Neural Network Cost Estimating Relationships

Presented
At the
2010 ISPA/SCEA Conference and Training Workshop

by
Edwin B. Dean
Consultant
designforvalue@att.net

Presentation Organization

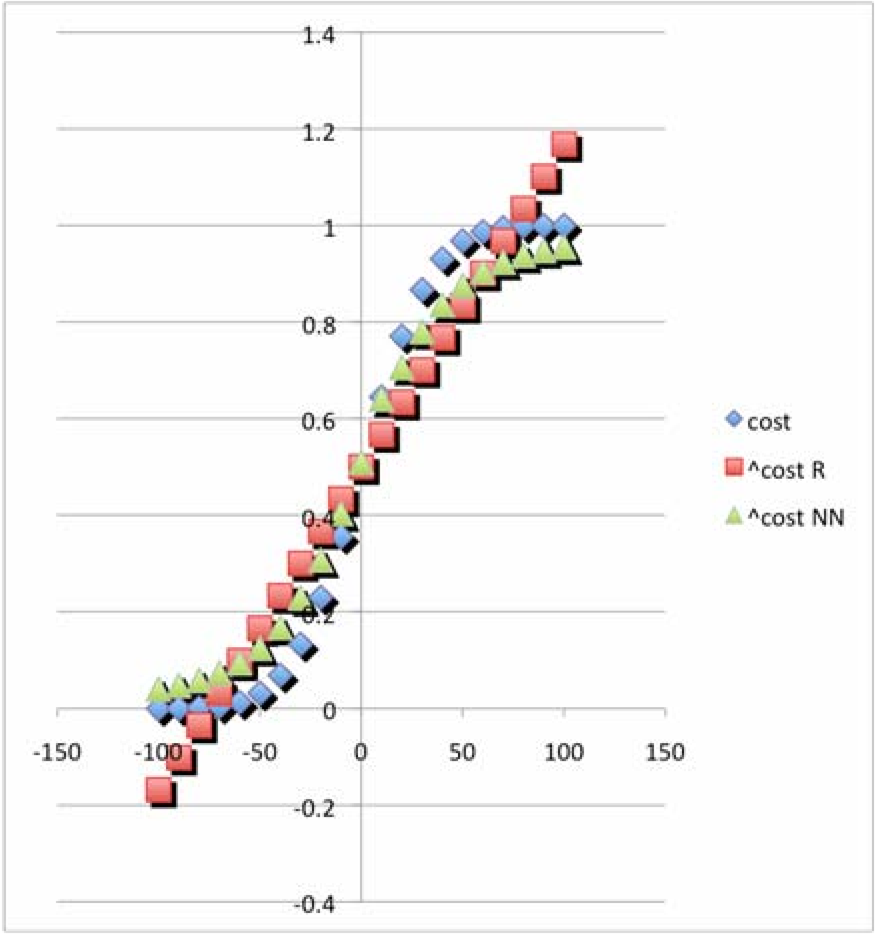
- This presentation is a road map for using neural networks for estimating cost
- Neural networks provide an alternate means of developing cost estimating relationships (CERs)
- Software exists that allows you to create a neural network CER (NNCER) without any knowledge of mathematics
  - But an understanding of the mathematics certainly helps
    - Lippmann (1987)
    - Smith (1993)
- The presentation will address
  - Pointers to prior use of neural networks for cost analysis
  - The basics of neural networks
  - Using neural networks to develop NNCERs
  - Comparing the goodness of fit with other types of models
  - Developing adaptive NNCERs
  - Performing cost risk with NNCERs

Prior Application of Neural Networks for Cost Analysis

- The primary publications of neural networks applications for cost analysis have occurred within the Association for the Advancement of Cost Engineering International (AACEI) community
- When one searches on the AACEI library using the keyword “neural network”, twenty abstracts appear
- Dean (2009)
  - Summarizes early applications of neural networks for parametric cost analysis
  - Provides two examples of neural network cost estimating relationships, and
  - Points to a number of resources for learning about neural networks

