2D.7 SUPPORT VECTOR MACHINE TECHNIQUES TO PREDICT TROPICAL CYCLONE REINTENSIFICATION FOLLOWING EXTRATROPICAL TRANSITION

Steven R. Felker[1], J. Scott Tyo[2], Elizabeth A. Ritchie[1], and Israel Vaughn[2]
[1] Department of Atmospheric Sciences, University of Arizona [2] College of Optical Sciences, University of Arizona

1. INTRODUCTION

Extratropical transition (ET) is a complex, multi-stage physical process during which a tropical cyclone (TC) interacts with the midlatitude environment and evolves from a warm-cored tropical system into a cold-cored midlatitude cyclone. A full review of extratropical transition and recent research is given in Jones et al. (2003). Although every tropical cyclone undergoes a unique evolution during and after the ET process, transitioning storms are often placed into one of two categories, intensifiers or dissipaters, based on their relative intensity evolution following ET (e.g., Demirci et al. 2007; Kofron et al. 2010). The transition process itself has been defined by Klein et al. (2000) as a dominantly two-phase process: the first being the transformation stage, and the second being the reintensification phase. Klein et al. (2000) found that intensifying and dissipating storms are especially difficult to distinguish during the first stage of ET, since all storms are generally moving into the midlatitude environment and weakening as they interact with the high levels of vertical wind shear associated with the midlatitude westerlies and low-level baroclinic zone in the midlatitudes (Figure 1). Thus the structure of all the storms moving poleward tends to appear very similar prior to the second stage of ET.

Despite this lack of differentiation between dissipating and re-intensifying classes in the storms themselves prior to the end of phase 1 of ET, recent studies have suggested that the spatial relationships between the TC and midlatitude upper-level trough are able to provide a capacity to predict the outcome of ET (Harr et al. 2000; Ritchie and Elsberry 2007). Based on these results, researchers have begun examining the practical application of these findings toward post-ET intensification prediction. Demirci et al. (2007) examined the problem from a spatial and spatiotemporal projection pursuit approach. The method involved empirical orthogonal function (EOF) analysis of the Navy's Operational Global Assimilation and Prediction System (NOGAPS) 500 hPa geopotential height analyses centered on each storm every 12 hours from 48 hours before ET to 48 hours after ET. Spatiotemporal techniques allowed for peak performance ~82% using fields from 12 and 24 hours before ET together, nearly matching the performance of the NOGAPS model itself at 24 hours prior to ET.

* Corresponding author address: Steven R Felker, Department of Atmospheric Sciences, Tucson, AZ, Email: felker@atmo.arizona.edu

In this study, we have built on the work of Demirci et al. (2007) by looking at transitioning storms in high dimensional spaces and using advanced support vector machine (SVM) classification techniques in the hope of further improving forecasting range and performance with respect to post-ET intensity change classification.

2. METHODOLOGY

Tracking Individual Storms and ET Time Determination

Overall, 108 western North Pacific tropical cyclones which underwent extratropical transition between January, 2000 and December, 2008 were identified and processed. For each storm, ET time was defined as the first time that the storm appeared as an open wave on the mid-latitude trough in the Global Forecast System (GFS) Final Analysis (FNL) 500-hPa geopotential height analyses using a 20 m contour (e.g. Figure 1).

After an ET time was established for each storm, the storm was then tracked in the model analysis from 72 hours before the ET time (ET-72) to 72 hours after the ET time (ET+72) at an interval of 6 hours. The storm center was determined by locating the lowest mean sea level pressure (MSLP) in an area surrounding the storm. These tracking data were archived and used to center the model data around the storm center at a distance of +/- 30 degrees longitude, and +/- 25 degrees latitude, and to save pressure-level fields of all available atmospheric model variables for each storm at each relative ET time. Thus, each storm variable field for each relative ET time was composed of 3111 (61 x 51) observations, corresponding to the storm itself and its surrounding environment before, during, and after ET.

