International Journal of Industrial Engineering, 15(1), 26-35, 2008.

Extracting Knowledge of Concrete Shear Strength From Artificial Neural Networks

William A. Young II[1], Gary R. Weckman[2], Michael D. Brown[3], and Jim Thompson[4]

[1,2]Department of Industrial and Manufacturing Systems Engineering
[3,4]Department of Civil Engineering
Ohio University
Stocker Center
Athens, Ohio 45701 USA

This article introduces an artificial neural network (ANN) to estimate the shear strength of reinforced concrete beams. Current methods for calculating shear strength use a model that is based on engineering mechanics and empirical values determined though testing of a beam failing due to shear. The current methods are intended to provide a conservative lower bound on the strength needed to prevent a shear failure. A database containing the results of over 1200 laboratory shear strength tests was used to train an ANN. The database contained the geometric and material property data from the test specimens and the recorded failure load. The ANN presented in this paper was able to predict the shear strength of reinforced concrete beams more accurately than the current approach. The ANN provides additional insight on the parameters that are most significant in estimating concrete shear strength, which may lead to a better understanding of the mechanism of shear failure.

Significance: Current methods for calculating shear strength use a model that should provide a conservative lower bound. A more accurate model has the potential to improve structural design of reinforced concrete beams with improved accuracy. The ANN provides additional insight on the parameters to better understand the mechanism of shear failure.

Keywords: Concrete shear strength, artificial neural networks, knowledge extraction, Levenberg-Marquardt learning algorithm, generalized feed-forward network.

(*Received 11 July 2006; Accepted in revised form 5 March 2007*)

1. INTRODUCTION

The current provisions of the American Concrete Institute's Building Code Requirements for Structural Concrete (called the ACI code in this paper) (ACI 318, 2005) use a model for predicting the shear strength of reinforced concrete beams that is based approximately on a physical understanding of the mechanisms of shear resistance in concrete beams. This model incorporates strength contributions from the concrete and the steel reinforcement. The model assumes a concrete beam without shear reinforcement has certain strength, and this is roughly the concrete contribution to shear strength. The shear strength estimate for the concrete contribution is based on tests of beams without shear reinforcement, which have a significant amount of variance (or scatter). The steel reinforcement contribution to the shear strength is taken as the force developed by the steel reinforcement that crosses an assumed inclined crack, with the force in the reinforcement assumed to be at its yield force. The contributions from the concrete and steel are then summed to determine the overall shear strength.

In attempting to estimate the response of a physical system to external inputs, engineers typically employ either mathematical or physical models. Physical models are the first model of choice when dealing with any phenomena, if the physical relationships are well known. However, the relationships between the parameters in the problem and the physical phenomena are not always clearly understood. Mathematical models are often limited because they do not capture necessary relationships to model accurately the physical interactions. Artificial Neural Networks (ANN)s are able to fill the void between the limitations of mathematical models and the unknowns of physical models.

Mathematical models that attempt to mimic the ability of the human mind to recognize patterns are called ANNs. McCulloch and Pitts developed the foundations of ANN as early as the 1940s, and Rumelart and McClelland (1986) applied ANN to practical applications in the 1980s. The capabilities of ANN are currently being realized in science, and are classified into three types of applications; function fitting, pattern recognition, and classification. Other common names for ANN include connectionism, adaptive systems, adaptive networks, neuro-computers, and parallel distribution processors (Itani and Yacoub, 2000). ANNs are also called vector maps. Vector maps accept a feature vector from one data space and produce a feature vector in another data space. This transformation of vector space has been referred to as an emergent computation because the input of the ANN disappears and becomes unidentifiable once inside the network, and emerges as an output vector space (Felker et al., 2004).

Multilayer feed-forward networks are the most commonly used ANN because of their effectiveness in generalizing a wide amount of data (Sanad and Saka, 2001). These multi-layered, feed-forward networks use supervised learning and an error-back propagation algorithm in order to minimize the error by propagating errors back through the network until the network has learned (Girish et al., 2003). An alternative to back-propagation is called counter-propagation. In these networks, training is much faster but is not as versatile in modeling abilities and is slower at producing results. Often, the key to accurately modeling a system depends on choosing the correct model based on the volume and quality of the training data (Umakant, 1999). Genetic algorithms are also used to increase accuracy of the ANN (Sanad and Saka, 2001).

