

Visualization of Uncertainty for Computationally Intensive Simulations Using High Fidelity Emulators

Ayan Biswas, Kelly R. Moran, Earl Lawrence, and James Ahrens, *Member, IEEE*

Abstract— Visualization of high-fidelity scientific simulations with high-dimensional inputs and outputs is an important task. Existing high-dimensional data visualization approaches generally assume a substantial amount of data are available or can be generated as needed. However, many of these simulations can be very computationally intensive, taking minutes or hours to run. Analysis and visualization of such expensive simulations poses a challenge. Statistical emulators are frequently used to approximate simulations for statistical analyses. In this work, we propose a visualization tool for an emulator of the simulator and describe how emulators can be used to create effective visualization systems. We choose Gaussian process emulators for this purpose as they enable fast and accurate prediction with uncertainty information. Using these predictions, we design a system that enables visualization of high-dimensional input and output spaces of complex physics simulations. Users of our system can get a detailed understanding of the uncertainties associated with the emulator predictions in both input and output space for a high-dimensional simulation.

Index Terms—Uncertainty, high-dimensional data, fast prediction, emulation

1 INTRODUCTION

Understanding properties of non-linear simulations from a limited set of high-dimensional input-output samples is a non-trivial problem. Despite recent advances in supercomputers, these simulations can still be very time consuming and it is often impractical to generate the desired number of runs. Domain experts from different sciences (e.g., astrophysics, cosmology, shock physics) frequently use data from these simulations and apply the knowledge gained from such sparse samples to their experiments (e.g., SLAC National Accelerator Laboratory for shock physics experiments). When new experiments deviate from the existing sparse simulation samples, it is often necessary to run simulations at new inputs to explore the *holes* in the input spaces. Because of the time expense, this generally cannot be done during the experimental window.

Emulators, i.e. statistical approximations to simulators, have emerged as an alternative to the existing work-flow. Emulators can quickly and accurately predict simulation output at untried inputs. Using this tool, domain experts can see which regions of input space produce output close to the experiment in near real-time. This makes the work-flow faster and facilitates new research and advancement. Among other choices, Gaussian process (GP) based emulators have gained popularity due to their fast and accurate prediction capabilities.

The role of visualization in such a data analysis pipeline is very important. After creating a high fidelity emulator, there are two tasks that warrant specialized visualization tools. First, a visualization tool could provide guidance about new simulations to run based on *importance* criteria. Second, visualization could be used to understand emulator prediction uncertainty. Simple yet descriptive visualizations can provide the domain scientists with clear knowledge of the uncertainties of the system and how much to trust the emulator.

In this work, we address the visualization of shock physics experiments using a Gaussian process-based emulator. Using this robust

emulator, we propose a visualization tool that provides detailed uncertainty quantification and visualization solutions for exploring the dataset and guiding the simulation scientists to new input parameters for next batch of simulation runs. Beyond this particular application, our proposed solutions should be applicable to other high-dimensional dataset that have similar data processing and visualization needs.

Our contributions in this work are following:

- We propose an interactive exploration tool for understanding how the input parameters are associated with prediction uncertainty of an emulator of a computationally extensive high-dimensional simulator
- We create a visualization system for understanding the GP-based emulation uncertainties in the high-dimensional output space

2 RELATED WORKS

Our proposed approach is closely related to the areas of uncertainty quantification and visualization, high-dimensional data visualization, and software emulation. We provide a brief summary and direct readers to some notable members of the the vast body of existing research.

Uncertainty and Ensemble Visualization: In the past decade, uncertainty analysis and visualization has become a very popular topic of research and ensemble datasets, one of the primary sources of uncertain datasets, are frequently used for creating new visualization techniques. The initial methods of uncertainty visualization were thoroughly discussed by Pang et al. [29]. Since then, researchers have proposed many techniques on this topic. We refer readers to the survey works by Potter et al. [30], Brodlie et al. [4], and Bonneau et al. [2] for an extensive survey of the recent uncertainty visualization methods.

