

“THEME ARTICLE”, “FEATURE ARTICLE”, or “COLUMN” goes here: The theme topic or column/department name goes after the colon.

An Interactive Exploration Tool for High-Dimensional Datasets: A Shock Physics Case Study

Ayan Biswas

Los Alamos National Laboratory

Christopher M. Biwer

Los Alamos National Laboratory

David J. Walters

Los Alamos National Laboratory

James Ahrens

Los Alamos National Laboratory

Devin Francom

Los Alamos National Laboratory

Earl Lawrence

Los Alamos National Laboratory

Richard L. Sandberg

Los Alamos National Laboratory

D. Anthony Fredenburg

Los Alamos National Laboratory

Cynthia Bolme

Los Alamos National Laboratory

Validating simulations with experimental results is a vital component of modern materials science. Existing workflows require substantial human intervention to interpret their results. We present a statistical approach to identify physically-meaningful features and we construct a visualization system that allows users to interactively and intuitively explore their datasets.

Materials scientists studying how materials respond to harsh environments involving materials failure (i.e. damage and shock) frequently work with ensemble datasets that are generated from the simulations as well as the experiments. These simulations often involve multiple input parameters (e.g., 10~20 parameters) and the output is a 1-dimensional time-series. In this high-dimensional dataset, generally there does not exist a uniquely resolved output that matches exactly with the experiments. In an effort to understand their datasets, the domain scientists have to go through the thousands of output curves manually, look for different features in these curves and try to find the input parameter combinations for which the simulation produces similar output as the experiments. This is a tedious and time-consuming task for the domain scientists and the existing data analysis tools that can help them achieve their goals, is mostly lacking.

Materials scientists studying how materials respond to harsh environments involving materials failure (i.e. damage and shock) frequently work with ensemble datasets that are generated from the simulations as well as the experiments. These simulations often involve multiple input parameters (e.g., 10~20 parameters) and the output is a 1-dimensional time-series. In this high-dimensional dataset, generally there does not exist a uniquely resolved output that matches exactly with the experiments. In an effort to understand their datasets, the domain scientists have to go through the thousands of output curves manually, look for different features in these curves and try to find the input parameter combinations for which the simulation produces similar output as the experiments. This is a tedious and time-consuming task for the domain scientists and the existing data analysis tools that can help them achieve their goals, is mostly lacking.

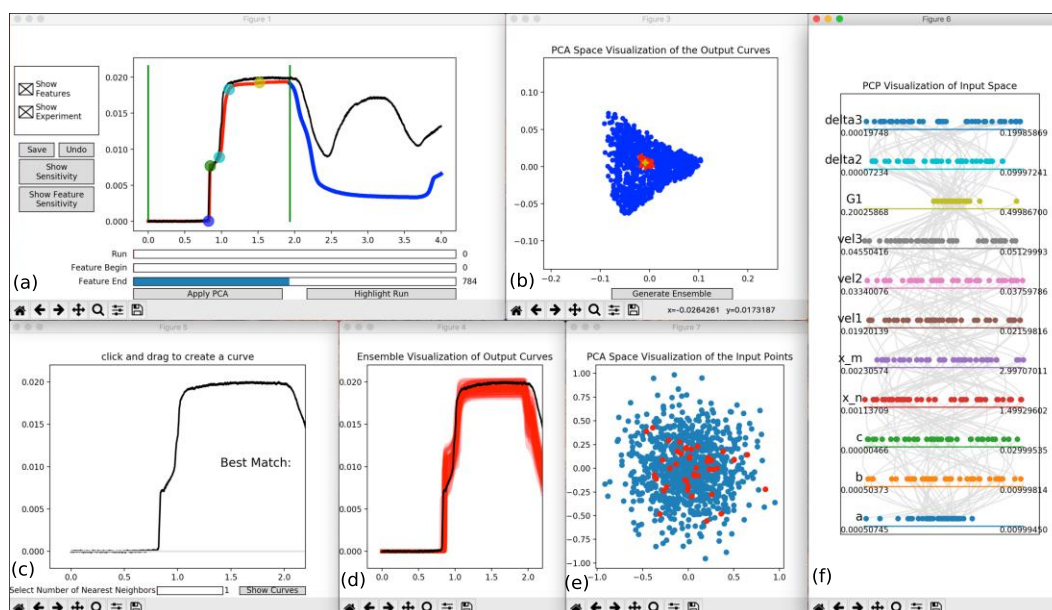


Figure 1. Overview of the proposed system: (a) the initial simulation/experiment explorer window, (b) output principal component analysis (PCA) space where proximity represents similarity in the velocimetry curves, (c) users can draw free-form curves for flexibly querying different velocimetry curves, (d) this window shows the output of the queries made in the window(c) by drawing, (e) PCA space showing the input parameter combinations, (f) visualization of input parameter combinations using parallel coordinates plot (PCP) for inter-parameter relationship exploration.

Interactive visual analysis of curves has been a well-researched topic in the past^{24,25}. Although there exist multiple data analysis and visualization tools that provide generic solutions to scientific problems (see “Related Tools” Section), there are certain problems - like the one above - that require specialized tools that utilize domain-specific knowledge or operations and still be lightweight. Examples of domain-specific software and tools are CoFlaVis¹, ImageVis3D²¹, and Seg3D²². The teams behind SCIRun and CoFlaVis have recognized and documented the usefulness of such domain-specific tools which make for easier maintainability, smaller codebases, and less time to solution for the users.

