

Classification of Emerging Extreme Event Tracks in Multivariate Spatio-Temporal Physical Systems Using Dynamic Network Structures: Application to Hurricane Track Prediction

Huseyin Sencan^{1*}, Zhengzhang Chen^{1*}, William Hendrix³, Tatdow Pansombut^{1,2}, Frederick Semazzi¹, Alok Choudhary³, Vipin Kumar⁴, Anatoli V. Melechko¹, and Nagiza F. Samatova^{1,2,†}

¹ North Carolina State University
Raleigh, NC 27695

² Oak Ridge National Laboratory
Oak Ridge, TN 37831

³ Northwestern University
Evanston, IL 60208

⁴ University of Minnesota
Minneapolis, MN 55455

[†] Corresponding author: samatovan@ornl.gov

Abstract

Understanding extreme events, such as hurricanes or forest fires, is of paramount importance because of their adverse impacts on human beings. Such events often propagate in space and time. Predicting—even a few days in advance—what locations will get affected by the event tracks could benefit our society in many ways. Arguably, simulations from first principles, where underlying physics-based models are described by a system of equations, provide least reliable predictions for variables characterizing the dynamics of these extreme events. Data-driven model building has been recently emerging as a complementary approach that could learn the relationships between historically observed or simulated multiple, spatio-temporal ancillary variables and the dynamic behavior of extreme events of interest. While promising, the methodology for predictive learning from such complex data is still in its infancy. In this paper, we propose a dynamic networks-based methodology for in-advance prediction of the dynamic tracks of emerging extreme events. By associating a network model of the system with the known tracks, our method is capable of learning the recurrent network motifs that could be used as discriminatory signatures for the event’s behavioral class. When applied to classifying the behavior of the hurricane tracks at their early formation stages in Western Africa region, our method is able to predict whether hurricane tracks will hit the land of the North Atlantic region at least 10-15 days lead lag time in advance with more than 90% accuracy using 10-fold cross-validation. To the best of our knowledge, no comparable methodology exists for solving this problem using data-driven models.

1 Introduction

The behavior of a complex physical system, such as the atmospheric-ocean system, is often characterized by patterns

or events (e.g., hurricanes) induced or inhibited by fluctuations associated with non-linearly coupled system parameters. These patterns are not static but dynamic, evolving both in space and time. For example, the fronts of plasma turbulence can propagate toward the edge of the fusion reactor and lead to fusion energy loss. Likewise, in climate, hurricanes might ultimately hit land, potentially causing detrimental side effects, or dissipate in the ocean. Even if the data about such systems is available, either through simulations or through observations, the problem of detecting and tracking such events in both space and time is often a non-trivial task due to inherent complexity of the system’s dynamics.

From a statistical learning perspective, the problem has been studied from two main perspectives. One approach is the *real-time, dynamic* detection and tracking of events. For example, detection of turbulent fronts has been enabled via sliding window linear approximation of spatio-temporal time series data followed by signal-to-noise ratio amplification through the product of the line’s anti-correlated slope and intercept [Shah *et al.*, 2010]. Alternatively, tracks may be analyzed during *post-processing*. For instance, *unsupervised learning* has revealed physically meaningful clusters of hurricane tracks with similar signatures [Gaffney *et al.*, 2007].

While both approaches deserve their own merits, in this paper, we draw readers’ attention to a slightly different, yet complementary, *supervised learning* problem. Namely, given a historic record about the spatio-temporal propagation of the event tracks, can an algorithm learn the complex non-linear relationships between system parameters so that the track’s *sink* location can be predicted from the knowledge about the track’s *source* and the system’s state at or before the event’s formation? Slightly more formally, if during the event e at time t_0 , the system’s state $S(t_0, e, F, L)$ and the source of the event track are described by some spatio-temporal multivariate feature set F over space L and time $t \leq t_0$, then the algorithm A should predict the sink of the event track at time $t_0 + \Delta t$, where the temporal resolution, Δt , is domain-specific (e.g., 5–15 days for hurricane tracks). For the sake of simplicity, we assume that the number of distinct modes for the track sinks is finite; for example, land-hitting and ocean-bound hurricane tracks. We call this problem a *supervised forecasting of source–sink track dynamics*.

