Prediction of Rainfall from WSR-88D Radar Using Kernel-Based Methods

Article *in* International Journal of Smart Engineering System Design · October 2003
DOI: 10.1080/10255810390243737

3 authors:

Theodore B. Trafalis
University of Oklahoma

Budi Santosa
Institut Teknologi Sepuluh Nopember

Michael Richman
University of Oklahoma

DRAFT

PREDICTION OF RAINFALL FROM WSR-88D RADAR USING KERNEL-BASED METHODS

Theodore B. Trafalis
Budi Santosa
School of Industrial Engineering,
University of Oklahoma,
202 West Boyd, Suite 124,Norman OK, 73019-063
405/325-4347, ttrafalis@ou.edu

Michael B. Richman
School of Meteorology
University of Oklahoma
1100 East Boyd, Suite 1310
Norman, OK 73019-0626
405/325-1853, mrichman@ou.edu

ABSTRACT

The main objective of this paper is to utilize standard Support Vector Regression (SVR), Least Squares Support Vector Regression (LS-SVR), and compared to traditional regression method and rain rate formula, to facilitate rainfall estimation. Ground truth rainfall data are necessary to apply intelligent systems techniques. A unique source of such data is the Oklahoma Mesonet. Recently, with the advent of a national network of advanced radars (i.e., WSR-88D), massive archived data sets are available for data mining. The applications of SVR and LS-SVR are new for estimation of rainfall by radar.

The WSR-88D records digital database contains information of reflectivity as one of the native variables. The primary focus of the proposed research is to capitalize on reflectivity at multiple elevation angles and multiple bins in the horizontal for precipitation prediction. The SVR and LS-SVR model are used for precipitation prediction. For direct comparison to the existing scheme rainfall totals from the Oklahoma Mesonet are utilized for the training and verification data.

The results show that SVR and LS-LSVR are better in term of generalization error than traditional regression and rain rate formula used in meteorology. Moreover LS-SVR in this case shows a better performance than SVR

Keyword List I : Data Mining Applications

Keyword List II : data analysis, data base, estimation, generalization error, histogram analysis, kernel function, MATLAB, mean squared error, neural networks, prediction, principle component analysis, radial basis functions, RBF, regression analysis, support vector regression, least squares support vector regression, time series analysis, training sample

INTRODUCTION

Heavy rainfall, leading to flash floods, kills more people than lightning, tornadoes or hurricanes. Despite this, our ability to estimate precipitation and flooding from current state of the science technology is frequently inaccurate and, hence, there is ample impetus for improvement. Much of this inaccuracy arises from poor precipitation estimates from the Weather Surveillance Radar 1988 Doppler (WSR-88D) algorithms, which use only empirical techniques (OFCM 1990, OFCM 1991, Fulton et al. 1998).

Regression models have been applied as an approach for modeling of linear and nonlinear systems. In regular regression, the errors assumed to be normally, identically and independently distributed. It is believed that the assumptions are not satisfied in real applications. In fact, the error distributions are likely to be contaminated by occasional bad values giving rise to outliers.

Alternative approaches to regression are support vector machines (SVMs) and Least Squares Support Vector Machines (LS-SVM). By these approaches some assumption about the distribution of the data can be ignored. Capturing the radar signature and classifying this signature using ANNs has been accomplished by Skapura (1996).

Support Vector Machine, introduced by Vapnik over the past 30 years, has been successfully applied to solve numerous problems in classification and regression. Müller et al. (1997) used support vector regression (SVR) to predict time series data on a noisy Mackey Glass equation data and Santa Fe competition (set D) and compared the results with radial basis function networks. They demonstrated that SVR had excellent performance. Santosa and Trafalis (2001) used SVR along with a feedforward neural network and radial basis function (RBF) networks to predict monthly flour prices in three cities. The results also showed that SVR outperformed two other methods.

Least squares Support Vector Machine for function estimation (LS-SVR) was introduced by Saunders et al (1998) as an interpretation of ridge regression in dual variables space. This approach is closely related to support vector machine. Johan Suyken at al.(1999,2002) then developed least squares support vector machine and weighted least squares support vector machine for function estimation.

The main objective of the paper is to develop an approach, new to radar meteorology, that uses SVR and LS-SVR to improve WSR-88D rainfall estimation over existing methods.

PROBLEM STATEMENT

For over forty years, rainfall estimation from radar has been an active area of research. For the most part, the issue has been addressed through reflectivity-rainfall relations (known as Z-R). The Z-R relation was pioneered by Marshall and Palmer (1948). A method, based on matching the probabilities of the two variables, of deriving Z-R relation is presented by Rosenfeld et al. (1993). Currently, radar rainfall estimates were computed from a parametric Z-R relation that can be demonstrated in various ways. The most common form of this relationship can be written as follows:

$$RR = 10\,C^{(0.0625)\,(Z)}$$

where $C = 0.036$ mm/hr and Z is reflectivity for the lower elevation angle.

The current WSR-88D default values for *a* and *b* are *300* and *1.4*, respectively (Fulton et al. 1998). However, the values of *a* and *b* vary from place to place, season-to-season, and time to time (Wilson and Brandes 1979). Smith and Joss (1997) suggested that *b = 1.5* is comfortably in the middle of any likely range of variation. They emphasized the importance of the value of *a*. It is obvious that an estimation technique based on this relation will not be very successful since no single values for *a* and *b* will give a good estimate of the rainfall over a broad range of conditions. What does occur is a large uncertainty in estimating rainfall from reflectivity. The bias is not a constant and cannot be corrected with existing algorithms. Therefore, it is essential to develop a new technique that reduces this variability of the Z-R relation.

WSR-88D records digital base data containing three variables, namely velocity (V), reflectivity (Z), and spectrum width (W). Current rainfall detection algorithms use only Z data (Fulton et al., 1998). Our paper proposes to

- determine the suitability of SVR and LS-SVR for precipitation prediction,
- compare the forecasts of precipitation given by SVR and LS-SVR to those of traditional regression and formula-based results,
- determine the characteristics of the forecast errors.

By application of the SVR and LS-SVR, we hope to improve the precipitation processing subsystem (PPS) by creating new algorithms.

DATA AND ANALYSIS

The rainfall data for calendar year 1998 is collected from the Chandler, Oklahoma Mesonet station. The Oklahoma Mesonet is a statewide network of 119 automated weather observing stations, designed with dense spatial coverage across the state, and locating at least one observing station in each of Oklahoma's 77 counties (Basara et al., 1999). Nine main parameters are measured at all stations including rainfall. Rainfall is measured at 5-minute intervals. The data are heavily quality controlled and archived for the period 1994 to present with 99.8 percent availability.

