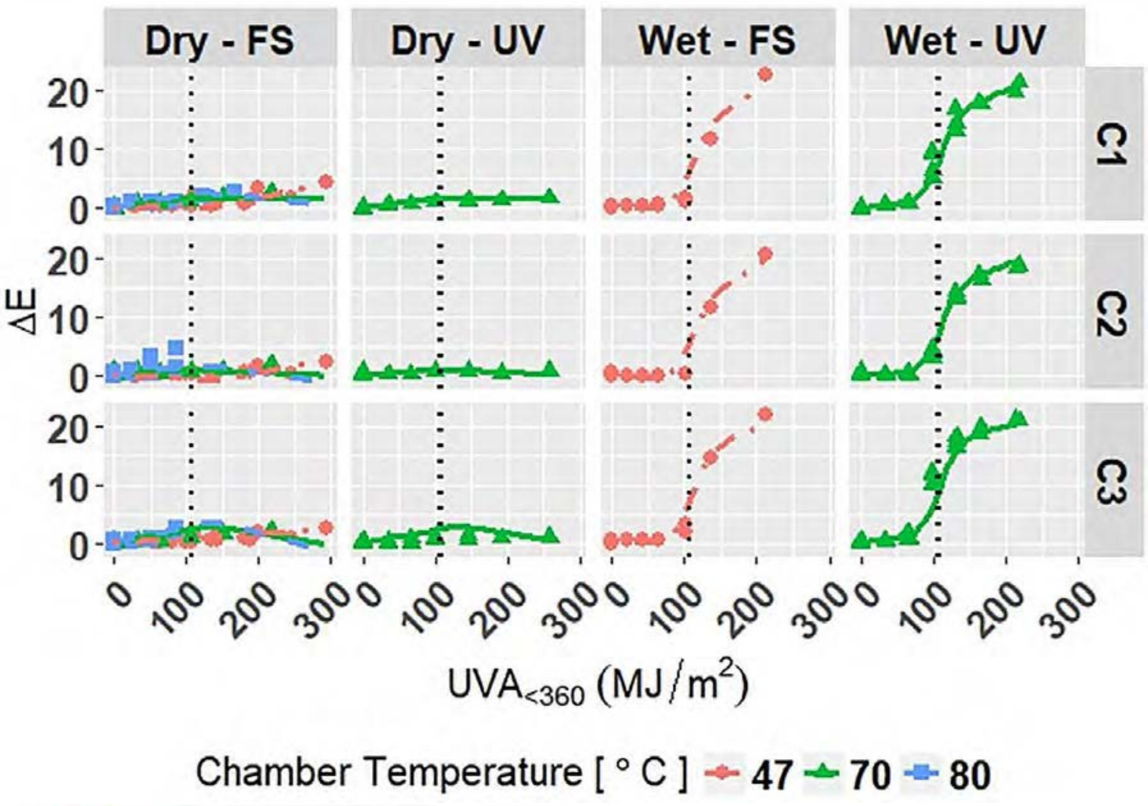
$$
X = \begin{cases} 1 & \text{if Exposure} = \text{HotQUV or FreezeThaw} \\ 0 & \text{otherwise} \end{cases}
$$

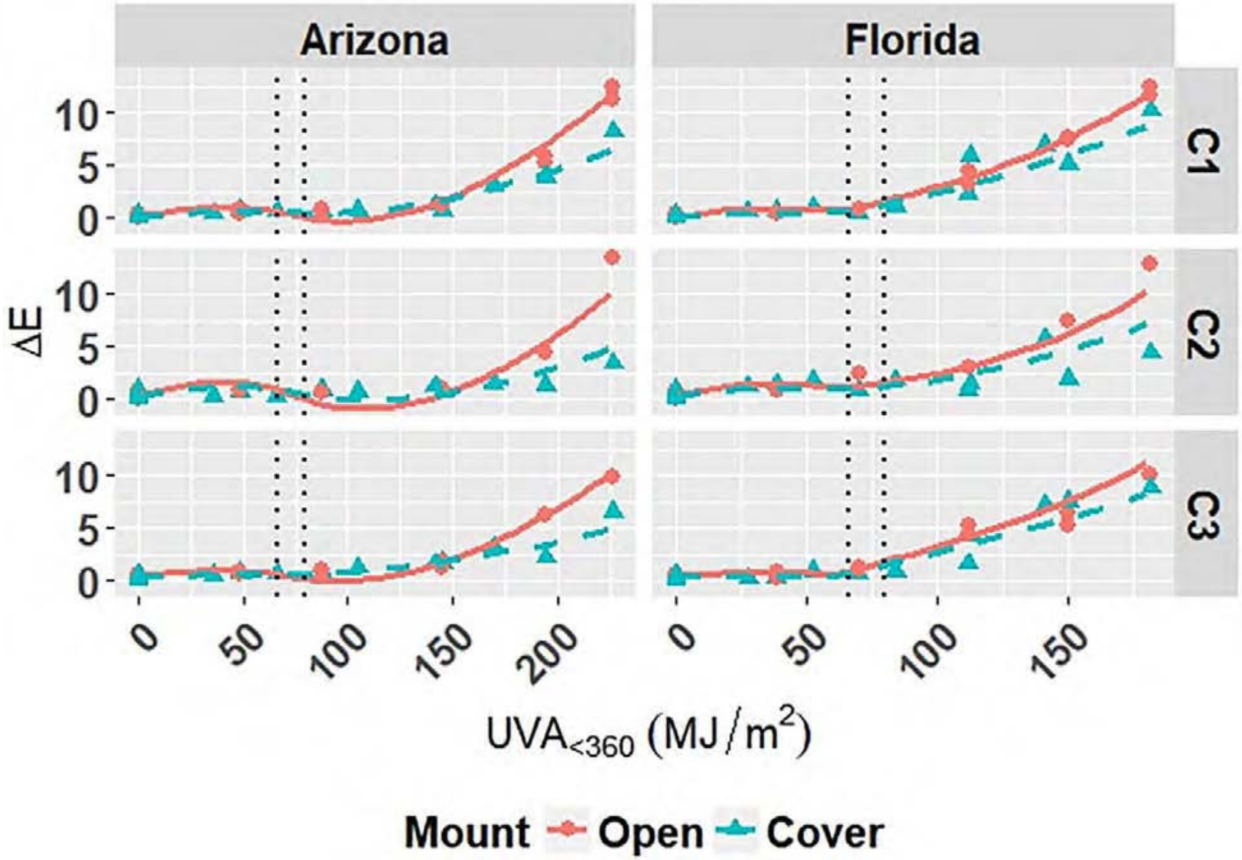
MATERIALS DATA SCIENCE RELY

Gok A, Ngendahimana DK, Fagerholm CL, French RH, Sun J, et al. (2017) Predictive models of poly(ethylene-terephthalate) film degradation under multi-factor accelerated weathering exposures. PLOS ONE 12(5): e0177614. https://doi.org/10.1371/journal.pone.0177614

Data Driven Model: PET Accelerated vs. Outdoor

Multi-variate multiple regression modeling



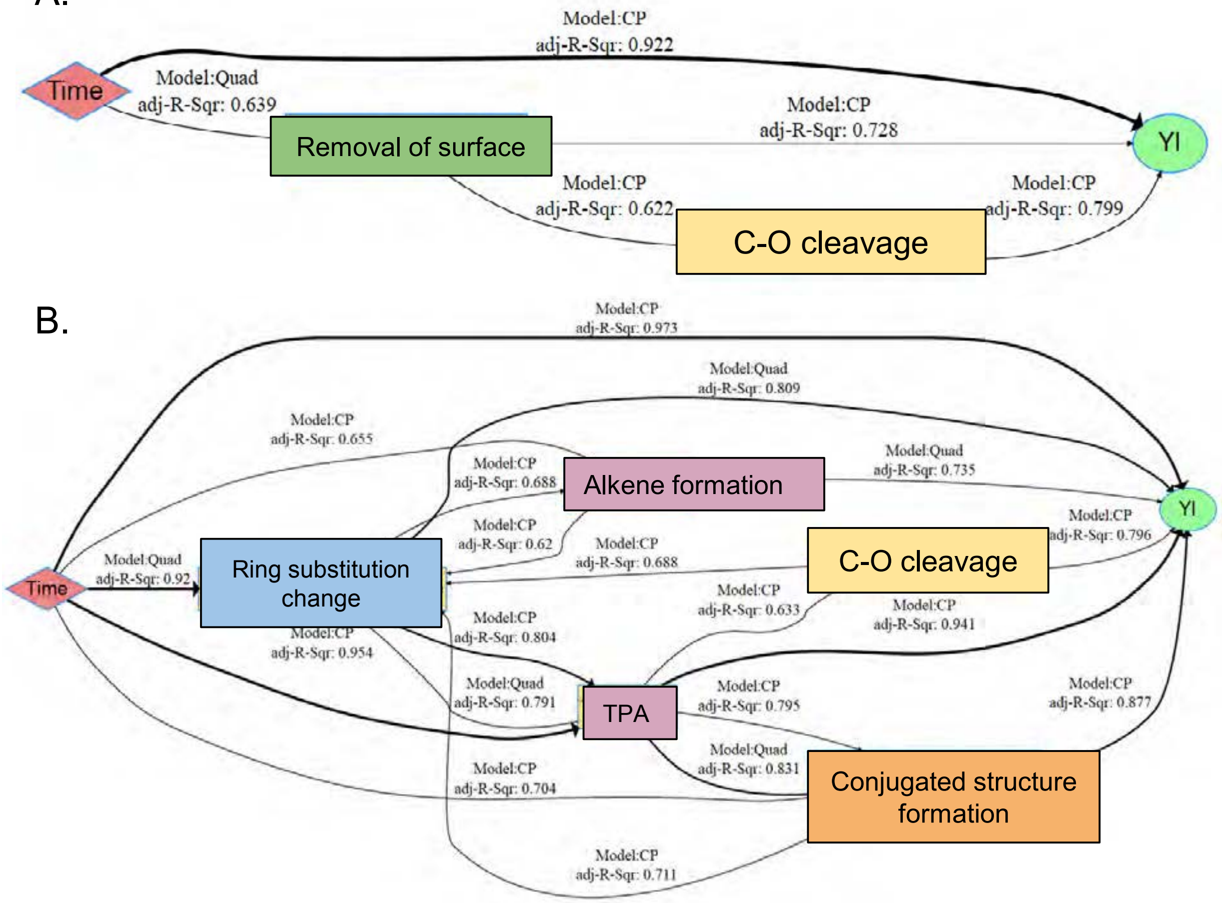


Gordon DA, Huang WH, Burns DM, French RH, Bruckman LS (2018) Multivariate multiple regression models of poly(ethylene-terephthalate) film degradation under outdoor and multi-stressor accelerated weathering exposures. PLOS ONE 13(12): e0209016.
https://doi.org/10.1371/journal.pone.0209016

netSEM modeling of PET

A.

B.



PET exposed to 0.8 w/m^2/nm at 340 nm, 80°C

- with water spray (A)
- without water spray (B)

FTIR indicated

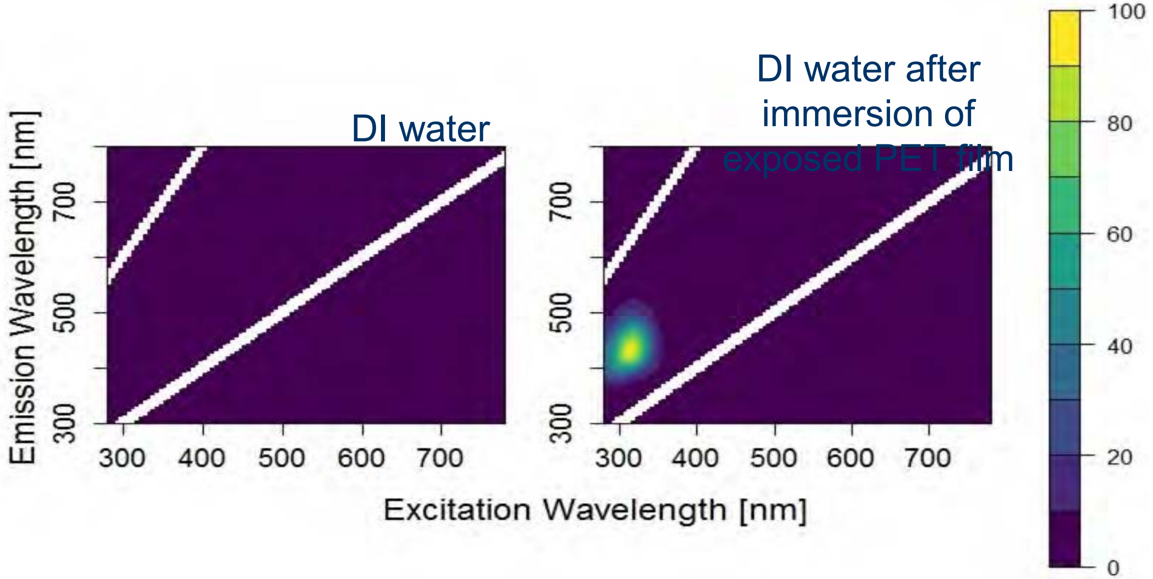
- Surface removal of degraded products with water spray
- Identify degradation products without water spray

Bruckman, Laura, et al. *Backsheets: Correlation of Long-Term Field Reliability with Accelerated Laboratory Testing.* No. DOE-UL-0007143. Renewable Energy, Underwriter's Laboratories, Northbrook, IL 60062, 2019.

Confirmation of Mechanisms

Parallel Factor Analysis

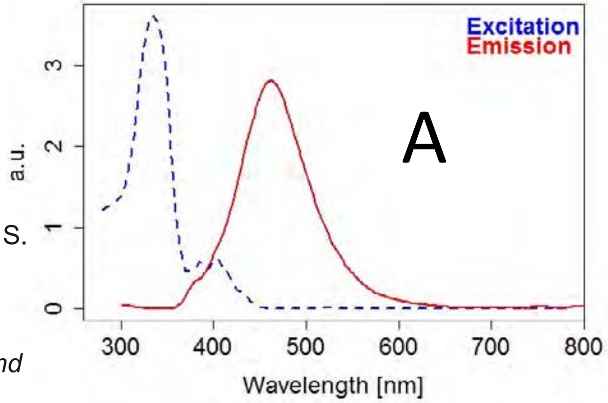
- Excitation/emission fluorescence spectra
- Determined surface removal of degraded PET
- Relative concentrations of
  - monohydroxy-terephthalate (A)
  - dihydroxy-terephthalate (B)

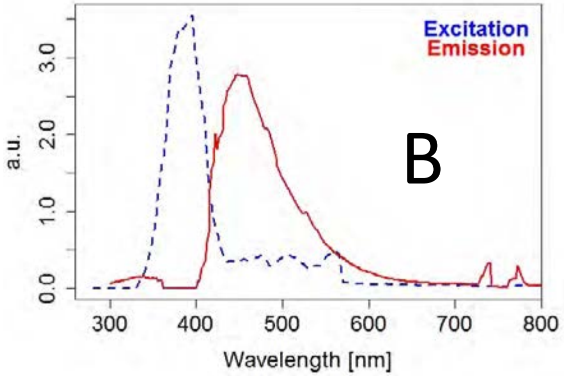


Size exclusion chromatography

Solid State NMR

Gordon, Devin A., Zhonghao Zhan, and Laura S. Bruckman. "Characterizing the weathering induced degradation of Poly (ethylene-terephthalate) using PARAFAC modeling of fluorescence spectra." *Polymer degradation and stability* 161 (2019): 85-94.





MATERIALS DATA SCIENCE RELY

Proposal Deliverables

1. Mechanistic and Performance Data Set
   a. Various polymer grades
   b. Under multiple stressors and stressor levels (parameter space)

1. Data driven and network degradation models
   a. Lifetime Prediction from model equations
   b. Uncertainty in model prediction
   c. Insights into behavior in a group

1. Resulting code to develop models
   a. Reproducible research methods
   b. Improved netSEM package

Proposal Tasks & Timeline

Tasks	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
1. Confirm polymer types and exposures and begin measurements for initial mechanistic evaluations								
2. Exposures and Evaluations of polymer types								
3. Data-driven and network degradation model development								
4. Service lifetime prediction of polymers under various stressors and stress levels								

MATERIALS DATA SCIENCE RELY

Thank You!

Project Proposal
(September 2019)

Project Title: Understanding Mechanical Abrasion Reliability of Multifunctional Coatings and Surfaces

Principal Investigator(s): Paul W Leu

New Project: X

Thrust Area: 3. Reliability Studies

Objective: The objective of this project is to provide for detailed understanding of the mechanical reliability of multifunctional coatings and surfaces with regard to tangential abrasion. Multifunctional coatings and surfaces offer tremendous potential for a variety of industry segments, having demonstrated new functionalities such as self-cleaning, anti-corrosion, anti-icing, stain-resistance, and microbial-resistance. However, the major challenge to their widespread adoption is their lack of reliability, and in particular, poor abrasion reliability. While some demonstrations of abrasion reliability have been demonstrated through self-healing and self-similar surfaces, our understanding of the mechanics of abrasion is still very limited. This project seeks to combine standard macroscopic linear abrasion tests with nanoindentation and nanoscratching experiments to provide for fundamental and mechanistic insight into the mechanics of wear abrasion of new multi-functional surfaces. This knowledge will enable the engineering and manufacture of coatings for abrasion-resistance through hardening and friction-reduction.

Standards Used: ASTM D6279 (Standard Test Method for Rub Abrasion Mar Resistance of High Gloss Coatings), ASTM D3363 (Standard Test Method for Film Hardness by Pencil Test), JIS L 0849 (Test Methods for Colorfastness to Rubbing), EN 3475-503 (Aerospace series - Cables, electrical, aircraft use - Test methods - Part 503: Scrape abrasion), ISO 8980-5 (Ophthalmic optics - Uncut finished spectacle lenses), ISO 15184 (Paints and Varnishes - Determination of Film Hardness by Pencil Test)

Code Developed: This project will leverage code from image machine learning projects for macroscale abrasion, nanoindentation, and nanoscratching experiments.

Datasets Produced: We will create optical, scanning electron microscopy, and scanning probe microscopy image datasets of various macroscale abrasion, nanoindentation, and nanoscratching experiments.

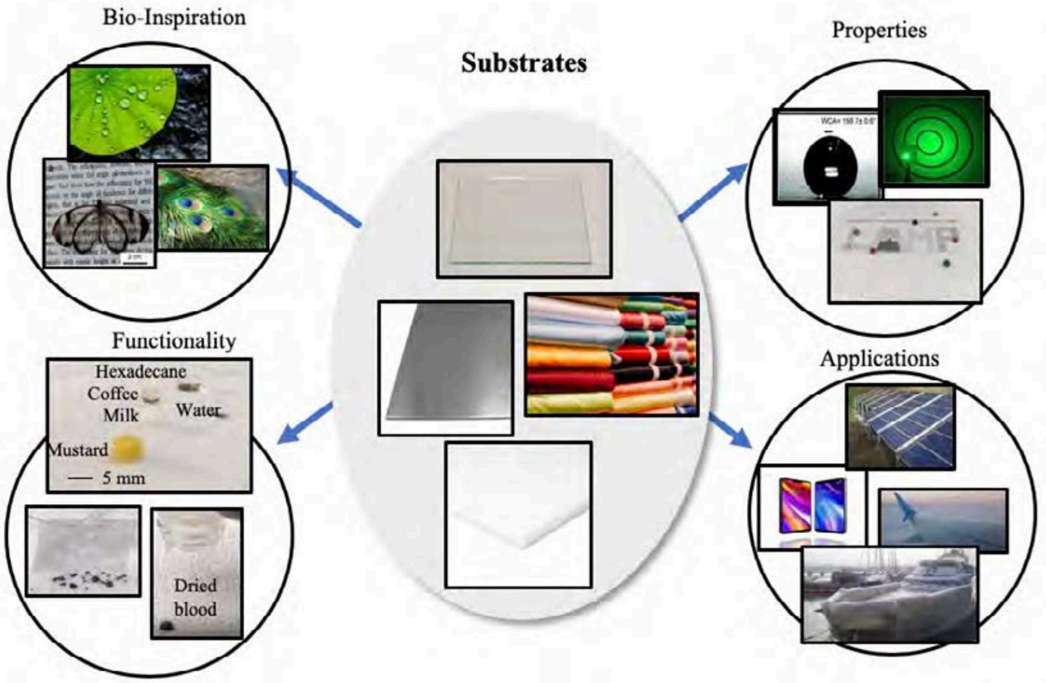


Figure 1: Multifunctional coatings and surfaces may have wide application to a wide variety of substrates such as glass, metals, textiles, and polymers. Multifunctional coatings are bio-inspired by examples from nature such as the lotus leaf, glasswing butterfly, or peacock feathers. Various properties are valued such as high transparency, high water contact angle, and high light scattering abilities as well as functionalities such as self-cleaning, stain-resistance, and microbial resistance. Applications for multifunctional coatings and surfaces include solar cells, consumer electronics, aerospace, and marine sectors among many.

Background: New materials, bioinspiration, manufacturing techniques and physics offer unprecedented opportunities in engineering coatings and surfaces for difficult-to-realize material–property combinations, novel photon management strategies, and new multi-functionality (**Fig. 1**). In the past ten years, discoveries in multi-functional coatings have led to a renaissance of activity in surface engineering, which may transform a variety of industry sectors such as automotive, aerospace, marine, construction, solar and wind, textiles, consumer electronics, food processing, and medical. New functionalities such as self-cleaning, anti-corrosion, anti-icing, stain-resistance, and microbial-resistance have been demonstrated. Many of these functionalities are closely related to superhydrophobicity or superomniphobicity, where water or wide variety of liquids, respectively, are repelled from the surface with both high static contact angles (> 150 °) and low contact angle hysteresis (< 10 °). These surfaces may also offer new opportunities with regard to photon management such as broadband antireflection, omnidirectional antireflection, and structural color.

