Mining of Process-Structure-Property Linkages Using Data Science Tools

R. Liu, T. Fast, A. Cecen, Y. Yabansu, A. Agrawal, S. R. Kalidindi
AFOSR-FA9550-12-1-0458
Homogenization

Microstructure (RVE) is one of the main inputs to homogenization linkages.
Template for Homogenization

Linking Features and Effective Properties

Dataset Acquisition and Preparation

Microstructure Feature Extraction

Statistical Description of Microstructure
Dataset Acquisition: Generation of Microstructures

Example Case: Inclusion in Steel Matrix

Assume Basic Inclusion Shapes

Volume Fractions between 0% and 20%.

1) Randomly Scattered
2) Vertical Bands
3) Horizontal Bands
4) Clustered
Properties of Interest

- **Effective Yield Strength** of the composite material system.

- **Effective Strain Hardening Exponent** is computed by fitting the averaged true stress-strain response to a power hardening law.

- **Localization Propensity** is defined as the area fraction of the matrix elements experiencing an equivalent strain greater than a prescribed cut-off strain.
Statistical Description of Microstructure

Example Microstructure

Corresponding 2-pt Statistics
Feature Extraction Using PCA

- Random Placement (300)
- Horizontal Bands (200)
- Vertical Bands (200)
- Clustered Placement (200)
Feature Extraction Using PCA
Structure – Property Linkages

- Average Particle Size
  - Power Law

- Volume Fraction
  - Power Law

- Data Based Method
  - Polynomial Fit
Structure – Property Linkages

![Graph showing simulation results for average particle size and volume fraction. The graphs display data points and trend lines illustrating the relationship between predictions and simulation results. The plots use color coding for different methods: green for average particle size power law, blue for volume fraction power law, and red for data-based method polynomial fit.]
Localization

- Calibration of $A_t^L$ is a one time computational cost.
- It serves for any microstructure in the material system.
- Combine best known physics with data science techniques.
Localization

\[ \mathbf{p}_s = \left( \sum_{L} \sum_{t=1}^{S} \frac{\Delta}{N_L} A_t^L M_{s+t}^L + \sum_{L} \sum_{L'} \sum_{t=1}^{S} \sum_{t'=1}^{S} \frac{\Delta^2}{N_L N_{L'}} A_{tt'}^{LL'} M_{s+t}^L M_{s+t+t'}^{L'} + \ldots \right) \langle \mathbf{p} \rangle \]

- **\( A_t^L \): influence function**
- **\( M_s^L \): local microstructure**
- **\( p_s \): local response**

Data driven framework

Materials Knowledge System (MKS)
### Microstructure Descriptor

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
</tr>
</tbody>
</table>

\( M_s^L \): volume fraction of local state \( L \) in spatial cell \( s \)

Phases: \( L = 1, 2 \)
- \( 1 \rightarrow \text{white phase} \)
- \( 2 \rightarrow \text{gray phase} \)

\[
\begin{align*}
M_{6}^{1} &= 1 \\
M_{6}^{2} &= 0 \\
M_{15}^{1} &= 0 \\
M_{15}^{2} &= 1
\end{align*}
\]
Illustration of higher order terms

3\textsuperscript{rd} order terms

Example local state combination

How many times does this combination appear in this microstructure?

\[
\begin{align*}
M_{s}^{h} & = M_{s}^{h_{1}} \\
\begin{array}{ccccccc}
0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 \\
\end{array} & \times & \begin{array}{ccccccc}
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
\end{array} & \times & \begin{array}{ccccccc}
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
\end{array} & \times & \begin{array}{ccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{array}
\end{align*}
\]
Influence functions

\[ p_s = \left( \sum_{L} \sum_{t=1}^{S} \frac{\Delta}{N_L} A_t^L M_{s+t}^L + \sum_{L} \sum_{L'} \sum_{t=1}^{S} \sum_{t'=1}^{S} \frac{\Delta^2}{N_L N_{L'}} A_{tt'}^{LL'} M_{s+t}^L M_{s+t+t'}^L + \cdots \right) \langle p \rangle \]

- \( A_t^L \): influence of local state \( L \) in the spatial cell \( t \) away from the reference cell \( s \)
- \( A_{tt'}^{LL'} \): combined influence of local states of \( L \) and \( L' \) in spatial cells \( t \) and \( t' \) away from the reference cell \( s \)

- \( A_t^L \) is independent of microstructure morphology.
- They can be applied to any microstructure in the material system.
Generalized Data Mining for Localization

$$p_s = \left( \sum_{L} \sum_{t=1}^{S} \frac{\Delta}{N_L} A_t^L M_{s+t}^L + \sum_{L} \sum_{t=1}^{S} \sum_{t'=1}^{S} \frac{\Delta^2}{N_L N_L'} A_{tt'}^L M_{s+t}^L M_{s+t+t'}^L + \ldots \right) \langle p \rangle$$

**Macro-Feature Extraction:**
RVE-level features (e.g., volume fraction, number and size of connected-same-phase clusters)

**Division/Resampling:**
A machine learning-based data division based on macro-features.

**Local-Feature Extraction:**
Voxel-level features, e.g. number of voxels from each phase at a certain neighbor level.
Problem:
What kind of neighboring information should we include in the calibration?

4 different approaches:
• No partition
• Partition according to volume fractions of data sets
• Partition according to macro features
• Partition of the microstructures in PCA and 2 point correlations
Results with No Partition Methods

No partition
Single model for the entire problem

Middle slice of a test microstructure with a volume fraction of 50.22%

FEM

DM Method

Average RVE Error: 17.01%
Results with P1 Partition: Volume fractions

P1 Partition
Ensemble of RVEs are grouped into categories according to their volume fractions.

Middle slice of a test microstructure with a volume fraction of 50.22%

FEM
DM Method
Avg. RVE Error: 8.89%
Results with P2 Partition: Macro Features

P2 Partition
Make partition of RVEs based on data clustering with a set of selected macro-features; Number of clusters (connected components), maximum, minimum, average size of clusters, dispersion (average of cluster center distances)

Middle slice of a test microstructure with a volume fraction of 50.22%

Avg. RVE MASE: 8.43%
Results with P3 Partition: Macro Features

P3 Partition
Make partition of RVEs based on the distribution of statistical descriptions of microstructures in principal component space.

A sample slice shown with VF of 50.22% (one of the hardest cubes)

FEM
Avg. RVE MASE: 8.03%
DM Method
Impact

• First generation of open source, open access, data-driven, templated protocols for mining PSP linkages
  ➢ Allow a systematic consideration of a large number of potential models
  ➢ Allow sharing of experience with different approaches and identification of best practices

• PyMKS Code repository
  ➢ Short Course at MS&T in Columbus, OH
  ➢ Parts of the code base have attracted a great deal of attention from the boarder community (several hundred downloads/month)

• Hierarchical Materials Informatics as a new field of study
Interactions

*Northwestern and Gatech:* 1 paper published, 2 in preparation

*CMU and Gatech:* 1 review paper published

*KIT and Gatech:* 1 paper submitted, 1 in preparation

*MPIE and Gatech:* 1 paper in preparation

*TRDDC and Gatech:* 1 paper published

*Gatech:* 1 paper submitted, 1 book in print, several new active collaborations on MATIN (as a part of the new course being taught in the FLAMEL program at GT)
Future Outlook

• Application of Homogenization and Localization templates to increasingly more complex, real-world, design problems in hierarchical materials

• Continued build out of open source, open access, data-driven, templated protocols for mining PSP linkages

• Promote community adoption of these tools