Downsampling inelastic cyclic stress-strain data with minimal points to preserve curvature

A Hartloper

July 21, 2021

Abstract

A downsampling algorithm for stress-strain data that uses few points but retains curvature within a pre-defined tolerance is developed in this document.

1 Introduction

The raw data (i.e., extensometer displacement, load cell force) in experiments used for material model calibration are typically sampled at 10 Hz in the EPFL test setup. As experiments range from around a few minutes (e.g., LP1) to hours (LP5), this results in thousands to hundreds of thousands of data points for each experiment. However, the sampled data is often denser than required to reasonably describe the stress-strain behavior of materials under the applied strain rates. When it comes to calibrating material models, the time required is proportional to the number of data points. Therefore, time can be saved by reducing the number of data points prior to running the calibration while maintaining the fidelity of the data.

Experience has shown that the fidelity of an experiment can be kept with around 10–100 times less points depending on the load protocol. Calibrating the UVC model with ten load protocols *without* reducing the number of points is, therefore, expected to take around several weeks to a year. A "good" method of reducing the number of data points, or *downsampling* the data, is essential to keep the calibration time reasonable (say, on the order of days to a week per material). The properties that define a good downsampling method in this context are now defined.

A good downsampler for inelastic, cyclic, stress-strain data of structural steels:

1. samples enough points in the initial elastic region to obtain a good estimate of the initial elastic modulus,

2. samples the upper yield point,

3. samples all the maxima and minima in each loading cycle (i.e., the *peaks* of the stress-strain graph),
4. samples enough points to retain the fidelity of regions in the stress-strain graph with high curvature,
5. samples as few points as possible in regions of low curvature,
6. can handle different load protocols,
7. can handle noise in the measurements,
8. can handle different strain rates,
9. has interpretable heuristic parameters,
10. is easy to use.

An algorithm that combines different techniques to satisfy these desirable properties is developed in this document.

2 Proposed algorithm

The proposed algorithm is composed of three parts. The first part is to find the peaks in the stress-strain data. The second part is to downsample the entire stress-strain history based on the curvature of the stress-strain data. The third part is to keep additional points in the initial elastic region. The final downsampled data is the unique set combining all three parts.

Mosts of the heavy lifting is done in the second part. The first part is essential to overcome *aliasing* in the sampled data. The third part is essential to ensure fidelity of the initial elastic region. Details are given for each part and a summary of the entire procedure is given afterwards.

2.1 Identifying peaks in the stress-strain graph

The peaks in the stress-strain graph are identified by the minima or maxima in each cycle. Cycles can be fairly reliably identified by the switching of signs in the stress signal. Therefore, the peaks are the minima and maxima between each change in sign of the stress signal. This method works well if the stress crosses zero in each cycle and there is no substantial noise in the stress signal around zero.

An attempt at identifying upper yield point, if it exists, is made in this first part. First, the 0.2 % offset yield stress is computed using the initial elastic modulus. The initial elastic modulus itself is computed using linear regression up to $0.65 f_{y,n}$, where $f_{y,n}$ is the nominal yield stress of the material. Then, the upper yield point is selected as the point of maximum stress up to the 0.2 % offset yield stress. If there is no upper yield point (due to the

material or imperfections in the specimen), then the maximum will typically be at the 0.2 % offset yield stress, and this point is sampled instead.

Another important point in stress-strain data tested according to the RESSLab protocols is the first instance 2 % strain amplitude is crossed. This point is important because the strain rate changes by a factor of around 30 at this point. Therefore, this point is identified as well. The final point of importance is the final point. Unless otherwise specified, the final point is selected at 12.5 % strain due to limitations in the small extensometer used for M8 experiments. Any points past the final point are neglected.

To summarize, the sampled points from the first part are:

- the upper yield point (or 0.2 % offset yield stress point),
- the point where 2 % strain amplitude is crossed (not applicable for LPs 4 and 5),
- the peaks in the stress-strain graph, and
- the final strain point (typically 12.5 % strain).

2.2 Curvature-based downsampling

2.2.1 The basic method

The idea in the second part of the overall procedure is to remove as many points as possible in the "straight" regions of the stress-strain graph and keep as many as necessary in the "curved" regions. The algorithm used to accomplish this task was based on a blog by Kaushik Ghose[1]. The key to this algorithm is to define what is "straight", and to sample a point every time that this condition is violated.

The method, named the "maximum deviation downsampler", is outlined in Algorithm 1. This method requires a set of points and a tolerance, ϵ. The algorithm starts at the first point and steps through each point in the set. A line is computed between the starting point and the current point. The perpendicular distance is computed between the line and each point in the range of the starting and current point. If the perpendicular distance is greater than the tolerance, the point before the current point is sampled; the sampled point is then set as the starting point.

