Published in the ***Proceedings of the Maintenance and Reliability Conference***, Knoxville, TN, May 1998

System Modeling and Instrument Calibration Verification with a Nonlinear State Estimation Technique

Christopher L. Black*, Robert E. Uhrig**, and J. Wesley Hines
Department of Nuclear Engineering
University of Tennessee
Knoxville, TN, 37996

* Now with DRYKEN Technologies, Inc., Oak Ridge, TN 37830
** Also with I&C Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

ABSTRACT

A nonlinear state estimation technique (NSET) has been developed to perform process modeling in a sensor and associated instrument channel calibration verification system. The model estimates the true process values, as functional sensors would provide them. The residuals between these estimates and the actual measurements (from sensors of unknown condition) are monitored using the sequential probability ratio test, a statistical decision method.

The NSET is a generalization of the multivariate state estimation technique (MSET) described by Singer, et al. [1]. The MSET itself is an extension of the least-squares minimization of the multiple regression equation, incorporating proprietary comparison operators. The theoretical introduction to these estimation techniques is provided, and the NSET is demonstrated with several different comparison (either similarity or distance) operators. The estimation performance of the NSET for each of these operators is evaluated, resulting in a recommended operator for general process modeling purposes. The NSET is demonstrated as the modeling engine for a calibration verification system.

INTRODUCTION

The NSET, MSET, and the multiple regression solution all utilize similarity operators to compare new measurements to a set of prototypical measurements or states. This comparison process generates a weight vector that is used to calculate a weighted sum of the prototype vectors. The sum is weighted to provide an estimate of the true process values.

For the similarity function, multiple regression utilizes the dot product, the MSET uses one of two proprietary similarity/distance operators [2], and the NSET can incorporate one of several possible operators, many of which are presented, demonstrated and evaluated in this paper.

The NSET functions as an autoassociative model, reproducing an estimate of each of a set of measured signals that are provided as inputs to the model. The training is single-pass (i.e., is not an iterative process), and consists of little more than the operations involved in a single matrix multiplication and an inversion or decomposition. Data selection plays an important role, as the number of operations (and processor time) required per recall is proportional to the product of the number of prototype measurements and the dimensionality of the measurements. Therefore, as a

function of the number of signals to be monitored, and the data availability rate, there is an upper limit upon the number of patterns that may be included in the prototype measurement matrix. As the purpose of the prototype matrix is to compactly represent the entire dynamic range of previously observed system states, the patterns that are included must be carefully chosen.

MULTIPLE REGRESSION

Let $\underline{\underline{A}}$, referred to as the prototype matrix, represent a matrix assembled from selected column-wise measurement vectors, and let $\underline{w}$ represent a vector of weights for averaging $\underline{\underline{A}}$ to provide the estimated state $\underline{y'}$ as follows:

$$\underline{y'} = \underline{\underline{A}} \cdot \underline{w} \qquad \text{(Eqn. 1)}$$

The column-wise measurement vectors which make up the prototype matrix $\underline{\underline{A}}$ are usually selected by a clustering data analysis technique. This selection is carefully performed to provide a compact, yet representative, subset of a large database of measurements spanning the full dynamic range of the system of interest. An example of a prototype matrix, constructed from the vertices of a unit cube is given:

$$\underline{\underline{A}} = \left[\underline{x}(1) \quad \underline{x}(2) \quad \ldots \quad \underline{x}(n)\right] = E.g., \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \qquad \text{(Eqn. 2)}$$

If $\underline{\varepsilon}$ represents the difference between an observed state $\underline{y}$ and the estimated state $\underline{y'}$, then the following relations may be constructed:

$$\underline{e} = \underline{y} - \underline{y'} = \underline{y} - \underline{\underline{A}} \cdot \underline{w} \qquad \text{(Eqn. 3)}$$

The least squares solution to the minimization of $\underline{\varepsilon}$ yields the following expression for $\underline{w}$, (where the left hand factor of the matrix product is known as the recognition matrix):

$$\underline{w} = \left(\underline{\underline{A}}^T \cdot \underline{\underline{A}}\right)^{-1} \cdot \left(\underline{\underline{A}}^T \cdot \underline{y}\right) \qquad \text{(Eqn. 4)}$$

A chief liability of this method is that linear interrelationships between state vectors in $\underline{\underline{A}}$ result in conditioning difficulties associated with the inversion of the recognition matrix. This shortcoming is avoided by the NSET by applying nonlinear operators, as shown later, in lieu of the matrix multiplication. These operators generally result in better conditioned recognition matrices and more meaningful inverses of the recognition matrices.

