GENERAL REGRESSION NEURAL NETWORKS FOR DRIVEN PILES IN COHESIONLESS SOILS

By M. A. Abu Kiefa[1]

ABSTRACT: Numerous investigations have been conducted in recent years to predict the load-bearing capacity of driven piles in cohesionless soils. The problem is extremely complex owing to the large number of factors that affect the behavior of piles. The available methods either oversimplify the nature of the problem or improperly consider the effect of certain governing factors. In this paper, a general regression neural network (GRNN) is used for predicting the capacity of driven piles in cohesionless soils. Predictions of the tip, shaft, and total pile capacities are made for piles with available corresponding measurements of such values. This is done using four different procedures as well as the GRNN. Comparisons of capacity component predictions (i.e., tip and shaft capacities) are only achieved for instrumented piles. The actual measurements of tip and shaft pile capacities were adjusted to account for residual stress using the wave equation analysis. For the remaining piles, the predicted total capacities are compared only with the measured value. The GRNN provides the best predictions for the pile capacity and its components. It is also shown that the GRNN is applicable for all types of conditions of driven piles in cohesionless soils.

INTRODUCTION

Although numerous investigations, both theoretical and experimental, have been conducted over the years to predict the behavior and load-bearing capacity of piles, the mechanisms are not yet entirely understood. The problem is particularly complex for piles driven into cohesionless soils owing to the sensitive nature of the factors affecting the behavior of the pile, which are difficult to quantify and involve considerable uncertainty. Among those factors are the stress and strain history of the soil, the installation effects, the effects of soil fabric and compressibility, and the difficulty of obtaining undisturbed soil samples.

To overcome the complications associated with the problem, correlations with in situ tests such as the standard penetration test, the cone penetration test, and the pressuremeter are being used in many cases. Although these tests reflect, to some extent, natural soil conditions, they have many limitations. The standard penetration test, for example, has substantially inherent variability and does not reflect soil compressibility. These tests, however, may give good predictions if correlated with load test data on a regional basis, rather than using general correlations (Darrag 1987).

Because of the limitations of the correlations obtained with the in situ tests, many empirical formulas have been developed between soil parameters and pile capacity in both end-bearing and friction piles, based on load test results to provide quick and yet reasonably accurate estimates of pile capacity. Among the most widely used empirical methods are those proposed by Meyerhof (1976), Coyle and Castello (1981), the American Petroleum Institute (*RP2A* 1984, 1991), and Randolph (1985, 1994). Review of these methods indicates that they are either oversimplifying or improperly considering the effects of certain factors. Among these factors are the effects of residual stresses, actual soil parameters existing after pile driving, and the stress history. Design guidelines provided by these methods are generally not consistent with the physical processes that dictate actual pile capacity (Randolph 1994). Therefore, there is a need for developing an alternative method that is capable of resolving the considerable uncertainties involved in predicting the pile load capacity.

Recently, artificial neural networks (ANNs) have been successfully applied to many applications in geotechnical engineering (Goh 1994, 1996; Ellis et al. 1995; Agrawal et al. 1994; Chan et al. 1995; Ghaboussi 1992). In particular, neural networks were used to predict the static pile capacity (Lee and Lee 1996; Teh et al. 1997). The neural networks were trained by dynamic stress-wave data (Teh et al. 1997) or from results of small model pile test (Lee and Lee 1996).

An ANN is usually defined as a network composed of a large number of simple processors (neurons) that are massively interconnected, operate in parallel, and learn from experience (examples). The present paper describes the application of an ANN in predicting the capacity of driven piles in cohesionless soils. The ANN is used to predict the total pile capacity as well as both of its components at the tip and along the shaft taking into consideration the important effect of the residual stresses. The general regression neural network (GRNN) model is used in the present study. An overview of the basic architecture of ANNs as well as a brief description of the algorithm used in this study is included. The basic architecture of the ANN has been widely covered in a variety of publications (Rumelhart and McClelland 1986; Lippmann 1987; Eberhart and Dobbins 1990; Hammerstrom 1993; Flood and Kartam 1994).

NEURAL NETWORK OVERVIEW

Neural networks are problem-solving programs modeled on the structure of the human brain. Neural network technology mimics the brain's own problem-solving process. Just as humans apply knowledge gained from experience to new problems or situations, a neural network takes previously solved examples to build a system of "neurons" that makes new decisions, classifications, and forecasts.

Neural networks look for patterns in training sets of data, learn these patterns, and develop an ability to correctly classify new patterns or to make forecasts and predictions. Neural networks excel at problem diagnosis, decision making, prediction, and other classifying problems where pattern recognition is important and precise computational answers are not required.

There are two basic types of neural networks: supervised and unsupervised. Supervised networks build models that classify patterns, make predictions, or make decisions according to other patterns of inputs and outputs they have "learned." They give the most reasonable answer based upon the variety of learned patterns. In a supervised network, the network is

[1]Dept. of Civ. Engrg., Univ. of Kuwait, P.O. Box 5969, Safat 13060, Kuwait; on leave from Dept. of Public Works, Cairo Univ., Giza, Egypt.

Note. Discussion open until May 1, 1999. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on June 6, 1996. This paper is part of the ***Journal of Geotechnical and Geoenvironmental Engineering***, Vol. 124, No. 12, December, 1998. ©ASCE, ISSN 1090-0241/98/0012-1177–1185/$8.00 + $.50 per page. Paper No. 13428.

J. Geotech. Geoenviron. Eng., 1998, 124(12): 1177-1185

shown how to make predictions, classifications, or decisions by giving it a large number of correct classifications or predictions from which it can learn. Backpropagation networks (BPNs), GRNNs, and probabilistic neural networks (PNNs) are examples of supervised network types. On the other hand, unsupervised networks can classify a set of training patterns into a specified number of categories without being shown in advance how to categorize. The network does this by clustering patterns. Kohonen networks are unsupervised (Haykin 1994).

Neither type of network is guaranteed to always give an absolutely "correct" answer, especially if patterns are in some way incomplete or conflicting. Results should be evaluated in terms of the percentage of correct answers that result from the model. In this regard, neural networks may not work at all with some applications. Some problems are well suited for the pattern recognition capabilities of a neural network and others are best solved with more traditional methods.

Models of ANNs are specified by three basic entities (Lin and Lee 1996): models of the neurons themselves; models of synaptic interconnections and structures; and the training or learning rules for updating the connecting weights. The basic building block of the neural network is the simulated neuron. Independent neurons are of little use, however, unless they are interconnected in a network of neurons. The network processes a number of inputs from the outside world to produce an output, the network's classifications or predictions. The neurons are connected by weights, which are applied to values passed from one neuron to the next.

A group of neurons is called a slab. Neurons are also grouped into layers by their connection to the outside world. For example, if a neuron receives data from outside the network, it is considered to be in the input layer. If a neuron contains the network's predictions or classifications, it is in the output layer. Neurons between the input and output layers are in the hidden layer(s). A layer may contain one or more slabs of neurons.

The network "learns" by adjusting the interconnection weights between layers. The answers the network is producing are repeatedly compared with the correct answers, and each time the connecting weights are adjusted slightly in the direction of the correct answers. Eventually, if the problem can be learned, a stable set of weights adaptively evolves that will produce good answers for all of the sample decisions or predictions. The real power of neural networks is evident when the trained network is able to produce good results for data which the network has never "seen" before.