High-dimensional Data Visualization: Visualization of high-dimensional datasets is a non-trivial research area with an active history. Notably, methods like PCA [19] and MDS [9] are frequently used for reducing the dimensionality of the data and projecting it into a more convenient space for user viewing and interaction. Traditional scatter plots may be augmented to scatter plot matrices [8] to provide a high-dimensional view of the data in which relationships could be explored. Interactive Parallel Coordinate Plots (PCP) [17], a popular method for high-dimensional data exploration, are currently included in many visualization tools. Of comparable popularity is the star-glyphs [34] method. Researchers have also looked into creating continuous representations of discrete scatter plots [1, 23] and parallel coordinates [24] for enhanced perceptual effectiveness. Along similar lines, in this work we propose to create continuous representations of the sparsely sampled high-dimensional input space for visualization and user interaction.

• Ayan Biswas is with Los Alamos National Laboratory.
E-mail: ayan@lanl.gov

• Kelly R. Moran is with Los Alamos National Laboratory.
E-mail: krmoran@lanl.gov

• Earl Lawrence is with Los Alamos National Laboratory.
E-mail: earl@lanl.gov

• James Ahrens is with Los Alamos National Laboratory.
E-mail: ahrens@lanl.gov

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

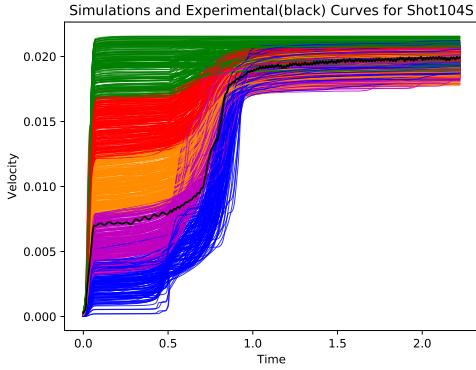


Fig. 1. An ensemble of velocity profiles originating from FLAG [6] shock physics simulations. The colors of the simulation generated curves are chosen based on k-means clustering of these output curves to show the trends. The corresponding experiment is shown in black.

Software Emulation and Gaussian Processes: Statistics has long been interested in approximating computationally intensive computer simulations for inference. McKay et al. [26] introduced the Latin hypercube and present a comparison of approaches for choosing the best inputs in order to analyze the output of a computer code. Sacks et al. [31] developed the concept of using a Gaussian process to approximate an intensive simulation. Jones et al. [20] use a GP to approximate an expensive black box function and sequential design to maximize the function. Kennedy and O'Hagan [21] introduce a Bayesian approach that approximates the expensive computer code using GPs, accounts for systematic discrepancies between the code and reality, and estimates the distribution of inputs that aligns simulations with data from physical experiments. Higdon et al. [14] present a Bayesian approach suitable for multivariate simulator output; this is the emulation method used in our work. Gaussian process emulators have been successfully used in a number of application areas, including cosmology [22], satellite navigation [27], and accelerator modeling [15]. Spline approaches (e.g. Francom et al. [11]) and filtering methods (e.g. [13]) have also been used for emulation. GPs have also recently been applied in the visualization community [32, 33].

3 PROBLEM OVERVIEW

This work is motivated by an application to shock physics. Materials scientists study the different deformation properties of a given material by sending a shock wave through it. The output of these experiments [3, 25, e.g.] and simulations [5–7, e.g.] is a time-series of velocity values, called a *velocity profile* or *velocimetry curve*. Examples of these curves are shown in Figure 1 (simulations shown in multiple colors, experiment in black) where wrought aluminum alloy Al-5083 was used to impact another Al-5083 metal block. The simulations, based on the FLAG [5, 6] hydrodynamic code, produce velocity profiles for an ensemble of input combinations. FLAG allows the direct manipulation of the five Johnson-Cook strength model [18] parameters (listed as A, B, C, X_n , and X_m), three shear modulus parameters (listed as G_1, δ_1 and δ_2), and three more input parameters, that describe the experimental impact velocities of the material that generate the shock wave (listed as v_1, v_2 and v_3); thus exposing to users a total of 11 tunable parameters. Using this accurate but time-consuming simulation, it takes hours to generate 1000 velocity profiles on powerful supercomputers (e.g., Moonlight supercomputing environment with 308 Intel Xeon E5-267 nodes and 4928 CPU cores).