The conditioning difficulty may also be ameliorated, as suggested by Masters [3], by solving the normal equations for the weight vector. Solving the normal equations involves a decomposition of the recognition matrix and then elimination or back-substitution.

$$\left(\underline{\underline{A}}^T \cdot \underline{\underline{A}}\right) \cdot \underline{w} = \underline{\underline{A}}^T \cdot \underline{y} \qquad \text{(Eqn. 5)}$$

Although the decomposition of the recognition matrix must only be performed once, up-front in a learning phase, the solution of the normal equations requires more operations per vector (than when simply using the matrix inverse) in the deployment or recall phase.

Unfortunately, the normal equations are not themselves immune to numerical instabilities, as zero or near zero pivot elements, due to correlation among prototype measurements, may be encountered during solution (elimination) of the normal equations. Zero and near-zero pivot elements result in, respectively, no solution, or one in which some of the fitted parameters (weights) have large magnitudes that delicately cancel one another out [4]. Singular value decomposition (SVD) is a particularly stable (with regard to rank deficiencies) method of matrix decomposition for finding the pseudoinverse of a matrix [4,5]. Prototype vectors whose combinations are irrelevant to the fit are assigned small magnitude weights, resulting in a fit that handles both overdetermined-ness (minimizes squared error) and underdetermined-ness (minimizing effect of correlation among prototypes) and as a result is numerically very dependable.

For this investigation of the NSET, the SVD algorithm is used to provide a pseudo-inverse of the recognition matrix, which is then employed in the manner of the true inverse of equation 4. This pseudo-inversion is performed in a training phase prior to deployment of the model, and is considered analogous to the up-front training of an artificial neural network (ANN) model.

Another cited drawback [2] (also necessitating the use of different similarity operators) to the least squares solution of the multiple regression equation is that perturbations outside the range observed in the prototype matrix in a single element of the observation vector result in large estimation errors. Whether or not this quality is truly a disadvantage is open to debate, as a conservative model might be expected to fail dramatically, as opposed to giving reasonable looking, but inaccurate, results. In other words, it is questionable whether any propensity to always supply a reasonable appearing, though possibly incorrect result should be considered a positive trait. This is an issue implicit in any empirically derived modeling application in which the general form of the relationship is not known or presumed. Extrapolation features of the NSET (or of data-derived models in general) are not investigated in this paper.

NSET: EXTENDING THE MULTIPLE REGRESSION EQUATIONS

NSET (as does MSET) springs forth from the multiple regression result as follows:

$$\underline{w} = \left(\underline{\underline{A}}^T \oplus \underline{\underline{A}}\right)^{-1}\left(\underline{\underline{A}}^T \oplus \underline{y}\right) \qquad \text{(Eqn. 6)}$$

or in normal equation form,

$$\left(\underline{\underline{A}}^T \oplus \underline{\underline{A}}\right)\underline{w} = \underline{\underline{A}}^T \oplus \underline{y}\ . \qquad \text{(Eqn. 7)}$$

The '⊕' symbol above represents any appropriate similarity or difference operator applied upon matrices. The definitions of a few of the many possible candidate operators provided below are all scalar-valued functions on vectors. The vector functions are easily extended to the required matrix functions by selecting vectors from the two input matrices and positions in the output matrix in exactly the familiar row and column manner used in performing individual vector dot products in matrix multiplication.

CANDIDATE DISTANCE/SIMILARITY OPERATORS

1. Bernoulli difference (BDIF) {requires y ∈ (0,1)}:

$$f(\underline{x},\underline{y}) = -\frac{1}{M\log(2)}\sum_{m=1}^{M}\left[x_m \log\left(y_m\right) + \left(1-x_m\right)\log\left(1-y_m\right)\right] \quad \text{(Eqn. 8)}$$

2. Relative entropy (RENT) {requires y>0}:

$$f(\underline{x},\underline{y}) = -\frac{1}{M\log(M)}\sum_{m=1}^{M} x_m \log y_m \quad \text{(Eqn. 9)}$$

3. Euclidean norm (DIST):

$$f(\underline{x},\underline{y}) = \sqrt{\sum_{m=1}^{M}\left(x_m - y_m\right)^2} \quad \text{(Eqn. 10)}$$

4. City block distance (CITY):

$$f(\underline{x},\underline{y}) = \sum_{m=1}^{M}\left|x_m - y_m\right| \quad \text{(Eqn. 11)}$$

5. Linear correlation coefficient (LCC):

$$f(\underline{x},\underline{y}) = \frac{\sum_{m=1}^{M}\left(x_m - \bar{x}\right)\left(y_m - \bar{y}\right)}{\sqrt{\sum_{m=1}^{M}\left(x_m - \bar{x}\right)^2\sum_{m=1}^{M}\left(y_m - \bar{y}\right)^2}} \quad \text{(Eqn. 12)}$$