All basic raw model output fields from the GFS FNL model were tested using an SVM classifier and 5 testing sets of storms. In particular, 850 hPa potential temperature fields captured the structure of the TC itself, but perhaps more importantly also captured the location and intensity of frontal zones and associated mid-latitude features (Figure 2). Therefore these fields were chosen as the inputs of choice for the results shown in this paper.

Storm Classification

Classification for each tropical cyclone was determined by observing changes in the mean sea level pressure (MSLP) at the center of each storm ($MSLP_{TC}$) through its ET evolution, as well as changes in the MSLP surrounding the storm ($MSLP_{ENV}$). Storms for which $MSLP_{ENV} - MSLP_{TC}$ increased more than 3 hPa between ET +/- 6 hours and ET+54 +/- 18 hours were categorized as positive storms; that is, storms which

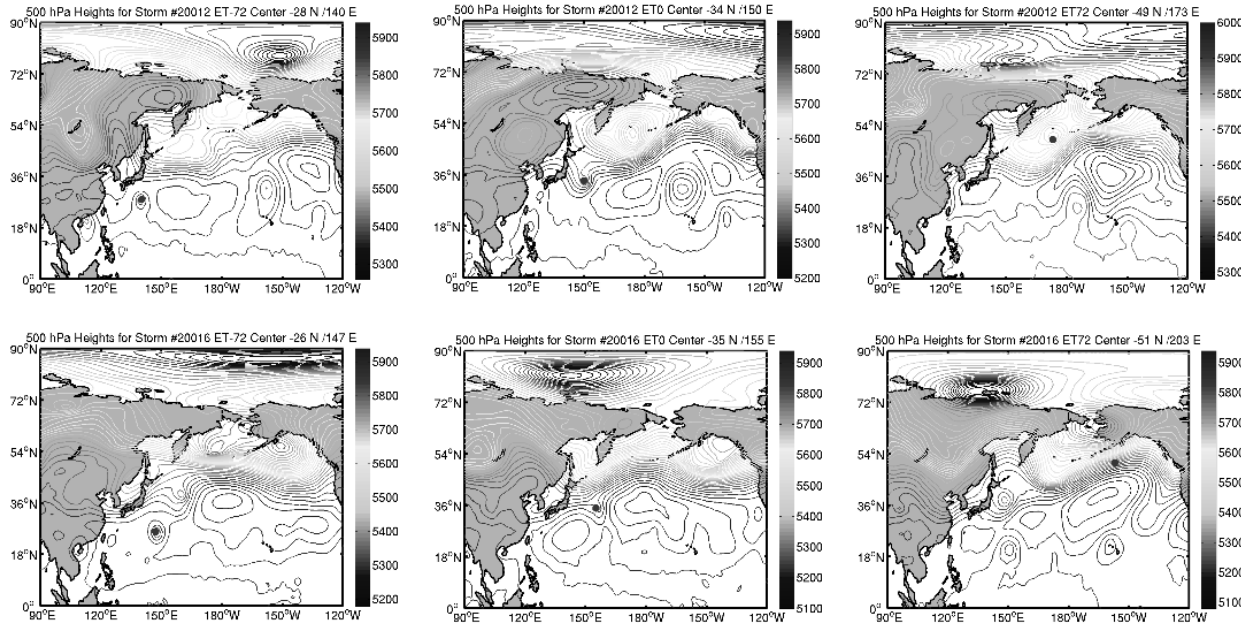

Figure 1. Panel of 500-hPa geopotential heights at ET-72 (left), ET (middle), and ET+72 (right) for Kong-Rey (2001 - reintensifyer) and Wutip (2001 - dissipater). In each panel, the black dot represents the determined storm center based on storm tracking and MSLP.

intensified substantially post-ET. All other storms were categorized as dissipating (negative) storms. This system of categorization resulted in 53 positive storms and 55 negative storms overall.