The WSR-88D radar has an effective range of approximately 230 km. The digital base data containing the three variables to be used herein is limited to a range of 180 km. The radar used is located in Norman, OK (KTLX) and has complete coverage over a large number of Mesonet sites. The radar performs approximately 10 elevation scans that comprise of a single volume scan. Each elevation scan is at a particular elevation angle between 0.5 and 15 degrees above the horizontal. For each elevation scan, the radar revolves a full 360 degrees about the vertical and makes about 366 azimuthal scans. For the azimuthal on the five lowest elevation scans, we will utilize Z. Z is measured once in every kilometer or at 230 points for every azimuthal scan. The radar used in this study is approximately 54 km west-southwest of the Chandler raingauge.

Given the 5-minute time resolution of both observing and measurement platforms, it is clear that there will be a large amount of data. For each elevation angle used (5), there are 25 radar spatial variables selected over and around the Chandler area. Therefore, the radar volume scan contains 125 input variables. There were 576 five minute-period samples used for both the radar and raingage information. Owing to the highly correlated structure of the data, it is convenient to preprocess these data and apply data reduction techniques. In the present study, we were able to reduce 125 input variables to 9 orthogonal dimensions accounting for 99 percent of the variance of the original data using principal components.

METHODOLOGY

Support Vector Regression (SVR)

Suppose we have been given training data $\{x_k, y_k\}_{k=1}^{\ell}$, with input data $x_k \in \mathbb{R}^n$ and ouput data $y_k \in \mathbb{R}$. By support vector regression, one wants to find a function *f(x)* that has at most ε deviation from the actual target y_i for all training data with small length of *w* (Fig. 1). For the linear case, suppose we have the following function as a regressor:

$$f(x) = \langle w,x \rangle + b, \tag{1}$$

where <.> denotes the dot product in *X*.

One should seek minimum norm of *w* that is equivalent with maximizing the distance between the function f to the farest points in the training set, and *b* is a threshold. One way to accomplish this objective is by minimizing the Euclidean norm $\| w \|^2$. Then we have the following problem in primal weight space:

$$\text{minimize } \frac{1}{2} \| w \|^2$$

$$\text{subject to} \begin{cases} y_i - \langle w,x_i \rangle - b \le \varepsilon \\ \langle w,x_i \rangle + b - y_i \le \varepsilon \end{cases} \tag{2}$$

We assume that there is a function *f* that approximates all pairs *(x_i, y_i)* with precision *ε*. In this case we assume that the problem is feasible. In the case of infeasibility where some points might deviate from *f* ± *ε*, one can introduce slack variables ξ, ξ* to cope with infeasible constraints of the optimization problem. Then the above problem can be formalized as (Vapnik, 1995):

$$\text{minimize } \frac{1}{2} \| w \|^2 + C\sum_{i=1}^{\ell} (\xi_i + \xi_i{*})$$

$$\text{subject to} \begin{cases} y_i - \langle w,x_i \rangle - b \le \varepsilon + \xi_i \\ \langle w,x_i \rangle + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi{*}_i \ge 0 \end{cases} \tag{3}$$

The constant C > 0 determines the trade off between the flatness of function *f* and the amount up to which deviation larger than ε are tolerated. Any deviation more than ε will be penalized with C. Solving (3) in dual space is more easily, especially when extended to nonlinear case. By using Lagrangian duality theory (Haykin, 1999), we get

$$w = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) x_i \tag{4}$$

and therefore $f(x) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) < x_i, x > +b = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) < x_i, x > +b,$ (5)

In (4) *w* can be completely described as a linear combination of training patterns x_i. A kernel function, *k(x, x')*, can be used to substitute the dot product $\langle x, x' \rangle$. In case of nonlinear function each data point x in the input space is mapped into the higher dimensional feature space with map φ. In the new space dot product $\langle x, x' \rangle$ becomes $\langle \varphi(x), \varphi(x)' \rangle$. A nonlinear kernel function, *k(x,x')*, can be used to substitute the dot product $\langle \varphi(x), \varphi(x)' \rangle$. In the higher feature space we can construct liner regression function that represents nonlinear function in te input space. There are three nonlinear kernel functions usually used in the SVR:

(1) polynomial: $(x^T x_i + 1)^p$,

(2) radial basis function (RBF): $\exp(-\frac{1}{2\sigma^2} \| x - x_i \|^2)$,

(3) tangent hyperbolic (sigmoid): $\tanh(\beta_0 x^T x_i + \beta_1)$.

The best kernel function which one can use to substitute the dot product depends on the data. One has to do some experiments to find the best kernel function that fits the data.

Loss Function

In order to define the differences between *f(x)* and the actual data *y*, we need to define a loss function C(.). There are several loss functions that can be imposed in the support vector regression. Specifically,

ε-insensitive loss function

$$C(f(x) - y) = \begin{cases} | f(x) - y | - \varepsilon & for\ | f(x) - y | \geq \varepsilon \\ 0, & otherwise \end{cases} \tag{6}$$

Using this loss function leads to the following solution (Gunn, 1998)

$$\max_{\alpha,\alpha^*} w(\alpha,\alpha^*) = \max_{\alpha,\alpha^*} -\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(x_i, x_j) + \sum_{i=1}^{l}\alpha_i(y_i - \varepsilon) - \alpha_i^*(y_i + \varepsilon) \tag{7}$$

or

$$(\overline{\alpha}, \overline{\alpha}^*) = \arg\min_{\alpha,\alpha^*} \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(x_i, x_j) - \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)y_i + \sum_{i=1}^{l}(\alpha_i + \alpha_i^*)\varepsilon \tag{8}$$

subject to

$$0 \leq \alpha_i, \alpha_i^* \leq C,\ i = 1,2\ldots,l$$

$$\sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0$$

Solving (7) or (8) determines the Lagrange multipliers α_i, , α_i* and the regression function is given by

$$f(x) = \sum_{SVs}(\overline{\alpha}_i - \overline{\alpha}_i^*)K(x_i, x) + \overline{b} \tag{9}$$

where

$$< \overline{w}, x >= \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) K(x_i, x_j) \tag{10}$$

$$\overline{b} = \frac{1}{NSV} \sum_{SVs} (Y_i - x_i^T \overline{w}) \tag{11}$$

The support vectors are points where exactly one of the Lagrange multipliers is greater than zero**.**

Quadratic loss function

$$C(f(x) - y) = \begin{cases} | f(x) - y |^2 \;\; for \; | f(x) - y | \geq \varepsilon \\ \frac{1}{2}(f(x) - y)^2, \qquad otherwise \end{cases} \tag{12}$$