In this work, we have developed a data analysis and visualization tool for exploring high-dimensional dataset by closely working with the materials scientists of the shock physics domain. Given high-dimensional input and output data, our tool provides interaction capabilities to the users for feature extraction, feature correction, simulation validation, parameter importance exploration, and intuitive drawing-based feature query within a unified interface. To facilitate the exploration process, our proposed tool uses ensemble spaghetti plots for initial data visualization, principal component analysis (PCA)-based similarity analysis for validation and interactive selection, parallel coordinates plot (PCP)-based visualization for high-dimensional input parameter combination, existing sensitivity methods for understanding parameter importance, and k-nearest neighbor (KNN)-based analysis for drawing-based curve selection. We work closely with domain scientists to understand their requirements, and their feedback shows the efficacy of our proposed exploration tool.

RELATED TOOLS

For general data analysis and visualization, there are different tools and software available. Visualization Tool Kit or VTK¹⁴ is a popular open-source software with multiple functionalities for data visualization such as isocontours, streamlines, and volume rendering etc. ParaView¹⁵ is built on top of VTK to provide a graphical interface for ease of use. Another open-source tool that can

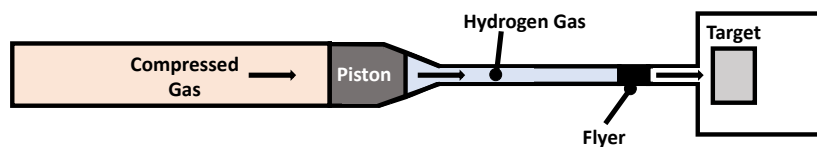


Figure 2. Schematic of the experimental setup. Compressed gas is released which launches a flyer down a barrel. At the end of the barrel, the flyer impacts a target and the impact causes a shock wave to travel through the target.

provide effective data visualization is VisIt¹⁶. MATLAB¹⁷ is a multi-paradigm numerical computing environment that is popular for basic data analysis and visualizations. Tools developed in Python¹⁸ are now becoming popular open-source alternatives to MATLAB scripts. MayaVi¹⁹ is a Python-based tool that uses VTK to provides effective data visualization. Amira 3D Software for Life Sciences²⁰ is another product that provides 3D and 4D visualizations for life science, CFD, and materials science datasets. CoFlaVis¹, ImageVis3D²¹, and Seg3D²² provide more domain-specific solutions for a set of targeted users. Our proposed tool is Python-based due to its popularity and cross-platform support.

DATA DESCRIPTION

Shock Physics Experiments and Simulations

In this work, the material properties of wrought aluminum alloy Al-5083 have been analyzed using experimental results and corresponding simulation outputs. The experiments are described in Boteler and Dandekar². They are performed with a gas gun which uses compressed gas to accelerate a flyer made of one material into a target made of another (or the same) material². A schematic of this experimental setup is shown in Figure 2. The impact has enough energy to generate a shock wave within the target material whose characteristics are governed by the material and the nature of the impact. If the impact is characterized well, then velocity measurements using optical interferometry⁴ are translated into information about the internal state of the material, such as stress, strain, pressure, and temperature using constitutive models for the elastic response⁵ (reversible deformation) and plastic response^{3,6} (irreversible deformation). In this work, there are three experiments under study that are labeled as Shot 104S, Shot 105S and Shot 106S and output of Shot 104S is shown in Figure 3 (in black).

Numerical hydrodynamic simulations of materials under intense thermodynamic conditions offer an alternative to shock physics experiments which require considerable infrastructure, time, and financial support. Simulations can provide a quick indication of whether a material is unsuitable for a particular application due to its plasticity. Hydrodynamic simulations are validated by comparing their output to experimental results. In this simulation, there are total 11 tunable input parameters: five of these are the popular Johnson-Cook strength model⁶ parameters (A , B , C , X_n , and X_m), three from varying the shear modulus (G , δ_2 , and δ_3) for the three experiments, and the three corresponding impact velocities (v_1 , v_2 , and v_3). Simulations of velocity profiles from the impact of two Al-5083 plates were conducted using the FLAG hydrodynamics code⁷ on the Los Alamos National Laboratory supercomputing platform Moonlight (308 Intel Xeon E5-267 nodes and 4928 CPU cores). The ensemble of simulated velocity profiles that were analyzed for Shot 104S are shown in Figure 3 (in red).

Features as Material Properties

Features in the velocity profiles are related to physical properties of the material and accurate identification of these features is of great importance to the scientists. There are primarily four features used to measure these properties and validate simulations against experimental results.

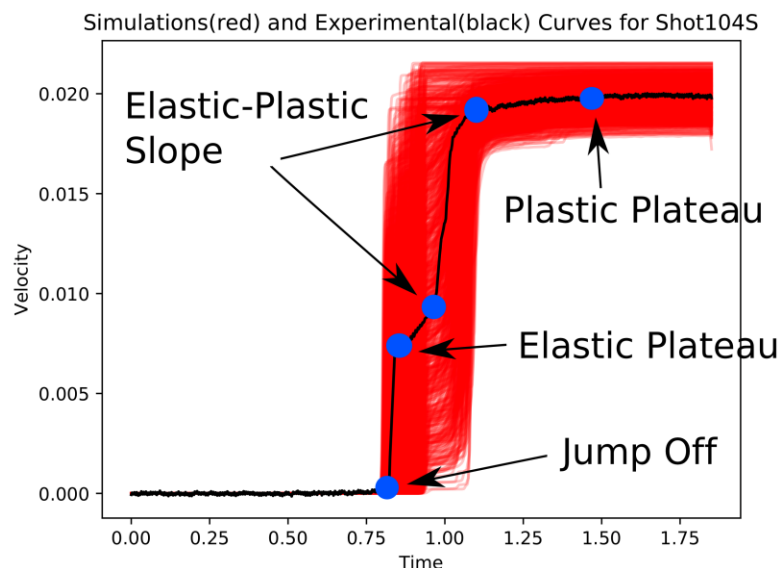


Figure 3. Velocity profile from one experimental setup (black) and an ensemble of simulated velocity profiles generated using the FLAG hydrodynamics simulation (red). The blue dots on the experimental curve represent the different features on the velocimetry curve.