Different natural surfaces have served as bioinspiration for various functionalities. For example, the leaves of the *Nelumbo nucifera* (sacred lotus) are superhydrophobic due to the combination of hierarchical surface morphology and hydrophobic epicuticular wax [1] and inspired a variety of synethic self-cleaning surfaces [2]. Peacock feathers [3] and *Morpho* butterflies wings [4] have brilliant iridescent colors due to structures instead of pigments. And glasswing butterfly wings have broadband and omnidirectional antireflection due to random height and spacing sub-wavelength nanostructures in their wings [5]. Towards demonstrating new functionalities, our research group has demonstrated surfaces and coatings with self-cleaning [2], stain-resistance [6], anti-reflection and superomniphobicity [7], and anti-microbial properties [8].

However, there is a need to improve the reliability of these new multi-functional coatings and surfaces. These coatings and surfaces may degrade mechanically through (1) delamination from the substrate, (2) normal impact with solids, liquids, or gas, (3) laundry and liquid baths, and (4) tangential abrasion to surface [9]. This proposal focuses on the major challenge to the widespread application of multifunctional coatings and surfaces, which is their poor reliability under tangential abrasion to the surface. Many multifunctional coatings are easily abraded, sometimes with little more than brushing with a tissue and need to be constantly reapplied.

One approach to improving abrasion durability has been to strengthen interactions between nanostructure coatings and substrates to keep the nanostructures fixed and prevent them from shearing off easily [10]. However, the nanostructures can still be completely abraded away relatively quickly. General strategies which have demonstrated better reliability have included creating (1) self-healing or (2) self-similar surfaces or some combination of the two [11]. Self-healing surfaces are bioinspired, where the surface is able to regenerate roughnesss and/or low surface-energy surface passivation. Self-similar surfaces consist of fractal-like network or hierarchical micro-/and nanostructures that retain underlying self-similar surfaces when the top surface is abraded away. Our research group has been involved with the demonstration of these types of surfaces as well, having demonstrated nonwoven polypropylene that is resistant to abrasion and scratching and slicing with a razer blade [12] and antireflective glass with self-healing properties [7].

However, we are just "scratching the surface" so to speak with regard to understanding the mechanics of abrasion in these new multi-functional coatings and surfaces. There is a need for more detailed understanding of how the mechanical properties and morphology of the surface impact scratching and wear during tangential abrasion. New strategies for increasing hardness and reducing friction must be evaluated.

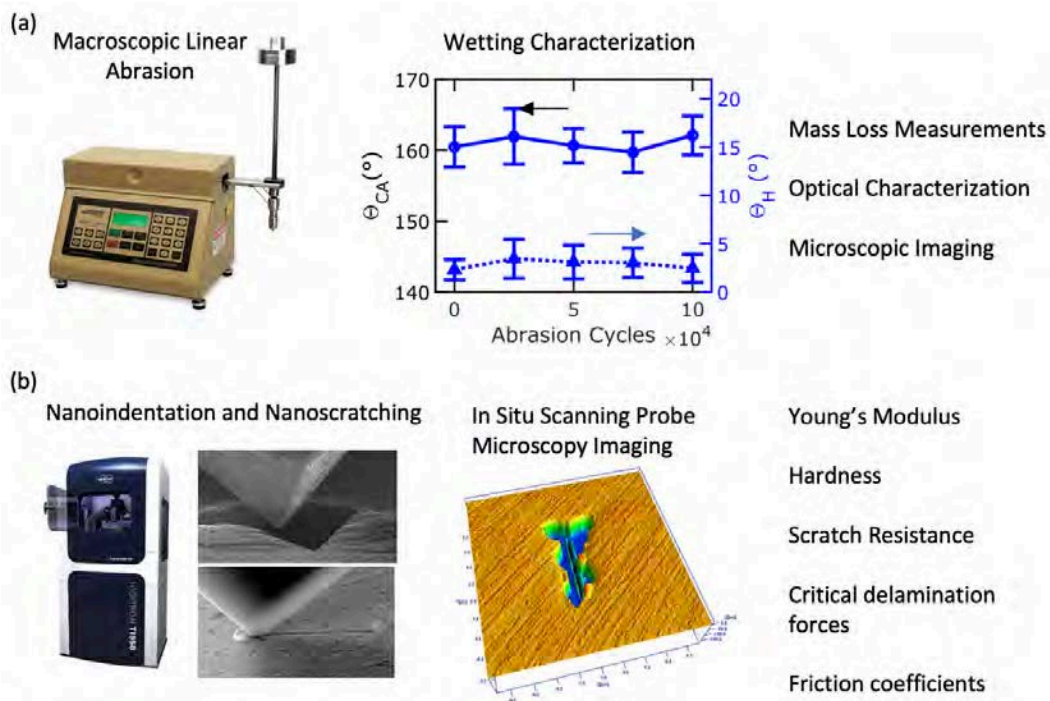


Figure 2: Abrasion characterization tasks of this research project include (a) macroscopic linear abrasion tests with (b) nanoindentation and nanoscratching experiments. Macroscopic linear abrasion tests will provide for qualitative evaluation of surfaces. The change in the surface will be assessed through wetting characterization of samples as well as mass loss, optical transmission change, and microscopic imaging. More detailed and precise testing will be performed through nanoindentation and nanoscratching tests. These tests will be complemented by in situ scanning probe microscopy imaging and enable the assessment of the coating/substrate for elastic modulus, hardness, scratch resistance, critical delamination forces, and friction coefficients.

Project Tasks: This proposal will seek to provide for more detailed mechanistic understanding of how various coatings and surfaces degrade from mechanical abrasion. Both standardized macroscopic tests and more detailed nanoindentation and nanoscratching experiments will be performed on coatings and surfaces to provide a comprehensive characterization of the surface (**Fig. 2**). Experiments will performed on some of the coatings and surfaces developed by the PI's own group or any surfaces or coatings that industry members are interested in.

Task 1. Macroscopic Linear Abrasion Tests. Macroscopic linear abrasion tests, which are currently the most popular reliability, evaluation method will be performed. Standard tests will provide for good reference. Different abrading surfaces consisting of a textile, a rubber, and vitrified material will provide for a comprehensive overview of the response to various abrasive actions. Downward pressures of at least 10 kPa will be applied, and hundreds of abrasion cycles will be performed. The static water contact angle as well as the contact angle hysteresis of these surfaces will be periodically characterized. Pencil hardness or blade tests will also be performed if relevant applications may include contact with sharp objects. Bare finger rubbing will used for surfaces that will be repeatedly handled by people. Mass loss measurements as well as optical haze measurements will also be utilized to assess changes to the surface. These tests will also be followed by microscopic imaging for evaluation of the surface. Imaging results may find good synergy with other proposed projects focused on image machine learning. These standard tests are widely performed by industry, but mostly phenomenological, and offer limited insight into understanding mechanical abrasion.

Task 2. Nanoindentation and Nanoscratching Experiments. To provide for deeper insight, we propose to perform nanoindentation and nanoscratching experiments. In particular, Pitt has a Hysitron TI 950 TriboIndenter that provides for quasistatic nanoindentation, scratch testing, nanowear, and in-situ imaging. The system consists of dual heads such that tests will be performed at both the micro- and nanoscale.

Nanoindentation will be performed various surfaces to determine the elastic modulus and hardness of the coating/substrate system from load-displacement curves. Ramping force nanoscratch experiments will be performed on samples and post-scratch images will be captured after each ramping force nanoscratch test using an in-situ scanning probe microscope. These images will provide for insight into different distinct regimes of damage such as plastic deformation, micro-cracking, chipping, and micro-abrasion. Normal displacement and lateral force versus time plots will be made for each scratch. Critical delamination forces and friction coefficients may be extracted from these curves.

Nanowear experiments will be performed to quantify wear behavior as a function of number of sliding cycles, sliding velocity, wear area, and applied force. These experimental results together with macroscopic linear abrasion tests will provide for insight into the mechanical reliability of various multifunctional coatings and surfaces.

Benefits to Members: The knowledgebase and understanding from these experiments will enable better engineering and manufacturing of multi-functional coatings with high mechanical abrasion reliability. The project also complements other projects that are focused on image machine learning.

Short Bio: Dr. Paul W. Leu received his Ph.D. from Stanford University in 2008. From 2008-10, he worked as a postdoctoral fellow at the University of California, Berkeley before joining the University of Pittsburgh as faculty in 2010. He is currently an Associate Professor and the John C. Mascaro Sustainability Faculty Fellow in the Industrial Engineering Department at the University of Pittsburgh. He directs the Laboratory for Advanced Materials at Pittsburgh (LAMP; `lamp.pitt.edu`). He has over 40 research publications and has been recipient of the Oak Ridge Associated University Powe Junior Faculty Enhancement Award, UPS Minority Advancement Award, and the NSF CAREER Award. His research has been showcased in Industrial Engineering magazine, Pittsburgh NPR, and Pittsburgh Magazine.

References

1. W. Barthlott and C. Neinhuis, “Purity of the sacred lotus, or escape from contamination in biological surfaces,” *Planta*, vol. 202, no. 1, pp. 1–8, 1997. [Online]. Available: http://www.springerlink.com/content/4m4jnt03qlvaywjl/

2. S. Haghanifar, P. Lu, M. Imrul Kayes, S. Tan, K.-J. Kim, T. Gao, P. Ohodnicki, and P. W. Leu, “Self-cleaning, high transmission, near unity haze OTS/silica nanostructured glass,” *Journal of Materials Chemistry C*, vol. 6, no. 34, pp. 9191–9199, 2018. [Online]. Available: https://pubs.rsc.org/en/content/articlelanding/2018/tc/c8tc02513d

3. J. Zi, X. Yu, Y. Li, X. Hu, C. Xu, X. Wang, X. Liu, and R. Fu, “Coloration strategies in peacock feathers,” *Proceedings of the National Academy of Sciences*, vol. 100, no. 22, pp. 12 576–12 578, Oct. 2003. [Online]. Available: https://www.pnas.org/content/100/22/12576

4. K. Chung, S. Yu, C.-J. Heo, J. W. Shim, S.-M. Yang, M. G. Han, H.-S. Lee, Y. Jin, S. Y. Lee, N. Park, and J. H. Shin, “Flexible, Angle-Independent, Structural Color Reflectors Inspired by Morpho Butterfly Wings,” *Advanced Materials*, vol. 24, no. 18, pp. 2375–2379, 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201200521

5. R. H. Siddique, G. Gomard, and H. Hölscher, “The role of random nanostructures for the omnidirectional anti-reflection properties of the glasswing butterfly,” *Nature Communications*, vol. 6, p. 6909, Apr. 2015. [Online]. Available: https://www.nature.com/articles/ncomms7909

6. S. Haghanifar, L. M. Tomasovic, A. J. Galante, D. Pekker, and P. W. Leu, “Stain-resistant, superomniphobic flexible optical plastics based on nano-enoki mushrooms,” *Journal of Materials Chemistry A*, May 2019. [Online]. Available: https://pubs.rsc.org/en/content/articlelanding/2019/ta/c9ta01753d

7. S. Haghanifar, M. McCourt, B. Cheng, J. Wuenschell, P. Ohodnicki, and P. W. Leu, “Creating Glasswing-Butterfly Inspired Durable Antifogging Omniphobic Supertransmissive, Superclear Nanostructured Glass Through Bayesian Learning and Optimization,” *Materials Horizons*, Jun. 2019. [Online]. Available: https://pubs.rsc.org/en/content/articlelanding/2019/mh/c9mh00589g

8. M. I. Kayes, A. J. Galante, N. A. Stella, S. Haghanifar, R. M. Q. Shanks, and P. W. Leu, “Stable lotus leaf-inspired hierarchical, fluorinated polypropylene surfaces for reduced bacterial adhesion,” *Reactive and Functional Polymers*, vol. 128, pp. 40–46, Jul. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1381514818301950

9. A. Milionis, E. Loth, and I. S. Bayer, “Recent advances in the mechanical durability of superhydrophobic materials,” *Advances in Colloid and Interface Science*, vol. 229, pp. 57–79, Mar. 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0001868615002377

10. C.-H. Xue and J.-Z. Ma, “Long-lived superhydrophobic surfaces,” *Journal of Materials Chemistry A*, vol. 1, no. 13, pp. 4146–4161, Mar. 2013. [Online]. Available: https://pubs.rsc.org/en/content/articlelanding/2013/ta/c2ta01073a

11. C. Peng, Z. Chen, and M. K. Tiwari, "All-organic superhydrophobic coatings with mechanochemical robustness and liquid impalement resistance," *Nature Materials*, vol. 17, no. 4, p. 355, Apr. 2018. [Online]. Available: https://www.nature.com/articles/s41563-018-0044-2

12. A. Galante, S. Haghanifar, N. A. Stell, R. M. Q. Shanks, and P. W. Leu, "Simple, super durable, superhemophobic nonwoven polypropylene fabric coating with serum protein repellency for smart medical textile applications," (to be submitted to Advanced Healthcare Materials).

Understanding Mechanical Abrasion Reliability of Multifunctional Coatings and Surfaces

Faculty Member(s): Paul W Leu

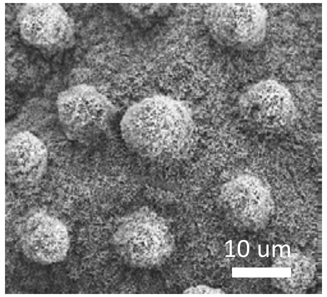
University / Department: Industrial Engineering

Multi-functional coatings and surfaces

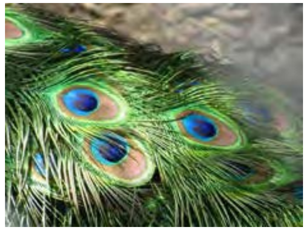
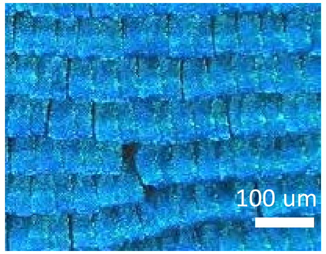
- Self-cleaning
- Anti-corrosion
- Anti-icing
- Stain-resistant
- Anti-microbial
- Photon management