Neural computation has a style. Unlike more analytically based information processing methods, neural computation effectively explores the information contained with input data, without further assumptions. Statistical methods are based on assumptions about input data ensembles (i.e., a priori probabilities, probability density functions, etc.). Neural networks, on the other hand, "discover" relationships in the input data sets through the iterative presentation of the data and the intrinsic mapping characteristics of neural topologies (normally referred to as learning). There are two basic phases in neural network operation. The first is the training or learning phase, where data are repeatedly presented to the network, while the weights of the data are updated to obtain a desired response. The second is recall, or retrieval, phase, where the trained network with frozen weights is applied to data that it has never seen. The learning phase is very time consuming owing to the iterative nature of searching for the best performance. But once the network is trained, the retrieval phase can be very fast, because processing can be distributed.

The biggest secret to building successful neural networks is knowing when to stop training. If the training is insufficient, the net will not learn the patterns. If the training is excessive, the net will learn the noise or memorize the training patterns and not generalize well with new patterns. A practical way to find a point of better generalization is to remove a random extraction of about 20% of the patterns in the training set before training and use it for cross validation. One should monitor the error in the training set and the validation set. When the error in the validation set increases, the training should be stopped because the point of best generalization has been reached. Cross validation is one of the most powerful methods to stop the training.

In general, neural networks offer viable solutions when there are large volumes of data available for training. When a problem is difficult (or impossible) to formulate analytically and experimental data can be obtained, a neural network solution is normally appropriate.

General Regression Neural Networks

GRNNs are known for their ability to train quickly on sparse data sets (Specht 1991). GRNN is a type of supervised network. It responds much better than a BPN to many types of problems. It is especially useful for continuous function approximation. GRNNs can have multidimensional input, and they will fit multidimensional surfaces through data. Unlike BPNs, which propagate training patterns through the network many times seeking a lower mean square error between the network's output and the actual output or answer, GRNN training patterns are propagated through the network only once. Because GRNNs evaluate each output independently of the other outputs, they may be more accurate than BPNs when there are multiple outputs.

A GRNN is a three-layer network with one hidden layer. The hidden layer, which sometimes refers to the regression network, consists of two slabs: pattern units and summation units. The overall block diagram of the GRNN in its adaptive form is shown in Fig. 1. The number of neurons in the input layer (first slab) is the number of inputs in the problem, and the number of neurons in the output layer (fourth slab) corresponds to the number of outputs. The number of neurons in the hidden layer is usually equal to the number of patterns (records) in the training set because the hidden layer consists of one neuron for each pattern in the training set. Choosing a larger number of neurons in the hidden layer can be advantageous in some problems if more patterns may be added later on.

When variables are loaded into a neural network, they must be scaled from their numeric range into the numeric range that the neural network deals with efficiently. There are two main

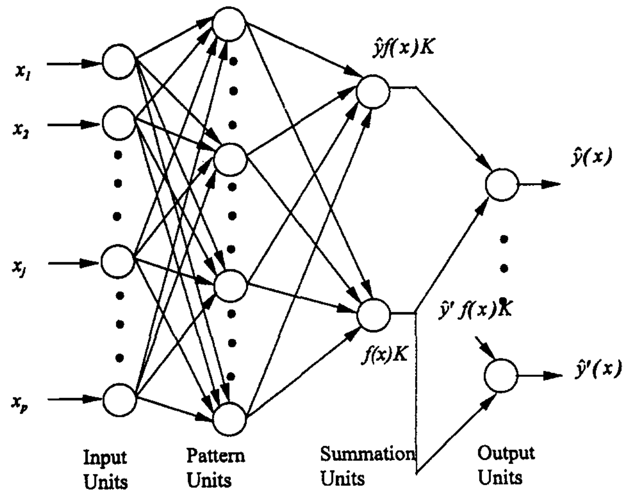

FIG. 1. GRNN Block Diagram

J. Geotech. Geoenviron. Eng., 1998, 124(12): 1177-1185

numeric ranges the networks commonly operate in, depending upon the type of activation functions in use: 0 to 1, denoted [0, 1], and −1 to 1, denoted [−1, 1]. If the numbers are scaled into the same range, but larger numbers are allowed later (i.e., they are not clipped at the bottom or top), then we will denote the ranges ⟨0, 1⟩ and ⟨−1, 1⟩. Thus [0, 1] and [−1, 1] denote that numbers below and above the ranges that it encounters later in new data will be clipped off. In other words, if data from 0 to 100 are scaled to [0, 1], then a later datum of 120 will get scaled to 1.

Fig. 1 shows a feedforward network that can be used to estimate a vector Y from a measurement vector X. The input units are merely distribution units, which provide all of the (scaled) measurements variables X to all of the neurons on the second layer (the pattern units). Each pattern unit is dedicated to one exemplar or one cluster center.

GRNNs work by measuring how far a given sample pattern is from patterns in the training set in N-dimensional space, where N is the number of inputs in the problem. When a new pattern is presented to the network, that input pattern is compared in N-dimensional space to all of the patterns in the training set to determine how far in distance it is from those patterns. The output that is predicted by the network is a proportional amount of all of the outputs in the training set. The proportion is based upon how far the new pattern is from the given patterns in the training set. For example, if a new pattern is in a cluster with other patterns in the training set, the outputs for the new pattern are going to be very close to the other patterns in the cluster around it.

One of two methods of computing this distance may be used: the Euclidean distance metric or the city block distance metric. The Euclidean distance metric is the sum of the squares of the differences in all dimensions between the pattern and the weight vector for that neuron. The Euclidean distance metric is recommended for most networks because it works the best. The city block distance metric is the sum of the absolute values of that difference. City block distance is computed faster than Euclidean distance. Once the summation is computed, it will be fired into the nonlinear activation function. The activation function normally used is the exponential, although different forms of the alternate activation function can also be used. The pattern unit outputs are passed on the summation units.

There are no training parameters such as learning rate and momentum as there are in BPN, but there is a smoothing factor that is used when the network is applied to new data. The smoothing factor is an essential parameter of both PNNs and GRNNs. The smoothing factor determines how tightly the network matches its predictions to the data in the training patterns. A higher smoothing factor causes more relaxed surface fits through the data. The success of GRNNs is dependent upon utilizing a suitable smoothing factor. Therefore, it is recommended to allow the network to choose a smoothing factor and try to find a better one through iterations or any optimization procedure. It should be noted, however, that the smoothing factor is only as good as the test set. If the problem has multiple outputs, some outputs may be more important than others may, and one smoothing factor may work better for one output while another smoothing factor may work better for another output.

In summary, the GRNN is a memory-bound, three-layer network that provides estimates of its variables and converges to an underlying linear or nonlinear regression surface. A GRNN is more advantageous with sparse and noisy data than a BPN (Specht and Shapiro 1991; Marquez and Hill 1993) and is much faster to train. Using substantial simulations, Marquez and Hill (1993) showed that the GRNN sees through noise and distortion better than the BPN. In addition, the BPN performance can be significantly hindered by the presence of local minim. Specht (1991) provided details of the GRNN paradigm, and a brief explanation of the algorithm is given here in Appendix I.

DATABASE FOR ANALYSIS

In this paper, the data used in GRNN development was drawn from 59 load test records compiled by Darrag (1987) for driven piles in sand. The records covered a wide spectrum of variation in soil relative density; stress history; geographic locations; and pile types, lengths, and diameters. Thirty-nine tests were performed on instrumented piles, so that separate data on tip and shaft capacities were available. For the remaining piles, only the total capacities were recorded. Table 1 shows the main data of the selected pile load tests and the references of such database. It should be noted that the available database is quite comprehensive in that it contains many parameters that are relevant to the determination of the pile capacity.