To address this problem, we propose a network-based methodology for identifying statistically significant, *sink*-

*Both authors contributed equally

biased network motifs by comparing more than a hundred system-scale networks associated with the event tracks. We then utilize a multivariate, iterative feature selection method that detects combinations of these motifs that collectively contribute to the decision on the particular mode of the track's sink. We successfully apply this methodology to a suggested problem of forecasting whether hurricanes at their genesis region would ultimately landfall or remain offshore. The forecast skill is 15%–20% higher than any other competitive machine learning methodologies could offer.

2 Motivating Example and Related Work

Hurricanes lead to major natural disasters in the regions of landfall. To better understand and predict regional hurricanes, it is important to consider the factors influencing the dynamics of their tracks. Data on hurricane locations from the “best-track” (HURDAT) data or from a simple tracking algorithm applied to 6-hourly mean sea-level pressure (SLP) fields from either a general circulation model (GCM) or an observed data set [Gaffney *et al.*, 2007] provide opportunities for data mining to contribute to understanding these factors.

Hurricane track research [Harr and Elsberry, 1991; Elsner *et al.*, 2000; Elsner, 2003] provided direct links of the sea surface temperature (SST) or vertical wind shear (VWS) to hurricane steering mechanisms and regional factors related to hurricane tracks. Recent cluster analysis of hurricane tracks [Gaffney *et al.*, 2007; Nakamura *et al.*, 2009; Kossin *et al.*, 2010] revealed distinct and physically interpretable clusters in the main development region of the central tropical Atlantic, including the following [Elsner, 2003] (see Figure 1, A): (a) “Recurving” hurricanes that threaten North America north of about 35°N but remain offshore and (b) “Straight Moving” (SM) hurricanes that make landfall in the Caribbean and North America south of this latitude.

Likewise, Poisson regression models [Kossin *et al.*, 2010] provided quantitative insights on the frequency trends for landfalling hurricanes, thus giving an estimation of annual rates for landfalling hurricanes given our knowledge about SST anomalies and other climate factors, such as the variability of the AMM, ENSO, NAO, and MJO indices.

In contrast, in this paper, we address an important and complementary problem—forecasting whether the hurricane in its genesis region will eventually strike land or remain offshore. Existing real time hurricane track prediction studies, such as CLIPPER [Neumann, 1972] and GFDL [Tuleya and Ross, 1995], focus on short-term (2–5 days) predictions. To complement these methods, we aim to make mid-term (5–10 days) predictions after hurricane formation and reveal important factors for possibly improving climate models or seasonal tropical cyclone activity projections. To the best of our knowledge, no literature has addressed this problem.

3 Method

3.1 Overview

In this section, we describe our methodology for classifying the dynamics of source-sink event tracks for emerging extreme events, with the goal of predicting the mode of the

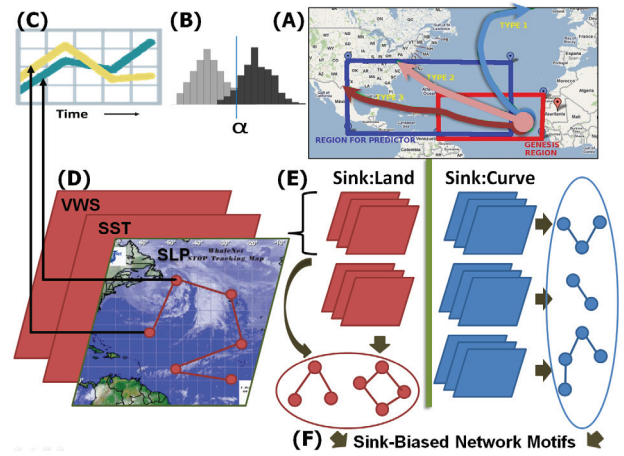


Figure 1: Sink-biased network motif discovery.

