Proceedings of the 5th International Conference on Fluid Flow and Thermal Science (ICFFTS 2024)
Lisbon, Portugal- November 21 - 23, 2024
Paper No. 142
DOI: 10.11159/icffts24.142

Manifold Learning and Graphical Statistical Characterization of Ocean Thermal-Momentum Structure Using Array Observations

Nicholas V. Scott[1], Jeffrey W. Book[2], and Katherine Seery[1]
[1]Apogee Engineering, LLC, Science and Technology Group
2661 Commons Blvd., Suite 200, Beavercreek, Ohio 45431, USA
nicholas.scott@apogeeusa.com; bookatsea@gmail.com; katherine.seery@apogeeusa.com
[2]Naval Research Laboratory, Oceanography Division
Stennis Space Center, MS 39529, USA

Abstract – Ocean drilling is a hazardous enterprise fraught with the possibility of oil spill accidents which can cause millions of dollars of damage to the environment. The 2010 Gulf of Mexico oil spill was responsible for detrimental changes to surrounding bio-chemical ecosystems over both short and long spatio-temporal scales. These sorts of oceanic disasters necessitate the need to understand the thermal-momentum dynamics which drive the spread of oil as well as how physical variables become modulated by fluid contaminants. Machine learning-based analysis is believed to be a valuable aid in this endeavor and was applied to temperature and velocity data acquired from an array of bottom mounted acoustic Doppler current profilers situated near Mobile Bay off the coast of Mississippi. Preliminary analysis of temperature data from 6 profilers covering the upper 6.5 m of the water column showed a 1-degree Celsius temperature increase over a 10-day spring time span in April 2016. This low thermal change was characterized by a slow stable temperature increase over 7 days followed by an increase in temperature variance over the last 3 days. The thermal change consistently appeared in two-dimensional manifolds generated by principal component analysis, Laplacian eigenmap analysis, t-distributed stochastic neighbor embedding, and non-negative matrix factorization as a separable, latent space two-cluster structure. Bayesian belief network analysis of the temperature array data spanning a 25 km spatial scale over the full time span displayed strong evidence of turbulent mixing with a temperature covariance structure suggesting heat transport from east to west along the Mississippi coast as well heat transport in both the offshore and onshore directions. Bayesian belief network analysis of the accompanying velocity magnitude showed a current velocity covariance structure suggesting significant momentum flux in the offshore direction into the deep waters of the Gulf of Mexico. Weak momentum flux in the east-to-west direction was suggested by the Bayesian belief network calculated over the first 7 days.
Link strength analysis of temperature and velocity magnitude Bayesian belief network edges supported the presence of the two-component manifold latent space thermal structure. The slow increase in the large spatial scale temperature field over the first 7 days was partially characterized by a momentum naïve Bayesian sub-network. Naïve Bayesian classification parameterized the velocity magnitude dynamics of a 3-node linear network situated perpendicular to the Mississippi coast with an average predictive error of 38%, providing additional evidence of the anisotropic offshore momentum flux over the spatial scale of the array. The turbulent mixing regime dynamics occupying the last 3 days was parameterized using a single mooring location via Markovian covariant statistical modeling of wave variance with temperature. The hidden Markov model transition matrix for wave variance demonstrated a trend where increasing wave variance has a tendency of transitioning into low wave variance. This was possibly indicative of the time-to-event statistical structure of the wave variance over the 10-day time span with the data showing the repetitive tendency of wave variance growth followed by a transition to a low variance. The emission matrix, where wave variance was modeled as the state variable and temperature as the observation variable, showed a weak pattern of low wave variance associated with low temperature and high wave variance associated with high temperature. Medium level wave variance was associated with a spread of temperature intervals which is possibly associated with turbulent mixing-based transport of heat.
The ability of machine learning analysis to provide a data driven summary of the coastal thermal and momentum dynamics suggests that it could support online machine learning-based data analytics for coastal environmental monitoring.

Keywords**:** acoustic Doppler current profiler, Bayesian belief network, coastal currents, hidden Markov modeling, manifold learning, naïve Bayesian classification, temperature, wave variance

1. Introduction

The 2010 Deepwater Horizon drilling rig explosion was the commencement of an 87-day oil spill responsible for the injection of over 5 million barrels of oil into the northern Gulf of Mexico. This contamination damaged surrounding ecosystems and coastal habitats of the states of Louisiana, Alabama, and Mississippi and made clear the need to understand the pathways of oil transport in the river-dominated Mississippi Bight. The incident also led to the formulation of the consortium for oil spill exposure pathways in the 25 km Mobile Bay. This consortium or the COastal River- Dominated Ecosystems (CONCORDE) possessed a research agenda centered on three scientific aims [1]. The first was the characterization of the 3-dimensional physical oceanographic structure of the coastal Mississippi area to understand potential oil pathways. The second was the description of the spatio-temporal distribution of planktonic organisms along with the interacting geochemical and bio-optical parameters relevant to oil transport. The third was the generation of a model to predict oil transport and estimate the ecological impact of future spill events on the surrounding coastal areas influenced by the Mississippi river. The third objective was and is of great interest to ocean physicists since the construction of Gulf of Mexico model seeks to include as much information as possible about the physical dynamics of momentum, heat, and particle fluxes which are inextricably tied to oil transport [1].

