

INTRODUCTION

- High-performance computing is creating massive amounts of simulation data.
- Scientists run simulation ensembles using different input conditions to investigate the relationships between parameters.
- Visual analysis of ensembles is a challenging task due to the high multidimensionality, complexity, and size of data.
- Current visualization approaches focus on either: (a) visualizing the ensemble members [1] or (b) the parameter space [2].

CONTRIBUTIONS

Our visual analysis tool intends to help scientists make sense of ensemble data through the interactive exploration of the parameter space of an ensemble. Our contributions are:

- a sensemaking approach that helps scientists investigate ensembles behaviors, patterns, and outliers for high dimensional data using input parameters as well as output variables.
- coupled interactive visual analysis and statistical summaries help users detect and analyze characteristics of parameter spaces across multi-dimensional ensembles.

USE CASES

Our tool is used to explore, compare, and analyze high-dimensional patterns and outliers in simulation ensembles, irrespective of the application domain. We have applied our techniques to a variety of domains including:

- Ecologists analyzing non-equilibrium relaxation features in a stochastic Lotka - Volterra predator-prey model
- Population health experts evaluating the allocation of resources in emergency situations (agent-based simulation)
- Geophysicists analyzing Carbon sequestration methods (Visualized w/ Cinema)

SYSTEM DESIGN AND INTERACTIONS

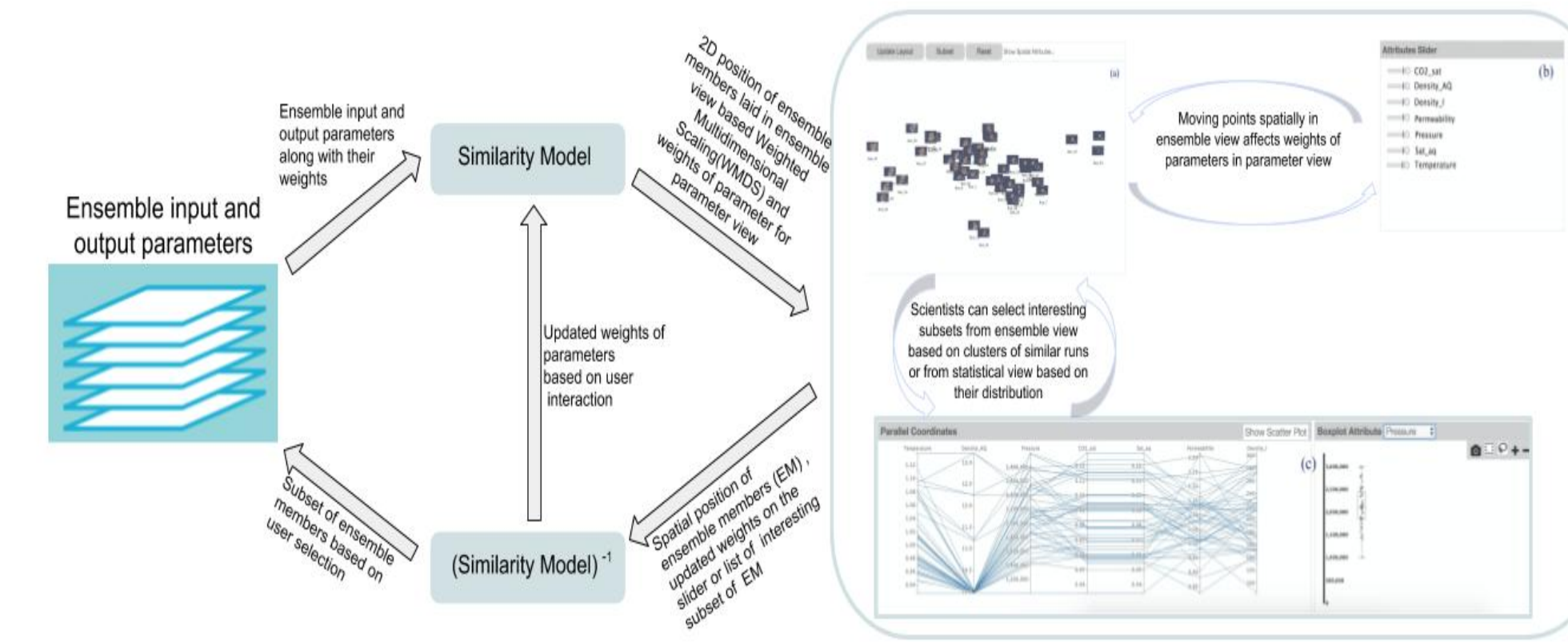


Figure 1: Pipeline of our comparative visualization tool, GLEE (Graphically-Linked Ensemble Explorer). GLEE main views are : (a) Ensemble view: each image represents an ensemble member laid out spatially using WMDS. The projected distances encode the similarities so that similar runs are placed near each other and dissimilar ones are set farther apart. (b) Parameter view: shows weight slider for ensemble inputs and outputs. This weight represents importance of this parameter. (c) Statistical view: displays statistical properties and distributions of ensemble using different graphs (i.e., parallel coordinates, box-plot, and scatter-plot).

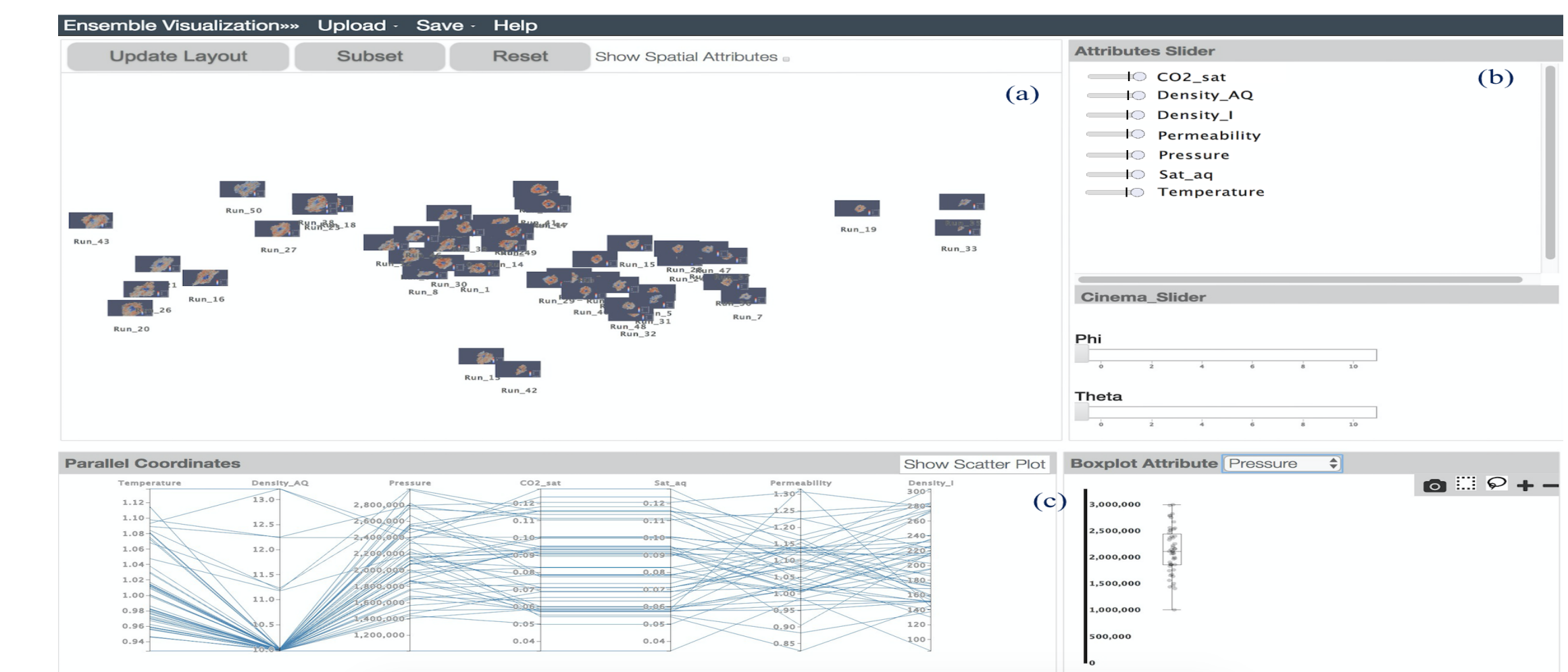


Figure 2: The main interface of our visualization tool. Our tool supports three main interactions: a) change spatialization layout by moving runs around in ensemble view (i.e. **Observation-Level Interaction (OLI)**) (Figure 3); these runs are interactive Cinema visualizations from Paraview b) changing parameter weights (i.e. **parametric interaction (PI)**) (Figure 4); c) subsetting the data in either ensemble view or statistical view (Figure 5).