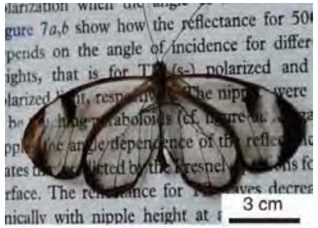
MATERIALS DATA SCIENCE
RELY

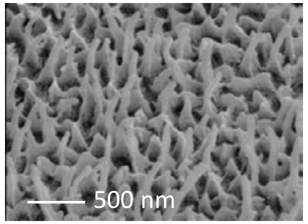
Bioinspiration





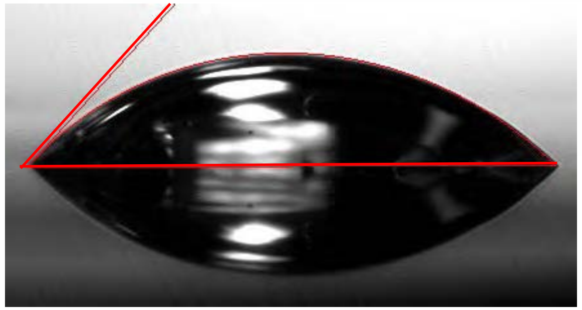




MATERIALS DATA SCIENCE RELY

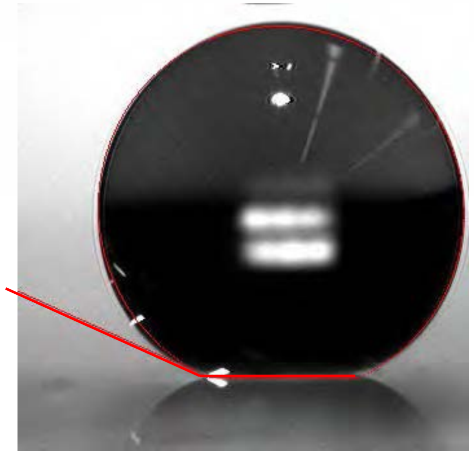
Self-Cleaning

Regular Glass



$\theta_{CA} = 42.9 \pm 1.1°$
$\theta_{H} = 35.5 \pm 2.7°$

Self-Cleaning Glass

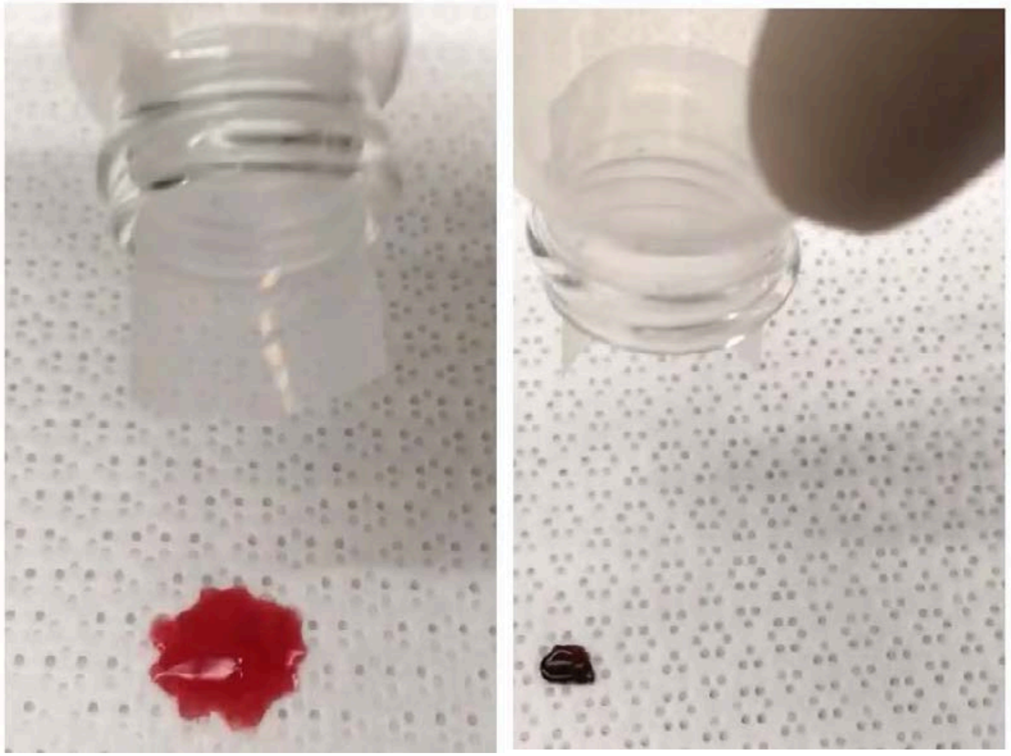


$\theta_{CA} = 159.7 \pm 0.6°$
$\theta_{H} = 4.9 \pm 0.6°$

Stain-Resistance

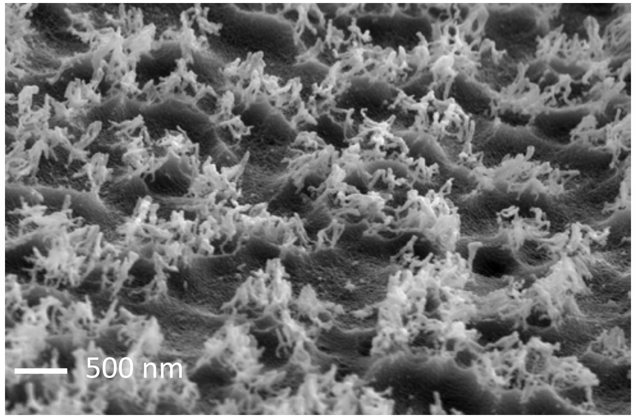
Regular PET

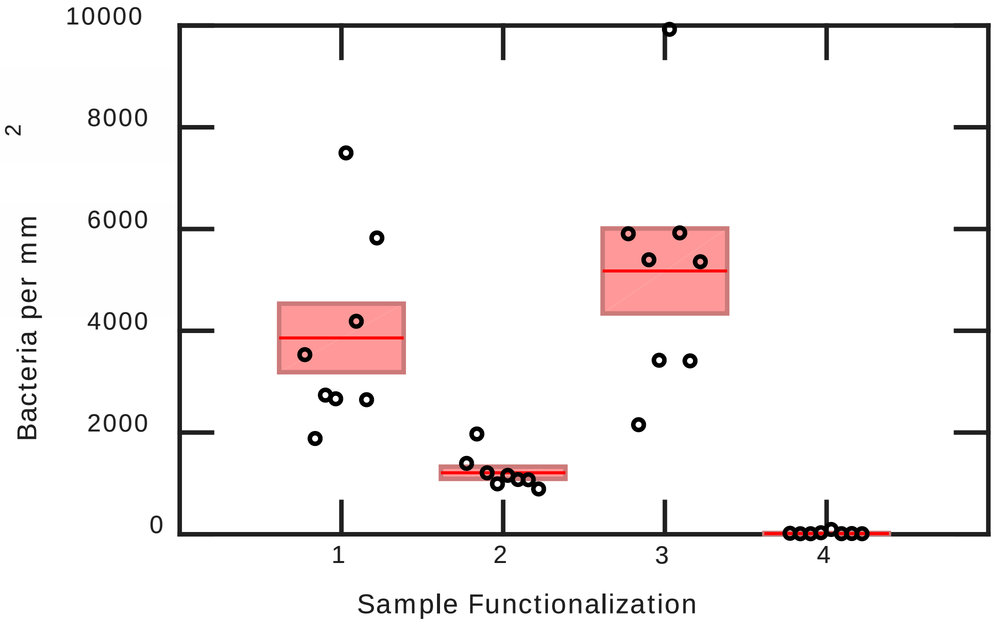
Superomniphobic PET



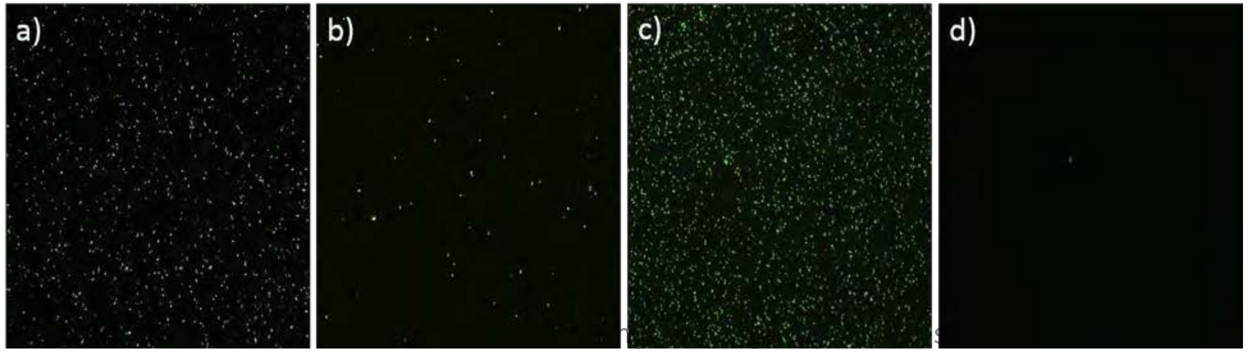
Stain-resistant Plastics

Anti-microbial





1-Control, 2-LP Oxygen, 3-HP Oxygen, 4-Hydrophobic



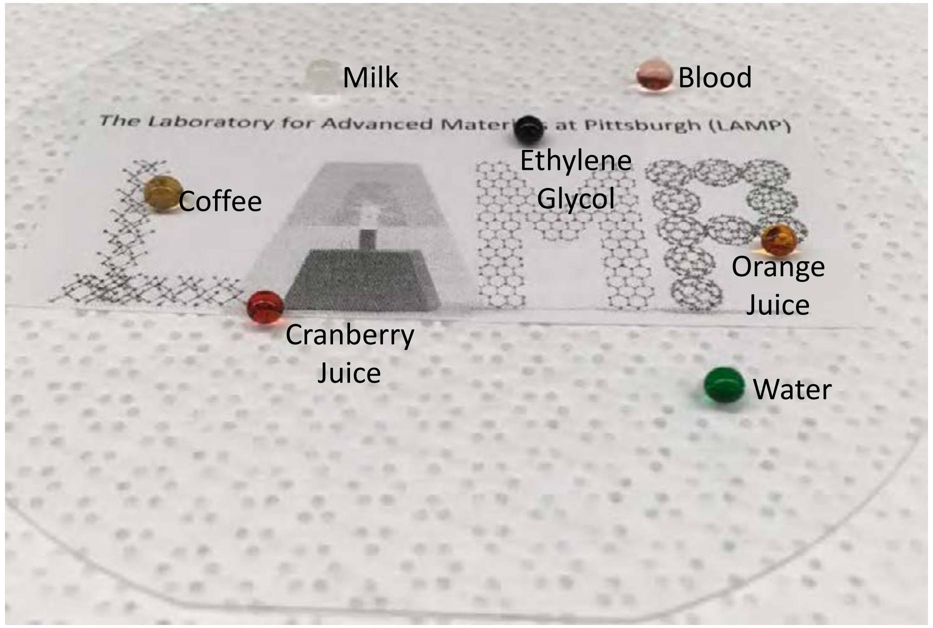
MATERIALS DATA SCIENCE RELY

Anti-reflective and Superomniphobic Glass

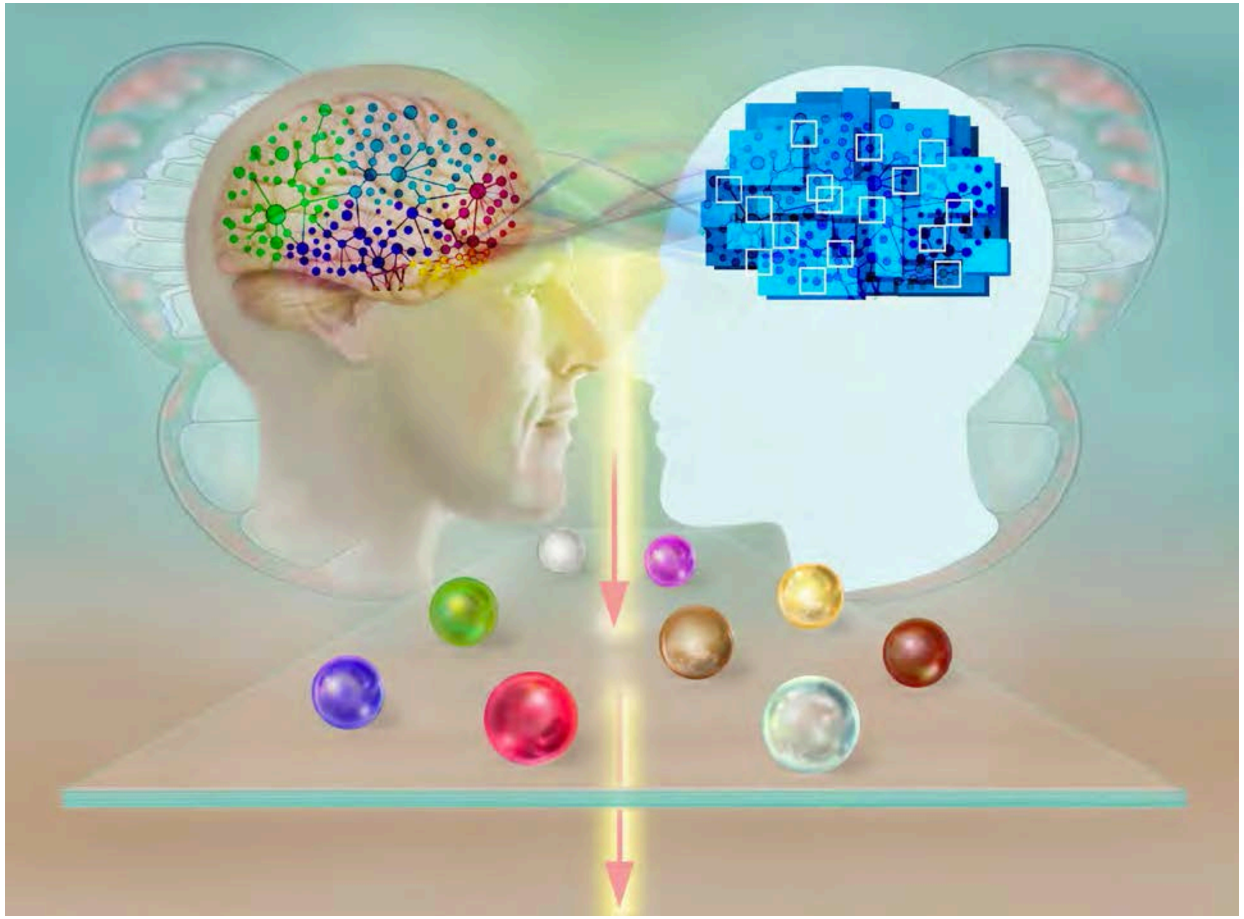
SIGOPT

Normal Glass

Nanostructured Glass



Cover Art



Mechanically Reliable?

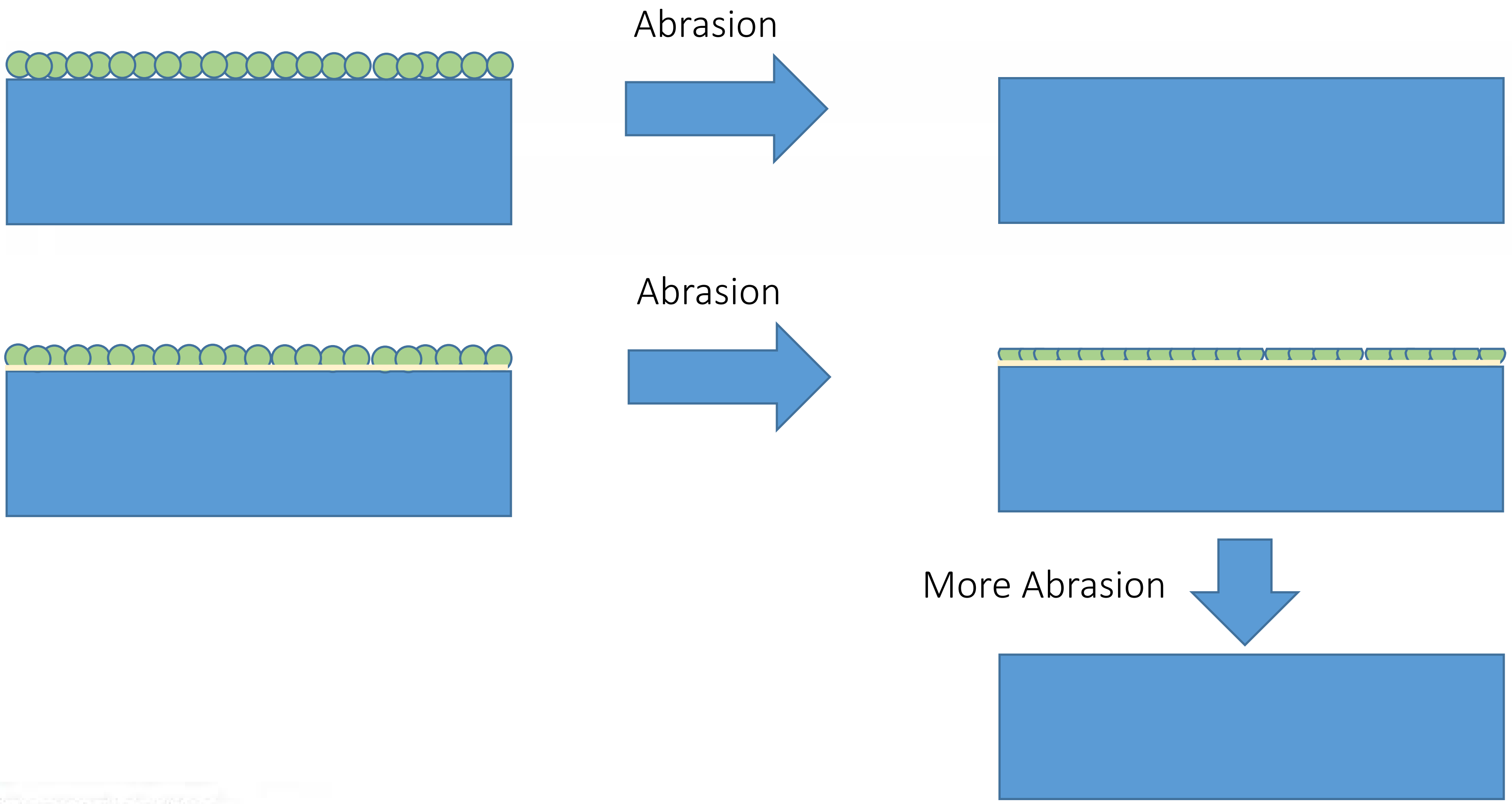
- Adhesion to substrate
- Normal impact with solids, liquids, or gas
- Laundry or liquid baths
- **Tangential abrasion**

Objective

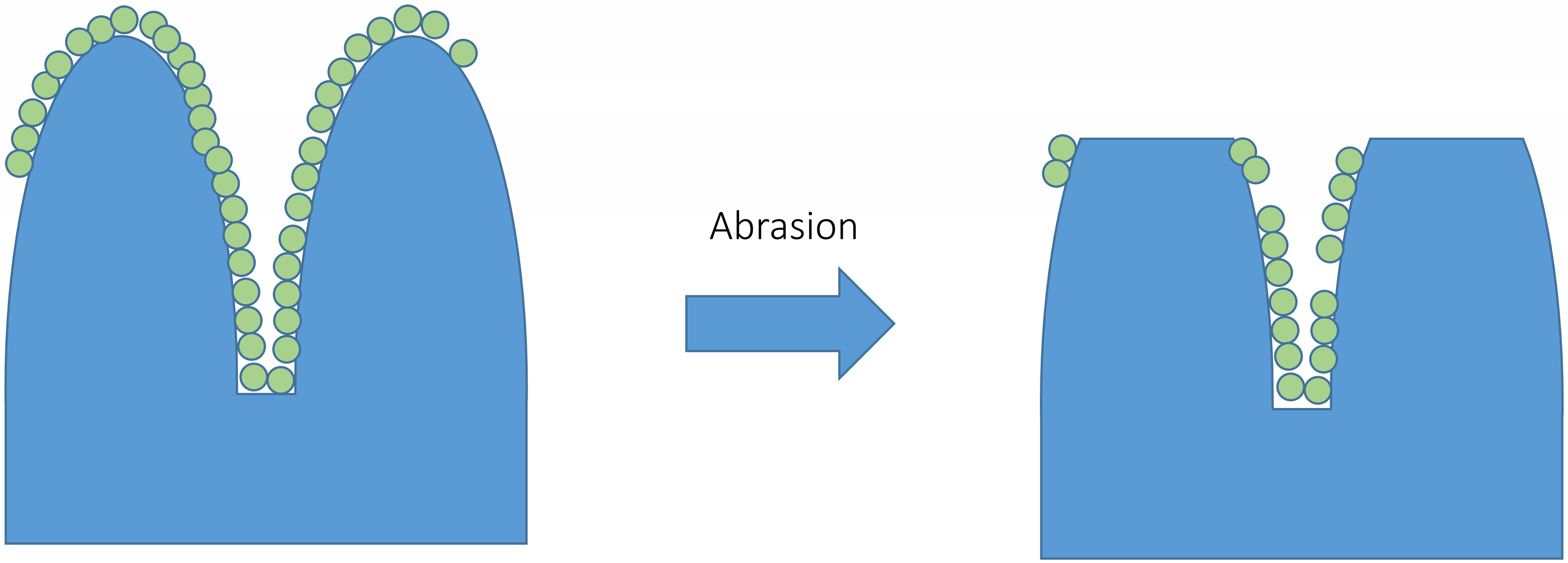
Provide for detailed understanding of the degradation mechanisms and mechanical reliability of multifunctional coatings and surfaces with regard to tangential abrasion.

Enable design of coatings and surfaces with reduced wear and better ability to maintain functionality