Given this simulation, we propose the use of GP-based emulators to create a visualization system that will help users:

1. Explore uncertainties in the functional output associated with regions of the high-dimensional input parameter space.
2. Effectively visualize uncertainties related to the emulation.
3. Select input parameter combinations for the next simulation run.

4 EMULATION WITH GAUSSIAN SPATIAL PROCESSES

Gaussian processes [35] are a popular choice for emulators [12, 14–16, 22]. Our emulation strategy follows the one laid out in [14] for multivariate simulation output. A vector of simulation outputs y (here, velocity values at selected time stamps) is modeled as

$$y = \mu + \delta K w. \quad (1)$$

The vector μ and scalar δ are an overall mean and scaling parameter, respectively. The matrix K is the set of orthogonal basis functions, obtained via singular value decomposition (SVD) of the standardized simulation output matrix. Each entry in the vector w represents the weight applied to the corresponding basis vector for the particular simulation output y . Different weights applied to the bases yield different shaped velocimetry curves. The weights for each basis vector are modeled with a Gaussian process independently of the weights for other basis vectors. Each GP treats observed input-output pairs as observations from a multivariate Gaussian distribution. Conditioning operations give a predicted mean and variance for the weights at new unobserved inputs given the observed simulation runs [35]. Gaussian processes have a number of hyperparameters that require estimation; we follow the Bayesian estimation approach in [14].

4.1 Uncertainty Quantification

An important component of our emulation approach is the ability to provide uncertainty quantification; although the simulation itself is deterministic, the output at new inputs is not known. Uncertainty appears in the Bayesian GP approach via two sources:

4.1.1 Distributional Uncertainty

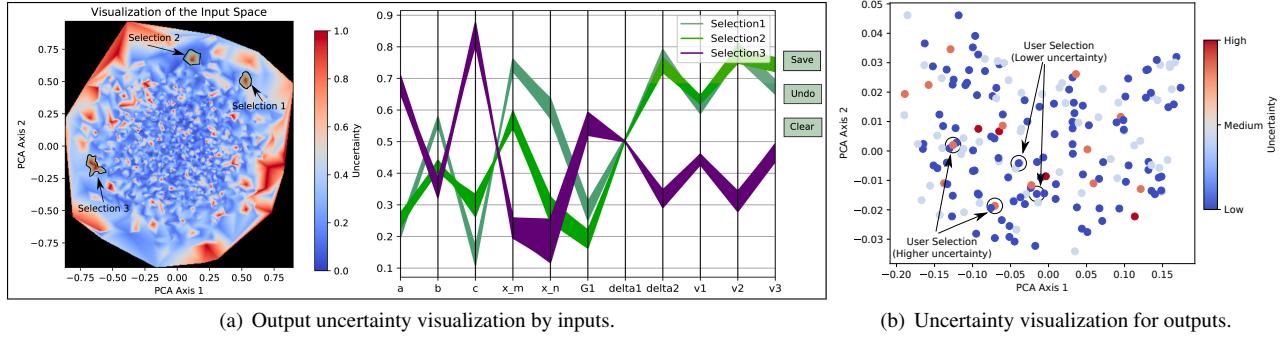
Assume we have fixed the GP hyperparameters (we use the posterior mean, but other choices are defensible). By modeling each component of w as independent zero-mean Gaussian process we are able to obtain a closed form solution for the distribution of each basis weight at new input points via conditional normal theory as described in [35]. The conditional (on the simulation runs) mean of w for a particular basis vector at a new input point is a weighted average of the w for that basis vector at simulation points, with weights depending on the GP hyperparameters and the distance between the new input point and each simulation input point. The conditional variance also depends on both the GP hyperparameters and these distances. Predictions have lower uncertainty at new inputs that are close to observed simulations. This shrinks to zero as a new input approaches an observed simulation.

We can describe the analytic form of the distribution of y at new inputs. Let M_w (V_w) be the vector (diagonal matrix) containing the conditional means (variances) for each basis weight. Based on Gaussian theory and Equation 1, the mean of the output is $\delta K M_w$. The covariance matrix for a whole output vector, i.e. the covariance matrix relating output points over time at a given (shared) input, is given by $\delta^2 K V_w K'$. If the length of w is less than the length of y , as is the generally the case so as to represent y using a reduced set of bases for computational purposes, this matrix will be degenerate. However, at a given time index we can still generate samples and/or compute desired quantiles using this distribution.

