A Fuzzy Logic System for Predicting Hurricane Intensity in the Eastern North Pacific*

Jonathan L. Vigh

SOARS Protégé

Science Research Mentor:
Kevin Petty

Research Applications Program
National Center for Atmospheric Research
Boulder, CO

11 August 2000

UCAR-SOARS
University Corporation for Atmospheric Research
Significant Opportunities in Atmospheric Research and Science
Sponsored by the National Science Foundation

* A CD-ROM supplement is available that contains the development and validation data, the programs used to reformat, analyze, and normalize the data, and the Simulink models

ABSTRACT

The primary goal of this research is to examine the efficacy of using a fuzzy logic/adaptive weighting (FLAW) technique to predict hurricane intensity change in the eastern North Pacific (ENP) basin. The intensity change forecasts, for 12-hour intervals ranging from 12 to 72 hours, produced by a model using FLAW are compared to a model developed using the standard statistical technique of multiple linear regression (MLR).

To this end, climatology and persistence variables, as well as observed intensity changes, are computed and extracted from the National Hurricane Center (NHC) best-track dataset and the weekly National Oceanic and Atmospheric Administration (NOAA) global optimum interpolation weekly sea surface temperature dataset for the years of 1982-99. The climatology and persistence predictors used for both models include the previous 6-hour intensity change, latitude of the hurricane center, current intensity at the time of the observation, and sea surface temperature interpolated to the time and location of the observed hurricane center. The FLAW technique is based on the Standard Additive Method (SAM) developed by Kosko. The method adapts the weights given to predictors to decrease the difference between the forecast and observed value.

Preliminary results suggest that the FLAW model produces errors comparable in magnitude to the MLR model. The bias is significantly less for the 36-, 48-, 60-, and 72-hour forecast periods. Several case studies show that the model is able to adapt the weighting appropriately when the one or more predictors become over-dominant. With more work and the inclusion of synoptic predictors, this method may eventually offer improved hurricane intensity change guidance for forecasters, thus reducing the loss of lives and property.

1. Introduction

W. Gray (Gray, et. al., 1997) has suggested that the Atlantic basin is now coming out of a multi-decade lull in hurricane activity. During this lull, the coastal population of the United States has grown exponentially, especially in Florida (Pielke and Landsea, 1998), exposing millions more people to the threat of landfalling hurricanes. The time needed to evacuate some highly vulnerable coastal areas, such as the Tampa Bay area, New York City, the Florida Keys, and especially New Orleans is on the order of 36-48 hours, while the current time frame for meaningful hurricane intensity change guidance is on the order of 12-36 hours. This means that effective early warnings and disaster mitigation planning for certain vulnerable and populated areas are difficult with the current intensity forecasting skill. The significant downward trend in fatalities experienced from the 1930’s to the present is not likely to continue unless the time envelope for meaningful intensity change guidance can be extended out to the 48- to 60-hour time frame.

L. Avila (1998), hurricane specialist at the Tropical Prediction Center/National Hurricane Center (TPC/NHC), states “The problem of forecasting tropical cyclone intensity change continues to be a real challenge for meteorologists despite the recent advances in numerical weather prediction” and “very little guidance is available to forecast intensity changes.” This problem is especially acute for rapidly intensifying storms: Avila notes that “the 72 hour intensity forecast valid for 1800 UTC 12 September was underestimated by 100 kts” during 1997’s Hurricane Linda, which occurred in the eastern North Pacific (ENP).

It is recognized that the changes in hurricane intensity and structure are related to synoptic and climatological factors (Frank, 1977; Merril, 1988). These relationships make the development of statistical hurricane intensity prediction tools possible. Some of the models in operational use include the Climatology and Persistence (CLIPER; Neumann 1972), Statistical Hurricane Intensity Forecast (SHIFOR; Jarvinen and Neumann 1979), Statistical Hurricane Intensity Prediction Scheme and (SHIPS; DeMaria and Kaplan 1994b, 1999). Climatological, persistence, and synoptic variables are correlated to future intensity change in each of these models through the use of multiple linear regression (MLR). SHIPS also incorporates synoptic information derived from the global Aviation model; thus it can be considered a statistical-dynamical model. Only with the inclusion of forecasted synoptic data, did SHIPS show significant forecast skill relative to a similar model using only climatological and persistance variables (DeMaria and Kaplan 1999). Other research models developed using similar methodology and synoptic/dynamic variables have shown a similar improvement in skill over climatology and persistance. An example is the Eastern Pacific Intensity Prediction model (EPIC, Petty and Hobgood, 1998).

The MLR technique used in all of these models is a standard statistical research method that develops an equation whose inputs are the predictors, variables believed to be correlated to the output, and whose output is the forecast. The great advantage of this method is that the linear (or quadratic) combination of the predictors obtained has the smallest possible error of any linear combination of those variables. The disadvantage is that the relationships between the predictors and the predictant (the forecast) are fixed. If these relationships change, the forecasts gain additional errors.

The goal of this research is to examine the efficacy of using a fuzzy logic/adaptive weighting (FLAW) technique to predict hurricane intensity change. We hypothesize that the FLAW technique will be able to take uncertainties in the degrees of relationships between predictors and intensity change into account by adapting the weights given to each predictor, thereby producing smaller errors than would be possible for a fixed linear (or quadratic combination) produced by MLR. This is an application of the Standard Additive Mode (Kosko, 1996), which incorporates fuzzy logic in the production of the forecast and the updating of the weights.

To accomplish this goal, climatological and persistance data are gathered for the eastern North Pacific (ENP) hurricane basin for the years 1982-99. Two models are developed on data from 1982-1996, one using the standard MLR technique, and the other using the FLAW technique. Then the models are run on three years of validation data from 1997-99. The forecasts are then examined to determine the differences, if any, in the characteristics of the model forecasts. Section two describes the data used, the third section describes the development of the models, results are discussed in section four, and conclusions and directions for future work are discussed in the last section.

2. Data

i. Discussion of the data sets

Two data sets are obtained and merged to obtain the climatological and persistence variables used to develop and validate the models. These include the best-track data set from the NHC and the global optimum interpolation (OI) sea surface temperature (SST) data set from the National Oceanic and Atmospheric Administration's National Centers for Environmental Prediction (NOAA-NCEP).

The NHC Eastern Pacific best track data set (Jarvinen, et. al., 1984) contains 6-hour observations (0000, 0600, 1200, and 1800 UTC) of all the tropical cyclones (depression, storm, and hurricane stages) observed in the ENP from 1949-1999. Each observation includes the latitude and longitude of the tropical cyclone (TC) center, the lowest sea level pressure (in mb) when available, and the intensity, defined as the highest 1-minute maximum sustained wind (rounded to 5 kt intervals) at an elevation of 10 m. Due to the lack of aircraft reconnaissance flights and sparse ship and buoy observations in the ENP, surface pressure and intensity are often estimated from satellite imagery using the Dvorak technique (Dvorak 1975). Changes in this technique, developed in the 1970's and improved in the 1980's (Dvorak 1980), may have caused inhomogeneities in the data. Since only best track data from 1982 onward are used in this study, the impact of changes in the Dvorak technique on the intensity estimations is deemed to be small. A more important inhomogeneity may have resulted when responsibility for the ENP basin transferred from the Eastern Pacific Hurricane Center in Redwood City, CA to the

National Hurricane Center in Miami, FL in 1988. Some researchers (Whitney and Hobgood, 1997) have noted a substantial increase of nearly 10% in the relative intensities for storms over the period 1988-1993, but it is unclear as to how much of this increase can be attributed to changes in the intensity analyses of these systems. Finally, because the area of responsibility for the ENP basin only extends out to 140° W, observations of the storms west of this line are excluded from consideration. In at least one case (Guillermo 1997), a storm crossed out of the basin and later recurved into the basin at a higher latitude. A few observations from a recurving storm east of 140°W are included for this storm and will be discussed in the analysis and results section. In addition to the longitude restriction, storm observations over land are eliminated, but observations near land but over water are included. This is in contrast to some researchers who throw out observations if they are too close to land. This research takes the view that an intensity prediction is most important from a societal perspective when the storm is close to land.

The global OI-SST data set contains weekly data on a one-degree grid using *in situ* and satellite SST data produced through the OI method of analysis (Reynolds and Smith 1994). The biases in the satellite data were first corrected (Reynolds, 1988; Reynolds and Marsico, 1993), which added in a small amount of noise in time. The recommended ¼-½-¼ binomial filter in time was not applied because the SST fields are not used in time comparisons. The weeks for the data are centered on Sunday from 1981-89 and on Wednesday from 19990-99. This change is taken into account. A Fortran program is used to linearly interpolate in time and space from the one-degree gridded SST data to the day and location (given in tenths of a degree longitude and latitude) of the tropical cyclone observations in the NHC best track data set.

ii. Discussion of climatology and persistence variables

A FORTRAN program that calculates persistence and climatology variables as well as observed future intensity changes was modified to include the interpolated SST at the TC center. Table 1 contains the list of variables computed and extracted for investigation in this study.

Variable Name	**Source**	**Description**
STNAME	Best track	The name given to the storm
DATE	Best track	The date of the observation (YYMMDDHH)
ABS(JDAY-237)	Calculated	# of days from the seasonal peak of the ENP hurricane season (days)
VMAX	Best track	Estimated intensity (maximum sustained wind in kts)
DTLF	Calculated	Closet distance to land from the center (km)
PVMAX06	Calculated	Intensity change in the previous 6-hour period (kts)
PVMAX12	Calculated	Intensity change in the previous 12-hour period (kts)
PVMAXAVE3	Calculated	The average of the past 6-, 12-, and 18-hour intensity changes (kts)
TCLAT	Best track	Latitude of the TC center (°)
TCLONG	Best track	Longitude of the TC center (°)
CSM	Calculated	Total storm speed based on change in storm location over the previous 12-hours (m/s)
USM	Calculated	Eastward component of storm speed based on the change in storm location over the previous 12-hours (m/s)
VSM	Calculated	Northward component of storm speed based on the change in the storm location over the previous 12-hours (m/s)
PUSM	Calculated	Change in USM from the USM calculated 12 hours ago (m/s)
PVSM	Calculated	Change in the VSM from the VSM calculated 12 hours ago (m/s)
N	Calculated	(Number of days since the storm formed)+2 (days)
NBRDAYS	Calculated	Total storm lifetime (days)
HOURS	Calculated	Number of hours since the storm formed (hrs)
SST	Interpolated from OI SST	SST interpolated in location and day of the storm observation (°C)
FINT = (VMAX(I+t)-VMAX(I))	Calculated	Future change in intensity from the current intensity, where I is the current observation and t=2,4,6,10, or 12 corresponds to the time periods 12,24,36,48,60, and 72 hours in the future from the current observation (kts)

Table 1. Description and source of the variables that are investigated.

To determine the correlation between predictors and future intensity change, Microsoft Excel is used to plot the predictor vs. the future intensity change for each of the six time periods (12, 24, 36, 48, 60, and 72 hours). The program automatically draws a best-fit line using single linear regression and displays the R^2 correlation statistic. Each predictor is chosen for inclusion into the model based on how well it explains the variance in intensity. Figures 1 and 2 contain an example using PVMAX06.