An example of these features is shown in Figure 3. Here, we provide a brief explanation of how each feature relates to properties of the material:

1. **Jump off:** The jump-off is the time when the elastic wave reaches the rear surface of the target material. This time depends on the propagation of the elastic wave through the material which is parameterized by the thickness, density, and speed of sound through the material.
2. **Elastic plateau:** Plastic deformation waves travel slower than elastic deformation waves. Following the impulse of acceleration from the elastic wave, the rear surface of the target material accelerates little until the arrival of the plastic wave; this is represented as a plateau in Figure 3. Materials with a higher yield stress will be able to convert more of the impact energy into elastic deformation resulting in a higher elastic plateau.
3. **Elastic-plastic slope:** The slope at the arrival of the plastic wave is related to characteristics of the material's plastic deformation (e.g. strain hardening/softening, strain rate effects, temperature effects, etc.). Generally, a smaller slope indicates that the plastically deforming material is much less stiff than the elastically deforming portion.
4. **Plastic plateau:** The velocities at the elastic and plastic plateaus can be used to determine the velocity of the shock wave in the material. The speed of the shock wave in combination with equations for the conservation of mass and momentum can be used to estimate the internal state of the material to help determine under what conditions the material begins to plastically (irreversibly) deform.

Although these four features can be easily pointed out in a regular shaped curve, it is possible for the simulation and experiment to generate curves that are very irregular shaped and where the feature locations overlap. Due to this, feature extraction from these curves become challenging and uncertain.

PROBLEM OVERVIEW AND SYSTEM REQUIREMENTS

Hydrodynamic modelers typically visually inspect an ensemble of velocity profiles, record the features for each velocity profile, perform validation against the experimental data, and analyze

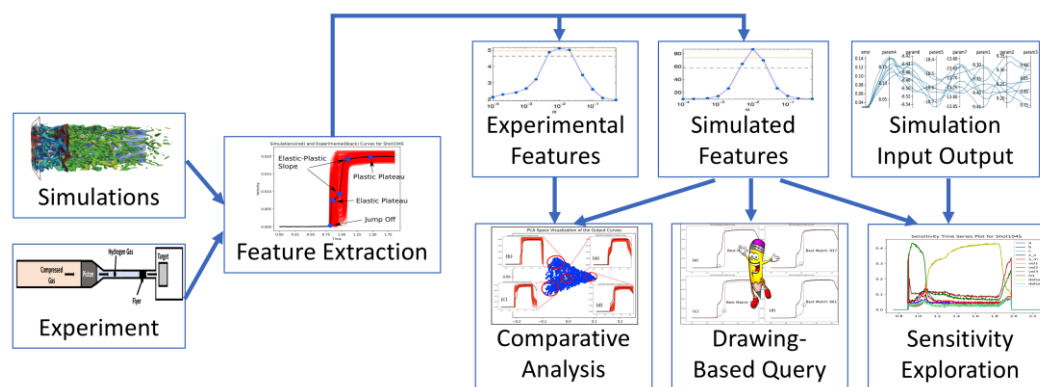


Figure 4. Overview of the proposed system. In this work-flow, simulations and experimental datasets are used for initial feature extraction using our GMM-based approach which are then further analyzed for understanding their sensitivity towards the input parameters. Our system also allows the users to query for features and hand-drawn curves in real time using PCA and KNN based approaches. PCP is used for high-dimensional visualization of the input space.

the input parameters to understand their effect on the output for tuning their values. This existing workflow can take up from days to weeks for each of the ensemble runs with moderate number of simulation runs (e.g., 1000 runs). A data analysis and visualization system that can enable the materials scientists to explore and understand these different aspects of their data is mostly lacking. To create a visualization system that can help answer the domain-specific questions, we first discussed with the domain experts regarding their existing data visualization needs and requirements. As noted from these discussions regarding the primary requirements of the intended system are as follows:

- 1) Accurate and automated identification of the features of the ensemble curves
- 2) Importance (i.e. sensitivity) analysis of the input parameters such that more important parameters can be tuned earlier
- 3) Exploration of simulated output to identify the suitable parameter combinations that produce data similar to the experiments
- 4) Fast and easy query mechanism of the features/curves

SYSTEM OVERVIEW

A schematic of our proposed workflow is shown in Figure 4. Given an ensemble of simulations and their corresponding experiments, we apply a feature detection algorithm to extract the material properties from the velocity profiles. Since these features are uncertain, we devise a statistical Gaussian mixture model-based method for fast and accurate identification of the features. Since, there exist multiple uncertain features, a comparison of different simulation runs with their experimental data requires a non-trivial feature comparison step. We employ principal component analysis (PCA) to reduce the dimensionality of the velocimetry time-series data and provide a 2D view of the simulations and the experiments for interactive exploration. With the extracted features and simulation input parameters, we now apply detailed sensitivity analysis within and across the range of input parameters and extracted features. Parallel coordinates plots (PCP) and an interactive drawing-based query mechanism aids in the exploration process. The use of innovative yet simple visualization and analysis tools help guide the exploration of the high-dimensional shock physics datasets.