Ultimate Capacity

The ultimate pile capacity has to be interpreted from the pile load test results since pile load tests are not carried out until plunging failure in many cases. Many techniques are available for such interpretation (Fellenius 1980; Darrag 1987). Fellenius (1980) provided a good review of some of the best known methods for load test interpretation. He interpreted some load test results by as many as 10 different methods. These methods result in a considerable scatter, which makes it very difficult to select a single value to be interpreted as the pile capacity.

The main objective of using the GRNN is to show its ability to learn the underlying relationship from incomplete and noise-oriented data records and estimate reliable, though not very accurate, pile capacities. Therefore, it is necessary to use one method throughout the development of the network(s) to provide consistent values. Darrag (1987) examined a wide variety of load-deformation curves for piles in sand. It was concluded that the procedure proposed by Chin (1970) provides a simple, reasonable, and consistent method. The method assumes that the load-movement curve is of hyperbolic shape, especially as the load approaches the failure value. Duncan and Chang (1970) recognized that the actual ultimate resistance is always lower than the asymptotic value of the hyperbola. In their constitutive model, Duncan and Chang (1970) recommended the application of a reduction factor, R_f, to the asymptotic value of the hyperbola to obtain the ultimate soil resistance. Darrag (1987) examined a wide variety of load test data for which plunging failure was clearly evident, and excellent agreement was obtained. An average value of R_f was found to be 0.85, with a coefficient of variation of about 10%. Therefore, it was decided to use this method for interpreting all load test results in the current study.

Residual Stresses

In order to build neural networks capable of predicting the pile shaft resistance and the pile tip resistance, it is necessary to observe and measure the development of such forces directly from pile load tests. However, a serious problem is posed by the phenomenon of residual stresses, which exist in any pile driven into the ground.

As a pile is driven, each blow causes the pile to penetrate a little farther into the ground and then rebound. This rebound is due partly to elastic unloading of the pile and partly to unloading of the soil beneath the pile toe. The rebound is sufficient to reverse the direction of the local shear stresses on

J. Geotech. Geoenviron. Eng., 1998, 124(12): 1177-1185

TABLE 1. Summary of Database for Pile Load Test and GRNN Predictions

	Input					Measured	Predicted	Reference		
Pile number (1)	ϕ_{Shaft} (2)	ϕ_{Tip} (3)	σ'_v (tip) (kN/m^2) (4)	L (m) (5)	A (m^2) (6)	Q (total) (kN) (7)	Q (total) (kN) (8)	Source paper (9)	Test number (10)	Site (11)
1	33	38	255	24.5	0.1310	2,615	2,494	Mansur and Kaufman (1958)	1	Low-Sill Structure, Old River
2	34	37.5	206	19.8	0.2230	3,674	3,638	Mansur and Kaufman (1958)	2	Low-Sill Structure, Old River
3	33	38	223	21.5	0.1310	2,164	2,389	Mansur and Kaufman (1958)	3-a	Low-Sill Structure, Old River
4	33	37.5	210	20.2	0.1468	3,042	2,881	Mansur and Kaufman (1958)	4	Low-Sill Structure, Old River
5	33	37	206	19.9	0.1821	2,856	2,831	Mansur and Kaufman (1958)	6	Low-Sill Structure, Old River
6	38	41	138	11.6	0.2090	3,558	3,558	Furlow (1968)	1	Lonesville, La.
7	38	40	164	13.7	0.2090	3,292	3,479	Furlow (1968)	2	Lonesville, La.
8	38	40	196	16.5	0.2090	3,923	3,735	Furlow (1968)	3	Lonesville, La.
9	35	37	158	16.2	0.1050	1,637	1,669	Hunter and Davisson (1969)	1	Arkansas River Project LD4
10	35	37	158	16.1	0.1644	2,233	2,276	Hunter and Davisson (1969)	2	Arkansas River Project LD4
11	35	36.5	158	16.2	0.2109	2,295	2,337	Hunter and Davisson (1969)	3	Arkansas River Project LD4
12	36.5	36.5	120	12.3	0.1654	1,779	2,182	Hunter and Davisson (1969)	4	Arkansas River Project LD4
13	34	38	475	47.2	0.2917	5,604	5,604	McCammon and Golder (1970)	1	Low Arrow Lake, B.C., Canada
14	34	34	38	3.0	0.1644	712	712	Vesic (1970)	H-11	Ogeeches River, Ga.
15	35	35	72	6.1	0.1644	1,735	970	Vesic (1970)	H-12	Ogeeches River, Ga.
16	35	35	100	8.9	0.1644	2,491	2,847	Vesic (1970)	H-13	Ogeeches River, Ga.
17	36	36	131	12.0	0.1644	3,158	3,096	Vesic (1970)	H-14	Ogeeches River, Ga.
18	36	36	161	15.0	0.1644	3,825	3,380	Vesic (1970)	H-15	Ogeeches River, Ga.
19	35.5	36	163	15.2	0.1301	2,695	2,831	Vesic (1970)	H-2	Ogeeches River, Ga.
20	34	38	146	11.3	0.0316	1,429	1,429	Koizumi (1971)	6C	Tokyo, Japan
21	35.5	35.5	89	9.1	0.0864	658	903	Tavenas (1971)	H2	St. Charles River, Que., Canada
22	35.5	35.5	119	12.2	0.0864	882	985	Tavenas (1971)	H3	St. Charles River, Que., Canada
23	35.5	35.5	148	15.2	0.0864	1,014	1,192	Tavenas (1971)	H4	St. Charles River, Que., Canada
24	35.5	35.5	178	18.3	0.0864	1,281	1,333	Tavenas (1971)	H5	St. Charles River, Que., Canada
25	35.5	35.5	89	9.1	0.0799	655	903	Tavenas (1971)	J2	St. Charles River, Que., Canada
26	35.5	35.5	119	12.2	0.0799	894	988	Tavenas (1971)	J3	St. Charles River, Que., Canada
27	35.5	35.5	148	15.2	0.0799	1,113	1,200	Tavenas (1971)	J4	St. Charles River, Que., Canada
28	35.5	35.5	178	18.3	0.0799	1,281	1,336	Tavenas (1971)	J5	St. Charles River, Que., Canada
29	31	31	134	16.0	0.0613	480	480	Gregersen et al. (1973)	D/A	Holemen Island, Drammen, Norway
30	31	31	134	16.0	0.0316	519	480	Gregersen et al. (1973)	B/C	Holemen Island, Drammen, Norway
31	33	33	111	12.2	0.0061	75	528	Bergdahl and Wennerstrand (1976)	Series II	Albysjon, Sweden
32	39	39.5	75	7.0	0.0999	2,439	2,371	Heins and Barends (1984)	A	North Sea, The Netherlands
33	39	39.5	72	6.7	0.0999	3,000	2,351	Heins and Barends (1984)	B	North Sea, The Netherlands
34	39	39.5	56	5.2	0.0999	1,950	2,251	Heins and Barends (1984)	C	North Sea, The Netherlands
35	32	34	198	21.0	0.2313	3,200	3,209	Carpentier et al. (1984)	P1	Zeebrugge, Belgium
36	37.5	40	301	29.9	0.3075	4,733	4,733	Gurtowski and Wu (1984)	A	West Seattle Freeway
37	37.5	39.5	258	25.6	0.3075	4,021	4,733	Gurtowski and Wu (1984)	B	West Seattle Freeway
38	33	35	169	18.0	0.6568	5,000	5,604	Mey et al. (1985)	P1(2)	Gadiz, Spain
39	28	39	213	16.8	0.1431	4,670	4,660	Reike and Crowser (1987)	1B-C	Tacoma, Wash.
40	34	35	111	12.2	0.1143	854	1,029	Thorburn and MacVicar (1970)	1	Clyde Valley–Glasgow, Scotland
41	32	35	241	23.3	0.1486	1,628	1,790	Thorburn and MacVicar (1970)	2	Clyde Valley–Glasgow, Scotland
42	34	35	92	9.1	0.1291	685	990	Thorburn and MacVicar (1970)	3	Clyde Valley–Glasgow, Scotland
43	32	37	176	17.5	0.3855	3,069	3,069	Milovic and Baci (1976)	C	Biograd Yugoslavia
44	35	37	246	23.8	0.0729	1,913	1,565	Blendy (1979)	SB-1	Missouri
45	35	37	183	17.7	0.0729	2,313	1,656	Blendy (1979)	SB-3	Missouri
46	35	37	260	25.3	0.0729	1,254	1,618	Blendy (1979)	53	Connecticut
47	39	39.5	56	5.2	0.0999	1,948	2,251	Heins and Barends (1979)	C	North Sea, The Netherlands
48	33	35.5	343	34.1	0.0325	1,761	1,830	Kessler (1979)	178	Cohasset, Minn.
49	34	36	319	31.7	0.0325	2,180	2,111	Kessler (1979)	4,571	Cohasset, Minn.
50	33	33	335	29.3	0.0557	3,203	3,211	Wan et al. (1979)	B.O.F.-8	Kaohsiung, China
51	33	34	354	31.1	0.0557	3,211	3,211	Wan et al. (1979)	C.O.-2	Kaohsiung, China
52	36	37	215	21.3	0.0409	1,779	1,583	Goble et al. (1982)	3-4	I-80 Winnemucca, Nev.
53	36	37	209	20.7	0.0827	1,868	1,611	Goble et al. (1982)	4-4	I-80 Winnemucca, Nev.
54	35	35	242	24.1	0.0660	1,779	1,643	Goble et al. (1982)	5-4	I-80 Winnemucca, Nev.
55	36	37	209	20.7	0.0929	1,913	1,670	Goble et al. (1982)	6-4	I-80 Winnemucca, Nev.
56	35	35	166	16.5	0.0613	2,100	1,803	Goble et al. (1982)	3-5	I-80 Winnemucca, Nev.
57	35	35	178	17.7	0.0827	1,509	1,653	Goble et al. (1982)	4-5	I-80 Winnemucca, Nev.
58	35.5	37	228	21.9	0.0660	922	1,513	Goble et al. (1982)	5-5	I-80 Winnemucca, Nev.
59	35	35	178	17.7	0.0929	1,779	1,621	Goble et al. (1982)	6-5	I-80 Winnemucca, Nev.

