EVALUATION OF NEURAL NETWORKS FOR MEANLINE MODEL DEVELOPMENT

10TH INTERNATIONAL SYMPOSIUM
ON TRANSPORT PHENOMENA AND DYNAMICS OF
ROTATING MACHINERY (ISROMAC-10)
HONOLULU, HAWAII 96753 U.S.A.
MARCH 7-11, 2004

Russell Japikse, Cambridge University
David Japikse, Concepts NREC
Robert Pelton, Concepts NREC/Brigham Young University

PRESENTED BY:
RUSSELL JAPIKSE
CONCEPTS ETI, INC.

EVALUATION OF NEURAL NETWORKS FOR MEANLINE MODEL DEVELOPMENT

10TH INTERNATIONAL SYMPOSIUM
ON TRANSPORT PHENOMENA AND DYNAMICS OF
ROTATING MACHINERY (ISROMAC-10)
HONOLULU, HAWAII 96753 U.S.A.
MARCH 7-11, 2004

Russell Japikse, Cambridge University
David Japikse, Concepts NREC
Robert Pelton, Concepts NREC/Brigham Young University

PRESENTED BY:
RUSSELL JAPIKSE
CONCEPTS ETI, INC.

The 10th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery
Honolulu, Hawaii, March 07-11, 2004

159

AN EVALUATION OF NEURAL NETWORKS FOR MEANLINE MODEL DEVELOPMENT

Russell Japikse, Cambridge University
David Japikse, Concepts NREC
Robert Pelton, Concepts NREC/Brigham Young University

Engineering Design Centre
Department of Engineering
University of Cambridge
Trumpington Street
Cambridge, CB2 1PZ, UK
Tel: +44-1223-332742/Fax: +44-1223-332662
R. Japikse: rdj26@cam.ac.uk

Concepts NREC
Corporate Headquarters
217 Billings Farm Road
White River Junction, VT 05001-9486, USA
Phone: 802-296-2321/Fax: 802-296-2325
D. Japikse: djapikse@conceptsnrec.com
R. Pelton: rjp9@et.byu.edu

ABSTRACT

A large database of meanline performance modeling parameters, covering a wide array of centrifugal compressors and pumps, plus a few axial machines, has been assembled by Concepts NREC. Neural Networks are one tool that might be used to extract substantive relationships from prior design and test experience. This paper presents an early investigation of neural network models applied to the available database. Resulting lessons are shared, and several resulting examples are given. Suggestions for future development are offered.

INTRODUCTION

Meanline design optimization and meanline analysis predictions for turbomachinery have been a mainstay of both design optimization work and general predictive capability since the beginning of turbomachinery design. On the meanline level, the engineer can model the bulk flow and thermodynamic parameters at each important station within the turbomachine and can carry out rapid, preliminary optimizations before detailed design work is conducted. Indeed, most of the important design decisions are actually made on a meanline evaluation basis; subsequent refinement is based on through-flow calculations and CFD analysis. To have a robust design system, one needs excellent models of meanline performance. For axial turbines, a thorough system was developed over a period of four or five decades and is well documented in numerous other references, see *Axial and Radial Turbines* (Moustapha, et al., 2003), Chapter 2; however, in all other areas of turbomachinery, the standards for comprehensive meanline modeling are more diffuse and, in many cases simply do not exist. Over the past decade, a modeling system has been developed for centrifugal pumps and compressors, with practical extensions to mixed-flow and axial stages as well (please see companion paper by Japikse, et al. 2004). Indeed, one manufacturer has achieved ±0.5% modeling accuracy over a significant range of industrial machines and operating conditions. However, the general modeling capability for arbitrary stages is less precise, as discussed in the reference cited.

Fortunately, a significant database has been developed by Concepts NREC over the past 30 years. This database is now being thoroughly evaluated and efforts are being made to obtain useful modeling relationships. In the cited companion reference, the classical approach of developing empirical modeling equations, or correlations, is pursued and evaluated. In that work, a substantial list of independent variables and a small set of dependent variables has been presented. The dependent variables are those cited as the output parameters in the nomenclature of this paper; sample independent variables are listed as the input as given in the nomenclature herein. However, the input parameters for this paper are the starting set established several years ago *without the refinements introduced in our companion paper* (Japikse, 2004).