The CONCORDE field research program consisted of oceanographic measurements from multiple modalities including satellite imagery; autonomous glider surveys of bottom boundary layer dynamics; ship-based sampling surveys of temperature, salinity, and current velocity; and moored platform observations [1]. During the spring of 2016 an array of six bottom moorings of upward looking acoustic doppler current profilers (ADCPs) were deployed south of Mobile Bay to study the Mississippi river plume dynamics and exchange with the Gulf of Mexico. It is believed that the analysis of the velocity and temperature data, using computational statistical methods techniques employed by machine learning and artificial intelligence community, can assist in setting the foundation for predictive analytics surrounding future environmental disasters. This work seeks to address a series of oceanographic questions using machine learning to provide an ancillary perspective on a past problem, illustrating the applicability of machine learning statistical techniques. The first query is whether temperature structure and its variability driven by currents, waves, and wind captured over medium sized, spatio-temporal scales admits robust and separable topological structure suggesting different parameterizations for different dynamical regimes. The second question is whether models of the joint probability distribution over the array can robustly characterize the statistical thermal and momentum-based covariance dynamics of these regimes. The third question is whether characterization of the dynamics within captured thermal and momentum dynamical regimes can be quantified allowing for prediction of future momentum and thermal structure. This paper attempts to address these questions using a variety of signal processing and statistical machine learning techniques including manifold learning, Bayesian belief networks (BBNs), naïve Bayesian classification, and hidden Markov models (HMM). After a brief description of the data and analytical methods used, results are given providing preliminary insight into the velocity and temperature spatial-temporal structure. It is through this analysis that evidence supporting the feasibility of the machine learning approach is demonstrated, paving the way for alternate ways of looking at the relationships of waves and currents and their influence on thermal variability as well as their impact on oil transport.

2. Field Data Structure and Data Analysis Methods

2.1. Temperature and Velocity Data Acquired from Array

Temperature and velocity data were acquired from an array of bottom mounted ADCPs deployed in a ‘V’ shaped configuration near Mobile Bay off the coast of Mississippi as shown in Figure 1. Each ADCP was attached to a mooring whose positions are spatially designated as M1-M6. The distance between each mooring spanned by M3 – M6 was approximately 5 miles. The distance between M1 and M2, and M2 and M3 was 3 miles and 6 miles respectively. Each ADCP recorded the average temperature over 10 days with a temporal sampling interval of 10 minutes. The temperature array data was interpolated onto a single common one-dimensional temporal grid to provide the semblance of 6 cotemporally sampled temperature time series each with 1547 samples spanning the time interval of 4/03/2016 to 4/14/2016.

Each ADCP measured the individual 3-dimensional velocity components over a set spatial resolution from the bottom of the water column to the ocean surface. The spatial resolution for M1, M3, M4, M5, and M6 was 0.5 m whereas the spatial resolution for M2 was 0.25 m. Each ADCP provided velocity component measurements over both depth and time. The ADCP data set at each mooring was processed by first calculating the Euclidean velocity vector magnitude at each spatial bin sample

of the water column depth and then averaging these values over the depth bins from the air-sea interface to the depth of 6.5 m. The depth-averaged velocity magnitude estimated in this way provided 6 velocity magnitude time series for the 6 moored ADCPs. The average velocity magnitude times series was interpolated to the same common temporal grid used for the temperature data. Two data matrices for the velocity magnitude and temperature acquired from approximately 4/03/2016 to 4/14/2016 were created in this way.