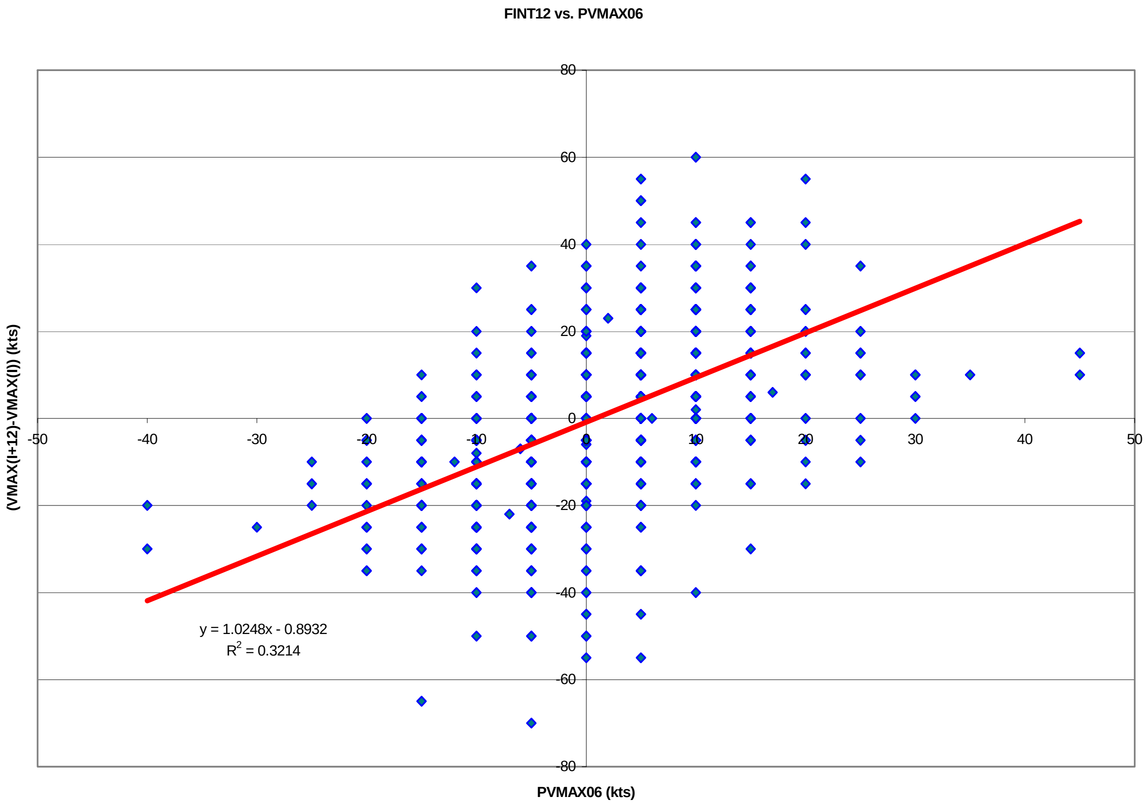

Fig. 1. A graph of the intensity change (kts) 12-hours from the current observation vs. the previous 6-hour intensity change (kts) for 5202 observations from 1982-1996.

The slope of the best-fit line determines the relationship, in this case positive, meaning that the previous 6-hour intensity change is positively correlated to the intensity change that occurs in the next 12-hours. The large number of points in the first quadrant suggest that if a storm intensified in the previous 6-hour period, it tends to keep on intensifying in the next 12-hour period. Similarly, if a storm weakened in the previous 6-hour period, it tends to keep on weakening in the next 12-hour period, as evidenced by the number of points in the third quadrant. The percent of the variance of intensity change explained by this variable is 32.1%. Thus, the previous 6-hour intensity change, or PVMAX06 has significant skill as a predictor of the 12-hour intensity change (hereafter FINT12) because the climatology and persistance factors that favored intensification or weakening will likely persist into the future to favor further intensification or weakening.

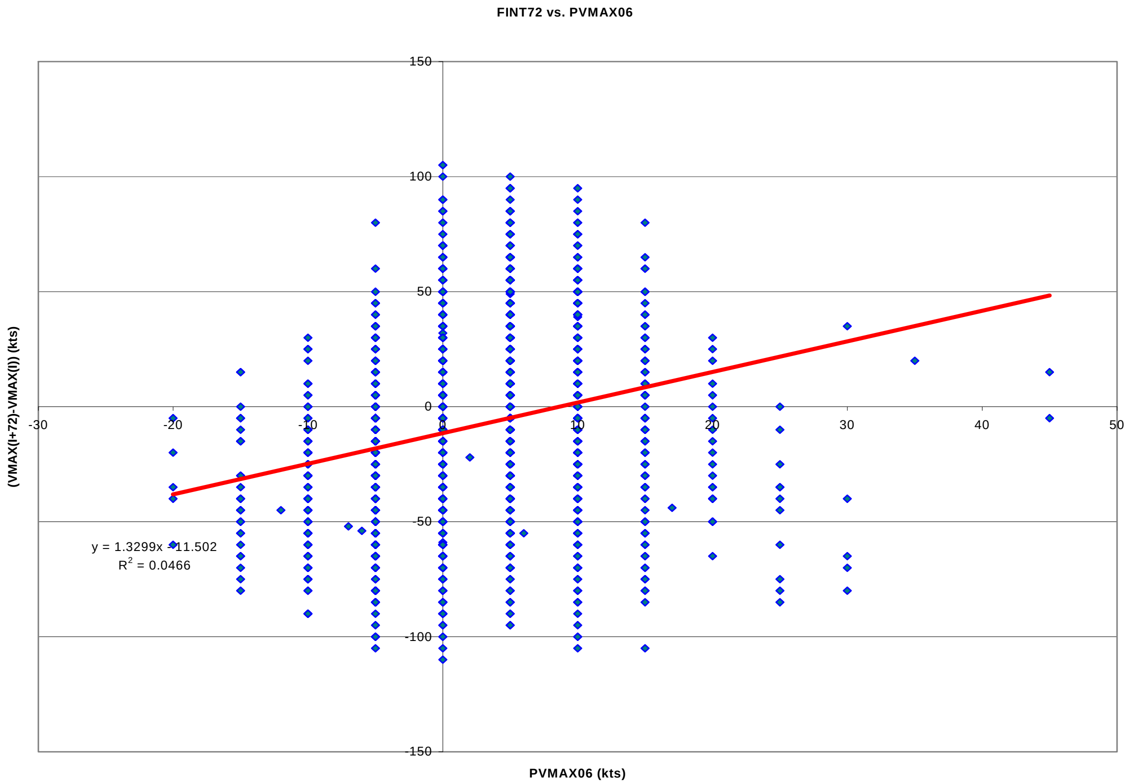

Fig. 2. A graph of the intensity change (kts) 72-hours from the current observation vs. the previous 6-hour intensity change (kts) for 3268 observations from 1982-1996.

Figure 2 shows a much weaker relationship, with only 4.66% of the variance in the 72-hour future intensity changes explained by PVMAX06. The factors that may have favored intensification at the current observation are not likely to be similar to the factors which the storm will experience in 72-hours. The probability that conditions will be similar at two time periods decreases as the time period increases. This is due partly to storm motion, which tends to carry storms into different SST and synoptic regimes, and the evolution of the synoptic pattern itself.

This type of analysis is done for many of the variables, and the resulting R^2 values from linear regression are graphed for the six time periods. This is shown in Figure 3.

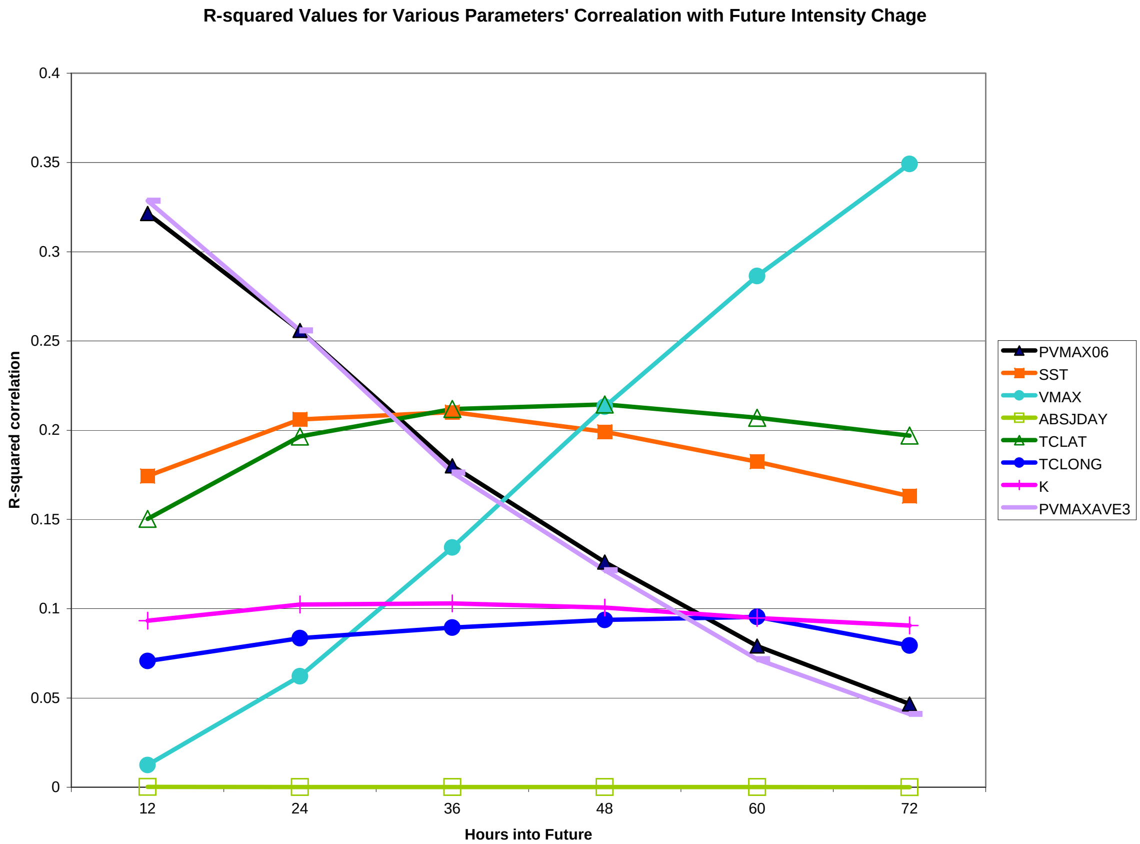

Fig. 3. R^2 values from linear regression analysis for eight variables at 12, 24, 36, 48, 60, and 72 hours into the future.

Figure 3 shows the decrease in strength over time of PVMAX06 as a predictor of future intensity change, and the strength of VMAX as a predictor at longer time periods. Since storms go through a life cycle of intensification and then weakening, each VMAX value could potentially be associated with both a positive and a negative intensity change. For example, a storm with an intensity of 50 kts could be either weakening or strengthening. The fact that its intensity is 50 kts says little about how the synoptic or internal factors may change in the future. For a more intense storm, say 120 kts, VMAX becomes a good predictor because it indicates at what part of the life cycle a storm is. A storm this strong is not likely to intensify much more and thus is more likely to weaken, especially at longer time periods. The synoptic factors that favor the rapid intensification usually undergone by the very intense storms are unlikely to persist for more than a couple days. Internal dynamical processes that tend to occur in intense storms, such as eye-wall contraction cycles, can start to have a detrimental effect on further intensification (Willoughby, 1978; Willoughby, et. al, 1982; Willoughby, 1990). Thus VMAX becomes a good predictor of future intensity change at longer time periods.