The data used for network training was divided into three parts: training, cross validation, and testing sets. Cross validation and testing data sets are independent of the data used for training. Cross validation techniques are used to stop the process of training if the network shows over-fitting, or memorization (Boulle et al., 2001). Memorization occurs when there is a lack of generality in the training data. The network training is discontinued when the network starts to fit the training data well, but large errors are seen in the cross-validation data. When the amount of data available is rather high, it is common practice to use 60% of the data for training, 20% for cross validation, and 20% for testing (De Alcantara and Gasparini, 2005). These proportions have been used for the analysis described herein.

2. LITERATURE REVIEW

The primary obstacle inhibiting the use of ANNs in practice comes from the lack of understanding and skepticism toward them. However, geo-mechanical and pavement systems have begun to recognize the advantage of ANNs because of their successful applications in other decision-making sciences (Adeli, 2001). Often scientific problems are too complex, poorly understood, resource-intensive, or a combination thereof to solve using traditional computational methods. ANNs have been used to solve these complex problems where traditional methods have failed (Umakant, 1999).

In recent studies, ANN were used, along with image analysis of the concrete structures, to optimize the natural gravel content in order to reduce material cost for a 5% to 15% savings (Kalliomaki et al., 2005) ANNs have also been used to estimate shear strength of reinforced concrete deep beams. For example, tests revealed that the developed ANN model produced an average ratio of measured to calculated shear strength of 0.99 which is considerably more accurate than other methods such as ACI, strut-and-tie, and Mau-Hsu which produced ratios of 2.08, 0.85, and 0.84 respectively (Sanad and Saka, 2001). Finally, a significant portion of the cost of pavement systems comes from the cost for rehabilitation, resurfacing, and further construction. Identifying key parameters for concrete reliability is very important in order to minimize cost. ANNs can help to determine design characteristics which would increase the reliability of a structure given various input variables (Felker et al., 2004).

There are benefits to using ANN in researching historic data of specific optimal design configurations of concrete structures. Efforts have been made to optimize the design of concrete beams under different loading conditions without the use of ANNs. These efforts are time consuming and complicated because designers must go through an iterative process to reproduce desired characteristics of concrete structures (Hadi, 2003). The Hadi research has found that adding additional hidden layers past one did not significantly increase accuracy but increased training time from 6 to 17 hours.

Depending on the complexity of the network being analyzed, the time needed for training can increase rapidly. There has been an effort to generate ANNs for physical systems that are concise based on dimensional analysis. Dimensional analysis helps unlock embedded information within a physical system by transforming the representation of the input data to the network. The theory behind concise network design is based on generating a set of dimensionless parameters, which is based on the Buckingham Pi theorem (Gunaratnam et al., 2003). This theorem stipulates the requirements for finding a number of independent dimensionless variables sufficient to model accurately a physical system. The theorem identifies the high-level estimating parameters needed for the ANN. These techniques use feature selection, where insignificant variables are discarded, and feature extraction, where variables are combined to reduce the number of input variables for the ANN. When dimensional analysis is used correctly it has been shown that performance indicators such as correlation coefficient and standard deviation are improved (Gunaratnam et al., 2003). This method achieves the realization of a concise network by reducing the amount of vector space the network must find to minimize the input-output relationships. However, generating concise ANNs may be difficult if expert knowledge is not known. The most significant benefit of using ANNs is that prior understanding about the physical nature of the input and output variables are not needed (Shahin et al., 2001).

3. PROBLEM BACKGROUND

The input and output variables used for shear strength prediction of a reinforced concrete beam can be found in Table 1. Many of the input variables are depicted in Figure 1.

Table 1. Concrete Input Variables

Input	Definition
b	width of the beam in inches
d	effective depth of the beam in inches
f'_c	compressive strength of concrete used in the beam in psi
f_y	yield strength of the shear reinforcement in ksi
ρ_l	longitudinal reinforcement ratio (steel reinforcement by area of concrete)
ρ_v	vertical shear reinforcement ratio
ρ_{vh}	longitudinal shear reinforcement ratio
a	shear span in inches (distance from applied load and the support)
$^a/_d$	shear span-to-depth ratio
l_n	clear span of the beam in inches
Output	Definition
V_{test}	measured shear strength of the specimen at the face of the support in kips
V_{pred}	predicted shear strength at the face of the support in kips