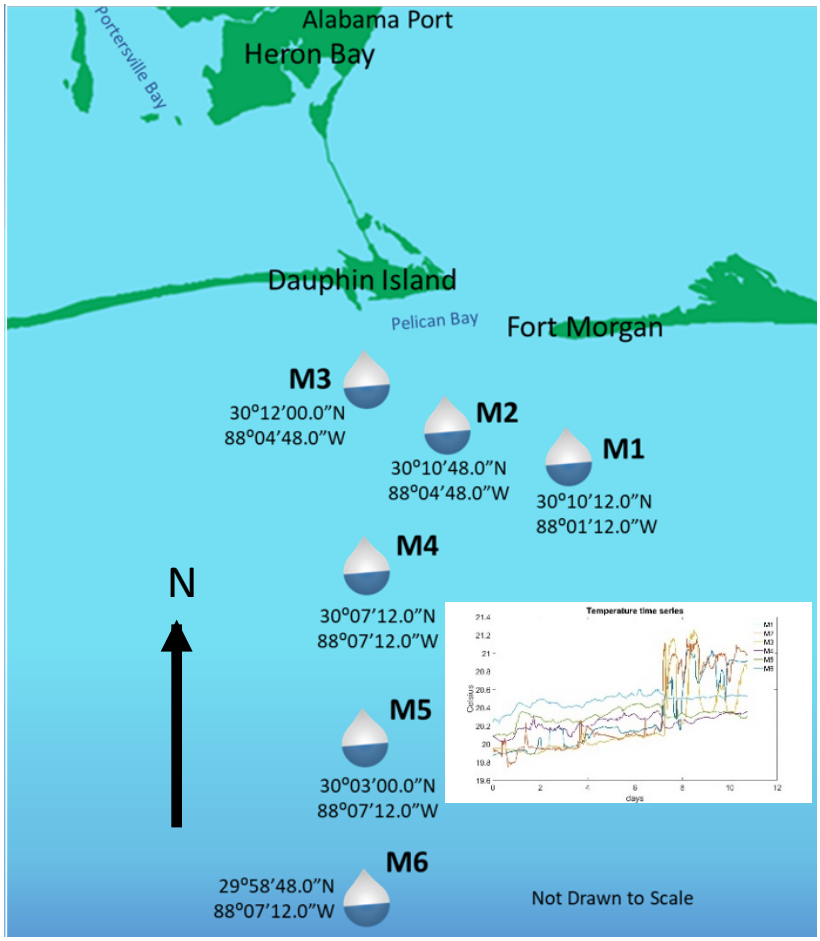

Figure 1: Map of 6 mooring nodes for ADCPS deployed in CONCORDE 2015-2016 field experiment. The average distance between each of the mooring nodes M3-M6 is 5 miles. The average distance between M1 and M2, and M2 and M3 is 3 and 6 miles respectively. The temperature time series recorded by each ADCP at the mooring nodes is shown in the lower right.

2.2. Summary of Major Data Processing Methodologies

A variety of machine learning algorithms were used to extract structural information from ADCP temperature and velocity data. Four different manifold learning algorithms were applied to the 6-dimensional temperature data array. These include principal component analysis (PCA), Laplacian eigenmap (LE) analysis, t-distributed stochastic neighbor (t-sne) embedding, and nonnegative matrix factorization (NMF). Principal component analysis is a dimensional reduction technique that preserves the variance of data via seeking a representation consisting of a linear combination of basis vectors. PCA transforms the data into a new coordinate system such that the projection of the data onto first coordinate has the greatest variance with decreasing amounts of variance allocated to the remaining sequentially ordered coordinate vectors [2]. LE analysis is a nonlinear dimension reduction technique that assumes that the high dimensional data structure lies on lower

dimensional manifold structure and estimates the coordinates of the lower dimensional structure in two parts. A graph is first built parameterizing the proximity of neighboring data points. The generated graph, being an approximation of the higher dimensional data structure, is embedded in a lower dimension and the coordinates found using an eigenfunction spectral technique [3]. LE embedding eigenvectors are created while preserving the data constraint of keeping dissimilar features segregated [3] providing data pattern lucidity.

The t-sne is an unsupervised, local, but nonlinear dimensionality reduction technique for embedding high-dimensional data for visualization in a two-dimensional space. The two-dimensional t-sne starts by calculating pairwise similarity between all data elements in the high-dimensional space using a Gaussian kernel, where data elements that are far apart have a lower probability of being selected than elements close together [4]. The algorithm maps higher dimensional data elements or temperature array data points in a time series onto a lower dimensional space while preserving the pairwise similarities. This is achieved by minimizing the divergence between the probability distributions of the original high-dimensional and lower-dimensional spaces. The optimization allows for the creation of clusters of similar data elements in the lower-dimensional space and provides visual understanding of structural relationships in the original higher-dimensional data.

Nonnegative matrix factorization (NMF) is a method that factors a matrix into the product of two matrices W and H containing nonnegative elements, and in the process, performs dimensional reduction [5]. The factor matrices are found via a minimization scheme for convex functions with non-differentiable constraint called the Proximal Alternating Linearized Minimization (PALM), producing solutions that are smooth and sparse. The columns of W are composed of r NMF eigenvectors of dimension l spanning the data space making the reconstruction of X from W and H a lower-rank approximation. The rank number r is determined a priori. The matrix H is composed of column vectors which are weights for the NMF eigenvectors.

BBNs are probabilistic graphical models which use network edges and nodes to model the joint probability distribution existing between a set of random variables describing a system [6]. The defined and parameterized network nodes and conditional probabilities provide for statistical inference where the effects of evidence at one or more random variable network nodes are propagated throughout the BBN to estimate its impact on others. BBN analysis was performed on temperature and velocity matrix data using the software package Bayes Server manufactured by Bayes Server Ltd. automating much of the statistical analysis including the BBN structural and parameter learning. The Chow-Liu structural learning algorithm was used and assumes a tree structural model appropriate to the data via seeking an edge-node structure consisting of a low number of dominant parental nodes which provide sub-dominant children nodes. Parameter learning provides numerical values to the nodal-edge structure allowing for statistical inference between nodes. This is done using a relevance tree algorithm which allows for exact statistical inference rather than approximate inference [7].

