DSCI 351-451
Hardcoat Acrylic Film Degradation:

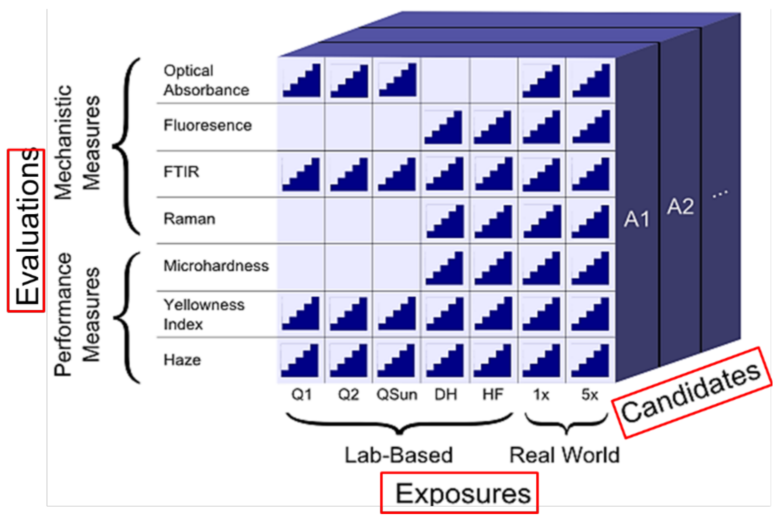
Project 1 Assignment
“Data Assembly, Cleaning, EDA and Summarizing”

netSEM modeling of Degradation Pathways

R. H. French, L. Bruckman, Alan Curran
Fall 2017

Purpose of Project 1 Assignment

In this film degradation Project 1 assignment, you will work on degradation data from outdoor exposure of hard-coat acrylic films on polyester (PET) and urethane (TPU) substrates. These data are taken step-wise, where the samples are exposure to real-world (outdoor) conditions or lab-based (accelerated) conditions of temperature, humidity and solar irradiance and then are evaluated using non-destructive techiques such as optical spectroscopy (see Figure 1.). In project 1 you will need to understand, document, clean, tidy, assembly and summarize single value and spectral data taken on these samples. You will then use the film degradation dataset, and the dataframe you assembled, EDA, Visualization and to determine structural equation model (SEM) network models of the degradation pathways.



In these project assignments you will need to develop code to manipulate, clean, tidy, summarize and visualize your data, and to answer various data science questions. You can develop straight R code scripts. But you will also need to use Rmd, with R code chunks to produce the report on your projects to turn in. You will turn in the R code scripts, the compiled Rmd pdf, and your final dataframe. **<u>It will not be acceptable to pre-process the csv files in excel to get an initial dataframe structure, everything must be done in R code</u>.** You will need to show your R code in the Rmd report that manipulates the individual files to build a single dataframe.

References:

1. Bruckman, L.S., Wheeler, N.R., Kidd, I.V., Sun, J., French, R.H., 2013. ***Photovoltaic lifetime and degradation science statistical pathway development: acrylic degradation*.** p. 88250D–88250D–8. doi:10.1117/12.2024717

2 Roger H. French, Rudolf Podgornik, Timothy J. Peshek, Laura S. Bruckman,Yifan Xu, Nicholas R. Wheeler, Abdulkerim Gok, Yang Hu, Mohammad A. Hossain, Devin A. Gordon, Pei Zhao, Jiayang Sun, Guo-Qiang Zhang **Degradation science: Mesoscopic evolution and temporal analytics of photovoltaic energy materials" Current Opinion in Solid State and Material Science**, doi:10.1016/j.cossms.2014.12.008

Sample Information

The file (acryhc-key.csv) contains information on each of the samples and each sample is identified with a unique sample number. These samples are either pure substrate (TPU and PET) or a substrate with a unique coating (identified by a #-substrate). There are a total of 240 samples, on two different substrate films (TPU and PET), with three different hard coat acrylic compositions (9006, 9013, 9025).