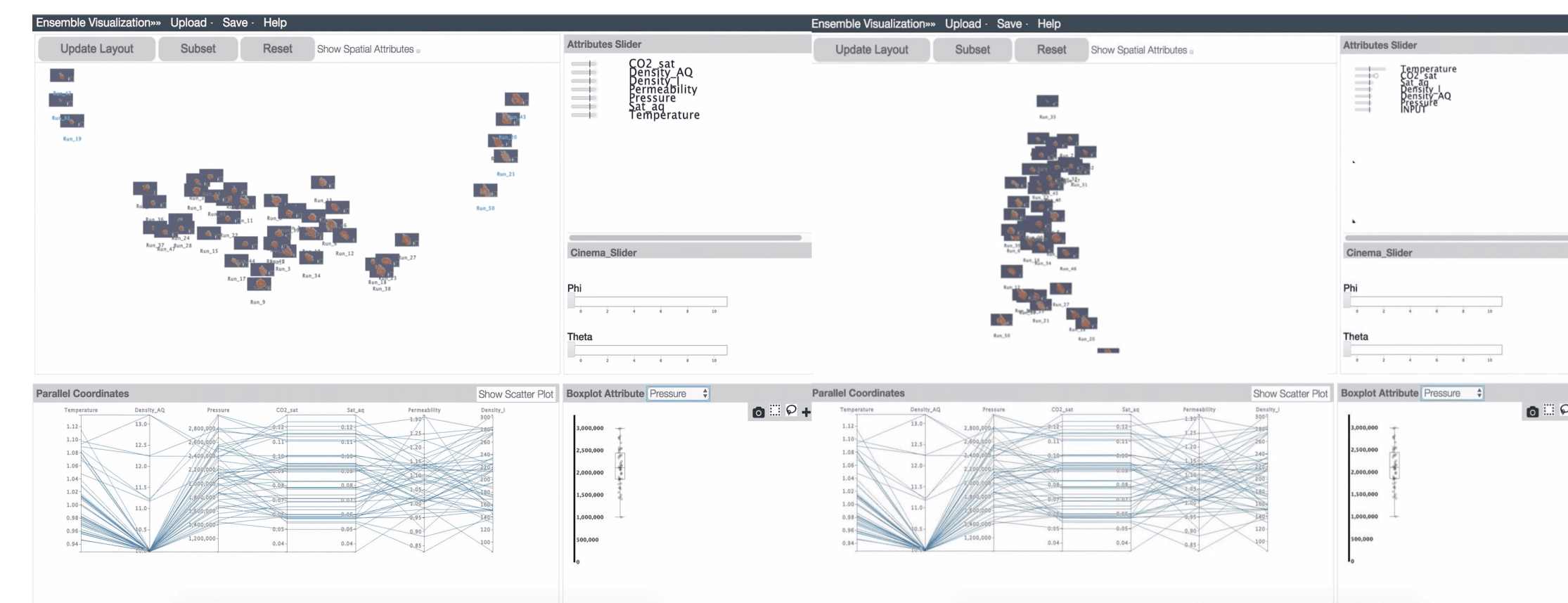


Figure 3: Scientists use OLI to understand effect of temperature by dragging high temperature runs together and low temperature runs together (i.e., highlighted)(left image). The result is an increase in the weights on the temperature and CO2 saturation parameters(right image). This shows that larger CO2 plumes exhibit more variability in temperature. The scientist sees this relationship between permeability and fluid pressure, and the ability of CO2 to expand (i.e., CO2 cools as it expands.)