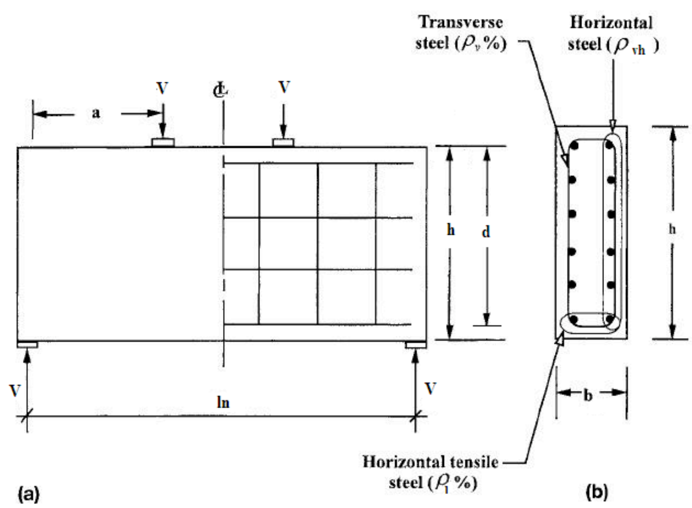

Figure 1. Basic Parameter for Deep Beam Model a) elevation b) Cross-Section (adapted from Sanad and Saka, 2001)

The current model that is used to calculate the strength of concrete is not intended to be a best-fit line through the data. The model attempts to estimate the lower bound of the beam strength as shown in Equation 1.

$$V_n = 2 \cdot \sqrt{f'_c} \cdot bd + \frac{A_v f_y d}{s} \qquad \ldots \qquad (1)$$

The objective of this research is to determine a model that estimates concrete shear strength more accurately. The accuracy of current efforts to model concrete shear strength can be improved. In order to improve the current model, an ANN was investigated. After the most accurate model was generated, further investigation was performed in order to determine if inputs to the model could be eliminated without sacrificing accuracy. After this point, an effort was made to extract knowledge from this network potentially to provide additional insight on the physical phenomena being modeled.

4. ARTIFICIAL NEURAL NETWORK

Determining a network structure to be used for prediction can be time consuming because there are several factors that can affect the accuracy of the model. In order to determine the network architecture, several configurations need to be studied. The structures tested differ in the network type, processing element used, and learning algorithms. The type of network tested was either a multi-layer preceptor or a generalized feed-forward network. The processing elements used in this study were tan-axon or sigmoid elements. Finally, the learning algorithms differed and were tested using either momentum or Levenberg-Marquardt techniques. Table 2 summarizes the combinations of learning, network types, and processing elements used in the testing. A combination of Levenberg-Marquardt learning algorithm, generalized feed-forward network type, and sigmoid processing elements produced the highest average value of R^2. This combination also produced the highest R^2 in any test throughout any of the trials.

Table 2. Training Combination Summary

Learning	**Network Type**	**PE Type**	**Avg. R^2**	**Max R^2**	**Min R^2**	**Range**
LM	GFF	SM	0.877	0.903	0.832	0.071
LM	GFF	TA	0.874	0.903	0.850	0.053
LM	MLP	TA	0.866	0.892	0.841	0.051
MOM	MLP	TA	0.834	0.835	0.832	0.004
MOM	GFF	TA	0.832	0.866	0.807	0.059
LM	MLP	SM	0.827	0.828	0.827	0.001

Since it takes some amount of effort to accurately collect input variables of concrete structures, there may be a need to reduce the network in size in terms of input parameters. Reducing the number of inputs to an ANN also has the potential of increasing the model's accuracy (Fonseca, et al., 2003). A network was constructed using each of the eleven input parameters in Table 1. Table 3 provides the results of the training of the network.

Table 3. Best ANN Performance with 11 Variables

Parameter	**Value**
Sum Sq.	394246.941
Avg. Error	5.916
Std.	52.613
R^2	0.892

In order to test whether limiting the number of input parameters will increase the ANN's prediction accuracy, sensitivity needed to be examined. Figure 2 shows what range each variable has in estimating the shear strength of concrete. This figure represents a cumulative sensitivity which is based on the average results of individual sensitivity analysis trials. From this figure, it appears that the specimen's shear strength is the least sensitive to the variable ρ_{vh} and most sensitive to d. Variable a/d was calculated from other input variables and an investigation was performed to see if this variable could be removed without losing accuracy in the model.