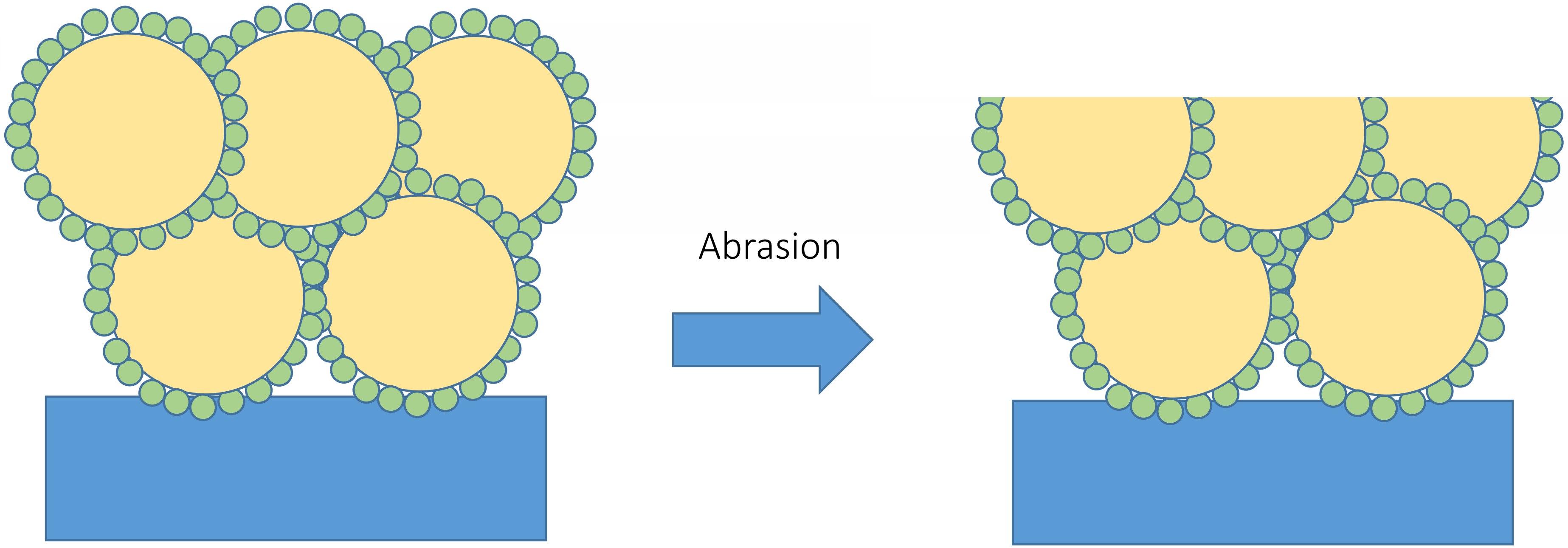
Design for Reduced Wear



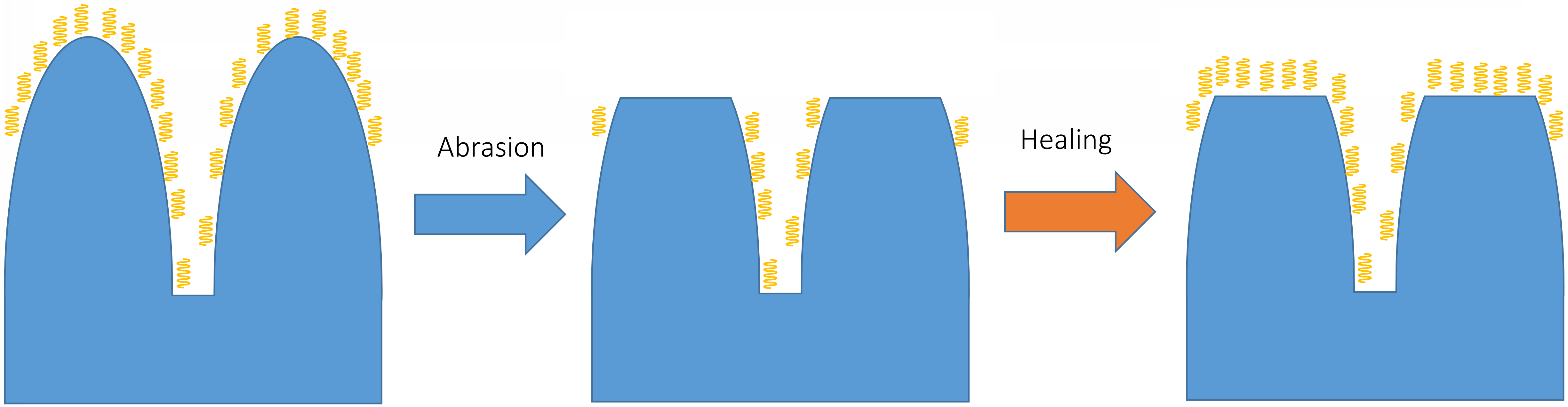
Hierarchical Micro-/Nanostructures



Self-Similar Structures

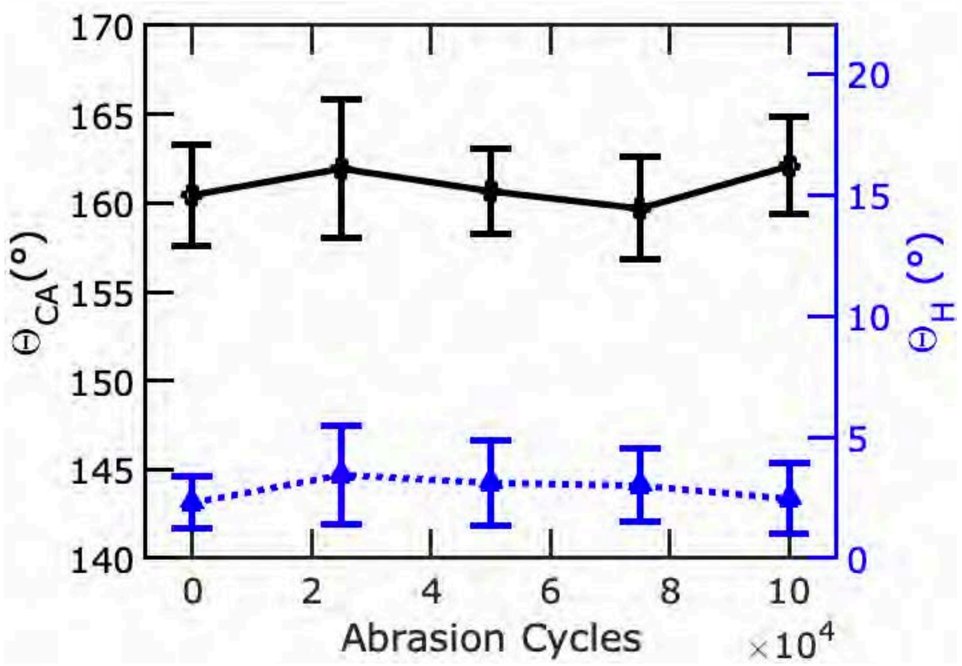


Self-Healing

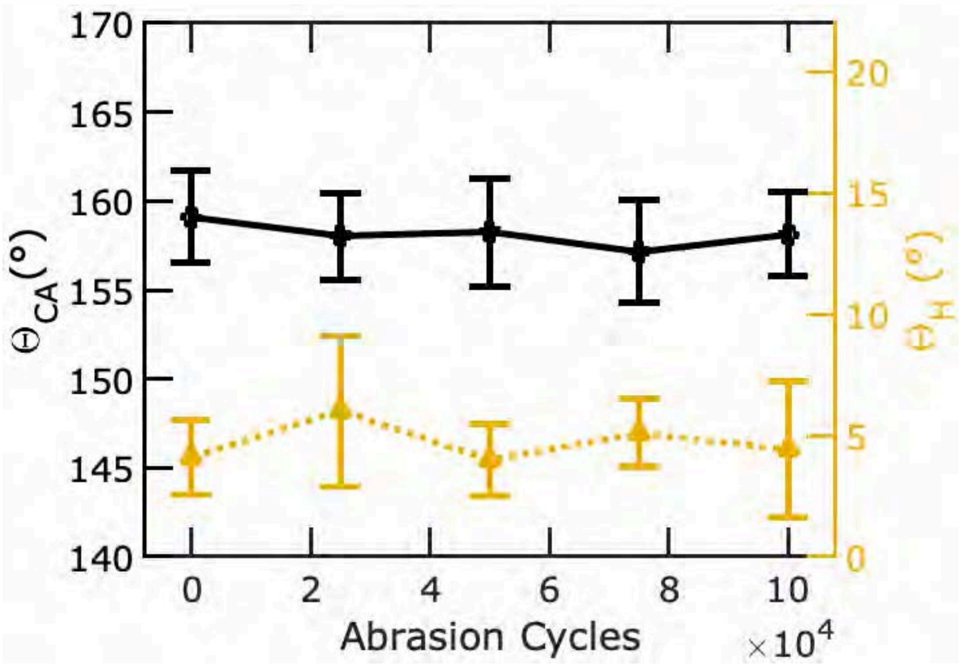


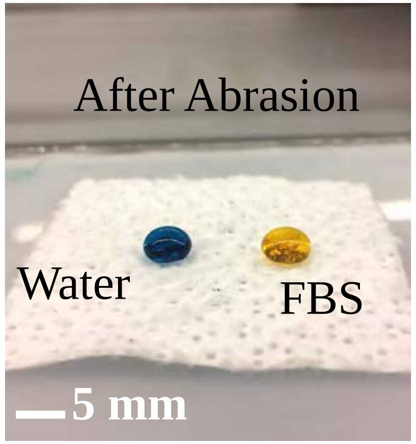
Durable Surfaces

Water



Fetal Bovine Serum

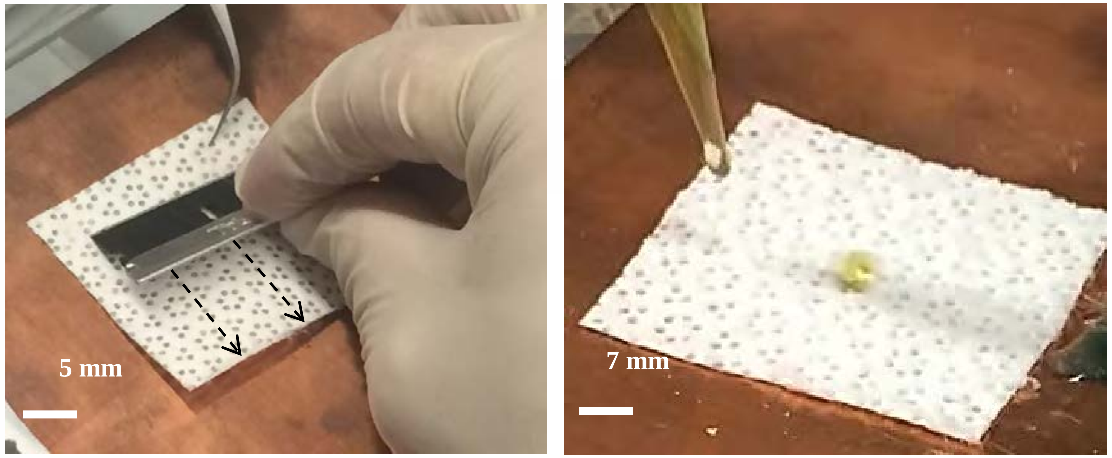




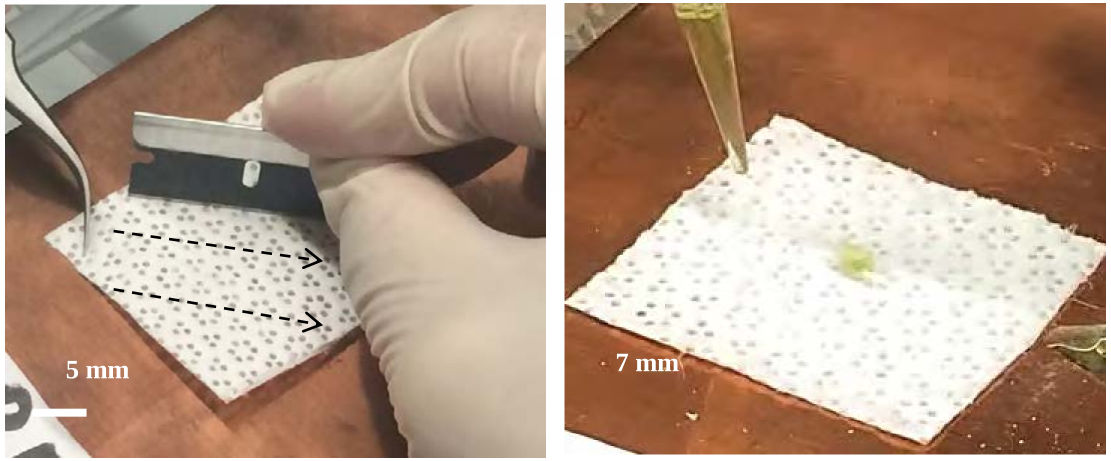
MATERIALS DATA SCIENCE RELY

Scratching and Slicing Tests

Scratching

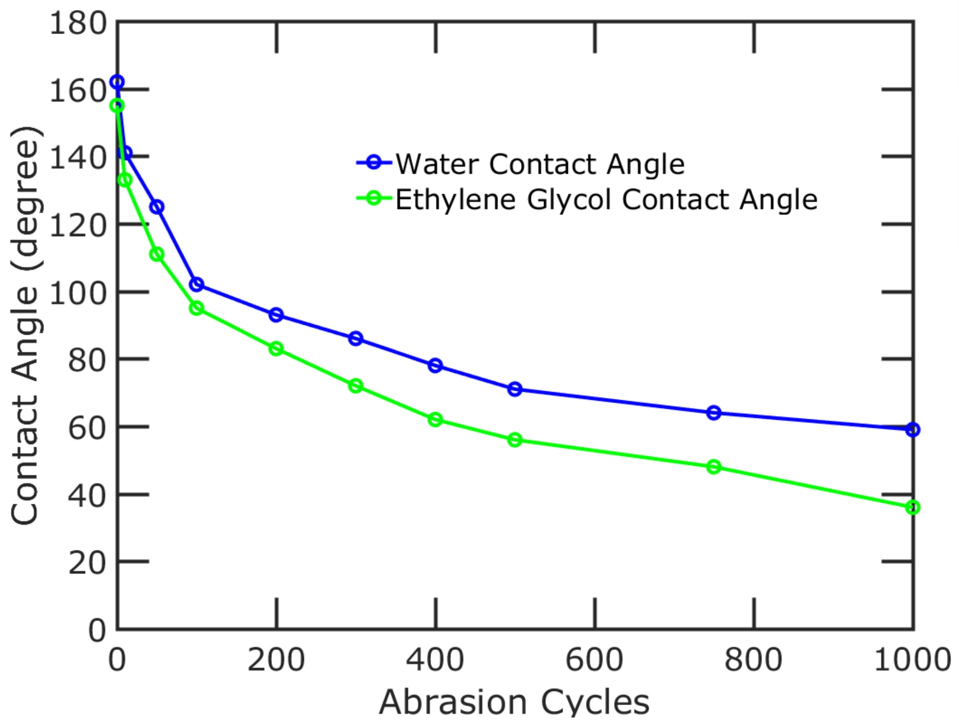


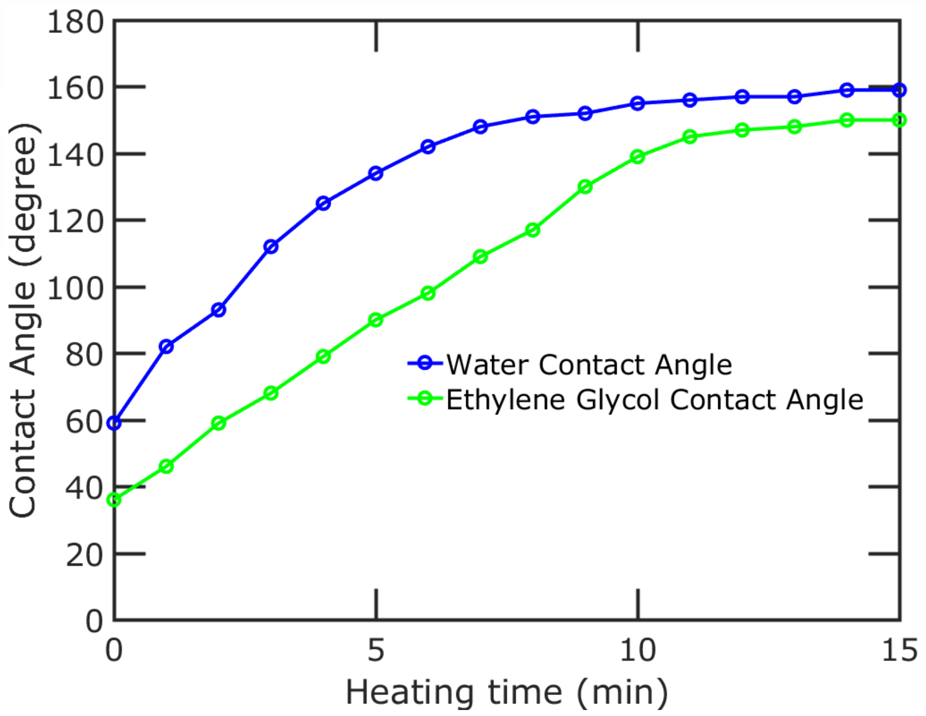
Slicing



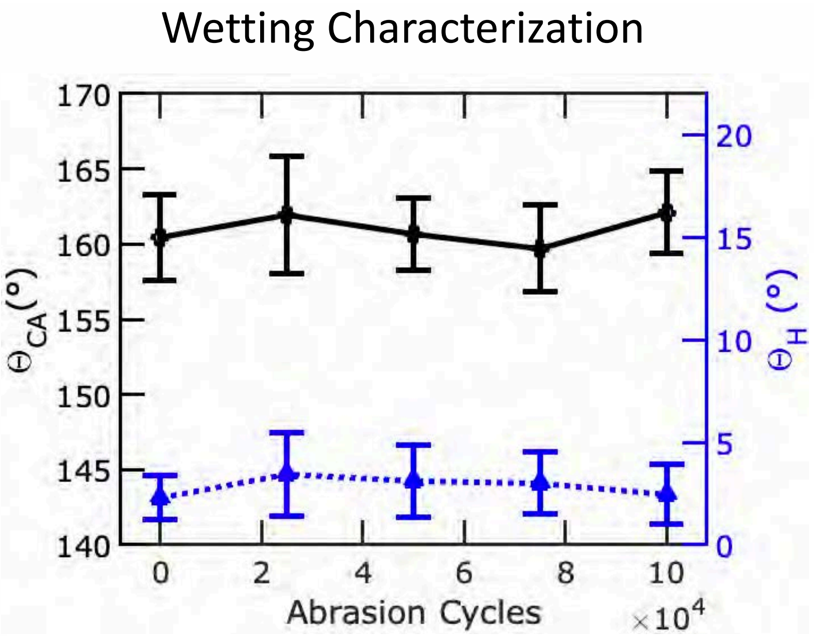
Scratching and Slicing Tests

Self-Healing





Task 1: Macroscopic Linear Abrasion



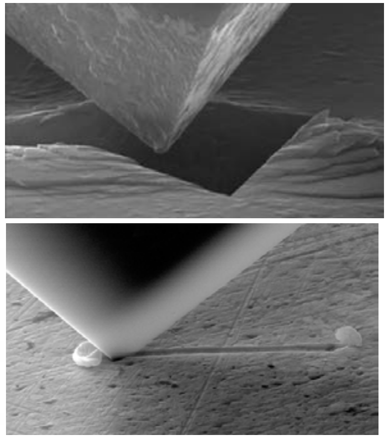
Mass Loss Measurements

Optical Characterization

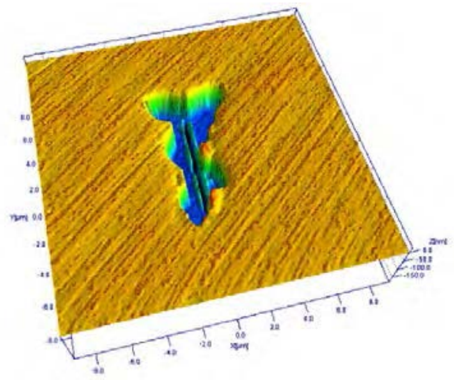
Microscopic Imaging

Pencil Hardness and Scratching or Slicing Experiments

Task 2: Nanoindentation and Nanoscratching



In Situ Scanning Probe Microscopy Imaging



Young's Modulus

Hardness

Scratch Resistance

Critical delamination forces

Friction coefficients

Study

- Effects of mechanical properties and surface morphology
- Hardening
- Friction-reduction
- Design for reduced wear

Deliverables

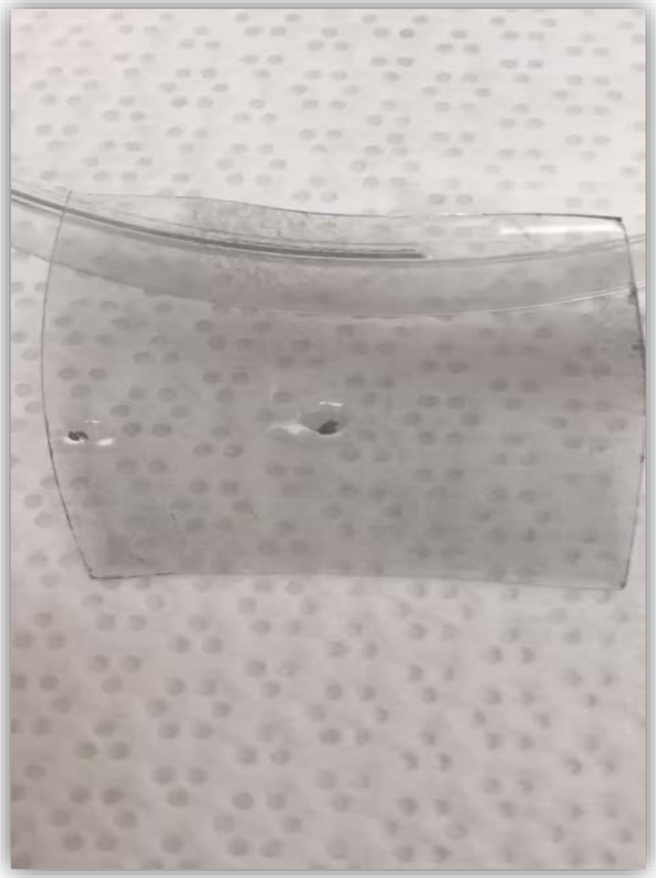
- Microscopy image datasets of various macroscale abrasion, nanoindentation, and nanoscratching experiments.
- New nanoindentation and nanoscratching protocols
- Mechanistic insight into how to make multifunctional coatings and surfaces more mechanical reliability with regard to abrasion

Thank-you

Please submit questions on life form or save for networking breaks.

Self-Cleaning Glass

Normal Glass



VS.

Self-Cleaning

(graphite flakes)

Project Proposal
(September 2019)

Project Title: Data Analytics for Benchmarking of Degradable Biopolymer Films and Tapes

Principal Investigator(s): Mostafa Bedewy

New Project: X

Thrust Area: 2. Materials Data Science

Abstract: With the increasing need for new flexible and degradable materials for wearable devices, biomedical sensors, and food packaging applications, a fundamental challenge is benchmarking of these materials against commercial polymers. A major roadblock however is that with biopolymers such as silk, there are differences in processing steps, some of which are not always reported in literature that impact the reported results. As a result, the causal relation between specific processing parameters and the degradation behavior are not well understood. For example, degradation of regenerated silk fibroins depends on the relative content of alpha- and beta-secondary structures relative to amorphous content. Hence, to be successfully incorporated in commercial products, such as tape-backings, food protection barriers, or wound-healing films, data analytics is needed to elucidate process-structure-property relationship towards creating a robust manufacturing process. In particular, it is proposed that time-series data for mass measurements is collected in a scaled set-up with parallel real-time measurements in standardized conditions such as in saline, water, or other liquid with recorded acidity. Moreover, multivariate analysis will be used to identify the process parameters that significantly influence the silk film degradation behavior. The statistical correlations obtained will be key for repeatable tunability of degradable silk products.

Keywords: degradable polymers, silk fibroins, flexible films, transparent materials, tape backings

Datasets Produced: The data sets that will be collected are time-series measurements of mass of samples that are processed according to a standardized set of meta data. Meta data for each sample will be shared on an online database in order to facilitate the adoption of the approaches developed in this project among the broader community Companies that could be interested: 3M could benefit from this project in creating degradable tape backings, food protection films, and perhaps flame-retardants. We will create optical, scanning electron microscopy, and scanning probe microscopy image datasets of various macroscale abrasion, nanoindentation, and nanoscratching experiments.

Short Bio: Dr. Mostafa Bedewy is an Assistant Professor of Industrial Engineering, Chemical & Petroleum Engineering (secondary appointment), and Mechanical Engineering & Materials Science (secondary appointment) at the University of Pittsburgh, where he leads the NanoProduct Lab (`www.nanoproductlab.org`). Before that, he worked as a Postdoctoral Associate at the Massachusetts Institute of Technology (MIT) in the area of bionanofabrication. In 2013, he completed his PhD at the University of Michigan in Ann Arbor, where he worked on studying the population dynamics and the collective mechanochemical factors governing carbon nanotube growth. Dr. Bedewy recently received the Outstanding Young Manufacturing Engineer Award from the Society of Manufacturing Engineers (SME) in 2018, the Ralph E. Powe Junior Faculty Enhancement Award from the Oak Ridge Associated Universities (ORAU) in 2017, the Robert A. Meyer Award from the American Carbon Society in 2016, the Richard and Eleanor Towner Prize for Distinguished Academic Achievement from the University of Michigan in 2014, and the Silver Award from the Materials Research Society (MRS) in 2013. His research interests include nano- and micro-manufactuing, biology-assisted fabrication, surface engineering and coating technology, in situ materials characterization and metrology, and Cybermanufacturing.

Data Analytics for Benchmarking of Degradable Biopolymer Films and Tapes

Mostafa Bedewy, PhD

Assistant Professor, Industrial Engineering,
Chemical and Petroleum Engineering,
Mechanical Engineering and Materials Science
University of Pittsburgh

mbedewy@pitt.edu | www.nanoproductlab.org

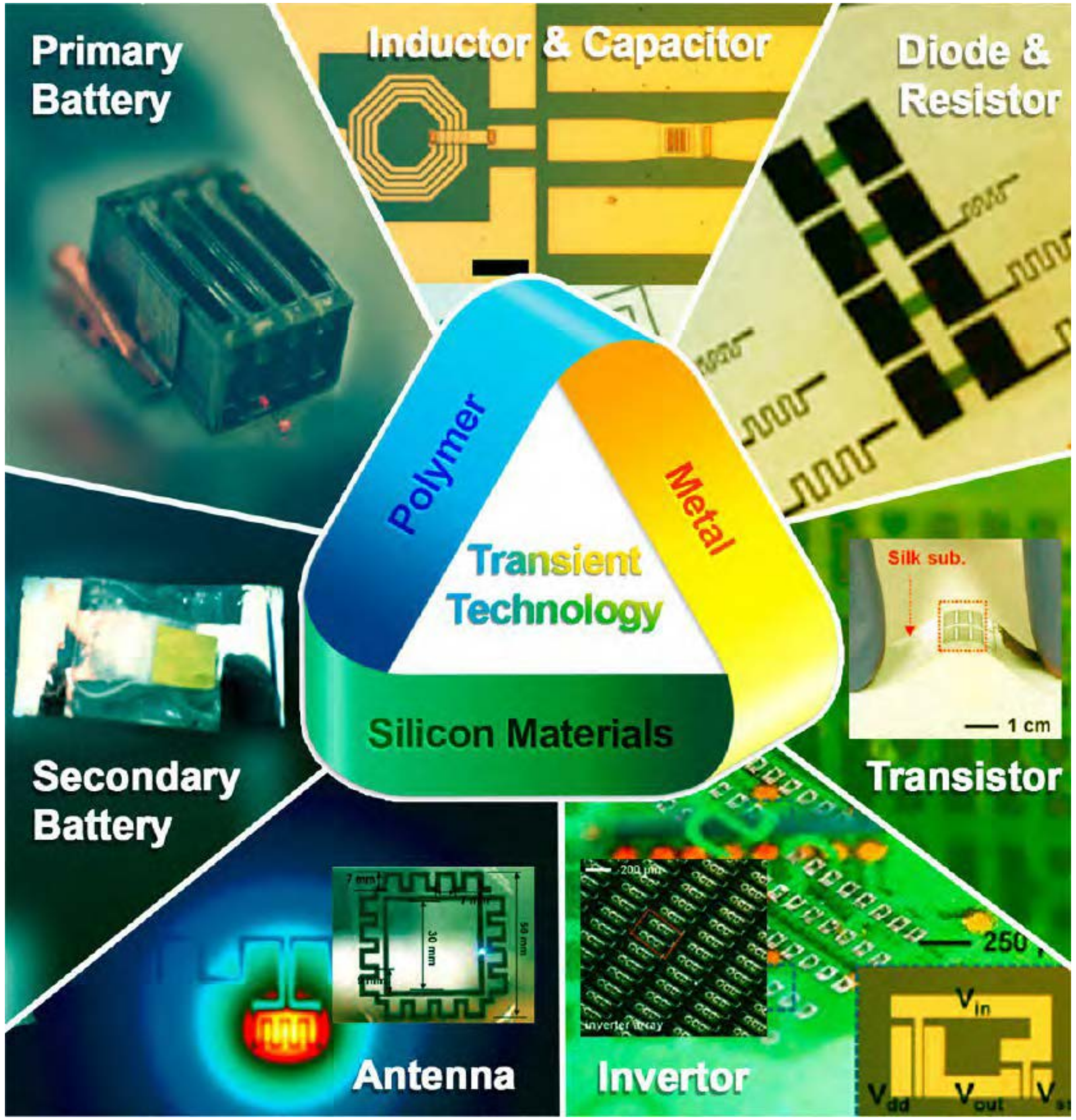
09/11/2019


Hi, I am Cathy!
I am really tall!