The next most important predictors of intensity change are SST and tropical cyclone latitude (TCLAT). Warm SST's are a necessary (but not the only) condition needed for intensification (DeMaria and Kaplan, 1994a). The energy source of a TC arises from the latent energy released by condensing water vapor. High SST's increase the potential energy available to the storm, but the rate at which condensation occurs is controlled by other synoptic and dynamic factors, so the actual intensity of the storm is not related to SST. Instead, SST serves to place an upper bound on the intensity. In the ENP basin, the warmest SST's tend to occur near the Mexican coast with SST gradually cooling to the west and north. Most ENP storms form over warm SST's (typically greater than 26.5°C) and move west northwestward toward decreasing SST's, leading to their eventual demise. Thus SST is positively correlated with intensification, or more accurately, cooler SST's are positively correlated with weakening.

TCLAT is partially correlated with SST, since SST decreases with higher latitudes in the ENP. This means that storms at higher latitudes tend to be over cooler SST's, which are detrimental to intensification. Another reason that TCLAT is a significant predictor stems from the fact that the synoptic regime at higher latitudes tends to also counter intensification due to the increased westerlies and associated wind shear. The thermodynamic instability, which affects the rate of latent energy release, also tends to be less at higher latitudes.

The four most important predictors, PVMAX06, VMAX, SST, and TCLAT were chosen for model development. Other predictors, such as the TC longitude and number of hours since TC formation were significant but were not used in this study. One predictor, the absolute difference between the Julian day of the observation and the climatological peak of the ENP basin (ABSJDAY), was found to have no significant correlation to future intensity change. This conflicts with some previous studies (DeMaria and Kaplan, 1999; Petty and Hobgood, 2000), which use it as a predictor of intensity change. To back up these conflicting findings, it should be pointed out that the relative intensity of ENP storms tends to be fairly constant from May through October, the period in which most storms occur. This suggests that for most of the season, the climatological and synoptic variables generally favor normal development and intensification (Whitney and Hobgood, 1998). Early or late season storms would certainly tend to be affected by a more hostile synoptic regime, but these storms are relatively infrequent. Indeed, ABSJDAY is more likely to be a predictor of TC genesis and frequency than of intensity change.

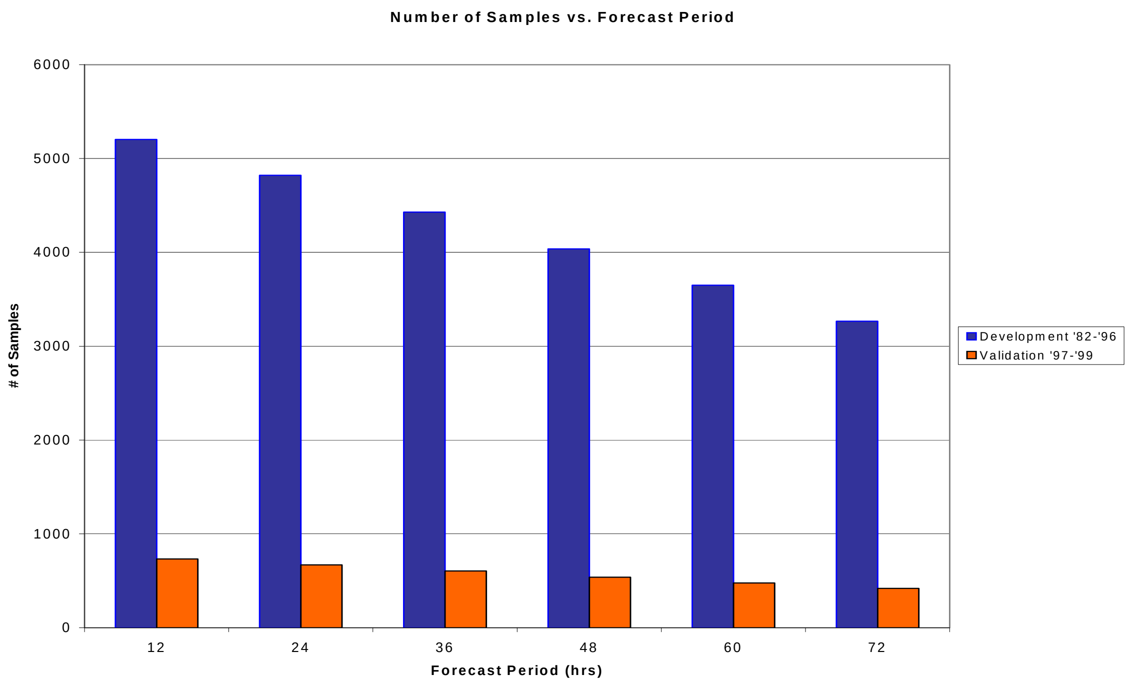

Fig. 4. The number of observations used in development and validation for the 12-, 24-, 36-, 48-, 60-, and 72-hour forecast periods.

3. Methodology

This section describes the theory behind the MLR and FLAW models and how they were implemented using Simulink, a Matlab tool.

i. The theory behind MLR

MLR is a statistical technique that uses linear algebra to solve a system of linear equations in order to come up with the equation that best relates the predictors to the desired predictant (the forecast). This unique equation consists of a linear combination of predictors multiplied by their respective correlation coefficients. Their sum gives a forecast of the predictant. The error is given by the difference between the forecast value and the observed value (called truth). MLR finds the linear combination that minimizes the square of the error. Thus, it is impossible to find a better linear combination with a smaller absolute cumulative error.

The advantages of MLR are clear: the method finds the optimum coefficients for a linear combination of predictors. There are disadvantages however. Predictors that have a high degree of colinearity (that is, they are correlated with each other), could cause the best possible linear combination of a set of predictors to have a larger cumulative error than another combination of predictors that are more independent. SST and TCLAT are somewhat collinear, although not to a great degree. This could cause the MLR method to over-predict weakening in the case of a storm both at high latitudes and over cold water.

Colinearity also limits the number of predictors that can be used in a MLR model. Too many predictors will cause the system to be over-determined, which means that it may do well with development data, but will forecast poorly on independent data (the kind encountered in operational use). Another disadvantage is that the relationships between the climatology and persistance predictors and the intensity, indicated by the correlation coefficients, are fixed. If the relationships between predictors and intensity change, or have non-linear interactions, the correlation coefficients produced by MLR using a development data set may not produce the smallest possible errors for a linear combination on independent data. Finally, to be effective, the predictors and predictant should have a normal, or Guassian, distribution.

ii. Implementing MLR with Simulink

The MLR model is implemented using Simulink, a symbolic programming toolset of Matlab often used for the simulation of dynamical systems and signal processing. The predictor and truth data for a time period such as 12-hours are ingested into the program. Each predictor and the truth values are normalized linearly so that values close to –1 indicate a weakening influence, while values close to 1 indicate that the predictor favors intensification. To normalize the data, the mean of each predictor for the entire development period of 1982-96 is subtracted from the observed value of the predictor. Then the predictor is divided by a value that places most of the data between –1 and 1. This number is chosen by analyzing the graphs and adjusting the number until the predictor distribution looks similar to the truth distribution. One of the predictors used, VMAX, does not have a normal distribution because all storms have a weak phase, but few storms reach a high intensity, so the VMAX data are quite skewed. A more complicated normalization might have been more appropriate for this variable, but for simplicity and consistency, a linear normalization is used.

The following figure shows the top-level of the Simulink program used to normalize the data and find the MLR correlation coefficients.

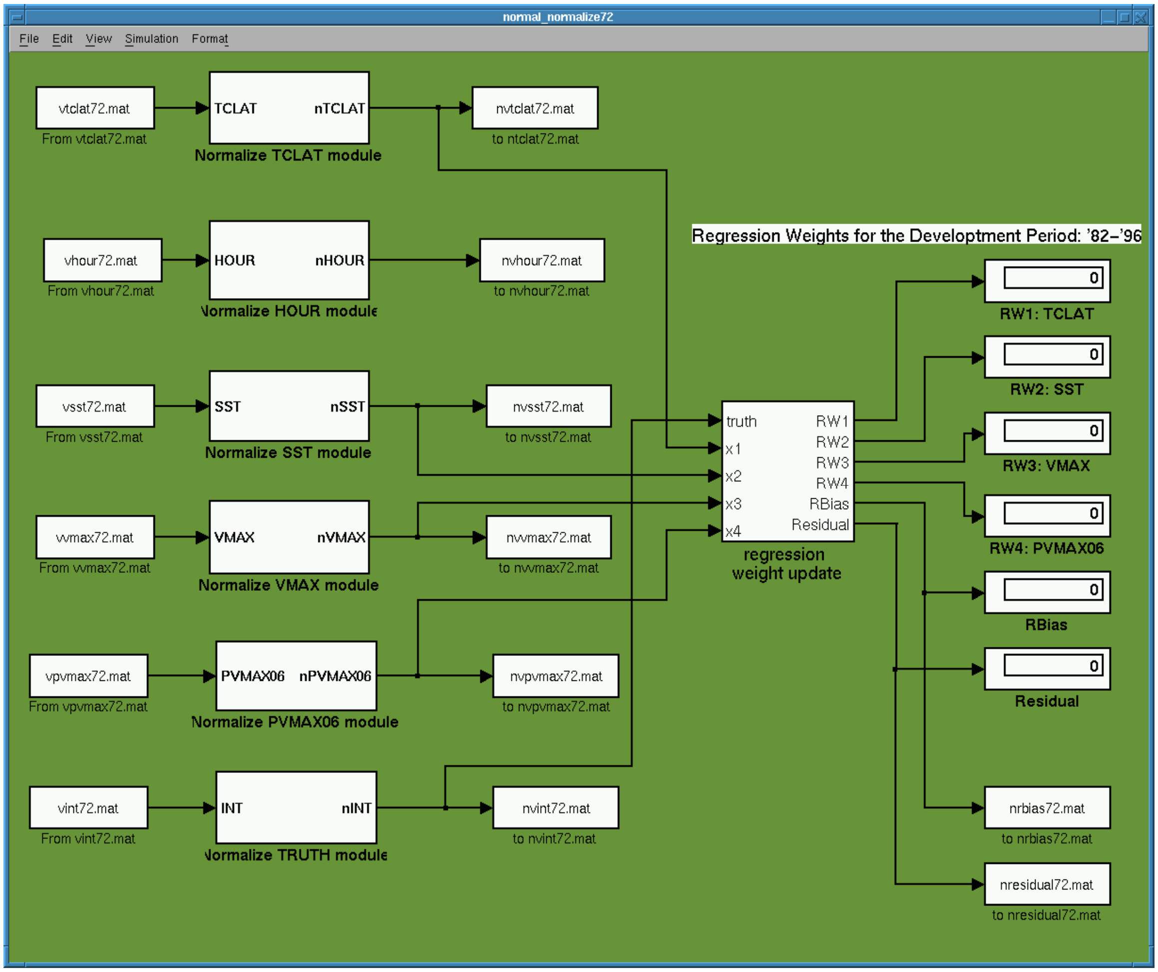

Fig. 5. The top-level diagram of the program that normalizes the data and finds the MLR correlation coefficients. The general flow of data is from left to right.

The next figure shows how TCLAT is normalized. The means were calculated using Microsoft Excel.