The keys to making Algorithm 1 in the current context are to: compute an appropriate perpendicular distance, to use an appropriate tolerance, and to remove noise from the data. One challenge in stress-strain data is that the "distances" in stress and strain have different units. This is handled by normalizing the stress data by the range in stress ($\sigma_i^* = \sigma_i/(\max\sigma - \min\sigma)$), and the strain by the range in strain ($\varepsilon_i^* = \varepsilon_i/(\max\varepsilon - \min\varepsilon)$). Therefore, the normalized values have a range of 1.0 in both axes.

[1] `https://kaushikghose.wordpress.com/2017/11/25/adaptively-downsampling-a-curve/`

Given this normalization, a tolerance of $\epsilon = 0.001$ has been found to be appropriate. This tolerance can be interpreted that "straight" is defined as 1/1000'th of the maximum "distance" spanned by the graph. Finally, noise in the data can cause spurious sampling of points. Therefore, a moving average filter (Savitzky-Golay, `scipy.signal.savgol_filter`) is applied to the stress-strain data prior to normalization. The moving average filter requires a window length, w and an interpolation order, p. A window length of 55 is used for the data prior to 2 % strain (low strain rate), and a window length of 5 is used for the data past 2 % strain.

Filtering the stress-strain data removes the noise, but leads to aliasing. Recall that the peaks of the stress-strain signal are already included in part 1 of the overall procedure, therefore, aliasing is not an issue in part 2. For this reason, linear interpolation ($p = 1$) is used because it is most effective at reducing the noise (the trade-off is that it increases aliasing—which we have already dealt with). Therefore, the set of 2-D points $\{x_i\}$ input to Algorithm 1 are the filtered, normalized stress-strain data.

Algorithm 1 Maximum deviation downsampler.

input: Set of points, $\{x_i\}_{i=0}^{N-1}$ with $x_i \in \mathbb{R}^n$ and $N \in \mathbb{Z}$; and tolerance, $\epsilon \in \mathbb{R}$.
output: Set of indices between $[0, N)$ to keep, k_{sample}.
$k_{sample} \leftarrow \emptyset$
$i \leftarrow 0$
for $j \in (0, \ldots, N)$ **do**
    $l \leftarrow x_j - x_i$ ▷ Line between i and j
    for $k \in (i, \ldots, j)$ **do**
        $d_k \leftarrow x_k \perp l$ ▷ Perpendicular line between x_k and l
    end for
    if $\max \|d_k\|_2 > \epsilon$ **then**
        $k_{sample} \leftarrow k_{sample} \cup \{j - 1\}$
        $i \leftarrow j - 1$
    end if
end for

2.2.2 The fast, less accurate method

As shown through the blog in footnote 1, an alternative to using the maximum deviation is to use the mid-point deviation. The mid-point, k_m, is defined as the index nearest to half-way between i and j in Algorithm 1. Only the mid-point perpendicular distance, d_{k_m}, is computed rather than d_k for $k \in (i, \ldots, j)$. Then, if $\|d_{k_m}\|_2 > \epsilon$ the point $j - 1$ is sampled as usual.

Using the mid-point is significantly faster, depending on the total number of points, the curvature, and the tolerance. Likely, it is around 10–100 times

faster. The mid-point method is worse than the maximum method in that it may result in more samples and higher error between the sampled and original data.

2.3 Additional sampling in the initial elastic region

The initial elastic region is assumed to be bounded by the first point in the data and the upper yield point obtained in part 1 of the overall procedure. It is important to have a few points in this region to obtain a good estimate of the initial elastic properties from the downsampled data. Therefore, $n_{elastic}$, additional points are sampled in the initial elastic region at approximately evenly spaced strain intervals. Each additional point is sampled using

$$j = \arg\min |x_i - x_t|, \tag{1}$$

where x_t are the target evenly spaced strain points.

2.4 Saturation in constant amplitude tests

Cyclic hardening saturates logarithmically under constant amplitude cyclic loading in structural steel materials. Therefore, many cycles may be associated with a small increase in stress and a large portion of data is associated with a small amount of information. One option to reduce this data is to cut the cycles after saturation has been reached. The saturation point is defined as the index of the first peak that satisfies $\sigma_i > \eta \max |\sigma|$ with $0 < \eta \leq 1$. The data after the saturation point is not sampled.

Cutting cycles after saturation in constant amplitude tests requires the definition of the saturation tolerance, η. The minimum number of cycles to include, n_{cyc}, is also specified.

