Correlation-Aware Uncertainty Visualization of Multivariate Data With Many-Core Acceleration

Alex Gorczowski,[1] Timbwaoga A. J. Ouermi,[2] and Tushar M. Athawale[2]
1: University of Illinois at Urbana-Champaign, 2: Oak Ridge National Laboratory

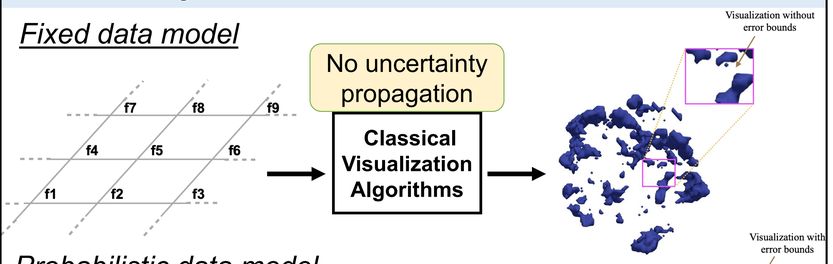
Introduction

Why Visualize Data Uncertainty?

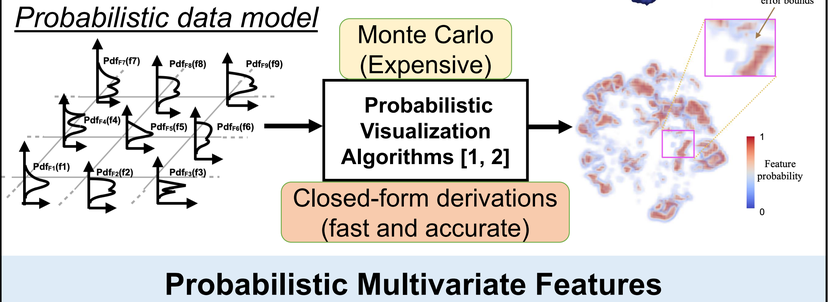
- Uncertainty exists in all simulation, observational, and experimental data
- Leads to inaccurate analysis, if neglected
- Knowing uncertainty is essential to gain trust in the data

Uncertainty Visualization: Extension of 1D Error Bars

Fixed data model

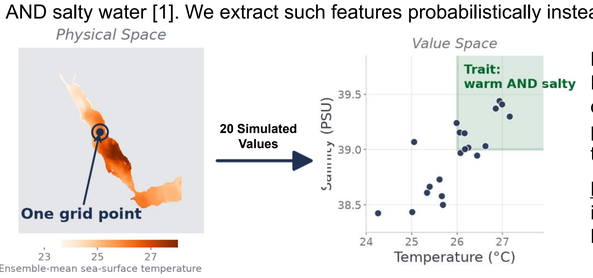


Probabilistic data model



Probabilistic Multivariate Features

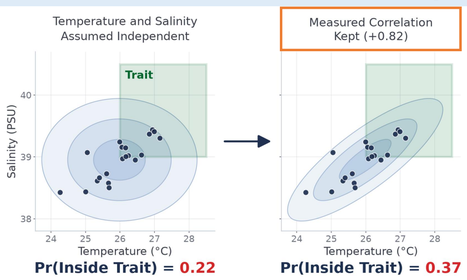
Extracting spatial features that satisfy a user-chosen bivariate range (trait), e.g., warm AND salty water [1]. We extract such features probabilistically instead of binary.



Pr(Inside Trait) = How likely 20 uncertain data values at a grid point belongs to the trait (shown in green)?

Here: 9 of 20 members inside. So,
Pr(Inside Trait) = 0.45

Barrier: The Restricted Independence Assumption



Many existing methods assume the variables are independent models (left image) due to their computational simplicity [1].

Real-world data often exhibits correlation, which we model (right image) and account for, *in closed form,* in this work!