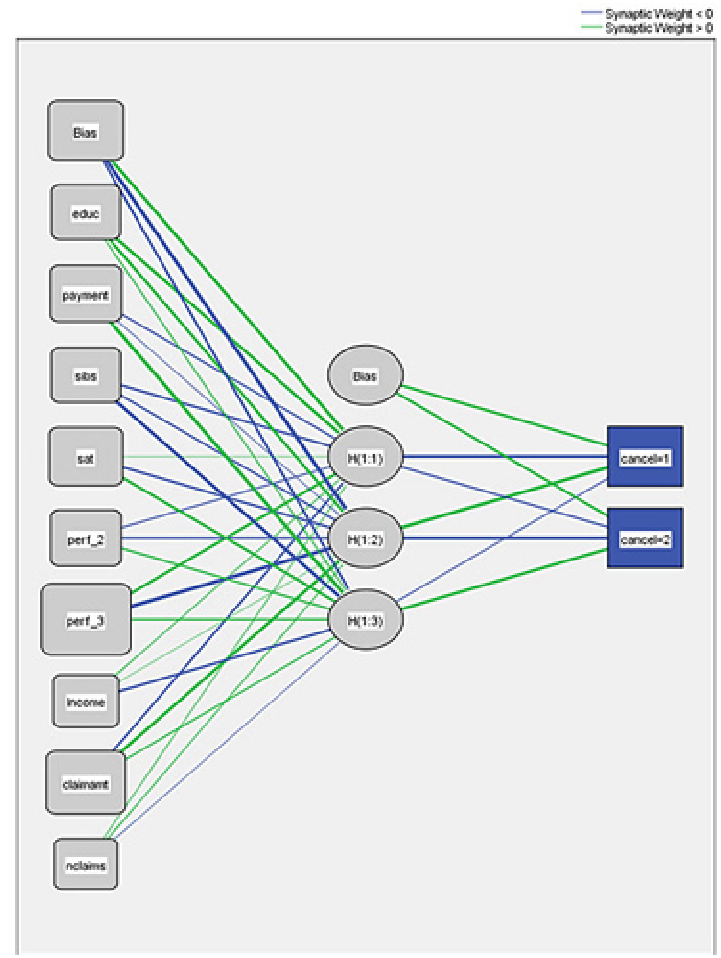
Neural Networks

- A neural network is a mathematical entity that simulates the learning capability of the brain
- It learns by approximating a set of outputs given a set of inputs
- There are many types of neural networks
- This paper only addresses the backpropagation network by
  - Werbos (1974) and
  - Rumelhart, Hinton, and Williams (1986)
- The result of the learning process is a model that predicts the desired outputs based upon a set of (input, output) data