Bayesian classification (BC) and naïve Bayesian classification are classification algorithms based on the implementation of Bayes theorem for the estimation of state or class conditioned on observations or feature information. Both algorithms are based on the allocation of class based on the evaluation of feature information structure. For Bayesian classification, the features are allowed to possess covariance whereas for naïve Bayesian classification features are modeled as being independent of each other, where each feature contributes equally to the designation of class [8]. In this work Bayesian and naïve Bayesian tri-class distributions were made where data points were classified into three classes. The prior probability distribution was set to 1/3 for each class [8,9]. The conditional probabilities used in the classification algorithm are mean and covariance-based Gaussian probability densities for Bayesian classification but mean and variance-based Gaussian probability densities for naïve Bayesian classification. This reflects the feature independent feature of naïve-Bayesian classification.

A hidden Markov model (HMM) is a double stochastic process consisting of a Markov chain hidden state stochastic process, which is not directly observable, and an observation stochastic process which is observable. The statistical relationship between states is parameterized via a state transition matrix A, while the relationship of states to observations is parameterized via an emission matrix B [10]. The HMM is applied to wave variance measurements X and temperature measurements Y, at a single mooring locations where X is considered the state variable and Y the observation variable. The parameters of the HMM are estimated by taking the values in each of these measurement variables and dividing them up into intervals. The transition frequency between state intervals in X is counted along with the calculation of the distribution over the observation variable intervals for Y when the state transition is made. This instance counting is simply a method for

counting the number of state transitions and observation emissions associated with each state transition [11]. A fully parameterized HMM, consisting of transition and emission matrices, allows for understanding how a random variable changes and it's probabilistic connection with another variable where the two possess a conditional relationship.

3. Results

Previous study of the current dynamics in the Mobile Bay outflow area connecting to the Gulf of Mexico revealed very complex dynamics. For example, despite the proximity of the area to the Mississippi River outflow, it was found that riverine outflow did not play a significant role in driving weekly variations of currents during the field study period [1]. Rather, local southeasterly winds drove south-westward currents which were responsible for offshore fluxes, and north-westerly winds drove north-eastward currents responsible for onshore fluxes [1]. Overall, the interaction of wind, waves, and currents was directly related to the levels of temperature-momentum variability and turbulent mixing in the study region. ADCP observations support this dynamical picture. The application of manifold learning in the form of PCA, LE analysis, t-sne, and NMF-PALM to the 1547 x 6 temperature matrix capturing the thermal dynamics measured by the 6 component ADCP array consistently demonstrates a 2-cluster latent space structure. PCA, LE analysis, and NMF-PALM showed two highly separable thermal state clusters which corresponds to the initial slow stable rise in temperature transitioning to a state of large amplitude thermal fluctuations as shown in lower right corner of Figure 1.