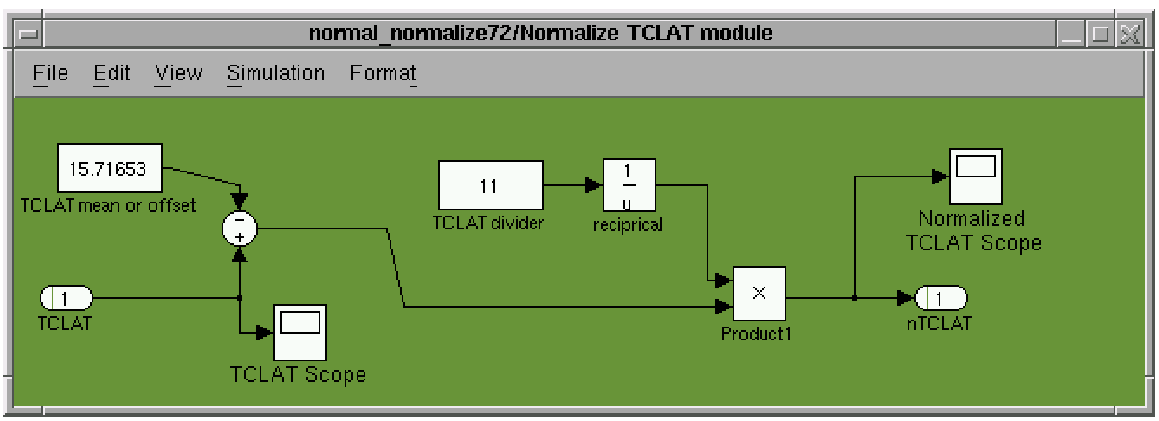

Fig. 6. The normalization of TCLAT.

The data enters the module through the oval at the left. The mean of 15.7 is then subtracted. Next, the data are divided by the number 11, which put most of the normalized latitude values between –1 and 1. Two scopes are also connected which allow the user to see the distribution of the data before and after normalization. Finally, the data exits the module via the oval and is then sent by the top-level 'code' out to file.

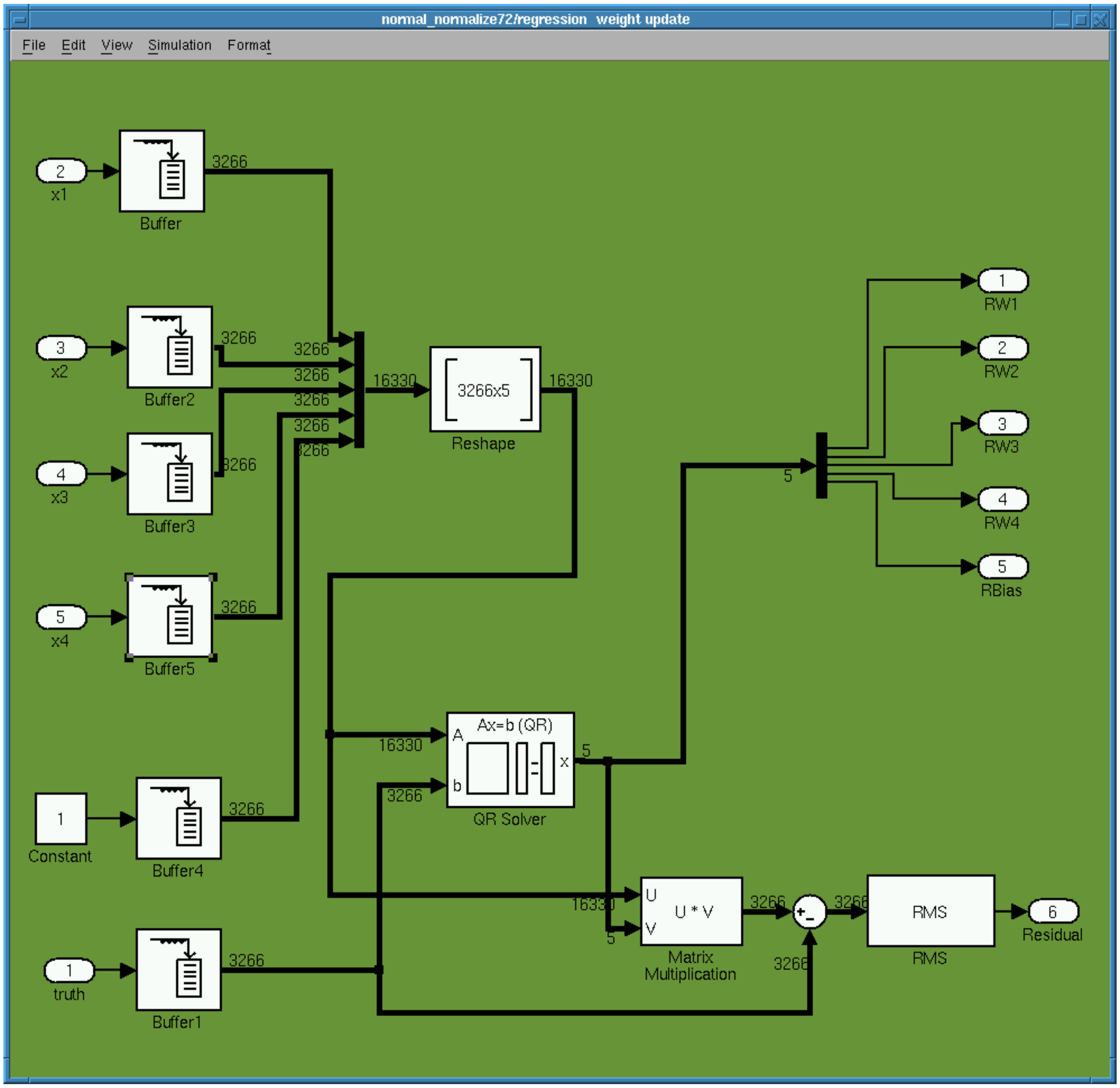

Fig. 7. The 'code' that calculates the correlation coefficients for the MLR.

In Figure 7, the data are ingested and placed in buffers, which store the data as vectors. The thick lines connecting the buffers to the bar indicate a vector flow instead of a scalar flow. The bar, called a *mux* in Simulink, combines vector or scalar flows into one large vector flow. Then the vector is reshaped into a matrix whose five rows correspond to the four predictor vectors and a vector of ones. The linear system **Ax**=**b** is then solved, where **A** is the predictor matrix, **b** is the truth vector, and **x** is the desired vector of correlation coefficients. These coefficients are sent out of the module on the right, and also to a matrix multiplier, which computes the matrix product **Ax**. This produces a regression forecast vector. By subtracting the truth vector, the residual errors are obtained. A module then computes the root-mean-square error and outputs the resulting residual. The correlation coefficients are displayed on the top-level diagram. The values are recorded for future reference and use in the next program, where the MLR forecasts and errors are computed along with the FLAW forecasts and errors.

iii. The theory behind FLAW

Fuzzy logic, an extension of multi-valued logic, is a controversial field (McNeil and Freiberger, 1994). It breaks the laws of traditional Aristotelian logic, which state that something either is or isn't: there is no middle ground. For instance, a certain piece of fruit is either an apple or it is not an apple – it cannot be both part apple and part non-apple. This gives rise to the idea of crisp sets, where 0 or 1 indicates membership in the set. Fuzzy logic deals with fuzzy sets, where membership can be a range of values from 0 to 1. For instance, a wooden carving of an apple may have partial membership in the set of apples because it has the shape and color of an apple. Another example would be a pear-apple. The degree of membership is determined by how apple-like an object is. Some things, like a chair, have no membership in the set of apples. Fuzzy logic is able to use fuzzy sets, and the ambiguities and uncertainties thus contained, to do operations whose results can sometimes exceed those possible using traditional logic methods.

The methods of fuzzy logic were first applied to engineering and control systems, with stunning success in the late 1980's in Japan, but were largely ignored in the corporate and university laboratories of America. After a full decade, American technology and science is now playing catch-up in the world of fuzzy logic.

Fuzzy logic is used in the FLAW model in several ways. First, the data is 'fuzzified' or normalized in a way so that the predictors give information about intensity change. A 'fuzzy set' is then created in degrees of possibility. In order to do this, simplistic assumptions are made about the relationship between the predictor and intensity change. The linear normalization used (the same as that of the MLR model) undoubtedly does not capture the precise relationship. Instead there is a certain amount of uncertainty. The values of the predictors themselves are also somewhat uncertain.

As an example, SST is believed to be related to intensification, since the SST provides an upper bound on the hurricane's intensity. A normalized value of 1 (corresponding in this case to an SST of 30°C) would indicate that the SST is definitely warm enough for further intensification. A normalized value of 0 would indicate that the SST is neutral for further intensification, that is, the storm has enough energy to sustain itself. On the flip side, a normalized value of –1 would indicate that the SST is definintely too cool to sustain the present intensity, thus the storm will likely weaken. There are actually two fuzzy sets here, one for intensification, and one for weakening. A normalized SST value of 0.5 would indicate the degree of possibility that exists for further strengthening (to do this properly, the SST fuzzy set confidences should be computed using information from the current storm intensity).

iv. The implementation of FLAW

Since the relationships between the climatology and persistance predictors and the future intensity change are uncertain, and may undergo changes and non-linear interactions, it might be possible to derive a scheme by which the weights are allowed to adapt to changes in truth. The weights given to the predictors are supposed to embody the importance of the predictor in relation to the forecast. As the relationships change, the weights would ideally change accordingly, thereby producing smaller errors than would be possible using a fixed linear combination as in MLR. This is implemented by modifying a Simulink program, called the 'Forecast Integrator', developed at the National Center for Atmospheric Research (NCAR) (Dalton, 1999). The original goal of the integrator is to combine temperature forecasts from three sources, climatology, persistence, and DMOS into a more accurate forecast which has less error than a linear combination of the three forecast sources (obtained by MLR). In the words of the program's creator, "The forecast integrator is an iterative approach to determining a weighted average with a small bias." The program follows a "format similar to that used in the fuzzy standard additive mode (SAM) (B. Kosko, 1997). It is beyond the scope of this paper to describe the Forecast Integrator algorithm in depth. Instead, an overview of the algorithm, much of which comes directly from Dalton's paper, is described. Some of the more germane Simulink modules are shown.

