Computer Methods in Applied Mechanics and Engineering 436 (2025) 117709

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

Analyzing fragmentation response of heterogeneous ring using the method of characteristics and machine learning techniques

Reza Abedi [a,*], Colin Furey [b], Farhad Pourkamali-Anaraki [b], Giang Huynh [a], Erdem Caliskan [a], Alireza V. Amirkhizi [c]

[a] Department of Mechanical, Aerospace and Biomedical Engineering, University of Tennessee Knoxville, 1506 Middle Drive, Knoxville, 37916, TN, USA
[b] Department of Mathematical and Statistical Sciences, University of Colorado Denver, 1201 Larimer St, Denver, 80204, CO, USA
[c] Department of Mechanical and Industrial Engineering, University of Massachusetts Lowell, 1 University Ave, Lowell, 01854, MA, USA

ARTICLE INFO

Keywords:
Fragmentation
Method of characteristics
Random fracture
Dynamic strength
Neural networks
Feature selection

ABSTRACT

The fragmentation of a 1D ring problem is analyzed using the method of characteristics. The fracture strength is treated as a random field. Quantities of interest (QoIs) are macroscopic dynamic strength and energy dissipation (toughness). As a loading length scale falls below the correlation and cohesive length scales, macroscopic strength transitions to its dynamic limit dictated by the mean strength. In contrast, for quasi-static loading rates, macroscopic strength follows the weakest link model and is equal to the minimum point-wise strength. The resulting fragmentation data was run through an end-to-end Machine Learning (ML) pipeline, enabling the construction of a Neural Network-based surrogate model for the QoIs requiring only 3 of the original 15 features as input (loading rate, minimum of the strength field and correlation length). Not only were many correlations among the features identified, but we revealed that these dependencies were consistent with the underlying physics. A major consideration was the feature selection process: surrogate models are inspired by computational efficiency, so a parsimonious model is essential. Of particular interest was the performance of a feature selection method based on Shapley Additive Explanations (SHAP), a recent development for interpreting the output of ML models. Although utilization of SHAP in the feature selection process is proven to be successful in the past, our results show that its naive applications can yield subpar results, while a combinatorial approach provided the most robust selection compatible with physical expectations.

1. Introduction

Understanding the mechanisms of dynamic fragmentation is a fundamental scientific problem in mechanics and, in recent years, has received significant engineering and applied attention [1]. Fragmentation studies often express input parameters in the nondimensional form, including nondimensional loading rate, fracture properties, and elastic properties. Key quantities of interests (QoIs), *i.e.,* the outputs of the fragmentation problem, are dynamic strength, energy dissipation (toughness), and fragment size distribution. One-dimensional (1D) studies are widely adopted due to their lower computational cost and the availability of several (semi)analytical and numerical models for fragment size distribution [1–6]. These 1D studies typically consider a straight bar subject to a constant strain rate loading, resulting in a spatially uniform, temporally linearly increasing stress field until the first

* Corresponding author.
E-mail address: rabedi@utk.edu (R. Abedi).

https://doi.org/10.1016/j.cma.2024.117709
Received 14 September 2024; Received in revised form 23 November 2024; Accepted 23 December 2024
Available online 7 January 2025

Nomenclature

$\dot{\mathcal{E}}_o, \dot{\epsilon}_o, \dot{\epsilon}_{o,coh}$	Applied strain rate: dimensional and two nondimensional forms based on domain and cohesive length scales
τ_{coh}	cohesive time scale
G_{coh}, g_{coh}	Work of separation and its nondimensional form
$\Delta_{coh}, \delta_{coh}$	Fracture displacement scale and its nondimensional form
r	Range parameter
ϱ	Density
$\Lambda, \lambda, \lambda_{coh}$	Correlation length: dimensional and two nondimensional forms based on domain and cohesive length scales
L_{coh}, l_{coh}	cohesive length scale and its nondimensional form by domain length
$l_{\dot{\epsilon}_o}$	Nondimensional loading length
$\mathcal{E}$	Strain
Cor	Correlation function
Tri	Triangular distribution
N_{frag}	Fragment count
ϕ_d	Nondimensional total dissipated energy per unit length
ρ	Spatial distance
S_{coh}, s_{coh}	Reference strength and its nondimensional form
$l_{0,s}$	Distance for which the correlation function stays positive
D_s	Higuchi fractal dimension of a realization
Σ	Stress
$l_{cross,s}$	Average segment length between all the crossing of strength through its mean value
σ_M	Nondimensional dynamic strength
$s_{mean}, s_{min}, s_{max}, s_{span}$	Mean, minimum, maximum and span of s
$\varsigma_s, \text{cov}_s$	standard deviation and coefficient of variation of s
s	Nondimensionalized strength field
c	Wave speed
E	Elastic modulus
h	Element size
L	Domain length
p	Number of features
T, t	Time and its nondimensional form
U	Displacement
V	Velocity
X, x	Spatial circumferential positions and its nondimensional form
Z	Impedance
CDF	Cumulative Distribution Function
COV	Coefficient of variation

fragmentation incident initiates. Alternatively, a ring problem subject to a radially constant velocity exhibits behavior similar to the bar problem while avoiding some boundary effect issues of the bar problem [6].

Most numerical fragmentation studies assume homogeneous material properties. However, some 1D studies sample the strengths of equally distant interfaces from different *Probability Density Functions* (PDFs) [7–9], treating strength as white noise, implying that spatially neighboring strength values are uncorrelated. A random field assumption for material properties can model such spatial correlations. Examples of such an approach can be found in [10–17].

In the present study, we employ the method of characteristics for fragmentation analysis of a ring problem with heterogeneous fracture strength handled as a random field. Background information on the method of characteristics can be found in [1,5]. Additionally, and in contrast with most numerical fragmentation studies, we use the Ortiz-Pandolfi [18] *Traction Separation Relation* (TSR), *i.e.,* a cohesive model with a different unloading path compared to those used in [1,6,19]. Moreover, instead of the commonly-used penalty approach, we use Riemann solutions [20,21] to prevent material inter-penetration under contact conditions. We refer the readers to [22] for a brief overview of the method of characteristics and specific aspects of our implementation.

The treatment of the fracture strength as a random field facilitates the analysis of the interaction of the length scale of material heterogeneity, represented by the correlation length of the strength field, with other length scales of the problem, including the fracture length scale, the loading length scale, and the domain size. As shown in [23], such interactions at different relative lengths

result in distinctly different wave propagation responses in linear media. Here we extend such analysis to the nonlinear fragmentation problem. While [13] also employed a random field approach, their work fixed the correlation length and pursued a different goal of studying fragment size distribution.

We study the statistics of the aforementioned QoIs for a wide range of nondimensional loading and material parameters. In addition, the random fields are represented and realized using the range and correlation length parameters, with each random field realization containing over ten thousand point-wise values. We extract various scalar values from individual field realizations and discuss the influence of each one on the fragmentation response, for example, in quasi-static and dynamic loading rates. Gaining such physical insight is the first goal of this study. As further discussed below, *Machine Learning* (ML) methods significantly facilitate such dimensional reduction techniques and enable exploring the relations between the input parameters and output QoIs. This is the second goal of the present work.

With recent advances in computing power and data storage capacity, numerical simulations provide valuable data to train ML algorithms for constructing surrogate models that assist in characterizing complex systems in a fraction of time. Oftentimes, physical simulations for real-world problems can take days or even weeks to complete, hindering the ability of researchers and engineers to conduct important analyses and limiting the scope of the study. ML surrogate models are becoming a widely used remedy for this problem. They are capable of emulating the behavior of simulations while providing significant savings in terms of the computational costs of evaluation compared to the numerical simulations.

For example, in [24], the authors successfully construct a surrogate model for parametric fluid dynamics problems. The authors were motivated by the computationally expensive nature of simulating fluid dynamics problems via solving variations of the Navier–Stokes equations. It is demonstrated that instead of solving *Partial Differential Equations* (PDEs) to obtain data on parametric fluid dynamics problems, the input–output relation of these systems could be well approximated with a deep neural network, which allows for easy evaluation. Similarly, in [25], the prospect of developing surrogate models for stochastic PDEs with high-dimensional, spatially-varying coefficients is explored by incorporating the physical equations governing the PDEs into the loss function; *e.g.,* physical constraints, initial conditions, boundary conditions, etc. It is shown that the inclusion of the underlying physics into the loss function allows for the training and development of surrogate models with less upfront simulation data than would otherwise be required.

In the same vein as the works described above, our main concern from an ML perspective was building an end-to-end pipeline for the construction of a surrogate model for the QoIs. The general outline for the pipeline takes the fragmentation response data as input, performs feature selection, tunes model hyperparameters, and outputs a trained NN surrogate model capable of precise prediction of both QoIs. However, we were initially faced with a high dimensional input space containing many highly correlated features, so our focus was quickly drawn to the feature selection stage of our pipeline. As will be discussed, features associated with how spatial and amplitude variation of the strength field are highly correlated among themselves, negatively impacting the predictive capability of our surrogate model. This prompted an investigation into which feature selection methods should be used at this stage of the pipeline.

Particular focus is devoted to a recent approach to feature selection that utilizes SHAP, a development from explainable artificial intelligence (AI). The merits of the SHAP approach are compared with the results of the baseline wrapper method, where it is determined that although the SHAP-informed feature selection method is useful, it has some important drawbacks as well. In the final analysis, we show that the best-performing subsets of features oftentimes include features for which the SHAP-informed approach attributed minimal importance. Finally, the performance of our final model is discussed, along with some potential areas of improvement.

In short, the main contributions of this manuscript are as follows:

1. Modeling the inherent heterogeneity and aleatory uncertainty of material properties by treating the strength as a random field.
2. The analysis of the interplay between different length scales of the model (correlation length, fracture length scale, loading length scale, and domain size) on the variation of mechanical QoIs.
3. Developing an end-to-end pipeline for training, refining, and deploying ML surrogate models that utilize a SHAP-informed and combinatorial feature selection approach.
4. Extraction of novel features by the integration of ML and physics-based model; some features, such as a certain form of normalized correlation length, were not evident to play a key role on QoIs from the physics model alone.

The remainder of this paper is organized as follows. The fragmentation problem is described in §2, including the overview of all input parameters and output QoIs. The interplay between input parameters and QoIs, such as dynamic strength and fracture energy, is discussed from the physics perspective in §3. Construction of an end-to-end ML pipeline with an emphasis on overcoming multicollinearity in the simulation data via a relatively novel feature selection implementation is discussed in §4. Finally, key findings are summarized, and the conclusions are drawn in §5.

2. Fragmentation problem

2.1. Method of characteristics

The solution of hyperbolic PDEs often involves a situation where, on the two sides of an interface we have different solution fields and possibly different material properties. In its most common and simplest form, the Riemann solution corresponds to the solution

of this problem when these initial condition fields and properties are assumed constant on either side. This Riemann problem setup uses the evolution of characteristic values along the characteristic lines to determine the solution in space–time. Some uses of these Riemann solutions or characteristics for elastodynamic and fracture problems are in characteristic *Finite Difference* (FD) [26,27], *Finite Volume* (FV) [28], and *Discontinuous Galerkin* (DG) [29–31] methods, where the two sides corresponds to two neighboring finite difference points, finite volumes, or finite elements, respectively.

The method of characteristics has been used for fragmentation problems, for example [1]. For the numerical solution of the fragmentation problem, we have developed a characteristic code for a broad class of problems and Riemann solutions [32]. Given the similarities of the two methods for 1D fragmentation problems, we do not provide the formulation here and refer the reader to [1] for more details. Still, a brief overview of the method and some notes on the differences between two methods in modeling the TSR unloading and contract enforcement is provided in the supplementary material.

2.2. Scalar input parameters

2.2.1. Nondimensional input parameters

We consider the fragmentation response of a ring with uniform material properties, except for the fracture strength. The input properties include: (1) Elastic modulus E and density ρ; (2) Fracture displacement Δ_{coh} and reference strength S_{coh}; (3) The domain size L; (4) The applied strain rate $\dot{\mathcal{E}}_\circ$ by which the ring expands radially with constant speed. Any quantity such as S_{coh} and Δ_{coh} associated with cohesive fracture is denoted by the subscript coh. $\dot{\mathcal{E}}_\circ$ induces a spatially uniform strain and stress field that start from zero at time $T = 0$ and linearly increase in time until the spatially uniform stress reaches the strength of the weakest point of the domain. The strain $\mathcal{E}$ and stress Σ before this time at any (circumferential) position X are given by,

$$\mathcal{E}(X) = \dot{\mathcal{E}}_\circ T \qquad \Rightarrow \qquad \Sigma(X) = E\dot{\mathcal{E}}_\circ T, \qquad \text{(before any crack opening)} \tag{1}$$

and once cohesive interfaces undergo the debonding process, such spatially uniform strain and stress fields will no longer be valid.

The elastodynamics problem with TSR-based (cohesive) fracture can be nondimensionalized in various ways. This is achieved by choosing scales (reference values) for three physical quantities. For example, the wave speed $c = \sqrt{E/\rho}$ and S_{coh} are chosen as velocity and stress/strength scales, respectively. Length is selected as the third physical scale. Two choices are considered for the length scale $\tilde{L}$. The first option sets $\tilde{L}$ equal to the length of the domain L. The second option sets $\tilde{L}$ equal to the so-called (Irwin's or Griffith's) fracture length scale $L_{\text{coh}} := G_{\text{coh}}E/S_{\text{coh}}^2$, where G_{coh} is the *work of separation* in the context of cohesive fracture. For the linear Ortiz-Pandolfi [18] cohesive model used, $G_{\text{coh}} = S_{\text{coh}}\Delta_{\text{coh}}/2$. The corresponding time scale is then $\tilde{\tau} := \tilde{L}/c$. Accordingly, for the first option the derived time scale $\tilde{\tau} := L/c$ corresponds to wave travel time in the domain. For the second option, $\tilde{\tau} = \tau_{\text{coh}} = L_{\text{coh}}/c$ is the cohesive time scale [33]. For the linear TSR model used $\tau_{\text{coh}} = \Delta_{\text{coh}}Z/2S_{\text{coh}}$, where $Z = \sqrt{E\rho}$ is the impedance of the material. The scales of other quantities, such as energy and displacement, are computed from the scales of velocity, stress, and time.

In the remainder of the manuscript, all the quantities are expressed in nondimensional form. Uppercase and lowercase symbols represent dimensional and nondimensionalized quantities, respectively. For example, $\dot{\epsilon}_\circ$ is the nondimensional form of the strain rate $\dot{\mathcal{E}}_\circ$. A subscript of coh is used for quantities nondimensionalized based on cohesive quantities, *e.g.,* length scale L_{coh}, to distinguish between the two nondimensional forms of quantities. For example, $\dot{\epsilon}_\circ$ and $\dot{\epsilon}_{\circ,\text{coh}}$ are the normalized versions of $\dot{\mathcal{E}}_\circ$ based on two different length scales discussed above (specific expressions are provided below). Expressing quantities such as strain rate in two nondimensional forms serves two purposes. First, designing the set of simulations, such as setting the terminal time, is more easily achieved using $\dot{\epsilon}_\circ$ ($\tilde{L} = L$). Second, while the majority of results in the literature are expressed in the latter nondimensional form ($\tilde{L} = L_{\text{coh}}$), certain nondimensional quantities, such as correlation length, will be shown to be more predictive of the response when nondimensionalized using $\tilde{L} = L$.

After nondimensionalizing the fragmentation problem, the only two relevant input parameters are the nondimensional loading rate (using the domain length scale L) $\dot{\epsilon}_\circ = \dot{\mathcal{E}}_\circ/(S_{\text{coh}}/E/\tau_L)$ and a nondimensional cohesive model displacement scale $\delta_{\text{coh}} = \Delta_{\text{coh}}/(LS_{\text{coh}}/E)$. The other form of nondimensional strain rate $\dot{\epsilon}_{\circ,\text{coh}}$ using the cohesive length scale is $\dot{\epsilon}_{\circ,\text{coh}} := \dot{\mathcal{E}}_\circ/((S_{\text{coh}}/E)/\tau_{\text{coh}}) = \dot{\epsilon}_\circ(\tau_L/\tau_{\text{coh}}) = \dot{\epsilon}_\circ\delta_{\text{coh}}/2$. Since the domain length L is scaled by the length scale L, in the nondimensional form the domain is of size one. We choose the spatial coordinate x in the range $[-0.5, 0.5]$.

2.2.2. Parameters defining the random field

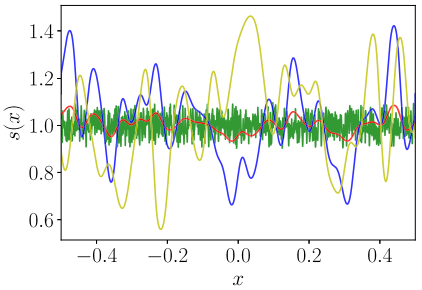
The fracture strength is assumed to be a random field. A realization of a nondimensionalized strength field $s(x)$ is given by,

$$s(x) = S(xL)/S_{\text{coh}}, \qquad \text{for} \quad x \in [-0.5, 0.5], \tag{2}$$

where, as mentioned before, S_{coh} is the scale of fracture strength and $S(xL)$ is the strength of this realization of a ring of length L.

We further discuss the properties and the realization scheme for the nondimensional strength field. We assume s to be ergodic and stationary in a strong sense, with a triangular distribution $\text{Tri}(1-r, 1, r)$, where r is the range parameter, for all $x \in [-0.5, 0.5]$. Clearly, for this triangular distribution, the minimum, maximum, and mean values are $1-r$, $1+r$, and 1.