Data Pre-processing

Due to the large amount of data available in the form of 3111 spatial data points, the first pre-processing step was to reduce the dimensionality of the inputs using feature selection, while attempting to best retain information content that would be useful for classification. For this task the Correlation-based Feature Selection (CFS) method developed by Hall (1999) was used. The main premise behind this selection method is that the best features for classification are those which are most highly correlated with the classes (intensifiers and dissipaters), and at the same time least correlated with other features. For this study, a forward selection version of the CFS method was used, where the system first chooses the best individual feature based on the metric:

$$M_s = \frac{k \cdot \bar{r}_{cf}}{\sqrt{k + k(k-1) \cdot \bar{r}_{ff}}} \qquad (1)$$

where M_s is the merit metric, k is the number of features in the subset, r_{cf} is the outer correlation between features and classes, and r_{ff} is the inner correlation between features. Subsequently, the first 'best' feature is combined with each other feature, and M_s is recalculated for each. The feature combination that maximizes M_s is then chosen iteratively up until a chosen number of input features is reached. In the interest of further reducing dimensionality and preventing over-training, as well as to promote consistency throughout all relative ET time, only 20 features were used as inputs into the classification system at all times during ET evolution.

The second pre-processing step involved dividing the storms into training, validation, and testing data subsets and centering and scaling all of the variable inputs between storms using the mean and standard deviation within each subset. To evaluate the classification system, K-fold cross-validation was used as a means of producing training and validation sets. This method is appealing because it allows all inputs to be used both in the training and in the validation of the classification system. For all of the evaluations, 20% (21 storms) of the data were held out for final testing and 10 folds were used for training and validation (remaining 87 storms).

Support Vector Machines and Model Selection

Support vector machines (SVMs) are a specific type of classifier that combines the concepts of a maximum margin with kernel techniques in an attempt to classify sparse data sets in a manner that maximizes generalization capacity. The basic premise is to find a hyperplane that perfectly classifies all inputs, while optimizing the normal margin between the classes to be as wide as possible (for a full review, see Burges,1998).