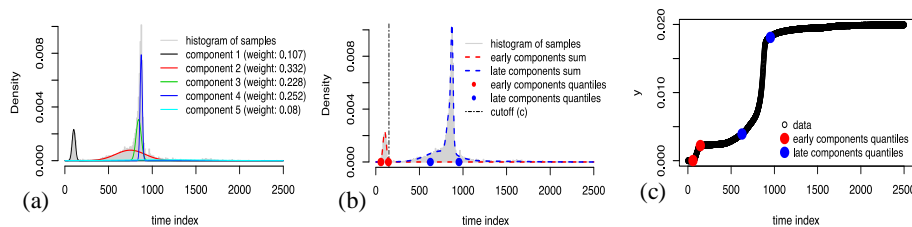


Figure 5. (a) We approximate the density function with a mixture of Gaussians. (b) We use a five-component mixture and separate the components into two groups - early and late components. Features correspond to certain quantiles from the groups. (c) The identified features on top of the original data.

UNCERTAIN FEATURE DETECTION

Velocity profiles are important to materials scientists since the location of features in this time series stand for the different properties of the material. There is no generally accepted method for finding features that will work for all cases, though there is some research in that direction⁸. In this work, we propose a Gaussian mixture model (GMM) algorithm to solve our domain-specific problem. Next, we briefly describe our novel GMM-based automated method to identify these features.

Feature Finding Method

We have developed a statistical approach to detect these four features. In our proposed method, the N velocity profiles $y_i(t)$, $i \in 0, \dots, N$ are treated as cumulative density functions (CDFs), and we treat their derivatives, $y_i'(t)$, as probability density functions (PDFs). The PDFs are then modeled as a mixture of Gaussians. While it may appear that two mixture components would be sufficient after visual inspection of Figure 5, we use more mixture components in order to (1) capture tail behavior, and (2) accurately model the CDF when there are clearly more than two jumps in $y(t)$. Figure 5a shows five mixture components for one velocity profile. The four features that describe the deformation properties of a material on a curve $y_i(t)$ are obtained by finding certain quantiles from groups of mixture components. We use quantiles in the tails of the components instead of means because they correspond to places where $y_i'(t)$ is nearing zero, and hence $y_i(t)$ is flattening out. These are the points of interest to the domain scientists.

We split the five mixture components into two groups which we call the early and late groups. If $\mu_j + 2\sigma_j < c$ (where μ_j is the mean and σ_j is the standard deviation of the j th Gaussian component), component j is assigned to the early group, indicating that most of the mass of the component occurs before c . The cutoff c is recognized visually when all the data are plotted together (a time point between the elastic plateau and the first elastic-plastic slope features). The jump-off and elastic plateau features are the 0.01 and 0.99 quantiles of the early components. These quantiles can be obtained by rescaling the early components so that the weights add up to one, integrating the summed components to get a CDF, and then inverting that CDF. The slope features are the 0.1 and 0.9 quantiles of the late components. An example of early and late groupings, as well as the corresponding feature locations, is shown in Figure 5b. The plastic plateau is the 0.995 percentile of the original CDF $y(t)$. The corresponding detected feature points are shown in Figure 5c.

Feature Visualization, Refinement, and Validation

Following the previously mentioned method, features are extracted from the ensemble of simulation and experimental velocity profiles and presented to the user as shown in Figure 1a. The blue

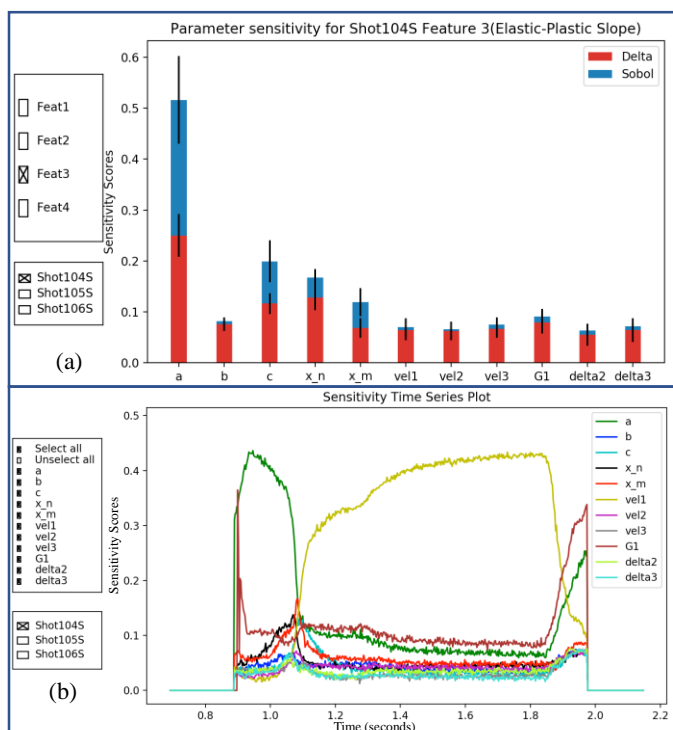


Figure 6. (a) This figure shows the sensitivity information based on the feature 3 (elastic-plastic slope). Here it is shown that the parameter 'a' (strength model parameter from the Johnson-Cook strength model) is the most important in deciding the elastic-plastic slope of the simulations velocimetry curves. (b) This figure shows the time-varying sensitivity information as a generalized version of the plot 6a. The higher sensitivity scores here indicate that the corresponding input variables become important in deciding the velocity values of the velocimetry curves at the varying time steps. This agrees with the results from (a), and also shows that the impact velocity (vel1) has a strong impact on the plastic plateau. Users can interact with plots of (a) and (b) using the toolbars on the left panel for more detailed exploration.

