Relating Product Specifications and Performance Data With a Neural Network Model for Design Improvement

Alice E. Smith

Bopaya Bidanda

Safouen Ben Brahim

Department of Industrial Engineering
University of Pittsburgh
1048 Benedum Hall
Pittsburgh, PA 15261
412-624-5045
aesmith@engrng.pitt.edu

Abstract: This paper presents research resulting in a neural network model relating product design specifications and performance testing results using data from a sanitary ware manufacturer. The main constraint of the work was the limited availability of actual data for neural network training and testing, a situation often found in real situations where *a priori* product knowledge is limited during the product design phase. The authors used two training techniques, the standard hold back and the leave-k-out, for the neural network model to leverage the sparseness of the data. Neural network results are compared and contrasted to statistical models relating product design and performance. This work is an exploration of the value of neural network models to assist with interactive product design.

Revised October 1992 for *Journal of Intelligent Manufacturing*

Relating Product Specifications and Performance Data With a Neural Network Model for Design Improvement

1.0 Introduction

The public concern and legal mandate to conserve water has motivated sanitary ware manufacturers to design new low consumption toilets. This design process presents two interesting issues: first, toilets have complex, irregular shapes with significant room for more creative design contribution and functional improvement. Second, the design process is time consuming and fraught with uncertainty since it is based mostly on trial and error. Toilets must be certified to meet a series of tests developed by the American National Standards Institute (ANSI) (ASME 1990). However, manufacturers seek to develop products which improve upon ANSI minimum standards, and iteratively modify their designs for the best performance, using expensive prototype products and performance testing.

We worked with a large U.S. manufacturer of sanitary ware products to improve the quality and timeliness of their toilet designs through development of neural network models relating multiple quantitative and qualitative design parameters to product performance over a series of ANSI tests. These models will allow product designers to iteratively and interactively test parameter specification changes and evaluate corresponding changes in product performance before a prototype product is actually built and tested.

Because supervised neural networks need training data, we obtained ANSI test results on the products currently produced by the manufacturer, however this was a limited set. To create the best possible neural network models, and to validate our results, we adapted strategies that maximize learning with sparse data. The models were compared against the statistical modeling techniques - multivariate linear regression and multivariate non-linear regression. We used the neural networks to build sensitivity models to be used interactively by product designers to identify design specification decisions that are robust.

Two previous works have used neural networks for product design by training a multi-layered perceptron to act as an expert system (Hung and Adeli 1991, Zarefar and Goulding 1992). Both efforts used documented design policies, heuristics and calculations to construct a rule base (or decision table). The neural network was then trained on representative examples adhering to this rule base. Previous work using neural networks for interpolative modeling in manufacturing design include (Chryssolouris et al. 1989, Liu and Liu 1990, Cariapa et al. 1991, Schmerr et al. 1991). The first paper dealt with design of manufacturing systems using simulation augmented with neural network interpolation. Liu and Liu used neural networks to interpolate between test points relating three circuit design parameters and two voltage conditions to one performance parameter - the variability of output current of the circuit. They wished to find the design settings which yielded the smallest variation over all voltage conditions. Their approach was not dissimilar from ours, though they concentrated on a smaller problem and did not isolate effects through sensitivity studies. The latter two works focused on interpolating between Taguchi design points using a neural network so that a full factorial design could be simulated to search for optimal design parameter settings. These works used small subsets of whole products to test their approach. The Taguchi approach was not appropriate for our products as they do not adhere to any known analytical description of performance. Our approach is distinct from these previous efforts by addressing the product as a whole, both in terms of design parameters and performance results. Another contribution of this work is development of full range sensitivity models to be used interactively during product design.

2.0 Description of Product Design Parameters

Toilets in use today have different designs and styles. One of the styles used in this study is shown in Figure 1 to illustrate the generic components. The product design parameters that affect toilet performance are water volume used per flush, the flushing mechanism, the rim and jet delivery system, the jet hole, the water seal, the pan contour, the trap seal depth, the trap size, and the water flow rate.

Figure 1. Typical Toilet Showing Design Parameters.