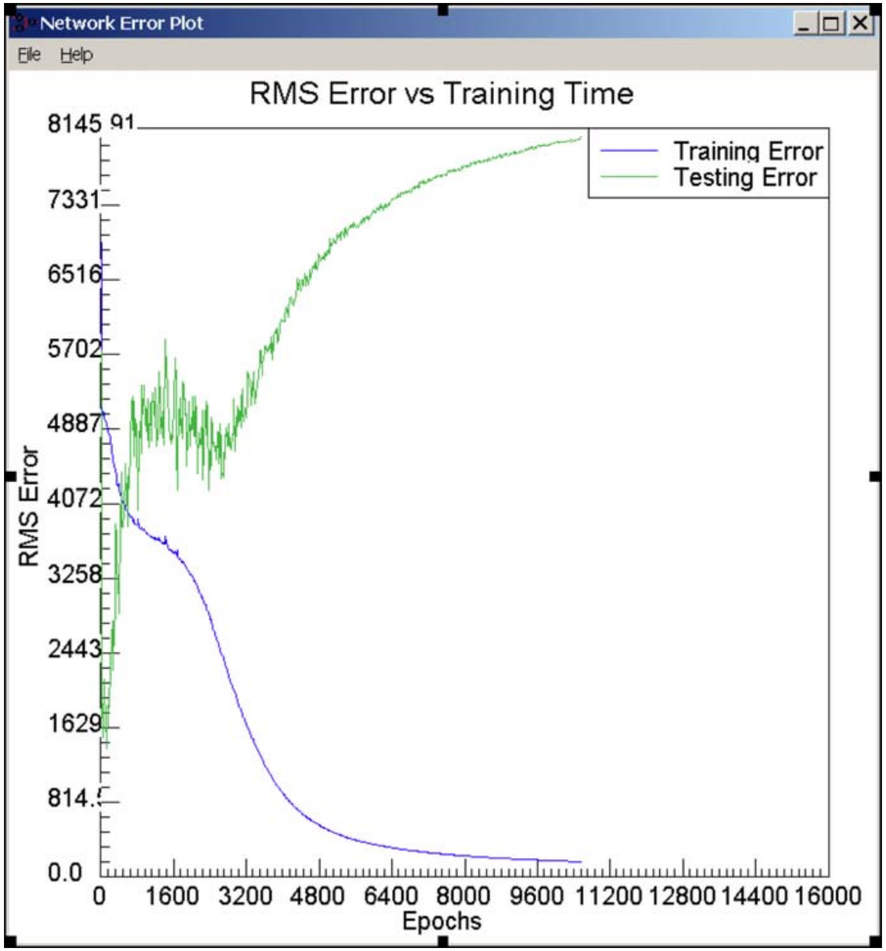
Neural Network Architecture

- A neural network has
  - an input layer
  - one or more hidden layers
  - an output layer
- Each layer contains one or more artificial neurons
- For each hidden and output layer artificial neuron, the learning process adjusts
  - the input weights and
  - the threshold weight
  - to reduce the root mean squared output error
- The nonlinear activation function enhances the approximation capability



Training the Neural Network

- A single application of all of the data is called an epoch
- The training data set is used to train the neural network
- The test data set is used to see how well the neural network generalizes to another independent set of data
- The test data set is also used to determine the stopping point for the training
- Epochs are applied repeatedly until the inputs provide a reasonable approximation of the outputs
- Data for following examples is from the JSC 1994 Advanced Mission Model (Econ and Cyr, 1994)
- Inputs (parameters) for following examples are
  - Empty weight
  - Payload weight
  - IOC date
  - R&D quantity
  - Production quantity