point is the jump off, green is the elastic plateau, the cyan points are the two end points for the elastic-plastic slope calculation, and the yellow point is the plastic plateau. Using our interface, users can choose to observe all these feature locations for different simulations and experiments by clicking buttons and cycling through the simulation runs. Since feature finding has uncertainty, we provide the option for users of this system to refine each feature. Users can zoom into the feature locations, drag the feature points to the desired locations, undo the last drag, and save the features to a file on the disk. Using this input from materials scientists as the ground truth, we compared our proposed feature extraction method with simpler automated methods (such as trying to find the feature locations following the feature definitions) and our proposed method has consistently showed higher accuracy, specifically in handling the irregular shaped curves.

EXPLORATION OF PARAMETER IMPORTANCE

Feature Sensitivity

We would like to understand the importance of each input parameter on the set of features we extract from the simulated velocity profiles. Sensitivity studies are a popular method for quantifying the uncertainty of the output due to each input parameter. Here, we use two popular global sensitivity methods: Delta¹⁰ and Sobol¹¹. For both of these methods, a higher value represents higher importance for the corresponding parameter in deciding the underlying feature. Knowing the sensitivity values, it becomes easier for the scientists to identify and tune the more important parameters, thus saving valuable time.

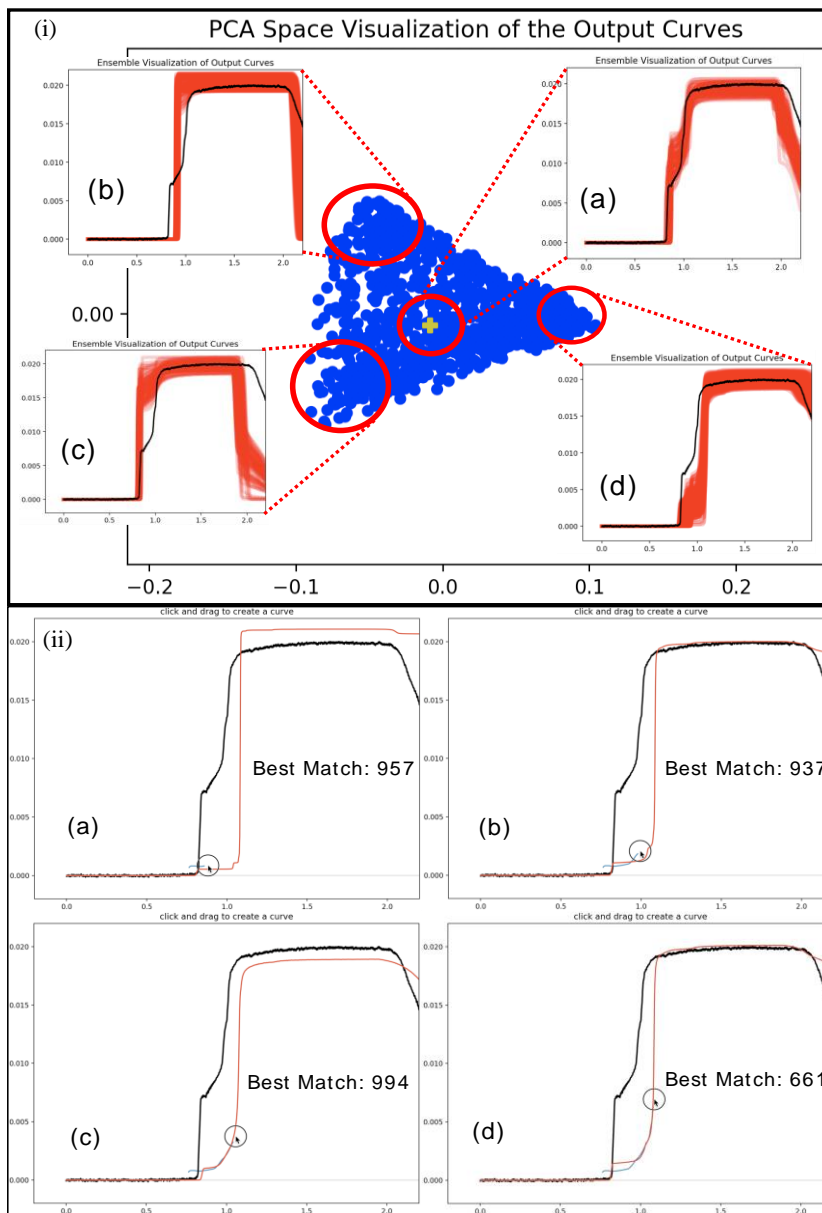


Figure 7. Different exploration and validation methods: (i) PCA space exploration for similarity-based validation and analysis. Using this tool, users can quickly identify which simulated velocimetry curves are similar/dissimilar to the experimental output (physical reference experiment shown as the yellow cross) and find out their corresponding input parameters. (ii) User drawing for querying velocimetry curves with desired features. In this example, the user is looking for the curves with low elastic plateau.

If we assume each time step of the velocity profile is an independent feature, then we can compute the sensitivity of the output due to each parameter at that time. Since analysts have a general idea where each feature occurs in time, this can be thought of as an implicit method to understand which parameters influence a particular feature the most. For each shot, we compute the sensitivity at each feature (Figure 6a) as well as the sensitivity as a function of time (Figure 6b).