There are two types of **flushing mechanisms**: the gravity flush closet and the pressurized flush closet. All toilets in this study used the gravity flush mechanism, which is the method of choice for home use. The **rim wash system** design can be either an open slot without a jet (a European style) or a set of punched holes around the rim which work with a jet to disperse water (Kinney 1977). The **jet hole** is located at the bottom of the bowl, facing the trap passage way. An excessively large jet hole causes less jet pressure; on the other hand, a small hole restricts water flow and produces noise in the rim. The **water seal** is the cross sectional area of standing water in the toilet. **Contour pan** styles are different from one bowl to another. The available models from this manufacturer are round and square pans, which were handled as qualitative parameters. **Trap seal depth** is the water depth in the trap. The **trap** itself is sized so that either a 2" or a 2 1/8" ball can pass through. The **water flow rate** from the toilet tank to the pan depends on the diameter of the rim supply hole, the height of the water inside the tank, the water pressure, and the type of the rim delivery system. The toilets in this study were all tested at a rate equal to 1.6 gallon per minute, and a flowmeter was used to maintain the same rate for different products.

3.0 Performance Evaluation of the Product

The performance of toilets in the U.S. is usually evaluated using tests developed by ANSI, and approved by the American Society of Mechanical Engineers (ASME). The overall performance is based on a variety of diverse criteria:

1. The ability to clean the bowl in a single flush.
2. The ability to handle floating paper waste.
3. The proportion of water exchanged during a flush.
4. The distance solid waste is propelled down a sewer drain line.
5. The aesthetics of the product to attract customers.
6. The durability, or how long the fixture lasts.

ANSI has developed two tests to evaluate flushing and handling of floating waste. In the first one, the **Ball Test**, the toilet flushes a load of 100 three quarter inch plastic balls. The fixture meets minimum standards if it passes at least 75 balls, though the objective is to pass as many as possible. The second test, the **Granule Test**, measures the ability to dispose of floating waste. Here, approximately 2500 buoyant polyethylene granules are thrown in the bowl and the fixture is considered acceptable if it leaves no more than 100 granules after flushing (ASME 1990).

The water exchange ability of a toilet affects the cleansing of the bowl after bulk waste disposal. A poorly designed bowl leaves streaks of residue that causes further flushing, and therefore more wasted water. One test to evaluate this ability, the **Ink Line Test**, uses a felt marker to draw a line around the inside of the bowl just under the rim, and measures how much of the ink stripe is left after flushing. Another test, the **Dilution Ratio Test**, starts with a blue dye poured into the bowl to turn the water color to dark blue. After flushing, water color is measured when the bowl is refilled. Measurement is done by comparing the residual water color to a reference color with labeled concentrations (ASME 1990). This is an important test, since complete water change in the bowl with each flush prevents bacteria growth.

To measure how well a toilet transports solid waste down the sewer pipe, a 60 foot long clear plastic drain pipe is connected to the toilet for the **Drain Line Transport Test** (ASME 1990). A well designed toilet can be expected to transport a whole simulated load down the pipe in three flushes or less (*Consumer Reports* 1990). This is an important quality to avoid clogging of the main sewer line from the dwelling to the sewer system.

For the last two categories of aesthetics and durability, there are no standardized tests. These qualities, along with those of manufacturability and cost, are important, but were unable to be included in our model due to lack of comparison data.

4.0 Building a Neural Network Model

4.1 Overview of Neural Networks

An artificial neural network is an aggregation of processing elements called nodes or neurons, whose function and organizational architecture is similar to biological neurons. A neural network model is defined mathematically as a directed graph that has the four following properties: (Muller and Reinhardt 1990)

1. For each node *i* there is an associated state variable.
2. For each link or connection between two nodes *i* and *j* there is a real-valued, trainable weight.
3. For each node *i* there is an associated real-valued bias.
4. For each node *i*, a transfer function is created, usually non-linear. The function determines the state of the node as a function of the bias, of the weights of the incoming connections, and of the states of the nodes connected to it by these connections.

Figure 2 is an illustration of the basic building blocks and processes of a multi-layered perceptron artificial neural network. We selected multi-layered perceptrons with a sigmoidal transfer function trained by error backpropagation because of their ability to handle both binary and analog inputs and outputs, the supervised learning technique (Rumelhart et al. 1986), and their proven value for modeling (Werbos 1988, Okafor et al. 1990, Andersen et al. 1990, Ho et al. 1991, Smith 1991, Steck et al. 1991, Lambert and Hecht-Nielsen 1991).