To generate random field realizations for $s(x)$, we use the *Cumulative Distribution Function* (CDF) inverse method; *cf. e.g.,* [34]. We initially generate a random field with point-wise *standard normal* (SN) distribution. The SN random field has a squared exponential correlation function with correlation length λ, such that the correlation function of two points of spatial distance ρ is given by $\text{Cor}(\rho) = e^{-(\rho/\lambda)^2}$. We use the Karhunen-Loève (KL) [35,36] method to generate realizations of the SN random field using a Julia package implementation of the KL method [37]. For each realization of the SN random field, we apply the CDF of SN and the inverse



	λ	r
1	0.001	0.1
2	0.0316	0.1
3	0.0316	0.9
4	0.0316	0.9

Fig. 1. Four sample random fields for nondimensional strength s illustrate the effects of varying the correlation length λ and range parameter r.

of the CDF of the triangular distribution to map it to the target random field realization with point-wise triangular distribution. More details on the process of generating the random fields can be found in [38].

We note that $\lambda = \Lambda/L$ is the nondimensionalized correlation length with respect to the domain size L. As mentioned earlier, in the nondimensional form of the problem, the domain length is scaled to unity. Similar to the two versions of the nondimensionalization of loading rate (using the domain and cohesive length scales), the correlation length can also be nondimensionalized both ways. The nondimensional correlation length with respect to the cohesive length scale is $\lambda_{\text{coh}} = \Lambda/L_{\text{coh}} = (\Lambda/L)(\tau_L/\tau_{\text{coh}}) = 2\lambda/\delta_{\text{coh}}$. Again the superbar indicates normalization with respective to the cohesive (length) scale. In §4, we will discuss how each of these nondimensional correlation lengths compares in forming predictive machine learning models.

Fig. 1 exhibits how correlation length λ, range parameter r, and realization number affect the realized nondimensional strength field s. A comparison of the first two lines reveals that increasing λ reduces the spatial variation of realized fields. The second and third lines are generated for the same λ and are mapped from the same underlying standard normal field realization (realization #1 generated by the KL method), but are mapped to triangular distributions with different r ranges. While not directly proportional, they exhibit the same maximum and minimum locations. Finally, the last two lines show two sample realizations (#1 and #2) for the same λ and r values.

2.2.3. Summary of all scalar input parameters

From §2.2.1, the fracture length scale δ_{coh} and the loading rate $\dot{\epsilon}_\circ$ (or its alternative expression $\dot{\epsilon}_{\circ,\text{coh}}$) are the only two nondimensional parameters for the fragmentation problem considered. From §2.2.2, the random fields for nondimensional fracture strength s are generated using only two scalar parameters λ and r. We consider the following range of values as scalar inputs:

1. $\log(\delta_{\text{coh}}) \in \{-2, -1.5, -1\}$. 3 values are considered for the base 10 logarithm of δ_{coh}, the cohesive length parameter.
2. $\log(\dot{\epsilon}_\circ) \in \{-3, -2.5, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4\}$. 15 values are considered for the base 10 logarithm of $\dot{\epsilon}_\circ$, the strain rate parameter.
3. $\log(\lambda) \in \{-4, -3.5, -3, -2.5, -2, -1.5, -1, -0.5\}$. In addition, a white noise case is considered wherein the values at computational domain points are independently sampled. Thus, 9 values are considered for the λ of the random fields and the white noise case.
4. $r \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$. 5 values are considered for r, the amplitude (point-wise) strength range parameter.

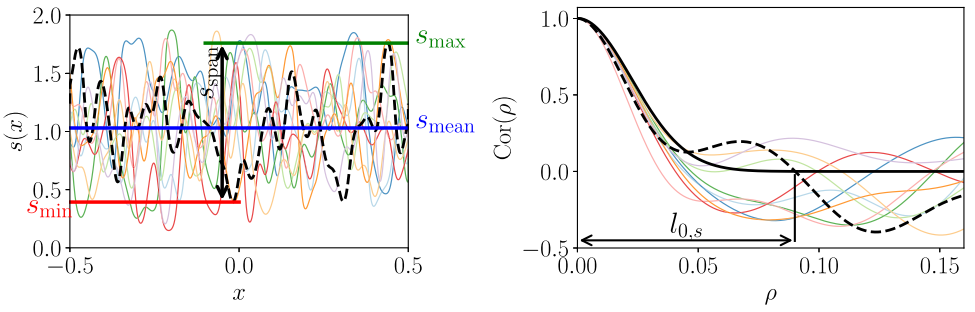
Thus, a combination of $2,025 = 3 \times 15 \times 9 \times 5$ scalar parameters is considered when the strength field is inhomogeneous. We also consider the homogeneous strength field case (corresponding to $r = 0$). For this case, only $\log(\delta_{\text{coh}})$ and $\log(\dot{\epsilon}_\circ)$ vary across the ranges mentioned above. Thus, for the homogeneous strength case, there are 45 scalar input combinations. In short, a combination of 2,070 scalar input parameters is considered.

To accurately capture the statistics of QoIs, fracture response of about 80 random field realizations were considered for each of the 2,025 combinations above. The wall clock time of simulations on the university of Tennessee (UTK) ISAAC clusters range from about 15 min ($\log(\dot{\epsilon}_\circ) = 4$) to around a day for the low loading rates. Long simulation times were needed for the QoIs such as fragment count and especially energy dissipation to stabilize and reach their terminal values. Given the large number of simulations and high computational costs, especially for the lower rates, we limit this study to 1D, paving the way for 2D and possibly 3D fragmentation studies in future; *cf.* §5. The existence of several high fidelity computational results for 1D fragmentation, *cf.* the Zhou-Molinari-Ramesh (ZMR) best-fit for homogeneous domains [1] in §3.1.1, was another motivation for the choice of 1D results.

2.3. Field input parameters

Since the nondimensional fracture strength s is a random field, each pair of $\log(\lambda)$ and r produces an infinite number of random field realizations using the KL method. In this study, a total of 178,915 simulations were carried out on the University of Tennessee Knoxville (UTK) ISAAC clusters. Specifically, for each of the 1,800 cases of fragmentation involving four input parameters for an inhomogeneous medium ($r > 0$), approximately 100 realizations of the random field were simulated.

Fig. 2 compares 10 sample realizations and corresponding correlation functions of the random field realizations for $\log(\lambda) = -1.5$ and $r = 0.9$. Since each realization shown is different, their fragmentation response under the same loading rate (same $\log(\dot{\epsilon}_\circ)$) and



(a) Realized fields (b) Correlation functions

Fig. 2. Comparison of 10 realizations of the random field for s with $\log(\lambda) = -1.5$ and $r = 0.9$ in (a) and their corresponding correlation function in (b). The first realization is highlighted by the black dot-dashed line to show the definition of certain parameters of a realization of s. In (b), the correlation function of the KL method ($\text{Cor}(\rho) = e^{-(\rho/\lambda)^2}$) is shown in a black solid line. The definitions associated with this function, including $l_{0,s}$, are provided in §Section 2.3.2.

for the same δ_{coh} will differ. The computational domain consists of $2^{14} = 16{,}384$ segments and fracture interfaces, which is enough to resolve the cases with the smallest correlation length scale. The random field realizations (*e.g.,* in Fig. 2(a)) are expressed in terms of 16,384 interface values. The fracture response is fully determined by these 16,384 field strength values, the loading rate (from $\log(\dot{\epsilon}_o)$), and the fracture length scale (from $\log(\delta_{\text{coh}})$). However, the large dimensionality of the field input clearly is not ideal for a useful model. Clearly, $\log(\lambda)$ and r indirectly influence the form of a random field realization; the correlation length affects the spatial variation of the field (*i.e.,* how rapidly the values go over peaks and valleys in x) and r affect the range of sampled values (shown in the y axis in Fig. 2(a)). In an attempt to reduce the dimension of the field values, in addition to $\log(\lambda)$ and r, we need to extract certain meaningful parameters from an actual random field realization and its corresponding correlation function that distinguishes one realization from the others for the same $\log(\lambda)$ and r.

The variations of values in space and amplitude can be characterized from different perspectives. Some of these parameters are shown in Fig. 2. These extracted values are discussed next.

2.3.1. Amplitude variation of the field

The parameters characterizing the amplitude variation are as follows:

1. s_{mean}: The mean strength, as shown in Fig. 2(a).
2. $s_{\min}$: The minimum strength, as shown in Fig. 2(a). Given the role of r in the triangular distribution for s, $s_{\min} > 1 - r$.
3. $s_{\max}$: The maximum strength, as shown in Fig. 2(a). Again, from the distribution of s we have $s_{\max} < 1 + r$.
4. s_{span}: The span of strength $s_{\text{span}} = s_{\max} - s_{\min}$, which characterizes the maximum amplitude variation of strength for one realization (simple function of the previous two).
5. ς_s: The standard deviation of strength, which characterizes the overall amplitude variation of strength across all points.
6. cov_s: The *Coefficient of Variation* (COV) of s, which is equal to ς_s divided by s_{mean}. Given that the mean values are close to one for all random field realizations considered, there is not a great difference in the values of ς_s and cov_s. We will, however, investigate which one will be a better predictor of output parameters.

2.3.2. Spatial variation of the field

The parameters characterizing the spatial variation are either directly calculated from the actual values of the field (Fig. 2(a)) or its corresponding correlation function (Fig. 2(b)). The latter is computed from discrete function values of a particular realization. As the spatial size of the domain relative to the correlation length λ used in the KL method increases, the computed correlation function from a particular realization tends to that used to generate the random field, *i.e.,* the black line corresponding to $\text{Cor}(\rho) = e^{-(\rho/\lambda)^2}$ in Fig. 2(b). The parameters representing the spatial variation of the field are,

1. $l_{\text{cross},s}$: The average segment length between all the crossing of s through its mean value. For example, in Fig. 2(a), the number of crossings over the mean value of realization one ($s_{\text{mean}} = 1.0294$) is 20. Hence, the average segment length is $1/21 = 0.0476$. The lower this value, the more rapidly the field varies spatially.[1]
2. $l_{0,s}$: As mentioned before, the correlation function calculated from individual realizations does not match $\text{Cor}(\rho) = e^{-(\rho/\lambda)^2}$ used in the KL method to generate the realizations. The distance for which the correlation function stays positive is denoted by $l_{0,s}$. For example, $l_{0,s} \approx 0.0901$ for realization 1 in Fig. 2(b), which is illustrated with the green line.
3. $\|\text{Cor}_s\|_2$: The l^2 norm of the correlation function.

[1] The number of crossing of a random field realization over its mean value (or other values) is itself a random variable. Further discussion on the statistics of this variable can be found in [39–41]. In a forthcoming publication, we plan to relate this parameter to the quasi-static strength of the ring.

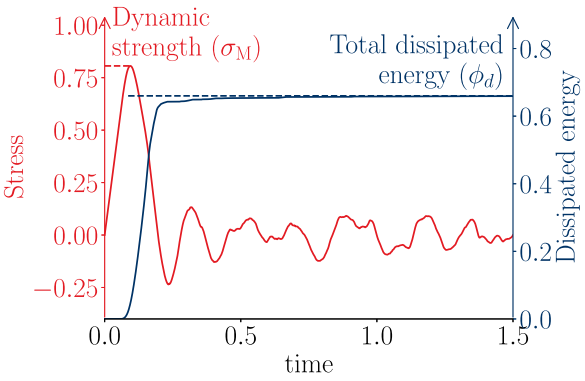


Fig. 3. Examples of output parameters from a fragmentation simulation: Maximum dynamic stress (σ_M) from averaged strain versus average stress response (red); Terminal value of total dissipated energy, ϕ_d (blue). The scalar input parameters for the shown results are $\dot{\epsilon}_\circ = 10$, $\delta_{\text{coh}} = 0.1$, $\lambda = 0.01$, and $r = 0.3$.

4. $\|\text{Cor}_s\|_{2,0}$: The l^2 norm of the correlation function from $r = 0$ to $r = l_{0,s}$.
5. D_s: Higuchi fractal dimension of a realization, computed using the Python code found in [42]. The fractal dimension represents the roughness of a field realization.

2.4. Output parameters

The simulations, numbering approximately 160,000, involve subjecting the ring to a spatially uniform and constant strain rate until all output parameters can be calculated accurately. Some of the key outputs of the fragmentation analysis are the maximum dynamic average stress σ_M and total dissipated energy per unit length ϕ_d. The average strain for each stage of the loading is simply the product of time and strain rate. The average stress is calculated as the volumetric average of point-wise stress from $x = -0.5$ to $x = 0.5$. As depicted in Fig. 3, the average stress linearly increases versus average strain until the cohesive surface in the weakest point in the domain (corresponding to $s_{\min}$) starts opening. The opening of the weakest and a few other weaker points of the domain represents the softening stress response to σ_M, the observed maximum dynamic stress. After this point, more interfaces undergo debonding until the average stress becomes zero and subsequently oscillates between positive and negative values.

The energy dissipation accumulates from the instant that the weakest material point starts debonding. In the example shown in Fig. 3, the energy dissipation is zero in the linear part of the stress response, then rapidly increases around the peak stress time. This is often followed by a much slower increase of energy dissipation until, as shown in the figure, the dissipation reaches its terminal value denoted by ϕ_d.

We note a few aspects pertaining to the calculation of ϕ_d. First, this energy refers to the part of energy dissipation that is not recoverable even upon closing of crack surfaces [1]. Second, since the problem is solved in the nondimensional form, the energy dissipation is per unit length of the ring. Third, due to the very slow convergence of energy dissipation, for example, in nondimensional times over 1000 units for the loading rate $\dot{\epsilon}_\circ = 10^{-3}$, the simulation cost of low loading rates was prohibitively high. Accordingly, we gradually decreased the mesh resolution from its default value of 2^{14} to 2^{10} for loading rates corresponding to $\log(\dot{\epsilon}_\circ) = 0$ to $\log(\dot{\epsilon}_\circ) = -3$. In doing so, for a vertex V in the coarser mesh, the strength value is taken as the minimum of strengths of vertices of the finer mesh ($h = 2^{-14}$) that are half edge length to the left and right of V. Our numerical analysis (not shown here for brevity) shows that for these low loading rates, ϕ_d and σ_M were not affected by this minimum strength coarsening operation and using coarser computational domains.

3. Fragmentation results

In this section, we study the physical behavior of σ_M and ϕ_d in the performed simulations. We first discuss the length scales of the problem, whose interactions result in distinct behaviors as similarly observed in other 1D random media [23]. Their influence on the fragment size and interaction with the element size is then discussed in §3.1.

The problem considered herein has four relevant input length scales: (1) The geometric length scale L, which corresponds to the length (perimeter) of the ring. (2) The fracture length scale L_{coh} corresponding to the fracture process zone size scale in fracture mechanics, which is related to the dimensional fracture displacement scale Δ_{coh}, *cf.* §2.2.1. (3) The length scale of material heterogeneity, which is represented by L_{cor}, the correlation length of the random strength field. (4) The loading length scale, which is defined using the loading time scale represented by the time at which the spatially uniform stress solution from (1) reaches the average strength of the material. The loading length scale corresponds to elastic wave travel length for such a time scale. As discussed in §2.2.1, for the majority of the paper, the domain size is chosen as a reference length scale. Accordingly, all the first three length scales can be nondimensionalized relative to L. Considering this and other physical dimensions of the problem, the relevant nondimensional length scales are: (1) Nondimensional fracture length $l_{\text{coh}} = \delta_{\text{coh}}$; (2) Nondimensional correlation length λ; (3) Nondimensional loading length $l_{\dot{\epsilon}_\circ} = 1/\dot{\epsilon}_\circ$.

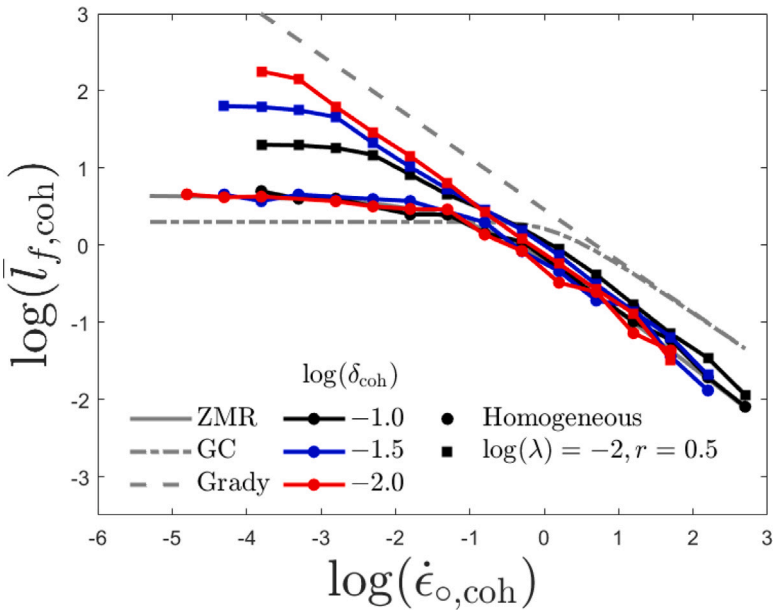


Fig. 4. Fragment size $\bar{l}_{f,\mathrm{coh}}$ versus loading rate $\dot{\epsilon}_{\circ,\mathrm{coh}}$, both nondimensionalized using the cohesive scales. Numerical results are presented for the homogeneous domain ($r = 0$) and ensemble average of inhomogeneous domains for $\log(\lambda) = -2$ and $r = 0.5$. Grady, GC, and ZMR correspond to the homogeneous domain fragmentation models [2,4], and [1], respectively.