the shaft, at least over the upper part of the pile, so that they act downwards. Relatively large displacements are required to unload the toe, whereas relatively small displacements are required to unload the shaft. Consequently, significant residual stresses remain beneath the toe, countered by the majority of the shaft resistance acting downwards.

When piles are instrumented, it is common to zero all instrumentation after driving, so that any instrumentation drift does not influence the test results during driving. The existence of residual stresses will often be apparent only if a tension test to failure is carried out, when the base resistance will register a negative load. In reality, there was a compressive load at the base before the test, and this has reduced to 0. It is important to account for residual stresses in interpreting data for compression tests; otherwise the base resistance will be underestimated and the shaft resistance will be overestimated (Lings 1997).

Hery (1983) modified the wave equation computer program

J. Geotech. Geoenviron. Eng., 1998, 124(12): 1177-1185

(WEAP), which was developed by Goble and Rausche (1976), so that it allowed for a multiblow analysis. The program was referred to as CUWEAP. Residual stresses were reasonably predicted by this program, which provides a better representation of the hammer assembly and soil damping parameters. Therefore, CUWEAP was used to obtain the magnitude and distribution of residual forces below the pile tip and along its shaft. The actual measurements of tip and shaft capacities were adjusted to account for residual stresses to perform the analysis of the current study (Darrag 1987).

NEURAL NETWORKS FOR ESTIMATING TOTAL, TIP, AND SHAFT PILE CAPACITY

The neural networks were developed using the neural network development program Neuroshell 2 (1993). This program implements several different neural network algorithms, including BPN, PNN, and GRNN. To use the program, a set of inputs and outputs must be defined, and a suitable training set has to be developed.

The present study introduces three general regression neural network models. The first model, GRNNM1, was developed to estimate the total pile capacity. The second model, GRNNM2, was introduced to estimate the tip pile capacity. The last model, GRNNM3, was utilized to estimate the shaft pile capacity.

GRNNM1

The database of the 59 records was used to develop GRNNM1 to estimate the total pile capacity. The data records were split in such a way that different references were randomly divided into two parts. The first group of 38 records was used as a training set, and the remaining 21 records (about 35% of the data) were used for testing the robustness of the developed neural network.

The first step in developing a trained network is to select a number of different combinations of input variables to evaluate the most reliable neural network model. Five variables [$\tan(\phi_{Shaft})$, $\tan(\phi_{Tip})$, σ'_v(tip), L, and A] were selected as input parameters in GRNNM1 as shown in Table 1. The angle of shear resistance of the soil around the shaft (ϕ_{shaft}) and at the tip of the pile (ϕ_{Tip}) provides information about the shear strength of surrounding and supporting soils. Since the effective overburden pressure at the tip of the pile is always thought to have a significant role in calculating the total pile capacity, the value of σ'_v(tip) will be sufficient to provide all needed stress information. The length of the pile (L) as well as the equivalent cross-sectional pile area (A) will provide the necessary information about the pile dimensions. Components of all the input vectors of the training and testing sets were scaled between 0 and 1 before being presented to the network. A single neuron representing the measured total pile capacity is used for the output.

The designed GRNNM1 (Fig. 2) has three layers. Five neurons in the input layer, 59 neurons in the hidden layer, and one neuron in the output layer. The number of neurons in the hidden layer is equal to the number of the load test records in the database. The city block norm is used for convergence with an initial smoothing factor (σ) of 0.3.

The GRNN used in this study was genetic adaptive; i.e., it uses a genetic algorithm to find an input smoothing factor adjustment. This is used to adapt the overall smoothing factor to provide a new value for each input. A genetic algorithm works by selective breeding of a population of "individuals," each of which is a potential solution to the problem. The larger the breeding pool size, the greater the potential of it producing a better individual. Genetic algorithms use a "fitness" measure to determine which of the individuals in the population survive and reproduce (Lin and Lee 1996). The fitness for GRNN is the mean squared error of the outputs for the entire data set. The genetic adaptive algorithm seeks to minimize the fitness. A genetic breeding pool size of 200 was used in GRNNM1.

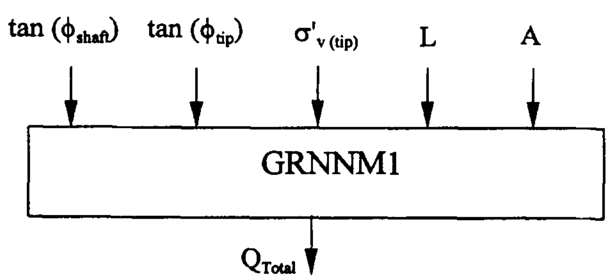

FIG. 2. Neural Network for Total Pile Capacity

GRNNM2 and GRNNM3

In this part, the application of GRNN to evaluate the tip and shaft pile capacities is presented. Thirty-nine load tests from the previous data records were performed on instrumented piles, so that separate data on tip and shaft capacities were available. These actual measurements of tip and shaft pile capacities were adjusted to account for residual stresses, as mentioned in the database section. The adjusted records were then used to build two different neural networks—one for the tip pile capacity (GRNNM2) and the other for the shaft pile capacity (GRNNM3).