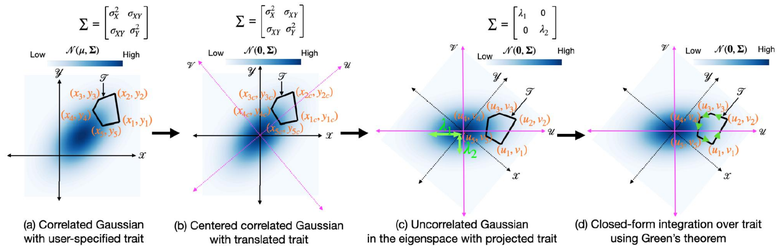
Research Objectives

- Create correlation-aware algorithms for visualizing uncertainty in multivariate data with traits of rectangular and arbitrary shapes
- Adopt kernel density estimation (KDE) noise models for the multimodal and correlated ensembles that parametric models cannot capture
- Replace expensive, traditional Monte Carlo methods with validated closed forms, delivered as production many-core Viskores filters for CPUs and GPUs

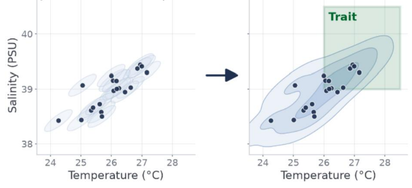
Methods

Correlation-Aware Uncertainty Visualization

Gaussian uncertainty: Each grid point's ensemble modeled as a correlated Gaussian and steps (a)-(d) reduce the trait probability to closed form using eigen decomposition and transforming the problem to the independent case, avoiding costly Monte Carlo sampling.



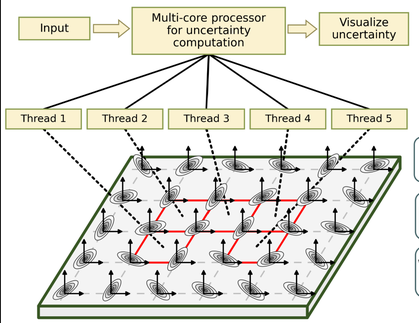
KDE noise model: A closed-form Gaussian kernel computation on every member (with Silverman's bandwidth) is summed.



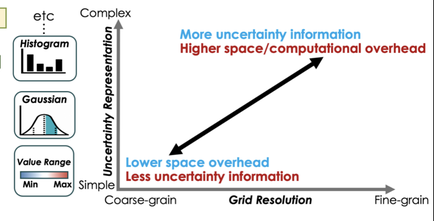
Bandwidth comes from the data and keeps the measured correlation.

Closed-form trait integration applies to every KDE kernel.