With this loss function the solution is given by (Gunn, 1998)

$$\max_{\alpha,\alpha^*} w(\alpha, \alpha^*) = \max_{\alpha,\alpha^*} -\frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i, x_j) + \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) y_i - \frac{1}{2C} \sum_{i=1}^{l} (\alpha_i^2 + \alpha_i^{*2}) \tag{13}$$

and let $\beta_i = \alpha_i - \alpha_i^*$, then we have

$$\min_{\beta} -\frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \beta_i \beta_j K(x_i, x_j) + \sum_{i=1}^{l} \beta_i y_i + \frac{1}{2C} \sum_{i=1}^{l} \beta_i^2 \tag{14}$$

$$\text{st} \; \sum_{i=1}^{l} \beta_i = 0$$

and the regression function is given by Equations 9.

$$\overline{b} = mean(Y_i - \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) K(x_i, x_j))$$

Least Squares Support Vector Regression (LS-SVR)

Suppose we have the same data set as in SVR, by using Least squares approach we consider the following optimization problem:

$$\min_{w,b,e} J(w, e) = \tfrac{1}{2} w^T w + \tfrac{1}{2} C \sum_{i=1}^{\ell} e_k^2 \tag{15}$$

such that

$$y_k = w^T \varphi(x_k) + b + e_k, \; k = 1, \ldots \ell,$$

where φ maps the data in the input space to the higher dimensional feature space. Term C in the cost function *J* is the regularization parameter for which the deviation of the predicted value from y_k will be penalized. In primal weight space our model is:

$$y(x) = w^T \varphi(x) + b, \tag{16}$$

The dimension of weight vector w can be infinite which makes a calculation of w impossible in general. To get rid of this problem, we need to convert the above problem into dual problem. Using Lagrange formulation we have:

$$L(w, b, e; \alpha) = J(w, e) - \sum_{k=1}^{\ell} \alpha_k \{w^T \varphi(x_k) + b + e_k - y_k\} \tag{17}$$

where α_k are Lagrange multipliers. The conditions for optimality are given by:

$$\begin{aligned}
&\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{k=1}^{\ell} \alpha_k \varphi(x_k) \\
&\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{k=1}^{\ell} \alpha_k = 0 \\
&\frac{\partial L}{\partial e_k} = 0 \rightarrow \alpha_k = Ce_k, \qquad k = 1,..\ell \\
&\frac{\partial L}{\partial \alpha_k} = 0 \rightarrow w^T \varphi(x) + b + e_k - y_k = 0, \; k = 1,..\ell
\end{aligned} \tag{18}$$

By some manipulations the solution of the above problem is given by

$$\left[\begin{array}{c|c} 0 & 1_v^T \\ \hline 1_v & \Omega + \frac{1}{C} I \end{array}\right] \left[\begin{array}{c} b \\ \hline \alpha \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline y \end{array}\right] \tag{19}$$

where y = $[y_1;..;y_\ell]$, $1_v = [1;..;1]$, $\alpha = [\alpha_1;\ldots;\alpha_\ell]$ and $\Omega_{kj} = \varphi(x_k)\varphi(x_j)$ for $k, j = 1..\ell$.

One solves a linear system (19) to find the optimal solution of (15). In case of large linear systems it is necessary to apply iterative methods as suggested in Golub et al. (1989).

We note that the values of α_k are proportional to the errors at the data points. This is different to that of SVR where one looses sparseness property in LS-SVR.

EXPERIMENTS

All experiments were performed using a computer Pentium IV. The SVR and LS-SVR experiments were performed in MATLAB environment. For SVR, an RBF kernel function was used with some values of parameters C and spread σ. Finally we found that the best parameters values are 10000 for C and 0.7 for σ. The codes used are from Gunn (1998). For LS-SVR polynomial kernel was used. LS-SVR experiments were performed using LS-SVMlab by Pelckman et al. (2002). Some statistics computations were done with Splus.

RESULTS

Results of Linear Modeling of Radar Inputs

Z was used as a predictor with the rainfall as the response variable. The V and W were not used herein based on poor performance in previous results (Trafalis et al., 2002). Using the full set of 576 observations, multiple linear regression of Z on rainfall rate used independent predictors from the set of the 125 highly correlated Z predictors. This was accomplished by using principal component (PC) scores as predictors. The principal component model written in terms of the original standardized variables (Z) is $Z = FA^T$, where F are the PC scores and A are the PC loadings (Richman, 1986). The PC scores are a new coordinate system that has uncorrelated dimensions and frequently is used as a data compression tool. Owing to their property of mutual uncorrelatedness, the PC scores satisfy assumption that the predictors are uncorrelated in regression. In the present case, for Z, 125 correlated variables are represented by 8 uncorrelated PC scores with 99 percent of the original variance explained (the last 1 percent is considered to represent pure noise

Results of Analysis

Results of analysis of the radar inputs for predicting rainfall rate are now presented. Three types of regression methods are used in this work, traditional linear regression, SVR and LS-SVR.

TRAINING

Linear regression

We apply linear regression to facilitate rainfall prediction. Reflectivity (Z) data at 8 PC scores representing 125 points in space obtained from WSR-88D. These were used as inputs to linear regression. There are 550 observations for the training and 26 observations for testing. The dimension of each data point is a 1 by 125 column vector. Each model is run ten times (replications). The 550 observations for training are selected randomly without replacement from 576 observations. Then, from ten replications, the average MSE is obtained. The average training MSE for the model is 0.40 whereas the average testing MSE is 0.52 as shown in Table 1.

Analysis of the scatterplot (Fig. 2) of the observed versus predicted values of rainfall illustrate how ineffective the model is in capturing any functional relation. The one-to-one line, representing a perfect fit, is far removed from the main groupings of variables. Analysis of the residuals, defined as observed minus predicted , for the linear regression model (Fig. 3) indicates a mean of -0.08191 with a minimum of –1.1032 and a maximum of 3.2194. The standard deviation is 0.4623 with a skewness of 3.6302 and a kurtosis of 18.1134.

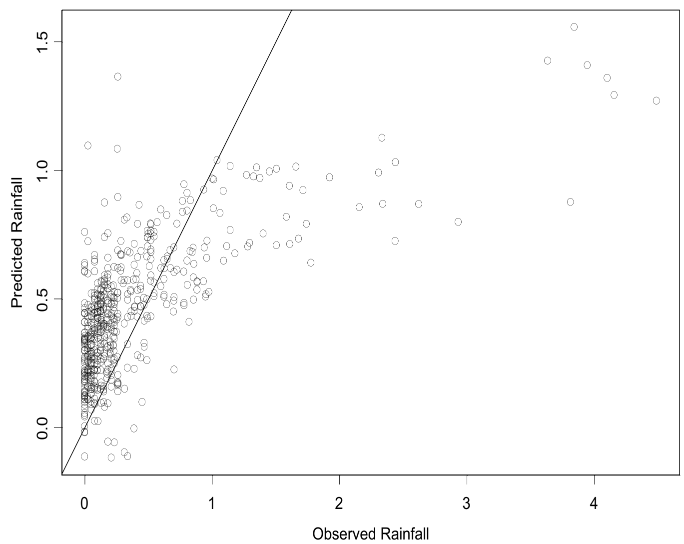

Figure 2. Scatter Plot of Observed vs Predicted Rainfall with traditional Regression. The solid line is perfect fit