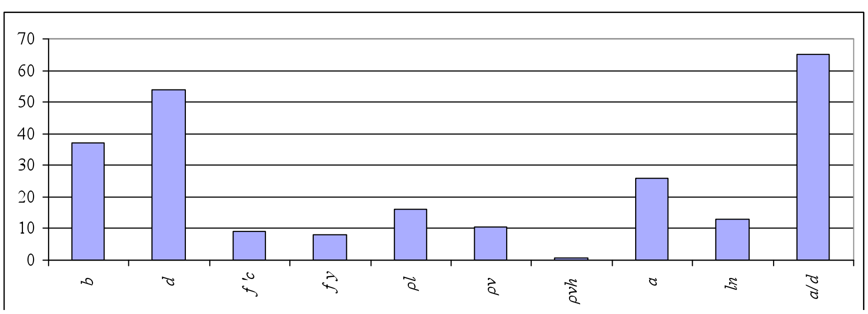

Figure 2. Cumulative Sensitivity Analysis for all 10 Variables

Table 4 shows the results of traning the network with the eight remaining input variables, after eliminating ρ_{vh} and a/d from the input parameters. From this table, the value of R^2 is 0.903, which indicates that this estimate is a good fit to the data and even better than the model trained with eleven input variables.

The sensitivity of the training with eight variables is shown below in

Figure 3. Further reduction of the number of input variables caused a loss of accuracy, so the number of inputs to the network was not reduced further.

Table 4. Best ANN with all 8 Variables

Parameter	Value
Sum Sq.	353411.642
Avg. Error	0.243
Std.	51.979
R^2	0.903

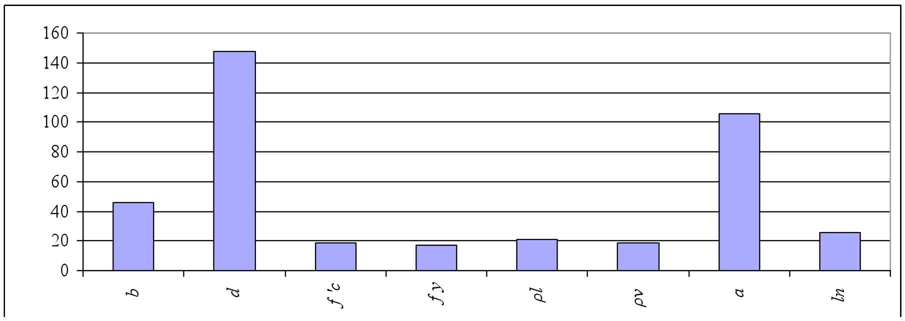

Figure 3. Variable Sensitivity with 8 Variables

Figure 4 presents a graph of the ANN estimate (*Vpred*) versus the actual values (*Vtest*) of concrete shear strength. This graph illustrates that plotted values below the actual line (45°) are estimated low, and the values above the line are estimated high. This figure indicates that lower values shear strength fit well along the line, and higher estimates fit more accurately than had previously been modeled with 11 variables.

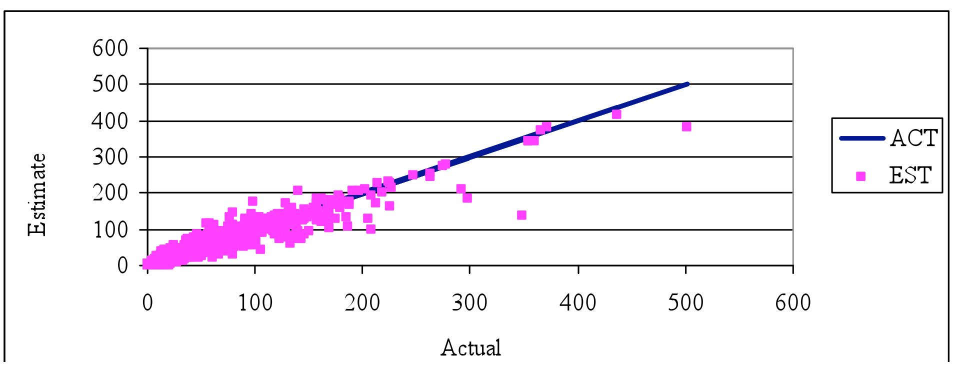

Figure 4. Neural Network Estimate with 8 Variables vs. Actual Concrete Shear Strength

5. INFORMATION EXTRACTION

Once the ANN is determined, information can be extracted from the results. In many cases, ANNs are asked to explain the unknown, and extracting information from the developed network helps to bridge this gap. Pulling information from the ANN can help researchers and engineers better understand the physical nature of the system being studied. It can also be a way of measuring the validity of the model if physical reactions are observed and known.