6. Common mean linear correlation coefficient (CMLCC):

$$f(\underline{x},\underline{y}) = \frac{\sum_{m=1}^{M}\left(x_m - \bar{x}\right)\left(y_m - \bar{x}\right)}{\sqrt{\sum_{m=1}^{M}\left(x_m - \bar{x}\right)^2\sum_{m=1}^{M}\left(y_m - \bar{x}\right)^2}} \quad \text{(Eqn. 13)}$$

7. Root mean power error (RMPE):

$$f(\underline{x},\underline{y},m) = \sqrt[m]{\frac{1}{M}\sum_{m=1}^{M}\left(x_m - y_m\right)^m} \quad \text{(Eqn. 14)}$$

8. Scaled mean power error (SMPE):

$$f(\underline{x},\underline{y},m) = \frac{1}{mM}\sum_{m=1}^{M}\left(x_m - y_m\right)^m \quad \text{(Eqn. 15)}$$

9. Matrix multiplication (MMLT):

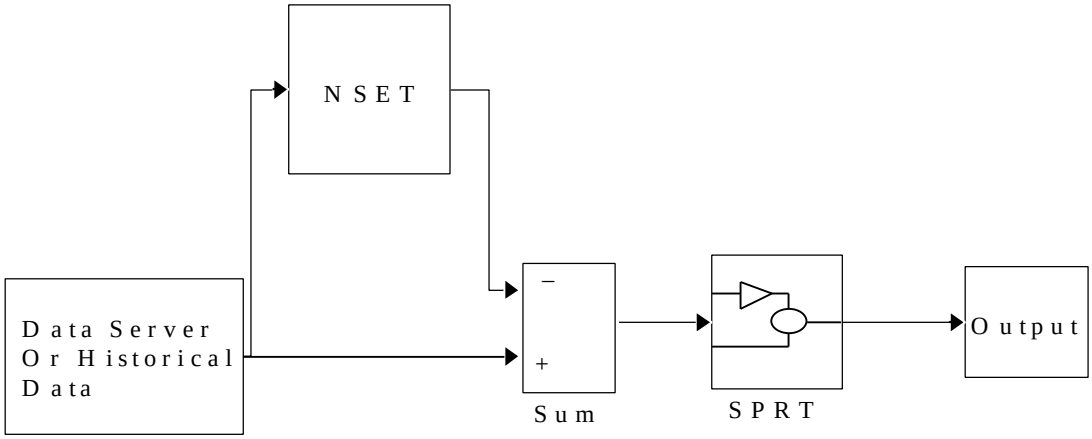
$$f(\underline{x}, \underline{y}) = \sum_{m=1}^{M} x_m \cdot y_m \qquad \text{(Eqn. 16)}$$

10-18. Versions of candidate operators 1-9, 'inverted' so that large similarities correspond to small distances and small similarities correspond to large distances, as concisely follows:

$$f(\underline{x}, \underline{y}) = \frac{1}{1 + f'(\underline{x}, \underline{y})} \qquad \text{(Eqn.17)}$$

INSTRUMENT CALIBRATION VERIFICATION SYSTEM

The instrument calibration verification system (ICVS) consists of three main components. The first component denoted "Data Server or Historical Data" in Figure 1, either obtains real-time samples of the measurements provided by a server program on a plant computer, or provides previously acquired and stored samples. The measurements are then provided to the second component, the NSET model, which produces the prediction (estimate) of the true process value. The final component is the sequential probability ratio test (SPRT) module. The SPRT, a statistical decision technique developed by Wald [8], evaluates the difference or residual obtained between measurement and model prediction, and determines when drift or failure has occurred. The SPRT has previously been applied to detection in a neural network model based ICVSs [7,9-11] and in signal monitoring applications incorporating other modeling techniques [1,2,12 and 13]. A full description of its role in the ICVS is not provided here, but may be found in detail elsewhere [11].



DEMONSTRATION OF APPLICATION OF NSET TO MODELING OF SELECTED HIGH FLUX ISOTOPE REACTOR (HFIR) VARIABLES

For the purposes of demonstrating system modeling by NSET (as well as characterizing the relative performance of the candidate operators relative to one another), a datafile was created containing process measurements from the Oak Ridge National Laboratory's (ORNL's) High Flux Isotope Reactor (HFIR). This datafile spans the research reactor's entire fuel cycle, and contains 1636 time-measurements. Each time-measurement contains 34 of the 56 analog signals provided (every 2 seconds) by the data server on the HFIR's plant computer. The 22 discarded analog

signals include a few process variables that never demonstrate any variation from a constant measurement value. Other variables, also discarded, possessed sufficient physical redundancy such that inclusion might serve to weaken a demonstration of the modeling power of an analytically redundant modeling system. Some only variables represented simple sums or averages of other measured analog quantities, and were not included.