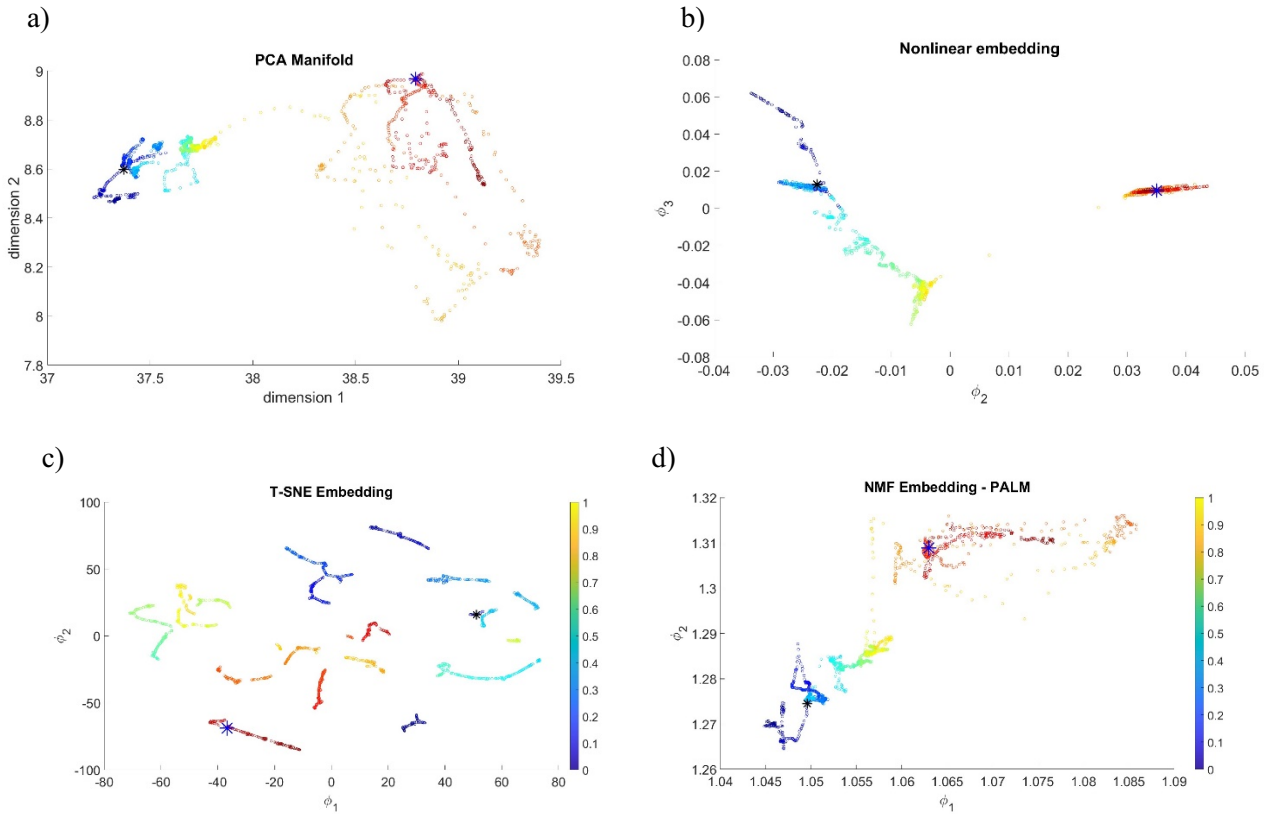

Figure 2: Manifold learning-based projections for a) PCA b) LE analysis d) T-SNE, and d) NMF-PALM. First two eigenvectors for temperature used for PCA and LE with correspond to eigenvectors associated with maximum and minimum variance respectively. NMF-PALM temperature eigenvectors used in plot selected to demonstrated maximum qualitative latent space cluster separation.

The average temperature values associated with the stable temperature rise and large temperature variance regimes, as they appear in their respective latent space clusters, are designated by black and blue asterisks respectively. These were calculated by taking the arithmetic mean values for the first 1000 point temperature segment and the second 547 point segment followed by plotting the latent space values which were closest to these calculated mean values in a Euclidean sense. The t-sne also shows strong latent space separability with strong latent space thermal pockets over the temperature time series.

BBN analysis using the Chow-Liu algorithm for the full-time span of temperature measurements, as shown in Figure 3a), reveals temperature covariance signified by arrow-based directed edges suggesting heat flux in the east-to-west along-coastal direction and in the offshore direction. ADCP moorings 1 to 3 are connected with mutual information values, parameterizing link strength, on the order of 0.5 (on a 0-1 scale), suggesting a medium but significant linkage level and heat flux continuity between the ADCP mooring stations. The heat flux in this area is towards the east in the direction of the Mississippi river which is noteworthy and unexpected. Linkages also exist between ADCP moorings 2 and 5, and moorings 5 and 6 suggesting offshore heat flux. It is worthy of note that a linkage also exists between ADCP moorings 6 and 4 suggesting onshore directed heat flux. The BBN demonstrates evidence of strong temperature variability with temperature information propagation existing in both the offshore and onshore direction as well towards the Mississippi over the spatial scale of the array.

BBN analysis of the concurrent velocity structure over the same temporal period reveals evidence of momentum flux in the offshore directions as shown in Figure 3b). Such advection-based momentum flux directions are partially consistent with the thermal flux directions indicated by the BBN analysis of temperature. Evidence of momentum flux in both the east-to-west and west-to-east directions also exists in the network graph. However, the mutual information link strengths range of [0.2 0.5] demonstrates that a coherent covariant velocity magnitude strength that is weak. The direction of the edge arrows in the along-coastal direction are partially consistent with thermal fluxes shown in Figure 3a) again suggest turbulent mixing.

The two cluster latent space states depicted in the manifold learning projections of Figure 2 prompted an examination of the dynamics over the two-time scales associated with these states. Shown in Figure 3c) is the BBN for the velocity magnitude over the slowly increasing temperature time scale of 7 days before it turns visibly turbulent as shown in Figure 1. The mutual information values for this 7-day time scale for all edges are small being on the order of 0.2 again suggesting turbulent mixing. Given the large spatial distances between the ADCPs, the mutual information value is totally reasonable. The BBN edge structure is very similar for the velocity over the complete 10-day time scale which suggests that the dynamics over the first 7 days makes a strong contribution to the overall momentum flux dynamics. The BBN for the velocity magnitude for the last 3 days is not shown but also has small edge mutual information values. Edge arrow structure for this BBN was more erratic and is consistent with the large amplitude thermal fluctuations over this time period again suggesting mixing.