Using the GP-based emulator, we can offer scientists multiple visualizations of the distributional uncertainty inherent in the model. As shown in Figure 3(b) and 3(e), we can get samples (curves shown in yellow) instead of only the mean prediction (shown in red). Also, using the covariance matrix, we can get a closed form estimate of the uncertainty along the curve as shown in Figure 3(a) and 3(d). Finally, for each prediction, a scalar estimate of uncertainty can be produced using this same covariance matrix by taking the sum of the diagonal entries as shown in Figure 2 (more details regarding the visualization are given in next section).

4.1.2 Parameter Uncertainty

Fixing the GP hyperparameters, as described above, ignores hyperparameter uncertainty in the model. A subtly different set of hyperparameters implies a GP model having slightly different behavior. To



(a) Output uncertainty visualization by inputs.

(b) Uncertainty visualization for outputs.

Fig. 2. a) Visualization of output uncertainty by input region. The left panel shows a continuous representation of the input space where the color represents the scalar summary of uncertainty of the emulator output (see Section 4.1.1). Red (blue) represents a high (low) level of uncertainty about the shape of the predicted curve at that region of the input space. The right panel is a linked PCP view of the input parameter combinations as selected by the user on the left image. b) Exploration of output curves in the PCA space for detection of curves having higher/lower uncertainty. The user's selections are highlighted and used for further exploration.

incorporate both distributional and hyperparameter uncertainty, we can use samples of the hyperparameters produced during GP training. For each sample, a mean and variance for each basis weight are available in closed form as described above. Thus, for a given sample we can take a single draw of each of the weights and use them to construct the output function according to Equation 1. The resulting set of samples of y incorporates both hyperparameter and distributional uncertainty and can be used to compute desired posterior quantities. This approach comes with more computational expense, as matrix quantities that depend on the GP hyperparameters must be computed for each sample (rather than once, as when the GP hyperparameters are fixed).

5 VISUALIZATION SYSTEM AND RESULTS

We develop a visualization system that uses the fast predictions and uncertainty calculations from our GP-based emulator to provide guidance to the users for effective data exploration.

5.1 Visualization of Input Space

As mentioned in Section 3, two tasks important to users are exploring uncertainties associated with the high-dimensional input parameter space and deciding at what region of the input space to next run the simulator. To allow users to perform these tasks, in our designed visualization system we allow for an interactive exploration of the input space. In this exploration mode, starting from the set of available sparse data samples we first compute a continuous representation of the input space (e.g., Figure 2(a)). Since the emulator makes fast predictions of the output and provides the uncertainty about those predictions, we can now use the emulator to sample the input space much more densely and compute the uncertainty at all these points. In our shock physics example, we made predictions at between 20,000 and 50,000 input parameter combinations depending on the available time.

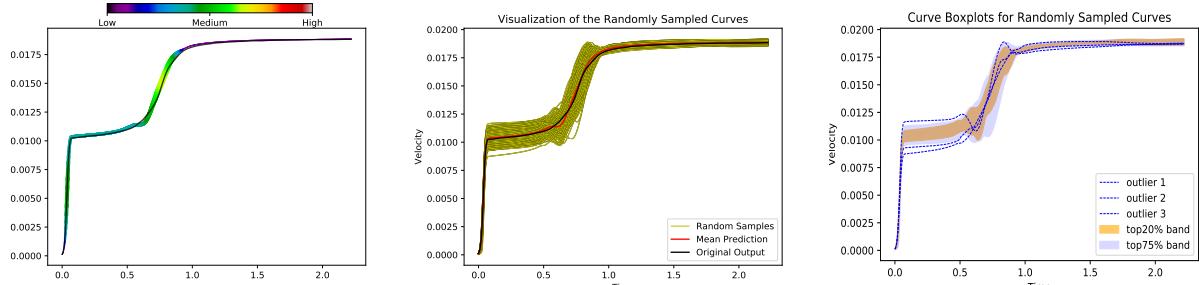
To facilitate user interaction and exploration of this high-dimensional input space, we use PCA-based dimensionality reduction on these new input parameter combinations and project them down to the 2D plane (note that pairs of inputs could also be used). PCA-based methods have the advantage of fast computation and are an oft-used dimension reduction technique. In the reduced PCA space, we map uncertainty (the closed form summary uncertainty of an emulator-predicted output corresponding to an input, computed as discussed in Section 4.1.1) to colors for each sampled location and create a continuous high-resolution 2D image. Using a suitable colormap, users can visually explore the regions of the input space that have high or low uncertainty due to their previously chosen sparse number of simulator runs. They can also interact with this image by selecting regions of interest (e.g., hot-spots showing pockets of high uncertainty) using a lasso-based selection tool, and get a Parallel Coordinate Plot (PCP) view of the high-dimensional input combinations that correspond to the user's selection.