3.1. Fragment and element size

3.1.1. Fragment size and validation

During the fragmentation process, the domain is broken to fragments of different sizes. Of particular interest is the dependence of the average of the fragment size, denoted by $\bar{L}_f$, on the loading rate. Analytical and computational models demonstrate that nondimensionalization by cohesive parameters are appropriate for homogeneous domains. The *cohesive nondimensional fragment size* $\bar{l}_{f,\mathrm{coh}} := \bar{L}_f / L_{\mathrm{coh}}$ is related to the cohesive loading rate $\dot{\epsilon}_{\circ,\mathrm{coh}} = \dot{\mathcal{E}}_\circ / \dot{\epsilon}_{\mathrm{coh}}$ [1,2,4],

$$\bar{l}_f(\dot{\epsilon}_{\circ,\mathrm{coh}}) = \frac{\sqrt[3]{24}}{\dot{\epsilon}^{\frac{2}{3}}_{\circ,\mathrm{coh}}} \qquad \text{Grady [2],} \tag{3a}$$

$$\bar{l}_f(\dot{\epsilon}_{\circ,\mathrm{coh}}) = \frac{4}{\dot{\epsilon}_{\circ,\mathrm{coh}}} \sinh\left[\frac{1}{3}\sinh^{-1}\left(\frac{3}{2}\dot{\epsilon}_{\circ,\mathrm{coh}}\right)\right] \qquad \text{Glenn–Chudnovsky (GC) [4],} \tag{3b}$$

$$\bar{l}_f(\dot{\epsilon}_{\circ,\mathrm{coh}}) = 4.5\left[1 + 0.77\dot{\epsilon}^{\frac{1}{4}}_{\circ,\mathrm{coh}} + 5.4\dot{\epsilon}^{\frac{3}{4}}_{\circ,\mathrm{coh}}\right]^{-1} \qquad \text{best-fit from Zhou-Molinari-Ramesh (ZMR) [1],} \tag{3c}$$

where $\dot{\epsilon}_{\mathrm{coh}} = S_{\mathrm{coh}}/E\tau_{\mathrm{coh}} = 2S^2_{\mathrm{coh}}/EZ\Delta_{\mathrm{coh}}$ is the *cohesive strain rate scale*; *cf.* §2.2.1.

Fig. 4 compares our numerical average fragment size with (3). At the quasi-static loading rate, the limiting normalized fragment size is $\bar{l}_f = 4.5$ and 2, for the ZMR and CG models, respectively, and the Grady's model is only appropriate for high loading rates. As the strain rate $\dot{\epsilon}_{\circ,\mathrm{coh}}$ exceeds one, the fragment size precipitously decreases, with the ZMR model predicting smaller fragment sizes than the Grady and GC models. The ZMR best fit is based on TSR numerical solutions and considers some aspects that could not be incorporated in the analytical Grady and CG models. Hence the ZMR solution is considered as the most accurate (semi-)analytical solution provided. As observed, our numerical results for the homogeneous domain ($r = 0$) closely match the ZMR solution for all loading rates.

The fragment sizes of a nonuniform strength field example is also shown in the figure. For each loading rate and δ_{coh}, the ensemble average of $\bar{l}_{f,\mathrm{coh}}$ of all realizations with $\log(\lambda) = -2$ and $r = 0.5$ is taken. The inhomogeneous domain fragment sizes for loading rates $\dot{\epsilon}_{\circ,\mathrm{coh}} \gtrsim 1$ are close to homogeneous domain fragment sizes and are generally slightly larger, *e.g.,* about a factor of two for those corresponding to $\log(\lambda) = -2$ and $r = 0.5$ in the figure. The main difference between the homogeneous and inhomogeneous domains, however, is for the low loading rates where the inhomogeneous domains yield larger fragment sizes. In fact, in all cases the inhomogeneous domains break in exactly one point (the "weakest link" [43]) and the corresponding fragment size is $\bar{L}_f = L = 1 \Rightarrow \bar{l}_{f,\mathrm{coh}} = 2/\Delta_{\mathrm{coh}}$. Thus, at low loading rates, nondimensionalization with respect to the domain length related scales is more appropriate, because regardless of the value of Δ_{coh}, $\bar{l}_f := \bar{L}_f / L = 1$. We will mainly use the latter nondimensionalization, as it also better represents other QoIs such as σ_M.

3.1.2. Element size

As mentioned in §2.3, the computational domain is discretized to $N^h = 2^{14} = 16{,}384$ segments (elements). To ensure accuracy of the solution, the element size $h := 1/N^h$ should be smaller than all relevant length scales of the fragmentation problem. Clearly, h is substantially smaller than the domain and cohesive lengths L and L_{coh}. We will elaborate on the relation between h and inhomogeneity and loading rate length scales.

Referring to §2.3.2, each realization of the random field has $1/l_{\text{cross},s} + 1$ crossings over the mean strength. The average segment length $l_{\text{cross},s}$ is itself a random variable and represents the resolution at which the random field oscillates around the mean strength of unity for s in (2). Using the statistical models in [39–41] and the form of covariance function $\text{Cor}(\rho) = e^{-(\rho/\lambda)^2}$, one can show that the expected value of $l_{\text{cross},s}$ is $\bar{l}_{\text{cross},s} = \mathcal{E}(l_{\text{cross},s}) = \pi\lambda/\sqrt{2}$ for random fields generated for a given λ. The element size should be chosen small enough so the average "wave length" of the realized random field ($2\bar{l}_{\text{cross},s}$) is spanned by several elements. Even for the most demanding case of $\lambda = 10^{-4}$, the ratio of $2\bar{l}_{\text{cross},s}/h = 7.28$ shows that the heterogeneities are resolved well.

The loading rate is the other parameter that should be considered in determining the element size. First, the time discretization is set such that the instant of first crack opening (related to $l_{\dot{\epsilon}_\circ}$) is captured accurately. Moreover, referring to (3), it is evident that the fragment size of a (homogeneous) domain decreases as the loading rate increases. For the highest loading rate $\dot{\epsilon}_\circ = 10^4$ and according to the ZMR model (3c), the ratio of fragment size to element size $\bar{L}_f/h$ ranges from 6.4 ($\delta_{\text{coh}} = 0.1$) to 3.5 ($\delta_{\text{coh}} = 0.01$). For the highest loading rates, our fragment sizes are within 3% and are often over two folds larger than the ZMR estimate for homogeneous and inhomogeneous domains, respectively; *cf.* Fig. 4. Accordingly, even for the highest loading rates, each fragment is spanned by about 3.5 to over 13 elements. This high discretization resolution of fragment segments results in a good match between our solutions and the ZMR estimate in Fig. 4 and ensures that the forthcoming QoIs of σ_M and ϕ_d are computed accurately.

3.2. Dynamic fracture strength σ_M

The interactions of the three nondimensional length scales, related to the value of σ_M, are discussed referring to the example in Fig. 5. Each subfigure shows the variation of σ_M versus $\log(\dot{\epsilon}_\circ)$ for the three values of δ_{coh}. For each δ_{coh}, there are three curves depicting the mean and mean plus/minus standard deviation of σ_M. Displayed values represent the mean and variation across all realizations for $r = 0.9$, with the correlation length λ specific to each plot, δ_{coh} indicated in the legend, and $\log(\dot{\epsilon}_\circ)$ shown on the x-axis. From Fig. 5(a) to Fig. 5(d) the correlation length decreases from $\lambda = 0.1$ to 0.0001.

Rate effect and rate regimes. For all except Fig. 5(d) we observe three distinct regimes of σ_M versus $\dot{\epsilon}_\circ$: (1) Quasi-static regime ($\log(\dot{\epsilon}_\circ) \lesssim -1$ to 0). This limit is qualitatively chosen as the strain rate for which the strength of each realization is close to its quasi-static (minimum point-wise) strength; (2) Transitional regime corresponding to a rapid increase of σ_M. This generally corresponds to loading rates higher than quasi-static rates, but rates for which $\lambda \lesssim l_{\dot{\epsilon}_\circ}$ or $l_{\text{coh}} \lesssim l_{\dot{\epsilon}_\circ}$; (3) Dynamic regime corresponding to rates for which $l_{\dot{\epsilon}_\circ} \lesssim \min(\lambda, l_{\text{coh}})$. The loading rate limits between these loading rate regions are shown by dot-dashed and dashed vertical lines. In the quasi-static regime, the mean and variations (plus/minus standard deviation) of σ_M are rather insensitive to the loading rate and are mainly driven by the specific point-wise distribution of strength across the ring. The influence of range parameter r and correlation length λ on σ_M will be further discussed below. In the transitional regime, strength is highly dependent on the loading rate and transitions from its quasi-static value to the dynamic range values. Finally, in the dynamic range, the mean strength is mostly around unity, meaning that the expected value of strength for the random fields generated.[2] The transition of behavior can be explained as follows. In the quasi-static regime, the condition $\log(\dot{\epsilon}_\circ) \lesssim -1$ to 0, means that during the time that stress ramps from its zero initial condition value to the (mean of) material strength, elastodynamic waves can traverse the domain from 1 ($l_{\dot{\epsilon}_\circ} = 1$) to 10 ($l_{\dot{\epsilon}_\circ} = 0.1$) times. The strength value in this regime tends to be the weakest link strength [43], *i.e.,* the $s_{\min}$ of each realization. As the loading rate tends to zero, the weakest point of the domain fails, and only one fragment forms. But as the loading rate increases, more weak points (local minima) with higher point-wise strengths will undergo (partial) damage, resulting in the rapid increase of σ_M versus $\log(\dot{\epsilon}_\circ)$ in the transitional region.[3]

On the other extreme, reaching the plateau in the dynamic regime is a consequence of $l_{\dot{\epsilon}_\circ}$ being smaller than two length scales. First, the relation between $l_{\dot{\epsilon}_\circ}$ and λ determines the strengths of the fragmented points. As the loading rate increases in the transitional region, the fragment size decreases. Eventually, not only do almost all local minima points of the strength field get fragmented, but also, points with intermediate and high strength values start to experience complete or partial debonding. It is beyond this stage that the average of the strength of these fragmented points, hence indirectly σ_M, tends to the mean of the strength of all points, that is s_{mean}. This transition of σ_M from $s_{\min}$ to s_{mean} for quasi-static to dynamic loading regimes and the relation between $l_{\dot{\epsilon}_\circ}$, λ, and average fragment size will be discussed in more detail below. Interestingly, our fragment size analysis (not presented here for brevity) shows that right around the loading rate for which $l_{\dot{\epsilon}_\circ} \approx \lambda$, the fragment size starts to get smaller than the average spacing between the local minima of the strength field, essentially validating the above discussion about the strength plateau.

Second, the relation between $l_{\dot{\epsilon}_\circ}$ and l_{coh} determines how the average stress within a fragment evolves in time. This interaction is rather complex and involves fragment size as well. It suffices to mention that the analysis of mean fragment size, $l_{\dot{\epsilon}_\circ}$ and l_{coh} from

[2] At very high loading rates, for each individual realization σ_M may become slightly larger than s_{mean} of that realization. This is attributed to higher-than-strength value stresses that can be experienced between any two cohesive interfaces. Given that the mean of s_{mean} over all realizations is one, this explains larger than unity values of the mean of σ_M at very high loading rates in Fig. 5. This is especially evident in Fig. 5a,b.

[3] The quasi-static σ_M has a Weibull distribution. We plan to further investigate the relation between the random field $s(x)$ and the distribution of σ_M for different loading rates in forthcoming publications.

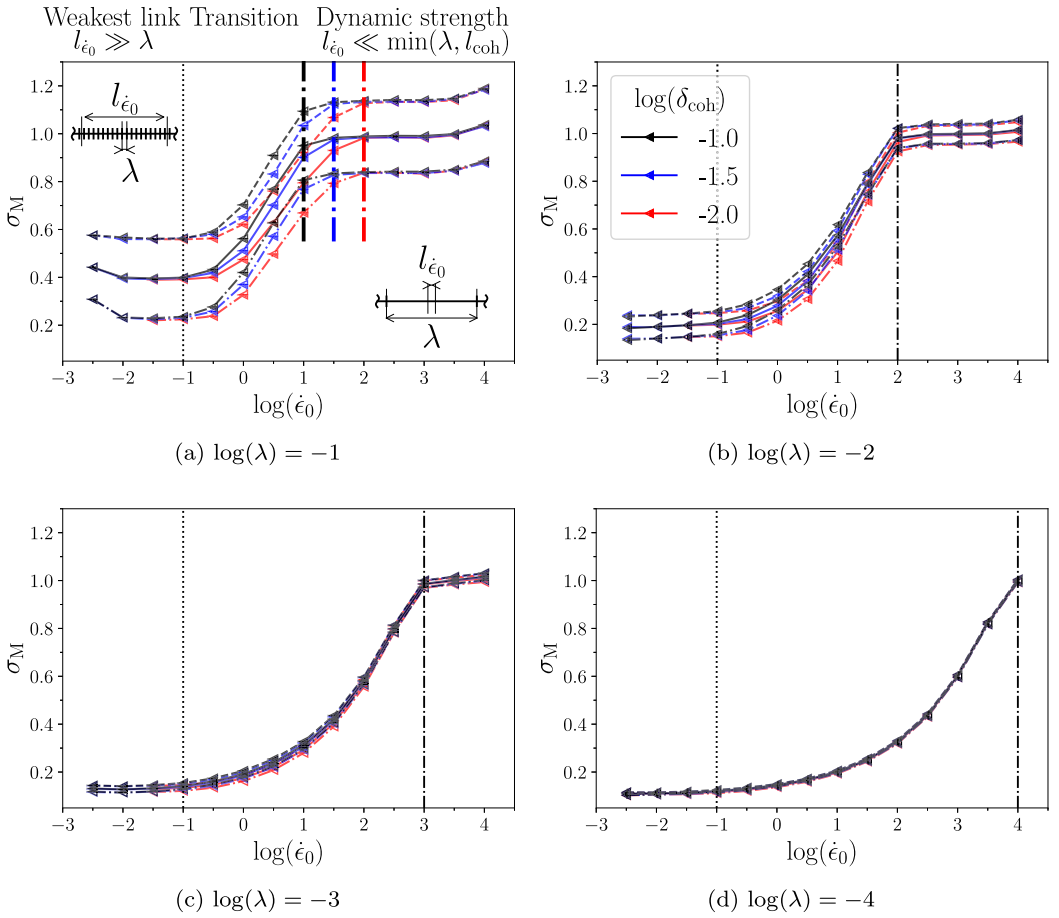


Fig. 5. The variation of σ_M versus $\log(\dot{\epsilon}_\circ)$ for $r = 0.9$, and (a) $\lambda = 0.1$, (b) $\lambda = 0.01$, (c) $\lambda = 0.001$, and (d) $\lambda = 0.0001$. For each δ_{coh}, the middle curve corresponds to the mean value, and the top and bottom curves correspond to the mean ± standard deviation of σ_M. The vertical dot-dashed and dashed lines separate quasi-static, transitional, and dynamic regimes. In (a) the dashed lines are colored based on δ_{coh} as the condition $l_{\text{coh}} = l_{\dot{\epsilon}_\circ}$ precedes $\lambda = l_{\dot{\epsilon}_\circ}$.

our results shows that the stress in the intermediate loading rate can increase if $l_{\dot{\epsilon}_\circ} < \lambda$ but $l_{\dot{\epsilon}_\circ}$ is still larger than l_{coh}. This feature is only relevant for the results in Fig. 5(a) as in the other subfigures $\lambda \leq l_{\text{coh}}$. Consequently, our results suggest that the condition $l_{\dot{\epsilon}_\circ} \approx \min(\lambda, l_{\text{coh}})$ determines the boundary between the transitional and dynamic strength regimes.

Influence of the correlation length. From hereon, we focus only on the influence of λ on σ_M. First, there is a connection between relative values of λ and $l_{\dot{\epsilon}_\circ}$ and homogenization theories. The length scale of material heterogeneity (λ) is much smaller than that implied by the loading in the quasi-static regime in Fig. 5. If the underlying random field were a material property such as elastic modulus, this would have corresponded to long wavelength or quasi-static homogenization that is appropriate for such loading rates [44]. In contrast, in the dynamic range, the loading rate is so high that within each segment, a highly variable stress field is observed. Again, referring to [44], the response of a random material in this regime is mainly driven by scattering events at the interfaces of heterogeneities (*e.g.,* grains), and generally, material response cannot be homogenized. While homogenizing the material response in this case (random fracture strength) is beyond the scope of this work, we note that the quasi-static/long wavelength and wave scattering nature of the problem in these two limits provides additional insight on the distinct responses observed in Fig. 5.

Second, we investigate how λ affects σ_M for all loading rates, especially for the quasi-static regime. Referring to Fig. 5, we can make two observations. First, as λ decreases from Fig. 5(a) to Fig. 5(d), the variation of σ_M around its mean value tends to zero. Second, again, as λ decreases, σ_M decreases in the quasi-static regime (and to some extent in the transitional regime) without much influencing σ_M in the dynamic regime. The second observation is more clearly illustrated in Fig. 6, where only the mean σ_M values are plotted and compared for different values of λ.