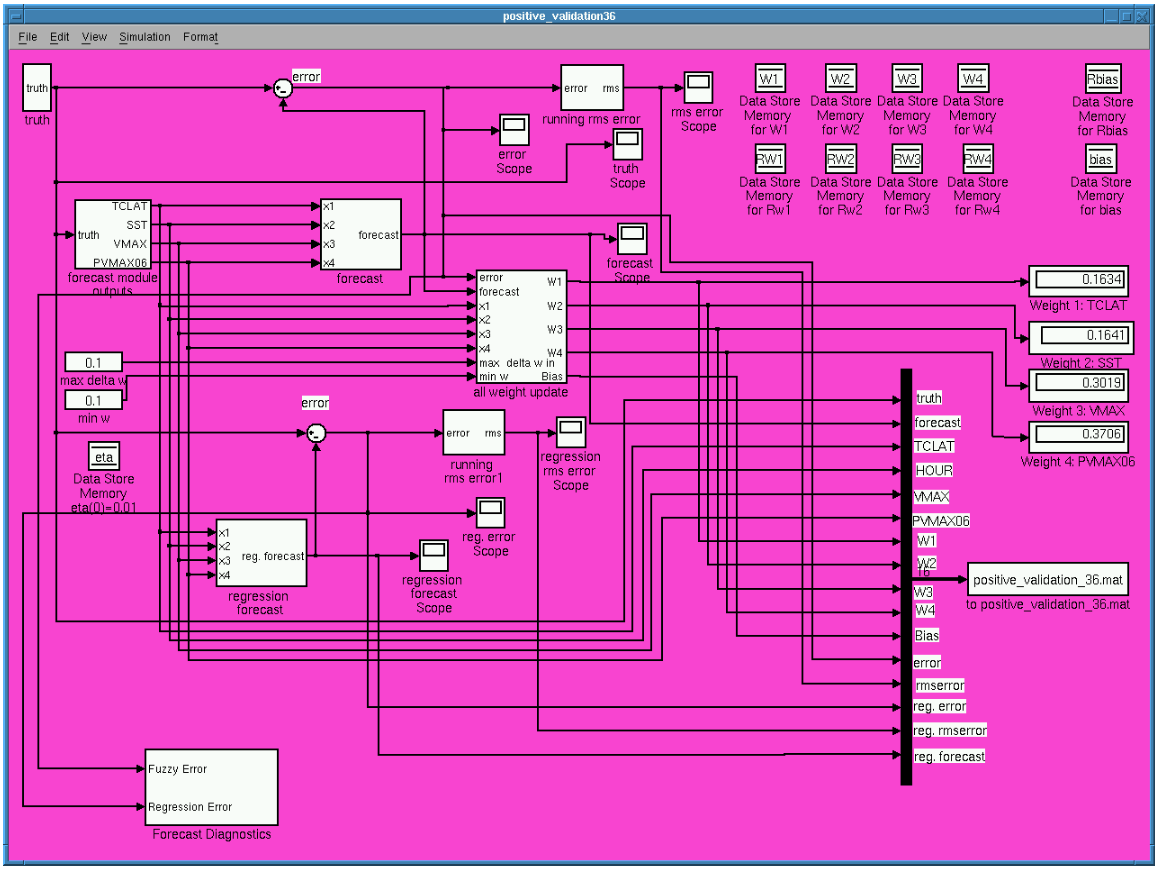

Fig. 8. The top-level diagram of the FLAW and MLR models.

Figure 8 shows the overview of the FLAW model. Truth data are ingested in the "truth" module in the upper left corner of the diagram. The normalized predictor data are ingested in the "forecast module outputs" nearby. These data are sent to the fuzzy "forecast" module and the "regression forecast" module where forecasts are produced. The forecast is then subtracted from the truth to obtain the error for the FLAW and MLR models. The forecast, error, and predictor values (the outputs of the forecast modules) are then sent to the "all weight update module" which updates the weights for the FLAW module. Finally, truth, forecasts and errors of both modules, as well as the predictor values, assigned weights, and bias are combined into a vector feed and stored in an output file. The forecast errors are also sent to a "forecast diagnostics" module, which computes the absolute cumulative error, overall forecast bias, and forecast variance and standard deviation for MLR and FLAW.

The first phase of the algorithm, which follows the fuzzy feedforward SAM system, computes the forecast. The forecast produced by the Forecast Integrator is the normalized weighted sum of the forecast module outputs, which in this case are the normalized values of the four predictors: TCLAT, SST, VMAX, and PVMAX06. This is summarized in equation 1:

$$X_f = \frac{\left(\sum_{i=1}^{n_f} w_{if} c_{xif} X_{if}\right)}{\left(\sum_{i=1}^{n_f} w_{if} c_{xif}\right)} + b_f \qquad (1)$$

where X_f is the forecast output of the Forecast Integrator for the forecast parameter f. X_{if} are the outputs of the forecast modules (the predictors), w_{if} is the iteratively updated weight for the ith forecast module, c_{xif} is the confidence value associated with the ith forecast module of the forecast parameter f, n_f is the number of forecast modules, and b_f is the iteratively updated bias to the forecast integrator. In this case, there are four forecast modules, and all the confidences are set to one. The advantage to using (1) instead of a simple linear combination is that the resulting sum gracefully handles missing values and zero confidences (Dalton, 1999).

"Following the feedforward SAM system, the confidences play the role of the degree of memberships of the outputs from the forecast modules in the set that contains the "true" forecasts. The outputs from the forecast modules play the role of the centroids of the rules. The rules in this case are equal to the statement that if the output of the forecast module has a certain value, then that is the value for the forecast. The volumes for the centroids are also set equal to one, so that they are not a factor in the equation. The bias is an extra degree of freedom added to the equation to compensate for any biases from the outputs of the forecast generators" (Dalton, 1999). Since this implementation of the model does not use confidence fields, the full advantage afforded by fuzzy logic is not realized.

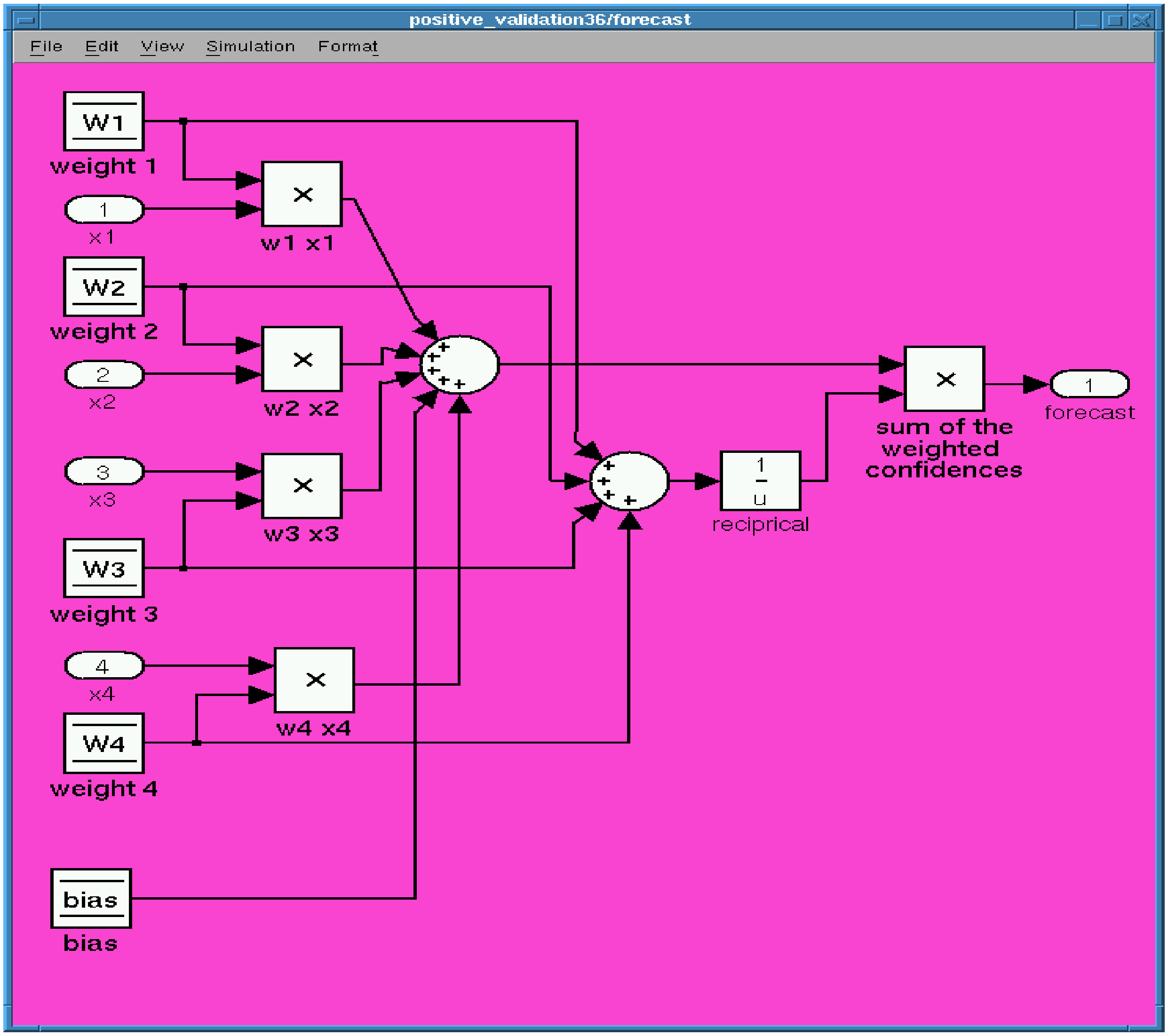

Fig. 9. The fuzzy "forecast" module that implements equation (1).

The second part of the algorithm iteratively updates the weights and bias so that the system adapts to changes in truth. In essence, this part of the algorithm allows FLAW to 'learn' on a storm, as the relationships between climatology predictors and intensity change. "The algorithm is based on the widely used least-means-squared (LMS) optimization, where the weights and bias are changed proportional to the negated derivative of the squared error with respect to the weights or bias" (Dalton, 1999). Essentially, if output from one forecast module output causes the overall forecast to be pulled too much in the wrong direction, the weight is updated to reduce the error. Weight changes (delta weight) are calculated and checked to ensure they do not exceed the set maximum delta weight change (for stability and convergence of the system), and that the resulting weights do not become non-negative (to avoid numerical problems).