An example of this is shown in Figure 2(a). In this figure, users have selected three regions of the input space with high output uncertainty; the corresponding PCP view reveals the selected input parameter combinations. Using our tool, users can save these selections in a text file (e.g., for use in running future simulations), clear the current selections, or undo the last selection. Although using exploration of input space remains the primary goal of this visualization technique, the same tool could be applied for the output space as well with some modification. Due to properties of the emulator, the input space near the boundary in Figure 2(a) is expected to have higher output uncertainty. As can be seen in this figure, there are regions of high uncertainty even towards the center of the input space. These high uncertainty regions indicate where users may want to generate more simulator runs.

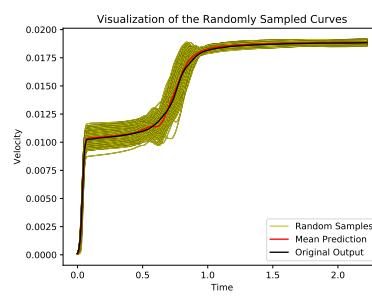
5.2 Visualization of Prediction Uncertainty

The other important task is to explore the uncertainty in the emulator prediction; how certain are we of the velocity profiles for a new input? To address this, we treat each predicted output curve (i.e., the mean prediction from the emulator) as a high-dimensional vector and use PCA-based dimensionality reduction (similar to the idea presented in streamline variability plots [10]) to present the results to the users in a 2D scatter plot view. In this view, each scatter plot point represents the projection of the mean predicted output curve to the 2D PCA-space, and similarity of the curves are represented by their proximity in this lower dimensional space. For uncertainty exploration, the color (and/or size) of these projected points are modulated according to the total uncertainty estimate as quantified by the GP emulator as discussed in Section 4.1.1. Examples of such PCA-space output exploration is shown in Figure 2(b). In this case, we have chosen to show around 200 data points that were held out from the available 1000 data samples. These hold-out samples allow us to compare the emulator output to the simulation to assess accuracy. As users working in the PCA-space may find it less intuitive than working with the original velocimetry curves, curve selection in the PCA-space is followed by visualization in the original high-dimensional output space.

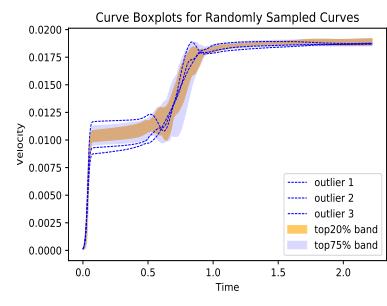
After the users have selected one or more points from the PCA space, we provide two different forms of uncertainty visualization in our proposed system that is linked with the previous PCA window. First, our emulator can provide the estimate of uncertainty at each time point of the predicted velocimetry curve (as mentioned in Section 4.1), and we use that information to modulate the color and thickness of mean predicted curves for easy identification of their uncertain segments. Second, using our emulator, we generate a collection of randomly sampled curves for each input which are used to provide a visual understanding of the spread of the uncertainty in the prediction. For visualizing such ensemble curves, multiple techniques already exist in the visualization literature. We chose to use spaghetti plots and curve boxplots [28] in this work. As an initial visualization, spaghetti



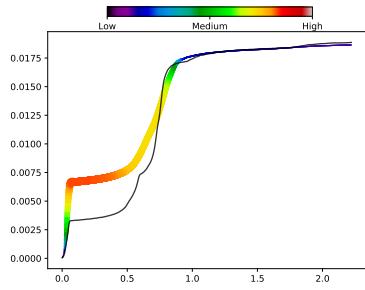
(a) Demonstration of prediction with good accuracy for run id 20.