These observations can be explained by referring to the interplay between the domain and heterogeneity length scales, reflected in the nondimensional correlation length λ. Referring to Fig. 1, we note that as λ decreases, the number of peaks and valleys

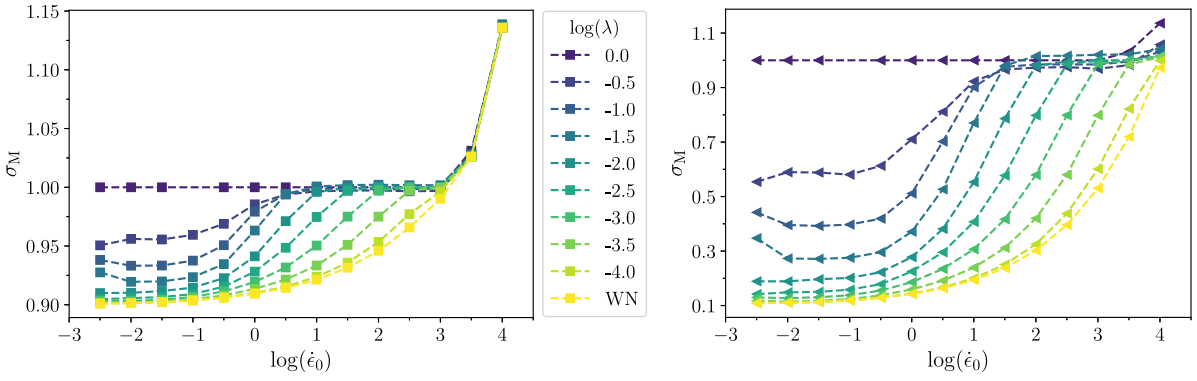


Fig. 6. The effect of loading rate on the mean value of σ_M for $\log(l_{\text{coh}}) = \log(\delta_{\text{coh}}) = -1.5$ σ_M increases for higher loading rates and larger correlation lengths.

of a random field realization increases. In fact, mathematically, one can show that the expected value of the number of peaks and valleys increases inversely proportional to λ [39,40].[4] Moreover, the standard deviation of relevant quantities (*e.g.,* number of peaks and valleys, $s_{\min}$ and $s_{\max}$) all tend to zero as λ goes to zero [39,40]. The fact that the standard deviation of almost any quantity extracted from the random field tends to zero as λ goes to zero implies that all domain realizations that are significantly larger than the heterogeneity length scale are almost identical (law of large numbers) and hence have almost identical physical responses. This explains the first observation above, wherein the variation σ_M decreases as λ decreases.

The second observation on the high sensitivity of σ_M on λ in the quasi-static regime is a direct consequence of the number of valleys (local point-wise minimums) of the strength field. For a static loading condition, the strength of the ring is equal to the minimum point-wise strength, $s_{\min}$, referred to as the weakest link model *cf.* [43]. $s_{\min}$ is the minimum of all local minima. Each local minimum clearly stays in the bound $(1-r, 1+r)$ given the triangular distribution of s. As λ decreases, the likelihood of $s_{\min}$ taking a lower value increases as there are more local point-wise minima. This explains why in Fig. 6, σ_M tends to $1-r$ for the white-noise and $\log(\lambda) = -4$ cases for both r considered in the quasi-static regime. It also explains why larger λ values result in higher mean values for σ_M.

Dependency on point-wise strengths. The dependency of σ_M on point-wise strength values is further clarified by referring to Fig. 7. In this figure, the macroscopic strength (*i.e.,* the ring's dynamic strength σ_M) is plotted versus the point-wise minimum $s_{\min}$ and mean s_{mean} strengths to determine which one affects σ_M to a greater extent. In the top row, the results are presented for a quasi-static loading rate. It is evident that σ_M is identical to $s_{\min}$ (Fig. 7(a)), validating the well-known weakest link model in this loading rate regime. Clearly, there is no correlation between s_{mean} and σ_M in this case (Fig. 7(a)).

In contrast, for a loading rate in the dynamic range (bottom row), σ_M is mainly driven by s_{mean}. In fact, $\sigma_M \approx s_{\text{mean}}$ in this case. This behavior was noted for material with white noise strength values [7], and these results extend such a claim from the white noise to the random field assumption for point-wise fracture strength. This conclusion of convergence of σ_M to s_{mean} is also consistent with the analysis of stochastic discrete element simulations provided in [45] for fracture analysis of 1D and 2D random media. It is evident that in the dynamic regime, unlike the quasi-static regime, $s_{\min}$ does not correlate with σ_M (Fig. 7(c)).

3.3. Energy dissipation ϕ_d

For brevity, we only discuss the rate effects on ϕ_d and defer the more complex dependency of statistical distribution of ϕ_d on other input parameters, especially λ and r, to future publications. Fig. 8 shows the variation of energy dissipation versus loading rate for different λ values. Similar to Fig. 5, for each δ_{coh}, the middle line shows the mean of values whereas the top and bottom lines show mean plus/minus standard deviation. Energy dissipation has two limiting behaviors; in the quasi-static regime, energy dissipation is rather insensitive to the loading rate. For the high loading rates, mainly in the dynamic and parts of transition regimes identified in §3.2, we observe straight lines with slopes close to 2/3. This means that as $\dot{\epsilon}_\circ$ increases, ϕ_d scales close to $\dot{\epsilon}_\circ^{2/3}$. The Grady's model [2,3] predicts that $\bar{l}_f$ to scale as $\dot{\epsilon}_\circ^{-2/3}$; *cf.* (3a). A simple analysis incorporating the energy dissipation at fully debonded interfaces show that ϕ_d scales as $\dot{\epsilon}_\circ^{2/3}$ at high loading rates, validating the 2/3 slope mentioned above.

[4] This number is $\sqrt{2}/\pi\lambda$ per unit nondimensional length for the random field correlation function considered; *cf.* §3.1.2.

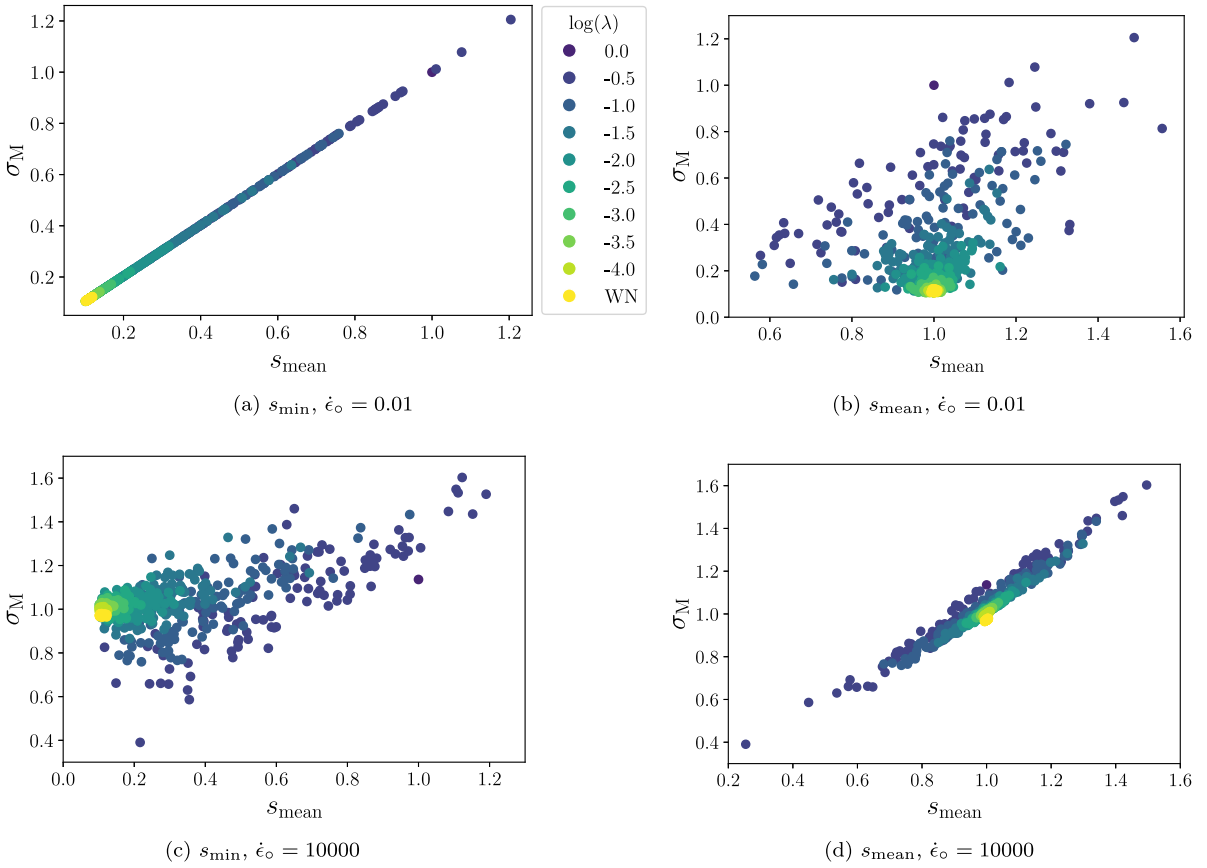


Fig. 7. Scatter plots for σ_M versus minimum point-wise strength $s_{\min}$ (a and c) and mean strength s_{mean} (b and d) for $\log(\delta_{\text{coh}}) = -1.5$, $r = 0.9$ and all correlation lengths. The top (a and b) and bottom (c,d) row shows the results for $\dot{\epsilon}_\circ = 0.01$ and $\dot{\epsilon}_\circ = 10000$, respectively.

4. Machine learning-enabled analysis

In this section, we focus on training a Neural Network (NN) model to accurately predict the output quantities of interest mentioned above, total dissipated energy, ϕ_d, and maximum stress, σ_M, simultaneously. Since both of these outputs are continuous quantities, the problem is well-suited for regression. However, due to the relative abundance of data provided by the fragmentation simulations and the appearance of nonlinear relationships in the data, it was determined that higher accuracy was obtained when using a neural network regression. We discuss how we constructed our NN regression model in the results section below.

The main bottleneck from a modeling perspective was the high dimensional input space. Originally containing 15 possible features (alternatively, predictor or input variables) such as loading rate $\dot{\epsilon}_{\circ,\text{coh}}$[5] and correlation length λ, we later determined that only a fraction of these were necessary to accurately predict ϕ_d and σ_M. Interchangeably referred to as the "curse of dimensionality" or Hughes phenomenon [46–49], this issue occurs when the inclusion of irrelevant features degrades, rather than improves, an ML model's predictive capabilities; *e.g.,* including redundant features can result in multicollinearity problems. This phenomenon also complicates the already difficult task of interpreting the black box function obtained following successful training, an additional focus of our analysis.

In this section, we will first provide a short overview of feature selection in the ML setting, followed by a discussion of SHAP (SHapley Additive exPlanations). SHAP is a development in explainable AI (XAI) that, in recent works, has been proposed as a feature selection mechanism itself. Our results demonstrate that although SHAP has been shown to be successful as a tool for feature selection in benchmark datasets [50–52], similar techniques performed on the scientific dataset described in the preceding sections show that SHAP-informed approaches have some previously overlooked limitations. Our results should not be interpreted as a condemnation of SHAP. It is an excellent tool for both interpretation and explanation of many ML models that would otherwise

[5] In §2.2.1 we introduced the two nondimensional strain rates $\dot{\epsilon}_\circ$ and $\dot{\epsilon}_{\circ,\text{coh}}$, for which the reference length scales of domain size L and cohesive length l_{coh} where used in normalization. While each of these values is more appropriate to explain certain phenomena, for ML we use $\dot{\epsilon}_{\circ,\text{coh}}$ as it is more commonly used in fragmentation problems [1,5,6]. Given the relation $\dot{\epsilon}_{\circ,\text{coh}} = \dot{\epsilon}_\circ \delta_{\text{coh}}/2$, the inclusion of $\log(\dot{\epsilon}_{\circ,\text{coh}})$ and $\log(\delta_{\text{coh}})$ can directly model any dependence on $\log(\dot{\epsilon}_\circ)$.

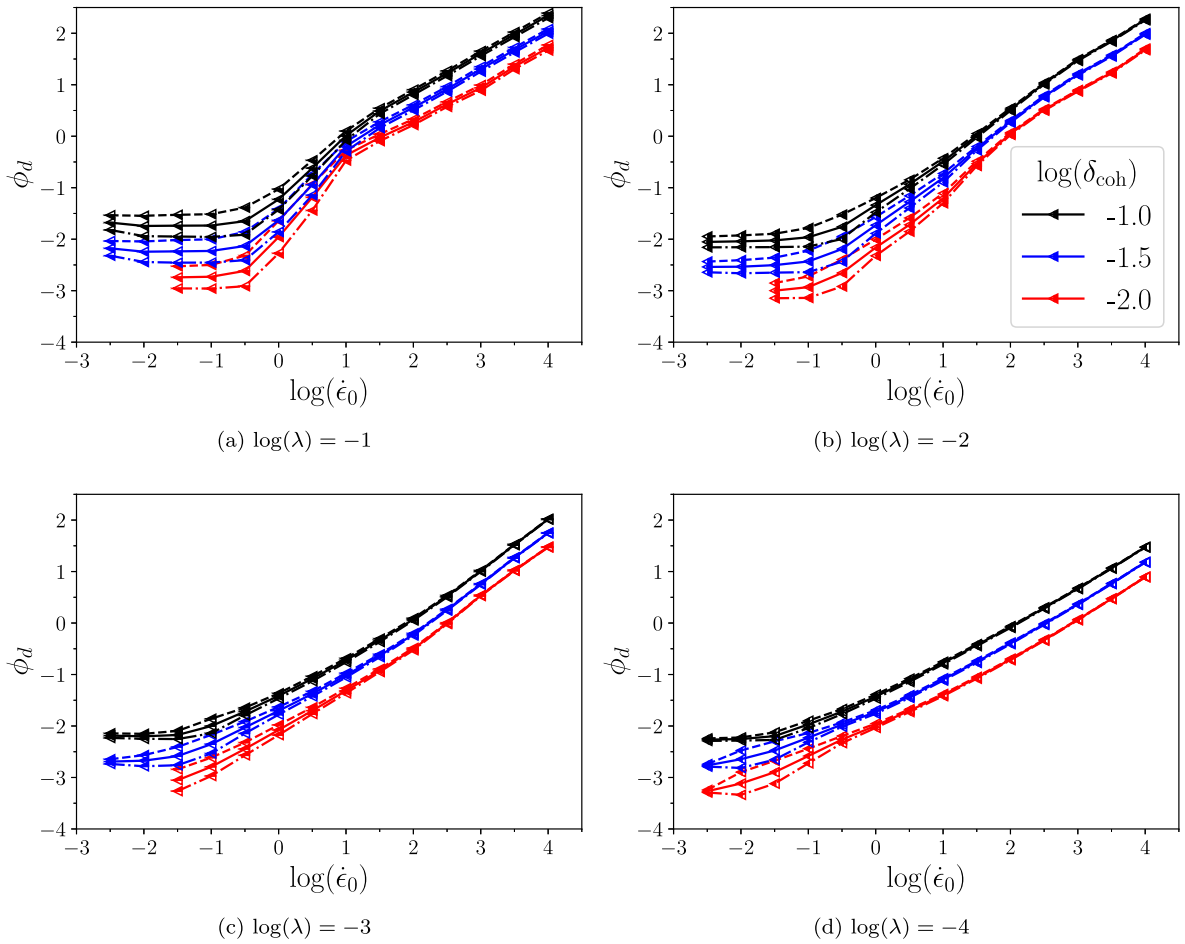


Fig. 8. The variation of ϕ_d versus $\log(\dot{\epsilon}_0)$ for $r = 0.9$ and (a) $\lambda = 0.1$, (b) $\lambda = 0.01$, (c) $\lambda = 0.001$, and (d) $\lambda = 0.0001$. For each δ_{coh}, the middle curve corresponds to the mean value, and the top and bottom curves correspond to the mean ± standard deviation of ϕ_d.

operate largely as black-box functions, capable of precise predictions but lacking any explanation as to why. Instead, our results are a cautionary tale, suggesting that while SHAP is certainly capable of making meaningful contributions to the feature selection process, using it alone to perform feature selection is not advisable.

4.1. Feature selection

Feature selection is a process whereby some of the features initially provided to predict an output are excluded, reducing the dimension of the model's input space. Although there are many motivations for performing feature selection, they generally fall into one of two categories: (1) a smaller input space significantly reduces training time and inhibits overfitting; (2) the features that are retained often provide insight into underlying relationships between the features themselves, between the features and a particular output of interest, or both [53].

Feature selection methods themselves fall into one of three broad categories [54–56]. The first, wrapper methods [57–60], are model-specific and by far the most computationally intensive. If we are initially provided with p features, wrapper methods consider each of the $2^p - 1$ nonempty subsets of features. For each subset, a model is fit to the training data using only those features in the subset and then scored using a corresponding holdout testing set. The subset that obtains the highest score on the holdout data is deemed to be optimal. Due to the dependence on labeled data to score each feature subset, this method only applies to supervised learning. Further, an exhaustive search requires fitting $2^p - 1$ models, which in most settings is simply too computationally expensive to perform. However, if an exhaustive search is feasible, the results can be treated as a baseline with which to compare other, more sophisticated selection methods and may disclose latent relationships between the feature(s) and a particular output.

The second category, filter methods [61–64], refer to the class of feature selection methods that are agnostic with respect to a particular model or output quantity. Filter methods are, in general, less computationally intensive, applicable in both supervised and unsupervised settings, and focus on dependencies among the features themselves (e.g., pairwise correlation). Oftentimes, filter methods yield a feature ranking as opposed to an optimal subset of features.

The last category is the embedded methods [65–67]. These methods are model-dependent and are characterized by performing feature selection and model construction simultaneously. The most common method is the least absolute shrinkage and selection operator or LASSO, which adds a penalty for nonzero regression coefficients to the mean squared error (MSE) loss function of a regression model. The penalty gives the model an incentive to choose nonzero weights for only a small number of features, implicitly selecting those features and excluding the others.