An alternative to classical empirical correlation is to use a neural network system with part or all of the database while investigating mathematical relationships between the independent variables specified and the resultant dependent or output parameters. Each method has it own advantages and disadvantages. The present work has not progressed to the point of selecting one approach over the other. Instead, this paper reports initial considerations of using neural networks, with appropriate training and validation. We attempt to study possible dependencies and relationships among the large field of parameters employed. Some interesting conclusions are reached even at this early, preliminary point. Opportunities for further improvement and refinement in the months ahead are also outlined. Even though the results are tentative and preliminary, already some successes have been achieved and useful value of the neural networks has been derived. The reader should understand that 1) only a preliminary set of independent variables has been employed at this time, 2) the database is still being refined concurrent with the development of this paper, 3) further experimentation with the neural network systems is possible, and 4) no effort has been made to scan for bad data cases with their consequent elimination.

NOMENCLATURE

In this paper, certain abbreviations and terms will be used. These are considered standard within the relevant fields.

NN(s)	Neural Network(s)
BR	Bayesian Regularisation
LM	Levenberg-Marquardt
Over Fitting	Training, or fitting, a NN to a subset of data so that the NN cannot be generalized to the entire domain
Meanline	1D or bulk flow, station-by-station modeling of mass, momentum, and energy fluxes covering all thermo-dynamic and kinematic variables.

Output Parameters

η_a	Effectiveness of Element ***a*** (between inlet and throat) of impeller (TEIS model)
η_b	Effectiveness of Element ***b*** (between throat and exit) of impeller (TEIS model)
DR_{stall}	Diffusion ratio stall limit (TEIS model)
χ	Mass flow fraction in the secondary zone
δ_{2p}	Deviation angle for the primary zone

Input Parameters

r_2	impeller exit radius
N_S (US)	Specific Speed (US)
ReMachine	Reynolds number based on impeller exit rotating speed and the exit diameter.
b_2/r_2	Impeller tip width to impeller exit radius
t_{clr}/b_2	Ratio of tip clearance and blade exit width
Z_i	Impeller inlet blade number dimensionless)
Z_r	Impeller exit blade number (dimensionless)
AS_1	Impeller inlet aspect ratio at RMS
Rot. Num. (C,A)	Rotation Number based on averaged absolute velocity $\text{Rot. Num. }(C{,}A) = \frac{(b_1 + b_2)}{(C_{m1} + C_{m2m})}\frac{N}{60}$
B_{LE}	leading edge blade blockage
$L/D_{hydro,ave}$	Passage length divided by average hydraulic diameter of impeller inlet and exit
AS_{1t}	Impeller inlet aspect ratio at tip
β_{2b}	Impeller exit blade angle
t_{clr}/r_{1t}	Ratio of inlet tip clearance and inlet tip radius

MEANLINE PARAMETER DATABASE

The meanline parameter database is believed to be the largest, most diverse collection of centrifugal pump and compressor performance parameters. It represents over 170 different pumps and compressors that were tested under laboratory conditions. Ninety percent of the cases are radial pumps and compressors with the remainder being axial and mixed-flow pumps. At least 60 different potential performance parameters have been recorded for each case. At the present, only a subset of those parameters is being scrutinized for relationships that may be used to predict design parameters. This is a proprietary database that has been developed by Concepts NREC through the course of 30 years of design and testing research. As with all test data, there is a certain amount of noise in these data. Most test cases of the last 30 years are considered to be less noisy than earlier historical reference cases. As no theoretical relationships are known to exist between the model's input and output parameters, the amount of noise is impossible to quantify. The ramifications of this are twofold. Any modeling methods employed must be able to tolerate data error and the verification of any models will rely upon their ability to predict design parameters with greater accuracy than current practice. It should also be noted that, while this is a large database with respect to any other existing databases, it is still small from a statistical modeling perspective given the perceived number of independent variables.

With the exception of the axial turbine case mentioned above, most areas of turbomachinery are beset with the problem of bringing together the extremely diverse and complex relationships of different physical parameters for the purpose of meanline representation and optimization, especially for radial flow machines. Even the process of determining an appropriate set of independent variables, and the frequently complex non-linear coupling of different variables together, is a very complex problem. Additionally, the processing of the data to render a correlation or a mathematical model, such as a NN, is a further level of considerable complexity. This study does not attempt to finally resolve either of these issues. Instead, it takes a preliminary look at the possibility of using NN systems to replicate the data. Upon establishing potential feasibility, it is expected to require considerable additional time to find the best NNs to serve different industrial design and analysis tasks.