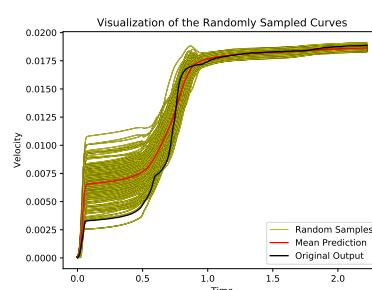
(b) Visualization of the all the mcmc sampled curves with highlighted boundary.



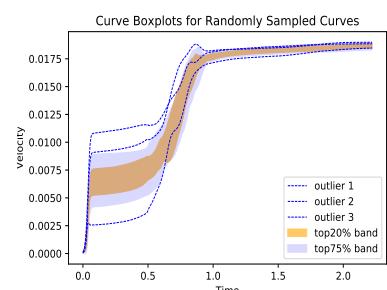
(c) Visualization of the uncertainty using curve boxplots.



(d) Demonstration of prediction with high uncertainty for run id 22.



(e) Visualization of the all the mcmc sampled curves with highlighted boundary.



(f) Visualization of the uncertainty using curve boxplots.

Fig. 3. Example of uncertainty visualization on the predicted curves. Here, the black curves are the simulation generated curves (ground truth) and the predicted curves are color-mapped to their prediction uncertainty. It is visible that for the cases shown in figure a-c, the prediction has higher accuracy and the GP-emu returns low uncertainty for the different regions of the predicted curves. Compared to this, the case shown in figure d-e, they consist of high prediction uncertainty in different regions of the predicted curves and the GP emulator accurately highlights this.

plots provide a good visualization of the randomness and spread of the curves. We use curve boxplots to provide a concise visualization of the spreads, spatial bands, and outliers in the set of random velocity profiles. To show trends in these ensemble profiles, the clustering-based approach as presented in streamline variability plots [10] could also be used.

Examples of the prediction uncertainty exploration are provided in Figure 3. To generate these results, out of the available 1000 simulations, we used 800 as the training set for the emulator. We used the remaining 200 simulations for prediction and uncertainty analysis. The number of basis components was set to 15 for an explained variance of around 99.99%. To represent a predicted curve at a given set of input points we generated 150 samples from the emulator; these samples are shown visually in the spaghetti plots and were used to create the curve boxplots. We give users options on what type of confidence band and how many outliers to show in the curve boxplots. The default was set to two bands showing the top 20% and 75% and three outliers.

In Figure 3, we have chosen to show two different user selections from the previously mentioned PCA spaces as examples. Please note that here we show the simulated curves along with the emulator predictions for illustrative purposes, but in most real uses these ground truths will not be available. In Figure 3, the first row represents a case where the emulator prediction has low uncertainty. Figure 3(a) shows the color-mapped segments for run id 20 of the test set. Using the color-mapped curve segments, it is easy to visualize that the predicted curve has low prediction uncertainty in most of its regions. Comparing it with the corresponding simulated output (shown in black), we can see that indeed the low uncertainty amounted to low error in prediction. The corresponding spaghetti plot is presented in Figure 3(b) that shows the general spread of the random MCMC curve samples. The curve boxplot for these random curve samples are shown in Figure 3(c). The prediction uncertainty can be well estimated from these visualizations. The next set of results, presented in Figure 3(d)-3(f), show a case where the emulator exhibits high uncertainty for predicting run id 22. In this

case, the emulation result was not very accurate, but the emulator was able to highlight the regions of high uncertainty successfully. This is a very desirable feature for our domain experts.