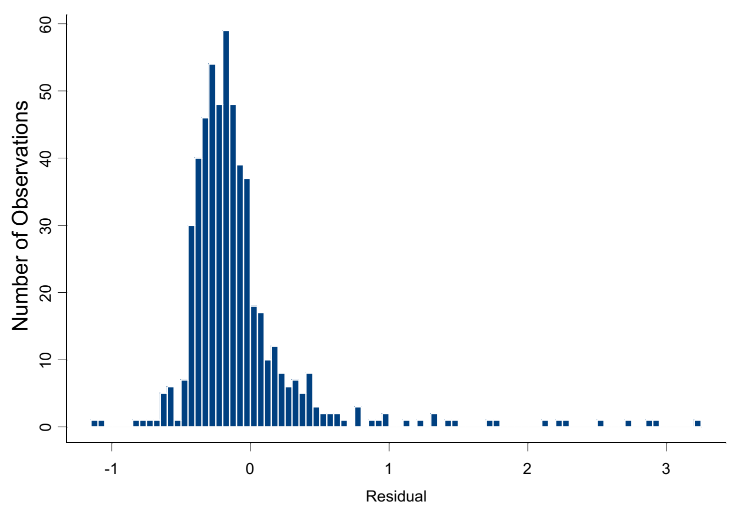

Figure 3. Histogram of Residuals with traditional Regression.

Analysis of the scatterplot (Fig. 4) of the observed versus predicted values of rainfall for testing illustrate how ineffective the model is in capturing any functional relation. The one-to-one line, representing a perfect fit, is far removed from the main groupings of variables. Analysis of the residuals, defined as observed minus predicted, for the linear regression model (Fig. 5) indicates a mean of 0.035 with a minimum of –.4279 and a maximum of .3275. The standard deviation is

0.1840 with a skewness of -0.6066 and a kurtosis of –0.0624. The average testing R^2 is 0.243.

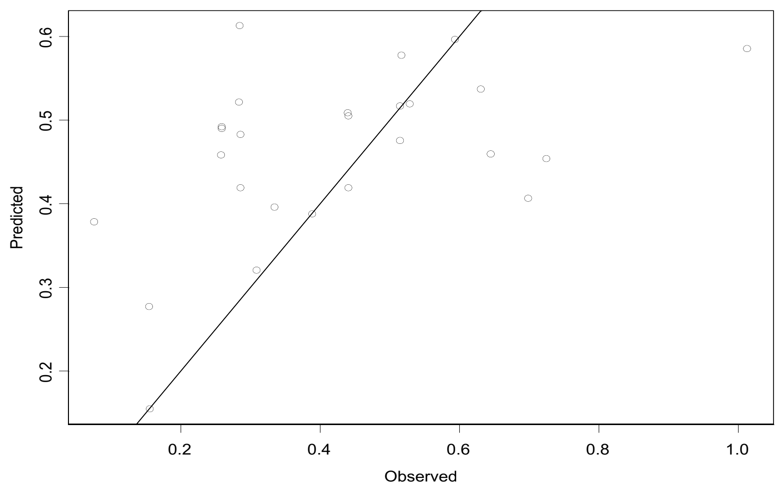

Figure 4. Scatter Plot of Observed vs Predicted Rainfall with traditional Regression. The solid line is perfect fit

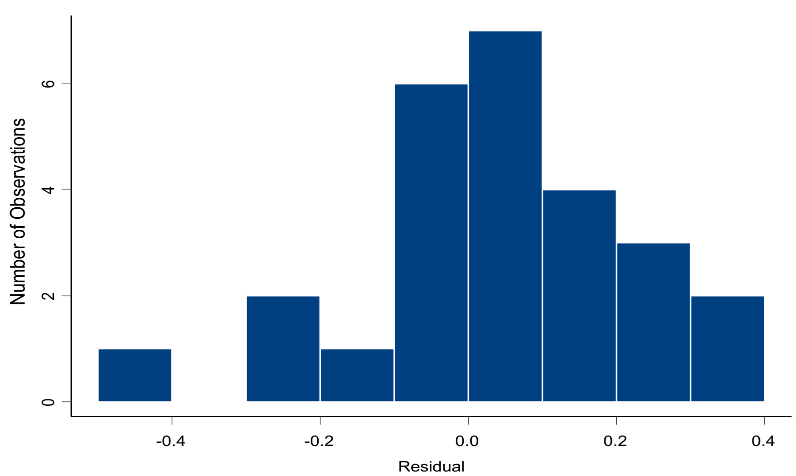

Figure 5. Histogram of Residuals with traditional Regression.

Support Vector Regression

We apply SVR since theoretically it is more efficient in terms of generalization (Vapnik, 1995). Additionally, the poor performance of the traditional regression warrants analysis with an alternative methodology. Reflectivity (Z) data at 8 PC scores representing 125 points in space were used as inputs to SVR. There are 550 observations for the training and 26 observations for testing. The dimension of each data point is a 1 by 125 column vector. Each model is run ten times (replications). The 550 observations for testing are selected randomly without replacement from 576 observations. Then, from ten replications, the average MSE is obtained. The average training MSE for the model is 0.20 whereas the average testing MSE is 0.32 as shown in Table 1. . Analysis of the scatterplot (Fig. 6) of the observed versus predicted values of rainfall illustrate how much more effective the SVR model is in capturing the functional relation. The one-to-one line, representing a perfect fit, is much closer to the main groupings of variables than in the case of linear regression. Analysis of the residuals, defined as observed minus predicted, for the support vector regression model (Fig. 7) indicates a mean of –0.0000 with a minimum of –2.2399 and a maximum of 0.9580. The standard deviation is 0.2499 with a skewness of –2.951 and a kurtosis of 20.0520.