Five variables [$\tan(\phi_{Shaft})$, $\tan(\phi_{Tip})$, σ'_v(tip), L, and A] as shown in Fig. 3 were selected as input parameters in GRNNM2 to evaluate the tip pile capacity. The variable $\tan(\phi_{Shaft})$ is included in GRNNM2, since the tip pile capacity is somehow affected by the shear resistance of soil around the lower part of the shaft which is very close to the tip of the pile. The 39 records were split at random into two groups: the first group of 30 records was used as a training set, and the remaining nine records (about 23% of the data) were used for testing the developed neural network.

To evaluate the shaft pile capacity, four variables [$SPT_{(shaft)}$, $\tan(\phi_{shaft})$, L, and D] were selected as input parameters in the GRNNM3, as shown in Fig 4. The 39 records were split at random into two groups: the first group of 31 records was used as a training set, the second group of eight records (about 21% of the data) was used for testing the developed neural network.

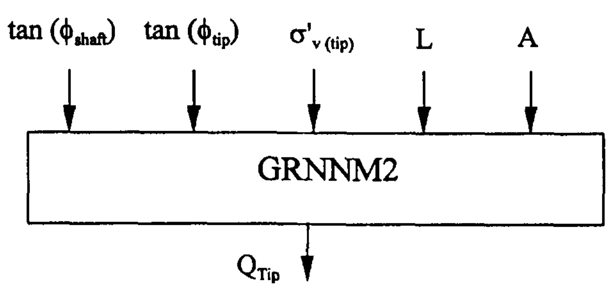

FIG. 3. Neural Network for Tip Pile Capacity

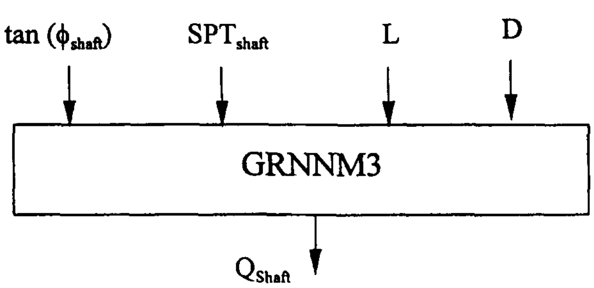

FIG. 4. Neural Network for Shaft Pile Capacity

J. Geotech. Geoenviron. Eng., 1998, 124(12): 1177-1185

GRNNM2 and GRNNM3 both had three layers, five or four neurons in the input layer depending on the number of input variables, 40 neurons in the hidden layer, and one neuron in the output layer with initial smoothing factor (σ) of 0.3. The city block norm was used for convergence. A genetic breeding pool size of 200 was utilized.

RESULTS AND DISCUSSION

During the training phase, the measured total pile capacities are compared with those obtained by GRNNM1. At the end of this phase, it is expected that the neural network will correctly reproduce the target output values provided that the errors are minimal. The next step is to examine the robustness of the trained neural network in generating correct responses for a new set of data. Once the training and testing steps are successfully accomplished, the neural network obtained can be used as a practical design tool for this type of application.

Table 1 presents a summary of the measured total pile capacities and the results predicted by GRNNM1 for all 59 records included in both the training and the testing data sets. A comparison between the measured total pile capacities and the predicted values is shown in Fig. 5. The results show that the regression surfaces generated by GRNNM1 from the 38 training records were sufficient to predict, with high accuracy, the total pile capacities for the 21 testing records. High coefficient of correlations (r^2) are obtained, as shown in Fig. 5 for both training and testing sets of data.

Pile capacity predictions were made using GRNNM1 as well as four other empirical techniques, and they were also compared with actual measurements. These other methods are those proposed by Meyerhof (1976), Coyle and Castello (1981), the American Petroleum Institute (*RP2A* 1984), and

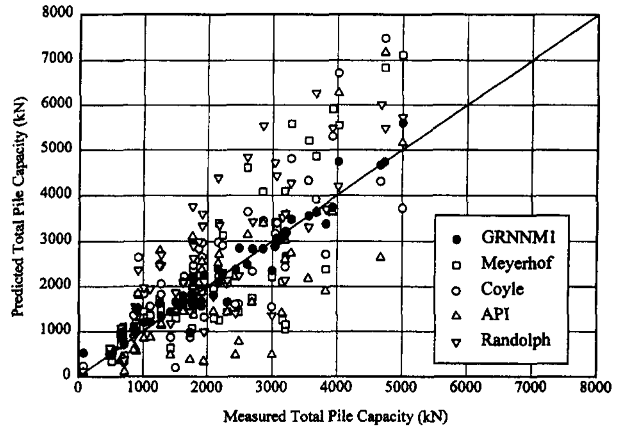

FIG. 6. Comparison of Predicted and Measured Total Pile Capacity

Randolph (1985). Fig. 6 shows the relationship between measured and predicted values for each of the five methods. GRNNM1 predictions in Fig. 6 show fewer scatters in the data points than the predictions of all the other methods. By choosing one method—Chin's procedure (1970), as mentioned before, for interpreting all test results—the dispersion caused by the different definitions of pile capacity should be minimized (if not eliminated). If the existing empirical methods are consistent, the obtained results should either overpredict or underpredict the pile capacity. However, by examining Fig. 6 considerable scatter can be observed between the measured pile capacities (according to the selected definition of pile capacity) and the predicted ones from these methods. This scatter, as seen on both sides of the line, indicates the inconsistency of the methods, which may yield overestimated