6 CONCLUSION AND FUTURE WORK

In this work, we have presented a visualization system that facilitates the analysis and exploration of high-dimensional simulations. Taking shock physics datasets as our example application, we addressed the issues of uncertainty quantification and visualization in the case of emulation of high-dimensional simulation data. Given only sparse data samples, we proposed the use of a Gaussian processes-based Bayesian model for creating reliable emulators. We investigated the different sources of uncertainty in this emulator and used different visualization methods to provide the users with a detailed understanding of their datasets and emulation process. Using our system, users can explore regions of the input space associated with high output prediction uncertainty, select regions of interest for future simulator runs to reduce the uncertainty in prediction, identify the emulator predictions which have high/low prediction uncertainty, and explore the prediction uncertainty for each individual predicted curve.

In the future, we would like extend this system such that it can handle 2D and 3D scientific datasets so we can apply it in other existing applications. We would also like to apply this system to analyze other popular visualization methods, e.g., streamlines, pathlines, isocontours etc. Finally, we would like to handle multi-resolution, time-varying ensemble datasets.

ACKNOWLEDGMENTS

The authors wish to thank Cynthia Bolme and Richard Sandberg from Los Alamos National Laboratory for the dataset and their valuable comments on the proposed system. This work was supported in part by a grant from Los Alamos National Laboratory LDRD (# 20170029DR).