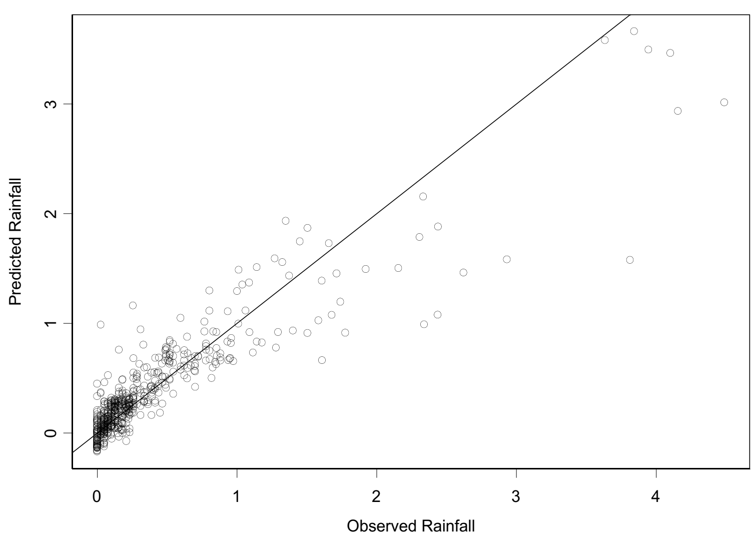

Figure 6. Scatter Plot of Observed vs Predicted Rainfall with SVR. The solid line is perfect fit

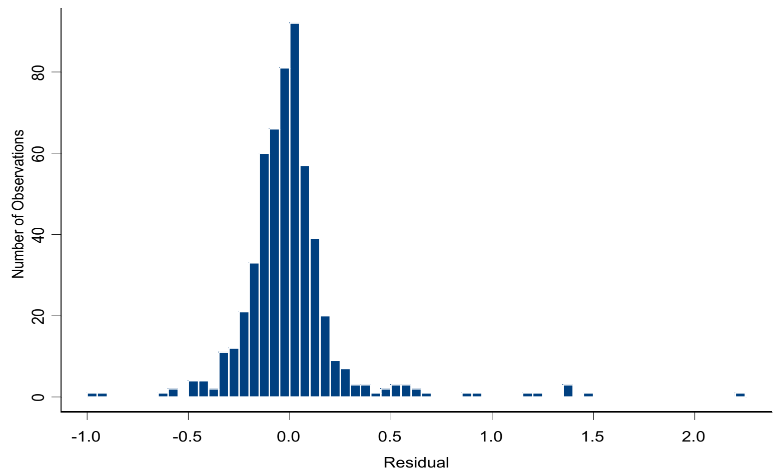

Figure 7. Histogram of Residuals with SVR

Analysis of the scatterplot (Fig. 8) of the observed versus predicted values of rainfall for testing illustrate how much more effective the SVR model is in capturing the functional relation. The one-to-one line, representing a perfect fit, is much closer to the main groupings of variables than in the case of linear regression. Analysis of the residuals, defined as observed minus predicted, for the support vector regression model (Fig. 9) indicates a mean of –0.0239 with a minimum of –0.4631 and a maximum of 0.3181. The standard deviation is 0.1924 with a skewness of –0.2766 and a kurtosis of –0.4826. The average testing R^2 is 0.309

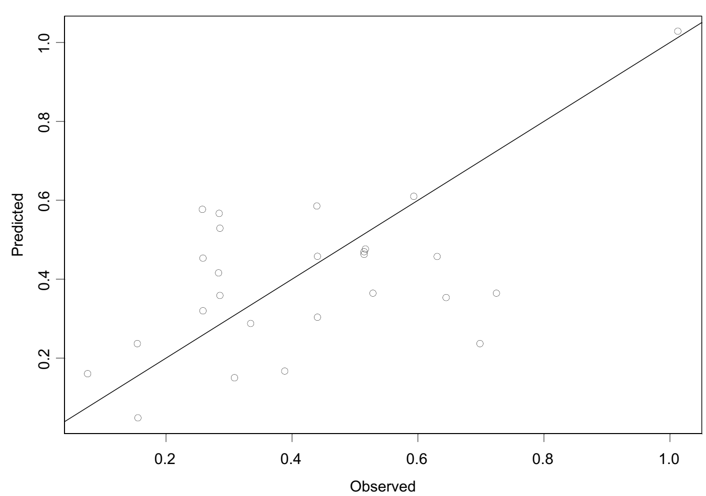

Figure 8. Scatter Plot of Observed vs Predicted Rainfall with SVR. The solid line is perfect fit

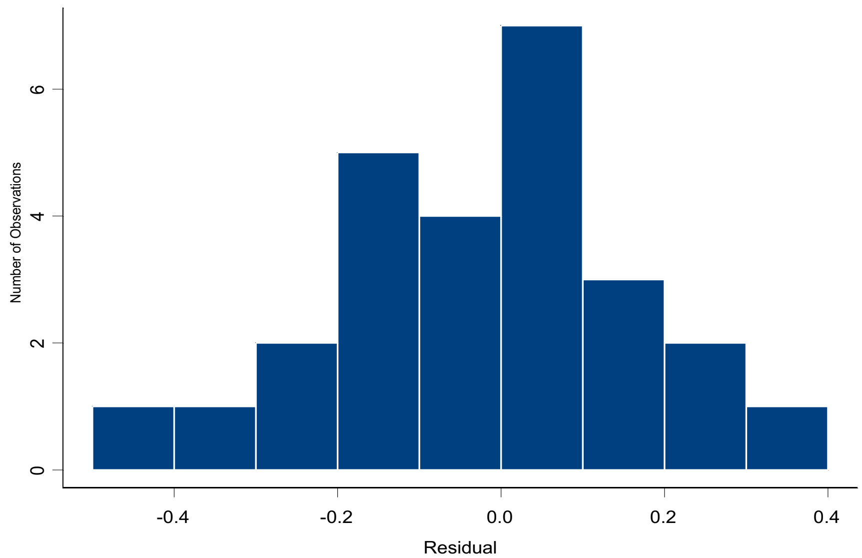

Figure 9. Histogram of Residuals with SVR

Least Squares Support Vector Regression

We apply LS-SVR as another approach similar to SVR. The input data are the same as those used in regression and SVR. From the experiments we obtained the average training MSE for the model is 0.18 whereas the average testing MSE is 0.27 as shown in Table 1. Analysis of the scatterplot (Fig. 10) of the observed versus predicted values of rainfall illustrate how much more effective the SVR model is in capturing the functional relation. The one-to-one line, representing a perfect fit, is much closer to the main groupings of variables than in the case of linear regression.

Analysis of the residuals, defined as observed minus predicted, for the LS-SVR model (Fig.11) indicates a mean of –0.0000 with a minimum of –2.1138 and a maximum of 0.9945. The standard deviation is 0.2565 with a skewness of –2.2742 and a kurtosis of 16.1147.