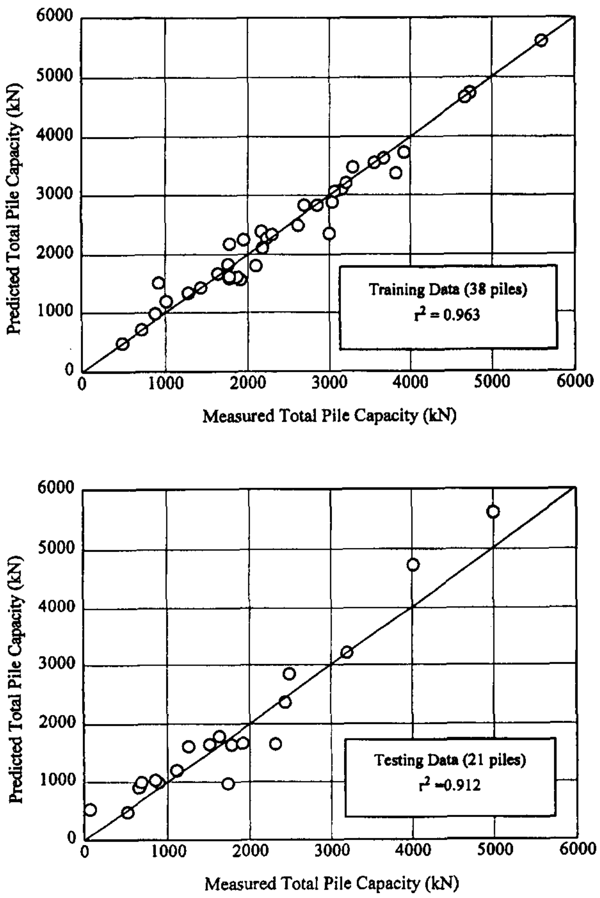

FIG. 5. Comparison of Predicted and Measured Total Pile Capacity

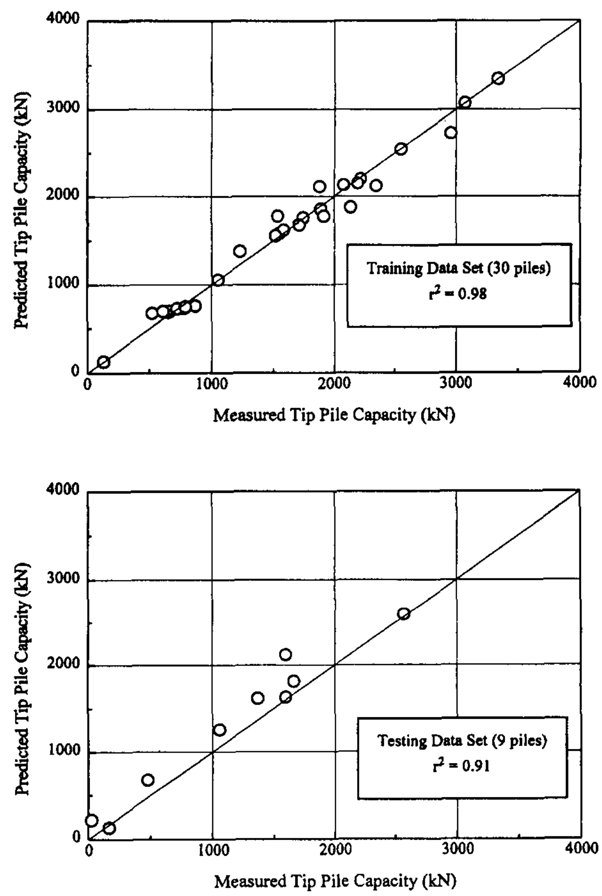

FIG. 7. Comparison of Predicted and Measured Tip Pile Capacity

J. Geotech. Geoenviron. Eng., 1998, 124(12): 1177-1185

predictions in some cases and underestimated predictions in others, even though the measured pile capacities are similar. The determination coefficient (r^2) for all data records equals 0.95 for the GRNNM1, while it ranges between 0.52 and 0.63 for other methods. This figure indicates that the new method provides the best prediction for total pile capacity. It is clear that GRNNM1 is also applicable for all conditions of driven piles in granular soils, in contrast to the other procedures, which are limited to certain conditions and usually produce inconsistent results.

A comparison between the measured tip pile capacities and the predicted values from GRNNM2 for both the training and the testing sets is shown in Fig. 7. Very high coefficients of correlation were obtained, as shown in figure for both training

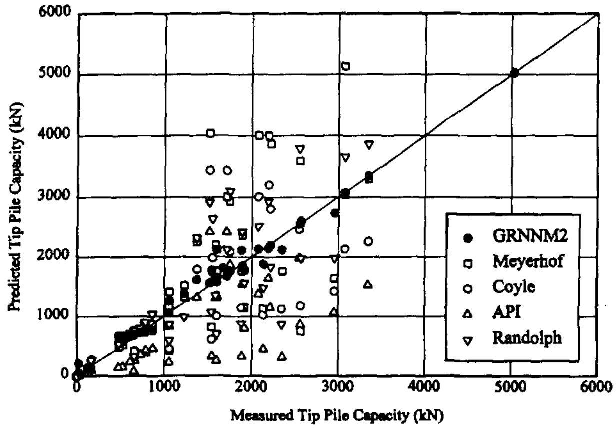

FIG. 8. Comparison of Predicted and Measured Tip Pile Capacity

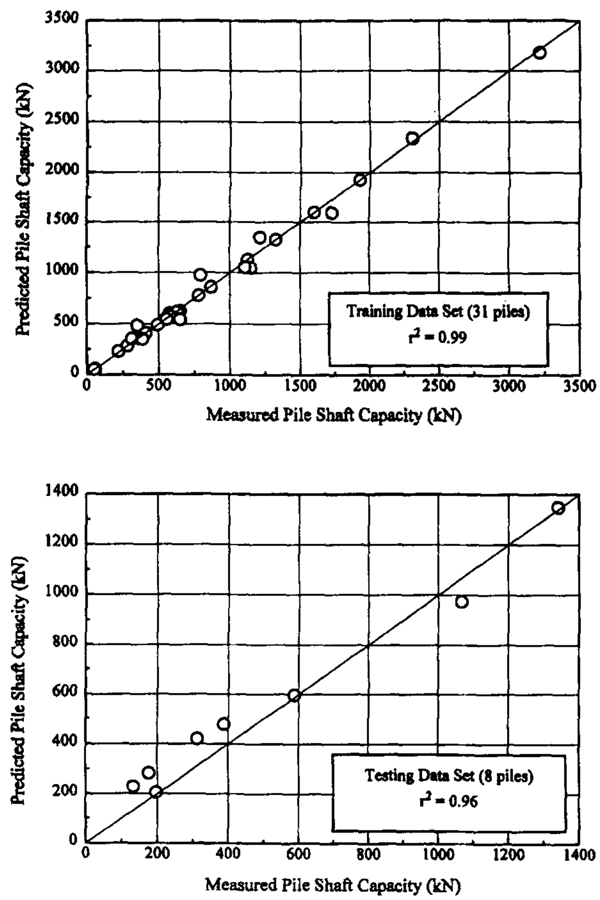

FIG. 9. Comparison of Predicted and Measured Shaft Pile Capacity

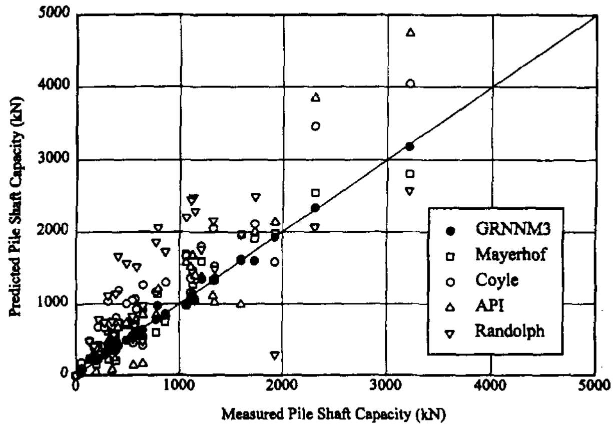

FIG. 10. Comparison of Predicted and Measured Shaft Pile Capacity

and testing sets of data. The tip pile capacity predictions using both the previous empirical procedures and GRNNM2 were compared with actual measurements. Fig. 8 shows the relationship between measured and predicted values for each of the five methods. Considerable scatter and inconsistency can be observed between the measured tip pile capacities and the predicted ones from the empirical methods. Clearly, the GRNNM2 prediction shown in Fig. 8 provides the best prediction of tip pile capacity.

A comparison between the measured shaft pile capacities and the predicted ones from GRNNM3 for both the training and the testing sets is shown in Fig. 9. For both the training and the testing sets of data, very high coefficients of correlation were obtained, as illustrated in the figure. The shaft pile capacity predictions using the previous empirical procedures as well as GRNNM3 were also compared with actual measurements. The relationship between measured and predicted values for each of the five methods is shown in Fig. 10. From this figure, it is most evident that GRNNM3 also provides the best prediction for shaft pile capacity.