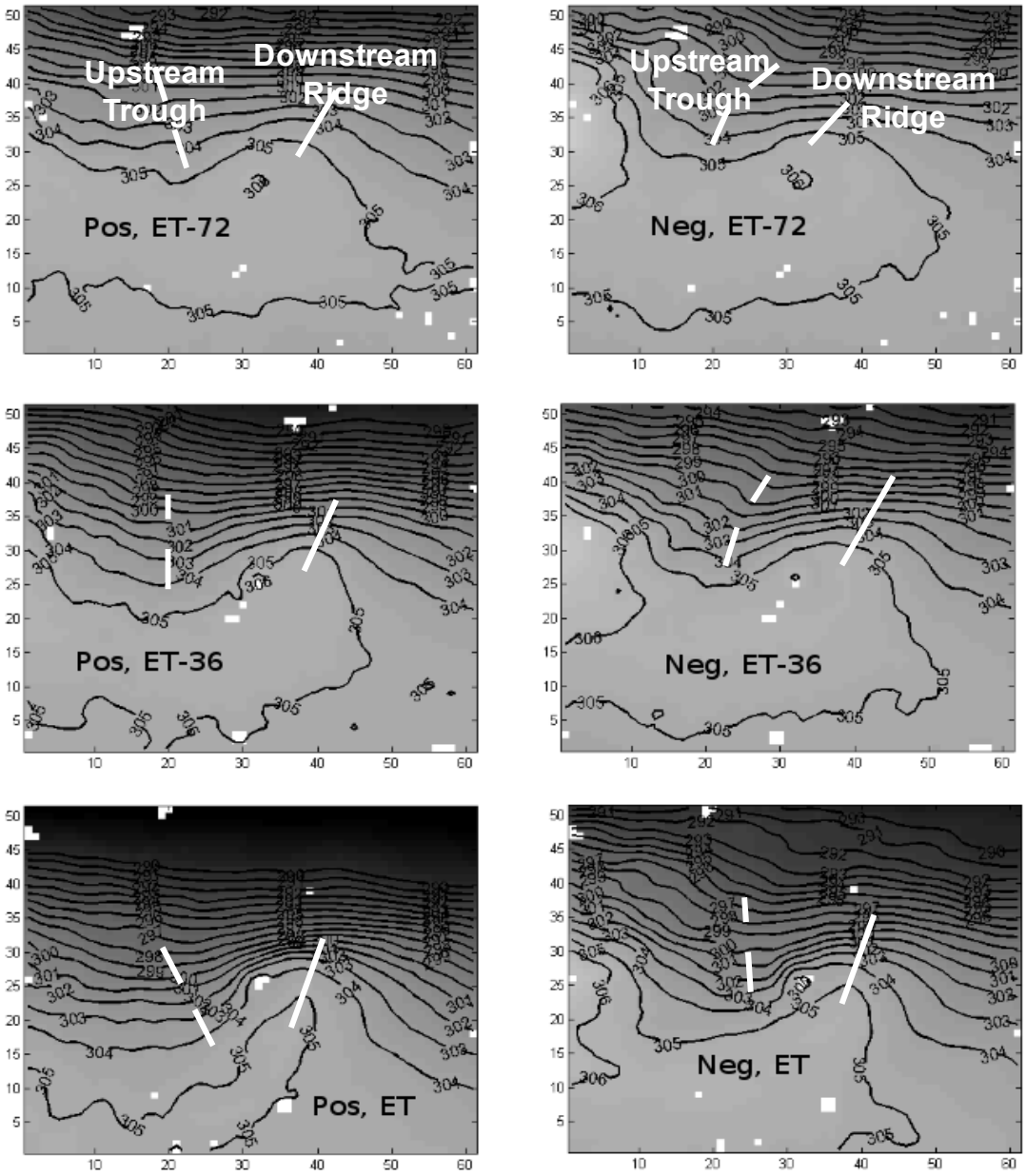

Figure 2. Mean 850-hPa potential temperature field for 53 post-ET intensifying storms (left) and 55 post-ET dissipating storms (right) at 72 hours prior to ET (top), 36 hours prior to ET (middle) and at ET (bottom). All frames are centered on the TC. The upstream trough (dashed line) and downstream ridge (solid line) generally associated with ET are indicated. Whitened pixels represent those chosen by the CFS feature selection system as most useful for class discrimination.

In this study, an RBF kernel function is used in the SVM calculations:

$$\mathbf{K}(x_i, x_j) = e^{\frac{\|x_i - x_j\|^2}{2\times\sigma^2}} \qquad (2)$$

where x_i and x_j are individual observation vectors and σ is a parameter, which controls the areal influence of each observation. Both the margin parameter and sigma influence the number of support vectors, which are chosen by the trained SVM. The support vectors are those observations (storms in this case), which, if removed, would alter the hyperplane solution. The best model parameters are chosen in the following manner: (1) iteratively test classification performance for values of C (1e-6, 1e-5, ..., 1e-1) and RBF kernel parameter sigma (2, 2.25, ..., 10), using a 2-norm soft-margin SVM system on each of 10 cross-validation folds; (2) determine mean estimated true error based on