Past Design Practice

Previously, the design input parameters of η_a, η_b, DR_{stall}, and δ_{2p} would be obtained by estimation and reference to past design cases (via chart look up). The actual values of these input parameters cannot be known until the pump or compressor is built and tested. Experts who are intimately familiar with the type of machine they are designing can typically produce much more accurate initial estimates. Compared with an engineer who infrequently designs pumps and compressors, the expert's design progresses faster and is more accurate (see companion paper, Japikse, 2004). Although many factors are present, accurate starting empirical modeling values for a design play a significant role.

Current Design Practice

Current state-of-the-art design practice combines the information contained in the meanline parameter database with the experience of experts. The information contained in the meanline parameter database can be accessed through a weighted, similarity matching and averaging program. In this approach, up to 14 input parameters (nomenclature section) may be selectively used to identify similar past designs. The output parameters are then combined in a weighted average. This simple approach has proven to be more effective than previous attempts to use statistical correlations. It is believed that this approach has proven to be robust because it builds upon the experience of seasoned engineers. The parameters that were used to identify similar past designs and the weights that were used in estimating the output parameters were all hand tuned by experienced engineers. In effect, these weights provide a rough representation of the engineers' experience with respect to past design cases.

In effect, the experience of others is used to approximate an unidentified relationship between the input and output parameters. This approach works well with noisy data and helps with sparse data sets. Unlike correlations or NNs, this method is likely to find some approximate relationship that covers the entire domain.

However, it is difficult to describe inherently implicit engineering knowledge, and much less accurately describe an unknown relationship between a large set of inputs and outputs. Although a significant advancement over educated guesses, this method leaves room for substantial improvement. It is most likely that these gains will be made by approximating the inherent relationship between the input and output values.

DATA MODELING OPTIONS

Statistical Correlations

Most engineers completed their academic education with the understanding that much of the physical world can be represented in simple empirical relationships to specify heat transfer, skin friction, drag, etc. We are used to seeing relationships between one or two independent variables and a desired dependent variable, which would give us the capability to do engineering modeling for a particular problem. Such correlations have been the mainstay of basic engineering work for more than a century. Engineers invariably hope to find such basic relationships. However, studies in the current century go far beyond the simple correlations of the period of our academic training. No longer are we working with simply two independent variables, in essentially a linearly independent set, but we now deal with a large multitude of geometric and fluid dynamic variables, some of which are highly coupled even in a non-linear manner.

One of the initial tasks when building any kind of model is determining which variables to use for predicting the output. If all of the important factors are not included, then a satisfactory model cannot be developed. Including irrelevant variables only introduces noise that will hinder the model's performance. Several statistical tools have been used to select the important variables to be used in predicting the TEIS and two-zone impeller modeling parameters (see Japikse 1996 and 1997).

The covariance is a basic statistical tool for determining if two parameters vary together. Scaling the covariance by the standard deviation of each parameter (Montgomery, 1998) yields a normalized result called the correlation. Since the correlation is non-dimensional it can be used to compare relationships between sets of parameters with different units. A matrix of correlations can then be developed including various input parameters and a broad range of possible variables. Relevant variables can then easily be identified from those factors that do not affect the modeling parameters.

Unfortunately, a correlation analysis only detects linear association between variables. Therefore, any strong non-linearity in the correlations must be identified using other methods. Like other statistical methods, correlation values are dependent on both the quality and quantity of the input data. Therefore care must be taken to ensure that correlations represent real trends in the data and not simply input bias.

Past research studies (not reported, but conducted at Concepts NREC) have attempted to investigate the non-linear character of this performance modeling problem. Strong relationships were found between particularly selected coupled terms, usually the product of at least two different variables, said variables raised to different powers. No mathematical method has yet been found to a priori select these variables or their power relationships. Such studies were conducted purely by intuition in the hands of a good mathematical modeling specialist. Although the results achieved a sensible level of statistical error, they did not rise to the level of reliability required for daily engineering optimization work. Thus, early studies for statistical correlation showed that the problem was potentially tractable, considerably complicated, and not ready for our reduction to practical equations. Instead, it was determined that more work was necessary for improving the database, reducing noise, and studying alternate numerical methods, such as NNs.

Neural Networks

NNs are an enticing option; they can handle noisy data, small or large data sets, approximate any function and require little a priori knowledge as to what the function is. Like all tools, they do have some real limitations. As they can approximate any function, it is quite possible to approximate a set of points, modeling the sub space in which the data exists, rather than the entire design domain. It is also possible that discontinuities will exist in the NN model. However, steps were taken to minimize this risk.

