Making Sense of Scientific Simulation Ensembles

Mai Dahshan*
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Nicholas Polys†
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

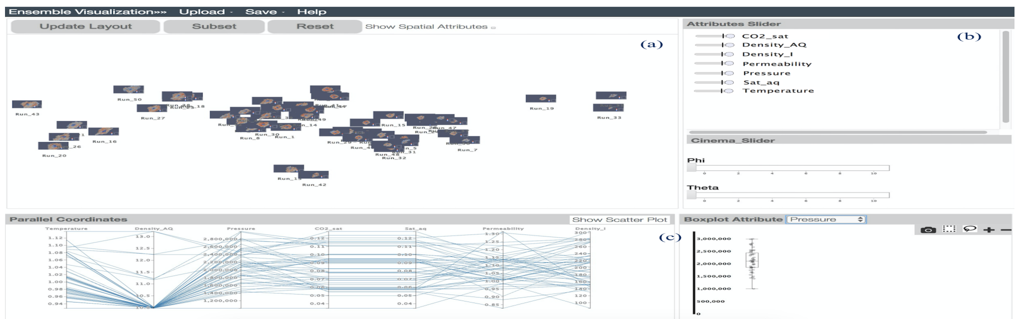

Figure 1: The main interface of GLEE for ensemble simulation analysis (a) Ensemble View: each image represents an ensemble member laid out spatially using WMDS (b) Parameter View: shows weight slider for ensemble inputs and outputs (c) Statistical View: displays statistical properties and distributions of ensemble using different graphs.

ABSTRACT

Scientists run many simulations with varying initial conditions, known as 'ensembles', to understand the influence and relationships among multiple parameters or ensemble members. Most of the ensemble visualization and analysis approaches and techniques focus on analyzing the relationships between either the ensemble members or output parameter space while neglecting the effect of input parameters and human in the analysis loop. Therefore, we developed an approach to the visual analysis of scientific data that merges human expertise and intuition with machine learning and statistics allowing scientists to explore, search, filter, and make sense of their high dimensional ensemble. Our tool, 'GLEE '(Graphically-Linked Ensemble Explorer), is an interactive visualization tool that consists of three visual views: Ensemble View, Parameter View, and Statistical View. Each view offers different functionality for exploration and interoperation of the relations and correlations between different runs, a subset of runs, and input and output parameters.

KEYWORDS

Ensembles, Sensemaking, Scientific Visualization

1 INTRODUCTION

The recent advancement in computing power and the availability of high-performance computing has led to the feasibility of running complex real-world simulations for large sets of parameters in an acceptable amount of time. Scientists usually run simulation ensembles using different input conditions, simulation parameters, and simulation models to interpret the distributions within the data, investigate the sensitivity of outputs to certain input parameters, and understand the similarities and dissimilarities between ensemble members. Therefore, the visual analysis of ensembles is a challenging task due to the high multi-dimensional complexity, and the size of data.

Current research in the visual analysis of ensembles relies on multiple techniques for showing the variability of the ensemble members, major trends, and outliers. Some of these techniques focus on statistically studying the parameter space and measuring the correlation between different parameters [3, 4]. Alternatively, other techniques study the shape and variability of the ensembles themselves [1, 2]. Most of these ensemble visualization techniques are tailored for specific application areas (i.e., specifically weather and climate).

The motivation behind our visual analysis tool is to help scientists make sense of ensemble data through the interactive exploration process regarding both the parameter space and the ensemble runs. Our contributions are to: 1) create a visual analysis approach that helps scientists make-sense of ensembles behaviors, patterns,

*mdahshan@vt.edu
†npolys@vt.edu

©

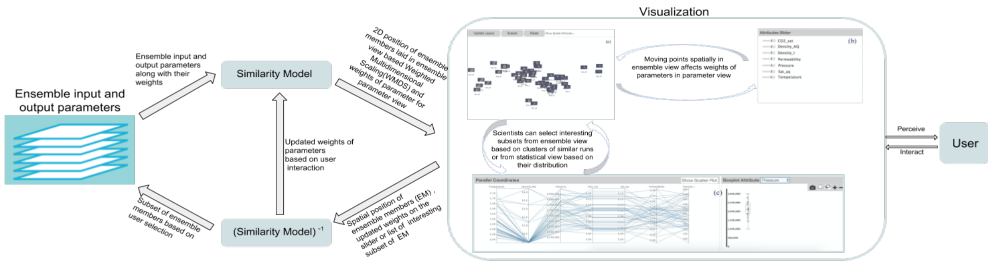

Figure 2: Pipeline of our comparative visualization tool

and outliers for high dimensional data using inputs as well as outputs; 2) couple interactive visual analysis and statistical summaries to detect and analyze characteristics of the parameter spaces across multi-dimensional ensembles.