Figure 2. A Typical Multi-layered Perceptron Neural Network With One Hidden Layer.

4.2 Building Product Design Models

To build the model, fifteen toilet designs from the manufacturer were evaluated. Both the design specifications and the product performance over the series of five ANSI tests conducted by a professional, independent testing service were obtained for each toilet model. The Ink Line Test was not included in the analysis because the results were identical; in other words, the toilets passed this test and thus the test did not discriminate between toilets. We used the mean of the three test runs conducted by the service for each toilet because neural networks become confused during training if identical inputs yield different outputs. By using the mean value of each test for a given toilet, we trained the network to output the expected performance value on the ANSI test. This expected response in performance to a particular product design specification is the core of the neural network model.

Our first approach used a single network to simultaneously estimate all ANSI test results. This is an efficient use of computing power and training time, however it was unsuccessful, with networks failing to learn within the prescribed error tolerance or iteration limits. The networks which did train, performed poorly. We believe the multiple decision approach was not viable because of the limited number of training vectors. Learning to correctly make four simultaneous estimations is much harder than making a single estimation, and requires much more training data.

In our next approach, every network had an input vector with seven input parameters, (each one a design parameter for that toilet), and one output vector (the test result for that toilet). We translated as many parameters as possible into continuously valued numeric measures so that toilets could be better compared. For example, the number and sizes of the holes in the rim design were converted to an area, as was the open slot rim design. This translation proved more successful than our early attempts which included many variables represented as binary qualitative parameters. All inputs were normalized between 0.1 and 0.9, typical values for sigmoidal based neural networks. A sample training vector (input and desired output) as raw data, and after being transformed, is shown in Table 1.

Table 1. Typical Raw and Transformed Training Vector

	Water Volume	Rim Design	Trap Seal Depth	Water Seal	Trap Size	Jet	Pan Contour	Ball Test Results
Raw Data	1.6 Gallons Per Flush	1 7/16" + 24 1/8" Holes	2 5/8"	10 1/4" x 12"	2 1/8"	Two 1 1/8" Holes	Square	93 Balls Passed
Transformed Vector	0.1410	0.1044	0.500	0.7614	0.10	0.748	0.10	0.6314

Our most significant handicap in developing the design model was the lack of more data points (i.e. products) due to the limited number of toilets designed and manufactured. Unlike many neural network manufacturing applications (such as machine vision or quality control) copious amounts of data cannot easily be generated in product design. We used the standard approach of holding back a percentage of randomly selected data points (in our case, 30%) for testing, but we also leveraged our sparse data by using the "leave-k-out" procedure for training (Lawrence 1991). In this approach, one out of the fifteen input/output vectors was held back each time for testing (i.e. k = 1), and networks were trained, changing each hold back vector each time. A total of fifteen different subsets were used for every ANSI test except for the Drain Line Transport Test, which was not performed on four of the toilets, so these networks had eleven subsets. This translated to a total of 56 neural networks (15 subsets * 3 ANSI tests + 11 subsets * 1 ANSI test) for the leave-k-out approach, while the traditional approach only required 4 neural networks, one for each ANSI test. Since the size of each network was modest (one hidden layer with three hidden neurons) and the number of training vectors small, training proceeded rapidly and creating the 60 networks was not burdensome.

We used a scaling factor to dynamically determine the learning rate during training. The scaling factor is a measure of the rate of change in the learning rate vis a vis the error generated in the training. Large scaling factors engenders quick changes in the learning rate, while smaller scaling factors slow it down, and a 0 scaling factor keeps the learning rate constant (Baffes 1989). We also used a momentum term which allows past weight changes to be taken into account as shown in the training formula (Sejnowski and Rosenberg 1987):

$$D_p W_{ij} = \eta (\alpha D_{p-1} W_{ij} + (1 - \alpha) \delta_{pi} O_{pi})$$

where $D_p W_{ij}$ is the change in weight connecting neuron j to neuron i for input vector p, O_{pi} is the output of neuron i for input vector p, δ_{pi} is the error of the output of neuron i for input vector p times the derivative of the transfer function, η is the training rate, and α is the momentum term. In this research, a scaling factor value of 0.06, a learning rate of 0.5, and a momentum value of 0.9 were chosen throughout all networks. Training was considered complete when either all training vectors were estimated to at least within 10% (a typical stopping criteria for backpropagation training) or when 10,000 iterations through the training set had passed.