Exposure Information

The samples were exposed, in equipment as shown in Figure 2., to 6 different exposures. Some samples were never exposed and are step "0" samples. The exposures include Humidity/Freeze (HF), ASTMG155, ASTMG154, modified ASTMG154 (mASTMG154), an outdoor exposure (1x), and a concentrated light outdoor exposure (5x). HF is the only exposure that does not contain an irradiance stressor because it is a cycle of 85C at 85% RH to -70C with uncontrolled humidity. ASTMG155 is TUV irradiance dose with sprayed water at 70C. ASTMG154 is constant UV irradiance at 1.55 W/m^2 at 340 nm at 70°C. ASTMG154 is a cycle of 8 hours of UV irradiance at 1.55 W/m^2 at 340 nm at 70°C and then 4 hours of condensing dark humidity at 50°C.

Fig. 2. QUV (ASTMg154) SPHS100 (DH, HF) QSun (ASTMG155)

Exposure Steps

HF, ASTMG155, ASTMG154, and mASTMG154 have data on Steps 0-4 while the two outdoor exposures (1x and 5x) have data on Steps 0-3. Samples were exposed to a varying number of hours for the different exposure conditions, which can be seen in the file acryhc-exp.csv. The numbers given are the hours per step and not a cumulative.

Retained Samples

Samples were retained at different Steps on exposure (0-4) (See Fig. 3 for an example). Therefore, some samples were only exposed in Step 1 and then held out of exposure and other samples were exposed to only through steps 2 and 3 and were not exposed to subsequent steps. So a sample maybe retained at 0-4 steps.

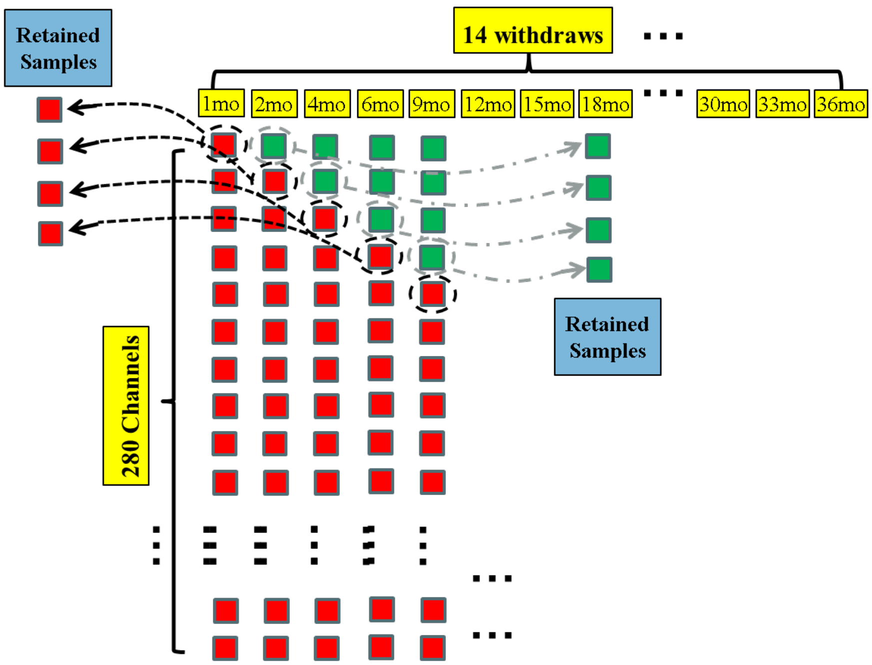


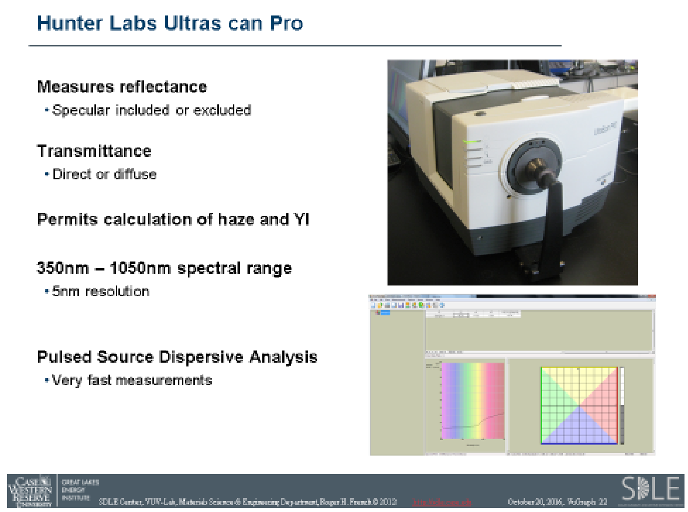
Fig. 3. An example of how a study protocol is structured to produce a retained sample library