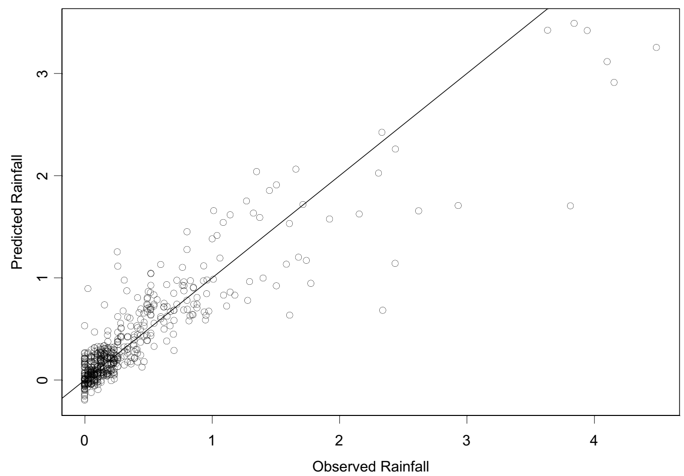

Figure 10. Scatter Plot of Observed vs Predicted Rainfall with LS-SVR.
The solid line is perfect fit

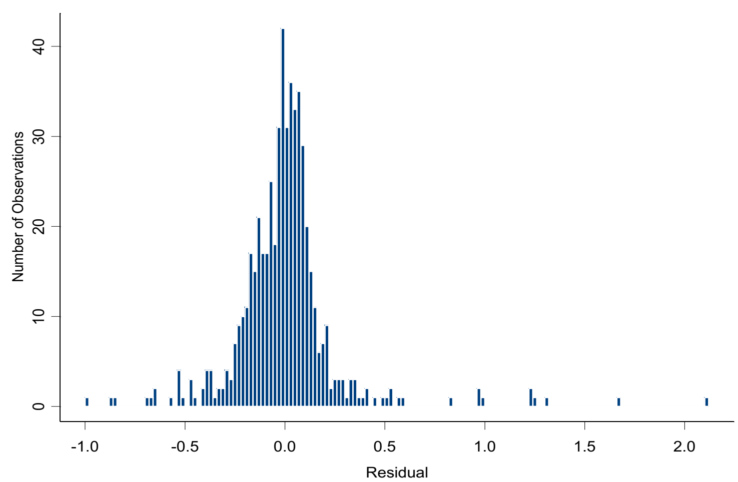

Figure 11. Histogram of Residuals with LS-SVR

Analysis of the residuals of testing data, defined as observed minus predicted, for the least squares support vector regression model (Fig. 13) indicates a mean of –0.0319 with a minimum of –.4405 and a maximum of 0.3733. The standard deviation is 0.1864 with a skewness of –0.1436 and a kurtosis of -0.2453. The average testing R^2 is 0.339.

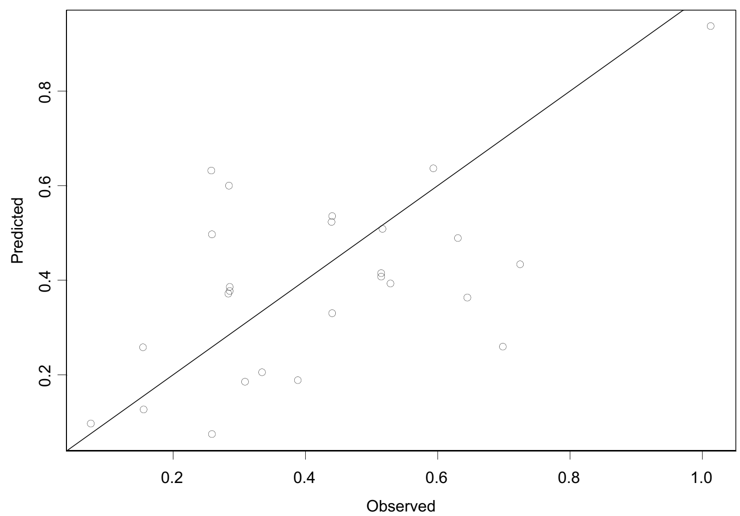

Figure 12. Scatter Plot of Observed vs Predicted Rainfall with LS-SVR.
The solid line is perfect fit

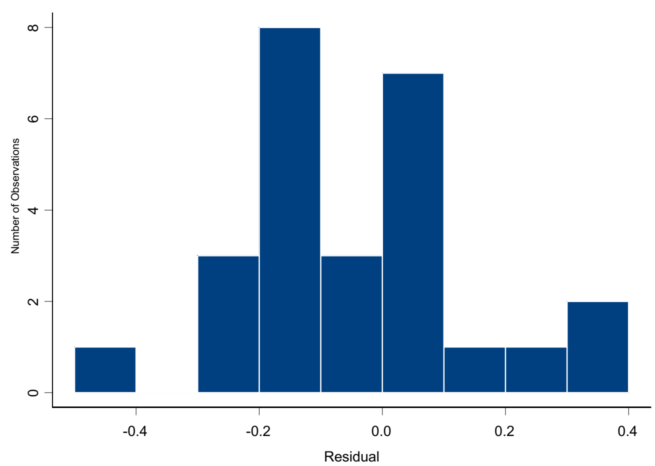

Figure 13. Histogram of Residuals with LS-SVR

Rate of Precipitation Formula

From the experiments we obtained the average training MSE for the model is 0.18 whereas the average testing MSE is 0.27 as shown in Table 1. Analysis of the scatterplot (Fig. 14) of the observed versus predicted values of rainfall illustrate how much more effective the SVR model is in capturing the functional relation. The one-to-one line, representing a perfect fit, is much closer to the main groupings of variables than in the case of linear regression.

Analysis of the residuals, defined as observed minus predicted, for the Formula model (Fig.15) indicates a mean of –0.0877 with a minimum of –1.2458 and a maximum of 2.6541. The standard deviation is 0.3227 with a skewness of 2.9824 and a kurtosis of 19.2189.

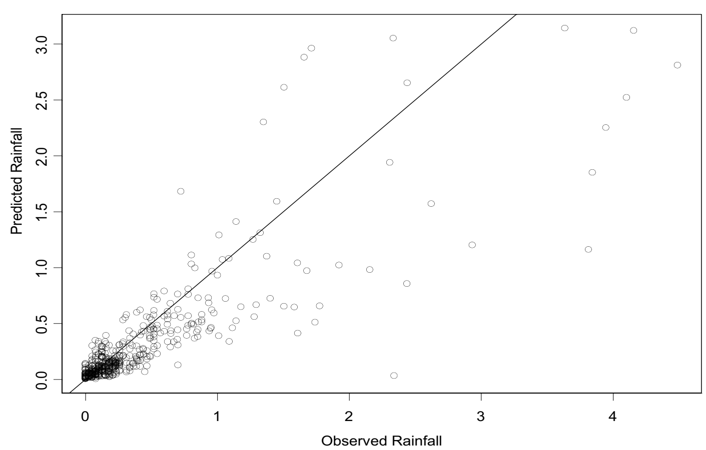

Figure 14 Scatter Plot of Observed vs Predicted Rainfall with Formula.The solid line is perfect fit

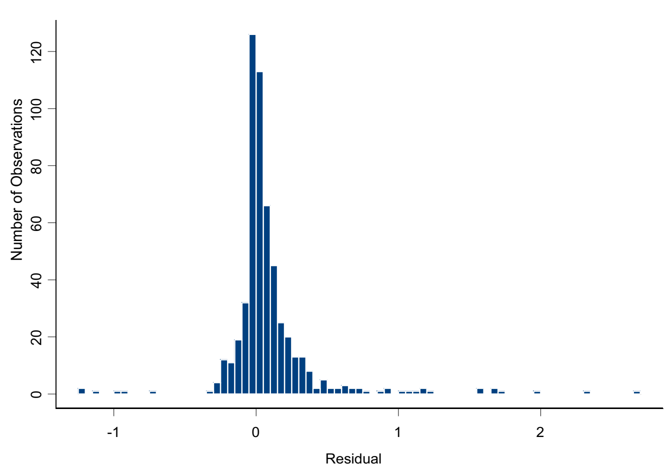

Figure 15. Histogram of Residuals with Formula

Analysis of the residuals, defined as observed minus predicted, for the Formula model (Fig.17) indicates a mean of –0.1336 with a minimum of –0.5161 and a maximum of 0.4292. The standard deviation is 0.1970 with a skewness of 0.3928 and a kurtosis of 0.8522. The average testing formula R^2 is 0.225.