For feature-based sensitivity, we have employed the popular stacked bar charts for its widespread popularity and ease of interpretation when comparing multiple quantities. For time-varying sensitivity, we use time-series sensitivity plots. Figure 6a and 6b show examples of this sensitivity analysis using simulations of Shot104S. Our interface allows the user to interactively display this sensitivity information for each shot. This enables the user to comparatively explore the importance of each parameter and across multiple shots, and as can be seen in Figure 6b, the importance of each parameter is time dependent.

Interactive Exploration

Using our visualization tool, users are able to interactively explore the different aspects of parameter importance via the sensitivity study. We provide simple selection tools (as shown in the left panel of Figure 6a and 6b) to facilitate the comparison across multiple shots, features (Figure 6a), and input parameters (Figure 6b) for comparative analysis. Users can choose to explore the time series or the feature-based sensitivity values in order to gain insight into the importance of the input parameters across the different shots and parameters. For example, using toolbars of Figure 6b, users may wish to select the shots 104S and 105S, and the parameters a and v_I to explore their sensitivity across the two different shots. It should be noted that different parameters show varying trends across time for different shots. Simple and intuitive interactions allow materials scientists to thoroughly explore this varying sensitivity information.

COMPARATIVE EXPLORATION AND SYSTEM INTEGRATION

PCA and PCP-based Exploration

PCA¹² is a popular method for high-dimensional data reduction and visualization²³, and it has been widely used for feature extraction, face recognition, variance analysis, and machine learning. If a set of high-dimensional data points are similar or have a small distance in the higher dimension, then this relative distance is often preserved in the lower dimensional PCA-space. We use PCA in our system to enable the users to compare and validate the experimental output with the simulated output in an interactive way.

In our system, each velocity profile $y_i(t)$ is a high-dimensional vector of size M . So, each velocity profile can be thought of as a point in the M -dimensional vector space. For validation purposes, we take all the simulation and the experimental outputs together in a matrix of $(N+k) \times M$ dimensions where there are N simulations and k experiments. Then we apply the PCA-based dimensionality reduction on this data matrix and project them down to 2-dimensions to present the results to the users and facilitate user interaction. In this PCA-based dimensionality reduction, the two principal PCA-dimensions retain more than 95% of the original information as found by computing the PCA_explained_variance and this produces a good quality visualization for comparison purposes.

An example of this is provided in Figure 7i. In this case, we have $N=1000$ simulations (shown by blue dots) and $k=1$ experiment (shown by the yellow plus) with $M=1600$ points on each velocity profile $y_i(t)$. If the users want to select the curves that are similar to the experiment, then using a mouse they can select the points near the yellow plus; as shown in window (a) of Figure 7i, all the curves similar to the experiment can be selected. On further inspection, since proximity in the PCA-space suggests similarity, users can also select the curves that are similar to each other but different from the experiment, e.g. as shown in windows (b), (c) and (d). For example, all the curves with a low elastic plateau are found in window (d), and the combination of the input parameters that generate these curves can be easily highlighted in a parallel coordinates plot (PCP). Similarly, velocity profiles with higher elastic plateau can also be quickly found from this visualization, as shown in windows (b) and (c).

The full system view can be seen in Figure 1. In Figure 1a, users can interactively select their region of interest for the curves. Then, the PCA method can be used to interactively compare the

selected regions of the simulated curves with the experimental curve. Users can move the green vertical lines in the Figure 1a to select a timespan and apply the PCA method to observe the output in the window shown in Figure 1b. When the user selects a set of points using a lasso-based selector in this window (Figure 1b), then those selected points are highlighted in red. After performing this selection, users can click on "Generate Ensemble" to visualize the collection of velocity profiles that correspond to their selected points in Window (d). The input parameter combinations of the selected outputs can be visualized in windows (e) and (f) in Figure 1. Window (e) provides the PCA reduced visualization of the input parameters such that users can explore which subspaces produce the desired velocity profiles. Window (f) is PCP for the corresponding input parameters. In this window, users can have a detailed view of all the parameter values that produced the user selected curves, and they can further interactively select subsets of the different parameter values to explore any hidden relationships.

Exploration Through User Drawing

Apart from using the PCA-based curve selection process and validation, our proposed system allows direct user input via free-hand drawing. Sketch-based image and shape retrieval has gained much popularity in the recent years^{12,13}. In this work, we take up a similar idea to aid the users of our proposed system in finding velocimetry curves with user-defined features from the simulation ensemble output curves. The domain experts often need to query for a certain kind of velocity profile (e.g., low/high elastic/plastic slope). Using a free-hand drawing-based query system allows free-form and easy feature description. This facilitates identification of the expected and the unexpected velocity profiles from the simulations, and understanding the input parameter combinations that generate such velocity profiles.

In window (c) of Figure 1, users can use their finger (on a touch enabled device) or computer mouse to draw the shape of the curve they would like to see in the simulations. At each stroke, our system finds the closest matching curves from the simulations database and displays them overlapped with the users drawing. In order to achieve this functionality, we use k -nearest neighbor (KNN) algorithm to find the best match (i.e., setting $k=1$). A more detailed example of this is shown in Figure 7ii. In the windows (a)-(d) of Figure 7ii, it is shown how the system generated prediction (shown in red) changes in real-time as the user draws more portions of the curve (shown in blue). The run number is also highlighted in the process. To enable more precise feature query, the time-axis and the experimental curve (shown in black) are shown in the drawing window.