NNs consist of a set of neurons (Figure 1) arranged in a series of layers with each output from the previous layer feeding every input of the next layer. A set of known inputs is supplied to a NN; the outputs of the NN are then compared to the known output. The error then drives the adjustment of the NN values.

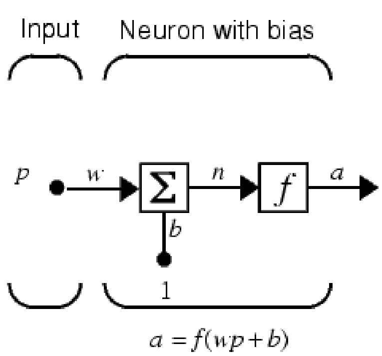

Figure 1. A Single NN Neuron (H. Demuth, 2000)

In the above neuron, an input (*p)* is supplied to the neuron which produces an output *a*. *(p)* is subject to a weighting value (*w)*, and is summed with a bias (*b*), resulting in $n = p \bullet w + b$. In turn, this feeds the transfer function. Most commonly, the transfer function is either log sigmoid, tan sigmoid, or pure linear. Given enough neurons, a two-layer NN with tan sigmoid neurons in the first layer and pure linear neurons in the second layer can approximate a given set of points. Three-layer NNs are the practical limit for modeling complex functions.

When a NN is created, the biases are typically set to one and the weights are given random values. At this point the network must be trained to model the desired function. Numerous methods for training exist, but the most pertinent ones for this application are Levenberg-Marquardt (LM) and gradient descent with Bayesian Regularization (BR).

The LM is a gradient descent method. It works very well and is a standard training method. Advantages of this method are its fast convergence and stability. Certain disadvantages occur due to the nature of our data set. Using LM training, it is easy to over train when using small data sets, particularly with noisy data. The data being used here are both. A large data set is on the order of 1000+ points. It is not feasible to expect enough data for this method to become the preferred method. This method may still be used in two different ways. The first is to gauge the potential of a NN architecture before training with a slower method. The second is to use LM training with early stopping. LM training is used with early stopping to avoid over training. In early stopping, the data are divided into three sets: a training set, an error comparison set, and a validation set. This method was tested and found to be unsuitable; refer to the results section for an explanation.

In BR a gradient descent training method is used along with regularization to improve the NN's generalization. In most training methods, the performance function is the mean sum of squares of the network errors (Demuth, 2000).

$$F = mse = \frac{1}{N}\sum_{i-1}^{N}\left(e_1\right)^2 = \frac{1}{N}\sum_{i-1}^{N}\left(t_1 - a_1\right)^2 \qquad (1)$$

To improve generalization, this function is modified "by adding a term that consists of the mean of the sum of squares of the network weights and biases" (Demuth 2000).

$$msereg = \gamma mse + \left(1-\gamma\right)msw \qquad (2)$$

where γ is the performance ratio and

$$msw = \frac{1}{n}\sum_{j-1}^{N} w_j^2 \qquad (3)$$

BR training "will cause the network to have smaller weights and biases, and this will force the network response to be smoother and less likely to overfit" (Demuth 2000). BR training also tends to take more time to converge. In contrast, LM training with early stopping will converge more quickly, but will result in a choppier function approximation.

For a function approximation problem such as this one, where the inputs are not known for certain, the data set is small and contains noise, and the expected function is unknown, NNs are the clear choice. Used judiciously, NNs may offer the best possibility for accurate estimates of pump and compressor design parameters.

Neural Network Development

Using a few educated guesses as a starting point, the NN(s) were developed by trial and error. The networks were implemented in MATLAB®[1] and built upon the NN toolbox. Two- and three-layer NNs were constructed and trained using LM, LM with early stopping, and BR. Initially, it was hoped that a two-layer NN would be sufficient to approximate the unknown relationship. An experienced engineer chose 13 input values upon which the output values are believed to be dependent. These values are listed as input parameters in the nomenclature. Networks with 15 to 35 neurons in the first layer were tested, while the training error was monitored. The second layer had one neuron, as separate networks were constructed for each output variable. The first layer used tan sigmoid transfer functions while the second layer was pure linear. This level of work reinforced the fact that no simple function exists to be approximated. The LM training methods were able to quickly and accurately approximate the data points, while BR was 23 times less accurate. This is a clear indication that the LM method was over fitting the data and, hence, of no value; if any useful relationship did exist, it would only be found using more complex networks.