In this work we examine a new approach to feature selection. Inspired by recent applications of SHAP (SHapley Additive exPlanations) to trained ML models in the field of mechanical engineering [68–70], our analysis focuses on a comparison between a SHAP-informed approach to feature selection and the baseline wrapper method described above. SHAP [71] is a recent development in XAI that utilizes the concept of Shapley values [72–75] from cooperative game theory to provide insight into how black-box ML models, such as the NNs considered here, arrive at their predictions.

It is difficult to say which of the three common categories described above a SHAP-informed feature selection method would fall into. A SHAP-informed method cannot be characterized as an embedded method since model construction and feature selection take place at distinct stages. Similarly, any method that uses SHAP cannot be said to be model agnostic either due to the fact that estimating the Shapley values assumes a trained model as the starting place. Hence, a SHAP-informed approach to feature selection cannot be characterized as a filter method either.

A SHAP-informed approach has the most in common with wrapper methods in that it involves attempts to score individual features by utilizing a trained predictive model. However, there are two important distinctions between a SHAP-informed approach and the wrapper methods: (1) SHAP is used to estimate the importance of individual features as opposed to subsets of features; (2) SHAP provides an estimate of an individual features' contribution to predictions made on a holdout testing set whereas the wrapper methods use holdout testing data to estimate the importance of one or more features to the training process. In other words, SHAP-informed feature selection methods bear the most resemblance to wrapper methods, and yet the two approaches focus on fundamentally distinct operations in the ML pipeline: the former provides a measure of the individual features' importance to a given model's predictions, while the latter focuses more on how different subsets of features affect the training of a given model.

SHAP has been utilized in some recent works in mechanical engineering, such as [76]. In particular, the authors there used SHAP to determine feature importance when estimating the shear strength for concrete squat RC walls. There, the authors used SHAP to argue that flange height and wall length were the most influential features when it came to predicting shear strength. Similarly, in [77] the authors were able to utilize SHAP to determine that the carbon content of low-alloy steels was especially influential when it came to estimating the steel's Vickers hardness. Drawing inspiration from these approaches, we set out to obtain similar explanations for our scalar outputs of interest, ϕ_d and σ_M. Before proceeding with our results, we first provide a more in-depth explanation of SHAP and how it can be applied to feature selection.

4.2. SHAP (SHapley Additive exPlanations)

SHAP was introduced in [71] where the authors illustrated the utility of applying a result by Lloyd Shapley from cooperative game theory to trained ML models. Lundberg and Lee's paper was intended to address the growing ubiquity of complex, often nonlinear, ML models employed in various applications in both academia and industry. More specifically, the authors noticed that the ML models capable of the highest accuracy for large modern datasets are often highly complex models that even experts struggle to interpret, creating a tension between *accuracy* and *interpretability*. Researchers recently used Shapley values to evaluate the input feature dependence of machine learning models for modeling process behavior and deposit geometry of friction surfacing [78].

In what follows suppose we have a set of n features $\mathbf{X} = \{X_1, \dots, X_n\}$ under consideration. Each feature X takes on values in $\mathrm{dom}(X)$ and an observation of our n features takes the form $\mathbf{x} = (x_1, \dots, x_n)$ where $\mathbf{x} \in \mathcal{X} = \mathrm{dom}(X_1) \times \cdots \times \mathrm{dom}(X_n)$. Further, $f : \mathcal{X} \to \mathbb{R}$ denotes a trained model that produces predictions $f(\mathbf{x})$. The crux of the SHAP method is the consideration of feature "coalitions" for the observed feature values in $\mathbf{x}$. Suppose we're interested in the effect of the ith feature, X_i, on predictions made using f. The coalitions then correspond to variations of the original input $\mathbf{x}$ where feature i, and possibly some other features, have been removed. Let $\mathbf{x}^*$ denote such a coalition. To determine the effect of feature X_i on a prediction, a comparison is made between the prediction of a model trained on the full data set, $f(\mathbf{x})$, and a model trained on the same data set except for the omission of feature X_i and possibly others, $\hat{f}(\mathbf{x}^*)$. If a feature is unimportant for a particular prediction, then the difference $f(\mathbf{x}) - \hat{f}(\mathbf{x}^*)$ will be small. Otherwise, $f(\mathbf{x}) - \hat{f}(\mathbf{x}^*)$ will be large , evidence that the feature is indeed important to the prediction $f(\mathbf{x})$ made using the full dataset.

To better understand the authors' motivation, consider a generic prediction from a simple linear model with features F. Given an unseen test vector $\mathbf{x}$, a prediction takes the form

$$f(\mathbf{x}) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_n x_n, \tag{4}$$

where $\hat{\beta}_0, \dots, \hat{\beta}_n$ are coefficients estimated from training data. With this class of ML models, the tension between accuracy and interpretation is less of a concern. To understand why, let α_i denote the contribution of feature X_i to the above prediction. We can calculate α_i by considering the difference between the effect of feature X_i in the above prediction, given by $\hat{\beta}_i x_i$, and the average effect of feature X_i in this model, given by $\mathbb{E}\left(\hat{\beta}_i X_i\right)$. Then, the linearity of expectation allows us to write

$$\alpha_i = \hat{\beta}_i x_i - \mathbb{E}\left(\hat{\beta}_i X_i\right) = \hat{\beta}_i x_i - \hat{\beta}_i \mathbb{E}(X_i).$$

If we sum up the n feature effects for this prediction and again leverage the linearity of expectation, we obtain

$$\begin{aligned}\sum_{j=1}^{p}\alpha_j &= \sum_{j=1}^{n}\left(\hat{\beta}_j x_j - \hat{\beta}_j \mathbb{E}(X_j)\right)\\ &= \sum_{j=1}^{n}\hat{\beta}_j x_j - \sum_{j=1}^{n}\hat{\beta}_j \mathbb{E}(X_j)\\ &= \left(\hat{\beta}_0 + \sum_{j=1}^{n}\hat{\beta}_j x_j\right) - \left(\hat{\beta}_0 + \sum_{j=1}^{n}\hat{\beta}_j \mathbb{E}(X_j)\right)\\ &= f(\mathbf{x}) - \mathbb{E}\left(f(\mathbf{X})\right)\end{aligned} \tag{5}$$

and so the sum of all the effects is equal to the difference between the predicted value for the observed feature values in $\mathbf{x}$, $f(\mathbf{x})$, and the average predicted value, $\mathbb{E}\left(f(\mathbf{X})\right)$. Such a concise interpretation of feature effects depends on the fact that the model is linear. However, [71] provides a similar interpretation for more complicated nonlinear models, such as the NN utilized in our study, with SHAP.

The Shapley value, ψ_i, is a measure of the effect of feature X_i on the prediction $f(\mathbf{x})$. As described above, the effect is computed by training one model with the feature present, f, and another, $\hat{f}$, where the feature is not included in the training, and computing the difference in their two respective predictions $f(\mathbf{x}) - \hat{f}(\mathbf{x}^*)$. Since the effect of withholding the feature X_i could depend on the other $n-1$ features, the final Shapley value is computed as a weighted average across all possible subsets $S \subseteq \mathbf{X}$ that exclude the feature X_i:

$$\psi_i = \sum_{S\subseteq\mathbf{X}} \frac{|S|!(n-|S|-1)!}{n!}\left[f(\mathbf{x}) - \hat{f}(\mathbf{x}^*)\right].$$

Note that Shapley values are specific to a given set of observed feature values $\mathbf{x}$, each of which yields a single prediction $f(\mathbf{x})$. In order to understand which features of a given model are most influential in prediction, we must compute Shapley values for many distinct samples. Training separate models for each possible feature coalition and repeating this procedure for many distinct samples quickly becomes intractable, even for small datasets.

Instead, in practice, the theoretical Shapley values are estimated using a weighted linear regression combined with a resampling method that makes use of a single trained model f, the data on which the model is trained, and the set of holdout data for which Shapley values are desired. In a recent work [50] the authors used benchmark data sets to perform feature selection according to the ranking given by the magnitude of Shapley values, averaged over a holdout testing set. In this work, it was demonstrated that this SHAP-informed approach to feature selection achieved better results as compared to three other methods. However, their analysis was restricted to benchmark datasets available in public repositories. Our analysis utilizes a similar approach on the fragmentation response data.

4.3. Results

For ML, we consider all input parameters listed in §2.2.3, §2.3.1, and §2.3.2. As discussed in §2.2.1, for the loading rate and correlation length, there are two forms of nondimensionalization. For the loading rate, we only consider $\dot{\epsilon}_{\circ,\text{coh}}$ as our preliminary analysis showed its superiority to $\dot{\epsilon}_\circ$ and marginal gains by including both. Referring to §2.3.1, we only keep $s_{\min}$ and s_{span} for ML as $s_{\max} = s_{\min} + s_{\text{span}}$. This results in a total of fifteen input parameters. Finally, for the ML we only considered the domains with an inhomogeneous strength field. Moreover, we removed the simulations where the strength is white noise rather than a random field. The rationale is that one cannot assign a reasonable correlation length for either homogeneous or white noise cases. Thus, only 1,800 out of 2,070 combinations of §2.2.3 are considered herein.

As mentioned in the feature selection overview above, our initial motivation for feature selection came from some interesting results obtained during exploratory data analysis (EDA) on the input parameters. As part of our EDA we checked for functional dependencies among the input parameters. To do so, we computed an agglomerative, hierarchical clustering of the fifteen inputs. Agglomerative techniques are iterative clustering techniques that take as input a vector of pairwise distances and outputs a hierarchy of clusters. Beginning with each feature serving as a single cluster, agglomerative techniques search for a pair of clusters to merge at each iteration until only a single cluster, referred to as the "root", remains. The determination of which clusters to merge on a given iteration is dependent upon the choice of distance metric and linkage method, both of which are described below.

Our distance metric was based on the pairwise Spearman rank-order coefficient. Given n observations of two random variables X and Y, let $R(x_1,\ldots,x_n)$ and $R(y_1,\ldots,y_n)$ denote the rankings of the two samples and let $s_{R(x_1,\ldots,x_n)}$ and $s_{R(y_1,\ldots,y_n)}$ the standard deviations of the rank variables. Then the Spearman rank-order coefficient between two samples $x_1,\ldots,x_n$ and $y_1,\ldots,y_n$ is given by

$$\eta = \frac{\operatorname{cov}\left(R(x_1,\ldots,x_n), R(y_1,\ldots,y_n)\right)}{s_{R(x_1,\ldots,x_n)} s_{R(y_1,\ldots,y_n)}}$$

i.e., the Spearman rank-order coefficient between X and Y is simply the Pearson correlation coefficient between the corresponding rank variables of X and Y [79]. The Spearman coefficient η not only measures the strength of the linear relationship between two given variables but also provides a measure of how well the relationship between the two can be described by a monotonic function. This property allows the metric to detect certain types of nonlinear dependencies as well as linear ones.

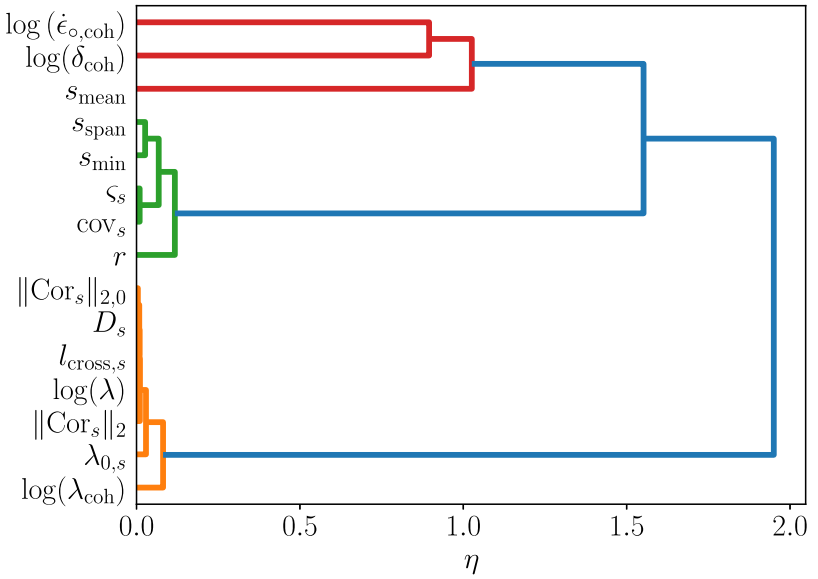


Fig. 9. Dendrogram depicting hierarchical clustering of inputs according to Spearman rank correlation distance metric. Short branch links indicate high correlation. The orange cluster corresponds to inputs that describe how the field changes spatially, and the green cluster corresponds to inputs that describe variations in the field's amplitude. These two clusters contain highly correlated features, indicated by the short branch links. Contrast this with the relatively long branches linking $\log(\dot{\epsilon}_{\circ,\mathrm{coh}})$, $\log(\delta_{\mathrm{coh}})$, and s_{mean} to each other and to the other inputs, an indication of weak correlation amongst themselves and with the other inputs.

Similar to the more standard Pearson correlation coefficient, η takes on values in $[-1, 1]$ and $|\eta| \approx 1$ implies that the two variables under consideration are well described by a monotonic function. This is important from a statistical perspective as two variables whose Spearman correlation is close to one in magnitude should not be assumed to be independent, an important assumption made by many statistical models. To create a distance metric using η, we compute the distance between two inputs as $d = 1 - |\eta|$ so that if $\eta \approx 1$, then the distance between two features is near zero, whereas if $\eta \approx 0$ the distance between two features is close to 1.

As mentioned above, agglomerative hierarchical clustering algorithms treat each feature as a single cluster on the first iteration, and then two clusters are merged at each following iteration. The algorithm terminates when only a single cluster remains. In our case, we utilized Ward's linkage, so the construction of clusters on each iteration was determined by Ward's variance minimization algorithm. This algorithm takes as input the pairwise distances between features generated by the metric defined above and outputs an agglomerative clustering of the features. Ward's algorithm is designed to minimize the total within-cluster variance, making it well-suited to detecting functional dependencies among a given set of features. At each iteration of Ward's algorithm, the distance between each pair of the current clusters is labeled with an objective function, and the next cluster is formed by merging whichever pair of clusters resulted in the minimal value of this function. Note that while the initial distances that serve as input to Ward's algorithm are computed using d, the distances between pairs of clusters computed at each iteration of the algorithm are determined by the algorithm's objective function (for more on Ward's variance minimization algorithm and its corresponding objective function see [80–83]).

Hierarchical clusterings are well described visually by a dendrogram, a diagram with a tree-like structure depicted below. Note that the axis in Fig. 9 extends out to 2, a reflection of the objective function utilized in Ward's algorithm as opposed to the distance metric defined above. Dendrograms also illustrate the order in which the clusters were formed and the distance between the two clusters at the time they were merged. For example, the features $\log(\dot{\epsilon}_{\circ,\mathrm{coh}})$ and $\log(\delta_{\mathrm{coh}})$ were merged into a single cluster when the objective function utilized in Ward's algorithm labeled the distance between the two as roughly 0.8. On the next iteration, the distance between this newly formed cluster and the cluster consisting solely of s_{mean} resulted in the minimal value of the objective function, a value slightly greater than 1, and the two were merged.

Not only did the clustering reveal that many of the input parameters are indeed functionally dependent upon one another, but the clusters also had a nice interpretation from a physical perspective. The inputs in the orange cluster correspond to those associated with the rate at which the field changes spatially, while those in green correspond to how the field varies in amplitude. The last inputs to be clustered are s_{mean}, $\log(\delta_{\mathrm{coh}})$, and $\log(\dot{\epsilon}_{\circ,\mathrm{coh}})$. Referring to §2.2.3, $\log(\delta_{\mathrm{coh}})$ and $\log(\dot{\epsilon}_{\circ,\mathrm{coh}})$ are two of the scalar input parameters that are completely uncorrelated ($d \approx 0.895$). Moreover, s_{mean} is indirectly influenced by the other two input scalar parameters, $\log(\lambda)$, and r, yet having no direct relation to point-wise (green) or amplitude (orange) variation clusters in Fig. 9. This explains the large distance between the three clusters.

The fact that the inputs could be clustered into three distinct and meaningful groups using a straightforward clustering approach prompted us to determine if the SHAP-informed feature selection approach could detect these relationships as well. This is our primary interest in this paper from the ML perspective. The results that follow focus on a comparison between a SHAP-informed approach to feature selection and that of the more robust wrapper method. In the publications [76,77], the authors employed

this approach with success, implementing feature selection informed by the estimated Shapley values of an estimator capable of accurate predictions when trained on all available features. We were interested in whether or not we could find similar success with the fragmentation data considered here, especially in light of the clear functional dependencies among the inputs revealed by the clustering of inputs discussed above and expected from physical considerations.

The starting point for our analysis of the SHAP-informed feature selection approach was to construct a reliable estimator for the two quantities of interest: ϕ_d and σ_M. Recall from above that we were motivated by the nonlinear nature of the data and the relative abundance of simulations to train a NN regression model to predict these two outputs. We obtained an accurate estimator by conducting a grid search over possible NN model hyperparameter configurations, including the number of hidden layers, number of nodes in each hidden layer, activation function, and batch size.