Why is it important to study degradable materials?

Degradable polymers for transient electronics

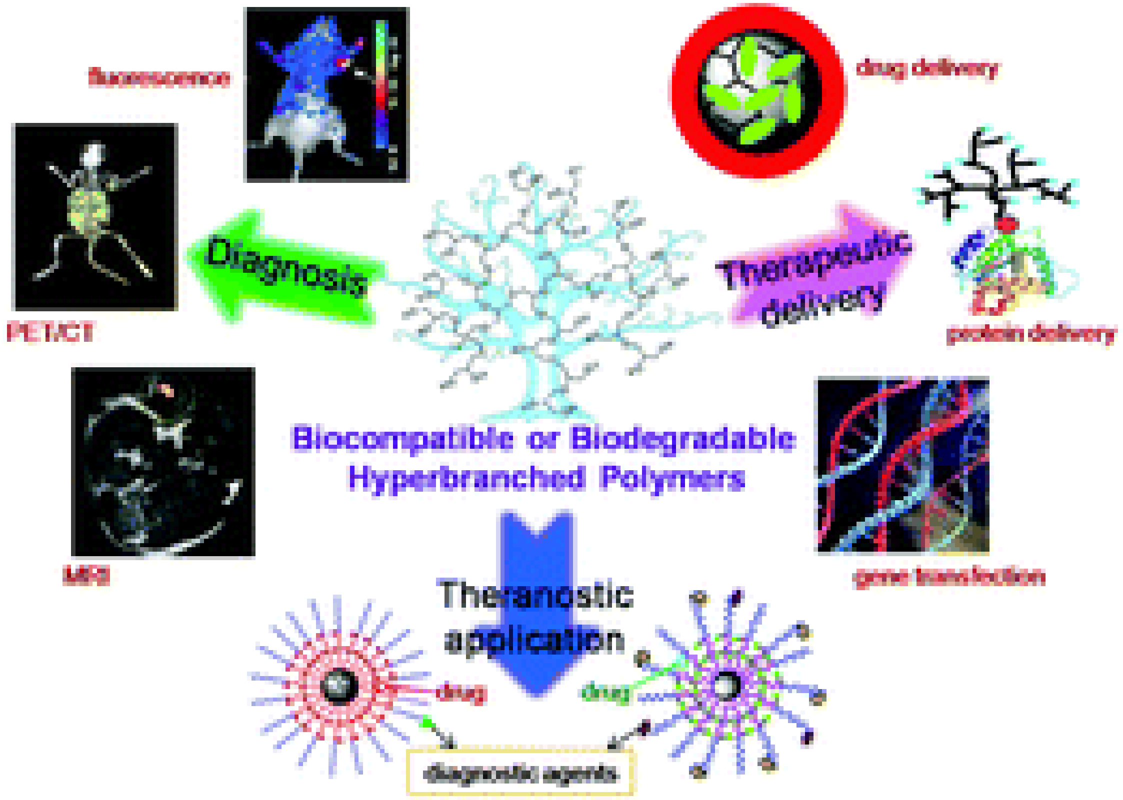
Devices that completely dissolve within ***a programmed period of time***



Fu el al. Chem. Mat. 2016

Degradable polymers for biomedical applications

Devices that completely dissolve ***in the body*** within ***a programmed period of time***

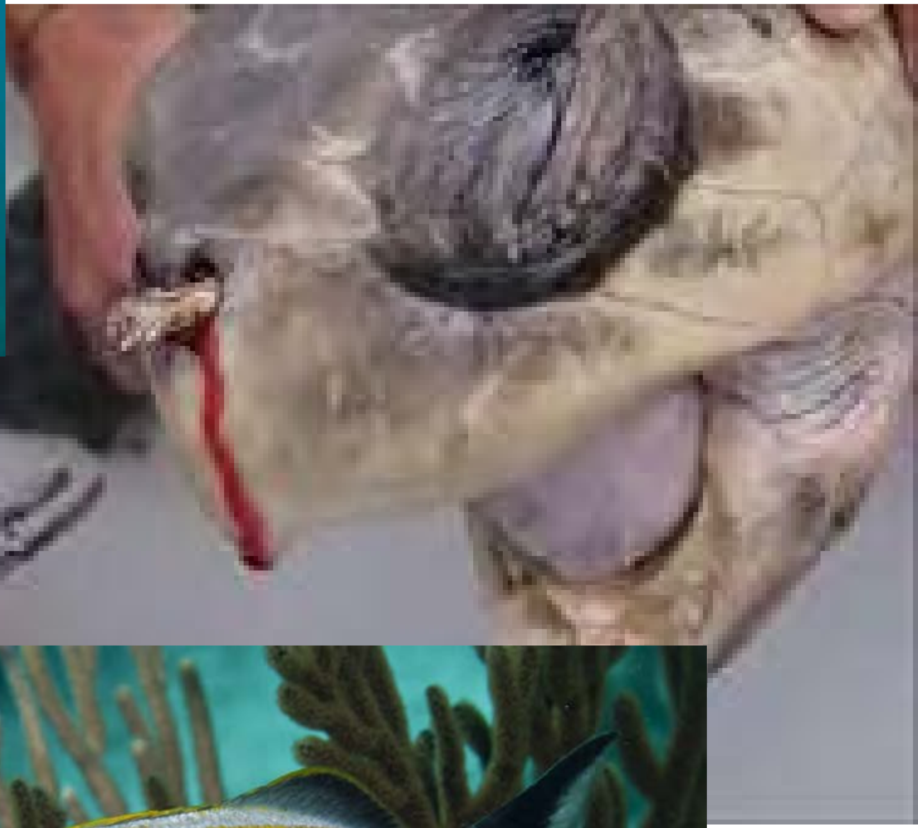


Huang el al. Pol. Chem. 2015

Earth’s Plastic Pollution



Plastic Pollution Affects Marine Life



What degradable materials this project focuses on?

Large portfolio of degradable materials

Semiconductors	Dielectrics	Conductors	Substrates
ZnO	SiO_x	Mg	Silk
IGZO	SiN_x	Zn	PLGA
Poly-Si	MgO	W	PLA
a-Si	SOG	Mo	PCL
Ge		Fe	POC
SeGe		pastes	Collagen
			Polyanhydride
			Metal foils

Adv. Mater. **26**, 7637 (2014).
Adv. Mater. **26**, 7371 (2014).
Adv. Mater. **26**, 3905 (2014).
ACS Nano **8**, 5843 (2014).
APL **105**, 013506 (2014).
Adv. Func. Mater. **24**, 4427 (2014).
Adv. Health. Mater. **3**, 515 (2014).
Small **9**, 3398 (2013).
Adv. Mater. **26**, 3905 (2014).
Adv. Func. Mater. **24**, 645 (2014).
Adv. Func. Mater. **23**, 4087 (2013).
Adv. Mater. **25**, 3526 (2013).

Large portfolio of degradable materials

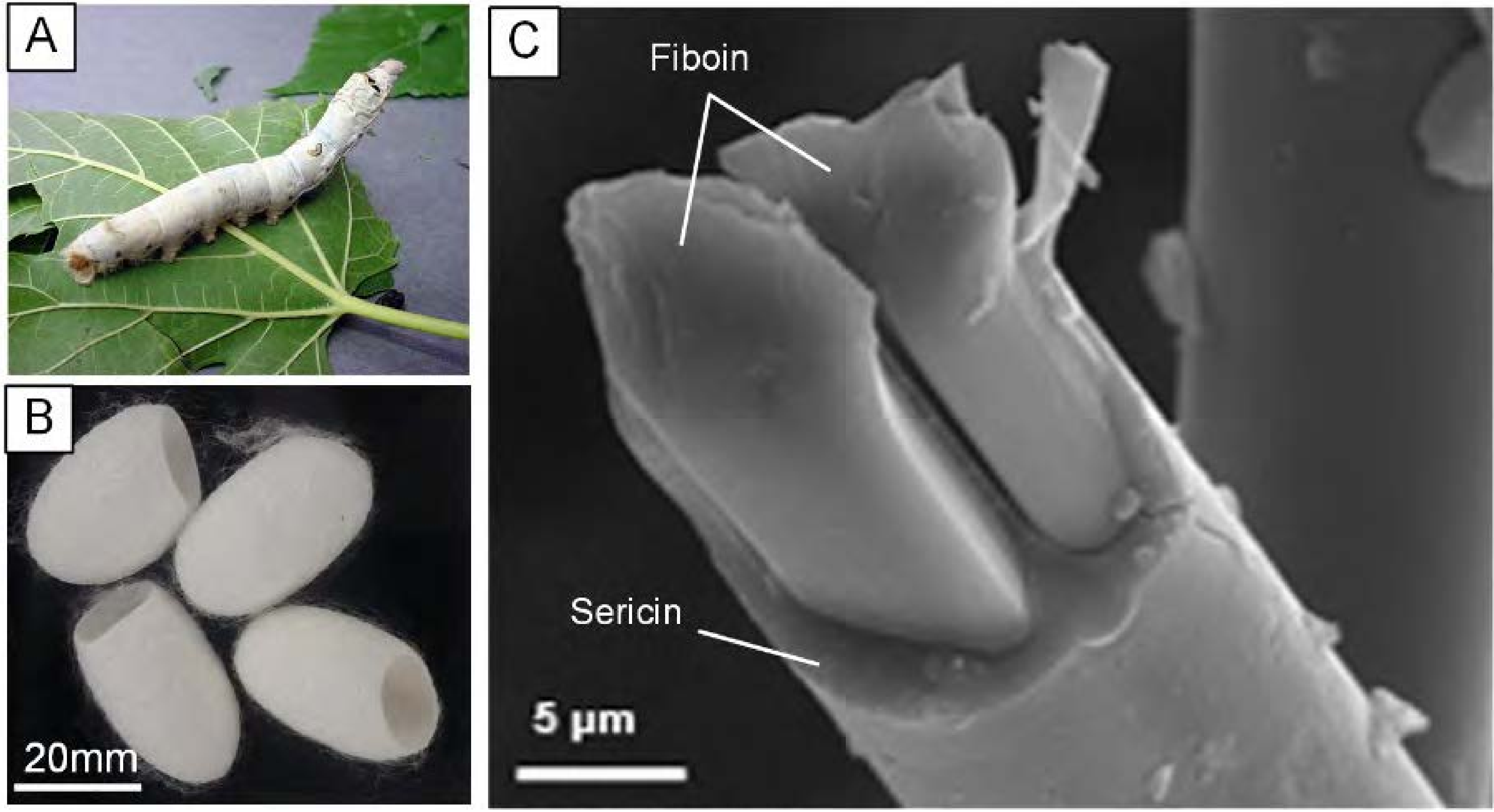
Semiconductors	Dielectrics	Conductors	Substrates
ZnO	SiO_x	Mg	**Silk**
IGZO	SiN_x	Zn	PLGA
Poly-Si	MgO	W	PLA
a-Si	SOG	Mo	PCL
Ge		Fe	POC
SeGe		pastes	Collagen
			Polyanhydride
			Metal foils

Adv. Mater. **26**, 7637 (2014).
Adv. Mater. **26**, 7371 (2014).
Adv. Mater. **26**, 3905 (2014).
ACS Nano **8**, 5843 (2014).
APL **105**, 013506 (2014).
Adv. Func. Mater. **24**, 4427 (2014).
Adv. Health. Mater. **3**, 515 (2014).
Small **9**, 3398 (2013).
Adv. Mater. **26**, 3905 (2014).
Adv. Func. Mater. **24**, 645 (2014).
Adv. Func. Mater. **23**, 4087 (2013).
Adv. Mater. **25**, 3526 (2013).

Why is silk an interesting material?

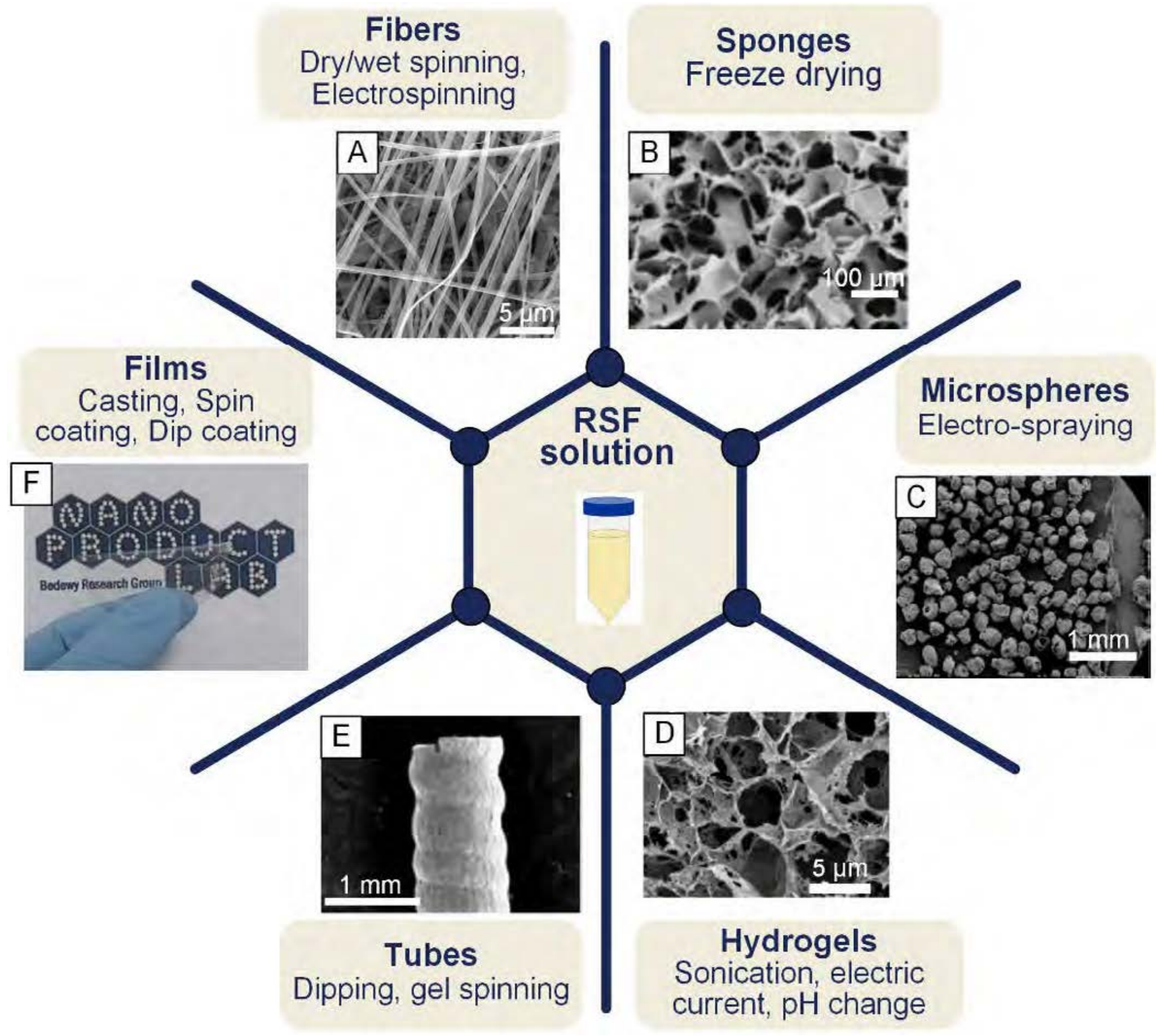
A closer look at silk from B. Mori (silkworm)

- Scalability
- sustainability



Bedewy et al., *in press*

Regenerated silk comes in many flavors

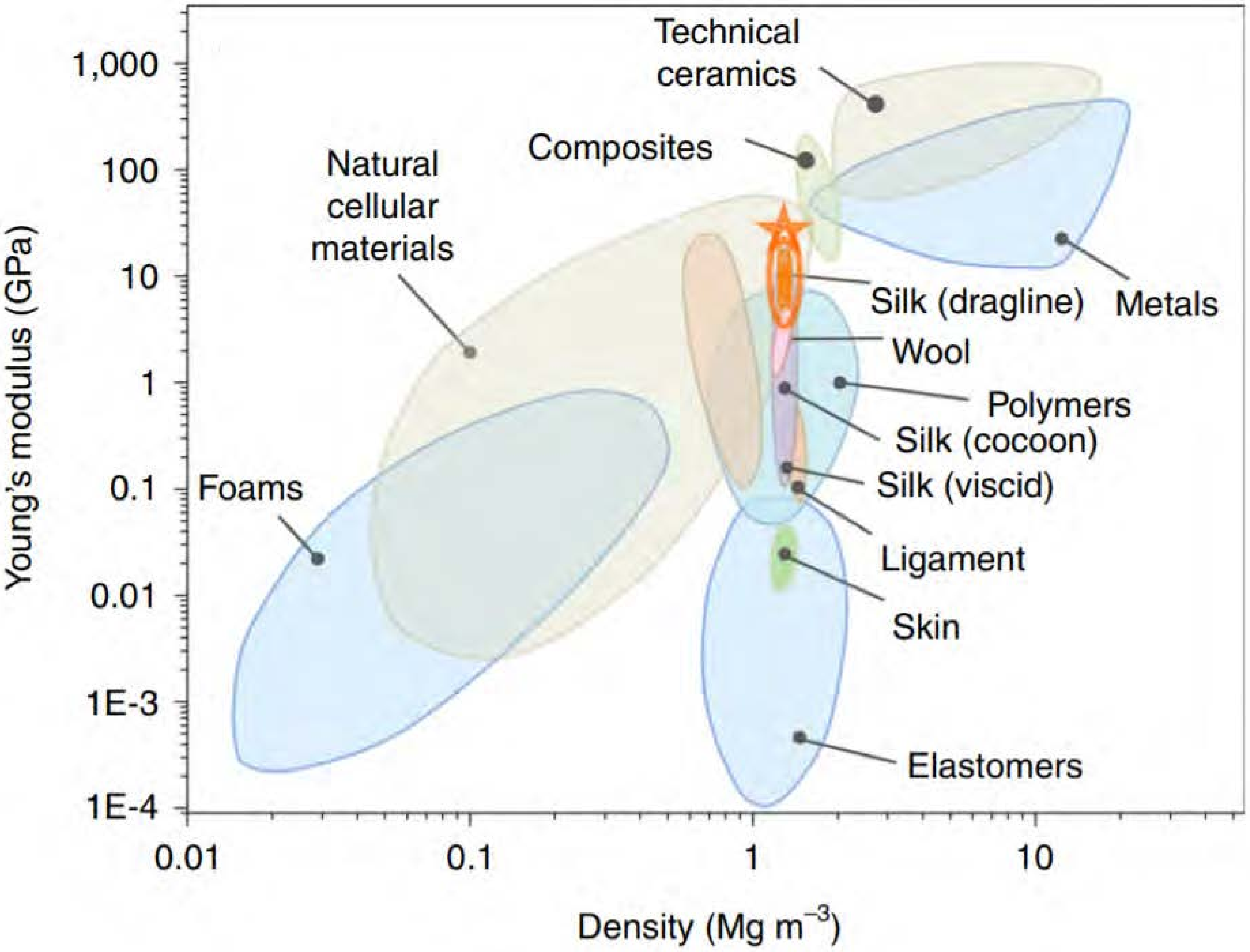


Bedewy et al., *in press*

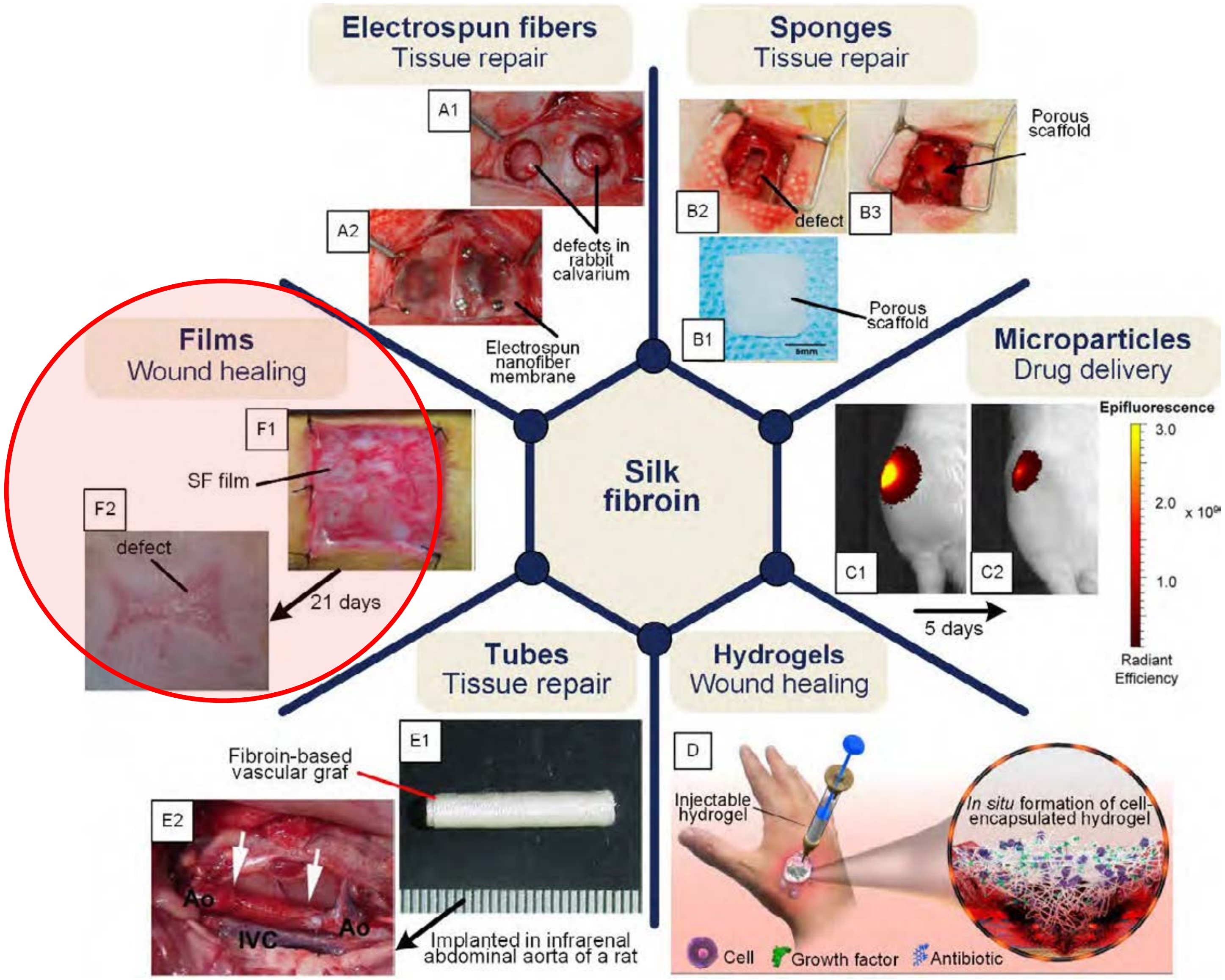
Silk's Mechanical Properties

- Flexible
- Lightweight
- High strength

Property		literature
Tensile modulus (single silk fiber with sericin, GPa)	5-12	(Viney et al. 1998)
Tensile modulus (single fibril without sericin, GPa)	15-17	(Pérez-Rigueiro et al. 2000)
Tensile strength (single silk fiber with sericin, MPa)	500	(Viney et al. 1998)
Tensile strength (single fibril without sericin, MPa)	610-690	(Pérez-Rigueiro et al. 2000)
Extensibility (single silk fiber with sericin, %)	5-12	(Viney et al. 1998)
Extensibility (single fibril without sericin, %)	4-16	(Pérez-Rigueiro et al. 2000)
Glass transition temperature (℃)	178	(Lewis TB and Nielsen LE 1970)
Thermal degradation (℃)	250	(Porter et al. 2009)
Density (g/cm^3)	1.3-1.8	(Shah et al. 2014)
Crystallinity (sheet content, %)	28-62	(Shah et al. 2014)
Moisture absorption (%)	5-35	(Shah et al. 2014)