Parallelization and Cost-Accuracy Trade-off Analysis

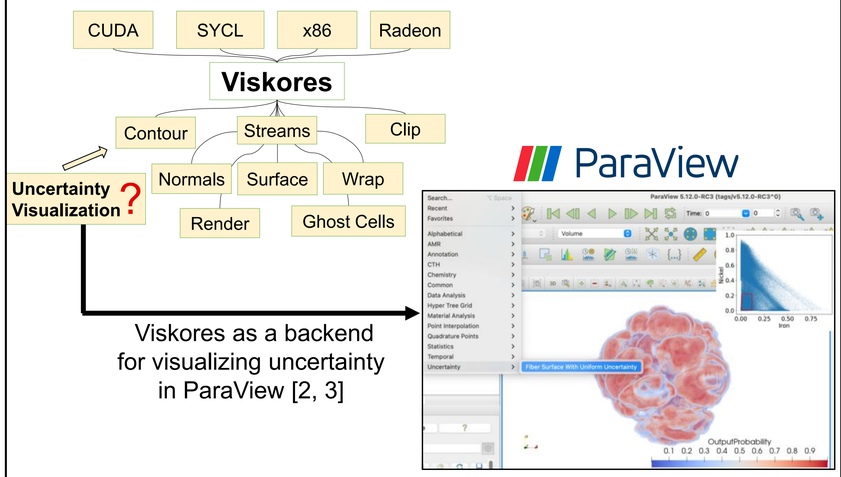


Many-core acceleration with parallel per-grid-point uncertainty computation



Cost-accuracy trade-off analysis

Integration with Production Visualization Tools

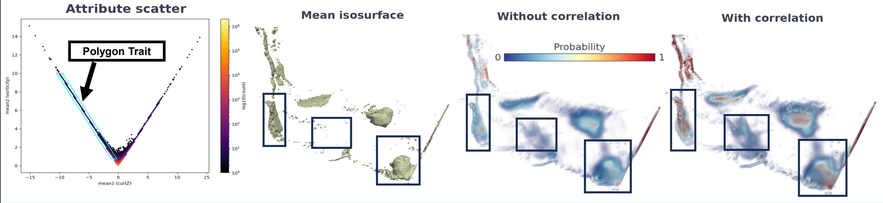


Viskores as a backend for visualizing uncertainty in ParaView [2, 3]

Results

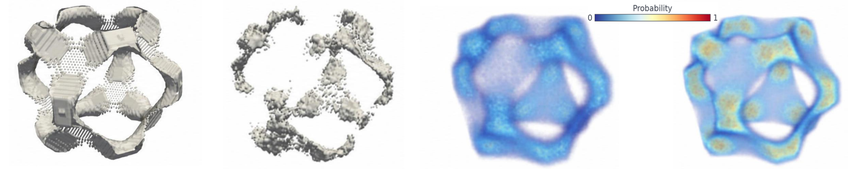
Multivariate: Correlation Accurately Captures Confidence

Vorticity vs. CurlZ trait (shown with the cyan box) on a **2.5M-point** Red Sea ensemble with correlated vs. independent Gaussian kernels for uncertainty visualization



Nonparametric: KDE Outperforms Parametric Models

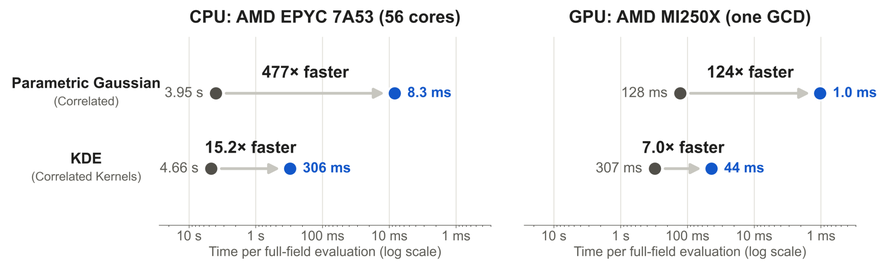
Synthetic ensemble of a trivariate sphere and tangle datasets. Statistical mean creates cracks, whereas KDE preserves confidence that independent Gaussian fails to capture.



(a) Original Synthetic Dataset
(b) Statistical Mean **RMSE: 0.2523**
(c) Independent Gaussian **RMSE: 0.2144**
(d) KDE with Gaussian kernel **RMSE: .1792**

Closed-Form Models Achieve up to 477× Speedup on Frontier

Closed-form vs. Monte Carlo evaluation on 2.5M-point Red Sea ensemble, parametric Gaussian and KDE models on the Frontier supercomputer, giving up to 477× speedup



Future Research

Closed-form, AI, and information-driven models for efficient uncertainty visualization of multivariate data and their topology.

Impact

Our closed-form filters replace expensive Monte Carlo methods on million-point grids, exhibiting massive speedups and providing reliable correlation-aware multivariate uncertainty visualizations.

Acknowledgements

This project was supported by the U.S. Department of Energy Office of Science, Advanced Scientific Computing Research (ASCR), Early Career Research Program, under the contract number DE-AC05-00OR22725. This work also used resources of the Oak Ridge Leadership Computing Facility (OLCF) under the contract number DE-AC05-00OR22725.

References

[1] Athawale, T. M., *et al.*, "Fiber Uncertainty Visualization for Bivariate Data With Parametric and Nonparametric Noise Models", *IEEE TVCG*, vol. 29, no. 1, pp. 613-623, 2023.
[2] Moreland, K., *et al.*, "Viskores: Integrating Parallel Scientific Visualization Research into Applications", *VisGap'26*, Nottingham, UK, 2026.
[3] Viskores Uncertainty Filters: https://github.com/Viskores/viskores/tree/main/viskores/filter/uncertainty.