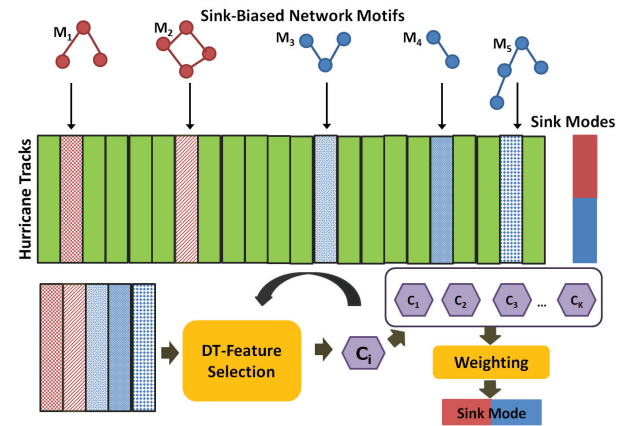


Figure 2: Sink-biased motif groups enumeration.

track's sink with lead lag time in advance of the track's formation at its source. Figure 1 and Figure 2 summarize the two main components of the proposed methodology. The first primarily deals with formation of system-scale networks and discovery of sink-biased network motifs. The second identifies the groups of these motifs that collectively provide the ability to predict different sink modes.

3.2 Sink mode–driven network formation

Physical systems, such as climate, have been recently studied from dynamic, structural, and functional properties of their complex networks. For example, the nodes of climate networks, or graphs, are identified with the spatial grid points and the edges between pairs of nodes exist depending on the degree of statistical interdependence between the corresponding pairs of anomaly time series taken from the climate data set [Tsonis and Roebber, 2004; Donges *et al.*, 2009].

Unlike traditional unsupervised methods for network formation, our supervised network construction distinguishes the mode of the track's sink at time t_{end} for the track at its source location at time t_{start} . Basically, our methodology is driven by the sink mode (Fig. 1(A), 1(E), and 1(F)). Given a

set of networks for each sink mode, our hypothesis is that the sink mode networks will exhibit distinctive network motifs, or subgraphs, that can be used to identify the target mode. For example, if the sink belongs to the Land or Offshore mode for the hurricane track, we call such biased network motifs Land-specific or Offshore-specific, depending on the bias type.

Because network topology affects our detection of such sink-specific network motifs, we will next discuss an important step underlying our network construction methodology—the assessment of statistical significance of network edges, namely, whether the edge is unlikely to be observed in random networks. Since the system state is described by multi-variate spatio-temporal data, we first compare univariate time series associated with pairs of spatial grid points for a time period of $(t_{start} - \Delta W, t_{start})$, where the time window ΔW is problem-specific (e.g., 90 days for hurricane tracks) (see Figure 1, Part C). We then adjoin the points with a weighted or unweighted edge, if these series are sufficiently correlated.

To determine whether the time series associated with two grid points are correlated, we randomly permute the time series associated with each data point and compare the distributions of the pairwise Pearson correlation between the permuted and original data across all pairs of points (see Figure 1(B)). In this figure, the black histogram represents the distribution of the pairwise Pearson correlation coefficients for the original data, and the gray histogram represents the correlation distribution for the permuted data. The edge weight threshold α is selected based on the desirable False Discovery Rate (FDR) defined from the overlap between the distributions. We connect pairs of data points with an edge if the Pearson correlation between their time series is greater than the threshold (see Figure 1(D)). (Note that the permutation mechanism is domain- and task-specific.)

3.3 Sink-biased network motifs detection

In a brute-force way, the enumeration of all sink-biased network motifs might first enumerate all frequent subgraphs [Inokuch *et al.*, 2000; Kuramochi and Karypis, 2001; Yan and Han, 2002] and then assess their statistical bias toward a particular sink mode. However, this approach quickly becomes impractical as there may be exponentially many such subgraphs. To deal with this computational complexity, we employ a much more efficient bottom-up approach that first identifies sink-biased edges as seed sets and then expands those sets via seed expansion algorithm. The seed expansion algorithm is outside of the scope of this paper, so we also show results only for the network motifs (edges) identified as part of the seed selection process. In this paper, we use edges in the climate networks as motifs and identify those edges that are biased towards a particular class of hurricanes.