A variety of three-layer NNs was implemented. These networks had 13 neurons in the first layer, 15 to 45 neurons in the second, and one neuron in the final layer. Respectively, the transfer functions were tan sigmoid, tan sigmoid, and pure linear. It is preferable to have the fewest number of neurons in a network. Those neurons are more likely to model the general trends, rather than the data noise (Demuth 2000). Although LM training will likely over fit this data, it was still found to be useful. It was used to determine if a network was likely to converge before the more time consuming BR training method was run.

LM training can also be used with early stopping techniques to avoid over fitting the data. As the network is being trained on the first subset, the outputs for the test subset are generated. When the error for these outputs starts rebounding, training is stopped as the network has begun over training. The third subset is subsequently used for validation. If there is significantly more error in the predicted validation values than in the other two data subsets, then the data were poorly partitioned. This method was also tried as it offers robust, fast training without over fitting. Unfortunately, this data set was too small to effectively use this method and the networks were unable to train below 100% error.

The final attempt used BR with the three-layer NNs. The data were randomly ordered, two-thirds of the data were used for network training, and the remaining third was saved for validation. Again, a range of networks was examined. The results for eight of these networks are given in Table 1.

[1] MATLAB is a registered trademark of the MathWorks, Inc.

Table 1. Neural Network Training Runs with Bayesian Regularization (2/3 of data sets)

Output	Number of Neurons			Percent
	Layer 1	Layer 2	Layer 3	
χ	13	15	1	6%
χ	13	20	1	5%
δ_{2p}	13	15	1	24%
δ_{2p}	13	20	1	24%
η_a	13	45	1	86%
η_a	13	30	1	86%
η_b	13	15	1	1%
η_b	13	20	1	1%

One can clearly see that a well-defined function has been approximated between the inputs and several of the output values. Reviewing the networks after they are trained, it is possible to study the weighting matrix. A variable that is entirely, or almost entirely weighted to zero is likely to be irrelevant. None of the inputs were trained to zero. This suggests that all of the inputs have some bearing upon the output.

Validation

The networks were validated by comparing the predicted values for the validation data set against the known values. Of interest here is the absolute error rather than the percent error. As one would expect, the performance of the networks varied greatly depending on which variable was being predicted. Table 2 shows several basic descriptions of the prediction error over the validation data set.

Table 2. Prediction Error over the Validation Data set (last 1/3 of data sets)

Output Variable	Number of Neurons	Absolute Error			
		Mean	Min	Max	Std. Deviation
χ	15	0.137	0.001	1.544	0.237
χ	20	0.136	0.001	1.580	0.252
δ_{2p}	15	3.391	0.067	16.441	3.437
δ_{2p}	20	3.237	0.063	16.523	3.224
η_a	30	0.141	0.001	0.502	0.127
η_a	45	0.141	0.001	0.502	0.127
η_b	15	1.101	0.003	20.498	3.021
η_b	20	1.246	0.004	24.442	3.543

A few things are worth noting from Table 2 and from Figures 2 through 5. The results may be scrutinized for over fitting by examining the validation plots, Figures 2-5, directly and by contrasting the results from each network. Examining the plots individually, one can see that the validation error (see Table 2) is modest and is consistent with the training error (e.g., it does not differ by orders of magnitude [see Table 1]). The separate networks not only differ in the number of their neurons, they also had different starting weights. Each pair of networks is trained to nearly identical end states. This means that it is extremely unlikely that these networks became trapped in a local minimum during their gradient descent training. It should also be mentioned that numerous other training runs of these three-layer networks, both with the same and different numbers of neurons, were performed. Each solution was consistent with those reported here.

The networks for χ and η_a both modeled the unknown relationship accurately enough to give commercially viable values. η_b was modeled with limited accuracy, and it is not yet acceptable for industrial use. δ_{2p} was predicted with modest accuracy, but may be usable for design work.

The error plots (Figures 2 through 5) give an indication of the probable accuracy of the NNs and provide a visual check for over fitting. The portion of the validation set with significant amounts of error (10% to 20%, Figures 2-5) could reflect poor quality data or local regions of over fitting. It will require further investigation to determine which scenario exists.