To implement drawing-based queries, we apply the Euclidean-distance based similarity scheme for only the timescale where the user has provided the drawn input. We re-train the KNN algorithm at each stroke and find the closest neighbor to the user's input. After the users have drawn the curve of their interest, they can also query for a collection of similar looking simulation curves by setting the slider "Number of Nearest Neighbors". The ensemble of these k matches will be shown in window (d) of Figure 1 in red. In both the windows (c) and (d), the experimental curve is always displayed in black.

EXPERT FEEDBACK

We received a very positive feedback from the users of this tool (the domain experts). According to them, previous attempts to parameterize the shock physics simulations of this study prior to using this new data visualization tool involved individually changing one parameter at a time. This process was time consuming since each change to a parameter involved rerunning the simulation and post-processing the data. This produced a new curve approximately every 10 minutes. This process would be repeated for each tunable parameter until a qualitatively good fit to the experimental data was achieved. The process may take only a few iterations for completely decoupled parameters. In this case, parameters had coupled effects, which required readjustment of previously optimized parameters. For this model, with only 5 tunable parameters, the entire process consumed 10-20 hours of effort.

The new process takes the same amount of time to setup the initial input deck to the simulations. However, hundreds or thousands of simulations with varied input parameters could be run simultaneously. Running

multiple simulations in parallel can take the same time as before, but the real gain is that there is no human interaction required between individual simulations. The old method required constant interaction with the simulations and their outputs. The new tool presents all of the simulation results together in a dynamic, interactive environment. Here, the effects of individual and coupled parameters on the resulting output can be easily explored. Finding the best match of input parameters to a desired output (through PCA-based visualization or using hand-drawn target curves) takes minutes or less. Sensitivities of individual parameters could also be quickly deduced: something that would be difficult or impossible to do using static images of the results. Additionally, parameter optimization would become intractable with more complex models involving perhaps dozens of input variables. Further, the previous manual feature finding endeavors required intimate knowledge of the data and used to cost in the order of minutes for each output curve. Our proposed feature finding method enabled a more precise material property extraction even in the case of the very complicated velocity profiles, taking only order of seconds for a set of thousand curves.

CONCLUSION

In this paper, we presented a workflow for analyzing the shock physics datasets for assisting materials scientists. Over multiple discussions with the scientists, we formulated the requirements of the workflow and designed a system for them. We enabled the extraction of the material properties as the uncertain features from the velocity profiles. We employed the sensitivity methods to compare the importance of the input parameters of shock physics simulations for multiple features as well as a function of time. We incorporated the PCA-based analysis for the validation of the simulations against the corresponding experiments. In our system, we enable the users to query for the features of interest via intuitive drawing using touch-screen enabled devices. We worked closely with materials scientists throughout the project and gathered their feedback regarding our proposed system that shows its efficacy.

REFERENCES

1. MICHALSKI M., RIETH M., KEMPF A., KRÜGER J.: Coflavis: A visualization system for pulverized coal flames. *Computing in Science Engineering* 19, 6 (November 2017), 72–78. doi:10.1109/MCSE.2017.3971156.
2. BOTELER J. M., DANDEKAR D. P.: Dynamic response of two strain-hardened aluminum alloys. *Journal of Applied Physics* 100, 5 (2006), 054902. doi:10.1063/1.2336492.
3. LUSCHER D., ADDESSIO F., CAWKWELL M., RAMOS K.: A dislocation density-based continuum model of the anisotropic shock response of single crystal acyclotrimethylene trinitramine. *Journal of the Mechanics and Physics of Solids* 98 (jan 2017), 63–86. doi:10.1016/j.jmps.2016.09.005.
4. BARKER L. M., HOLLENBACH R. E.: Laser interferometer for measuring high velocities of any reflecting surface. *Journal of Applied Physics* 43, 11 (1972), 4669–4675. doi:10.1063/1.1660986.
5. SALAS M. D.: The curious events leading to the theory of shock waves. In *17th Shock Interaction Symposium* (Rome, Italy, 2006), p. 19. doi:10.1007/s00193-007-0084-z.
6. JOHNSON G. R., COOK W. H.: A Constitutive Model and Data for Metals Subjected to Large Strains, High Strain Rates and High Temperatures. In *Seventh International Symposium on Ballistics* (The Hague, The Netherlands, apr 1983), pp. 541–547.
7. BURTON D. E.: Lagrangian Hydrodynamics in the FLAG Code. In *Symposium on Advanced Numerical Methods for Lagrangian Hydrodynamics* (Los Alamos, NM, USA, LA-UR-07-7547, 2007).
8. J. STRAIT AND S. KURTEK, "Bayesian Model-Based Automatic Landmark Detection for Planar Curves," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, 2016, pp. 1041–1049. doi: 10.1109/CVPRW.2016.134
9. BORGONOVO E.: A new uncertainty importance measure. *Reliability Engineering and System Safety* 92, 6 (2007), 771–784.