5.1 2D Relationships

Through sensitivity analysis relationships can be made that demonstrate how shear strength reacts to each variable as other input variables are held fixed. Figures 5, 6, 7 and 8 are examples illustrating the relationships between four different input variables (d, a, b, and ln) and the shear strength ($Vtest$). The shapes of the relationships are sometimes linear but also often non-linear. In general, the relationships and trends shown in the figures are consistent with currently accepted knowledge of the shear strength of reinforced concrete beams.

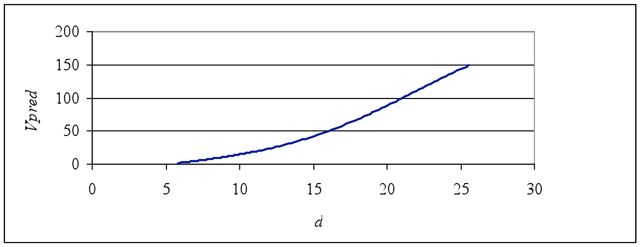

Figure 5. V_{pred} vs. d

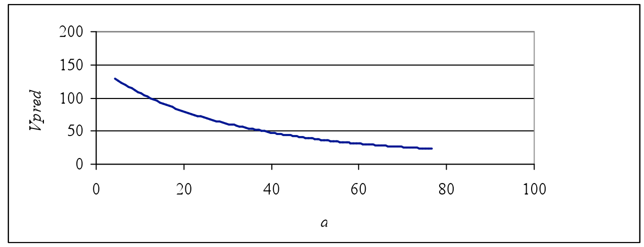

Figure 6. V_{pred} vs. a

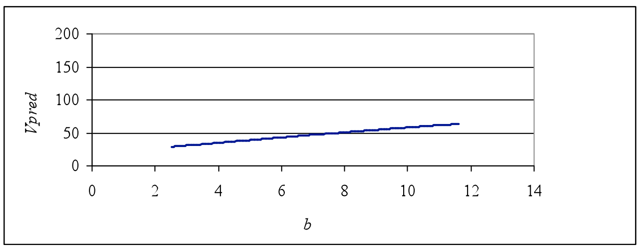

Figure 7. V_{pred} vs. b

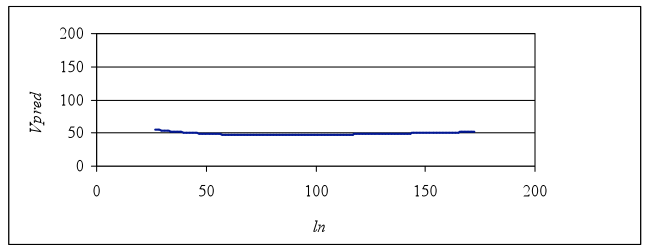

Figure 8. V_{pred} vs. ln

5.2 3D Relationships

Similar to the two-dimensional analysis described, a three-dimensional analysis can also be constructed in order to investigate the surface relationship of input parameters and the shear strength of concrete. Figure 9 shows how the input variables shear span (a) and depth of the beam (d) change the surface of the shear strength. From this figure, it can be concluded that having specimens with low shear span values and high depth values produce the largest shear strength.

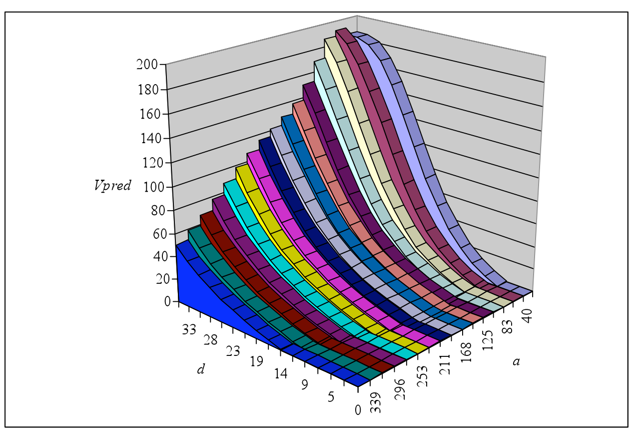

Figure 9. 3D Surface of d vs. a vs. $Vpred$

Another example of how input variables change the surface of shear strength in concrete can be seen in

Figure 10. In this figure, one can observe that the combination of a high clear span (l_n) and a large width of the concrete beam (b) will produce high shear strengths in concrete.