5.0 Results of Design Models

The first results presented are the performance of neural networks for both the standard hold back training/testing procedure and the leave-k-out procedure. Table 2 shows the percent mean absolute error (MAE) for both approaches, where the leave-k-out MAE is the average MAE over all the networks (15 for the Ball, Dilution Ratio and Granule Tests, and 11 for the Drain Line Test). Although it might appear that the hold back training strategy obtains lower error rates than the leave-k-out, this is misleading because the leave-k-out results are for testing on all observations instead of a 30% subset. Therefore, we believe the leave-k-out results are better representations on how the neural network approach does in general.

Table 2. Percent MAE for Neural Network Estimation Models on Test Set.

Training Strategy	ANSI Performance Test			
	Ball	Granule	Dilution Ratio	Drain Line
Hold Back 30%	2.052	45.09	41.05	2.241
Leave-k-out	1.860	95.48	65.85	0.756

We next compared these two neural network training strategies to multi-variate linear regression. The statistical significance (p value) and the fits (adjusted Coefficient of Determination - r^2) are shown in Table 3. Since the Granule Test had the worst fitting linear model, we also fitted a non-linear regression model, whose significance and fit are also shown in

Table 3. With the non-linear model of the Granule Test, all regressions had significance surpassing the $\alpha = 0.10$.

Table 3. Regression Significance and Fit for the Performance Tests.

	ANSI Performance Test				
Statistic	Ball	Granule - Linear	Granule - Non-linear	Dilution Ratio	Drain Line
p value	0.041	0.114	0.010	0.008	0.076
Adj. r^2	0.607	0.447	0.801	0.769	0.694

To compare the regression estimations to the neural network estimations, regressions using the identical training vectors of both neural network training approaches, the hold back 30% and the leave-k-out, were built. Therefore 60 regressions were created - 4 for the standard hold back approach and 56 for the leave-k-out. The estimation errors on the test data for these regression models and their neural network counterparts are shown in Table 4. For the leave-k-out models, the standard deviation of the estimations over the set of 15 (11 for the Drain Line Test) are also shown.

Two sample statistical tests were performed to test whether the central values of the two populations (neural network and regression) were significantly different for the leave-k-out training strategy. These included the parametric Paired T Test and the Two Sample T Test, and the nonparametric Rank Sum (Mann-Whitney) Test, the Wilcoxen Signed Rank and the Median Test. None of the test results were statistically significant to $\alpha = 0.05$. This was due to the difficulty of establishing significant population differences with small samples.

Table 4. Percent MAE for Neural Networks and Regressions on Test Set.

Model	ANSI Performance Test			
Strategy	Ball	Granule	Dilution Ratio	Drain Line
Hold Back Neural Net	2.052	45.08	41.05	2.241
Hold Back Regression*	2.978	40.70	52.41	#

Leave -k-out Neural Net	1.860 (σ=1.547)	95.48 (σ=97.75)	65.85 (σ=78.12)	0.756 (σ=0.557)
Leave-k-out Regression *	3.454 (σ=2.501)	92.40 (σ=128.2)	135.8 (σ=131.2)	5.460 (σ=4.443)

* The Granule Test regressions are non-linear.
# Not enough data points existed to build a 30% hold back regression model.

Since the percent MAE for the Granule and Dilution Ratio Tests were large for all estimation models, we examined how the neural networks ranked different toilet designs according to performance. We found the neural networks still estimated the performance results for different toilet designs in the correct rank order for all ANSI tests. This is important because the model can help discriminate between designs effectively by accurately estimating how toilet designs will fare relative to each other. Therefore the relative inaccuracy of the performance estimations themselves for these two ANSI tests is mitigated by the complete accuracy of performance rankings, which still allow product designers to select the best design alternative.