SUMMARY AND CONCLUSIONS

The principal objective of this paper was to demonstrate the feasibility of using ANN to predict the capacity of driven piles in cohesionless soils. Many factors affecting the capacity predictions are ignored, oversimplified or improperly introduced by the existing methods.

In this paper, three general regression neural network models (GRNNM1, GRNNM2, and GRNNM3) were introduced and verified as new and stand-alone procedures. Data from 59 good-quality pile load tests in granular soils were utilized to construct the networks. Pile capacity predictions were made using the GRNNM1 as well as four other empirical methods, and were compared with actual measurements. It may be concluded that the GRNNM1 is applicable for all different conditions of driven piles in cohesionless soils. The proposed method is superior to the empirical ones when compared with actual measurements.

The other models—GRNNM2 and GRNNM3—were developed to predict the tip and the shaft pile capacities in cohesionless soils, respectively. The actual measurements of tip and shaft pile capacities were adjusted to account for residual stresses using the wave equation analysis and then used for training and testing the neural networks. Predictions made using the GRNN models and other empirical methods were compared with the adjusted measurements, and the GRNN models proved to provide the best predictions for both shaft and tip pile capacity.

To improve the proposed prediction method, it is recommended that additional load tests on instrumented piles be

J. Geotech. Geoenviron. Eng., 1998, 124(12): 1177-1185

performed and that direct measurements of residual stresses along the pile shaft and below the tip be made before performing the load tests. Reliable instrumentation should be developed for making such measurements. These direct measurements should be used for training the GRNN models rather than using adjusted measurement by the wave equation analysis.

APPENDIX I. GENERAL REGRESSION NEURAL NETWORK

Let x be a vector of random variables and y a scalar random variable. Let X be a particular measured value of x. In other words, X is the finite vector of noisy measurements of x, and y represents the associated scalar value. The conditional mean of y given X is defined by

$$E[y|X] = \frac{\int_{-\infty}^{\infty} yf(X, y)\, dy}{\int_{-\infty}^{\infty} f(X, y)\, dy} \tag{1}$$

where $f(X, y)$ denotes the joint continuous probability density function (pdf) of a vector random variable X and a scalar random variable y. If $f(X, y)$ is not known, an estimate $\hat{f}(X, y)$ must be used. Parzen (1962) proposed a class of consistent estimators for $f(X, y)$. Using window estimation, $\hat{f}(X, y)$ can be written as

$$\hat{f}(X, y) = \frac{1}{(2\pi)^{p+1}\,\sigma^{p+1}} \cdot \frac{1}{n} \sum_{i=1}^{n} e^{(-(x-x^i)^T(x-x^i))/2\sigma^2} \cdot e^{(-(y-y^i))/2\sigma^2} \tag{2}$$

where $x \in R^p$ and n is the number of sample observations. Using (2), (1) becomes

$$\hat{y} = \frac{\sum_{i=1}^{n} e^{(-(x-x^i)^T(x-x^i))/2\sigma^2} \int_{-\infty}^{\infty} ye^{-(y-y^i)^2/2\sigma^2}\, dy}{\sum_{i=1}^{n} e^{(-(x-x^i)^T(x-x^i))/2\sigma^2} \int_{-\infty}^{\infty} e^{-(y-y^i)^2/2\sigma^2}\, dy} \tag{3}$$

Let $D_i^2 = (X - X^i)^T (X - X^i)$; then the estimate of the expected value $\hat{y}$ becomes

$$\hat{y} = \frac{\sum_{i=1}^{n} y^i e^{-D_i^2/2\sigma^2}}{\sum_{i=1}^{n} e^{-D_i^2/2\sigma^2}} \tag{4}$$

These density estimators are consistent; i.e., they asymptotically converge to the underlying joint pdf $f(x, y)$ at all points (x, y) provided that $\sigma = \sigma(n)$ is chosen such that

$$\lim_{n\to\infty} \sigma(n) = 0 \text{ and } \lim_{n\to\infty} n\sigma^p(n) = \infty$$

Here σ can be visualized as a smoothing parameter. That is, if σ is made large, the estimated density is forced to be smooth and the limit becomes a multivariable Gaussian with covariance $\sigma^2 I$. On the other hand, a smaller value of σ allows the estimated density to assume non-Gaussian shapes. In such a case, wild points may have too great an effect on the estimate.

Thus, given the joint pdf of X and y, the conditional pdf can be determined and, hence, the expected value can be computed. Cacoullos (1966) extended Parzen's method to the multivariable case. The resulting regression procedure is implemented via three layers of parallel neural network architecture (Specht 1990, 1991), where it called a general regression neural network.

ACKNOWLEDGMENTS

This research was supported by a grant from the Kuwait University Research Administration (No. EPV 090).

APPENDIX II. REFERENCES

Agrawal, G., Weeraratne, S., and Khilnami, K. (1994). "Estimating clay liner and cover permeability using computational neural networks." *Proc., First Congr. on Computing in Civ. Engrg.*

Bergdahl, U., and Wennerstrand, F. (1976). "Bearing capacity of driven friction piles in loose sand." *Proc., Sixth Eur. Conf. SMFE*, 355–360.

Blendy, M. M. (1979). "Rational approach to pile foundations." *Proc., Symp. on Deep Foundations*, ASCE, Reston, Va., 38–47.

Cacoullos, T. (1966). "Estimation of a multivariable density." *Ann. Inst. Statist. Math.*, Tokyo, Japan, 18(2), 179–189.

Carpentier, R., Van Damme, L., and Maertens, L. (1984). "Results of loading tests on two vibrated steel H-Piles in sand." *Proc., Sixth Conf. SMFE*, 593–600.

Chan, W. T., Chow, Y. K., and Liu, L. F. (1995). "Neural network: An alternative to pile driving formulas." *Comp. and Geotechnics*, 17, 135–156.

Chin, F. K. (1970). "Estimation of the ultimate load of piles not carried to failure." *Proc., Second Southeast Asian Conf. on Soil Engrg.*, 81–90.

Coyle, H. M., and Castello, R. R. (1981). "New design correlations for piles in sand." *J. Geotech. Engrg.*, ASCE, 107(7), 965–986.

Darrag, A. A. (1987). "Capacity of driven piles in cohesionless soils including residual stresses," PhD thesis, Purdue Univ., West Lafayette, Ind.

Duncan, J. M., and Chang, C. (1970). "Non-linear analysis of stress and strain in soils." *J. Soil Mech. and Found. Div.*, ASCE, 96(5), 1629–1653.

Eberhart, R. C., and Dobbins, R. W. (1990). *Neural network PC tools: A practical guide*. Academic Press, San Diego, Calif.

Ellis, G. W., Yao, C., Zhao, R., and Penumadu, D. (1995). "Stress-strain modeling of sands using artificial neural networks." *J. Geotech. Engrg.*, ASCE, 121(5), 429–435.

Fellenius, B. H. (1980). "The analysis of results from routine pile load tests." *Ground Engrg.*, London, 13(6), 19–31.

Flood, I., and Kartam, N. (1994). "Neural networks in civil engineering. I: Principles and understanding." *J. Computing Civ. Engrg.*, ASCE, 8(2), 131–148.

Furlow, C. R. (1968). "Pile tests, Jonesville lock and dam, Quachita and Black Rivers, Arkansas and Louisiana." *Tech. Rep. S-68-10*, U.S. Army Engineer Waterways Experimental Station, Vicksburg, Miss.