We specifically distinguish between treating the networks as unweighted graphs (or using *binary* edge weight values) or weighted graphs (using *continuous* edge weights). For the binary case, we calculate a cumulative hypergeometric probability to assess whether a given edge is biased towards a particular sink mode or another. We utilize the Fisher's exact test that estimates the cumulative hypergeometric probability, used to measure the likelihood of randomly selecting events

with the same (or less balanced) class distribution:

$$h(x \geq x_C; N, n_C, k) = \sum_{i=x_C}^k \frac{\binom{k}{i} \binom{N-k}{n_C-i}}{\binom{N}{n_C}},$$

where x_C represents the number of events of class C that exhibit the edge, N represents the total number of events, n_C represents the total number of events of class C , and k represents the number of events that exhibit the edge. We apply a cutoff p -value of 0.05 in order to eliminate motifs with low sink mode bias. Note that the test is used as a filtering method but with the conservative p -value of 0.05 used to increase the sensitivity of putative discriminatory features that while being weak individually could provide a strong signal collectively.

For the continuous case, we apply the Student's t -test over the sets of edge weights from two network classes (e.g., Land and Offshore) to assess the edges whose average weights are significantly different between two network classes (modes).

3.4 Feature selection and ensemble classification

Once we have a set of sink-biased network motifs, we use them as features to build a feature selection-based ensemble of classifiers in order to forecast the track's sink mode at time $t_{start} + \Delta t$ for the track whose formation stage is defined by its source location at time t_{start} . For our target application, described in Section 2, Δt may range between 10 to 15 days.

The intuition behind building the ensemble is that the outcome of the track is likely dependent on a combination of the sink-biased network motifs. Moreover, it is likely not the only combination that is important for track source—sync dynamics. Therefore, an ensemble of well-performing predictors could be effectively used as a mechanism to enumerate all or most of such combinations of sink-biased network motifs (see, Figure 4 for an example of two combinations, one for SST and the other for SLP climate factors). Moreover, understanding the physical interpretation of such multi-variate groups of sink-biased motifs and their possible cross-talks may provide additional insights about the factors influencing the dynamics of extreme event tracks.

To provide this enumerative capability, we propose an iterative feature selection technique based on decision trees. We use the iterative decision tree method for feature selection for a number of reasons: (1) to enumerate the sets of possible discriminatory features for future insights about their physical relevance and (2) to effectively deal with underdetermined problem with a much larger number of features than events by identifying significantly fewer features in each set for the events available for model building.

Our strategy iteratively selects the set of network motif features that correspond to the internal nodes of a decision tree constructed using the C4.5 [Quinlan, 1993] algorithm. The resulting feature sets and data are then used with the traditional machine learning classification algorithm like C4.5 or generic SVM. The resulting classification models are then filtered based on their performance so that only the “best” AIC [Akaike, 1973] performing models could vote on the combined value for the sink mode. For simplicity, we utilize majority voting to combine the predictions of models in the ensemble.

Table 1: Performance on binary Sea Level Pressure (SLP) and Sea Surface Temperature (SST) data (*unweighted graph*), using Leave-One-Out (LOO) and 10-Fold cross-validation

	LOO:Binary			10-Fold:Binary	
	SLP	SST	SLP+SST	SLP	SST
Accuracy	0.88	0.90	0.92	0.90 ± 0.004	0.90 ± 0.005
Sensitivity	0.91	0.96	0.97	0.95 ± 0.005	0.97 ± 0.003
Specificity	0.77	0.76	0.81	0.80 ± 0.025	0.74 ± 0.030
Precision	0.90	0.90	0.92	0.92 ± 0.003	0.90 ± 0.005
F_1 -measure	0.90	0.93	0.94	0.93 ± 0.002	0.93 ± 0.003

Table 2: Performance on continuous Sea Level Pressure (SLP) and Sea Surface Temperature (SST) data (*weighted graph*), using Leave-One-Out (LOO) and 10-Fold cross-validation

	LOO:Continuous			10-Fold:Continuous	
	SLP	SST	SLP+SST	SLP	SST
Accuracy	0.90	0.86	0.92	0.88 ± 0.004	0.82 ± 0.004
Sensitivity	0.96	0.97	0.97	0.93 ± 0.006	0.94 ± 0.005
Specificity	0.78	0.61	0.81	0.75 ± 0.040	0.54 ± 0.027
Precision	0.91	0.85	0.92	0.91 ± 0.005	0.83 ± 0.003
F_1 -measure	0.93	0.91	0.94	0.92 ± 0.002	0.88 ± 0.002

4 Results and Discussion

Data

The North Atlantic hurricane data from 1950 to 2009 is obtained from the best track database (HURDAT) at the NCDC (National Climatic Data Center), covering the genesis region between 10–17.5°N and 345–300°E. We distinguish those hurricanes that strike land using the “Hit” feature of the HURDAT. We also include the HURDAT hurricanes that made landfall in Mexico in our analysis.