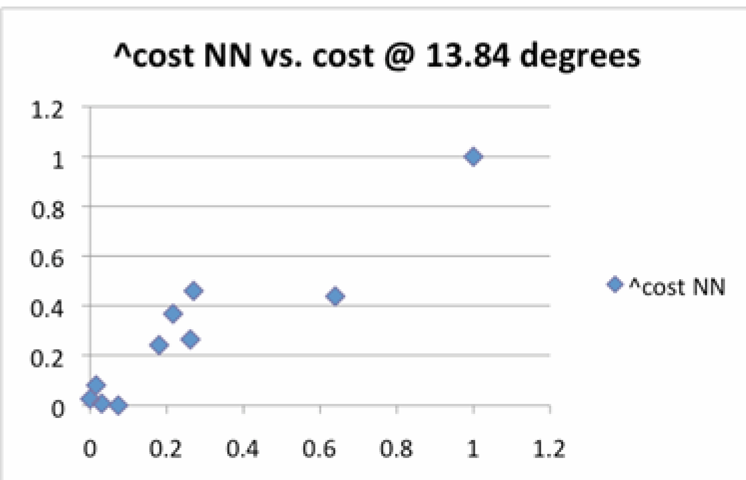


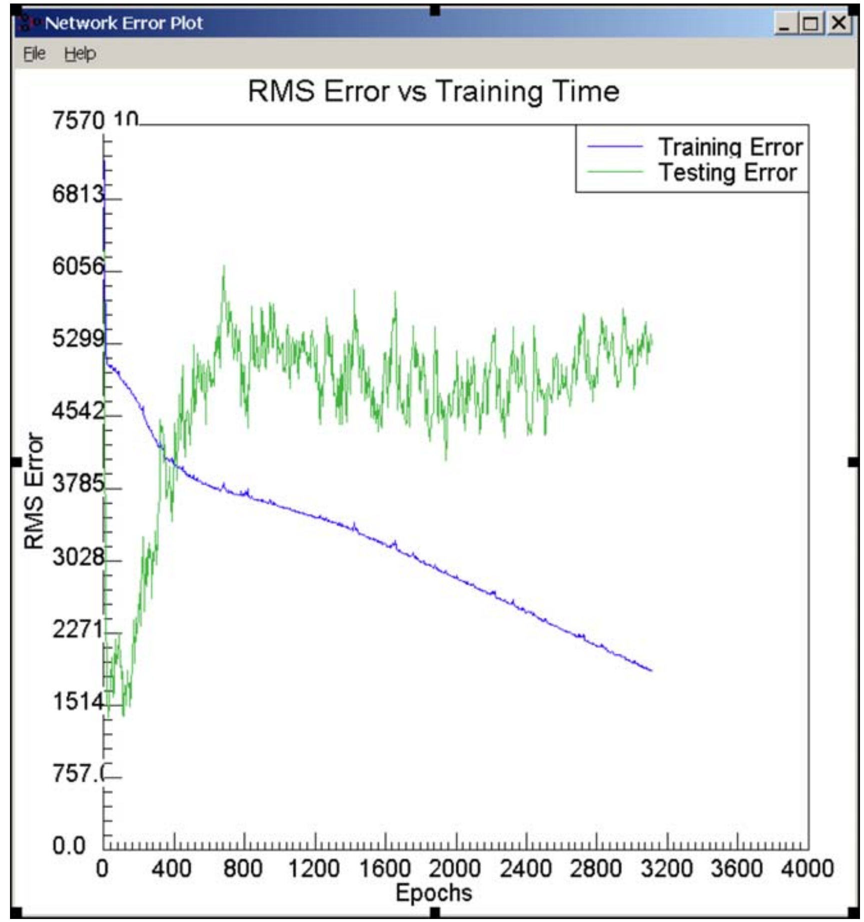
Goodness of Fit

- R^2 is used as the goodness of fit measure for regression models
- Root mean square error is the goodness of fit measure for the backpropagation neural network
- A goodness of fit measure for different types of predictive models (Dean, 2008) is the angle between
  - the predicted data vector and
  - the actual data vector
  - which is
    - distribution-independent and
    - method-independent
- The smaller the angle the better the fit
- Based upon experience, the fit is
  - Excellent for angles between 0 and 5 degrees
  - Good for angles between 5 and 10 degrees
  - So So for angles between 10 and 15 degrees
  - Poor for angles above 15 degrees

Generalization Example

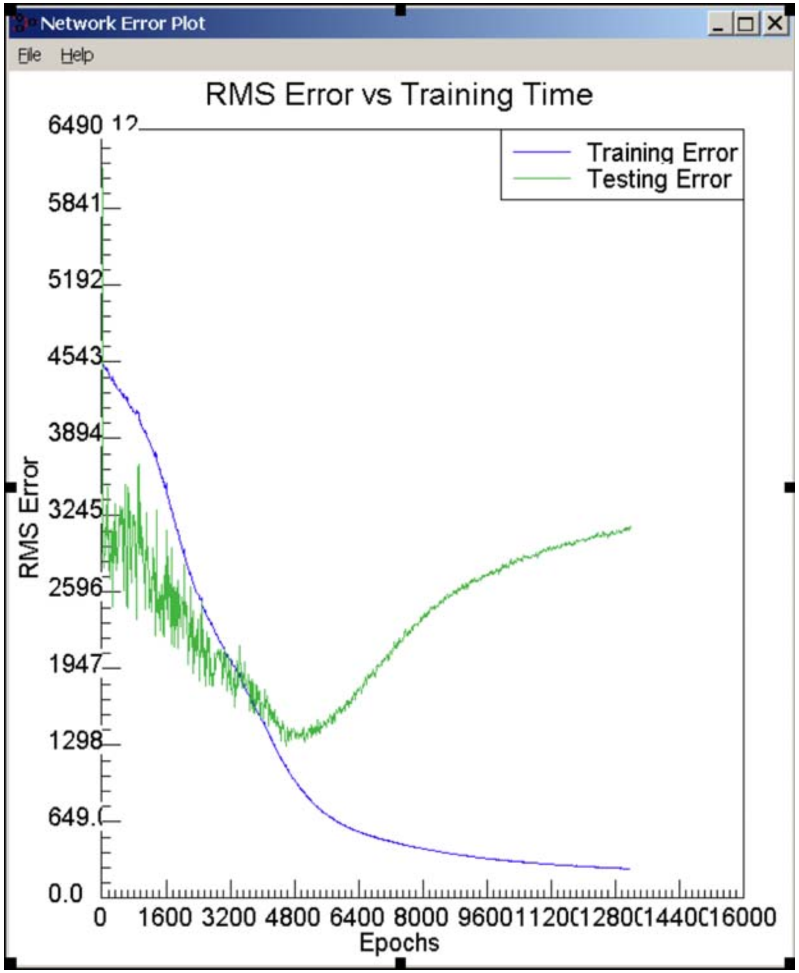
- All surface-to-air missiles in training set
- All ship-to-air missiles in test set
- Training angle = 13.84 degrees
- Test angle = 8.26 degrees
- Ship-to-air angle =8.26 degrees