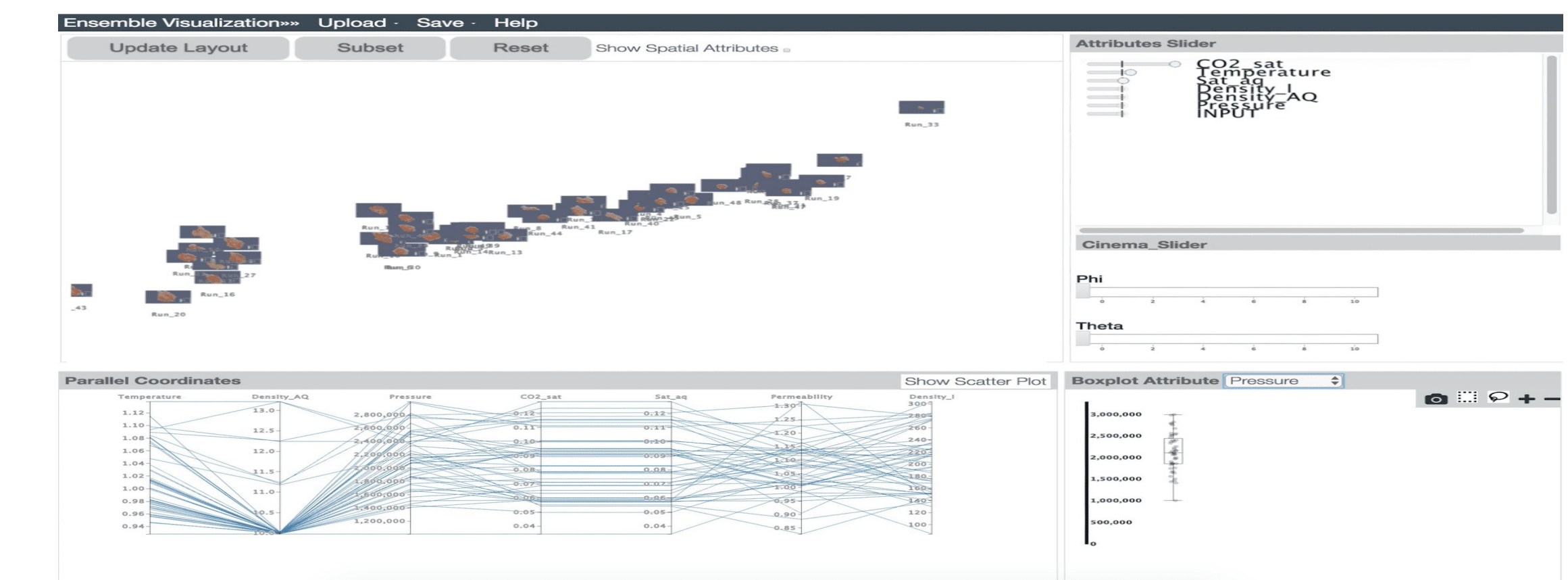


Figure 4: Scientists are also interested in exploring the effect of CO2 saturation, so they used PI to increase the weight of CO2 saturation on the attribute slider (top right). They noticed that smaller CO2 plumes in the lower left all appear to be more circular than the more oblique large plumes. This suggests that the anisotropic permeability correlation structure of this geologic formation exerts more control on plume geometry when the CO2 reaches longer radial distances from the injections well.



Figure 5: During analysis of the ensemble, scientists can subset runs using brushing and linking (i.e., multiple selection) in any visualization (parallel coordinates, scatter-plot, box-plot). Through this process, experts can discover patterns, trends, and relationships across groups of runs in the ensemble.

CONCLUSION AND FUTURE RESEARCH

- We have developed a visualization tool that merges human expertise and intuition with machine learning and statistics for sensemaking, analyzing, and exploring any multi-dimensional scientific ensembles irrespective of the application domain.
- Our linked views and Cinema visualizations present a unique opportunity for evaluating simulation ensembles
- Our evaluation with scientists demonstrates the potential of this tool for discovery and insight
- Future work includes extending our tool to incorporate in-situ workflows, and to support sub-mesh groups and time-series ensembles in our statistical modeling.

REFERENCES

- [1] Alexey Fofonov, Vladimir Molchanov, and Lars Linsen. 2016. Visual analysis of multi-run spatio-temporal simulations using isocon-tour similarity for projected views. IEEE transactions on visualization and computer graphics 22, 8 (2016), 2037 - 2050.
- [2] 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.