6.0 Using the Models to Build an Interactive Design Tool

The neural networks models described in Section 5.0 were the basis of the prototype interactive design tool to see how incremental specification changes in design parameters affect product performance on a given ANSI test. Unlike a statistical model, which assumes the relationship between any given input and output is always based on the parametrics of the model (e.g. linear, log or quadratic), the neural network model is flexible across the range of input values for which it was trained. We put this ability to use by building a sensitivity model for each ANSI test by changing each design parameter in small increments across its range, holding all other parameters constant. These models can be used interactively by design engineers to test the effects of toilet design changes on the resulting product performance. In this way, designs can be optimized for performance given cost and manufacturability constraints before prototype products are built and tested.

The neural network trained by the leave-k-out strategy that best estimated held back data for each test was selected to perform sensitivity analysis. All the test results were plotted versus each normalized design parameter, resulting in a total of 24 sensitivity studies (4 ANSI tests * 6 design parameters, leaving out pan contour). To illustrate this approach, we present the results for the ANSI Granule Test in Figures 3 through 8. The number of granules that remained in the toilet after flushing was estimated by the neural networks as a function of the normalized design parameters in increments over their ranges. For instance, as water volume is increased, fewer granules will remain in the toilet, and this same trend occurs in changing the seal depth, trap size and the jet area. On the other hand, increments in the rim design has little effect on the granules estimation, while the water seal estimations suggest that the smaller this specification is, the more granules will be flushed out.

Figure 3. Sensitivity Analysis of Granule Test vs. Water Volume.

Figure 4. Sensitivity Analysis of Granule Test vs. Rim Design.

Figure 5. Sensitivity Analysis of Granule Test vs. Seal Depth.

Figure 6. Sensitivity Analysis of Granule Test vs. Water Seal.

Figure 7. Sensitivity Analysis of Granule Test vs. Trap Size.

Figure 8. Sensitivity Analysis of Granule Test vs. Jet Area.

These performance estimating models can be used to design a superior toilet in several ways. First, they show which design parameters are important in determining the performance on each ANSI test, and whether the performance improves or deteriorates as the parameter specification is made larger or smaller. This can be used as a meta-analyzer to identify the product specifications which most impact the toilet's performance over the battery of ANSI tests. The second way in which the neural network models can used during design is to refine design parameters to obtain the best mix of specifications. For example, a trap size of 0.3 increments yields the best performance on the Granule Test, however for the Ball Test, performance monotonically degrades from a trap increment of 0.0 to 0.9. For the Dilution Ratio Test, performance improves as the trap size increases from increment 0.0 to 0.9, while for the Drain Line Test, trap size is inconsequential. The designer can select a value for trap size which provides the best overall mix of performance, while considering cost, manufacturability and aesthetics. While we would recommend that this be done interactively by the designer due to the incompleteness of our model (i.e. without cost, manufacturability or aesthetics) and the multi-criteria performance tests, a neural network could be inverted as described in (Hwang et al. 1990, Hoskins et al. 1992) to analytically obtain the optimum specification of any design parameter for any given criteria. Of course, the neural network tool could be enriched by adding models which relate cost, manufacturability and aesthetics to product performance.

7.0 Conclusions and Discussion

New product design is a key to the success and survival of manufacturing enterprises. During the past few years, competitiveness has become a major stimulating agent in the design of better products. In the sanitary ware industry, the design of new products is a long and expensive process due to the design, building and testing of product prototypes. Hence, the need for new techniques that will help shorten product design cycle time, improve productivity, and produce better quality products.

The neural network model, described and used in this research, is a promising technique for improving product design in different manufacturing industries. Designers can interactively evaluate product design in terms of both absolute and relative quality performance without building mock-ups. Neural network models do not require analytical relationships or parametric assumptions, and can handle multiple qualitative and quantitative variables simultaneously. The sanitary ware manufacturing arena was a tangible example of using this approach, and we believe it can also be applied to other manufacturing sectors using different product design parameters and different performance evaluation criteria.

Acknowledgments

The authors would like to acknowledge the assistance of James E. Thorne, Manager - Research and Development and Quality Assurance of the manufacturer who kindly supplied the design parameter and performance data, and the helpful comments of two anonymous referees.

References

Andersen, Kristinn, George E. Cook, Gabor Karsai and Kumar Ramaswamy, 1990, Artificial neural networks applied to arc welding process modeling and control. *IEEE Transactions on Industry Applications*, 26, 824-830.

ASME, 1990, *Vitreous China Plumbing Fixtures* (ASME A112.19.2M-1990).