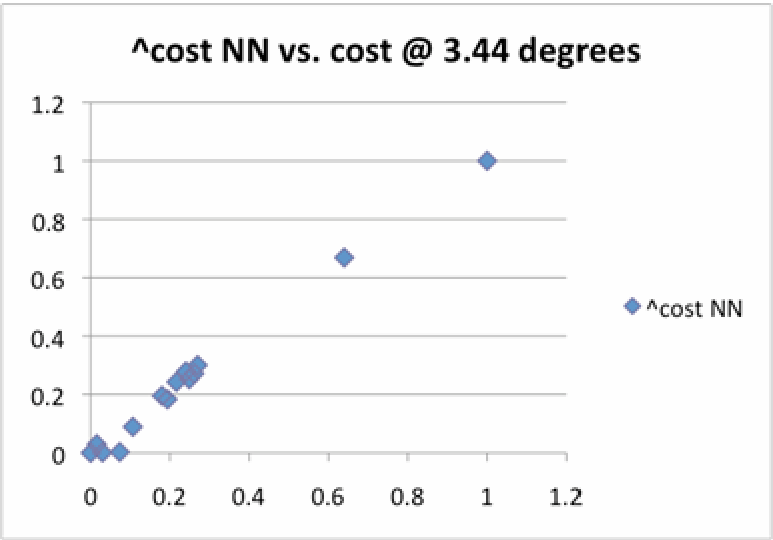
Adaptive CERs

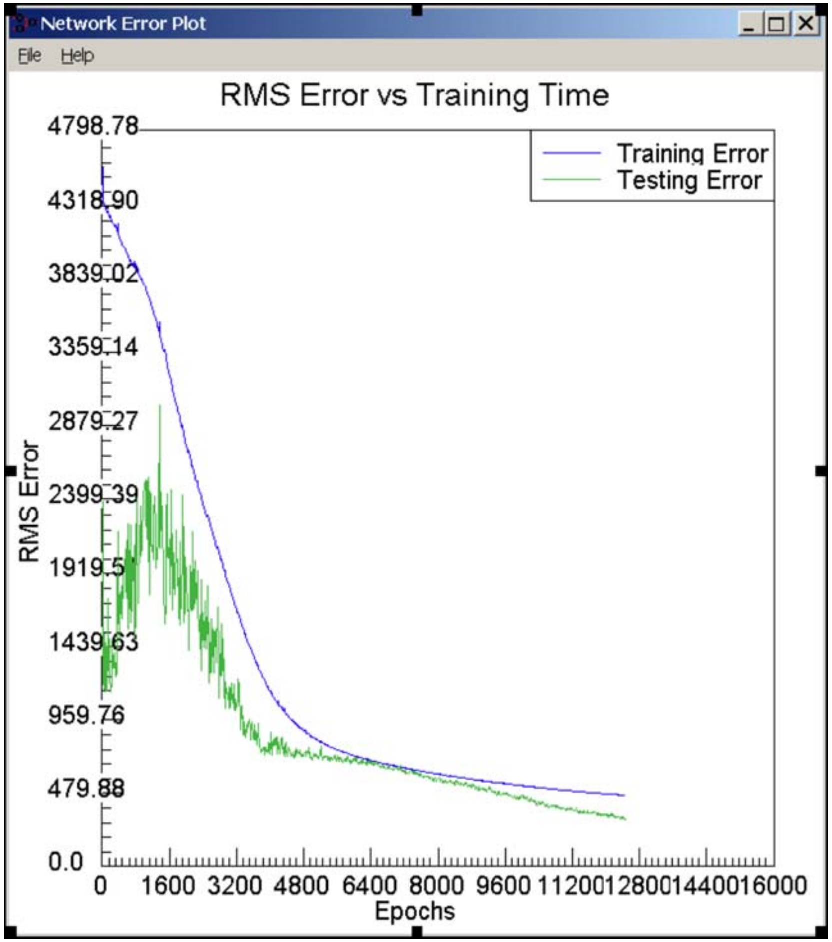
- An adaptive CER is one that ensures that a set of data points carries more weight in the model than other data points (Book, Broder, and Feldman, 2009)
- The neural network training process provides a simple way to develop an adaptive CER
- Instead of using a random set of data points in the test set, place the points for which you desire the most weight in the test data set
- Use the test data set to stop the training as the test data error curve turns up
- At that point the neural network has the least error for the data points desired to have the most weight in the manifold fitting process