To begin this process, we first randomly partition the fragmentation data (consisting of 140,808 simulations) into three separate datasets. The first and largest contains 68,995 simulations and is used as the training data. The next two subsets have a more nuanced purpose. One, commonly referred to as evaluation or validation data, consists of 29,570 simulations. This is used when utilizing the wrapper feature selection technique. Using the training dataset, a model is first trained on a given subset of the original features and then is scored by computing the model's coefficient of determination (defined below) on the evaluation data. . We also utilize this subset in a similar manner when we look at our model trained on some of the features in isolation, discussed below. The third subset is the holdout testing data, containing 42,243 simulation results. This subset is set aside only to be used at the very end of our analysis to gauge the effectiveness of our final model on unseen observations. A common practice at the final stage of the analysis, after model selection, hyperparameter tuning and feature selection have been completed, is to combine the evaluation data and training data into a larger training data set, which in this case consists of 98,565 observations. This is then used to train the final model, and afterward, we use the holdout testing data to determine the model's efficacy. Similar to the wrapper method described above, the hyperparameter tuning process we employed entails selecting a large number of possible hyperparameter configurations and training a NN using each arrangement and our training data. In order to get a better idea of how the configurations perform on unseen observations, we implemented 5-fold cross-validation when training each configuration. This is a process whereby, for each hyperparameter configuration under consideration, the training data is randomly partitioned into 5 datasets. The configuration is then trained on a union of four of them and subsequently scored using the fifth as a holdout set. This process is repeated the same number of times as the folds, in our case resulting in five model score values. Each configuration's final score is then the average of these five. The hyperparameter configuration that yields the highest average score is then selected for further analysis. In this paper, we are using a NN for regression, so we decided to score each configuration with the coefficient of determination metric

$$R^2 = \frac{\sum_{i=1}^{n} \left(y_i - \hat{f}(\mathbf{x}_i)\right)^2}{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2}$$

where n is the number of observations used and $\bar{y}$ is the average value of the response in the n observations. This metric measures the amount of variation in the response variable explained by the model at hand, relative to simply using the average value of the response variable for prediction. It is a commonly used metric in the regression setting.

We noticed early on in our search that NNs with a single layer performed well, so for the sake of computational efficiency, we began focusing on hyperparameter configurations with a single hidden layer.

The final model selected has a single hidden layer containing 80 nodes, utilizes the ReLU activation function, and has a batch size of 32. When conducting the grid search, we initially planned on training two separate NN estimators for each ϕ_d and σ_M. However, we noticed early on that similar hyperparameter configurations were working well for each. For the sake of computational efficiency, we then decided to train a single NN capable of predicting the two outputs of interest simultaneously, outputting predictions of the form $\hat{\mathbf{y}} = (\hat{\phi_d}, \hat{\sigma_M}) \in \mathbb{R}^2$.

In [76,77] the authors produced a ranking of their features by considering the magnitude of the Shapley values for each feature, averaged over a holdout testing set. Since the Shapley value associated with a feature can be interpreted as the effect of that feature on a given prediction, a Shapley value that is large in magnitude implies that its corresponding feature had an out-sized impact when it came to that particular prediction, independent of the sign. For this reason, a common method of interpreting Shapley values when they are estimated across many observations is to look at their average absolute value. These averages can then be used to rank the features in descending order; *i.e.,* the feature with the largest average Shapley magnitude is ranked first, the feature with the second largest average Shapley value is ranked second, and so on. Such a ranking (which we will refer to as the SHAP-induced ranking) for our data set is depicted in Fig. 10.

Also depicted in Fig. 10 is the R^2 score for our model when it is trained on the features in isolation. By this, we mean that for each of the fifteen features, we take the same hyperparameter configuration discovered via the grid search described above and train a model that takes only a single feature as input on our training data. We then score the model by computing R^2 on the evaluation data. The ordering of the features in both is identical for purposes of comparison. The clear discrepancy between the two plots caught our attention and motivated us to further scrutinize the utility of using Shapley values for feature selection.

In [76,77], the authors selected features according to the ranking induced by the average Shapley magnitudes alone. Mimicking this approach, we considered the top eight features according to the SHAP-induced ranking , illustrated in Fig. 10(a). The remaining seven features were left out so that we could implement the combinatorial wrapper method to compute a baseline for the feature selection process. Without removing some of the features computing such a baseline was not feasible, as the reduction allows us to consider the $2^8 - 1 = 255$ possible subsets of these eight features, as opposed to the $2^{15} - 1 = 32{,}767$ possible subsets of all fifteen features.

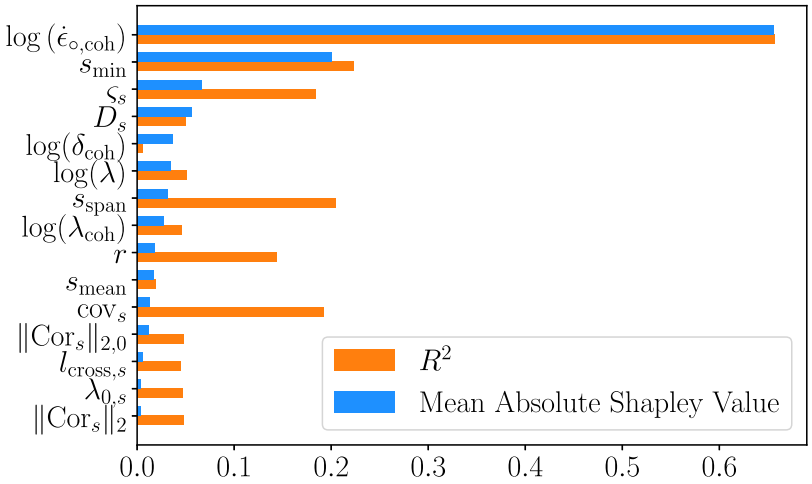


Fig. 10. Comparison of mean absolute Shapley values (blue) and coefficient of determination by feature (orange). For purposes of comparison the features have been ordered according to their mean absolute Shapley values over 42,243 predictions. Notice the discrepancy between a given features contribution to prediction versus.

Table 1
Combinatoric feature selection results. The highlighted row shows the suggested subset of input parameters.

No. Features	Highest Scored Feature Combinations (SHAP Rank)	R^2
1	$\log(\dot{\epsilon}_{o,coh})(1)$	0.6606
2	$\log(\dot{\epsilon}_{o,coh})(1),\ s_{min}(2)$	0.9356
3	$\log(\dot{\epsilon}_{o,coh})(1),\ s_{min}(2),\ \log(\lambda_{coh})(8)$	0.9812
4	$\log(\dot{\epsilon}_{o,coh})(1),\ s_{min}(2),\ s_{span}(7),\ \log(\lambda_{coh})(8)$	0.9879
5	$\log(\dot{\epsilon}_{o,coh})(1), s_{min}(2), \log(\lambda)(6), \log(\lambda_{coh})(8),\ s_{span}(7)$	0.9919
6	$\log(\dot{\epsilon}_{o,coh})(1), s_{min}(2), \varsigma_s(3), \log(\delta_{coh})(5),\ \log(\lambda_{coh})(8), s_{span}(7)$	0.992
7	$\log(\dot{\epsilon}_{o,coh})(1), s_{min}(2), \varsigma_s(3), D_s(4), s_{span}(7),\ \log(\delta_{coh})(5), \log(\lambda_{coh})(8)$	0.9958

The results from the baseline method are summarized in Table 1 . The subsets considered by the wrapper method ranged in size from one to eight, and the table displays the best-performing subsets of size 1–7. After size three, the best-performing subsets of size four and larger obtained only marginally higher scores than those displayed here. In the middle column of Table 1, the features in the highest scoring combination are listed along with their corresponding rank from the SHAP-induced ranking in Fig. 10.

The results from the baseline wrapper method illustrate a concerning trend when contrasted with the results from feature selection based on a SHAP-induced ranking alone. If the latter method were robust, we would expect to see a closer correspondence between the SHAP-induced ranking and the highest-scoring combinations by subset size. Instead, there were numerous discrepancies for subsets of three or more features. The limited success of the Shapley values for feature selection can largely be attributed to the interdependence of the inputs. This can cause various problems with Shapley values, two of which were borne out in our results. The first problem results from the combined explanatory power of two or more features being greater than the sum of their individual effects considered in isolation. When this phenomenon is present, the features are said to interact "synergistically". Because Shapley values are obtained by averaging the marginal contribution of an individual feature across the many possible feature coalitions, Shapley values for a given feature can be inflated when it is included in a coalition of features with which it interacts synergistically.

This effect is observed in the case of ς_s. As can be seen in Fig. 10, this feature obtained a relatively high average Shapley magnitude and would be selected for an optimal subset of three features using the SHAP informed approach. However, this can be attributed to the dependence of the feature on s_{min}. Demonstrating its robustness, the wrapper method picked up on this. After determining the optimal two feature subset of s_{min} and normalized loading rate, the wrapper method next selects a feature from the last unused cluster in Fig. 10, $\log(\lambda_{coh})$, thereby avoiding the inclusion of redundant information from the other two clusters.

Similarly, when a group of features is strongly correlated, Shapley values tend to be somewhat evenly distributed among the features in the group. When one of the features in the group is truly more influential to prediction of the QoI than the others, this can obfuscate the feature selection process. The phenomenon is manifested in our results when considering the feature $\log(\lambda_{coh})$. The features in the orange cluster to which $\log(\lambda_{coh})$ belongs all received relatively small, and comparable, average Shapley magnitudes, with D_s receiving the largest and obtaining a rank of fourth overall. However, the more robust wrapper method made a different selection. Beginning with the optimal subset of three features, it chooses $\log(\lambda_{coh})$ (ranked eighth overall) for all feature combinations of size three or larger rather than D_s. This implies that $\log(\lambda_{coh})$ provides more explanatory power than any other feature in the orange cluster when it comes to predicting the QoIs, more in line with the physical interpretation and an insight not provided by examination of Shapley values alone.

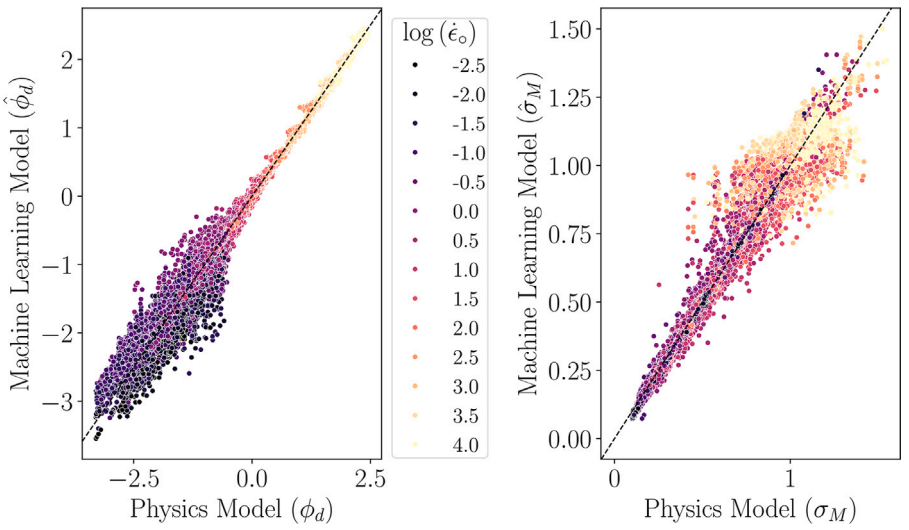


Fig. 11. Parity plots for the optimal 3 feature subset: Left: ϕ_d, $R^2 = 0.993$. Right: σ_M, $R^2 = 0.98$. Notice that for ϕ_d there is significantly more spread in the ML predictions for $\log(\dot{\epsilon}_\circ) < -1$ ($\dot{\epsilon}_{\circ,\text{coh}} < 0.05\delta_{\text{coh}}$) as compared with $\log(\dot{\epsilon}_\circ) \geq -1$, indicating that quasi-static loading cases are more difficult for the surrogate model to predict. Contrast this with σ_M, where simulations with $\log(\dot{\epsilon}_\circ) \geq -1$ (mostly dynamic regime) are more difficult to predict.

A couple of conclusions are drawn from the comparison of all physics-based ML results. First, as the number of features in a subset grows in Table 1, one distinct entry from each of the dendrogram clusters in Fig. 9 is drawn to efficiently represent the full data. The order is loading rate $\log(\dot{\epsilon}_{\circ,\text{coh}})$, minimum strength $s_{\min}$ (amplitude variation cluster), and correlation length $\log(\lambda_{\text{coh}})$ (spatial-variation cluster). Second, the addition of $\log(\lambda_{\text{coh}})$ in the third row of Table 1 highlights the higher importance of spatial variation of the strength field over other input parameters. Notably, the SHAP-induced ranking did not pick up on this. $\log(\lambda_{\text{coh}})$ was ranked eighth.

Although the single most important parameter for any loading regime is well understood a priori, ML help identify the importance and effect of other secondary parameters whose impact on the QoIs is not clear from theory alone. For example, despite our clustering technique correctly organizing the features associated with the random field into meaningful groups, *i.e.,* those associated with the field's spatial variation (orange) and those associated with variation in the field's amplitude (green), it is difficult to know *which* feature to pick from each cluster in order to optimize the surrogate model's predictive capabilities. Some of the examples above illustrate this. As another example, the physics of fragmentation shows a high influence of $s_{\min}$ and s_{mean} on QoIs in quasi-static and dynamic loading rates, respectively. Still, for a surrogate model incorporating all loading rates, $s_{\min}$ ranks much higher than s_{mean} in Fig. 10 and Table 1, an aspect that could not have been known a priori.

To conclude the Machine Learning-Enabled Analysis we consider the features in the third row of Table 1 since, as was mentioned above, after this point the inclusion of more features results in only a marginal increase in R^2. This subset achieves the balance sought after when implementing feature selection: the number of features is significantly reduced without a substantial reduction in accuracy (R^2 on the holdout testing data was 0.996 when all fifteen features are included in training). A parity plot for this model is displayed in Fig. 11, with each prediction colored according to the corresponding simulation's loading rate $\dot{\epsilon}_\circ$.

The predictions from the ML model mirror phenomena of the physics-based model discussed in §3 and the coloring in the parity plot is intended to highlight this. Namely, the standard deviation of dissipated energy predictions is much higher for values of $\log(\dot{\epsilon}_{\circ,\text{coh}}) < 0$ (mostly the quasi-static regime) as compared with higher values. As this variability is largely a result of the physical simulations, reducing it in the ML model would be difficult. However, using only the three features highlighted in the feature selection table above, one can accurately predict ϕ_d and σ_M utilizing a single network, greatly reducing the dimensionality of the input space.

As mentioned earlier, we by no means condemn SHAP itself or the utility it provides when factored into the feature selection process. However, we do caution against conflating a SHAP-induced feature ranking with feature selection more generally. Although the SHAP-induced ranking was able to consistently identify the two most important features in the baseline wrapper method, $\log(\dot{\epsilon}_{\circ,\text{coh}})$, and $s_{\min}$, the discrepancies between this ranking and the combinatorial approach are significant enough to warrant skepticism of feature selection using the SHAP-induced ranking alone. Further, a common motivation for implementing feature selection in the first place is to address problems associated with multicollinearity. Rather than alleviating such concerns, the SHAP-induced ranking includes both $s_{\min}$ and ς_s in the top 3, two features that exhibit a high degree of correlation ($\eta = -0.938$). All feature selection techniques have drawbacks, especially those such as the SHAP-informed approach, which is capable of providing relatively quick results. Our analysis is intended to outline some of these limitations in the hopes of informing researchers interested in utilizing SHAP-induced rankings for feature selection in the future.

5. Discussion and conclusions

We presented the fragmentation response of 1D rings with a heterogeneous fracture strength. The physical observations can be summarized as follows:

1. The fracture strength σ_M is almost equal to the minimum ($s_{\min}$) and mean (s_{mean}) point-wise strengths for quasi-static and higher loading rates, respectively, with a more complex behavior for intermediate loading rates.
2. The transition of σ_M to mean-based high loading regime occurs when the nondimensional loading length scale ($l_{\dot{\epsilon}_o}$) gets smaller than the nondimensional material heterogeneity length scale (λ).
3. The homogeneous material has a higher strength and toughness than all the combinations of material heterogeneity (in strength) considered.
4. As material heterogeneities occur at a finer length scale (λ decreases), there are more local point-wise strength minima in the ring. This results in lower scatter in both σ_M and ϕ_d and lower mean values for σ_M. The mean of ϕ_d is more complex as it depends on σ_M and fragment count. At high loading rates, the mean of ϕ_d decreases for smaller λ, whereas the effect at low rates depends on the strength of point-wise variations (r).

The results from the ML are consistent with these physical observations and reveal the following observations,

(1) The dendrogram obtained based on pairwise Spearman's rank correlation distances accurately categorized input features into nondimensional loading rate, fracture length scale, mean of fracture strength, and parameters that correspond to amplitude versus spatial variations of the inhomogeneous strength field. This exactly matches what is expected from the physics of the problem.
(2) The highlighted row of Table 1 summarizing the results from the wrapper method reveals a specific dimensionality reduction that only keeps the correlation length of all characteristics of the random field plus the minimum strength in a ring, in doing so accurately capturing the fracture response of the ring.
(3) The ordering of the inputs induced by their mean absolute Shapley values failed to identify this specific subset of features. We, therefore, caution against using this ordering to select a predetermined number of features naively. Although this proved successful in [76], our analysis showed that this ranking failed to fully capture key characteristics of the random field.
(4) By including those features that do capture key characteristics of the random field, rather than only those features that obtained a relatively high mean absolute Shapley value, we were able to reduce the total number of features substantially (from 15 to 3) while still retaining a model capable of accurately predicting both ϕ_d and σ_M simultaneously. Our final model using the three inputs $\log(\dot{\epsilon}_{o,\text{coh}})$, $s_{\min}$, and $\log(\lambda_{\text{coh}})$ obtained an $R^2 = 0.98$, only encountering difficulty when predicting ϕ_d for quasi-static loading rates.