A three-dimensional analysis can be created in order to demonstrate how certain input variables affect other selected input variables. For the following example shown in Figure 11, shear span (a) and the depth of the concrete beam (d) are varied to create a surface graph of the calculated width of a concrete beam (b). To solve for this desired input variable, the remaining inputs as well as concrete strength are held at an average value. Thus, this figure suggests that there are several combinations to the input values that will result in the desired concrete shear strength by adjusting *a, d,* and *b* according to the surface response shown below.

Another example of three-dimensional surface analysis is shown below in Figure 12. This surface illustrates what values the vertical reinforcement ratio (ρ_v) and the longitudinal reinforcement ratio (ρ_l) have on the depth of the beam (d) when based on average concrete shear strength.

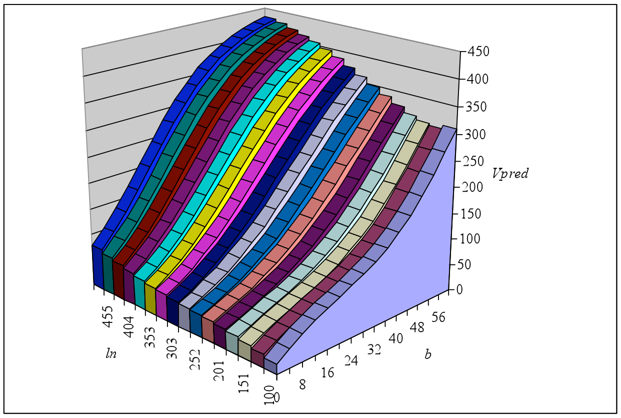

Figure 10. 3D Surface of *ln* vs. *b* vs. *Vped*

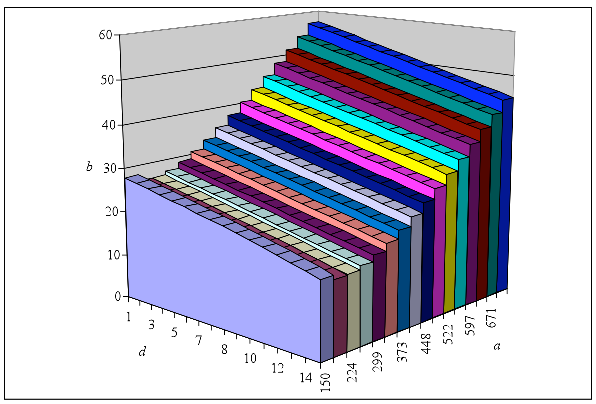

Figure 11. 3D Surface of *d* vs. *a* vs. *b*

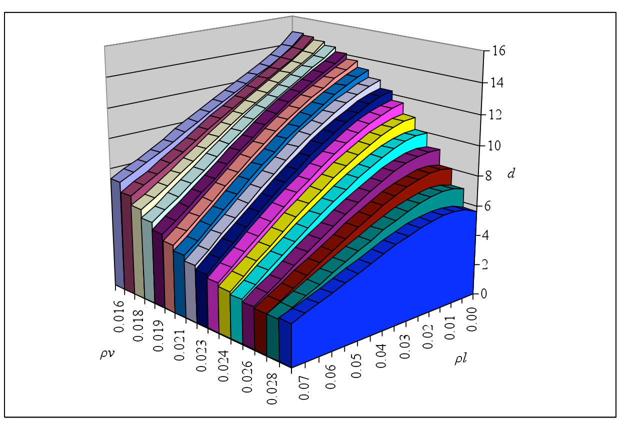

Figure 12. 3D Surface of ρ_v vs. ρ_l vs. d

6. CONCLUSIONS

An ANN for predicting the shear strength of reinforced concrete beams with and without transverse shear reinforcement was developed. Using ANNs has vastly improved estimation of beam strength. The correlation coefficient from prior estimates (the ACI Code) was equal to 0.424. The ANN produced very accurate predictions, with an R^2 value of 0.903. A diagram of the final network can be seen below in Figure 13. Originally, eleven input variables were used to estimate concrete strength. After a sensitivity analysis was performed on the input variables, two were eliminated without losing accuracy. The ANN structure that was found to produce the best R^2 value was a network with two hidden layers, Levenberg-Marquardt learning, sigmoid processing elements, and genetic optimization. Two hidden layers were found to be the minimum number of layers that produced results that did not sacrifice accuracy.