Adaptive CER Example

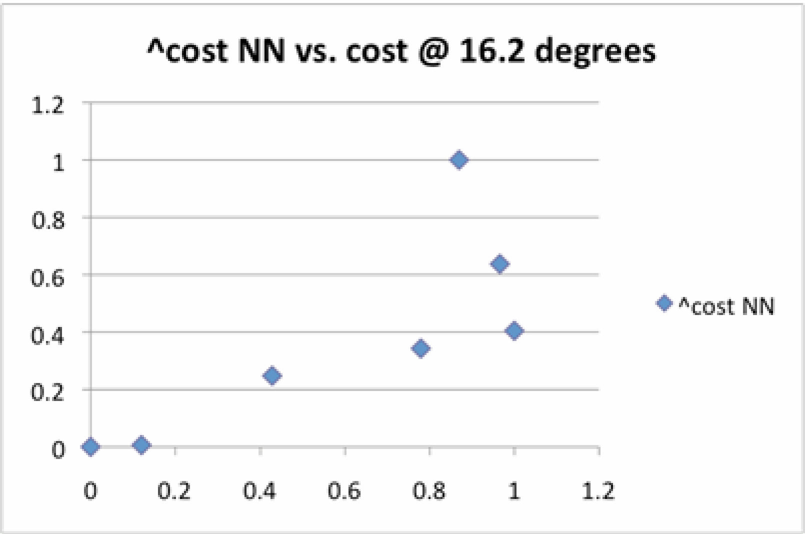
- All surface-to-air and ship-to-air missiles in training set
- All ship-to-air missiles in test set
- Training angle = 3.44 degrees
- Test angle = 3.71 degrees
- Ship-to-air angle =3.71 degrees

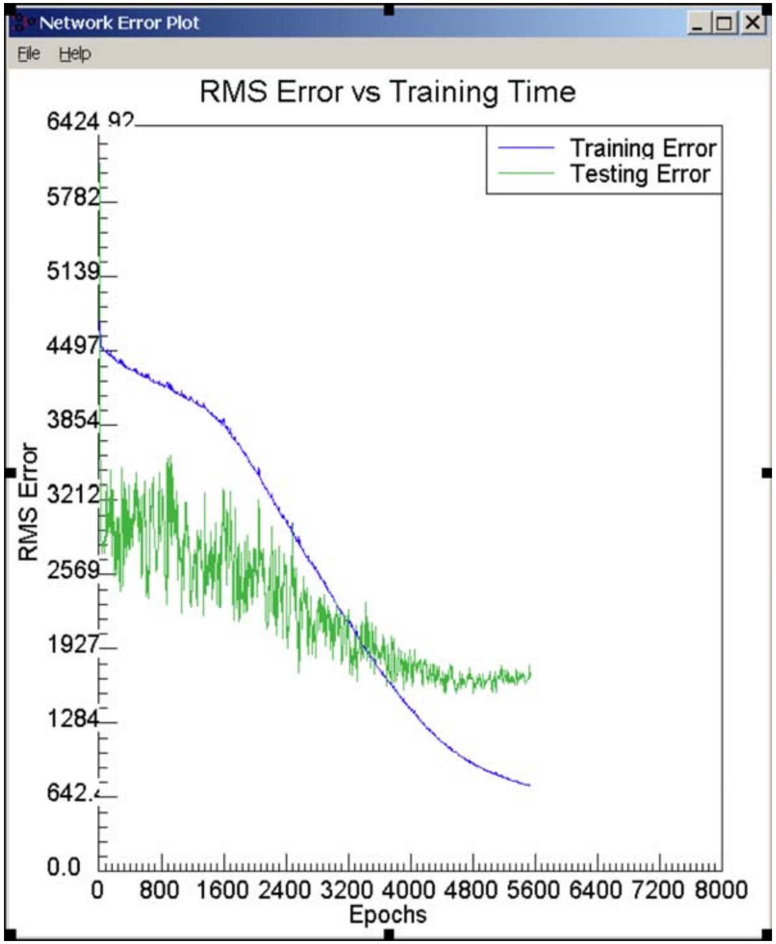




Adaptive CER with Generalization Example

- Some surface-to air-missiles and some ship-to-air missiles in training set
- Rest of surface-to-air and all ship-to-air missiles in test set
- Training angle = 5.00 degrees
- Test angle = 16.2 degrees
- Ship-to-air angle =8.14 degrees