2.5 Overall procedure

The overall procedure is summarized in Algorithm 2. The procedure consists of parts 1, 2, and 3 outlined earlier, with the incorporation of cycle cutting in constant amplitude tests. The original stress-strain data input to the procedure is $X = \{x_i\}_{i=0}^{N-1}$. Algorithm 2 produces a set of indices, k_{sample}. The downsampled stress-strain data is $X_{ds} = \{x_i \mid i \in k_{sample}\}$. I emphasize that the downsampled data is sampled directly from the original data, therefore, $X_{ds} \subset X$. I.e., the filtered, normalized data is only used to determine the indices and the added elastic points are the closest indices to the targets.

I recommend using the following input parameters in Algorithm 2:

- f_{yn}: based on the material
- $\eta = 0.99$

- $n_{cyc} = 20$
- $\epsilon = 0.001$
- $w = 5$
- $\alpha = 11$
- $p = 1$
- $n_{elastic} = 7$

The basis for these recommendations will be shown through later examples.

Algorithm 2 Overall summary of the proposed downsampling method.

input data: set of stress-strain data, $\{x_i\}_{i=0}^{N-1}$
input parameters: nominal yield stress, f_{yn}; saturation tolerance in constant amplitude loading, η; minimum cycles in constant amplitude loading, n_{cyc}; max deviation downsampler tolerance, ϵ; filter window length post-2 % strain, w; filter window length pre-2 % factor, α; filter interpolation order, p; number of additional points in the initial elastic region, $n_{elastic}$.
output: k_{sample} the indices of the sampled points.

Compute the initial elastic modulus and initial 0.2 % offset yield stress using the nominal yield stress, f_{yn}
$k_1 \leftarrow$ indices of the peaks in the stress-strain graph
if constant amplitude loading **then**
    Determine n_{sat} cycles to reach $\sigma_i > \eta \cdot \max|\sigma|$
    if $n_{sat} < n_{cyc}$ **then**
        $n_{sat} \leftarrow n_{cyc}$
    end if
    Truncate data past n_{sat} cycles
end if
Filter stress data pre-2 % strain with window length $\alpha \cdot w$ and order p
Filter stress data post-2 % strain with window length w and order p
Normalize the stress-strain data
$k_2 \leftarrow$ indices from the max deviation downsampler with tolerance ϵ
$k_3 \leftarrow$ indices of $n_{elastic}$ evenly spaced points in the initial elastic region
$k_{sample} \leftarrow k_1 \cup k_2 \cup k_3$

3 Results

The results shown in this section were used to establish the recommended parameters.

4 Remarks

4.1 Different distance measures

As outlined, a challenge in the max deviation downsampler is the difference in units between stress and strain. I propose to re-scale the stress-strain data to then use a basic 2-norm when computing the perpendicular distance. I outline two other options that could be used rather than rescaling:

1. use a different inner product: $d_k = \sqrt{x_k \cdot Y \cdot x_k}$, and
2. use an energy criterion.

In the first option, Algorithm 1 remains essential the same except for how d_k is computed. To use a matrix-vector notation:

$$x_k = \begin{bmatrix} \varepsilon_k \\ \sigma_k \end{bmatrix}, \quad Y = \begin{bmatrix} \frac{1}{(\max \varepsilon - \min \varepsilon)^2} & 0 \\ 0 & \frac{1}{(\max \sigma - \min \sigma)^2} \end{bmatrix},$$

where the values of Y would in this case give a similar effect to the scaling that I already use. Other options for Y could be selected. Essentially, we compute a scaled distance rather than scale the data itself. This method would require a selection of tolerance ϵ and sample $j-1$ based on the same criterion.

In the second option, the perpendicular distance criterion is replaced by the energy bounded by the polygon. In this case, instead of d_k's we compute

$$e = \int_{\varepsilon_i}^{\varepsilon_j} (\sigma - l_\sigma)^2 \mathrm{d}\varepsilon,$$

where l_σ represents the stress signal of the line between x_i and x_j. If the stress-strain graph is nearly straight, then $(\sigma - l_\sigma) \approx 0$. The point $j-1$ would be sampled if $e > \epsilon$.

The benefits of the second method are that it does not require any scaling of the data and the error is closely related to the calibration problem. However, the link between the energy, e, and the curvature in the stress-strain graph is less clear, making it more difficult to reason about an appropriate tolerance. Everything considered, a tolerance still needs to be selected regardless of the method. I went with the data scaling option because, in my opinion, the link between the algorithm and the geometric interpretation of curvature is most clear.