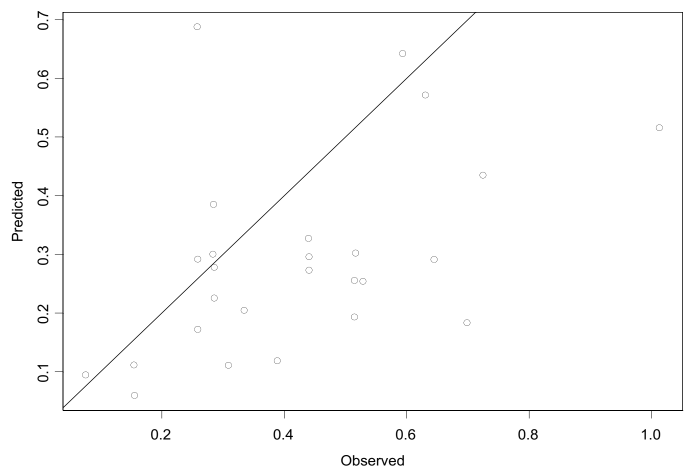

Figure 16 Scatter Plot of Observed vs Predicted Rainfall with Formula. The solid line is perfect fit

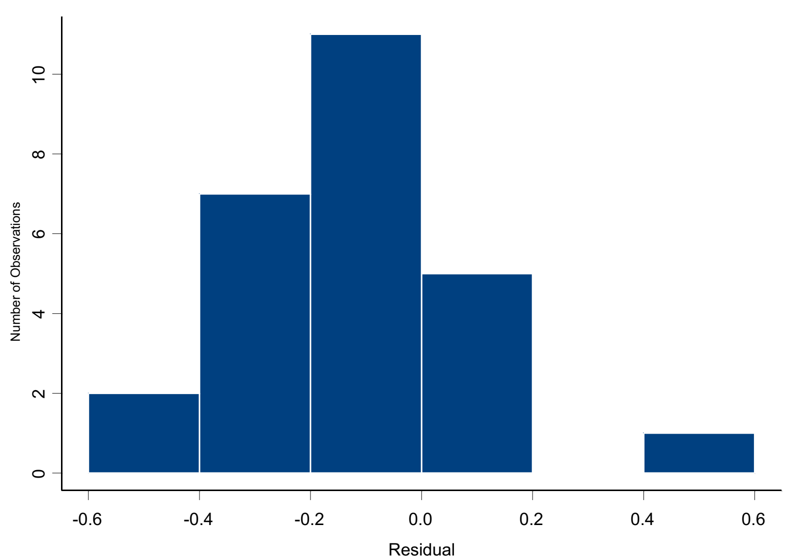

Figure 17. Histogram of Residuals with Formula

Comparisons of the two regression models

Comparison of Figs. 2 and 4 present a compelling picture where the SVR is superior to traditional linear regression. The one-to-one perfect fit line is much closer to the data points with large observed rainfall. This is particularly important if one wants to predict flooding situations with the heaviest rainfall rates. Similarly, comparison of the residuals of both models (Figs. 3 and 5) indicates that the SVR regression residuals have a lower dispersion about the mean and are less skewed. This is in direct agreement with the difference in MSE values for both models and the R^2 that is 24.8 percent larger for SVR.

Given the good results using Z as the input variables, both Z and W were attempted in SVR. The MSE (Table 1) was much larger, so no further investigation of those results will be presented at this time.

Comparisons of regression models and artificial neural networks

Compared to previous results (Trafalis et al, 2001), the SVR behaved similarly to the best ANN architecture (ZW) (Table 1). However, with the SVR, there is an added benefit of locating a global optimum solution. Therefore, we do not have to replicate the experiments, several times, as in the ANNs to find the best ANN architecture.

SUMMARY AND CONCLUSIONS

Analysis of the native variables from the WSR-88D radar in Norman, OK and experiments for predicting rainfall rates are undertaken with regression, SVR and ANNs. A 5 by 5 grid of 1 km boxes for the radar data is constructed, centered on the Chandler, OK Mesonet rain gauge. The importance of using these Mesonet data are that they are recorded in 5 minute intervals which nearly match the volume scan time of the radar. Additionally, the lowest five elevation angles are used to provide a three-dimensional snapshot of the atmosphere above the raingauge using 125 discrete boxes. The goal of the research is predictive data mining. We attempt to describe how the 125 grid boxes interrelate for the native variables during periods of rainfall. SVR modeling of the radar data suggests that reflectivity is the most predictive of the three variables examined. The SVR results were superior to traditional linear regression model and comparable in accuracy to the best ANN architectures tested. In the case of SVR, the best architecture is automatically determined, as opposed to the ANN case where extensive experimentation is required to identify the best architecture. However, in SVR, we need to tune several parameters and select the best kernel. The generalization properties of the used architectures are promising. In the near future, we plan to investigate more efficient methods to select the parameters and the most efficient kernel related to the specific data. In that regard, we plan to extract additional rainfall data and expect that by increasing the number of data will provide better estimation results. Experimentation with ensemble techniques will be investigated also.

Acknowledgements. This research was supported by DOC/NOAA Grant NA67RJ0150. We extend thanks to Tim O'Bannon for assistance in interpreting the radar tapes and providing program to interpret Level 2 radar data and to Tim Kwiatkowski for programming assistance. Additional gratitude is extended to Suat Kasap for data processing assistance and Jean Shingledecker for administrative assistance. This research was performed on The University of Oklahoma ECAS Cray-J90

REFERENCES

Basara, J. B., J. A. Brotzge, and K. C. Crawford, 1999: Investigating land-atmosphere interactions using the Oklahoma Mesonet. *Third Symposium on Integrated Observing Systems,* January 10-15, Dallas, TX, Amer. Meteor. Soc. 120-123

Bazaraa, M. S., H. D. Sherali, and C. M. Shetty, 1993: *Nonlinear Programming: Theory and Algorithms.* John Wiley & Sons, Inc., New York.