Preliminary remarks concerning accuracy will be limited to Tables 1 and 2, with further remarks offered in the following section. The parameter which controls the secondary mass fraction, χ, is shown to have a percentage error of about 5%, as shown in Table 1, where the training was conducted; when the validation was completed with the remaining 1/3 of data sets, the mean error was limited to less than 0.15, which is not outstanding, but certainly very workable. At this point, for purely a proof-of-concept study, this is quite acceptable. There is a clear indication of a relationship developing which, presumably, would be improved in further work. For the deviation of the flow leaving the impeller, δ_{2p}, the error is ±24% according to the training runs, Table 1, which appears to be a large error, but is actually deceptive. For example, if the deviation number was approximately 4°, then the error would be ±1° which is quite workable in design practice. The absolute error is somewhat higher as the validation study of Table 2 shows, approximately 3°. This may still be workable, but is at least ballpark relevant (see below).

Moving to the effectiveness of the inlet portion of the impeller, η_a, we detect very high apparent percentages in the training runs, Table 1. This is believed to be unintentionally deceptive. A few ancient data sets are included, which have effectiveness levels in the range of 0.00 to 0.3; however, meaningful design problems really only have values from at least 0.5 up to 1.1. The very low values may be skewing this examination badly and contributing to very strange percentages. This would suggest that we should look very carefully at the data sets and perhaps refine them to specific problem areas (for example, only centrifugal compressors or other particular