Some of the limitations, and accordingly, areas for future improvements, are as follows. The initial set of 15 input features was manually selected. Ideally, a closed-loop system should be used to add and remove features that best represent the random field. This can also alleviate the problems with multicollinearity, where, for example, several closely related parameters represent spatial or amplitude variations. Finally, given the distinct physical behaviors of the problem (*e.g.,* low versus high loading rates), a clustering algorithm and subsequent predictive models for each cluster can improve the performance of the ML models. In Appendix A we provide a preliminary study on this aspect, wherein based on the three loading rate regime definitions of quasi-static, transitional, and dynamic from §3, we train three separate ML models. These models are now capable of capturing the most influential input parameters for each cluster. For example, s_{mean} becomes the second most important parameter for the dynamic regime, matching the results in §3 such as Fig. 7(d). Interestingly, s_{mean} appeared only in row 6 of Table 1, by which time the R^2 of the model for all conditions was already very high. Ideally, the clustering itself should be performed purely based on data in the ML process, rather than driven by physical insight. Developing such a framework for clustering, feature identification/selection, and modeling will certainly be of great value in scientific machine learning and should be explored in future research.

2D and 3D fragmentation problems are other natural extensions of this work. Due to their lower costs, oftentimes semi-analytical studies of fragmentation focus on the 1D case. Similar reasoning applies to our ML approach: we first verify that we can train a surrogate model for the 1D case, leveraging the manageable cost of the simulations to perform model selection tasks; *e.g.,* grid searching hyperparameter configurations using cross-validation. Following successful construction of a surrogate model in 1D, we can utilize a similar strategy for the 2D and 3D cases since 1D to 3D fragmentation problems are not qualitatively different. We then take advantage of many of the model selection results to warm start our model for the 2D and 3D cases, allowing us to bypass model selection and instead focus on fine-tuning a surrogate model for the new scenarios. One approach would be to use the 1D model to extract uncertainty information regarding a significantly smaller number of 2D and 3D simulations. This uncertainty information can then be used to inform an active learning approach, whereby the model's uncertainty regarding certain inputs informs researchers on which specific simulations need to be conducted, and subsequently trained on, to improve the precision of the model. This has the potential to greatly reduce the number of costly 2D or 3D simulations.

CRediT authorship contribution statement

Reza Abedi: Writing – review & editing, Writing – original draft, Supervision, Software, Funding acquisition, Conceptualization. **Colin Furey:** Writing – original draft, Visualization, Software, Methodology, Formal analysis. **Farhad Pourkamali-Anaraki:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization. **Giang Huynh:** Visualization, Software. **Erdem Caliskan:** Writing – review & editing, Visualization. **Alireza V. Amirkhizi:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Reza Abedi reports financial support was provided by US Army Research Laboratory. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors wish to thank the US Army Research Laboratory for continued support throughout this effort. This research was supported by CC DEVCOM Army Research Laboratory, USA through Cooperative Agreement W911NF-22-2-0102. The authors would also like to thank the University of Tennessee Infrastructure for Scientific Applications and Advanced Computing (ISAAC). A portion of the computation for this work was performed with ISAAC computational resources.

Appendix A

After conducting the analysis above we realized that further insight could be garnered from applying the same techniques conducted above to three partitions of the fragmentation data according to the loading regime present in the simulations. This resulted in a partition of the original dataset analyzed above into three subsets provided in §3.2: the quasi-Static loading regime where $\log(\dot{\epsilon}_\circ) < -1$, the dynamic loading regime where $\log(\dot{\epsilon}_\circ) > \max\{-\log(l_{\text{coh}}), -\log(\lambda)\}$ ($l_{\dot{\epsilon}_\circ} \lesssim \min(\lambda, l_{\text{coh}})$), and the transitional regime which consisted of all remaining simulations. Although we applied the same analysis to each of these three partitions, significantly different conclusions can be drawn from the results.

A.1. Quasi-static loading regime

Consisting of 21,886 simulations, this partition is the smallest of the three described above and was also the most challenging to model. Although we were able to model σ_M with a high degree of accuracy for this loading rate regime $R^2 \approx 1$, the surrogate model was unable to achieve the same accuracy when it came to ϕ_d where $R^2 = 0.89$ (both scores are computed on a holdout testing set in the same manner discussed in the ML section above). Similar discrepancies between a naive feature selection according to the mean absolute Shapley values and the wrapper method were observed in this partition as well, as evidenced by Fig. 12 and Table 2.

Several observations are made from the quasi-static dataset. First, the relatively poor performance of the ML model, especially the low R^2 score for ϕ_d is contributed to high level of variability observed for ϕ_d at quasi-static loading rates; see §3.3 for further discussion. Second, the normalized loading rate $\log(\dot{\epsilon}_{\text{o,coh}})$ has a low Shapley rank of 10 in Fig. 12(a) and is absent in Table 2. This is well expected by definition in the quasi-static regime dynamic mechanisms are supposed to be inconsequential. Third, the response is mainly driven by $s_{\min}$. Again, this is well explained by the physics of quasi-static fracture, where the strength is driven by the weakest link model. That is, $\sigma_M \approx s_{\min}$, *cf.* Fig. 7(a). Fourth, the appearance of $\log(\delta_{\text{coh}})$ as the second entry in Table 2 reflects the fact that energy dissipation is proportional to δ_{coh}. Fifth, given the appearance of both versions of the correlation length in row 3 of the table is equivalent to adding one of the correlation lengths to the entries of row two, $s_{\min}$ and $\log(\delta_{\text{coh}})$. This arises from the linear dependency of $\log(\lambda)$, $\log(\lambda_{\text{coh}})$, and $\log(\delta_{\text{coh}})$ in the form $\log(\lambda_{\text{coh}}) = \log(2) + \log(\lambda) - \log(\delta_{\text{coh}})$ (recall $\lambda_{\text{coh}} = 2\lambda/\delta_{\text{coh}}$). Thus, for row 3 and other cases that $\log(\lambda)$ and $\log(\lambda_{\text{coh}})$ simultaneously appear as model inputs, one can consider the more physically-motivated pair of $\log(\delta_{\text{coh}})$ and $\log(\lambda)$.

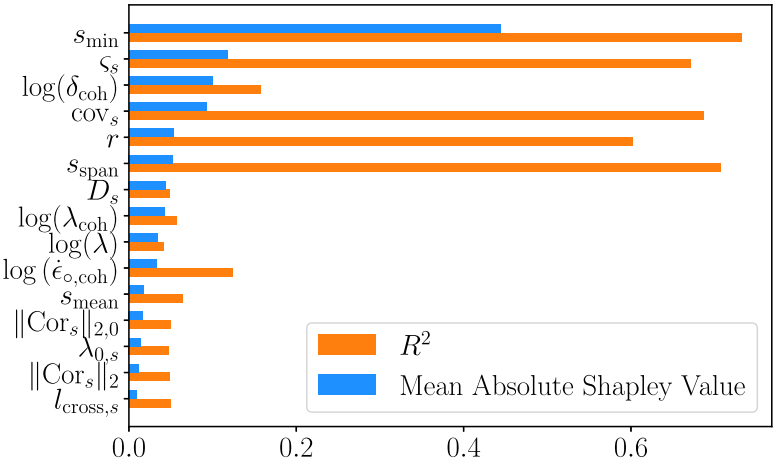


Fig. 12. Quasi-static loading regime: In a significant departure from the results obtained on the full dataset, $\log(\dot{\epsilon}_{\text{o,coh}})$ has significantly less influence.

Table 2
Quasi-static loading regime: Combinatoric Feature Selection Results.

No. Features	Highest Scored Feature Combinations (SHAP Rank)	R^2
1	$s_{\min}(1)$	0.73168
2	$s_{\min}(1)$, $\log(\delta_{\text{coh}})(3)$	0.89614
3	$s_{\min}(1)$, $\log(\lambda_{\text{coh}})(7)$, $\log(\lambda)(13)$	0.92432
4	$s_{\min}(1)$, $\log(\delta_{\text{coh}})(3)$, $\text{cov}_s(4)$, $D_s(7)$	0.92477
5	$s_{\min}(1)$, $\log(\delta_{\text{coh}})(3)$, $s_{\text{span}}(6)$, $D_s(7)$, $\text{cov}_s(4)$	0.92693
6	$s_{\min}(1)$, $\log(\delta_{\text{coh}})(3)$, $s_{\text{span}}(6)$, $D_s(7)$, $\text{cov}_s(4)$, $r(5)$	0.9258
7	$s_{\min}(1), \text{cov}_s(4), r(5), s_{\text{span}}(6), D_s(7), \log(\lambda_{\text{coh}})(7), \log(\lambda)(13)$	0.92615

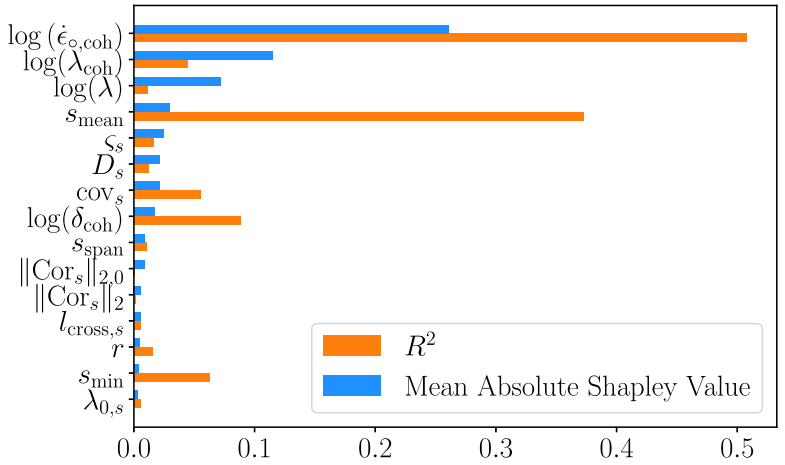


Fig. 13. Dynamic loading regime: Although $\log(\dot{\epsilon}_{\text{o,coh}})$ has retained its dominance, s_{mean} has replaced $s_{\min}$ as the most influential random field feature in this loading regime. Note s_{mean}'s relatively high SHAP induced ranking compared with the full data surrogate model, along with its relatively high R^2 score when the surrogate model for the dynamic loading regime is trained on a single feature.

Table 3
Dynamic loading regime: Combinatoric feature selection results.

No. Features	Highest Scored Feature Combinations (SHAP Rank)	R^2
1	$\log(\dot{\epsilon}_{\text{o,coh}})(1)$	0.50741
2	$\log(\dot{\epsilon}_{\text{o,coh}})(1)$, $s_{\text{mean}}(4)$	0.87769
3	$\log(\dot{\epsilon}_{\text{o,coh}})(1)$, $s_{\text{mean}}(4)$, $\text{cov}_s(7)$	0.91774
4	$\log(\dot{\epsilon}_{\text{o,coh}})(1)$, $s_{\text{mean}}(4)$, $\text{cov}_s(7)$, $\log(\delta_{\text{coh}})(8)$	0.97399
5	$\log(\dot{\epsilon}_{\text{o,coh}})(1), s_{\text{mean}}(4), \text{cov}_s(7), \log(\delta_{\text{coh}})(8), \log(\lambda_{\text{coh}})(2)$	0.98265
6	$\log(\dot{\epsilon}_{\text{o,coh}})(1), s_{\text{mean}}(4), \text{cov}_s(7), \log(\delta_{\text{coh}})(8), \log(\lambda)(2), s_{\min}(14)$	0.98591
7	$\log(\dot{\epsilon}_{\text{o,coh}})(1), \log(\lambda_{\text{coh}})(2), \log(\lambda)(3), s_{\text{mean}}(4), D_s(6), \text{cov}_s(7), \log(\delta_{\text{coh}})(8)$	0.98076

A.2. Dynamic loading regime

The dynamic loading regime consisted of 35,130 simulations in total. In contrast with the quasi-static loading regime, the surrogate model was able to predict both ϕ_d and σ_M with a high degree of precision (the surrogate model obtained $R^2 \approx 1$ for each QoI). More similar to the results obtained on the full dataset, $\log(\dot{\epsilon}_{\text{o,coh}})$ was the dominant feature when modeling this partition.

Some observations for the dynamic loading rate are as follows. First, as expected, the loading rate is the most important input parameter. Second, the mean value is the second most important parameter in Table 3. As discussed in §3.2 in highly dynamic fracture almost all points experience (partial) damage, thus the response is driven by mean properties. For example, in Fig. 7(d), we observed a clear linear correlation between σ_M and s_{mean} for this regime. The ML model captures the importance of s_{mean} for this subset, even though it was not highly ranked for the full dataset in Table 1. In fact, the combinatoric feature selection gives s_{mean} a higher importance than its Shapley rank of 4 in Fig. 13(a). Third, the new entry in the third row of the table is ς_s. Due to the lack of importance of $s_{\min}$ in the dynamic regime, ς_s turns out as a more important representative of the amplitude variation cluster (*cf.* Fig. 9). Fourth, the added entries in rows four and five are $\log(\delta_{\text{coh}})$ (influencing ϕ_d) and $\log(\lambda_{\text{coh}})$, which agree with the high Shapely ranks of 2 and 3 of $\log(\lambda_{\text{coh}})$ and $\log(\lambda)$ in Fig. 13(a).

A.3. Transitional loading regime

By far, the largest of the three subsets is the transitional loading regime, which consists of 83,792 simulations. Results for this partition are much closer to the full data results than those of the quasi-static and dynamic loading regimes. Further, the surrogate

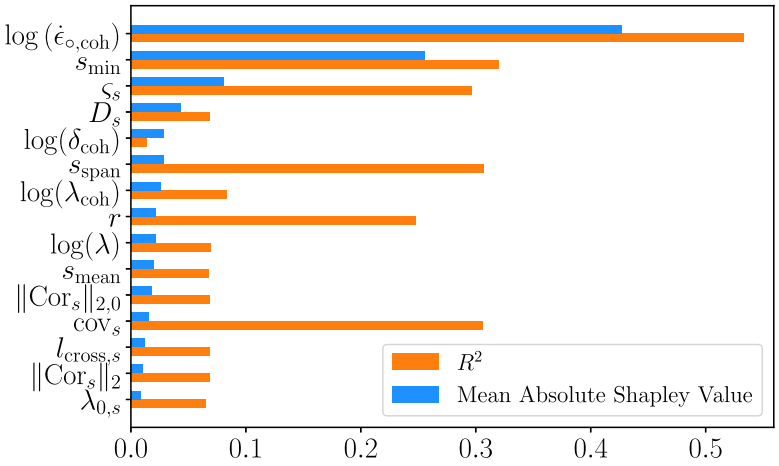


Fig. 14. Transitional loading regime: Similar to the full data surrogate model, $\log(\dot{\epsilon}_{\circ,\mathrm{coh}})$ and $s_{\min}$ are the most influential features for this loading regime.

Table 4
Transitional loading regime: Combinatoric feature selection results.

No. Features	Highest Scored Feature Combinations (SHAP Rank)	R^2
1	$\log(\dot{\epsilon}_{\circ,\mathrm{coh}})(1)$	0.53263
2	$\log(\dot{\epsilon}_{\circ,\mathrm{coh}})(1)$, $s_{\min}(2)$	0.9216
3	$\log(\dot{\epsilon}_{\circ,\mathrm{coh}})(1)$, $s_{\min}(2)$, $\log(\lambda_{\mathrm{coh}})(7)$	0.98436
4	$\log(\dot{\epsilon}_{\circ,\mathrm{coh}})(1)$, $s_{\min}(2)$, $\log(\lambda_{\mathrm{coh}})(7)$, $\log(\delta_{\mathrm{coh}})(5)$	0.98849
5	$\log(\dot{\epsilon}_{\circ,\mathrm{coh}})(1)$, $s_{\min}(2)$, $\log(\delta_{\mathrm{coh}})(5)$, $s_{\mathrm{span}}(6)$, $\log(\lambda)(9)$	0.99188
6	$\log(\dot{\epsilon}_{\circ,\mathrm{coh}})(1)$, $s_{\min}(2)$, $\log(\delta_{\mathrm{coh}})(5)$, $s_{\mathrm{span}}(6)$, $\log(\lambda)(9)$, $r(8)$	0.99131
7	$\log(\dot{\epsilon}_{\circ,\mathrm{coh}})(1)$, $s_{\min}(2)$, $\log(\delta_{\mathrm{coh}})(5)$, $s_{\mathrm{span}}(6)$, $\log(\lambda)(9)$, $r(8)$, $D_s(4)$	0.99213

model for this subset was able to accurately model each QoI with a high degree of accuracy, obtaining $R^2 \approx 1$ for ϕ_d and σ_M. The surrogate model for this subset, as with the full data surrogate model, was heavily influenced by the features $\log(\dot{\epsilon}_{\circ,\mathrm{coh}})$ and $s_{\min}$ (see Fig. 14 and Table 4).

The combinatoric feature selection also yielded similar results as the full data model, with both $\log(\dot{\epsilon}_{\circ,\mathrm{coh}})$ and $s_{\min}$ featuring prominently. In fact, in both cases the top four rows of the Table are populated in order by $\log(\dot{\epsilon}_{\circ,\mathrm{coh}})$, $s_{\min}$, log of correlation length ($\log(\lambda_{\mathrm{coh}})$ or $\log(\lambda)$), and $\log(\delta_{\mathrm{coh}})$. Recalling the linear dependence of $\log(\lambda_{\mathrm{coh}})$, $\log(\lambda)$, and $\log(\delta_{\mathrm{coh}})$, this again confirms that the ML model for the transitional regime closely resembles that for the full dataset. Interestingly, in both cases the high impact of s_{mean} on the response toward higher loading rates is shadowed by the influence of $s_{\min}$ that was the most important input parameter in the quasi-static loading regime.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cma.2024.117709.