Fulton, R. A., J. P Breidenbach, D.-J. Seo, and D. A. Miller, 1998: The WSR-88D rainfall algorithm. *Weather and Forecasting, 13,* 377-395.

Golub G.H., Van Loan C.F. *Matrix Computations*. Third Edition, Ed. Johns Hopkins, 1996.

Gong, X-F., and M. B. Richman, 1995: On the application of cluster analysis to growing season precipitation data in North America east of the Rockies. *J. Climate, 8,* 897-931.

Haykin, Simon, 1999: *Neural Network: A Comprehensive Foundation*. Prentice Hall, New Jersey.

J.A.K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, 2002 :*Least Squares Support Vector Machines*, World Scientific, Singapore.

K.R. Müller, A.J. Smola, G. Rätsch, B. Schölkopf, J. Kohlmorgen, and V.N. Vapnik. Predicting time series with support vector machines. In W. Gerstner, A. Germond, M. Hasler, and J.-D. Nicoud, editors, *Articial Neural Networks* | ICANN'97, pages 999 - 1004, Berlin, 1997. Springer Lecture Notes in Computer Science, Vol. 1327.

Marshall, J. S., and W. M. Palmer, 1948: The distribution of raindrops with size. *Journal of Meteorology, 5,* 165-166.

Montroy, D. L., 1997: Linear relation of central and eastern North American precipitation to tropical Pasific sea-surface temperature anomalies. *J. Climate, 10,* 541-558.

Montroy, D. L., M. B. Richman, and P. J. Lamb, 1998: Observed nonlinearities of monthly teleconnections between tropical Pacific sea-surface temperatures and central eastern North American precipitation. *J. Climate, 11,* 1812-1835.

Navonne, H. D. and H. A. Ceccatto, 1994: Predicting Indian monsoon rainfall: a neural network approach. *Climate Dynamics, 10,* 305-312.

OFCM, 1990: *Doppler Radar Meteorological Observations, Part B, Doppler Radar Theory and Methodology*. Federal Meteorological Handbook 11, FCM-H11B-1990, Office of the Federal Coordinator for Meteorological Services and Supporting Research, Washington, D. C.

OFCM, 1991: *Doppler Radar Meteorological Observations, Part C, WSR-88D Products and Algorithms*. Federal Meteorological Handbook 11, FCM-H11C-1991, Office of the Federal Coordinator for Meteorological Services and Supporting Research, Washington, D. C.

Pelckmans K., Suykens J.A.K., Van Gestel T., De Brabanter J., Lukas L., Hamers B., De Moor B., Vandewalle J., 2002: *LS-SVMlab : a Matlab/C toolbox for Least Squares*

Support Vector Machines, Internal Report 02-44, ESAT-SISTA, K.U.Leuven (Leuven, Belgium).

Richman, M. B., 1986: Review paper: Rotation of principal components. *International J. Climatology, 6,* 293-335.

Richman, M. B., and X-F. Gong, 1999: Relationships between the definition of hyperplane width to the fidelity of principal component loading patterns. *J. Climate, 12,* 1557-1576.

Santosa, B. and Trafalis T.B., Predicting Monthly Flour Prices through Neural Networks, RBFs and SVR, *Proceeding of ANNIE Conference 2001*, St. Louis, Missouri, November 14-17, 2001, ASME Press.

Smith, P. L., and J. Joss, 1997: Use of a fixed exponent in "adjustable" Z-R relationships. *Proceedings of the 28th Conference on Radar Meteorology,* September 7-12, Austin, TX, Amer. Meteor. Soc., 254-255.

S.R. Gunn, Technical Report, Dept. of Electronics and Computer Science, *Support Vector Machines for Classification and Regression*, University of Southampton, (Southampton, U.K.), 1998.

Trafalis, T.B., A. White, B. Santosa and M. Richman, 2001: Development of an intelligent system for improved WSR-88D rainfall estimation", *Intelligent Engineering Systems Through Artificial Neural Networks*, (C.H. Dagli, A.L. Buczak, J. Ghosh, M. Embrechts, O. Ersoy and S. Kercel, eds.), ASME Press, 11, 703-708.

Trafalis, T.B., A. White, B. Santosa and M.B. Richman, 2002: Data mining techniques for improved WSR-88D Rainfall estimation. *Computers in Industrial Engineering*, *43*, 775-786 .

Vapnik, V.N, 1995: *The Nature of Statistical Learning Theory*. John Wiley & Sons, New York

Wilson, J. W., and E. A. Brandes, 1979: Radar measurement of rainfall- a summary. *Bulletin American Meteorological Society, 60,* 1048-1058.

Xiao, R., and V. Chandrasekar, 1997: Development of a neural network based algorithm for rainfall estimation from radar observations. *IEEE Transactions on Geoscience and Remote Sensing, 35*, 160-171.

Table 1. MSE of training and testing for Regression, SVR and LS-SVR (mm^2)

Method	Predictors	Training Average MSE	Testing Average MSE
Regression	PC Z	0.4006	0.5184
SVR	PC Z	0.1507	0.3099
LS-SVR	PC Z	0.1754	0.2709
Formula	Z	0.2887	0.3728

Table 2. Residual training and testing for Regression, SVR and LS-SVR (mm^2)

Method	**Min**		**Max**		**Mean**		**Standard Deviation**		**Skewness**		**Kurtosis**	
	train	test	train	test	train	test	train	test	train	test	train	test
Regression	-5.939	-0.428	1.043	0.3275	0.034	0.035	0.720	0.1840	-4.917	-0.6066	32.627	-0.0624
SVR	-2.239	-0.4631	0.9580	0.3181	-0.000	-0.0239	0.2499	0.1924	-2.951	-0.2766	20.0520	-0.4826
LS-SVR	-2.1138	-0.4405	0.9945	0.3733	-0.000	-0.0319	0.2565	0.1864	-2.2742	-0.1436	16.1147	-0.2453
Formula	-1.2458	-0.5161	2.6541	0.4292	-0.0877	-0.1336	0.3227	0.1970	2.9824	0.3928	19.2189	0.8522

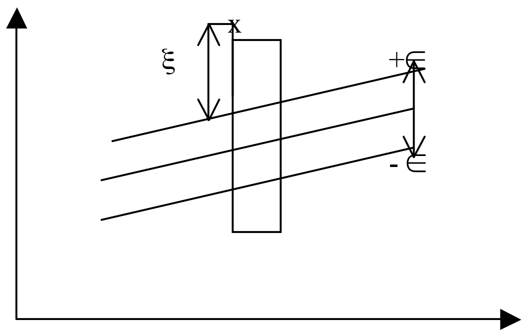

Figure 1. Linear non-separable SVR