

# **Bayesian Data Characterization and State Prediction for a Large Eddy Turbulent Flow Simulation: A Revisitation**

**Nicholas V. Scott<sup>1</sup>, Jack McCarthy<sup>2</sup>**

<sup>1</sup>Riverside Research, Open Innovation Center, Dayton Research Center  
2640 Hibiscus Way, Beavercreek, Ohio, 45431, USA  
nscott@riversideresearch.org; jack.mccarthy@duke.edu

<sup>2</sup>Duke University, Department of Statistical Science  
214 Old Chemistry, Box 90251, Durham, NC, 27708, USA

## **Extended Abstract**

Environmental engineering remote sensing platforms using hyperspectral imagery and other multidimensional modalities are often responsible for monitoring coastal regions in order to safeguard national waters. This objective requires determining sub-surface turbulent structure from surface water flow spatial measurements for state assessment and decision-making. The inability of remote sensing platforms to penetrate the water column at depth because of turbulence-induced sediment-concentration modulation necessitates using models that dynamically link surface and sub-surface flow structure. Large eddy simulations (LES) are a useful proxy for the analysis of hyperspectral imagery due to the tri-dimensional structure of both information carrying modalities. Bayesian statistical models are used to revisit the analysis of a large-eddy simulated three-dimensional turbulent shear flow [1]. The purpose is the exploration of the feasibility of creating data characterization and state prediction system models for sub-surface vorticity and stress, and surface root mean square (rms) velocity and rms sediment concentration which could then be utilized in the analysis of environmental hyperspectral imagery.

Nine turbulent feature time series were derived from forty-one three dimensional LES data cube realizations taken at distinct times during steady state conditions. Generative topographic mapping (GTM) [2] is a non-linear latent variable mapping which furnishes a two-dimensional organized representation of nonlinear multidimensional data in terms of latent variables. The nine-dimensional turbulent feature data points including vorticity, stress, rms cross-mean flow velocity, and rms sediment concentration were mapped using the GTM providing a clear segmentation of data points in latent space. This segmentation was characterized by a strong demarcation between low rms and high rms value data points. Vorticity time series and rms sediment concentration time series were selected for Gaussian mixture modeling and hidden Markov model parameter estimation to understand their statistical structure. Gaussian mixture modeling [2] results depict a covariance-based relationship between vorticity and rms sediment concentration which has a positive linear trend for low values of both variables, but which exhibits a negative linear trend at high values. Hidden Markov model parameter estimation [3] results substantiate this relationship between the two variables, in particular the nonlinear dynamics where high vorticity levels are associated with decreasing rather than increasing levels of rms sediment concentration. This trend is due to turbulence dampening where sediment concentration increases the effective viscosity of the sediment-laden fluid flow system to such a level that turbulent fluctuations can no longer sustain large sediment particle concentrations. Gaussian process [4] modeling is able to provide semi-linear extrapolations of future vorticity and rms cross-mean flow velocity values over time. A strong negative linear trend prediction, however, is only possible for rms sediment concentration because of not only the limited amount of data but also the limited dynamic range of data values.

## **References**

- [1] N. V. Scott, and Z. Cheng. (2019, May 10), “Structural field analysis of a large eddy turbulent flow simulation using probabilistic graphical modeling”, in *Proceedings of the Society of Photo-Optical Instrumentation Engineers 11014, Ocean Sensing and Monitoring XI, 110140S*, [Online]. Available: <https://doi.org/10.1117/12.2517112>

- [2] W. L. Martinez, and A. R. Martinez, *Computational Statistics Handbook with Matlab*. London, UK: Chapman and Hall/CRC, 2001.
- [3] L. E. Sucar, *Probabilistic Graphical Models*. London, UK: Springer, 2015.
- [4] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. London, UK: The MIT Press, 2005.