Baffes, Paul T., 1989, *NETS User's Guide, Version 2.0* (Lyndon B. Johnson Space Center).

Cariapa, Vikram, Kunter S. Akbay and Rajesh Rudraraju, 1991, Application of neural networks for compliant tool polishing operation. *Computer-Aided Production Engineering* (V. Venkatesh and J. McGeough, Editors) (Amsterdam: Elsevier Science Publishers) 271-280.

Chryssolouris, G., M. Lee, J. Pierce and M. Domroese, 1989, Use of neural networks for the design of manufacturing systems. *Proceedings of the Winter Annual Meeting of the American Society of Mechanical Engineers*, 57-63.

Consumer Reports, 1990, How to save water. *Consumer Reports*, 465-466.

Ho, Tuan T., Hai T. Ho and Jan T. Bialasiewicz, 1991, Stochastic neural adaptive control for nonlinear time varying systems. *Intelligent Engineering Systems Through Artificial Neural Networks* (C. Dagli, S. Kumara and Y. Shin, Editors) (New York: ASME Press) 531-536.

Hoskins, D. A., J. N. Hwang and J. Vagners, 1992, Iterative inversion of neural networks and its application to adaptive control. *IEEE Transactions on Neural Networks*, 3, 292-301.

Hung, S. L. and H. Adeli, 1991, A model of perceptron learning with a hidden layer for engineering design. *Neurocomputing*, 3, 3-14.

Hwang, Jenq-Neng, Chi-Hou Chan and Robert J. Marks II, 1990, Frequency selective surface design based on iterative inversion of neural networks. *Proceedings of the International Joint Conference on Neural Networks*, I-39-44.

Kinney, B. H., 1977, *The Eljer Caster Manual* (Eljer Industries) 34-37.

Lambert, Jean-Michel and Robert Hecht-Nielsen, 1991, Application of feedforward and recurrent neural networks to chemical plant predictive modeling. *Proceedings of the International Joint Conference on Neural Networks*, I-373-378.

Lawrence, Jeanette, 1991, Data preparation for a neural network. *AI Expert*, 6, 34-41.

Liu, Lihsin and G. H. Liu, 1990, Neural network computing applied in design for manufacturability. *Proceedings of the IEEE/SEMI International Semiconductor Manufacturing Science Symposium*, 107-110.

Muller, B. and J. Reinhardt, 1990, *Neural Networks: An Introduction* (Berlin: Springer-Verlag), 62-70.

Okafor, A. Chukwujekwu, Michael Marcus and Ramesh Tipirneni, 1990, Multiple sensor integration via neural networks for estimating surface roughness and bore tolerance in circular end milling. *1990 Transactions of NAMRI/SME*, 128-137.

Rumelhart, David E., J. L. McClelland and the PDP Research Group, 1986, *Parallel Distributed Processing: Exploration in the Microstructure of Cognition, Volume 1* (Cambridge, MA: MIT Press), 318-362.

Schmerr, Lester W., Stephen M. Nugen and Babak Forourachi, 1991, Planning robust design experiments using neural networks and Taguchi methods. *Intelligent Engineering Systems Through Artificial Neural Networks* (C. Dagli, S. Kumara and Y. Shin, Editors) (New York: ASME Press) 829-834.

Sejnowski, T. J. and C. R. Rosenberg, 1987, Parallel networks that learn to pronounce English text. *Complex Systems*, 1, 145-168.

Smith, Alice E., 1991, Quality control using backpropagation: An injection molding application, in *Intelligent Engineering Systems Through Artificial Neural Networks* (C. Dagli, S. Kumara and Y. Shin, Editors) (New York: ASME Press) 729-734.

Steck, J., K. Krishnamurthy, B. McMillin and G. Leininger, 1991, Neural modeling and control of a distillation column. *Proceedings of the 1991 Joint Conference on Neural Networks,* II-771-774.

Werbos, Paul J., 1988, Generalization of backpropagation with application to a recurrent gas market model. *Neural Networks*, 1, 339-356.

Zarefar, Hormoz and John R. Goulding, 1992, Neural networks in design of products: a case study, in *Intelligent Design and Manufacturing* (A. Kusiak, Editor) (New York: John Wiley & Sons) 179-201.