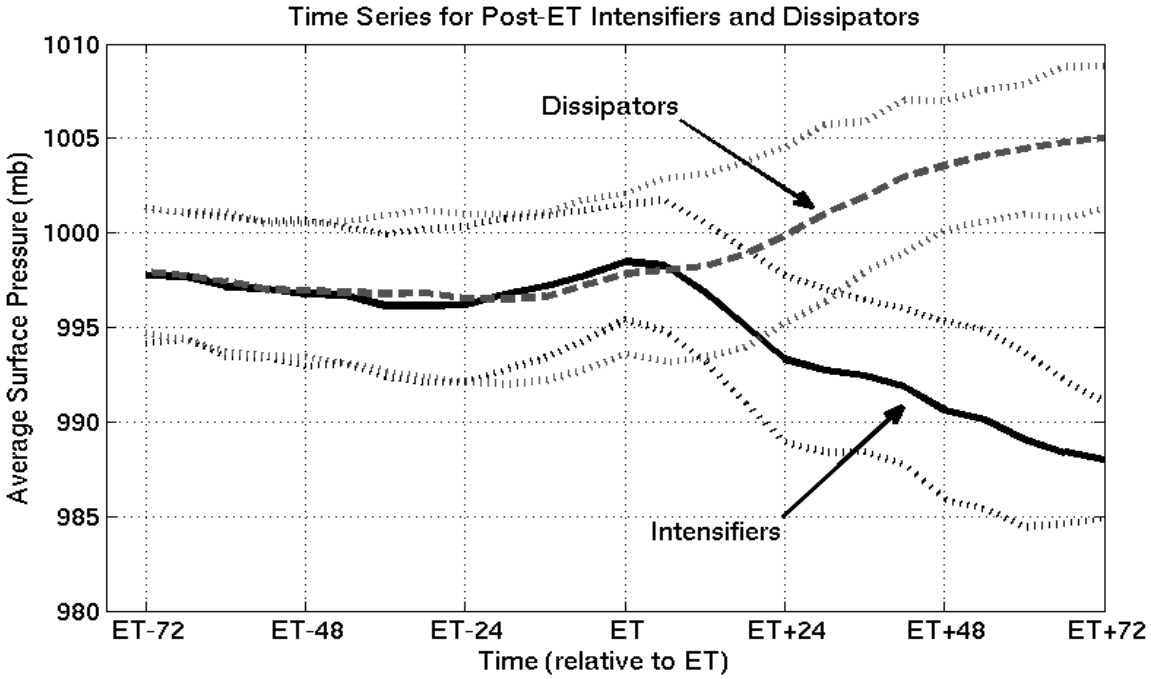

Figure 3. Time series of average intensity for post-ET intensifiers (solid black line) and dissipaters (dashed gray line). One standard deviation is plotted for each set. Note that, on average, the two classes of storms undergo a very similar intensity evolution before ET, and do not fully statistically diverge until ~ET+30 hours.

generalization for each parameter combination for each fold, calculated as follows:

$$\overline{E} = \frac{1}{k} \times \sum_{i=1}^{k} E_i \qquad (3)$$

where k is the number of folds and E_i is the generalization error for each fold; and (3) choose 'best' model parameters based on mean estimated true error, and use these model parameters to train an SVM and classify the held out test data. From this procedure, ideal model parameters, mean training set error, mean estimated true error, and actual test set error are determined at each relative ET time (ET-72 to ET+72).

3. RESULTS

A time-series analysis of class-averaged MSLP for intensifying versus dissipating storms (Figure 3) suggests that the two classes display almost identical intensity evolution during the 72 hours leading up to ET. This trend is in agreement with the findings of Klein et al. (2000). In both classes, storms on average tend to slowly weaken from ~ET-24 to ET time, before quickly diverging in behavior after ET. Note that by ~30 hours after ET, the two classes have statistically diverged. These results both suggest that the method of storm classification used in this study is reasonable, and that, as previously suggested, storms which intensify and weaken post-ET are very difficult to differentiate prior to the completion of the ET process using only data about their intensities leading up to ET.

Figure 2 shows sample features that were chosen prior to ET for 850-hPa potential temperature fields, overlaid onto intensifier and dissipater mean fields at ET-72, ET-36 and ET, respectively. Note that the correlation-based feature selection method tends to select features that describe the mid-latitude environment surrounding the storms over features of the storms themselves, particularly prior to ET. Observation locations that correspond to the relative position and intensity of the midlatitude trough to the northwest of the storms tend to be chosen as the most useful for differentiating classes at ET-72. Interestingly, data to the south and southeast of the storm centers are also chosen, representing differences in thermodynamic properties of the tropics themselves. Despite their lack of selection here, features of the downstream ridge have been identified in past studies as important to the eventual outcome of ET. Since the downstream ridge is usually identified as building ahead of the TC from the anticyclonic outflow, its presence and impact may be better identified in upper-level fields.