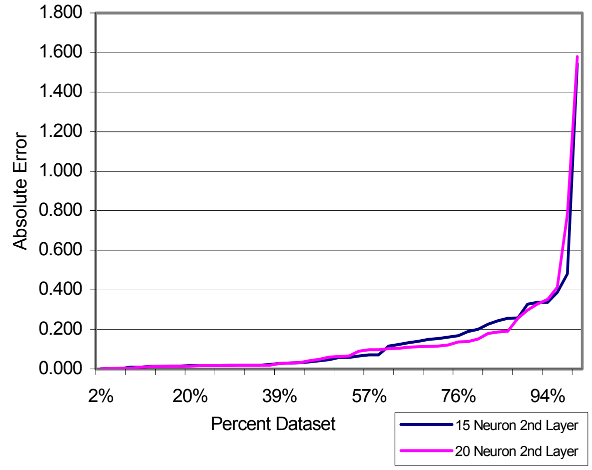

Figure 2. Sorted χ Prediction Error

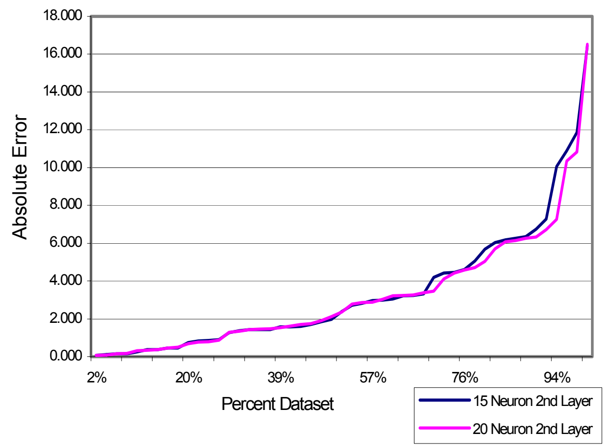

Figure 3. Sorted δ_{2p} Prediction Error

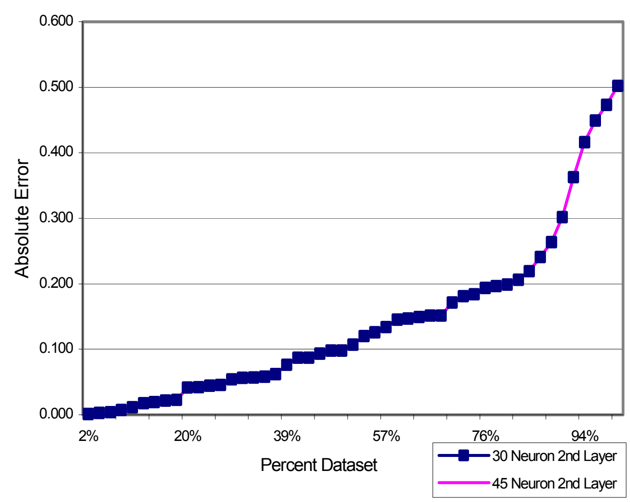

Figure 4. Sorted η_a Prediction Error

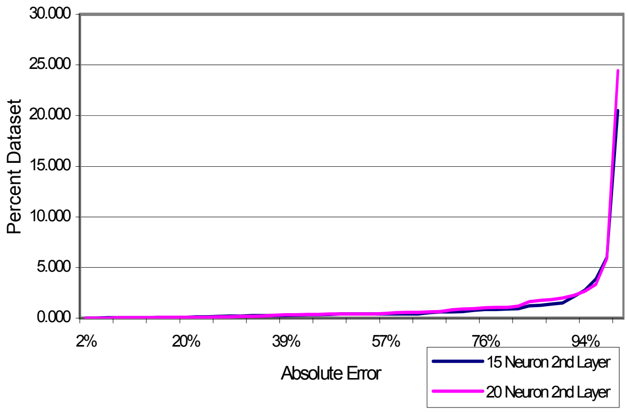

Figure 5. Sorted η_b Prediction Error

subsets). By contrast, the validation data set shows far superior characteristics with a mean error of only 0.14 and a standard deviation of 0.13. These numbers are quite workable for a preliminary model, although past correlation experience indicates that we will undoubtedly be able to reduce these values as more work is done with the NN methods. The maximum error for η_a of 0.5 is unacceptable and may reflect a single data point or just a few (to be reviewed). Finally, the very low training error for η_b when contrasted with the very high validation error for η_b is very disturbing. It appears that the NN system has not worked well for these parameters yet. It is quite possible that we have highly non-linear effects involved here and have not yet captured the essence of their impact in our modeling study.

Several steps may be taken to improve the accuracy of the networks. The accuracy of all of the networks may be improved by pruning spurious data points from both the validation and training data set. Initial investigation of these spurious points may be made by examining those points in the training and validation data sets that had a high amount of prediction error. For instance, those points may simply correspond to radial or mixed-flow pumps or open versus covered impellers. With a small data set such as this one, it is best to proceed cautiously for two reasons. If the data are pared too far, the resulting data may not be representative of pumps and compressors in general. And, with such a small amount of data, it is important to retain as much data as possible.

DISCUSSION

The results of the first figure are very interesting. Figure 2 gives a relationship for χ, the secondary flow mass fraction for the impeller, which is quite workable in practice. Any correlation that hits the proper value of χ with ± 0.05 is truly excellent for modeling purposes and most cases within ± 0.1 are still very good. Thus, approximately 60% of the cases meet this criteria. There is still considerable utility in many cases that fall within 0.2 absolute error, especially at low N_s, thus approximately 80% of the cases show the probability of being useful. From 0.2 to 0.4 the usefulness depends very much on the range of other parameters which accompany it and some cases will still be useful, but most cases above 0.4 will, in fact, be useless. Thus, it appears that there is a general usefulness of this relationship covering approximately, let us say, 75% to 85% of all the data. There is some small gain by using the 20-neuron second layer over the 15-neuron second layer, but one would tend to prefer the simpler relationships and, hence, we would be inclined to abandon the higher order model and might even consider lower order models as well.

Somewhat comparable results exist for the impeller deviation angle (δ_{2p}). In this case, for the design problem, there is an additional degree of latitude, in that an experienced designer will always shade the deviation a little bit extra toward the negative side to give a conservative design, i.e., a design which tends to produce a little higher head rather than one that could fall short. For analysis work, where one must model an entire map, you cannot bias results to one side. Since design is frequently the preeminent problem, one has an extra measure of safety. Any agreement within approximately 2° or 3° is generally workable and a designer can bias these results very well, indeed. This brings approximately 60% of the cases into a good relationship. In other cases, with even a few degrees more absolute error, the designer will still be able to bias the case safely, particularly if they fall within classical, well-known regimes. For the sponsoring company, this has not been a problem area for the past several decades. Thus, again, a substantial majority of the data appears to be reasonably correlated, if not ideally at this point and time.

η_a and η_b are the important modeling parameters for the impeller passage effectiveness divided between the inlet and the passage sections respectively. η_a has been fairly easily correlated by various past correlation models. It is possible to obtain workable and useful models with approximately ± 0.1 absolute error with some of the better agreements at half this level. Again, this reflects decently on the NN model since approximately 50% of the data meet this level. Errors at 0.15 are not too bad as shown around 60% or 65% level of the data and are still workable. Higher values are far more problematic and will lead to more significant errors in design and performance modeling. It would appear that more work could be done to improve this parameter. By contrast, η_b fails to meet a useful criteria for entering design work. An engineer with a decade of design experience can guess values better than the η_b correlation in Figure 5.