We use daily sea level pressure (SLP) and monthly sea surface temperature (SST) reanalysis data over the North Atlantic (10–37.5°N and 255–357.5°E) from NCEP/NCAR [Kalnay *et al.*, 1996; Kistler *et al.*, 2001]. The data is available at a horizontal and vertical resolution of $2.5^\circ \times 2.5^\circ$ for SLP and $2^\circ \times 2^\circ$ for SST. For each hurricane, we base our analysis on the preceding $\Delta W = 90$ days’ worth of data for SLP (daily averages) and 7 months’ data for SST (monthly averages).

Univariate vs. Multivariate, Weighted vs. Unweighted

Table 1 and Table 2 summarize the skill of the forecasting models to predict whether hurricane tracks forming at their genesis region in Western Africa will strike the land of North America or remain offshore. For model validation we employ both the leave-one-out (LOO) and 10-fold cross validation methodologies. For the latter, the mean of 10 models and their variance are reported. The tables show five skill metrics: accuracy, sensitivity, specificity, precision, and F_1 -measure. Note that the values of the four metrics reported in these tables are all based on the weighted majority voting result of the ensemble classifiers.

We build an ensemble of univariate classifiers to test the predictability of a specific variable, while we use an ensemble of multivariate classifiers to achieve the best overall performance. For univariate data, SST and SLP both perform well in both cases, with nearly 90% accuracy. SST performs slightly better on the unweighted graph data, while SLP per-

forms slightly better on the weighted graph data. Combining data for both variables and applying the multivariate classification, we achieve an accuracy of 92%, improving the accuracy over the single variate cases.

Besides accuracy, sensitivity, precision, and F_1 -measure of both univariate and multivariate cases are relatively high for both weighted and unweighted climate system networks. However, the specificity is slightly lower than the other four skill metrics. Specifically, in the case of univariate SST, there is a 15 to 20% difference in the specificity between the weighted and unweighted networks. This observation may be the indication that SST alone may not be as reliable as SLP or the two combined to predict the hurricane tracks if we employ the weighted data.

It is important to note that for the univariate case, the skill metrics, such as accuracy, sensitivity, precision, and F_1 -measure of both SLP and SST are relatively high (more than 85%) for both weighted and unweighted climate system networks. However, the specificity for both variables is relatively low (less than 80%). For binary classification, the accuracy of the classifier alone may not be sufficient to evaluate the predictability of a particular class. The specificity and sensitivity metrics in Tables 1 and 2 provide additional insights on the performance of the algorithm with respect to each class. The specificity is the fraction of the hurricanes that actually hit the land that the model is able to predict correctly. The sensitivity is the fraction of the curving hurricanes that the model is able to predict correctly. Higher specificity over sensitivity means better recognition of landfalling hurricane tracks.

In addition, in the case of SST, there is a 15% difference in the specificity between the weighted and unweighted networks in the leave-one-out experiments. This observation may be the indication that SST may not be as reliable as SLP in predicting the hurricane tracks. Therefore, for variable SST, the unweighted classifiers have advantage of being more accurate about predicting whether a hurricane will hit

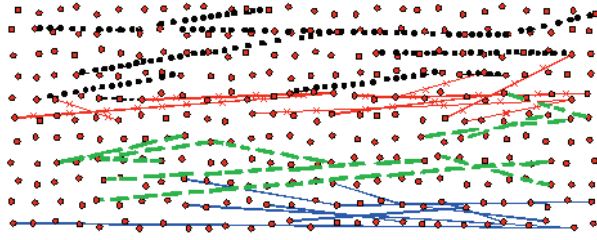


Figure 3: Correlations between the nodes with p -value increases in the order of black/dotted ($0 < p \leq 0.005$), red/(-x-) ($0.005 < p \leq 0.01$), green/dashed ($0.01 < p \leq 0.015$), and blue/solid ($0.015 < p \leq 0.02$)

the land over the weighted graph. The difference in performance between SST- and SLP-based classifiers could also be attributed to a number of other factors, including differences in time resolution. Future sensitivity analysis would aim to explain the reliability of such factors.