Lots of biomedical applications for silk



Bedewy et al., *in press*

Other applications for regenerated silk fibroins
Films

Conformal Electronics

Nat. Mater. 9 (2010) 511

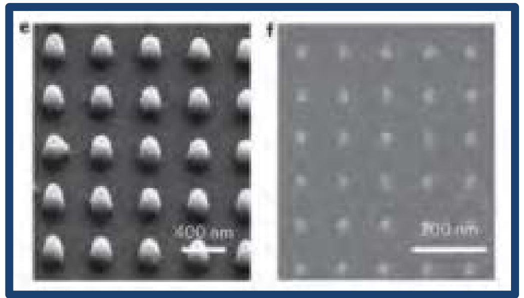
Flexible OTFT

Adv. Mater. 23 (2011) 1630

Bio-memristor

Adv. Funct. Mater. 22 (2012) 4493

Bio-resist



Nat. Nanotech. 9 (2014) 306

Edible Coating

Sci. Rep. 6 (2016) 25263

2009 2010 2011 2012 2013 2014 2015 2016 2017

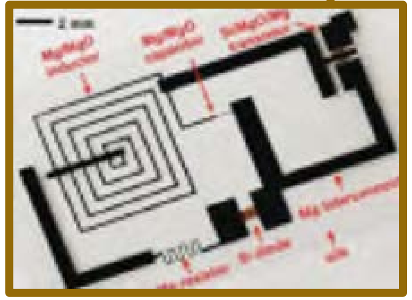
Appl. Phys. Lett. 95 (2009) 133701

Science 337 (2012) 1640

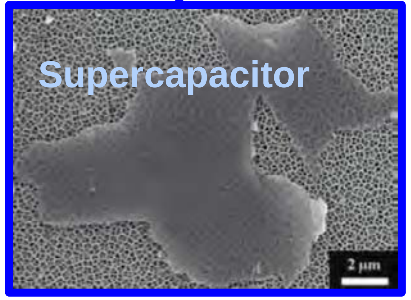
Nat. Commu. 3 (2012) 763

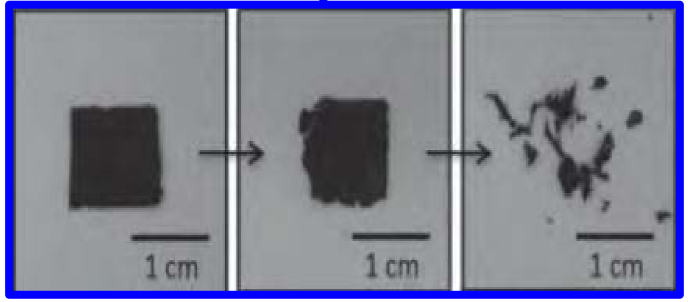
Adv. Mater. 25 (2013) 1993

Adv. Funct. Mater. 26 (2016) 1454



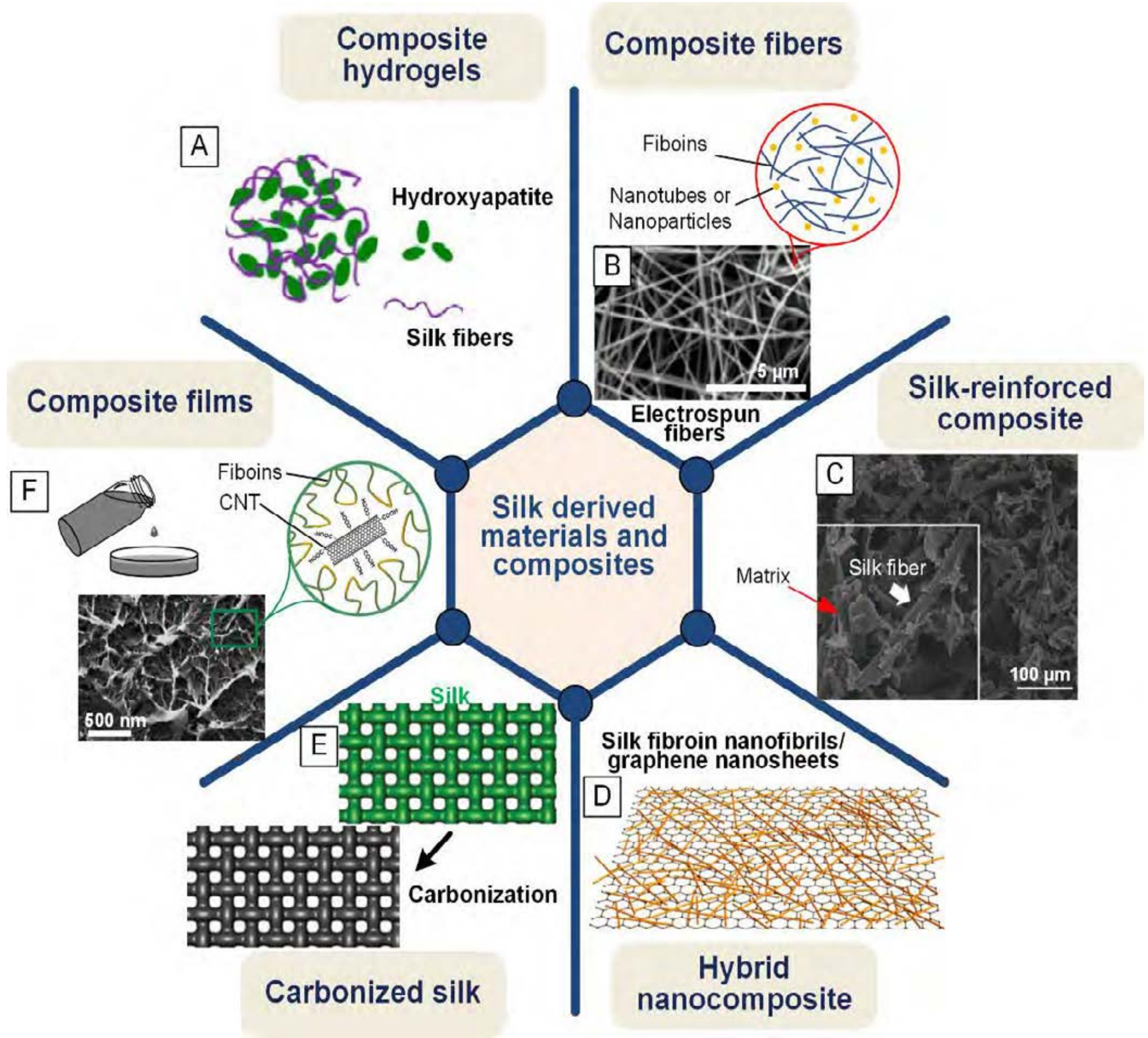
Bioresorbable & Transient electronics





Electrode for energy storage

Beyond silk ➔ silk-derived materials

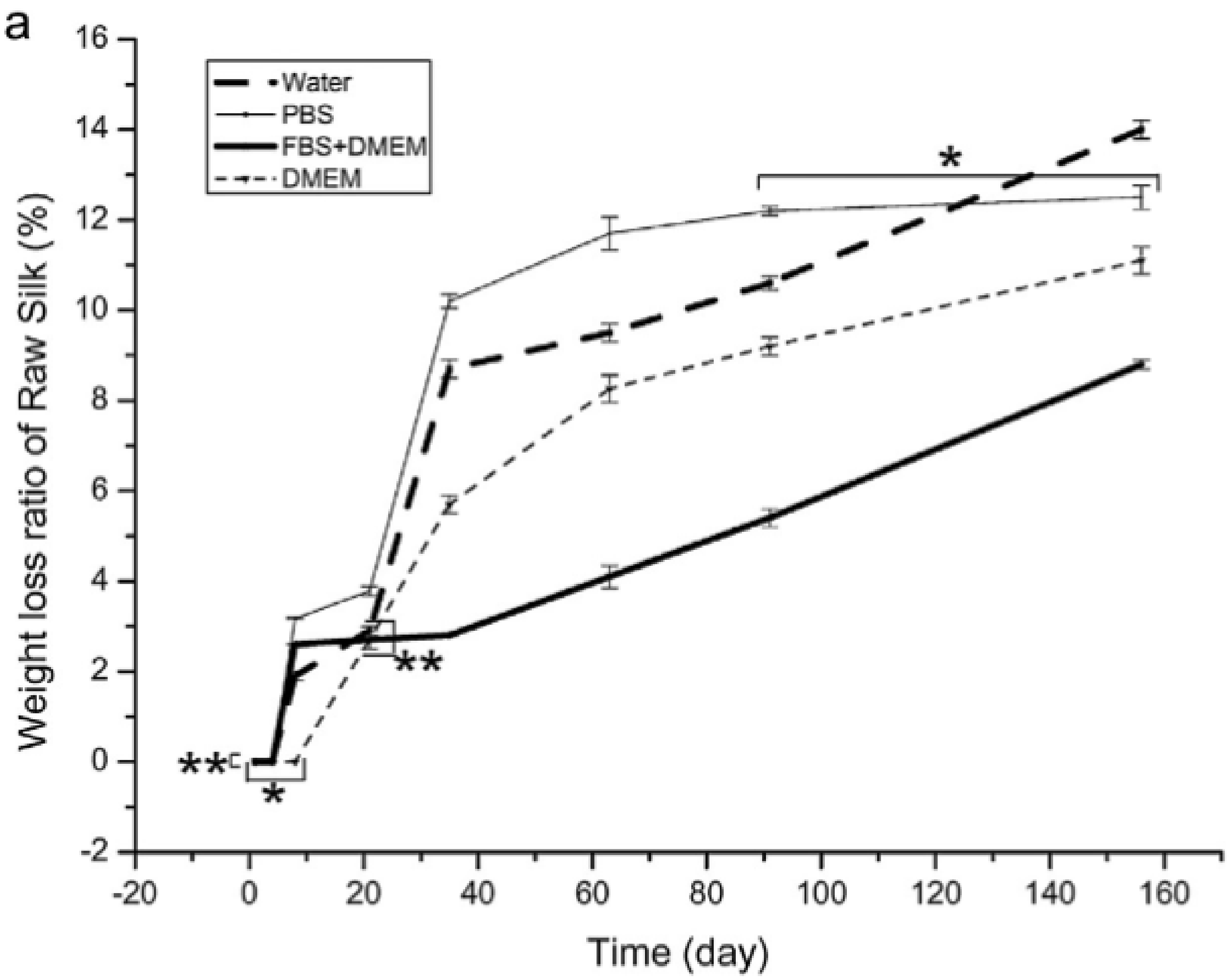


Bedewy et al., *in press*

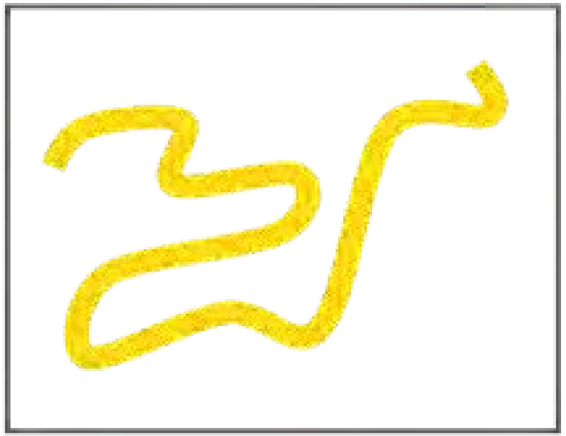
Challenge of
tuning the degradation behavior?

How silk degradation dare are reported

- Process-structure-property relation are not always clear
- Missing a lot of meta data
- Difficult to compare or build on published data



Secondary structure governs film properties



Random coils

Post-treatment

dehydration

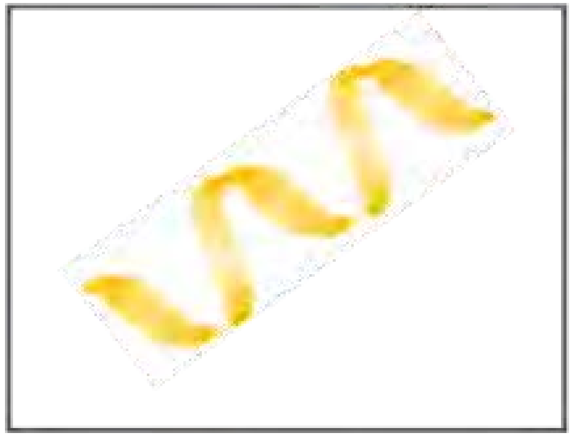
pH, Temp., Alcohols, aqueous solution with salts, hear force…

α-helix rich structure

Hydrophilic

Flexible

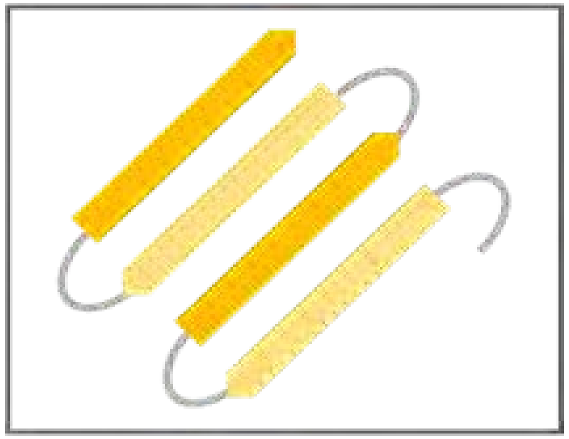
Faster degradation



Hydrophobic

Robust

Slow degradation



β-sheet rich structure

slow

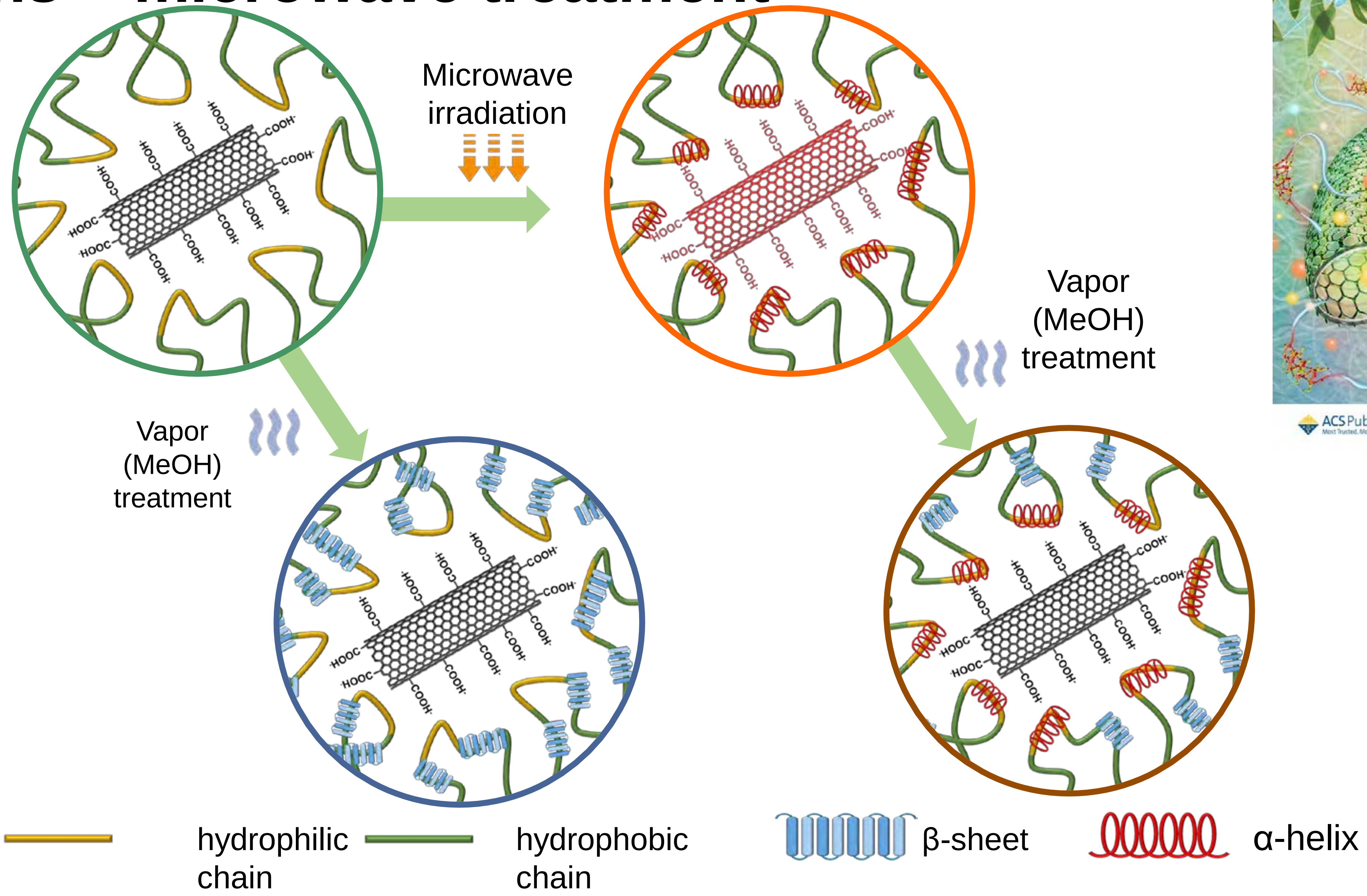
dehydration rate

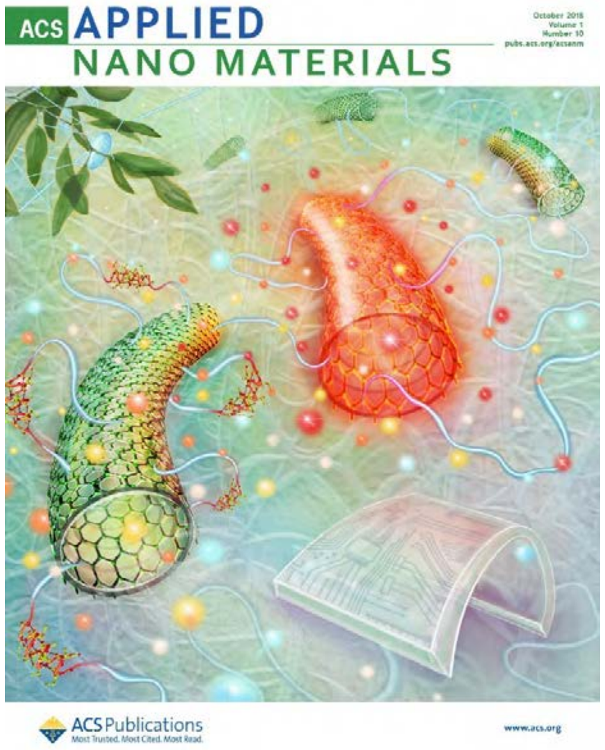
fast

Different processing methods lead to different content percentages of secondary structures

Treatment	Processing stage	Dominant resulting secondary structure	Notes
Formic acid dissolution	Solvent treatment	β-sheet	Requires two steps of dissolution and film casting
1,1,1,3,3,3-hexafluoroisopropyl alcohol	Solvent treatment	α-helix	Risk of chemical residues for biological application
Glycerol	Solvent treatment	α-helix	Glycerol can be totally dissolved in water
Room temp. drying	During-casting	α-helix	Drying time more than 48 hrs
Casting with controlled temp. in oven	During-casting	β-sheet	Slow drying in oven is needed
Water annealing	Post-processing of films	α-helix	High conc. of sol. (8 wt%) and thick film (100-200 μm) is required to have water-insoluble films
Temp. controlled water vapor treatment	Post-processing of films	Temp. dependent β content	High reproducibility Requires vacuum, Requires 12 hrs of treatment
Methanol treatment	Post-processing of films	β-sheet	Fast (10 min to 1 hour)
Controlled heating	Post-processing of films	β-sheet	Transition to β-sheet starts at 140 °C