In short, the NNs have shown that some of these parameters have already been decently modeled, quite well in the case of χ, with one of them (η_b) falling far short of a useful design value. For an initial evaluation, these results are encouraging. As the team works further to compare the results of the correlation studies and the NN studies together, it appears quite certain that the database can be improved further, spurious cases identified and eliminated, select problem areas (for example, compressors only, or pumps only) chosen, and higher order models developed. Whether non-linear fitting of higher order correlation procedures or NNs prove out to be the best is not the important point. Having several different mathematical procedures by which to analyze data and search for relationships is the real value of this work. This particular methodology has the cleanest overall numerical method for searching for statistically optimized models in a coherent fashion. Presently, NNs are confirming that multi-variable relationships do exist and some encouraging models are resulting.

The work presented herein is an initial scoping study to determine the possible appropriateness of NNs to address the problem of synthesizing the core physical relationships embedded in a large set of related data for a class of turbomachinery performance. Based on the initial success of this investigation, it is clear that more work should be done and additional ideas should be brought into the study. Several ideas have already been indicated (above) of how the process, even in its present level, can be improved by working with both the data and the tools. Additionally, more powerful mathematics can be brought to bear on the present level of evaluation. In a companion paper we have looked at processes of developing correlations, particularly emphasizing linear regression analysis. It is clear that both avenues should be pursued much further. At a more advanced level of NN investigation, we wish to look at particle swarm optimization as an alternative method for training networks (van den Bergh, 1999 and Kennedy and Eberhart, 1995). Further, the distinctive roles of genetic algorithms, fuzzy logic, kriging etc., remain for further examination, particularly following the lines of contrast set by E. T. Jaynes, 2003). The present work has established a valuable initiation; additional attention over the next several years will presumably lead to further breakthroughs.

CONCLUSIONS

The present work was an exploratory study to determine whether NNs might work with the database of approximately 170 builds of test data describing meanline performance of centrifugal compressors and pumps. The initial results are encouraging, but not without some disappointments as well as successes. The following observations may be made:

1. The present study seems to be a first of its kind investigation of the use of NNs for meanline performance model development and is investigative in nature;
2. From a statistical and practical point of view, a reasonably accurate prediction of the impeller secondary mass flow fraction, χ, has been achieved with a decent first model for an impeller exit flow deviation, δ_{2p}, and perhaps for the inlet effectiveness, η_a, parameter. The impeller passage effectiveness, η_b, parameter has not yet met with success.
3. The NN approach is a flexible one which is easily able to retrain to accommodate additional data.
4. This NN study has confirmed the existence of predictive relationships with a large base of independent variables and rational mathematical, statistical behavior.
5. Future work must include refining the input parameters to be better conceived parameters, perhaps drawing on the sensitivity studies conducted for the companion correlation study. Data sets can be improved and extended, spurious data points can be eliminated, and more detailed systematic studies of the type of NNs to be employed can be pursued.

The objective of studying the feasibility of NNs on a preliminary basis has been achieved. The initial success indicates that this tool should be used further to give both an overall numerical understanding of the database at large, a possible primary set of models for design and analysis predictions, and at least an alternative set of models for cross checking any correlations which might be developed by other means.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the contributions of Messrs. K. N. Oliphant and O. Dubitsky plus Dr. P. S. Weitzman through suggestions and review.

REFERENCES

Demuth, H., and Beale, M., *Neural Network Toolbox, User's Guide, Version 4,* for use with MATLAB®, The MathWorks, Inc., 2000.

Japikse, D., *Centrifugal Compressor Design and Performance*, Concepts NREC 1996.

Japikse, D., Marscher, W.D., Furst, R., *Centrifugal Pump Design and Performance*, Concepts NREC 1997.

Japikse, D., Oliphant, K.N., Pelton, R., "Optimization in Turbomachinery Data Reduction," The 10th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery, Honolulu, Hawaii, March 2004.

Jaynes, E. T., *Probability Theory: The Logic of Science*; Cambridge 2003.

Kennedy, J., and Eberhart, R. D. "Particle Swarm Optimization," Proceedings IEEE International Conference on Neural Networks, IV, 1942-48. Piscataway, NJ: IEEE Service Center.

Montgomery, D.C., Runger, G.C., Hubele, N.F., *Engineering Statistics*, John Wiley & Sons, New York, 1998.

Moustapha, H., Zelesky, M.F., Baines, N.C., Japikse, D., *Axial and Radial Turbines*, Concepts NREC, 2003.

Van den Bergh, F., "Particle Swarm Weight Initialization in Multi-layer Perception Artificial Neural Networks", In Development and Practice of Artificial Intelligence Techniques, pp. 41-45 (Durban, South Africa). Sept. 1999.