Cost Risk Using Neural Networks

- Unfortunately, my understanding is that statistics do not yet exist analogous to regression
  - This eliminates using the uncertainty associated with the development of the CER itself
  - This is a research topic that could provide one or more excellent papers
    - Hint: the cosine of the angle is the equivalent of R for non mean-adjusted data
- The void of statistics mandates that one must use what I call push cost risk
  - In push cost risk, distributions are placed on the input parameters and are fed as inputs to the trained neural network
- When subsystems are involved, one must use push cost risk with a network of networks
  - Each subsystem has a separately trained network that is input
    - to the appropriate subsystem output summing junction or
    - to the system output summing junction
- When activities or processes are involved, one must use push cost risk with a network of networks
  - Each activity has a separately trained network that is placed at the appropriate location in a PERT/CPM type network
  - Typically, a process is a network of activities. However, a process may have a trained network
- Risk register distributions
  - may be used as training inputs or
  - may be attached to any summing junction

Implementing Neural Network Cost Risk

- Good neural network software permits access to the weights of the trained neural network
  - Some neural network software provides C or C++ code as well as weights
  - QuikNet v2.23 by Jensen (2003) is an easy to use and inexpensive neural network software shareware download for Windows
    - It works well under Windows 2000 and Windows XP
    - It should work well under Windows 7 in XP compatibility mode
- The cost risk network of networks can be implemented
  - in a spreadsheet or
  - as code
- Regression based or other types of CERs can be integrated into the cost risk network of networks
  - For other types of CERs read Meisl (1989)

Adding Duration to Cost

- Note that a neural network may have multiple outputs
- A second output on the neural network may be used for a duration output
- If so, then cost and duration may be trained simultaneously and used simultaneously for
  - Estimating and
  - Risk

Summary

- This paper demonstrates that neural network CERs are a practical alternative to regression based means of developing CERs
- They provide a non assumption based CER derived from data
- The training process
  - is easy to use
  - provides a means of developing adaptive CERs
  - provides a means of training and simulating
    - Cost risk
    - Cost and duration risk
- A roadmap has been provided for the use of neural network CERs within the parametric cost estimating process

Resources

- AACEI. Association for the Advancement of Cost Engineering International, http://www.aacei.org/.
- Book. S., M. Broder, and D. Feldman (2009). "Statistical Foundations of Adaptive Cost-Estimating Relationships," Proceedings of the 2009 ISPA/SCEA Professional Development and Training Workshop, 2-5 June, Saint Louis, MO.
- Dean, E. B. (2008). "The Angle: A Goodness of Fit Metric for Cost Analysis", Proceedings of the ISPA/SCEA 2008 Joint International Conference, 12-16 May, Noordwijk, The Netherlands.
- Dean, E. B. (2009). "Parametric Cost Analysis Using Neural Networks", Proceedings of the ISPA/SCEA 2009 Professional Development and Training Workshop, 2-5 June 2009, Saint Louis, MO.
- Econ and K. Cyr (1994). Advanced Mission Cost Model, http://cost.jsc.nasa.gov/AMCM.html.
- Jensen, C. (2003). QuikNet v2.23, http://qwiknet.home.comcast.net/~qwiknet/.
- Lippmann, R. P. (1987). "An Introduction to Computing with Neural Nets", IEEE ASSP Magazine, pp. 4-22.
- Meisl, C. J. (1989). "Advanced Life Cycle Cost Modeling", *Journal of Parametrics*, Vol. 9, No. 1, January, pp.74-102.
- Rumelhart, D. E, G. E. Hinton, and R. J. Williams (1986). "Learning Representations by Back-Propagating Errors", *Nature*, 323, pp. 533-536.
- Smith, M. (1993). Neural Networks for Statistical Modeling, Van Nostrand Reinhold, New York.
- Werbos, P, J. (1974). Beyond Regression: New Tools for Prediction and Analysis in The Behavioral Sciences, PhD Dissertation, Harvard University, Cambridge, MA.