In the case of the unweighted graph data, applying the hypergeometric test to the data provides the most significantly biased motifs between two networks associated with each sink mode, so choosing a very small p -value may seem to be an appropriate choice for feature selection. However, features with similar p -values frequently come from the same spatial region, as seen in Figure 3. Figure 3 shows the biased network motifs (edges) for the SLP networks, color-coded by p -value. The highlighted edges in the figure have p -value less than 0.02, and the p -value increases in the order of black, red, green, and blue.

As geographic data often exhibits strong spatial correlation, though, choosing a very small p -value mainly captures redundant signals and reduces classification accuracy. Thus, we choose a standard p -value threshold of 0.05, selecting 150 biased edges from the SLP networks constructed. The networks were initially complete graphs consisting of $O(n^2)$ edges, where n is the number of nodes in the geographic region network (i.e., $n = 300$), and 85 to 90 percent of the edges were removed through the correlation analysis described in Section 3.2. This feature set is broad enough to capture biased network motifs across several geographic regions, while still representing a strong bias towards a particular sink mode. As a result, the supervised decision tree based feature selection chooses the features from geographically distinct locations, as can be seen in Figure 4.

Comparison with Alternative Classifiers

Due to inherently ensemble-driven methodology for building the forecasting model, as rationalized in Section 3.4, in this study, we compare two ensemble classification techniques to our iterative decision tree construction procedure on the biased network motifs for the SLP data. The results, which appear in Figure 5, show that the ensemble based on our decision tree feature selection algorithm has the best accuracy among these three.

Since the number of hurricane tracks may be signifi-

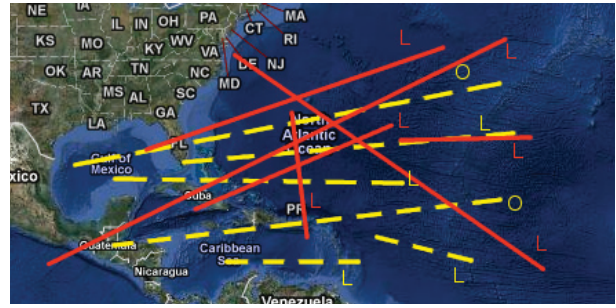


Figure 4: Features that emerge in a model; SLP (yellow/dashed) and SST (red/solid) positively correlated teleconnections; L —biased toward land-hitting tracks and O —biased toward offshore remaining hurricane tracks.

cantly smaller than the number of features (biased network motifs), we have a problem of building a classifier for a high-dimensional dataset. A support vector machine (SVM) [Cortes and Vapnik, 1995] is a classification learning technique that has been successfully used for high-dimensional datasets [Li *et al.*, 2000]. For this experiment, we employ an SVM feature selection method [Guyon *et al.*, 2002] to select a subset of features and then build an ensemble of SVM classifiers. An alternative technique for building a classifier ensemble for high-dimensional datasets is random forest [Breiman, 2001] [Ho, 1998]. In random forest, the ensemble consists of a set of “base” classifiers, each one constructed from all data points and a small subset of features sampled with replacement from the entire feature set. For purposes of comparison, we use decision tree as the base classifier for random forest.

Figure 5 shows our results comparing the performance of SVM feature selection, random forests, and decision tree feature selection. The SVM feature selection technique produces 7 “base” classifiers, while our decision tree feature selection technique produces 8 “base” classifiers. Since the decision tree feature selection chooses 5–16 features for unweighted dataset and 11–15 features for weighted datasets, we perform the experiments for random forest by selecting 8 sets of 5–16 features for unweighted and 8 sets of 11–15 features for weighted datasets. The number of features selected by SVM is about 25 for each “base” classifier. The accuracy reported for random forest is the highest accuracy obtained by varying the number of features in these ranges. The results show that the decision tree feature selection yields higher accuracies than SVM feature selection and random forest for both leave-one-out (LOO) and 10-fold cross-validation for both weighted and unweighted datasets by 10–25%.