Features from the downstream ridge are, however, chosen by the CFS system at ET-36, suggesting that differences in these features, such as the low-level frontal development, between positive and negative storms may become more evident as ET time approaches. Features near the storm center, particularly to the southwest where the cold frontal development is occurring, also become more important in differentiating classes at this time. At ET time itself, features closest to the storm center begin to become the best for class

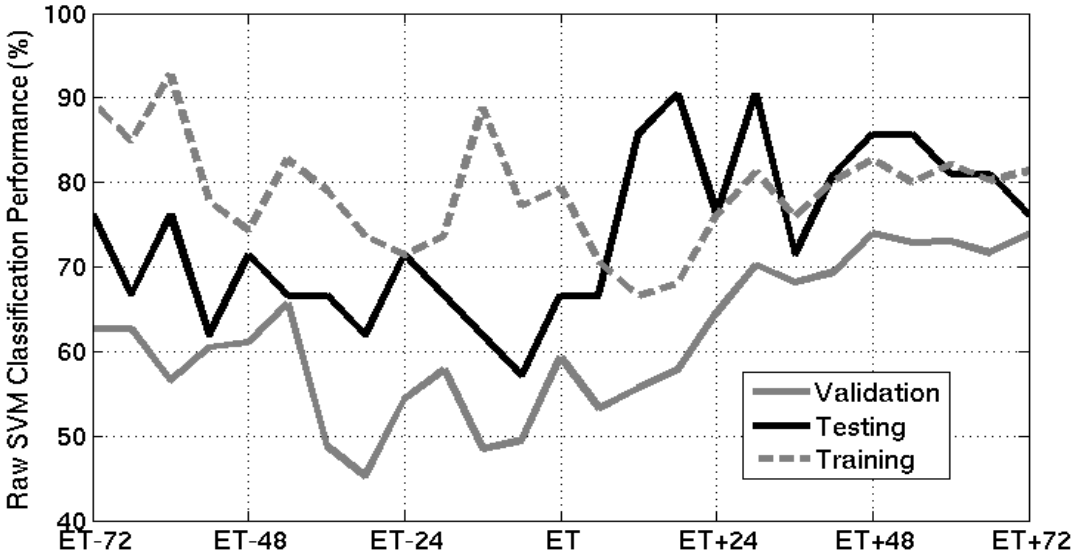

Figure 4. Time series of raw SVM classifier performance for 850-hPa potential temperature data and 20 features chosen by CFS. Validation results are plotted in solid gray and provide an estimated model error based on a 10-fold K-fold cross-validation procedure. Training results are plotted in dashed gray and provide a metric of how well the model is forced to fit the original data from which it is trained. Testing results are plotted in solid black and represent model performance on a randomly selected fully independent test set of 21 storms (20% of all data).

differentiation, probably because by this time the details of the trough interaction with the TC itself are the important differentiators. This sequence completes a movement from concentration on the upstream trough, to the downstream ridge, to the storm itself from ET-72 to ET time.

Overall classification performances (for the test set and validation sets) for 850-hPa potential temperatures through time (ET-72 to ET+72) are given in figure 4. These performance numbers are based on raw outputs of the SVM classifier. It can be seen that the testing data have slightly poorer performance than the training data prior to the ET time and are comparable or even slightly better after the ET time. This is not an unexpected result, and improvements over this may be obtained simply by increasing the training set to ensure that a more representative sample of ET cases exist in the training statistics. Examination of the receiver operator characteristics (ROC) curves for times leading up to ET (figure 5) suggests that better performance could also potentially be obtained by changing the threshold of the classifier. The raw SVM output performance reflects only a single point on the ROC curve, and purposefully allowing a certain amount of false detection may provide for a more robust detection system overall. For example, one could potentially achieve a positive detection rate of 80%, at 72 h prior to the ET time if a false alarm rate of 27% were an acceptable risk (Fig. 7). Despite this fact, even using raw SVM output, this prediction system correctly classifies ~76% of storms (16 out of 21) at 72 hours prior to ET.