REFERENCES

- [1] S. Bachthaler and D. Weiskopf. Continuous scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1428–1435, Nov 2008. doi: 10.1109/TVCG.2008.119
- [2] G.-P. Bonneau, H.-C. Hege, C. R. Johnson, M. M. Oliveira, K. Potter, P. Rheingans, and T. Schultz. Overview and state-of-the-art of uncertainty visualization. In *Scientific Visualization: Uncertainty, Multifield, Biomedical, and Scalable Visualization*, Mathematics and Visualization, pp. 1–25. Springer, 2014.
- [3] J. M. Boteler and D. P. Dandekar. Dynamic response of two strain-hardened aluminum alloys. *Journal of Applied Physics*, 100(5):054902, 2006. doi: 10.1063/1.2336492
- [4] K. Brodlie, R. A. Osorio, and A. Lopes. A review of uncertainty in data visualization. In D. K. J. V. John Dill, Rae Earnshaw and P. C. Wong, eds., *Expanding the Frontiers of Visual Analytics and Visualization*, pp. 81–109. Springer Verlag London, 2012.
- [5] D. E. Burton. FLAG, a Multi-Dimensional, Multiple Mesh, Adaptive Free-Lagrange, Hydrodynamics Code. In *Nuclear Explosives Code Developers Conference*, 1992.
- [6] D. E. Burton. Lagrangian Hydrodynamics in the FLAG Code. In *Symposium on Advanced Numerical Methods for Lagrangian Hydrodynamics*. Los Alamos, NM, USA, LA-UR-07-7547, 2007.
- [7] E. J. Caramana, D. E. Burton, M. J. Shashkov, and P. P. Whalen. The Construction of Compatible Hydrodynamics Algorithms Utilizing Conservation of Total Energy. *Journal of Computational Physics*, 146:227–262, 1998.
- [8] W. Cleveland. The elements of graphing data. *Nature*, 318(6045):417–417, 2017.
- [9] T. F. Cox and M. A. A. Cox. *Multidimensional Scaling*. Chapman & Hall, London, 2001.
- [10] F. Ferstl, K. Brger, and R. Westermann. Streamline variability plots for characterizing the uncertainty in vector field ensembles. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):767–776, Jan 2016.
- [11] D. Francom, B. Sansó, A. Kupresanin, and G. Johannesson. Sensitivity analysis and emulation for functional data using bayesian adaptive splines. *Statistica Sinica*, 2016.
- [12] K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon. The coyote universe. i. precision determination of the nonlinear matter power spectrum. *The Astrophysical Journal*, 715(1):104, 2010.
- [13] D. Higdon, J. Gattiker, E. Lawrence, C. Jackson, M. Tobis, M. Pratola, S. Habib, K. Heitmann, and S. Price. Computer model calibration using the ensemble kalman filter. *Technometrics*, 55(4):488–500, 2013.
- [14] D. Higdon, J. Gattiker, B. Williams, and M. Rightley. Computer model calibration using high-dimensional output. *Journal of the American Statistical Association*, 103(482):570–583, 2008.
- [15] D. Higdon, M. Kennedy, J. C. Cavendish, J. A. Cafeo, and R. D. Ryne. Combining field data and computer simulations for calibration and prediction. *SIAM Journal on Scientific Computing*, 26(2):448–466, 2004.
- [16] D. Higdon, E. Lawrence, K. Heitmann, and S. Habib. Simulation-aided inference in cosmology. In *Statistical Challenges in Modern Astronomy V*, pp. 41–57. Springer, 2012.
- [17] A. Inselberg. The Plane with Parallel Coordinates. *The Visual Computer*, 1(2):69–91, Aug. 1985. doi: 10.1007/bf01898350
- [18] G. R. Johnson and W. H. Cook. A Constitutive Model and Data for Metals Subjected to Large Strains, High Strain Rates and High Temperatures. In *Seventh International Symposium on Ballistics*, pp. 541–547. The Hague, The Netherlands, apr 1983.
- [19] R. A. Johnson and D. Wichern. *Multivariate analysis*. Wiley Online Library, 2002.
- [20] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998.
- [21] M. C. Kennedy and A. O’Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.
- [22] E. Lawrence, K. Heitmann, J. Kwan, A. Upadhye, D. Bingham, S. Habib, D. Higdon, A. Pope, H. Finkel, and N. Frontiere. The mira-titan universe. ii. matter power spectrum emulation. *The Astrophysical Journal*, 847(1):50, 2017.
- [23] D. J. Lehmann and H. Theisel. Discontinuities in continuous scatter plots. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1291–1300, Nov 2010. doi: 10.1109/TVCG.2010.146
- [24] D. J. Lehmann and H. Theisel. Features in continuous parallel coordinates. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1912–1921, Dec 2011. doi: 10.1109/TVCG.2011.200
- [25] D. Luscher, F. Addessio, M. Cawkwell, and K. Ramos. A dislocation density-based continuum model of the anisotropic shock response of single crystal α -cyclotrimethylene trinitramine. *Journal of the Mechanics and Physics of Solids*, 98:63–86, jan 2017. doi: 10.1016/j.jmps.2016.09.005
- [26] M. D. McKay, R. J. Beckman, and W. J. Conover. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.
- [27] P. M. Mehta, A. Walker, E. Lawrence, R. Linares, D. Higdon, and J. Koller. Modeling satellite drag coefficients with response surfaces. *Advances in Space Research*, 54(8):1590–1607, 2014.
- [28] M. Mirzargar, R. T. Whitaker, and R. M. Kirby. Curve boxplot: Generalization of boxplot for ensembles of curves. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2654–2663, Dec 2014.
- [29] T. A. Pang, M. C. Wittenbrink, and K. S. Lodha. Approaches to uncertainty visualization. *The Visual Computer*, 13(8):370–390, 1996. doi: 10.1007/s00310050111
- [30] K. Potter, P. Rosen, and C. R. Johnson. *Uncertainty Quantification in Scientific Computing: 10th IFIP WG 2.5 Working Conference*, chap. From Quantification to Visualization: A Taxonomy of Uncertainty Visualization Approaches, pp. 226–249. Springer Berlin Heidelberg, 2012.
- [31] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn. Design and analysis of computer experiments. *Statistical science*, pp. 409–423, 1989.
- [32] S. Schlegel, N. Korn, and G. Scheuermann. On the interpolation of data with normally distributed uncertainty for visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2305–2314, 2012.
- [33] T. Torsney-Weir, A. Saad, T. Moller, H.-C. Hege, B. Weber, J.-M. Verbavatz, and S. Bergner. Tuner: Principled parameter finding for image segmentation algorithms using visual response surface exploration. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1892–1901, 2011.
- [34] M. O. Ward. Xmdvtool: Integrating multiple methods for visualizing multivariate data. In *Proceedings of the Conference on Visualization ’94*, VIS ’94, pp. 326–333. IEEE Computer Society Press, Los Alamitos, CA, USA, 1994.
- [35] C. K. Williams and C. E. Rasmussen. Gaussian processes for machine learning. *the MIT Press*, 2(3):4, 2006.