Physical Interpretation

In Figure 4, we highlight the biased network features from one of the classifiers, built by applying our decision tree feature selection technique to the biased network motifs from the unweighted SLP and SST networks. In this set, a pattern can be clearly seen within the given feature set: both of the classifiers highlight the positively correlated teleconnections between the genesis and landfall regions of the hurricanes. We also observe this pattern in several of the other classifiers.

Our analysis of the nearly east-oriented SLP edges suggests horizontal pressure gradient configuration in the same direction. Although wind flow anomaly is not directly incorporated into our model, based on Buys Ballot's law, this pressure gradient would be associated with wind flow in the perpendicular direction, i.e., in the north-south direction. The actual sign of the anomaly flow would be dictated by the sign of the pressure gradient along the SLP edge. Onshore wind anomaly flow would promote favorable conditions for landfall, while opposite flow anomaly conditions would be more favorable for hurricanes trajectories that result in no-landfall. A similar argument may be considered for interpreting the significance of the orientation of the SST edges since there is a strong but non-trivial relationship between SLP and SST.

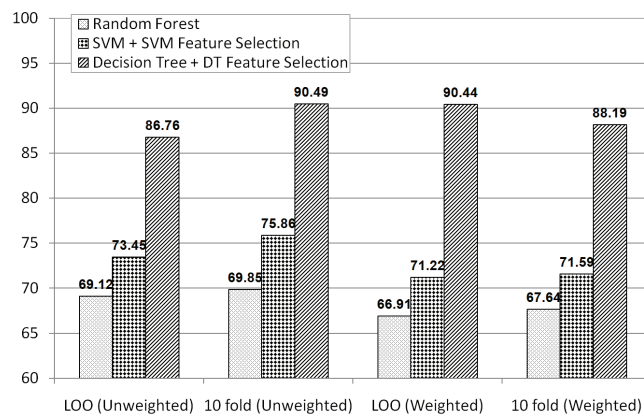


Figure 5: Comparison of different classification algorithms.

5 Conclusion

Performing analysis and prediction on dynamically changing complex systems is a challenging problem for statistical machine learning. For such systems, using a network may be a natural choice to condense and summarize the relationships inherent in the data. In this paper, we propose a technique to analyze the underlying dynamics of complex systems in order to predict the sink mode of extreme events like hurricanes or forest fires. By identifying the spatio-temporal relationships that characterize the sink modes of past events, we are able to construct classifiers capable of predicting the future track of an emerging extreme event from current data. Our results demonstrate that our technique is able to predict the landfall of North Atlantic hurricanes with up to 92% classification accuracy, demonstrating the practical value of our findings.

In the future, we plan to generalize edge-based features to subgraph-based features in climate networks. Inspired by the value of such network motifs in other application domains, such as biology, we envision that such biased subgraph-based features could not only improve the skill of our models but also improve our understanding of complex physical systems, such as climate. The current study primarily focuses on the North Atlantic region, however, the proposed methodology is not specific to a particular region and could be viewed as a general methodology. As future work, the scope of the study

will be extended to other regions subject to similar hurricane development. Finally, since we deal with a highly underdetermined problem, additional performance metrics more suitable for small sample size evaluation (see [Braga-Neto and Dougherty, 2004]) will be considered as part of our future work.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. 1029711 and by the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy. This work was supported in part at Northwestern University by NSF CCF-1029166 (expeditions), DOE DE-SC0005340, NSF IIS-0905205, and DOE DE-FG02-08ER25848. Oak Ridge National Laboratory is managed by UT-Battelle for the LLC U.S. DOE under contract no. DEAC05-00OR22725.