4. DISCUSSION

Our SVM-based classification system for post-ET intensity forecasting displays encouraging results, with over three quarters of storms being correctly classified at 72 hours prior to ET in the randomly chosen testing set. Performance is consistently in the range of 60-75% leading up to ET, with the exception of the few times immediately before the start of the ET process. Performance falls off at this point, which is potentially an artifact of the lack of spatial pattern differences at these times. Once the storms are beginning the ET process, both post-ET intensifiers and dissipators are both becoming features on the midlatitude trough and generally weakening and undergoing structural changes associated with interaction with the baroclinic zone. The decrease in performance is thus likely due to the loss of large-scale spatial pattern differences at this point, and a movement toward differences in TC structure itself. However, as expected, model performance quickly increases after the completion of ET (around ET+12), as the process has completed and the storm is now either intensifying or dissipating. At this point, the pattern recognition system is able to focus on features within and immediately surrounding the TCs which suggest considerable differences in relative size and strength of the disturbance between classes.

Of the 4 storms that are consistently

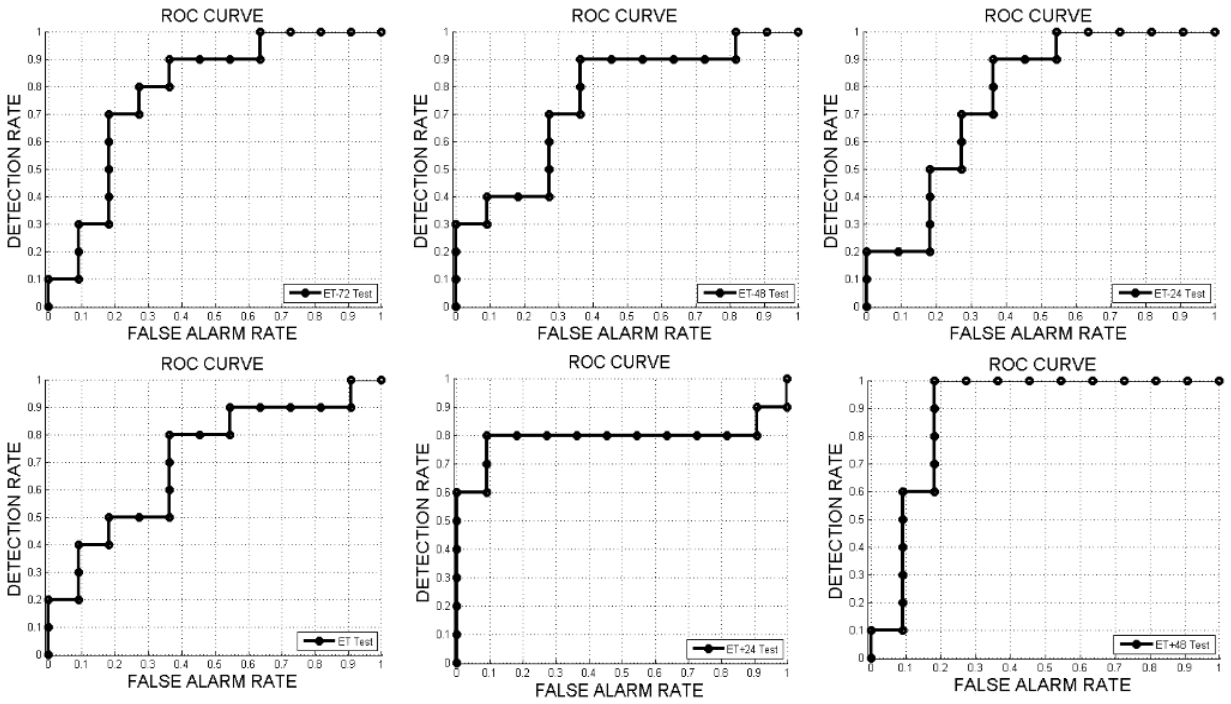

Figure 5. Receiver Operator Characteristics (ROC) curve for 850-hPa potential temperature test storms at 72 hours prior to ET (upper left) through 48 hours after ET (lower right) at 24 hour intervals.