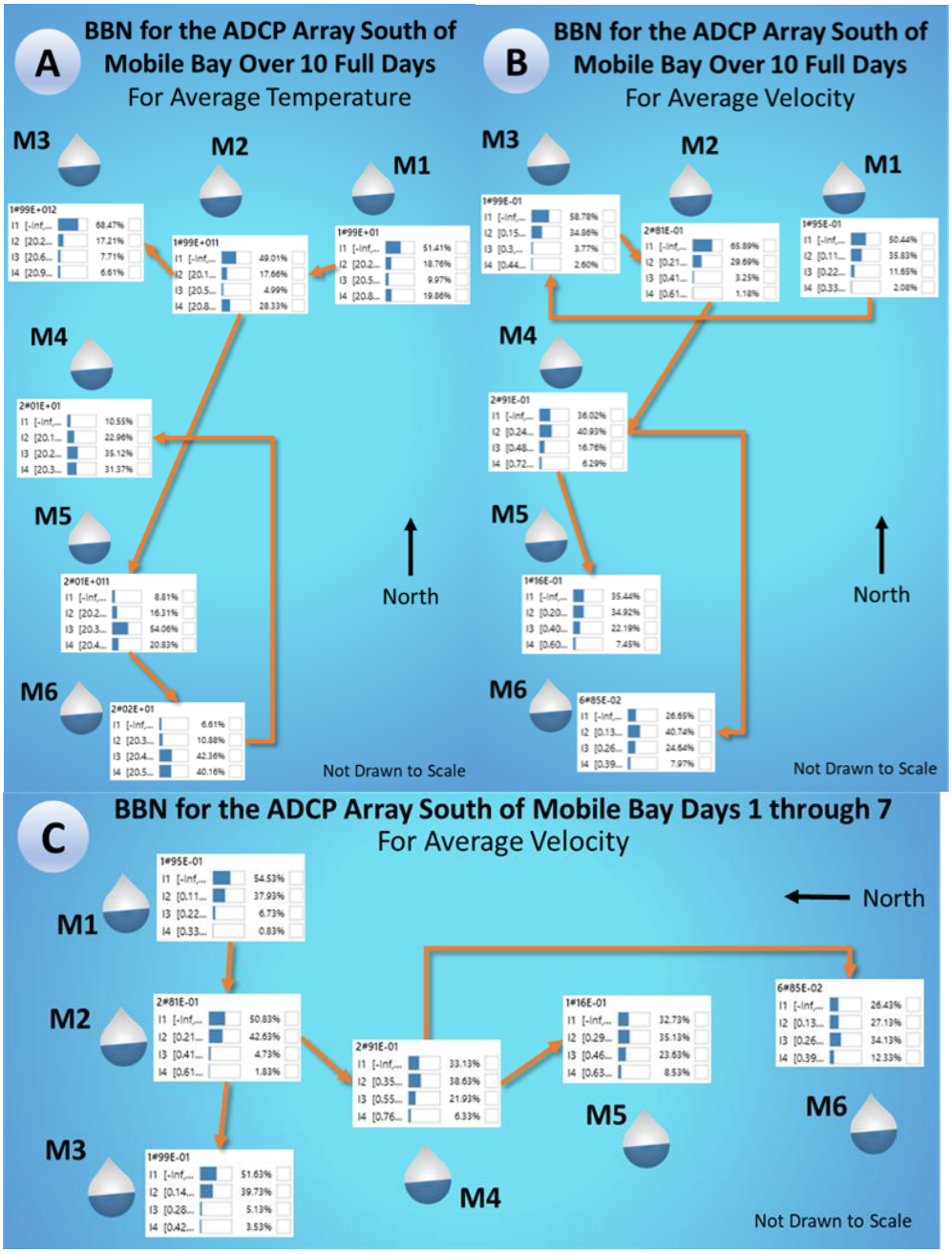

Figure 3: BBN for the ADCP array situated south of Mobile Bay over the full 10 days for a) the average temperature and b) the average velocity. c) BBN for average velocity captured by ADCP array over the first 7 out of 10 days. Marginal distributions shown with no instantiations. Mooring labels appear next to BBN nodes.

Further analysis of the two different thermal regimes was performed by examining unique momentum and thermal flux processes using two different statistical techniques. Bayesian and naïve Bayesian classification was performed to understand if the offshore momentum flux dynamics over the first 7 days characterizing the slowly increasing, stable thermal dynamics and captured by a spatial velocity sub-network, could be reasonably modeled. The Mississippi river outflow is anisotropically directed from north to south, and ADCP at moorings M4, M5, and M6 comprise this sub-network which were used to model

this flow using Bayesian and naïve Bayesian classification. For the Bayesian and naïve Bayesian classification, the ADCP velocity magnitude at M4 was modeled as the state variable and the ADCP velocity magnitudes at M5 and M6 were modeled as observation variables. The first 500 points of the velocity data array were used as training data allowing for the parameterization of the Bayesian and naïve Bayesian models. The velocity magnitude range was divided into 3 equal intervals where the range of 0.1 to 0.2 m/s was designated as interval 1. The range of 0.2 to 0.3 m/s and 0.3 to 0.4 m/s were designated as intervals 2 and 3 respectively. Given velocity information at each of the observation variables, the classification algorithms provide optimal estimates of the velocity state intervals in the state variable. This is performed using new velocity data values which serve as algorithmic test values. Figures 4a)- 4b) show rough agreement between classification values and the true values for the last 500 velocity data points. The error for both algorithms was approximately 38% which is reasonable suggesting that neither algorithm is significantly better in the estimation of velocity magnitude for the first thermal regime.