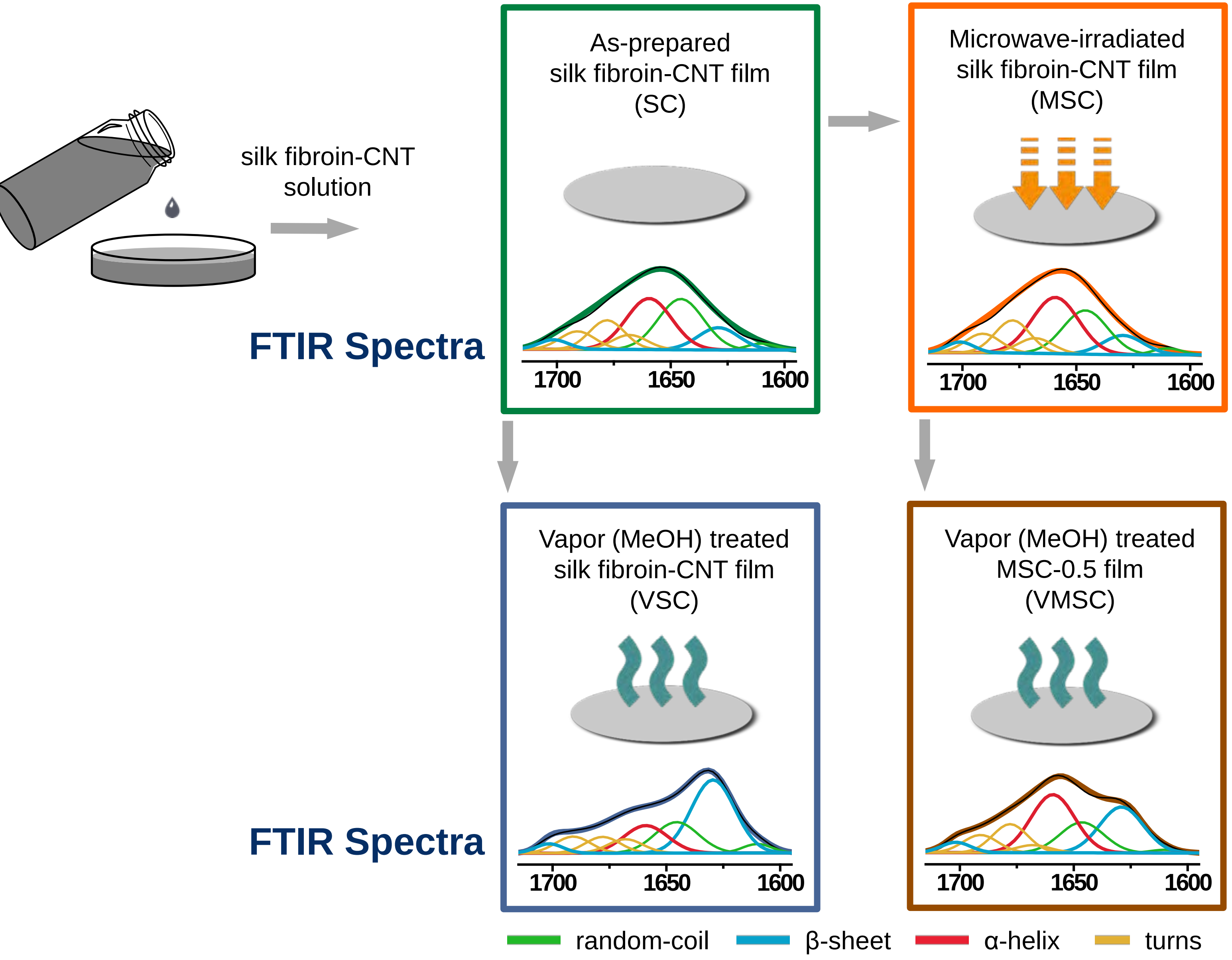
Molecular interaction between CNT and fibroins + microwave treatment

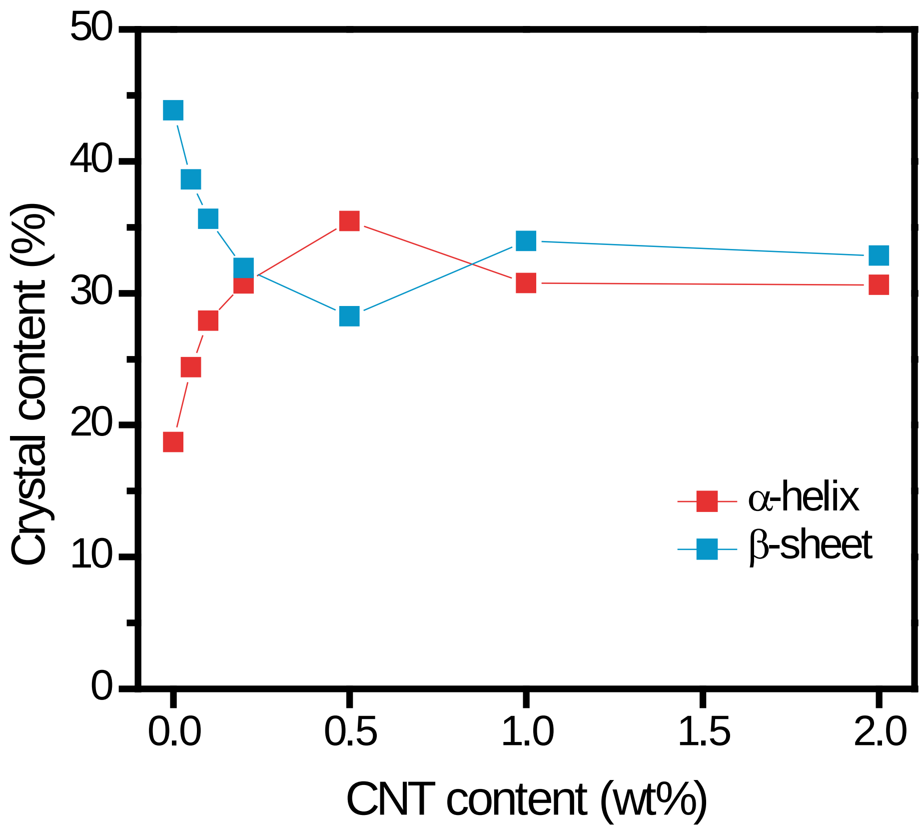




How structure is characterized?
FTIR, XRD, …etc.

FTIR: Microwaves + CNT induce α-helix formation in RSF

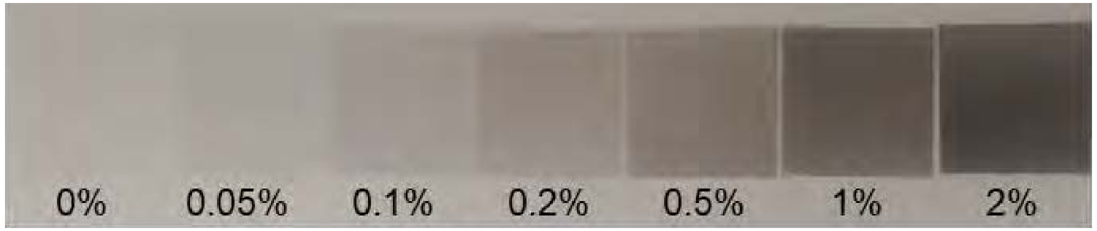


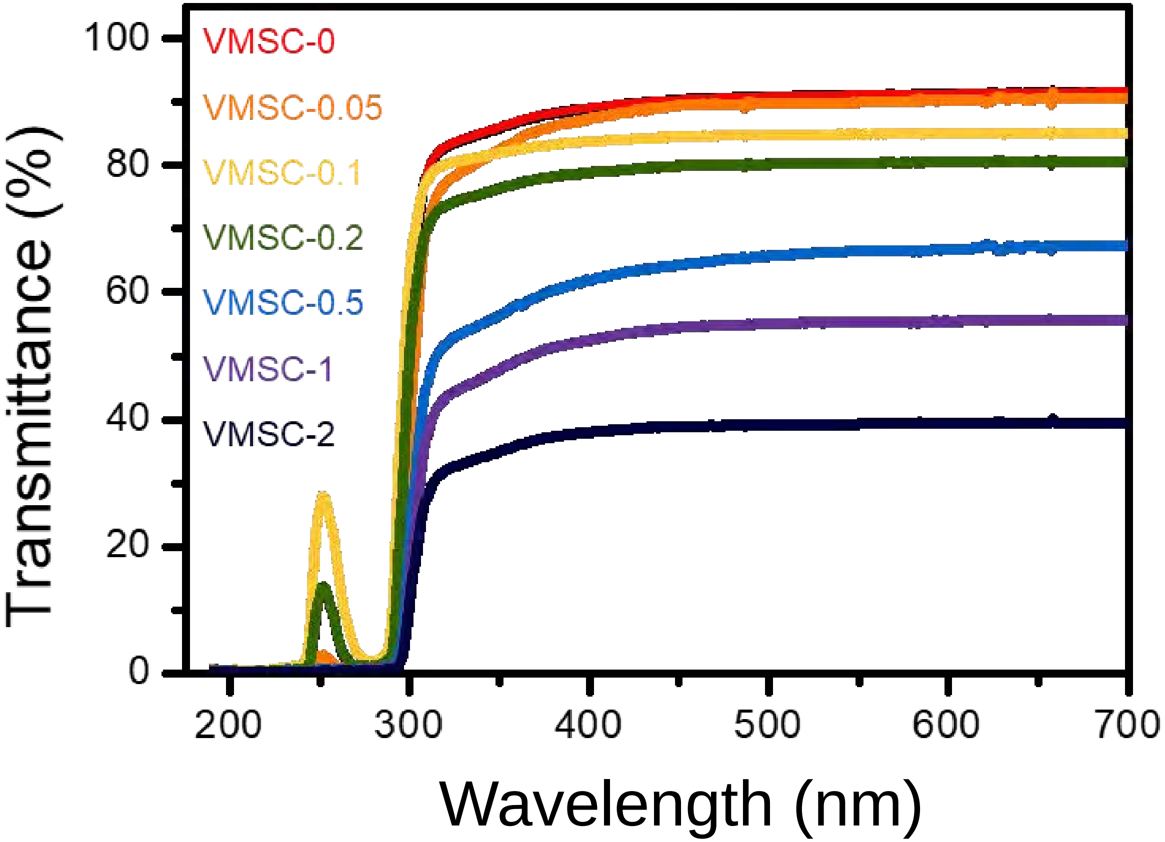


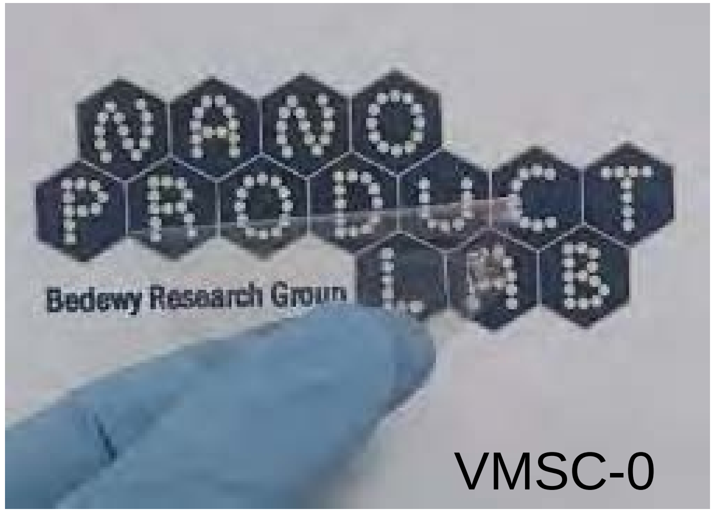
Standardization of data analysis is needed

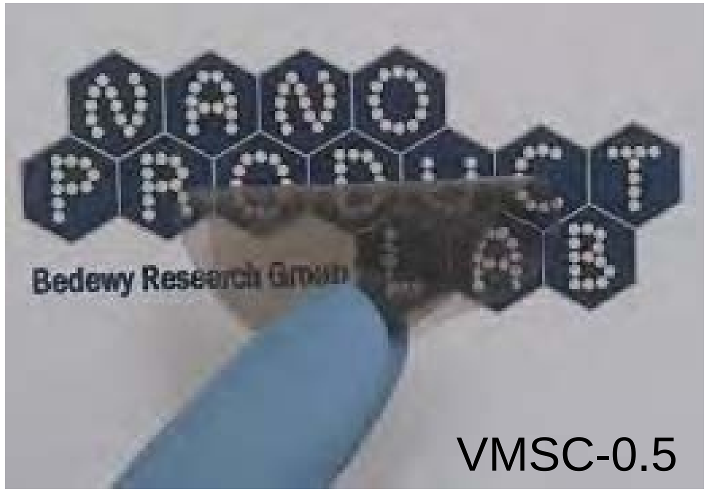
Our silk films

Highly transparent + flexible films

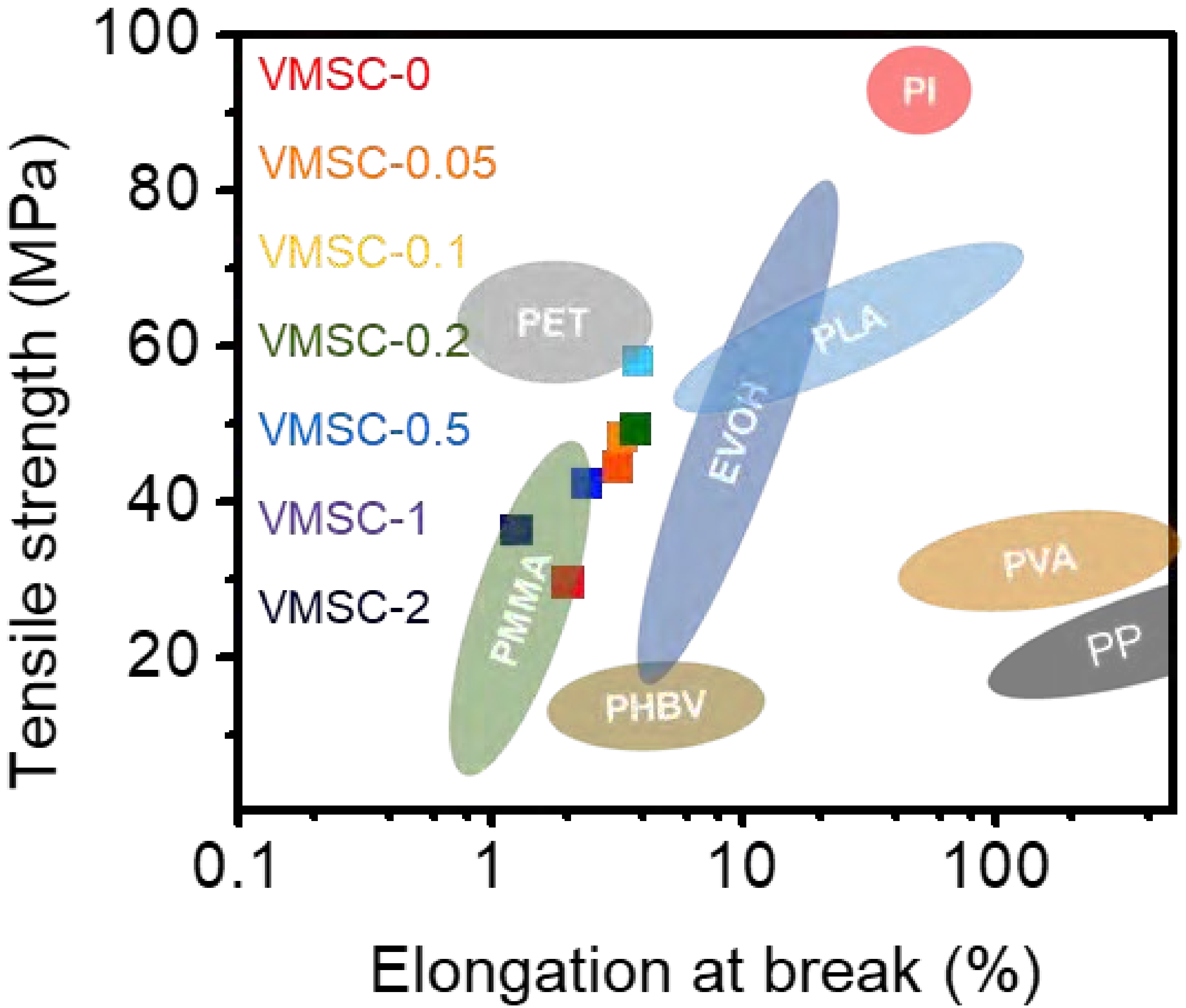








Enhancements make RSF films competitive



Tuning degradation by microwaves

SC-0.2

5 sec

MSC-0.2

30 sec

VMSC-0.2

10 min

Proposed project
2 years

Data to be collected

- Time-series data for mass measurements collected in a scaled set-up with parallel real-time measurements
- Standardized conditions of degradation environment such as in saline, water, or other liquid with recorded acidity.
- Fourier Transform Infrared (FTIR) spectroscopy data for processed silk films
- Detailed processing parameters (meta data)

Proposed activities

1. Create a standardized framework for reporting processing parameters of regenerated silk
2. Standardizing the structure of characterization approach for reliable quantification of secondary structures
3. Identifying quantifiable property to be measured and optimized (depending on the intended application)
4. Statistical Analysis: Analysis of variance (ANOVA), Multivariate analysis to identify the structural and processing parameters that correlate significantly to the measured properties of interest ➔ build process-structure-property relation

Deliverables

- Data driven understanding of process-structure property relation of silk-based films and tape backings
- Creating silk films with tunable properties for specific applications
- Benchmarking against commercial degradable polymers

Potential applications

- Degradable tape backings
- Fire retardant films
- Food protection barriers
- Wound healing films
- Other films with multifunctionality (optical, electrical, surface energy, biointerfaces)

Bedewy Research Group

NANO PRODUCT LAB

Website:
www.NanoProductLab.org

Email:
mbedewy@pitt.edu

/NanoProductLab

/NanoProductLab

/channel/UCYTIjKLZjj0M_ig9IK3Onig

NanoProduct Lab is supported by

UNIVERSITY OF PITTSBURGH

NSF

Center for ENERGY
A University Center within the Swanson School of Engineering

PITT | SWANSON ENGINEERING
INDUSTRIAL
CHEMICAL & PETROLEUM

PITT | UPCAM
CENTER FOR ADVANCED MANUFACTURING

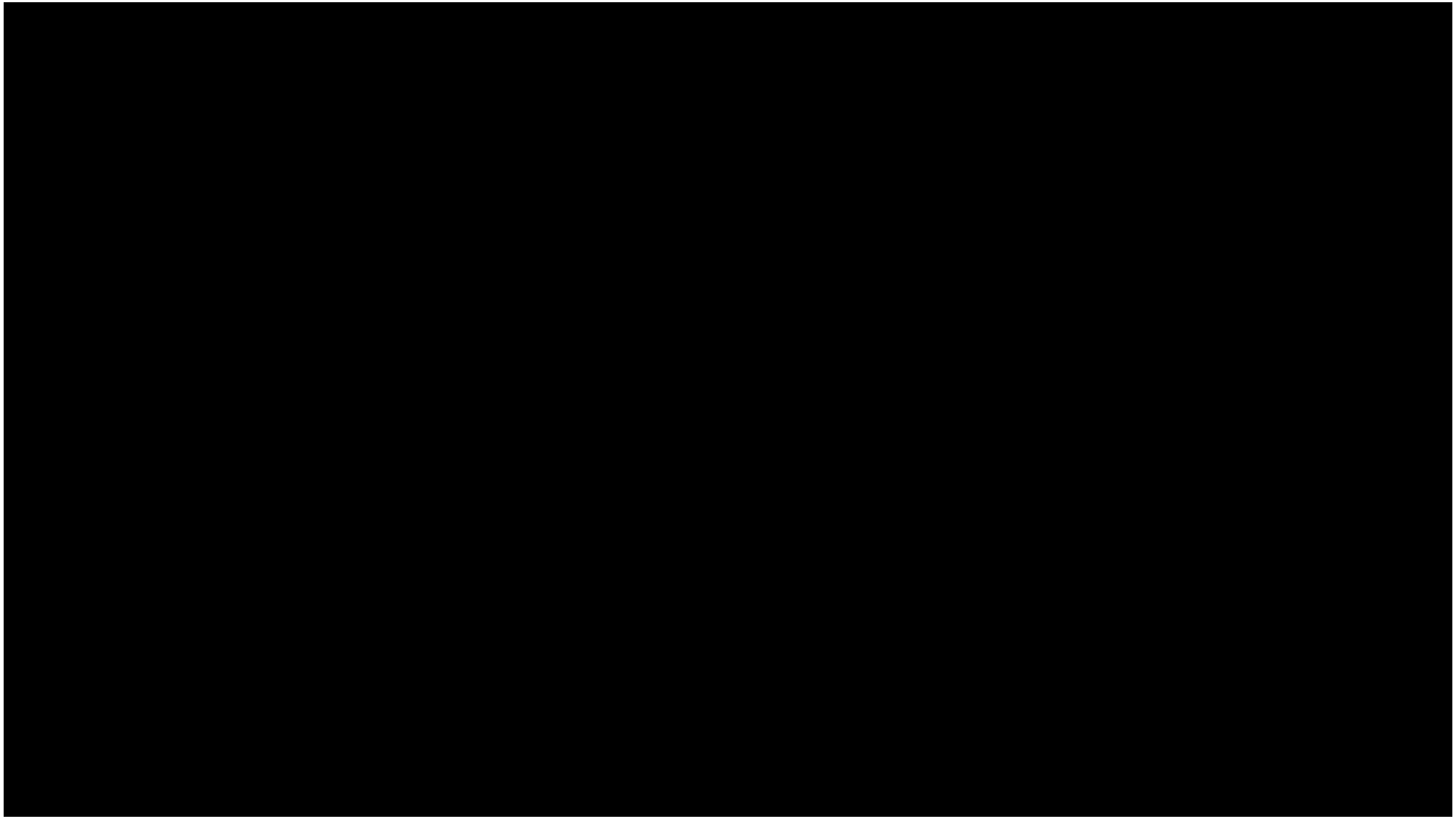
UNIVERSITY OF PITTSBURGH
Mascaro Center for Sustainable Innovation