Ghaboussi, J. (1992). "Potential applications of neuro-biological computational models in geotechnical engineering." *Numerical models in geotechanics*, G. N. Pande and S. Pietruszezak, eds., Balkema, Rotterdam, The Netherlands, 543–555.

Gregersen, O. S., Aas, G., and DiBiaggio, E. (1973). "Load tests on friction piles in loose sand." *Proc., VIII Int. Conf. Soil Mech. Found. Engrg.*, 19–27.

Goble, G. C., Cochran, D., and Marcucci, F. (1982). "Foundation design and evaluation for Winnemuca viaduct." *TRR No. 884*, 37–45.

Goble, G. G., and Rausche, F. (1976). *Wave equation analysis of pile driven, WEAP program*, Vols. 1–3, Office of Res. and Devel., Fed. Hwy. Admin., U.S. Department of Transportation, Washington, D.C.

Goh, A. T. C. (1994). "Seismic liquefaction potential assessed by neural networks." *J. Geotech. Engrg.*, ASCE, 120(9), 1467–1480.

Goh, A. T. C. (1996). "Neural networks modeling of CPT seismic liquefaction data." *J. Geotech. Engrg.*, ASCE, 122(1), 70–73.

Gurtowski, T. M., and Wu, M. J. (1984). "Compression load tests on concrete piles in alluvium." *Proc., Symp. Anal. and Des. of Pile Foundations*, ASCE, Reston, Va., 138–153.

Hammerstrom, D. (1993). "Working with neural networks." *IEEE Spectrum*, July, 46–53.

Haykin, S. (1994). *Neural networks: A comprehensive foundation*. Macmillan College Publishing Co., New York.

Heins, W. F., and Barends, F. B. J. (1979). "Pile test program in overconsolidated sand." *Proc., Conf. on Des. Parameters in Geotech. Engrg.*, Vol. 3, British Geotechnical Society, London, 75–78.

Hery, P. (1983). "Residual stress analysis in WEAP," MSc thesis, Dept. of Civ. Engrg. and Arch. Engrg., Univ. of Colorado.

Hunter, A. H., and Davisson, M. T. (1969). "Measurements of pile load transfer." *Performance of Deep Foundation*, ASTM, STP 444, 106–117.

Kessler, K. A. (1979). "Pile foundation in flood plain soils." *Proc., Symp. on Deep Foundations*, ASCE, Reston, Va., 215–234.

J. Geotech. Geoenviron. Eng., 1998, 124(12): 1177-1185

Koizumi, Y. (1971). "Field tests on piles in sand." *Soils and Found.*, Tokyo, Japan, 11, 29–49.
Lee, I. M., and Lee, J. H. (1996). "Prediction of pile bearing capacity using artificial neural networks." *Comp. and Geotechnics*, 18(3), 189–200.
Lin, C. T., and Lee, C. S. G. (1996). Neural fuzzy systems: A neuro-fuzzy synergism to intelligent systems. Prentice Hall Inc., Upper Saddle River, N.J.
Lings, M. L. (1997). "Predicting the shaft resistance of driven pre-formed piles in sand." *Proc., Instn. Civ. Engrs.*, London, 125, 71–84.
Lipmann, R. P. (1987). "An introduction to computing with neural nets." *Acoust. Speech & Signal Processing*, 4(2), 4–22.
Mansur, C. I., and Kaufman, R. I. (1958). "Pile tests, low sill structure, Old River, La." *J. Soil Mech. and Found. Div.*, ASCE, 82(4), 1–33.
Marquez, L., and Hill, T. (1993). "Function approximation using back-propagation and general regression neural networks." *Proc., 26th Hawaii Int. Conf. on Sys. Sci.*, 607–615.
McCammon, N. R., and Golder, H. Q. (1970). "Some loading tests on long pipe piles." *Géotechnique*, London, 20(2), 171–184.
Mey, R., Oteo, C. S., Rio, J. S., and Soriano, A. (1985). "Field testing on large driven piles." *Proc., 11th Int. Conf. Soil Mech. Found. Engrg.*, Vol. 3, 1559–1564.
Meyerhof, G. G. (1976). "Bearing capacity and settlement of pile foundations." *J. Geotech. Engrg.*, ASCE, 102(3), 196–228.
Milovic, D., and Baci, S. (1976). "The ultimate bearing capacity of piles determined by load tests." *Proc., Sixth Eur. Conf. Soil Mech. Found. Engrg.*, 519–522.
Neuroshell 2 manual. (1993). Word Systems Group, Inc., Frederick, Mass.
Parzen, E. (1962). "On estimation of a probability density function and mode." *Annals of Math. Statistics*, 33, 1065–1076.
Randolph, M. F. (1985). "Capacity of piles driven into dense sand." *Rep. Soils TR 171*, Engrg. Dept., Cambridge University, Cambridge, U.K.
Randolph, M. F., Dolwin, J. and Beck, R. (1994). "Design of driven piles in sand." *Géotechnique*, London, 44(3), 427–448.
Reike, R. D., and Crowser, J. C. (1987). "Interoperation of pile load test considering residual stresses." *J. Geotech. Engrg.*, ASCE, 113(4), 320–334.
RP2A: Recommended practice for planning, designing and constructing fixed offshore platforms, 15th Ed. (1984). American Petroleum Institute, Washington, D.C.
RP2A: Recommended practice for planning, designing and constructing fixed offshore platforms, 19th Ed. (1991). American Petroleum Institute, Washington, D.C.
Rumelhart, D. E., and McClelland, J. L. (1986). *Parallel Distributed Processing—Explorations in the microstructure of cognition*, Vols. 1 and 2, Massachusetts Institute of Technology Press, Cambridge, Mass.
Specht, D. F. (1990). "Probabilistic neural networks." *IEEE Trans. Neural Networks*, 3(1), 109–118.
Specht, D. F., and Shapiro, P. D. (1991). "Generalization of probabilistic neural networks compared with backpropagation networks." *Proc., Int. Joint Conf. on Neural Networks*, 887–892.
Specht, D. F. (1991). "A general regression neural network." *IEEE Trans. Neural Networks*, 2(6), 568–576.
Tavenas, F. A. (1971). "Load test results on friction piles in sand." *Can. Geotech. J.*, Ottawa, Canada, 8, 7–22.
Teh, C. I., Wong, K. S., Goh, A. T. C., and Jaritngam, S. (1997). "Prediction of pile capacity using neural networks." *J. Comp. in Civ. Engrg.*, ASCE, 11(2), 129–138.
Thornburn, S., and MacVicar, R. S. L. (1970). "Pile load tests to failure in the Clyde alluvium." *Proc., Conf. on Behavior of Piles*, Instn. Civ. Engrg., London, 1–8.
Vesic, A. S. (1970). "Tests on instrumented piles, Ogeechee river site." *J. Soil Mech. and Found. Div.*, ASCE, 96(2), 561–584.
Wan, W. C., Engelng, P., and Hacker, E. A. (1979). "High capacity pile foundations for China Steel Corporation integrated steel mill." *Proc., Symp. on Deep Foundations*, ASCE, Reston, Va., 495–519.

APPENDIX III. NOTATION

The following symbols are used in this paper:

$\phi_{(Shaft)}$ = average angle of internal friction of soil around pile shaft;
$\phi_{(Tip)}$ = angle of internal friction of soil beneath pile tip;
$\sigma'_{v(tip)}$ = effective overburden pressure of soil beneath pile tip;
L = pile length;
A = equivalent cross-sectional pile area; and
D = pile diameter.

J. Geotech. Geoenviron. Eng., 1998, 124(12): 1177-1185