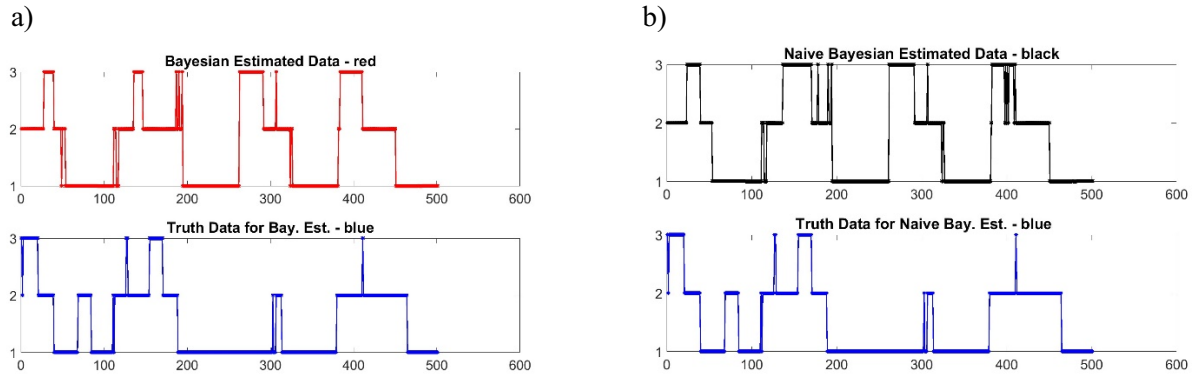

Figure 4: a) Bayesian classification results and true values for the last 500 points of the 1000 velocity samples for the velocity magnitude sub-network of M4, M5, and M6. b) Naïve Bayesian classification results and true values for the same data segment and sub-network. State variable class divided up into equidistant velocity intervals over the range [0.1 0.4] m/s. Average percent error difference between estimates and true values on the order of 38% for both classification algorithms.

The momentum-thermal dynamics of the second thermal regime sensed by the ADCP array, characterized by large temperature variance over the last 500 points or 3 days as shown in the bottom right part of Figure 1, can also be modeled. HMM parameter estimation was performed where the concurrent ocean surface gravity wave standard deviation (std) and temperature structure measure by the ADCP at mooring 1 was modeled. For the HMM the wave std and the temperature were modeled as the state and observation variables respectively. The choice of state and observation variable is based on the physical concept of the velocity field being responsible for the advection of heat. Figures 5a)-b) displays the transition and emission matrices for the HMM. The wave std transition matrix does not display the typical transition matrix structure where large values are sequestered near the diagonal. Such an energy structure delineates state transition dynamics where one wave std level on the vertical is strongly probabilistically related to an adjacent state level to the right and left horizontally only. The transition matrix possesses a structure where large state value levels gradually curve downward indication that large wave std are gradually associated with low wave std values. This transition matrix structure describes time-to-event dynamics where low and medium wave std values often transition into themselves, but as high wave std values are approached, there is an increasing tendency of transitions into low wave std intervals. This model feature describes the wave variance time series over the 10-day time span which displays a cyclical process of wave std growth followed by a transition into low wave std. Wind modulations were responsible for wave growth and death occurring roughly over a 2.5-day time scale with the wave pressure transition structure captured by the HMM transition matrix.

The emission matrix, parameterizing the relationship of temperature structure conditioned on wave pressure variance state transitions, depicts very noisy probability density functions for each wave std state level. (The rows of the emission matrix sum to 1). The wideness of the pressure probability distribution for each row suggests that large amounts of mixing occur over the 10-day time span. The emission matrix shows a weak pattern of low wave pressure std associated with low

temperature and high wave pressure std associated with high temperature. Medium level wave pressure std is associated with a spread of temperature intervals which is possibly associated with turbulent mixing based transport of heat.

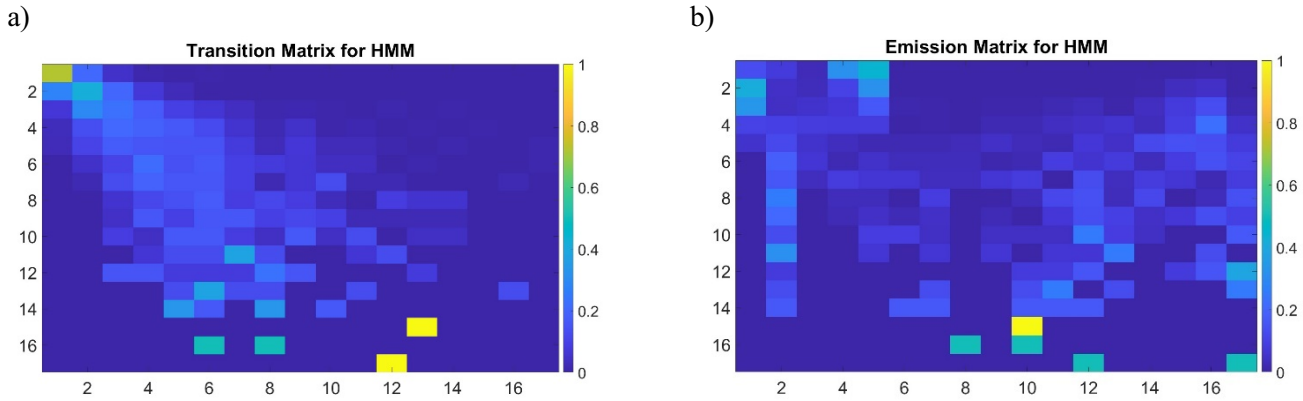

Figure 5: Transition and emission matrices for normally distributed wave pressure std (state variable) and temperature (observation variable) for ADCP measurements at mooring 1. a) Wave pressure std transition matrix where vertical and horizontal axis are variance intervals. b) Wave pressure std – temperature emission matrix where vertical axis designates wave std intervals and horizontal axis temperature intervals. State and observation intervals divided up into 16 linearly equidistant wave std and temperature intervals. The complete wave std interval span is from 0.0042 to 0.2420. The complete temperature interval span is from 19.87 to 21.05 degrees Celsius. Each row for the emission matrix sums to 1 making it a true probability distribution.