misclassified leading up to ET (1 intensifier and 3 dissipators), 2 seem to undergo intensity changes that are fairly representative of dissipaters, and it is not yet clear why they were misclassified at so many time periods. One other dissipater actually re-intensifies almost enough to have fallen into the re-intensification class and so the classifier may have difficulty with this storm. The one misclassified reintensifying TC actually weakens for a considerable time after ET, and only begins to reintensify around ET+48. Therefore it is not surprising that our pattern recognition system has some difficulty in classifying it as a re-intensifier prior to ET.

Although the method presented in this paper provides encouraging results for pattern recognition use in the forecasting of post-ET re-intensification, there are several limitations which will be addressed in the future. First, the overall system is a highly non-stationary system (the Earth's atmosphere), and seasonal and inter-annual variability of the input variables (850-hPa potential temperature) are not accounted for. Therefore the mean which is removed from the positive and negative storms individually may not be representative for many off-season storms or storms from strong El Niño or La Niña years. A movement toward either a removal of a seasonal mean or inclusion of temporal information as a feature in the classification system is needed, and is likely to improve the system's performance. Second, it will be vital to test more atmospheric variables both individually and in combination to see if there are certain fields which provide even better predictive ability. Lastly, the system simply needs a larger set of training data for improved performance. The CFS method has suggested that ~40 features are useful in class discrimination, but this number is simply unreasonable in a pattern recognition problem with under 100 training samples. It will therefore be vital to continue growing the training set as new model data becomes available for TCs which have undergone ET.

Acknowledgements: The GFS-FNL analyses were kindly provided by Mr. Bob Creasey at the Department of Meteorology, Naval Postgraduate School, Monterey CA. This research was supported by the National Science Foundation under grant ATM-0730079

5. REFERENCES

Burges, C. J.C., 1998: A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discover,* **2,** 121-167.

Demirci, O., J.S. Tyo, and E.A. Ritchie, 2007: Spatial and Spatiotemporal Projection Pursuit Techniques to Predict the Extratropical Transition of Tropical Cyclones. *IEEE Trans. Geosci. Remote Sens.,* **45**, 418-425.

Hall, M.A., 1999: Correlation-based Feature Selection for Machine Learning. *Thesis, University of Waikato.*

Harr, P.A., and R.L. Elsberry, 2000: Extratropical

Transition of Tropical Cyclones over the Western North Pacific. Part I: Evolution of Structural Characteristics during the Transition Process. *Mon. Wea. Rev.*, **128**, 2613–2633.

Jones, S.C., P.A. Harr, J. Abraham, L.F. Bosart, P.J. Bowyer, J.L. Evans, D.E. Hanley, B.N. Hanstrum, R.E. Hart, F. Lalaurette, M.R. Sinclair, R.K. Smith, and C. Thorncroft, 2003: The Extratropical Transition of Tropical Cyclones: Forecast Challenges, Current Understanding, and Future Directions. *Wea. Forecasting*, **18**, 1052–1092.

Klein, P.M., P.A. Harr, and R.L. Elsberry, 2000: Extratropical Transition of Western North Pacific Tropical Cyclones: An Overview and Conceptual Model of the Transformation Stage. *Wea. Forecasting*, **15**, 373–395.

Kofron, D. E., E. A. Ritchie, and J. S. Tyo, 2010: Determination of a consistent time for the extratropicl transition of tropical cyclones. Part I: Examination of Existing Methods for Finding "ET-Time", *Mon. Wea. Rev. (In Review).*

Ritchie, E.A., and R.L. Elsberry, 2007: Simulations of the Extratropical Transition of Tropical Cyclones: Phasing between the Upper-Level Trough and Tropical Cyclones. *Mon. Wea. Rev.*, **135**, 862–876.