Data availability

Data will be made available on request.

References

[1] F. Zhou, J.-F. Molinari, K. Ramesh, Effects of material properties on the fragmentation of brittle materials, Int. J. Fract. 139 (2) (2006) 169–196.
[2] D. Grady, Local inertial effects in dynamic fragmentation, J. Appl. Phys. 53 (1) (1982) 322–325.
[3] D. Grady, M. Kipp, Mechanisms of dynamic fragmentation: Factors governing fragment size, Mech. Mater. 4 (3–4) (1985) 311–320.
[4] L. Glenn, A. Chudnovsky, Strain–energy effects on dynamic fragmentation, J. Appl. Phys. 59 (4) (1986) 1379–1380.
[5] W. Drugan, Dynamic fragmentation of brittle materials: analytical mechanics-based models, J. Mech. Phys. Solids 49 (6) (2001) 1181–1208.
[6] F. Zhou, J.-F. Molinari, K. Ramesh, Analysis of the brittle fragmentation of an expanding ring, Comput. Mater. Sci. 37 (1–2) (2006) 74–85.
[7] F. Zhou, J.-F. Molinari, K. Ramesh, A cohesive model based fragmentation analysis: effects of strain rate and initial defects distribution, Int. J. Solids Struct. 42 (18–19) (2005) 5181–5207.
[8] S. Levy, J.-F. Molinari, Dynamic fragmentation of ceramics, signature of defects and scaling of fragment sizes, J. Mech. Phys. Solids 58 (1) (2010) 12–26.
[9] S. Levy, J.-F. Molinari, I. Vicari, A. Davison, Dynamic fragmentation of a ring: Predictable fragment mass distributions, Phys. Rev. E 82 (6) (2010) 066105.
[10] M. Ostoja-Starzewski, X. Wang, Stochastic finite elements as a bridge between random material microstructure and global response, Comput. Methods Appl. Mech. Engrg. 168 (1–4) (1999) 35–49.

[11] J. Guilleminot, A. Noshadravan, C. Soize, R.G. Ghanem, A probabilistic model for bounded elasticity tensor random fields with application to polycrystalline microstructures, Comput. Methods Appl. Mech. Engrg. 200 (17–20) (2011) 1637–1648.
[12] J. Guilleminot, C. Soize, Generalized stochastic approach for constitutive equation in linear elasticity: a random matrix model, Internat. J. Numer. Methods Engrg. 90 (5) (2012) 613–635.
[13] D. Cereceda, L. Graham-Brady, N. Daphalapurkar, Modeling dynamic fragmentation of heterogeneous brittle materials, Int. J. Impact Eng. 99 (2017) 85–101.
[14] B. Bahmani, M. Yang, A. Nagarajan, P.L. Clarke, S. Soghrati, R. Abedi, Automated homogenization-based fracture analysis: Effects of SVE size and boundary condition, Comput. Methods Appl. Mech. Engrg. 345 (2019) 701–727.
[15] T. Hu, J. Guilleminot, J.E. Dolbow, A phase-field model of fracture with frictionless contact and random fracture properties: Application to thin-film fracture and soil desiccation, Comput. Methods Appl. Mech. Engrg. 368 (2020) 113106.
[16] V. Dubey, A. Noshadravan, A probabilistic upscaling of microstructural randomness in modeling mesoscale elastic properties of concrete, Comput. Struct. 237 (2020) 106272.
[17] S. Ranftl, M. Rolf-Pissarczyk, G. Wolkerstorfer, A. Pepe, J. Egger, W. von der Linden, G.A. Holzapfel, Stochastic modeling of inhomogeneities in the aortic wall and uncertainty quantification using a bayesian encoder–decoder surrogate, Comput. Methods Appl. Mech. Engrg. 401 (2022) 115594.
[18] M. Ortiz, A. Pandolfi, Finite-deformation irreversible cohesive elements for three-dimensional crack-propagation analysis, Internat. J. Numer. Methods Engrg. 44 (9) (1999) 1267–1282.
[19] F. Zhou, J.-F. Molinari, K. Ramesh, Characteristic fragment size distributions in dynamic fragmentation, Appl. Phys. Lett. 88 (26) (2006) 261918.
[20] R. Abedi, R.B. Haber, Riemann solutions and spacetime discontinuous Galerkin method for linear elastodynamic contact, Comput. Methods Appl. Mech. Engrg. 270 (2014) 150–177.
[21] R. Abedi, R.B. Haber, Spacetime simulation of dynamic fracture with crack closure and frictional sliding, Adv. Model. Simul. Eng. Sci. 5 (1) (2018) 22, Equal contribution authorship.
[22] R. Abedi, G. Huynh, F. Pourkamai-Anaraki, A. Amirkhizi, C. Furey, H. Lee, C.J. Hansen, AI-assisted statistical analysis of fragmentation response of heterogeneous layered media, in: Proceedings of the American Society for Composites: Thirty-Eights Technical Conference, 2023, 004684.
[23] J.-P. Fouque, J. Garnier, G. Papanicolaou, K. Solna, Wave Propagation and Time Reversal in Randomly Layered Media, vol. 56, Springer Science & Business Media, 2007.
[24] L. Sun, H. Gao, S. Pan, J.-X. Wang, Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data, Comput. Methods Appl. Mech. Engrg. 361 (2020) 112732.
[25] Y. Zhu, N. Zabaras, P.-S. Koutsourelakis, P. Perdikaris, Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data, J. Comput. Phys. 394 (2019) 56–81.
[26] M.N. Guddati, J.L. Tassoulas, An efficient numerical algorithm for transient analysis of exterior scalar wave propagation in a homogeneous layer, Comput. Methods Appl. Mech. Engrg. 167 (3–4) (1998) 261–273.
[27] M.N. Guddati, J.L. Tassoulas, Characteristics methods for transient analysis of wave propagation in unbounded media, Comput. Methods Appl. Mech. Engrg. 164 (1–2) (1998) 187–206.
[28] N. Favrie, A. Renaud, D. Kondo, Hyperbolic modeling of gradient damage and one-dimensional finite volume simulations, Comput. Methods Appl. Mech. Engrg. 419 (2024) 116643.
[29] R. Abedi, P.L. Clarke, A computational approach to model dynamic contact and fracture mode transitions in rock, Comput. Geotech. 109 (2019) 248–271.
[30] Q. Zhan, Q. Ren, M. Zhuang, Q. Sun, Q.H. Liu, An exact riemann solver for wave propagation in arbitrary anisotropic elastic media with fluid coupling, Comput. Methods Appl. Mech. Engrg. 329 (2018) 24–39.
[31] Q. Zhan, Q. Sun, M. Zhuang, Y. Mao, Q. Ren, Y. Fang, W.-F. Huang, Q.H. Liu, A new upwind flux for a jump boundary condition applied to 3d viscous fracture modeling, Comput. Methods Appl. Mech. Engrg. 331 (2018) 456–473.
[32] R. Abedi, Characteristics1d: Solver for advection equation using method of characteristics, 2023, URL https://github.com/rabedi/Characteristics1D.
[33] R. Abedi, R. Haber, Spacetime dimensional analysis and self-similar solutions of linear elastodynamics and cohesive dynamic fracture, Int. J. Solids Struct. 48 (13) (2011) 2076–2087.
[34] E. Vanmarcke, Random Fields: Analysis and Synthesis, World Scientific, 2010.
[35] K. Karhunen, Über lineare Methoden in der Wahrscheinlichkeitsrechnung, vol. 37, Universitat Helsinki, 1947.
[36] K. Karhunen, I. Selin, On Linear Methods in Probability Theory, Rand Corporation, 1960.
[37] P. Robbe, Gaussianrandomfields.jl: A julia package to generate and sample from gaussian random fields, J. Open Source Softw. 8 (89) (2023) 5595.
[38] K.A. Acton, S.C. Baxter, B. Bahmani, P.L. Clarke, R. Abedi, Voronoi tessellation based statistical volume element characterization for use in fracture modeling, Comput. Methods Appl. Mech. Engrg. 336 (2018) 135–155.
[39] S. Kotz, S. Nadarajah, Extreme Value Distributions: Theory and Applications, world scientific, 2000.
[40] I. Rychlik, Five lectures on reliability applications of rice's formula for the intensity of level crossings, in: AMAS Lecture Notes, Reliability-Based Design and Optimisation RBO, Vol. 3, 2004.
[41] J. Masoliver, M. Palassini, Level-crossing counting for inertial random processes. rice's approach revisited, 2023, arXiv e-prints arXiv–2301.
[42] INuritdino, INuritdino higuchi Fractal Dimension, 2024, URL https://github.com/inuritdino/HiguchiFractalDimension. (Accessed 29 March 2024).
[43] Z.P. Bazant, J.-L. Le, Probabilistic Mechanics of Quasibrittle Structures: Strength, Lifetime, and Size Effect, Cambridge University Press, 2017.
[44] M. Ostoja-Starzewski, J. Trebicki, On the growth and decay of acceleration waves in random media, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 455 (1987) (1999) 2577–2614.
[45] J.-L. Le, J. Eliáš, A. Gorgogianni, J. Vievering, J. Květoň, Rate-dependent scaling of dynamic tensile strength of quasibrittle structures, J. Appl. Mech. 85 (2) (2018).
[46] E. Debie, K. Shafi, Implications of the curse of dimensionality for supervised learning classifier systems: theoretical and empirical analyses, Pattern Anal. Appl. 22 (2019) 519–536.
[47] M.C. Alonso, J.A. Malpica, A.M. de Agirre, Consequences of the hughes phenomenon on some classification techniques, in: ASPRS 2011 Annual Conference, 2011, pp. 1–5.
[48] L.O. Jimenez-Rodriguez, E. Arzuaga-Cruz, M. Vélez-Reyes, Unsupervised linear feature-extraction methods and their effects in the classification of high-dimensional data, IEEE Trans. Geosci. Remote Sens. 45 (2) (2007) 469–483.
[49] A.A. Joy, M.A.M. Hasan, M.A. Hossain, A comparison of supervised and unsupervised dimension reduction methods for hyperspectral image classification, in: 2019 International Conference on Electrical, Computer and Communication Engineering, ECCE, IEEE, 2019, pp. 1–6.
[50] W.E. Marcílio, D.M. Eler, From explanations to feature selection: assessing shap values as feature selection mechanism, in: 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images, SIBGRAPI, Ieee, 2020, pp. 340–347.
[51] J. Verhaeghe, J. Van Der Donckt, F. Ongenae, S. Van Hoecke, Powershap: A power-full shapley feature selection method, in: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer, 2022, pp. 71–87.
[52] F. Mokdad, D. Bouchaffra, N. Zerrouki, A. Touazi, Determination of an optimal feature selection method based on maximum shapley value, in: 2015 15th International Conference on Intelligent Systems Design and Applications, ISDA, IEEE, 2015, pp. 116–121.

[53] S. Saha, Z. Gan, L. Cheng, J. Gao, O.L. Kafka, X. Xie, H. Li, M. Tajdari, H.A. Kim, W.K. Liu, Hierarchical deep learning neural network (hidenn): an artificial intelligence (ai) framework for computational science and engineering, Comput. Methods Appl. Mech. Engrg. 373 (2021) 113452.
[54] G. Chandrashekar, F. Sahin, A survey on feature selection methods, Comput. Electr. Eng. 40 (1) (2014) 16–28.
[55] P. Dhal, C. Azad, A comprehensive survey on feature selection in the various fields of machine learning, Appl. Intell. 52 (4) (2022) 4543–4581.
[56] J. Miao, L. Niu, A survey on feature selection, Procedia Comput. Sci. 91 (2016) 919–926.
[57] G. Chen, J. Chen, A novel wrapper method for feature selection and its applications, Neurocomputing 159 (2015) 219–226.
[58] Z. Liu, J. Yang, L. Wang, Y. Chang, A novel relation aware wrapper method for feature selection, Pattern Recognit. 140 (2023) 109566.
[59] U.F. Njoku, A. Abelló Gamazo, B. Bilalli, G. Bontempi, Wrapper methods for multi-objective feature selection, in: 26th International Conference on Extending Database Technology (EDBT 2023): Ioannina, Greece, March 28-March 31: Proceedings, OpenProceedings, 2023, pp. 697–709.
[60] J. Linja, J. Hämäläinen, P. Nieminen, T. Kärkkäinen, Feature selection for distance-based regression: An umbrella review and a one-shot wrapper, Neurocomputing 518 (2023) 344–359.
[61] F. Masood, J. Masood, H. Zahir, K. Driss, N. Mehmood, H. Farooq, Novel approach to evaluate classification algorithms and feature selection filter algorithms using medical data, J. Comput. Cogn. Eng. 2 (1) (2023) 57–67.
[62] B. Parlak, A.K. Uysal, A novel filter feature selection method for text classification: Extensive feature selector, J. Inf. Sci. 49 (1) (2023) 59–78.
[63] D.Ö. Şahin, O.E. Kural, S. Akleylek, E. Kılıç, A novel android malware detection system: adaption of filter-based feature selection methods, J. Ambient Intell. Humaniz. Comput. (2023) 1–15.
[64] A. Bommert, X. Sun, B. Bischl, J. Rahnenführer, M. Lang, Benchmark for filter methods for feature selection in high-dimensional classification data, Comput. Statist. Data Anal. 143 (2020) 106839.
[65] B.Q. Lap, H. Du Nguyen, P.T. Hang, N.Q. Phi, V.T. Hoang, P.G. Linh, B.T.T. Hang, et al., Predicting water quality index (wqi) by feature selection and machine learning: A case study of an kim hai irrigation system, Ecol. Inform. 74 (2023) 101991.
[66] Q. Wang, C. Yue, X. Li, P. Liao, X. Li, Enhancing robustness of monthly streamflow forecasting model using embedded-feature selection algorithm based on improved gray wolf optimizer, J. Hydrol. 617 (2023) 128995.
[67] M. Yousef, F. Ozdemir, A. Jaber, J. Allmer, B. Bakir-Gungor, Pripath: identifying dysregulated pathways from differential gene expression via grouping, scoring, and modeling with an embedded feature selection approach, BMC Bioinformatics 24 (1) (2023) 60.
[68] W. Zhao, S. Feng, J. Liu, B. Sun, An explainable intelligent algorithm for the multiple performance prediction of cement-based grouting materials, Constr. Build. Mater. 366 (2023) 130146.
[69] Y. Chen, A. Cheng, C. Zhang, S. Chen, Z. Ren, Rapid mechanical evaluation of the engine hood based on machine learning, J. Braz. Soc. Mech. Sci. Eng. 43 (7) (2021) 345.
[70] E. Brusa, L. Cibrario, C. Delprete, L.G. Di Maggio, Explainable ai for machine fault diagnosis: understanding features' contribution in machine learning models for industrial condition monitoring, Appl. Sci. 13 (4) (2023) 2038.
[71] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst. 30 (2017).
[72] M. Sundararajan, A. Najmi, The many shapley values for model explanation, in: H.D. III, A. Singh (Eds.), Proceedings of the 37th International Conference on Machine Learning, in: Proceedings of Machine Learning Research, vol. 119, PMLR, 2020, pp. 9269–9278.
[73] H. Chen, I.C. Covert, S.M. Lundberg, S.-I. Lee, Algorithms to estimate shapley value feature attributions, Nat. Mach. Intell. 5 (6) (2023) 590–601.
[74] J. Zhang, Q. Sun, J. Liu, L. Xiong, J. Pei, K. Ren, Efficient sampling approaches to shapley value approximation, Proc. ACM Manag. Data 1 (1) (2023) 1–24.
[75] R. Mitchell, J. Cooper, E. Frank, G. Holmes, Sampling permutations for shapley value estimation, J. Mach. Learn. Res. 23 (43) (2022) 1–46.
[76] D.-C. Feng, W.-J. Wang, S. Mangalathu, E. Taciroglu, Interpretable xgboost-shap machine-learning model for shear strength prediction of squat rc walls, J. Struct. Eng. 147 (11) (2021) 04021173.
[77] J. Jeon, N. Seo, S.B. Son, S.-J. Lee, M. Jung, Application of machine learning algorithms and shap for prediction and feature analysis of tempered martensite hardness in low-alloy steels, Met. 11 (8) (2021) 1159.
[78] F.E. Bock, Z. Kallien, N. Huber, B. Klusemann, Data-driven and physics-based modelling of process behaviour and deposit geometry for friction surfacing, Comput. Methods Appl. Mech. Engrg. 418 (2024) 116453.
[79] E. Temizhan, H. Mirtagioglu, M. Mendes, et al., Which correlation coefficient should be used for investigating relations between quantitative variables, Acad. Sci. Res. J. Eng. Technol. Sci. 85 (2022) 265–277.
[80] F. Nielsen, Hierarchical clustering, in: Introduction to HPC with MPI for Data Science, 2016, pp. 195–211.
[81] P. Govender, V. Sivakumar, Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019), Atmospheric Pollut. Res. 11 (1) (2020) 40–56.
[82] A.M. Harsha, V.K. Rao, A comprehensive analysis of clustering methods for portfolio optimization, Int. J. Intell. Syst. Appl. Eng. 12 (10s) (2024) 27–36.
[83] F. Murtagh, P. Legendre, Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? J. Classification 31 (2014) 274–295.