4. Conclusion

Ocean drilling accidents, responsible for the inundation of ocean waters with oil, has detrimental effects to surrounding bio-geochemical systems that are only now starting to be fully understood. Part of the full comprehension of the effects of oil transport emanates from understanding of the oceanic momentum and temperature structure which allows for the prediction of the ultimate effects of contaminants over both small and large spatio-temporal scales. Machine learning can play an integral role in this process deep understanding. Preliminary results emanating from statistical machine learning provide corroboration of previously studied coastal dynamics. Dimensional reduction and manifold learning of temperature array data reveals a two-cluster latent space structure capturing the slow, stable temperature increase regime followed by a regime characterized by large temperature variance. BBN analysis for these two temperature states shows evidence of heat flux in both onshore and offshore directions as well heat flux directed towards the Mississippi River outflow. BBN analysis of velocity magnitude captured over the same spatial and temporal scale for temperatures depicts a velocity field with significant offshore and along-coastal advection in the direction the Mississippi river. Temperature and velocity covariant structure suggests the impact of turbulent mixing over the 25 km scale of the ADCP array with BBN-based evidence suggesting that large amounts of offshore and along coast Mississippi river directed fluxes occurring over the first 7 days.

Analysis of the dynamics within each temperature regime state diagnosed by manifold learning reveals unique structure. Bayesian and naïve Bayesian classification of a sub-network of ADCP velocity array measurements taken in a linear array aligned perpendicular to the Mississippi coast demonstrates velocity magnitudes over the first thermal state regime which can reasonably be predicted. Percent error values on the order of 38% for both classification algorithms confirm this. Large temperature variance over the second thermal state regime suggests the presence of wave induced mixing. The wave induced momentum-temperature structure extracted from a single mooring during the second thermal state regime was modeled using HMM analysis. A wave pressure std transition matrix structure exists consistent with a time-to-event statistical model. The birth and the decay of wave pressure variance is most likely driven by wind variability which also causes the temperature variance. The emission matrix characterizing the conditional relationship of wave pressure std as the state variable to temperature as the observation variable displays large amounts of variance for each wave pressure std interval. Even with large amounts of pressure variance in the emission matrix, there is a weak pattern of low temperatures associated with low wave pressure std and high temperatures associated with high wave pressure std.

All machine learning and signal processing algorithms employed in this work possess the capability to be employed in an automated fashion for array processing of oceanographic measurements for momentum-temperature field diagnostics. Algorithmic results strongly suggest the utility for the machine learning approach for environmental monitoring to improve understanding of the dynamic structure of variables which control contaminant transport.

References

[1] J. W. Book, N. L. Jones, R. J. Lowe, G. N. Ivey, C. R. Steinberg, R. M. Brinkman, A. E. Rice, C. E. Bluteau, S. R. Smith, T. A. Smith, and S. Matt, "Propagation of Internal Tides on the Northwest Australian Shelf Studied with Time-Augmented Empirical Orthogonal Functions," in *20th Australasian Fluid Mechanics Conference*, Perth, Australia, December 5-8, 2016.

[2] B. Ghojogh, M. Crowley, F. Karray, and A. Ghodsi, *Elements of Dimensionality Reduction and Manifold Learning*, Cham, Switzerland: Springer, 2023.

[3] A. J. Izenman, Modern Multivariate Techniques: Regression, Classification, and Learning, New York, NY: Springer, 2008.

[4] N. Gillis, Nonnegative Matrix Factorization, New York, NY: SIAM-Society for Industrial and Applied Mathematics, 2020.

[5] L. van de Maaten, and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 1, pp. 1-48, 2008.

[6] K. B. Korb and A. E. Nicholson, *Bayesian Artificial Intelligence*. Florida, USA: CRC Press, 2010.

[7] U. B. Kjaerulff and A. L. Madsen, *Bayesian Network and Influence Diagrams: A Guide to Construction and Analysis.* New York, NY: Springer, 2008.

[8] S. Theodoridis, A. Pikrakis, K. Koutroumbas, and D. Cavouras, *Introduction to Pattern Recognition: A Matlab Approach*, Burlington: MA, Academic Press, 2010.

[9] W. L. Martinez and A. R. Martinez, *Computational Statistics Handbook with MATLAB, Third Edition*. Boca Raton, FL: Chapman and Hall/CRC, 2015.

[10] O. C. Ibe, *Markov Processes for Stochastic Modeling*, Second Edition, New York: NY: Elsevier, 2013.

[11] G. Fink, *Markov Models for Pattern Recognition: From Theory to Applications*, New York, NY: Springer, 2014.