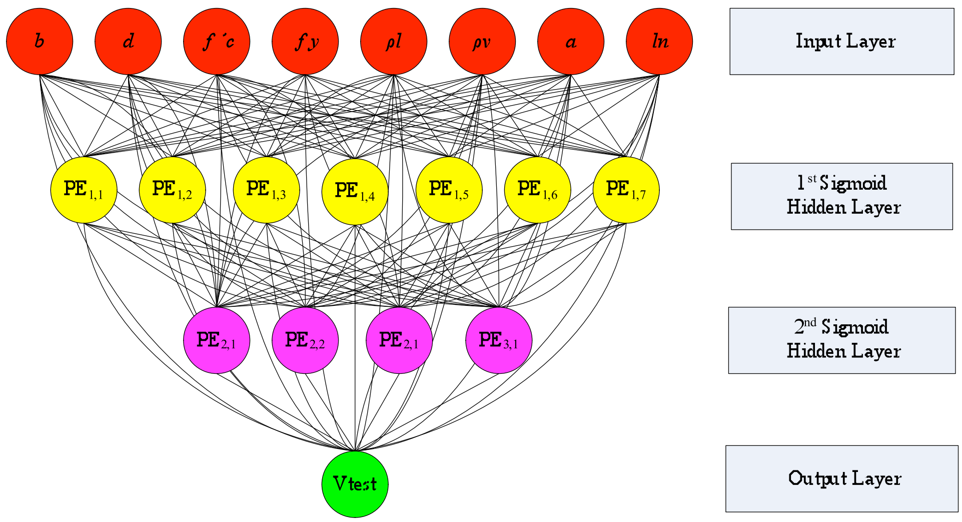

Figure 13. Final Network Diagram

The first layer contains 7 sigmoid processing elements while the second layer contains 4 processing elements. Momentum networks were trained extensively but did not provide better results than the Levenberg-Marquardt learning algorithm. This learning algorithm was found to train faster than momentum networks and arrived at a significantly lower error.

One benefit in using ANNs is that expert knowledge of the system is not needed to estimate accurately a physical relationship. An attempt to extract information to better understand the relationships was performed in this analysis. These techniques are in the form of sensitivity analysis, 2D and 3D factor relationships. Further investigation of the interactions of input and output parameters used to model the shear strength of reinforced concrete beams is needed to truly understand all of the physical relationships of the concrete structure. The techniques discussed in this research have the potential to increase knowledge of the physical relationships and prediction accuracy.

7. REFERENCES

1. ACI 318 (2005). Building Code Requirements for Structural Concrete and Commentary. American Concrete Institute. Farmington Hills, Michigan.
2. Adeli, H. (2001). Neural Networks in Civil Engineering. Computer-Aided Civil and Infrastructure Engineering. 16: 126-142.
3. Boulle, A., Chandramohan, D. and Weller P. (2001). A Case Study of Using Artificial Neural Networks for Classifying Cause of Death from Verbal Autopsy. International Journal of Epidemiology. 30: 515-520.
4. De Alcantara, N.P. and Gasparini M.E.L. (2005). Steel Bars Identification in Reinforced Concrete Structures by Using ANN and Magnetic Fields. Progress in Electromagnetics Research Symposium. Hangzhou, China, pp. 22-26.
5. Felker, V., Najjar, Y. and Hossain M. (2004). Modeling the Roughness Progression on Kansas Portland Cement Concrete Pavements, Kansas Department Of Transportation Topeka, Kansas, Kansas State University Manhattan, Kansas. Report No. K-TRAN:KSU-00-6.
6. Fonseca, E., Vellasco, P., Andrade, S., and Vellasco, M. (2003). A patch load parametric analysis using neural networks. Journal of Constructional Steel Research. 59: 251-267.