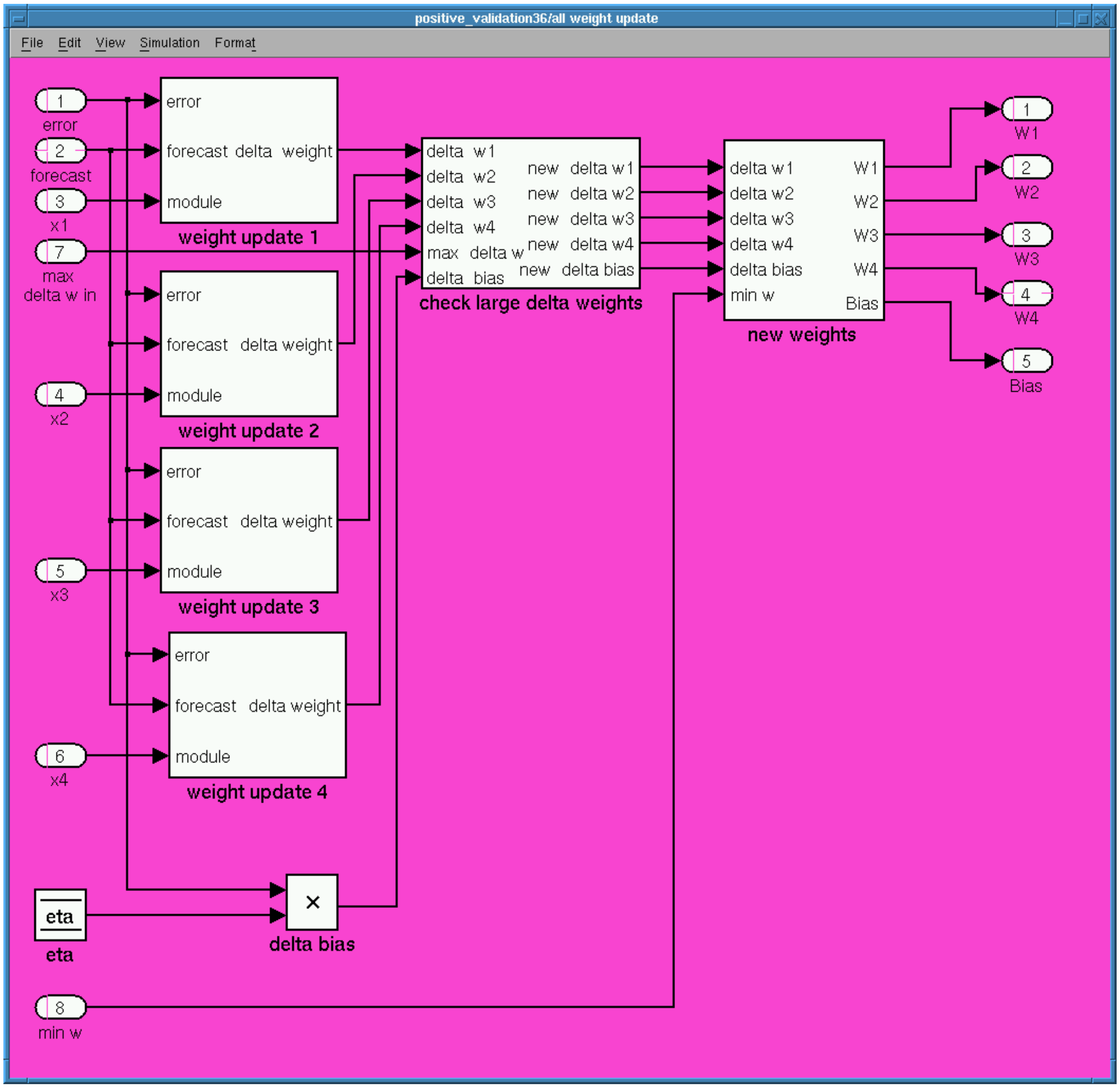

Fig. 10. The "all weight update" module, where the delta weights are computed, checked, and then added to the weights to produce the new weights to be used in the next forecast cycle.

The details of the "all weight update" module can be found in Dalton's paper. An important issue that arises is how to determine the step size, which controls the 'learning rate' of the system. Choosing the step size so that the maximum weight change is kept below a set value solves this problem. "A larger maximum weight change would make the system more responsive, but also more susceptible to noise. A smaller maximum weight change would make the system more robust, but may cause it to converge too slowly" (Dalton, 1999).

In order to be useful for operational forecasting, the output of the 'Forecast Integrator" program had to be 'de-normalized' so that the value of the forecast would be in knots. To do this, the reverse operations of the original normalization were simply applied to the output. The bias could not be de-normalized as easily, since it depends on a renormalization of the weights at each time step.

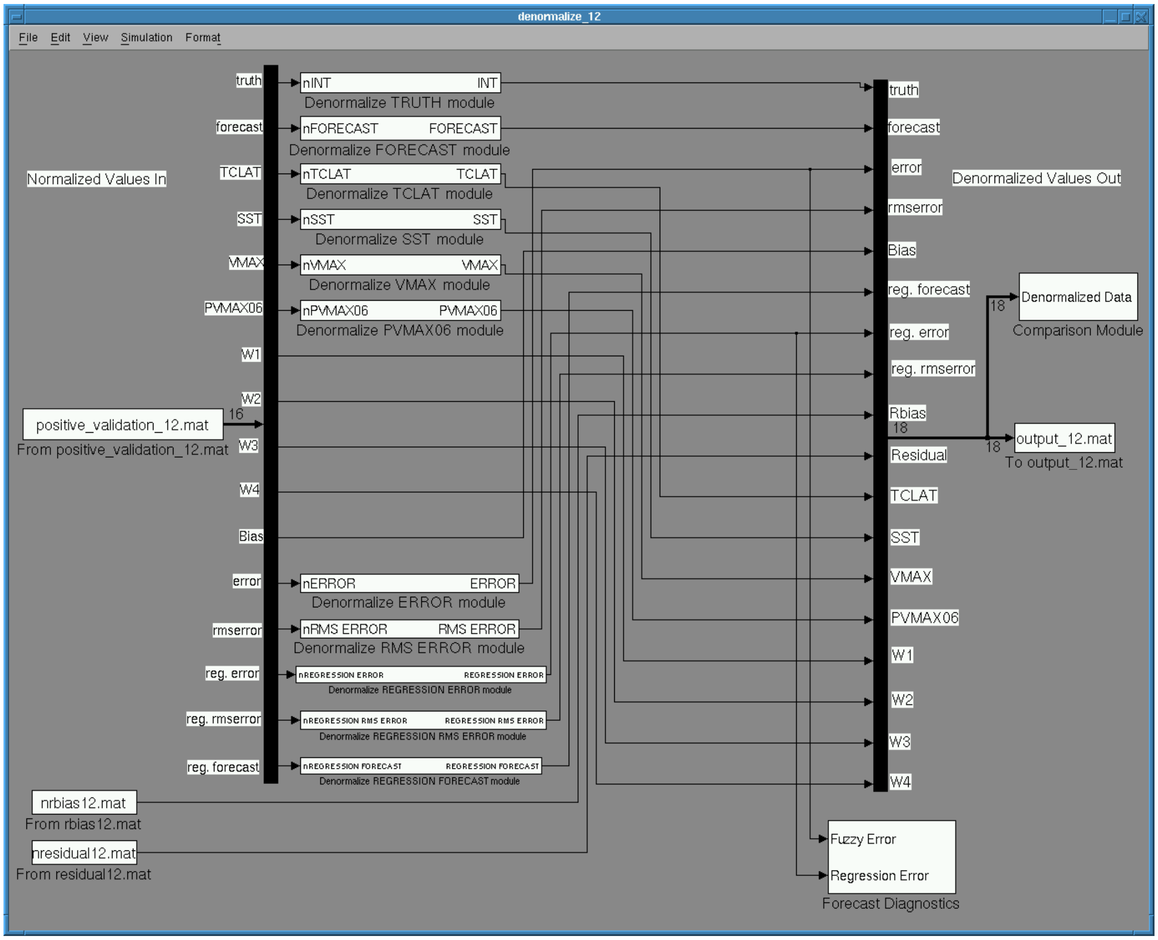

Fig. 11. The de-normalization overview. Denormalized values are then sent to scopes for viewing, to an output file, and to a “forecast diagnostics” module, which calculates the cumulative sum of the absolute errors, the overall forecast bias, and the variance of the forecasts.

4. Analysis and Results

The results are analyzed by examining the average absolute error and the mean of the errors (the overall forecast bias) for the validation period of 1997-99. Several case studies of individual storms are also examined to look at the characteristics of the model errors with respect to the storm development and subsequent decay.

i. Overall results

Overall, FLAW produces average absolute errors similar in magnitude to MLR. At 12-, 24-, and 72-hours, FLAW produces smaller errors than MLR. At other time periods, MLR has a smaller average absolute error. These results are shown in Figure 12.

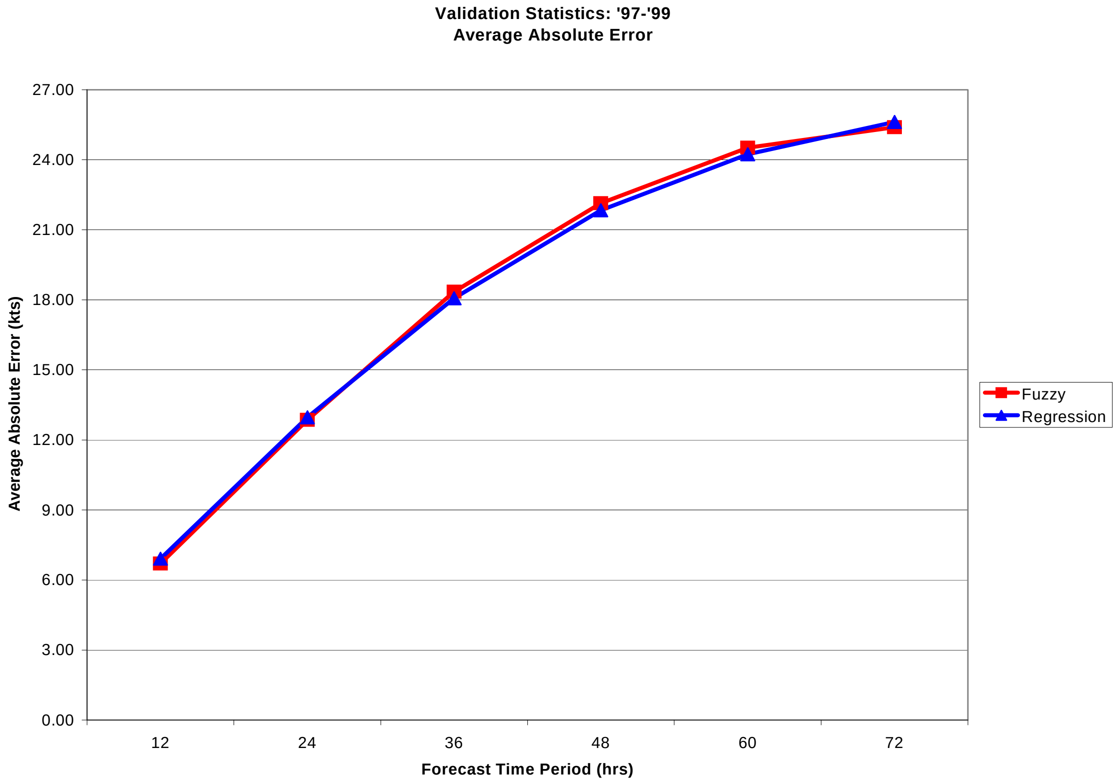

Fig.12. A graph of the average absolute error (kts) by forecast time period (hrs).

The differences between the average absolute error for FLAW and MLR are not significant. The percent improvement of FLAW over MLR is show in Figure 13. As the graph shows, the results are clearly mixed. FLAW appears to have done about as well as MLR in forecast intensity change.