HunterLab Data

HunterLab colorimetry measurements were taken on every sample prior to exposure and on every sample that was exposed in subsequent steps. These measurements include L*, a*, b*, YI (ASTM E313) and percent haze. This data is contained in the folder called "color" for steps 0-4. The data has not been cleaned and probably contains sample number mistakes. The data is contained in multiple csv files with multiple different naming conventions. These data are single value data integrated over a spectra.

FTIR Data

FTIR measurements were taken on retained samples only except for Step 4 for HF, ASTMG155, ASTMG154, and mASTMG154 and Step 3 for 1x and 5x. Therefore, there will be less data than with HunterLab and not all sample number will have FTIR associated with it. The data has not been cleaned and probably contains sample number mistakes.



You can use the Hyperspec R package to view the FTIR spectra, (library(hyperSpec). And the package has a vignette on File IO (hyperSpec::fileio).

Dose Basis

You will be comparing exposures and data on dose basis. Full spectrum and UVA-340 doses are important. The exposure (except HF) contain irradiance the irradiance levels and spectra are not the same'; therefore, in order to compare them, they must be converted to a dose basis. The table below will help you convert irradiance levels into Full Spectrum and UVA-340. For ASTMG155, the last column applies (QSun-TUV at 70 W/m^2). For ASTMG154 and mASTMG154, the second to last column applies (QUV 340 at 1.55 W/m2). Be aware that ASTMG154 is not constant irradiance and has a dark cycle when converting to a dose basis. To convert to dose basis take the number from the appropriate column and row and multiple by the exposure time.

	AM 1.5	Full Spectrum 50X	QUV-340 at 1.55 W/m2	QSUN-TUV at 70 W/m2
Full Spectrum Intensity (kW/m2)	1.00000	50.40000	0.08454	0.39071
TUV (280-400nm) Intensity (W/m2)	34.34662	1731.06947	84.54211	70.00000
UVA-340 Intensity (W/m2)	12.77404	643.81168	60.65379	26.06036
Exposure Time (hours)	1.00000	1.00000	1.00000	1.00000
Full Spectrum Dose (MJ/m2)	3.60000	181.44000	0.30435	1.40655
Equivalent Outdoor Full Spectrum Exposure (months)	0.00465	0.23412	0.00039	0.00181
TUV (280-400nm) Dose (MJ/m2)	0.12365	6.23185	0.30435	0.25200
Equivalent Outdoor TUV Exposure (months)	0.00465	0.23412	0.01143	0.00947
UVA-340 Dose (MJ/m2)	0.04599	2.31772	0.21835	0.09382
Equivalent Outdoor UVA-340 Exposure (months)	0.00465	0.23412	0.02206	0.00948

Project 1-Tidy Data Frame

Tidy clean data into a single data frame with multiple simultaneous observations on a single sample as a single row. You will want to use the tidyverse packages (`“library(tidyverse” ).`

You will want the dataframe structure to be appropriate for net-SEM analysis to be done in Project 4, so review the requirements below and in the user manual of the net-SEM package.

The data frame should contain numeric columns for the stressors, mechanisms, and responses, as well as categorical variables including sample, step, exposure type, and material type. Stressor (Dose) is the first column, second column is system response, and the following columns are mechanistic responses. In this case, irradiance on a particular dose basis is the stressor. The system response is either YI or Haze. The mechanistic variables are the FTIR integrated peak value. You will need to add columns containing information about the sample number and material type.

In the case of the FTIR data, you will need to make single value measures from the spectra from particular wavenumbers. Find the peak area at 1250 cm^{-1}, 1700 cm^{-1} for PET substrates and at 1700, 2900 (doublet), and 3350 cm^{-1}. Please look to see if the x- axis (wavenumbers) in the data were taken at the same intervals.