References

- [Akaike, 1973] H. Akaike. Information theory and an extension of the maximum likelihood principle. In *2nd Int. Symp. Info. Theory*, pages 267–281, 1973.
- [Braga-Neto and Dougherty, 2004] Ulisses M. Braga-Neto and Edward R. Dougherty. Is cross-validation valid for small-sample microarray classification? *Bioinformatics*, 20(3):374–380, 2004.
- [Breiman, 2001] Leo Breiman. Random forests. *Mach. Learn.*, 45:5–32, October 2001.
- [Cortes and Vapnik, 1995] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20:273–297, 1995.
- [Donges et al., 2009] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths. The backbone of the climate network. *Europhysics Letters*, 87:48007, August 2009.
- [Elsner et al., 2000] J. B. Elsner, K.-B. Liu, and B. Kocher. Spatial variations in major U.S. hurricane activity: Statistics and a physical mechanism. *Journal of Climate*, 13:2293–2305, 2000.
- [Elsner, 2003] J. B. Elsner. Tracking hurricanes. *Bulletin of the American Meteorological Society*, 84:353–356, 2003.
- [Gaffney et al., 2007] S. Gaffney, A. Robertson, P. Smyth, S. Camargo, and M. Ghil. Probabilistic clustering of extratropical cyclones using regression mixture models. *Climate Dynamics*, 29:423–440, 2007.
- [Guyon et al., 2002] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46:389–422, 2002. 10.1023/A:1012487302797.
- [Harr and Elsberry, 1991] P. A. Harr and R. L. Elsberry. Tropical cyclone track characteristics as a function of large-scale circulation anomalies. *Monthly Weather Review*, 119:1448, 1991.
- [Ho, 1998] T.K. Ho. The random subspace method for constructing decision forests. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(8):832–844, 1998.

- [Inokuch *et al.*, 2000] A. Inokuch, T. Washio, and H. Motoda. An Apriori-based algorithm for mining frequent substructures from graph data. In *Proc. of the 4th European Conference on Principles of Data Mining and Knowledge Discovery*, pages 13–23, Lyon, France, 2000.
- [Kalnay *et al.*, 1996] E. Kalnay, M. Kanamitsu, and et al. The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, 77:437–472, 1996.
- [Kistler *et al.*, 2001] R. Kistler, E. Kalnay, and et al. The NCEP-NCAR 50-Year Reanalysis. *Bulletin of the American Meteorological Society*, 82:247–267, February 2001.
- [Kossin *et al.*, 2010] J. P. Kossin, S. J. Camargo, and M. Sitkowski. Climate modulation of North Atlantic hurricane tracks. *Journal of Climate*, 23(11):3057–3076, 2010.
- [Kuramochi and Karypis, 2001] K. Kuramochi and G. Karypis. Frequent subgraph discovery. In *First IEEE Int’l Conference on Data Mining*, page 313, Los Alamos, CA, USA, 2001. IEEE Computer Society.
- [Li *et al.*, 2000] Y. Li, S. Gong, and H. Liddell. Support vector regression and classification based multi-view face detection and recognition. 2000.
- [Nakamura *et al.*, 2009] J. Nakamura, U. Lall, Y. Kushnir, and S. J. Camargo. Classifying North Atlantic tropical cyclone tracks by mass moments. *Journal of Climate*, 22(20):5481–5494, 2009.
- [Neumann, 1972] C. J. Neumann. An alternate to the hurricane tropical cyclone forecast system. *NOAA Tech. Memo.*, NWS(SR-62):32, 1972.
- [Quinlan, 1993] J.R. Quinlan. *C 4.5: Programs for machine learning*. 1993.
- [Shah *et al.*, 2010] N. Shah, Y. Shpanskaya, C-S Chang, S-H Ku, A. V. Melechko, and N. F. Samatova. Automatic and statistically robust spatio-temporal detection and tracking of fusion plasma turbulent fronts. In *SciDAC 2010*, 2010.
- [Tsonis and Roebber, 2004] A. A. Tsonis and P. J. Roebber. The architecture of the climate network. *Physica A: Statistical and Theoretical Physics*, 333:497 – 504, 2004.
- [Tuleya and Ross, 1995] R. E. Tuleya and R. J. Ross. Improvements in the gfdl hurricane prediction system. *Mon. Wea. Rev.*, 123:27912801., 1995.
- [Yan and Han, 2002] X. Yan and J. Han. gSpan: Graph-Based substructure pattern mining. In *Proceedings of the 2002 IEEE International Conference on Data Mining*, page 721. IEEE Computer Society, 2002.