10. SOBOL I.: Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation* 55, 1-3 (2001), 271 – 280. doi:[http://dx.doi.org/10.1016/S0378-4754\(00\)00270-6](http://dx.doi.org/10.1016/S0378-4754(00)00270-6).
11. XU Y., ZHANG D., YANG J.-Y.: A feature extraction method for use with bimodal biometrics. *Pattern Recognition* 43, 3 (2010), 1106 – 1115. doi:<https://doi.org/10.1016/j.patcog.2009.09.013>.
12. LEE Y. J., ZITNICK C. L., COHEN M. F.: Shadowdraw: Real-time user guidance for freehand drawing. In *ACM SIGGRAPH 2011 Papers* (New York, NY, USA, 2011), SIGGRAPH '11, ACM, pp. 27:1–27:10. doi:10.1145/1964921.1964922.
13. XIAO C., WANG C., ZHANG L., ZHANG L.: Sketch-based image retrieval via shape words. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval* (New York, NY, USA, 2015), ICMR '15, ACM, pp. 571–574. doi:10.1145/2671188.2749360.
14. SCHROEDER W., MARTIN K. M., LORENSEN W. E.: *The Visualization Toolkit (2Nd Ed.): An Object-oriented Approach to 3D Graphics*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1998.
15. AYACHIT U.: *The ParaView Guide: A Parallel Visualization Application*. Kitware, Inc., USA, 2015.
16. CHILDS H., BRUGGER E., WHITLOCK B., MEREDITH J., AHERN S., PUGMIRE D., BIAGAS K., MILLER M., HARRISON C., WEBER G. H., KRISHNAN H., FOGAL T., SANDERSON A., GARTH C., BETHEL E. W., CAMP D., RÜBEL O., DURANT M., FAVRE J. M., NAVRÁTIL P.: VisIt: An End-User Tool For Visualizing and Analyzing Very Large Data. In *High Performance Visualization—Enabling Extreme-Scale Scientific Insight*. Oct 2012, pp. 357–372.
17. MATLAB: version 9.0 (R2017b). The MathWorks Inc., Natick, Massachusetts, 2017.
18. ROSSUM G.: *Python Reference Manual*. Tech. rep., Amsterdam, The Netherlands, The Netherlands, 1995.
19. RAMACHANDRAN P., VAROQUAUX G.: Mayavi: 3D Visualization of Scientific Data. *Computing in Science & Engineering* 13, 2 (2011), 40–51.
20. STALLING D., WESTERHOFF M., HEGE H.-C.: Amira: a highly interactive system for visual data analysis, 2005.
21. CIBC:, 2016. ImageVis3D: An interactive visualization software system for large-scale volume data. Scientific Computing and Imaging Institute (SCI), Download from: <http://www.imagevis3d.org>.
22. CIBC:, 2016. Seg3D: Volumetric Image Segmentation and Visualization. Scientific Computing and Imaging Institute (SCI), Download from: <http://www.seg3d.org>.
23. JEONG, D. H., ZIEMKIEWICZ, C. , FISHER, B. , RIBARSKY, W. AND CHANG, R. (2009), iPCA: An Interactive System for PCA-based Visual Analytics. *Computer Graphics Forum*, 28: 767-774. doi:10.1111/j.1467-8659.2009.01475.x
24. Z. KONYHA, K. MATKOVIC, D. GRACANIN, M. JELOVIC AND H. HAUSER, "Interactive Visual Analysis of Families of Function Graphs," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 6, pp. 1373-1385, Nov.-Dec. 2006. doi: 10.1109/TVCG.2006.99
25. MUIGG, P. , KEHRER, J. , OELTZE, S. , PIRINGER, H. , DOLEISCH, H. , PREIM, B. AND HAUSER, H. (2008), A Four-level Focus+Context Approach to Interactive Visual Analysis of Temporal Features in Large Scientific Data. *Computer Graphics Forum*, 27: 775-782. doi:10.1111/j.1467-8659.2008.01207.x

ABOUT THE AUTHORS

Ayan Biswas is a scientist in the Data Science group (CCS-7) at Los Alamos National Laboratory. His research focuses on uncertainty quantification, statistical analysis and high-dimensional data visualization. He received his PhD in Data Visualization from The Ohio State University. Contact him at ayan@lanl.gov.

Christopher Biwer is a postdoc at Los Alamos National Laboratory. His research interests include data analysis and signal processing. Biwer received a PhD in physics from Syracuse University. Contact him at cmbiwer@lanl.gov.

David J. Walters is a postdoc research associate in the Theoretical Division at Los Alamos National Laboratory. His research focuses on microstructural modeling and simulation in addition to modeling shocked materials. He received his PhD in Applied Mechanics from Montana State. He is a member of the American Physical Society. He can be reached at djwalters@lanl.gov.

James Ahrens is a senior scientist at Los Alamos National Laboratory. His research interests include large scale data analysis and visualizations. He received his PhD from University of Washington. Contact him at ahrens@lanl.gov.

Devin Francom is a scientist in the Statistical Sciences group at Los Alamos National Laboratory. His research interests include uncertainty quantification and Bayesian analysis. He received a PhD in statistics and applied mathematics from the University of California Santa Cruz. He is a member of the American Statistical Association. Contact him at dfrancom@lanl.gov.

Earl Lawrence is a member of the Statistical Sciences group at Los Alamos National Laboratory. He likes learning things from data and has done so in numerous application areas including cosmology, space weather, power grids, and materials science. Lawrence received a PhD in statistics from the University of Michigan. Contact him at earl@lanl.gov.

Richard Sandberg is a staff member at Los Alamos National Laboratory. He received his PhD from the University of Colorado at Boulder with an emphasis in optics. His research focuses on X-ray coherent diffraction imaging and studying materials dynamics at the nanometer scale with X-rays light sources. Contact him at sandberg@lanl.gov.

D. Anthony Fredenburg is a research scientist at Los Alamos National Laboratory, specializing in materials properties under high strain rate loading conditions. Fredenburg received his PhD from the Georgia Institute of Technology, and is an active member of the American Physical Society. He can be reached at dafreden@lanl.gov.

Cynthia Bolme is a research scientist in the Shock and Detonation Physics group at Los Alamos National Laboratory. She received her PhD in Physical Chemistry from the Massachusetts Institute of Technology. Her research focuses on material dynamics with an emphasis in ultrafast laser and X-ray diagnostics. She can be reached at cbolme@lanl.gov.