7. Girish, T., Lam, S.W. and Jayaram, J.S.R. (2003). Reliability Prediction Using Degradation Data: A Preliminary Study Using Neural Network-based Approach. Proceedings of European Safety and Reliability Conference. Maastricht, The Netherlands. pp - 681-688.
8. Gunaratnam, D.J., Degroff, T. and Gero, J.S. (2003). Improving neural network models of physical systems through dimensional analysis. Applied Soft Computing 2: 283-296.
9. Hadi, M. (2003). Neural networks applications in concrete structures. Department of Civil, Mining and Environmental Engineering, Computers and Structures. 81: 373-381.
10. Itani, O. and Y. M. Yacoub. (2000) 3-D Modeling of Spatial Properties via Artificial Neural Networks. Transportation Research Record 1709, pp. 50-59.
11. Kalliomäki, I., Vehtari A. and Lampinen, J. (2005). Shape analysis of concrete aggregates for statistical quality modeling, Machine Vision and Applications. Machine Vision and Applications, 16:3: 197-201.
12. Rumelhart, D. E. and McClelland, J. L. (1986). On Learning the Past Tenses of English Verbs. Rumelhart, McClelland, and the PDP Research Group. 2: 216-271.
13. Sanad, A. and Saka M. (2001). Prediction of Ultimate Shear Strength of Reinforced Concrete Deep Beams Using Neural Networks. Journal of Structural Engineering, 127(7): 818-828.
14. Shahin, M.A., Jaksa M.B. and Maier, H.R. (2001). Artificial Neural Network Applications in Geotechnical Engineering. Australian Geomechanics. 36(1): 49-62.
15. Umakant Dash, Chairman, (1999) Use of Artificial Neural Networks in Geo-mechanical and Pavement Systems," Subcommittee on Neural Nets and Other Computational Intelligence Based Modeling Systems. Transportation Research Circular, Number E-C012.

BIOGRAPHICAL SKETCH

William Young is currently a doctoral candidate in the Integrated Engineering program at Ohio University, where he was awarded a Stocker Fellowship. Mr. Young's dissertation research has been awarded with Ohio University's 2007 Student Enhancement Award. It is focused on developing strategies for team compatibility through a hybrid intelligent system using artificial neural networks and cost modeling. William received both his Master's (MSEE) and Bachelor's (BSEE) degrees in Electrical Engineering at OU in 2002 and 2005 respectively. His primary research focus involves knowledge extraction from neural network prediction for decision support systems.

Gary Weckman was a faculty member at Texas A&M University-Kingsville for six years before joining the Ohio University faculty in 2002 as an associate professor in Industrial and Systems Engineering. Dr. Weckman's primary research focus has been multidisciplinary applications utilizing knowledge extraction techniques with artificial neural networks. He has used ANNs to model complex systems such as large scale telecommunication network reliability, ecological relationships, stock market behavior, and industrial process scheduling. In addition, his research includes industrial safety and health applications and is on the Advisory Board for the University of Cincinnati NIOSH Occupational Safety and Health Education and Research Center Pilot Research Project.

Jim Thompson joined Ohio University's Civil Engineering Department in September 2002. He earned a Ph.D. in Civil Engineering at Lehigh University (2004) where his doctoral work focused on precast, pre-stressed concrete inverted tee girders. Prior to coming to OU, he worked as a Visiting Research Scientist at Lehigh University, testing composite ship hull sections. After earning his Master of Science in Engineering at The Johns Hopkins University (1992), Dr. Thompson spent approximately four years working as a structural engineer in Baltimore, MD, designing steel, masonry, and wood buildings. Dr. Thompson spent four years in the U. S. Navy's Civil Engineer Corps after earning his Bachelor of Mechanical Engineering degree at Villanova University (1985). While in the Navy, he served as the Public Works Officer for the Naval Facility in Adak, Alaska and the Officer-in-Charge of CBU-420 in Mayport, Florida.

Michael Brown graduated with a Ph.D. degree from The University of Texas at Austin in 2005. The research for his dissertation topic focused on the shear strength of reinforced concrete members and strut-and-tie modeling concrete structures. Dr. Brown received a Master's degree (MSE) and Bachelor's degree (BSCE) at The University of Texas at Austin in 2002 and 2000 respectively. The research performed during the completion of his Master's degree examined restrained shrinkage cracking in concrete bridge decks and the effects of concrete mixtures on such cracking. Dr. Brown is an associate member of ACI-ASCE Committee 445 Shear and Torsion.