GERTRUDE E. & JOHN M. PETERSEN INSTITUTE OF
NanoScience and Engineering

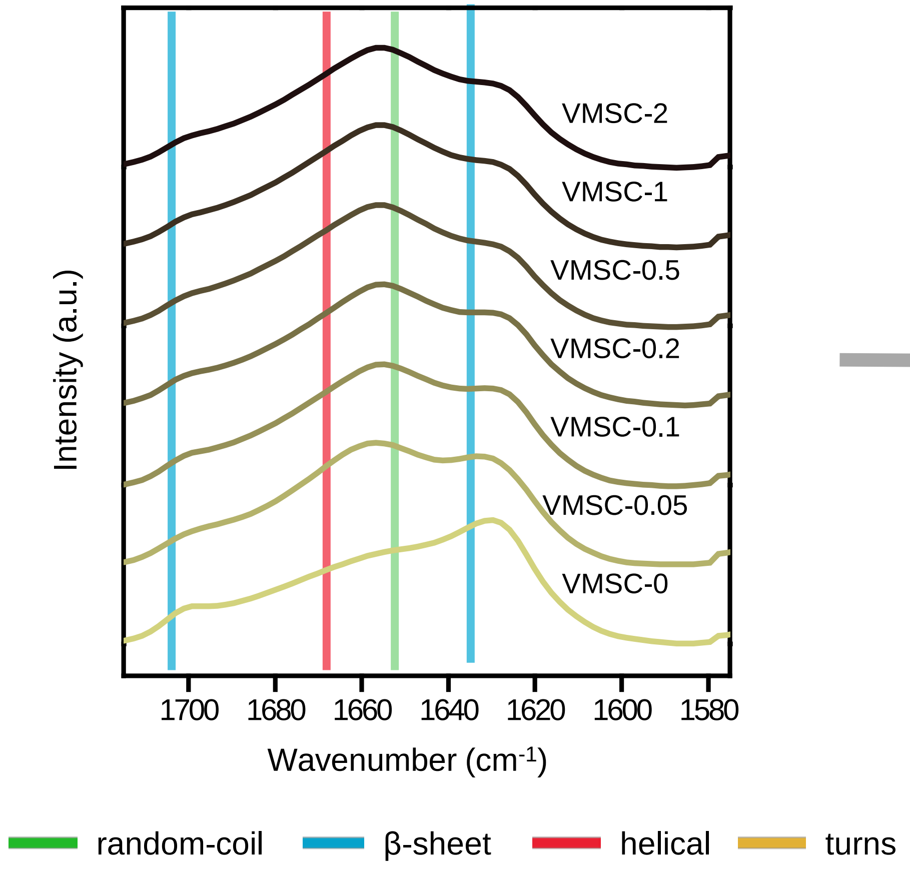
ORAU

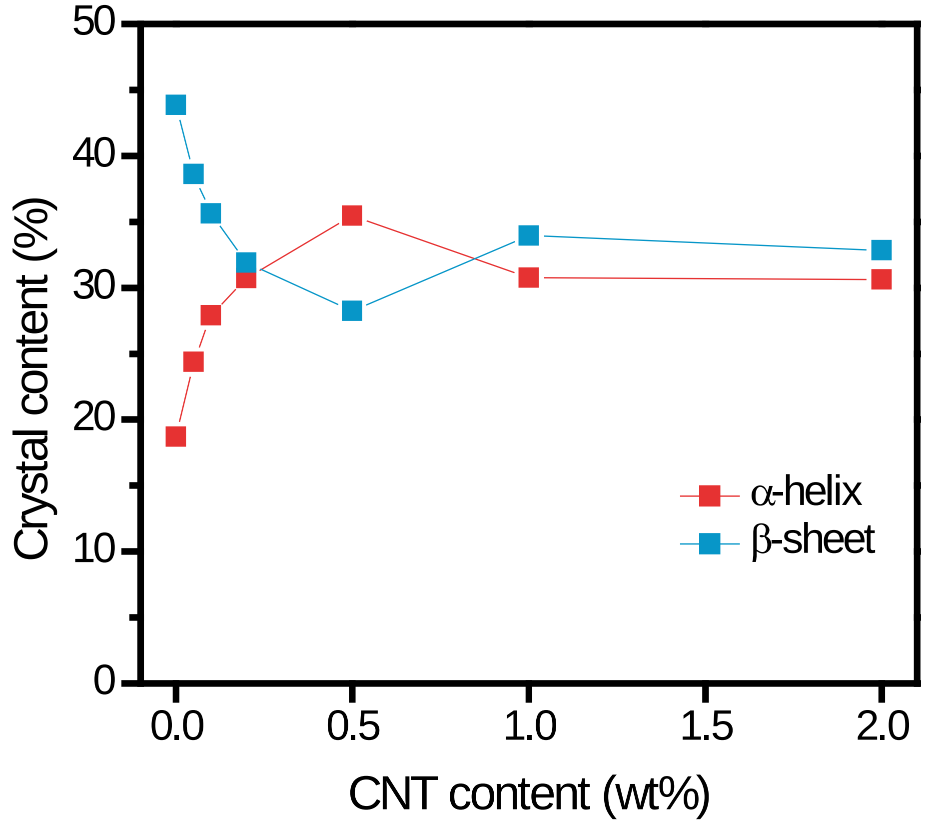
PITT CYBER
University of Pittsburgh
Institute for Cyber Law, Policy, and Security

Tuning degradation by microwaves

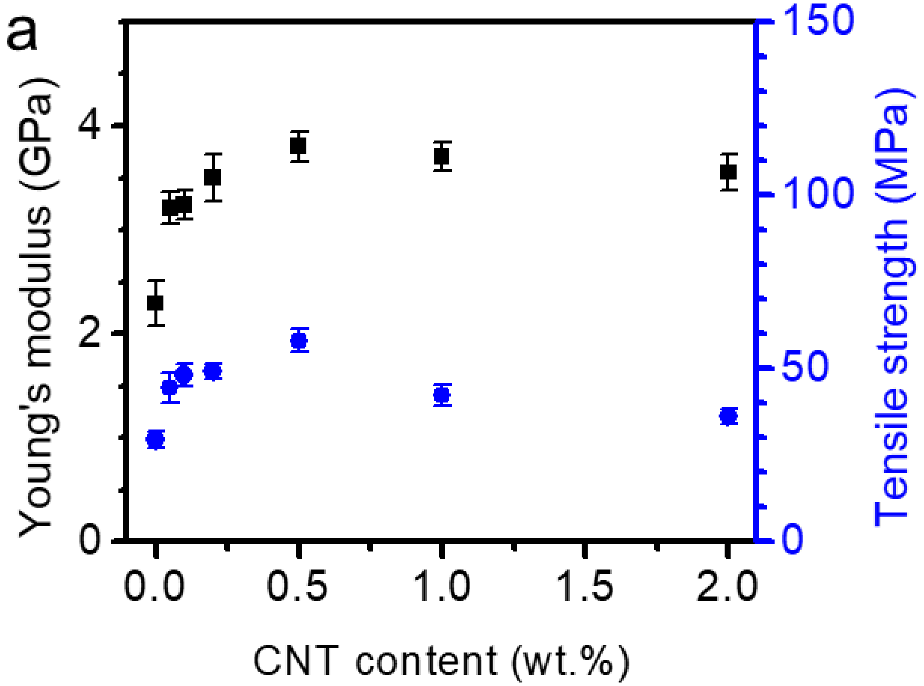


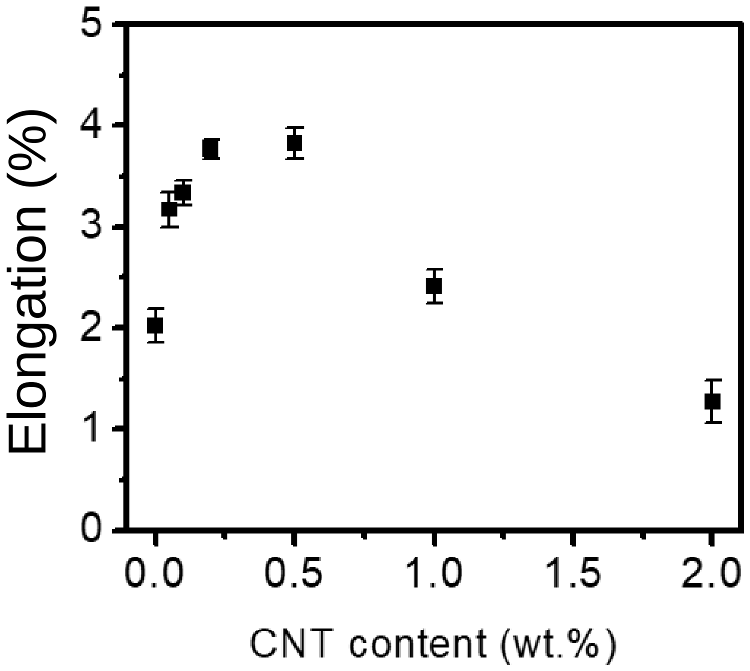
Influence of CNT% on secondary structure

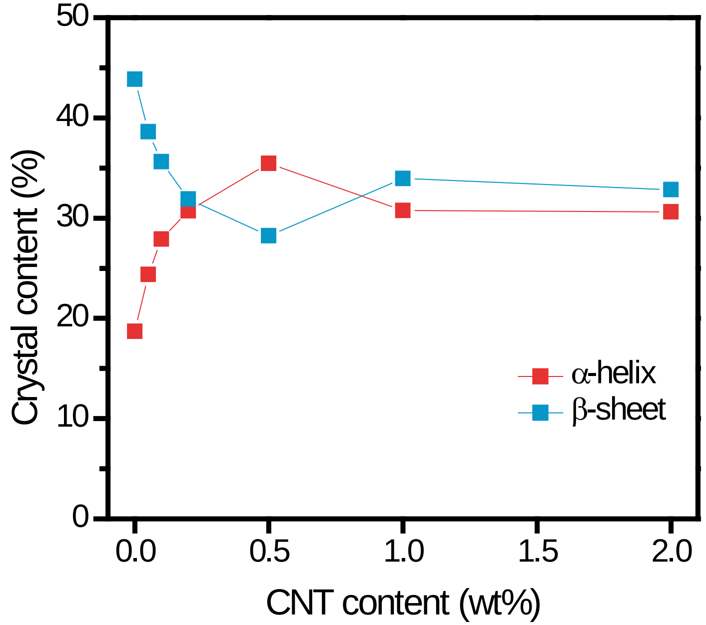




CNT increase mechanical properties

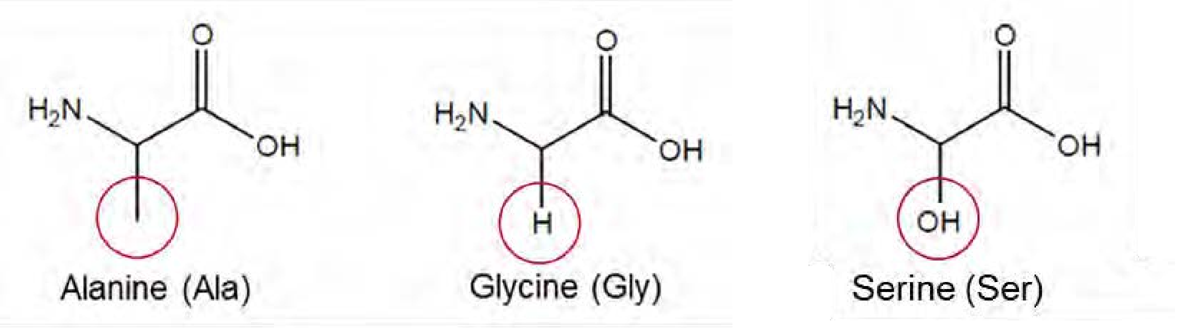




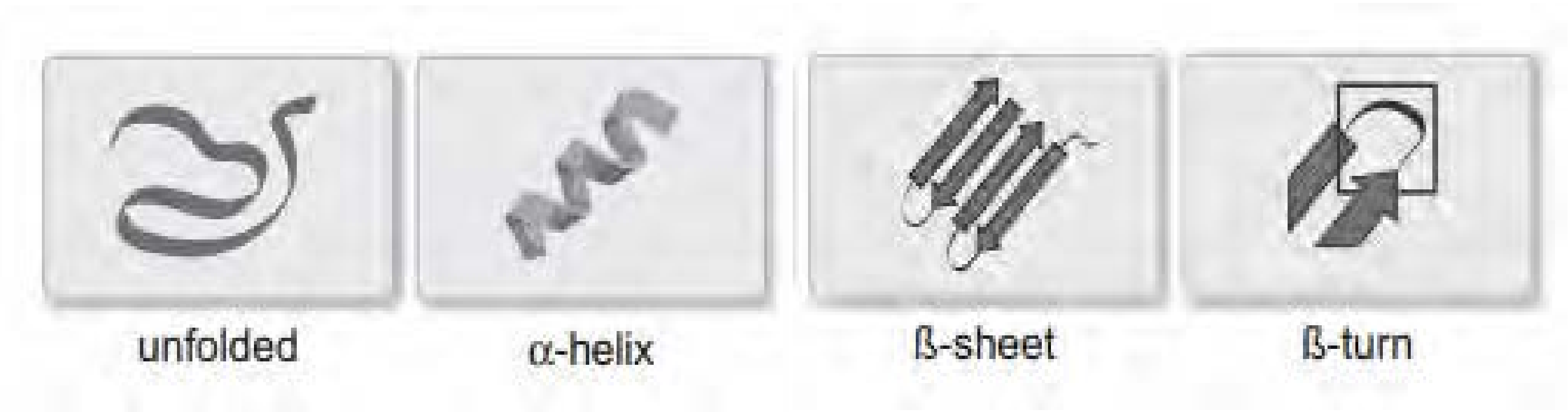


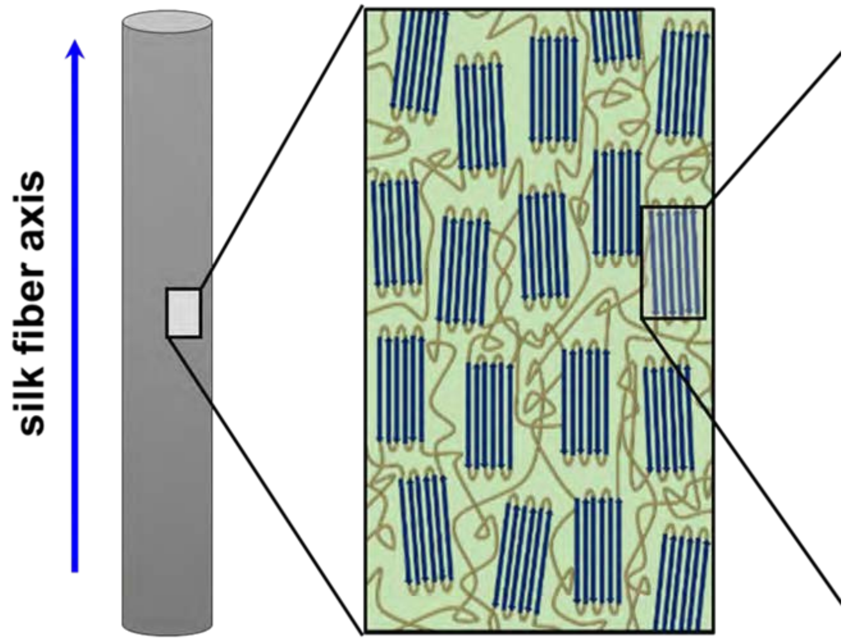
Silk Fibroin Primary Structure

Amino acid	Mol%
Alanine (A)	31.3
Arginine (R)	0.3
Asparagine (N)	0.4
Aspartate (D)	0.5
Cysteine (C)	0.1
Glutamine (Q)	0.2
Glycine (G)	45.9
Histidine (H)	0.1
Isoleucine (I)	0.2
Leucine (L)	0.1
Lysine (K)	0.2
Methionine (M)	0.1
Serine (S)	12.1
Threonine (T)	0.9
Tryptophan (W)	0.2
Tyrosine (Y)	5.3
Valine (V)	1.8



Silk Fibroin Secondary Structure



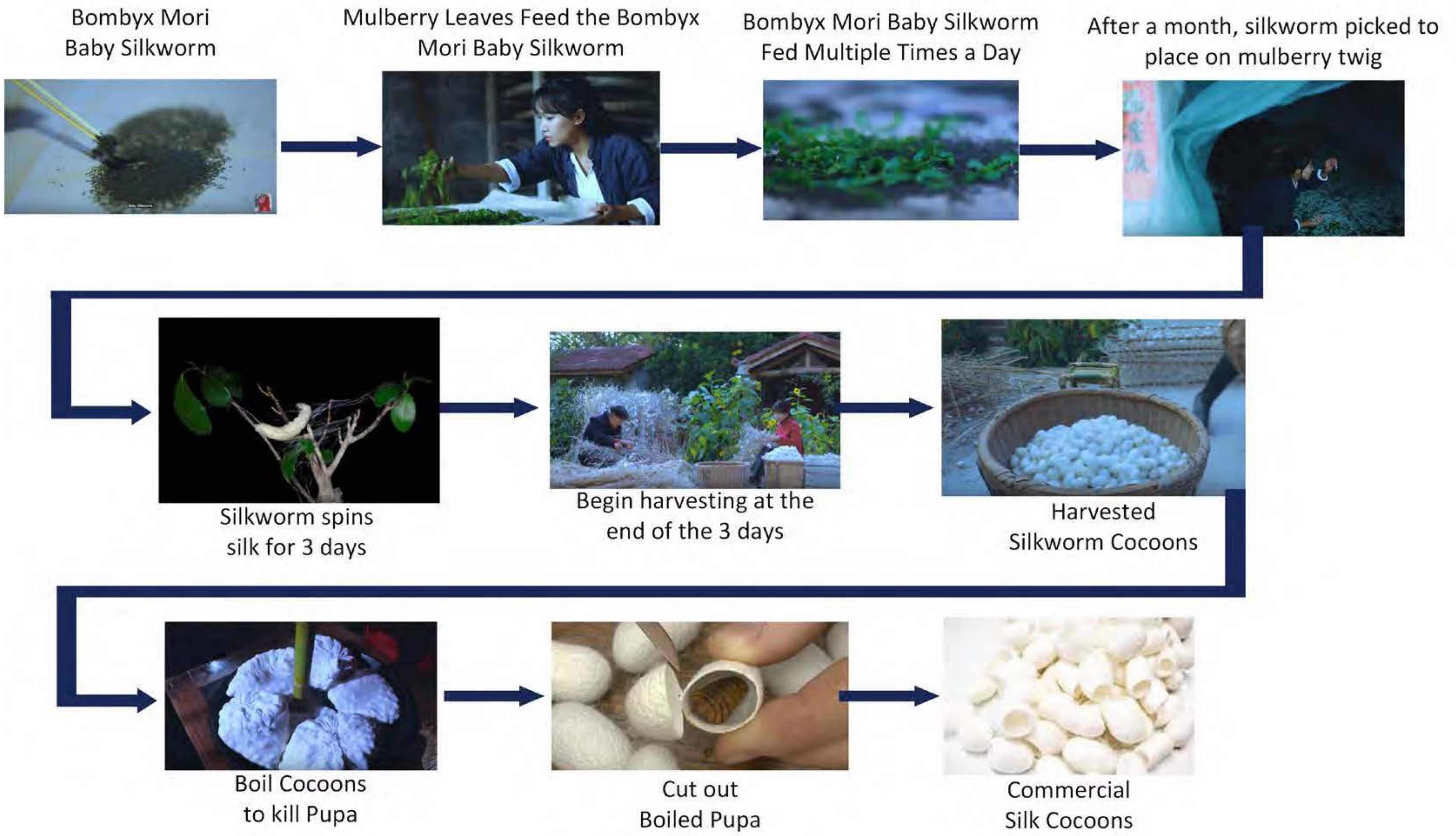


Crystalline Region

Silk Applications for Sustainability to Alleviate Plastic Pollution

Solution	Advantages	Limitations	References
Natural Silk Replace oil-based plastic for Facial Scrubs and Toothpaste	Replacing microbeads of plastic in face scrubs and toothpaste with silk decreases nanoparticles of plastic into ocean Can use the cocoons that would otherwise be wasted Secured for next few years	Expensive because the (1.7 oz bottle is $95)	https://www.fastcompany.com/40505636/this-natural-liquid-silk-is-starting-to-replace-oil-based-plastic
Silk Dental Floss	Biodegradable dental floss replacing the plastic dental floss that doesn't degrade. Also packaging biodegradeable. Affordable ($3.99 for one case)	More expensive than dental floss	https://madebyradius.com/products/natural-biodegradable-silk-floss
Silk Fabrics	Multiple uses of pillowcases, hair accessories, clothing, fabrics, and blended with other fabrics Degrades fully when thrown out	Expensive Smaller percentage of the world can afford these products	https://www.leaf.tv/articles/uses-of-silk-fiber/
Silk Facial Cleanser Balls	Replaces cotton balls that are normally always used so replacement allows for full degradation	100 times more expensive than one cotton ball (~.001 cent each)	http://www.sciencebeautygal.com/tips-and-explainers/the-benefits-and-risks-of-using-silkworm-cocoons-on-your-face-a-guide-to-why-they-work-and-how-to-use-them

Sericulture Process



Silk Cocoons to Silk Regenerated Silk Films