An example dataframe has been provided for a structure reference, T**HE DATA IN THE EXAMPLE IS INCOMPLETE AND INCORRECT, IT MERELY GIVES AN EXAMPLE OF HOW THE FINAL DATAFRAME SHOULD BE ORGANIZED.**

Please answer the following 3 questions. Make the necessary plots to show your thoughts.

1. What are the dimensions of your data frame?
2. Show the head and tail of your data frame
3. Plot the YI and Haze as a function of Dose for each material for the 1x exposure, do you notice any difference between substrates and the coatings on the substrates?

Project 1 Part 2 net-SEM

net -SEM is a network modeling methodology to develop < Stressor | Mechanism | Response > inferential models to look at relationships and their significance between different variables which include stressors, mechanistic/mode and response such as seen in Figure 1 below

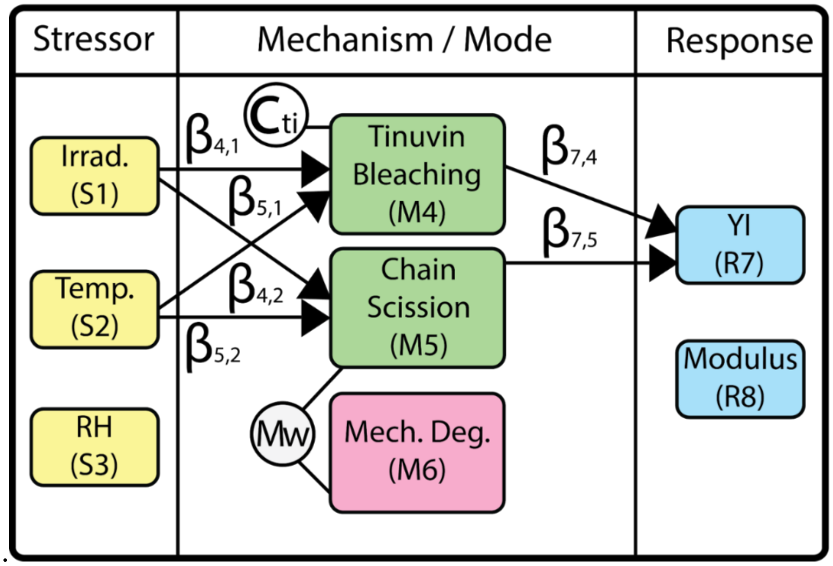


Fig. 1. <S|M|R> network model for the stressor/mechanism-mode/response for acrylic degradation, which shows the contributions of each stressor on a measured response where C_{ti} (concentration of Tinuvin) and Mw (molecular weight) are latent variables shown is circles. β_{ij} is the coupling strength. Yellow boxes are stressors. Green boxes are mechanisms. Pink boxes are modes and blue boxes are responses.[2]

The current 0.4.9 netSEM package should be installed on your ODS-VDI. If not, then you can install the current 0.4.9 production version of the netSEM package from the sgsem bitbucket repo. It is in the feature-0.4.9 branch of the repository. You can use devtools to install this package in your R/Rstudio environment. You can see [how to install R packages from bitbucket in this post](#).

The devtools package gives you the ability to install R packages from GitHub or BitBucket.

A command to install the semi-gSEM version 0.4.2.01 package is

```
devtools::install_bitbucket("cwrusdle/sgsem/packages/netSEM/netSEM@netSEM", auth_user = "sdle", password = "cwru")
```

Develop a data frame to run a semi-gSEM package on. The data frame should be purely numeric. Stressor is the first column, second column is system response, and the following columns are mechanistic responses. In this case, irradiance on a particular dose basis is the stressor. The system response is either YI or Haze. The mechanistic variables are the FTIR integrated peak value.

1. Run a netSEM model for 1x outdoor data. Define your stressors|mechanisms/modes|responses, and show the model and its coupling strengths. What does this tell you?
2. Run a netSEM model for the exposure (HF, ASTMG155, ASTMG154, mASTMG154) you thought was most closely resembled the degradation from the 1x outdoor exposure. What

insights do these give you about the impact of different exposure conditions and stressors on degradation of hardcoat acrylics?