2 SYSTEM DESIGN AND INTERACTION

2.1 Ensemble Visualization

Our multi-linked visualization tool is centered on three views (figure 1): Ensemble View, Parameter View, and Statistical View.

- Ensemble View -> displays the low-dimensional (2D) projection of ensemble member. The projected distances among ensemble members encode the similarities of the runs in high-dimensional space including input and output parameters; thus, similar runs are placed near each other and dissimilar ones are set farther apart.
- Parameter View -> displays all the ensemble parameters using sliders. Each slider represents the weight of this parameter. This weight represents the relative importance of this parameter.
- Statistical View -> displays graphs (i.e., box plot, scatter plot, and parallel coordinates) that show regions of variability and distributions of the data.

2.2 Ensemble Interaction

Our tool allows scientists to navigate, filter and manipulate ensemble members within the visualization layout. All user interactions are processed by a bidirectional interaction pipeline using simulation input and output parameters (Figure 2). Our approach starts with an ensemble of N 2D images, visualized outputs from each run. Initially, we spatialize the ensemble using its input and output parameters, where each parameter represents a dimension forming high dimensional data, in two dimensional space using weighted multidimensional scaling (WMDS).

An iteration begins in the pipeline with a user interaction (i.e., change spatialization layout by moving ensembles around in ensemble view (i.e. Observation-Level Interaction (OLI)), changing parameter weights (i.e. parametric interaction (PI)), subsetting the data in either ensemble view or statistical) which in turn runs the inverse algorithms from the inverse similarity model through to the ensemble data. After that, the forward model is executed and the results are projected on the visualization layout.

Users can move ensemble members within the ensemble (i.e., OLI operation) to express knowledge or test hypotheses using the "Update Layout" button. OLI is an automated procedure that transforms user interactions with data visualizations (visual feedback) to the parametric feedback that in turn adjusts an entire visual space. OLI operation sends an update message down the pipeline for inverse model to calculate the new set of weights that closely reflect the relative pairwise distances between the transferred members which in turn passes this information to forward model for re-projection on all views.

Similarly, moving parameter view slider (i.e., increasing or decreasing)to increase or decrease the importance of specific parameters (i.e., PI) sends an update message down the pipeline for inverse model to recalculate the weights based on slider movement parameters which in turn calls forward model for re-projection. Users can also send investigate interesting patterns by selecting either Ensemble Members (EMs) in ensemble view or selecting regions of variability in statistical view graphs. Consequently, an update message with these EMs or regions are sent down the pipeline for inverse model to pull the corresponding runs that are passed to forward model for re-projection.

3 CONCLUSION AND FUTURE WORK

We developed a new visualization tool, GLEE, that merges human expertise and intuition with machine learning and statistics for sense-making, analyzing, and exploring any multi-dimensional scientific ensembles irrespective of the application domain. Initial results with computational scientists have demonstrated new discoveries, insights, and hypotheses. Future work includes extending our tool to incorporate in-situ workflows, to support spatial subsets and to support time-series ensembles.

REFERENCES

[1] Alexey Fofonov, Vladimir Molchanov, and Lars Linsen. 2016. Visual analysis of multi-run spatio-temporal simulations using isocontour similarity for projected views. *IEEE transactions on visualization and computer graphics* 22, 8 (2016), 2037–2050.

[2] Lihua Hao, Christopher G Healey, and Steve E Hutchinson. 2015. Ensemble visualization for cyber situation awareness of network security data. In *Visualization for Cyber Security (VizSec), 2015 IEEE Symposium on*. IEEE, 1–8.

[3] Kristin Potter, Joe Kniss, Richard Riesenfeld, and Chris R Johnson. 2010. Visualizing summary statistics and uncertainty. In *Computer Graphics Forum*, Vol. 29. Wiley Online Library, 823–832.

[4] Rainer Splechtna, Mai Elshehaly, Denis Gračanin, Mario ÆĽuras, Katja Bühler, and Krešimir Matković. 2015. Interactive interaction plot. *The Visual Computer* 31, 6-8 (2015), 1055–1065.