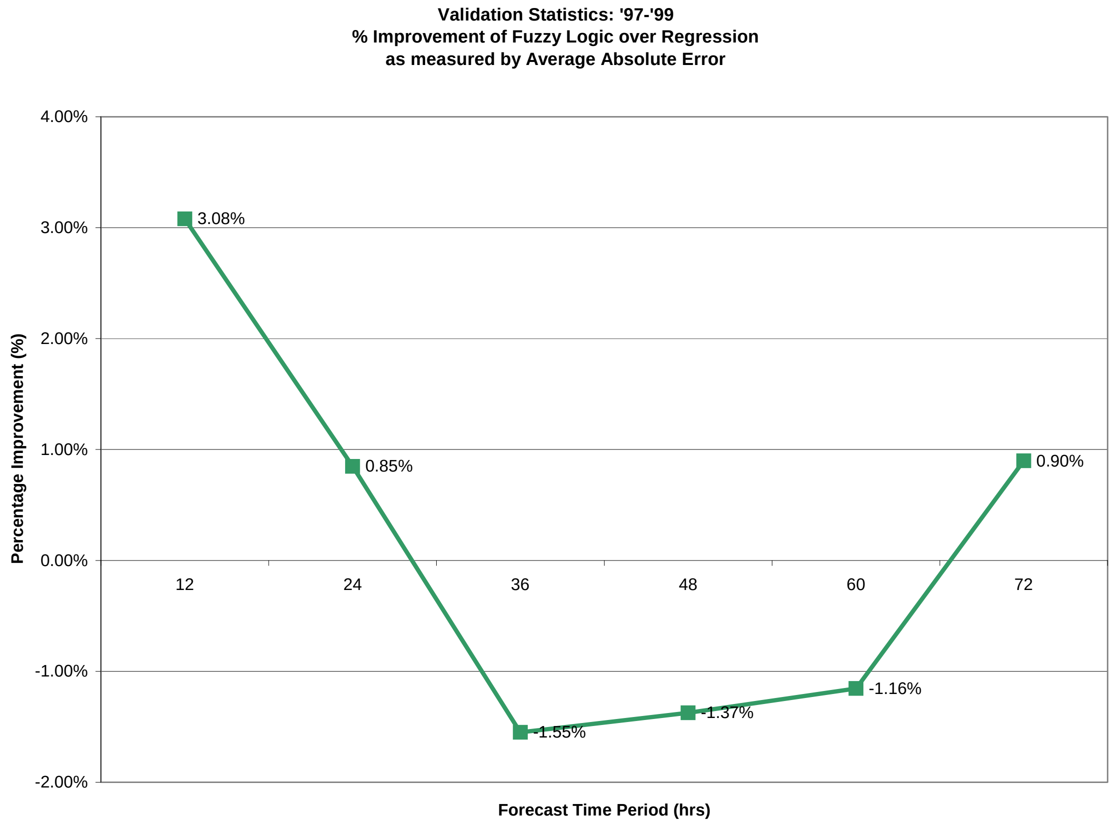

Fig.13. The percent improvement of FLAW over MLR for the forecast time periods (hrs).

One area in which FLAW substantially improves over MLR is in the average error, or overall forecast bias. At the 36-,48-, 60-, and 72-hour forecast periods, FLAW had an overall forecast bias at least 30% smaller than MLR.

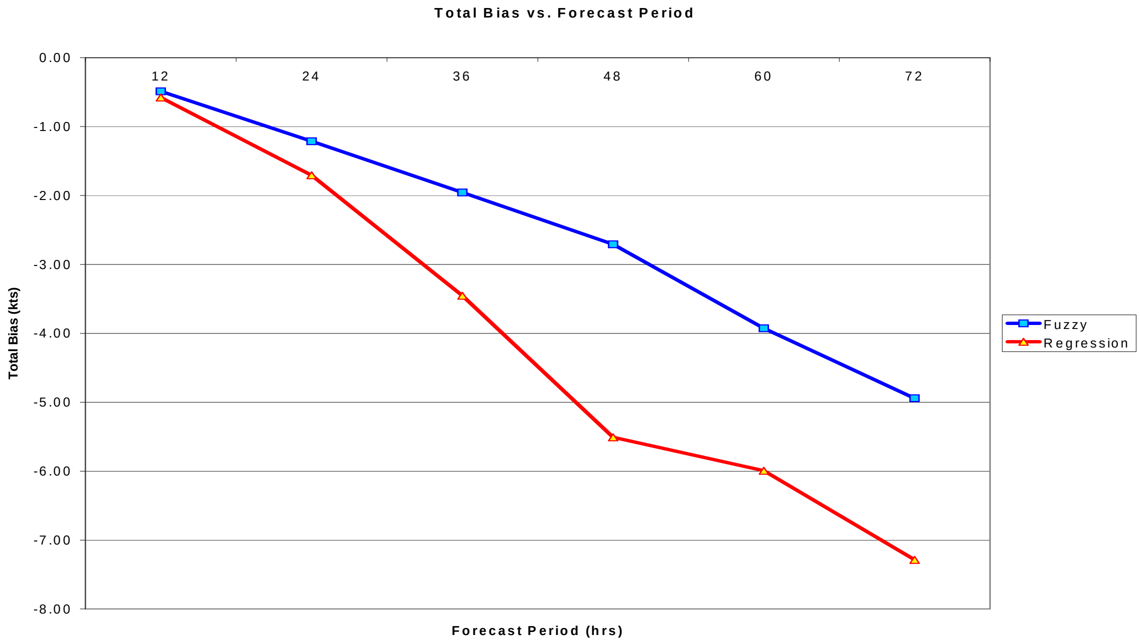

Fig.14. The mean forecast error (kts), or overall forecast bias, for the forecast time periods (hrs).

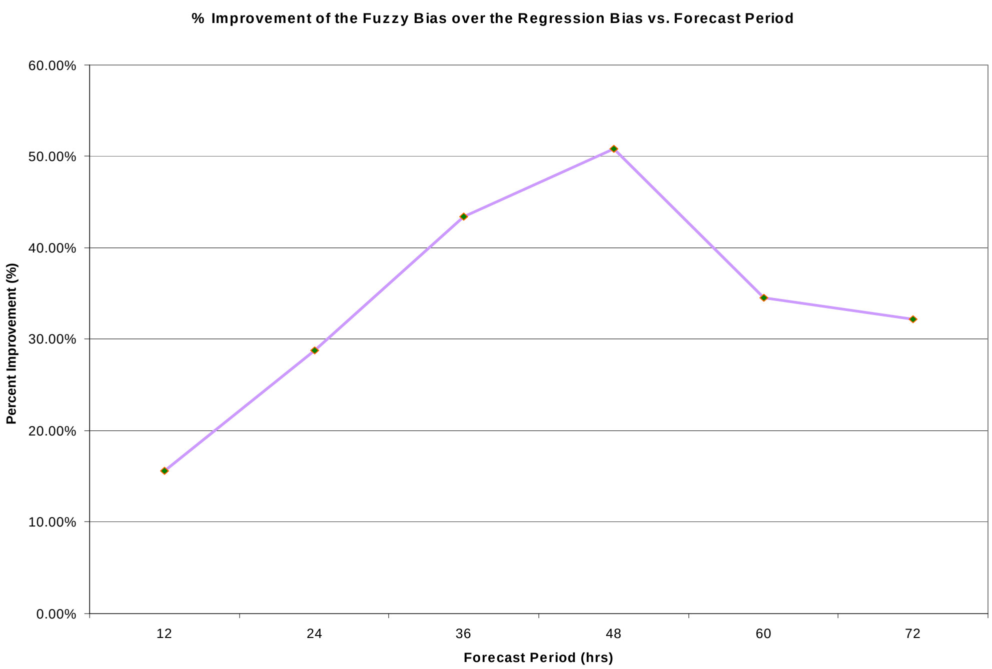

Fig. 15. The percentage of improvement of the FLAW bias over the MLR bias (%) for the forecast period (hrs).

From these results, we can conclude that while the average absolute errors of FLAW were comparable in magnitude to MLR, FLAW does produce a smaller overall forecast bias than MLR, which is an attractive benefit in operational forecasting.

ii. Case study of Hurricane Guillermo

The first year of the validation period witnessed two very intense storms. The first storm, Guillermo, was long-lived, lasting fifteen days. During this time, it underwent rapid intensification and then moderate weakening. Guillermo crossed out of the ENP basin into the Central Pacific basin and later recurved into the ENP at high latitude. The data during the recurvature was included once the storm traveled east of 140°W. At this point the storm was both at high latitudes and over very cold SST's. Both models produced large errors as the TCLAT and SST predictors went far below –1.

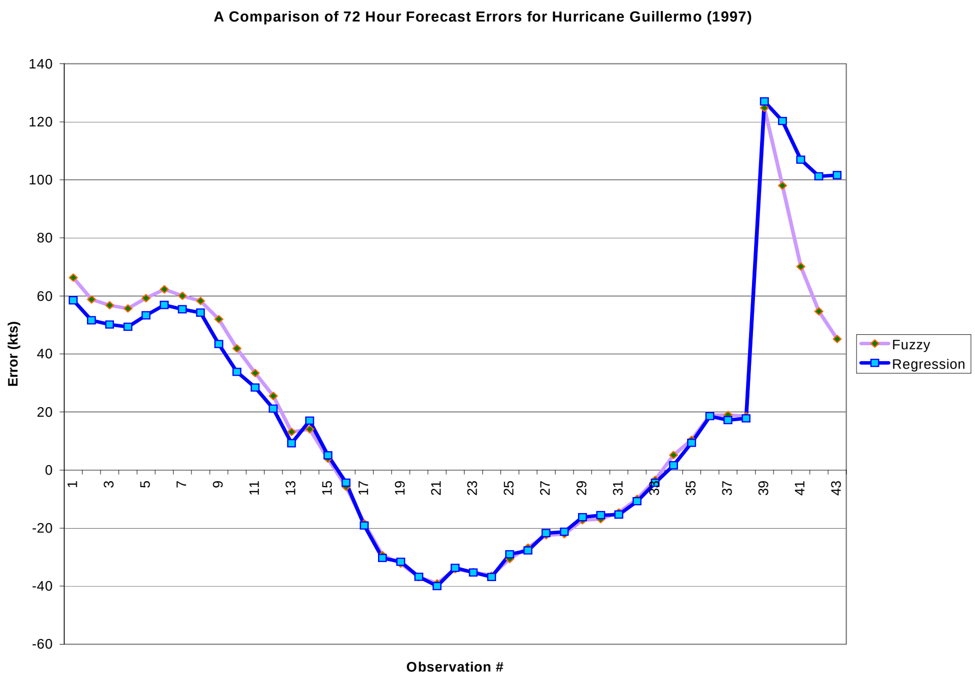

Fig.16. The 72-hour forecast errors (kts) for each observation of Hurricane Guillermo.

At observation number 39, the storm suddenly re-enters the ENP basin. Both models first forecast an extreme weakening (in this case, much more than the current intensity of the storm). Since the relationships between the predictors and intensity change are fixed for MLR, its error remains high. The FLAW model, on the other hand was able to 'learn' that storms at high latitudes and over very cool waters do not actually weaken that much, and after a few forecast periods, was able to tone down it's errors by decreasing the weights given to TCLAT and SST and increasing the weights given to VMAX and PVMAX. Thus, this example shows effective adaptation taking place. This also suggests that a check be placed on the forecasts so as not to allow the storm to obtain negative intensities.

iii. Case study of Hurricane Linda

Hurricane Linda, the strongest hurricane in the ENP, offers a challenging test to intensity change guidance, due to its extremely rapid intensification. 72-hours after formation, it was already nearing it's peak intensity of 160 kts. Since both the MLR model and FLAW model depend on data from at least one observation, these models mostly missed the intensification. The weakening stage of such a powerful storm provides a good test however.

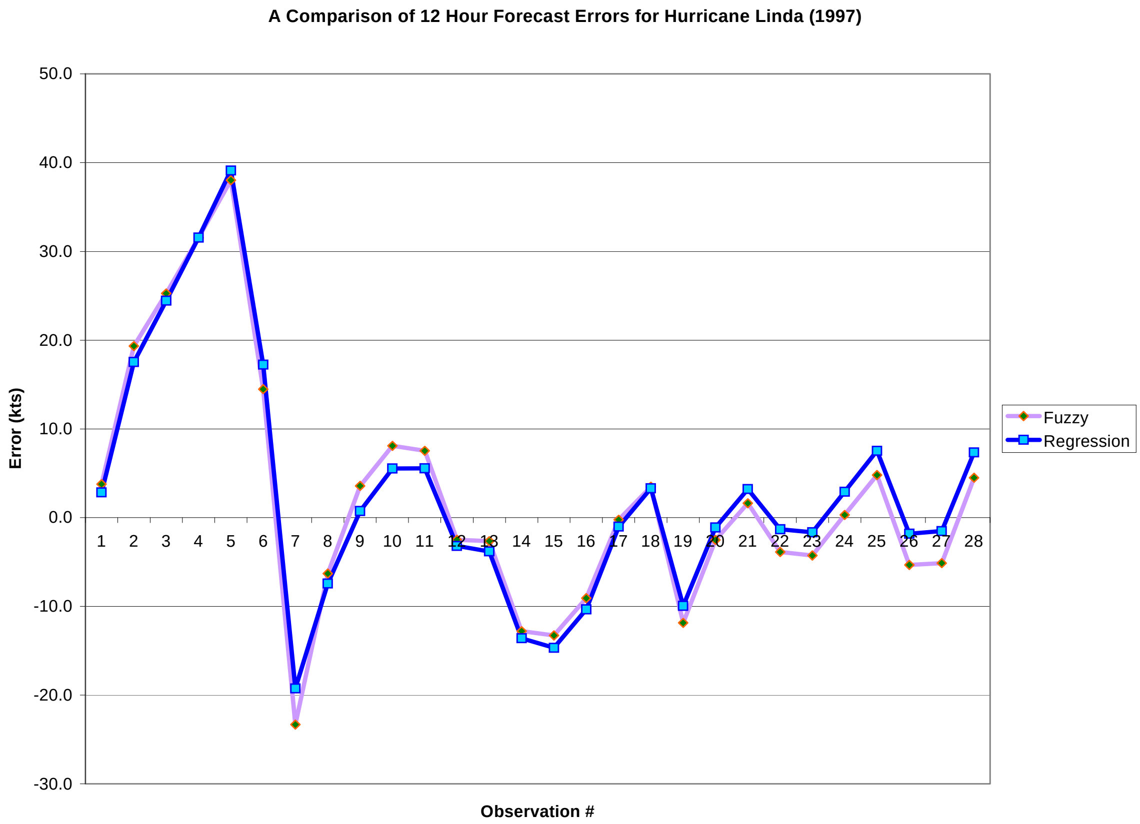

Fig. 17. The 12-hour forecast errors by observation #.

At 12-hours, both FLAW and MLR have similar error curves. At 72-hours, substantial differences appear. FLAW initially underestimates the initial intensification, but later does not also underestimate the weakening. It adapts well through the transition from intensification to weakening, and produces substantially smaller errors than MLR until near the final dissipation.

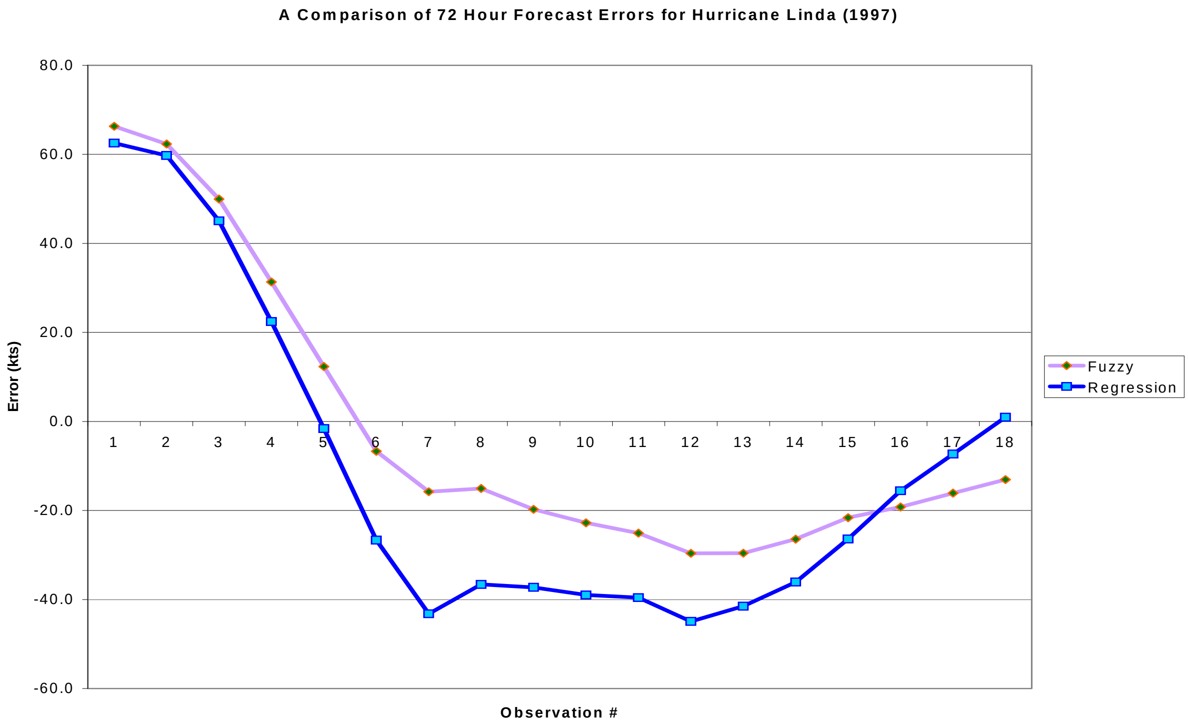

Figure # 18 shows the same as in Figure 17, except at the 72-hour forecast period.

The case study of Linda points to a weakness in both FLAW and MLR: the initial 72-hour forecast often completely misses rapid intensification events. MLR tends to forecast the 'normal' development rate, while FLAW does not have a chance to appreciably "learn" the storm until after the first couple forecast cycles.

5. Conclusions and Directions for Future Work

The goal of this research was to examine the efficacy of FLAW by comparing its forecasts to MLR. These models were implemented using Simulink and developed on the climatological and persistence variables: TCLAT, SST, VMAX, and PVMAX06 for the years 1982-96. They were then run during the 3-year validation period of 1997-99 and the average absolute error and the average error, or overall forecast bias, were compared. Several case studies were also examined.

Overall, FLAW produced errors comparable in magnitude to MLR, but with a smaller bias. FLAW demonstrated effective 'learning', or adaptation of the weights, in several case studies. This suggests that the model may have the potential to be a valuable tool for intensity change forecasting if further research is conducted.

There are many changes that might improve the accuracy of FLAW. One of the most important would be to properly "fuzzify" the predictor and truth data by mapping the data onto fuzzy sets following standard fuzzy techniques, instead of linear normalizations. Perhaps confidence fields could be calculated on the data, and a version of the algorithm that uses confidences could be used, and thereby take full advantage of the fuzzy feedforward SAM method. The sensitivity to changes in the step size and maximum delta weight should be examined. Also, constraints should be placed on the forecast (for example, do not allow an intensity change that gives the storm a negative intensity). A method of predicting formation and future predictor data could be used to help the model initialize for the first few forecast periods. Also, the maximum delta weight change could be linked to observation number, allowing the model to learn faster at the beginning of the storm. The model should be initialized at the start of each new storm. It would be nice if the algorithm allowed for negative weights.

If these changes result in substantial gains over MLR, then the model should be developed for use in other basins, with synoptic and dynamic data. Other fuzzy methods should also be tested, for example, to use in an expert forecast system to spot rapidly intensifying storms. Other avenues, such as neural networks, or fuzzy neural networks, should also be explored. Biak and Hwang have used a neural network to predict the intensity (lowest surface pressure) of western North Pacific typhoons with favorable results. With proper implementation, fuzzy logic may turn out to be an invaluable tool in the forecasting of hurricane intensity change.

ACKNOWLEDGEMENTS

This challenging and ambitious project would not have succeeded without the help and encouragement of numerous people at the NCAR. I owe a debt of gratitude to science research mentor, Kevin Petty, for the considerable time and effort he spent guiding my research and helping me to find and extract bugs from my programs. Kevin provided the FORTRAN program used to read the NHC best-track data set and Bill Thorson provided a code example used to read the OI-SST data set. I also owe much thanks to Shel Dalton for providing the original Simulink code for the Forecast Integrator. She spent considerable time helping me learn how to use and modify the code for the purposes of my project. I also would like to thank Diane Rabson for the constructive feedback she gave for this paper. Finally, I am grateful for the computing support provided by the RAP helpdesk from Jeff Stolte, Terri Eads, and Anne-Marie Tarrant.

CREDITS

This work was done under the auspices of the Significant Opportunities in Atmospheric Research and Science program of the University Corporation for Atmospheric Research, with funding from the National Science Foundation, the U.S. Department of Energy, the National Oceanic and Atmospheric Administration, the Cooperative Institute for Research in Environmental Sciences, and the National Aeronautics and Space Administration.

REFERENCES

Avila, L., 1998: Forecasting tropical cyclone intensity changes; An operational challenge. Preprints, *Symp. On Tropical Cyclone Intensity Change*, Phoenix, AZ, Amer. Meteor. Soc., 1-3.

Baik, J., and Hwang, H., 1998: Tropical cyclone intensity prediction using regression method and neural network. *J. Meteor. Soc. of Japan*, **76, 5**, 711-717.

Dalton, S., June 18, 1999: Forecast Integrator. Internal document, Research Applications Program, NCAR.

DeMaria, M., and J. Kaplan, 1994a: Sea surface temperature and the maximum intensity of Atlantic tropical cyclones. *J. Climate*, **7**, 1324-1334.

----, and ----, 1994b: A statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic Basin. *Wea. Forecasting*, **9**, 209-220.

----, and ----, 1999: An updated statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **14**, 326-337.

Frank, W., 1977: The structure and energetics of the tropical cyclone, I. Storm Structure. *Monthly Wea. Review*, **105**, 1119-1135.

Gray, W., Scheafer, J., and Landsea, C., 1997: Climate trends associated with multidecadal variability of Atlantic hurricane activity. *Hurricanes, Climate and Socioeconomic Impacts*, H.F. Diaz and R.S. Pulwarty, Eds.; Springer, 15-53.

Kosko, B., 1996: *Fuzzy Engineering*, Prentice Hall; pp. 549.

McNeill, D., and Freiberger, P., 1994: Fuzzy logic. Simon & Schuster.

Merril, R., 1988: Environmental influences on hurricane intensification. *J. Atmos. Sciences*, **45, 11,** 1678-1687.

Petty, K., and Hobgood, J., 2000: Improving tropical cyclone intensity guidance in the Eastern North Pacific. *Wea. Forecasting*, **15**, 233-244.

Pielke, R. and Landsea, C., 1998: Normalized hurricane damages in the United States 1925-95. *Weather and Forecasting*, **13**, 621-631.

Whitney, L., and Hobgood, J., 1997: The relationship between sea surface temperatures and maximum intensities of tropical cyclones in the eastern North Pacific Ocean. *J. Climate*, **10**, 2921-2930.

Willoughby, H., 1978: The vertical structure of hurricane rainbands and their interaction with the mean vortex. *J. Atmos. Sci*, **35**, 849-858.

----, 1990: Temporal changes in the primary circulation in tropical cyclones. *J. Atmos. Sci*., **47**, 242-264.

----, Clos, J., and Shoreibah, M., 1982: Concentric eye walls, secondary wind maxima, and the evolution of the hurricane vortex. *J. Atmos. Sci.*, **39**, 395-411.