The constructed datafile contained 1636 data vectors representing a stratified sample (every nth sample) of an entire fuel cycle (startup transient, full power operation, and shutdown transient). This data was clustered using a hyperbox clustering technique [6], so that a pre-specified number (100) of data clusters were created. The datapoint closest to the center of each cluster centroid was included in the prototype matrix.

For the prototype matrix created, the 18 candidate operators were each explored by computing the sum-squared error (SSE) between NSET model estimation and measurement of the original datafile. As previously mentioned (specifically in equations 8 and 9), some of the operators require a particular domain. In addition, demonstration of the effect of scaling upon estimation was desired. To these ends, each operator was first used and evaluated with raw (non-scaled) data. Linear scaling of the data to [0.1,0.9] was then performed. The range (the central region of the sigmoid function) of the linearly scaled data was arbitrarily chosen, as previously existing routines for artificial neural network input/output scaling were conveniently available. The estimation was repeated for each operator, and the estimates were 'de-scaled' for evaluation. Finally, this estimation was performed similarly as before with the linear scaling case, but with mean-centered unit-variance (MCUV) scaling and de-scaling. MCUV scaling results in each individual scaled time-series possessing a mean of zero and variance of one. The results of this demonstration are given below in Table 1. The columns and rows of Table 1 respectively correspond to the scaling method and the similarity/distance operator utilized.

Table 1. Test Prediction Error Results for Forward and Inverse Distance/Similarity Operators

	Forward Test SSE			**Inverse Test SSE**		
	None	**Linear**	**MCUV**	**None**	**Linear**	**MCUV**
BERR		2.52E+06			7.43E+03	
RENT		1.10E+07			6.93E+04	
DIST	6.68E+03	1.25E+04	1.45E+04	1.29E+07	4.83E+04	2.60E+05
CITY	4.85E+03	1.53E+04	1.51E+04	4.22E+07	1.54E+05	7.86E+05
LCC	1.27E+06	4.02E+06	6.16E+05	6.99E+04	3.02E+07	7.97E+07
CMLCC	3.09E+03	5.23E+05	1.54E+05	2.64E+03	1.10E+08	3.42E+08
RMPE	6.68E+03	1.25E+04	1.45E+04	7.61E+05	1.63E+04	4.17E+04
SMPE	3.58E-12	1.42E-13	1.72E-15	2.65E+06	3.27E+02	6.48E+03
MMLT	9.38E-10	5.84E-13	2.33E-16	6.33E+04	3.95E+06	2.94E+09

Table 2, following, provides the recall estimation rate on a 200 MHz Pentium PC for only the 'forward' version of each operator, as the inverse versions were found to differ by less than 1%. The operators were initially formulated such that all arithmetic was performed as a series of scalar operations upon the elements of two vectors. The 'vectorized' versions are enhanced such that the operators involve a series of vector operations upon matrices, and are thereby significantly

improved in MATLAB, an interpreted mathematical programming environment. Future recoding in C++ is expected to bring about further and more significant recall estimation rate improvements.

Table 2. Recall Estimation Rate for Various Distance/Similarity Operators

	Estimations / Second	
	Original	**Vectorized**
BERR	16.73	45.70
RENT	51.79	87.67
DIST	39.77	142.19
CITY	49.31	123.56
LCC	16.56	55.22
CMLCC	15.21	41.74
RMPE	43.50	90.16
SMPE	46.32	95.29
MMLT		551.61

Illustration of inversion difficulties is provided below in Table 3. The last three columns consist of a reciprocal condition number associated with the recognition matrix for each combination of operator and scaling combination. A value near one is associated with an easily invertible matrix, while values approaching zero are associated with near singular matrices. The first three columns of Table 3 correspond to an index created to describe the quality of the pseudo-inverse obtained from the recognition matrix. The index is obtained by pre-multiplying the recognition matrix by its pseudo-inverse. The sum square difference is then computed between the result and the same size identity matrix. Of particular interest is the fact that the result obtained, in all cases, closely approaches an integer. Recognition matrices that possess a reciprocal condition number very close to zero ($< 10^{-10}$) result in positive integer values (within several decimal places), while greater condition numbers foretell indices near zero. This is consistent with the SVD's method of disregarding more singular values when presented more poorly conditioned matrices.

Table 3. Recognition Matrix Inversion and Invertability Indicators

	Pseudo-inverse Error Index			Reciprocal Condition Number		
	None	**Linear**	**MCUV**	**None**	**Linear**	**MCUV**
BERR		65			3.18E-19	
RENT		66			6.94E-20	
DIST	0	0	0	1.60E-04	4.99E-04	4.09E-04
CITY	0	0	0	6.08E-05	1.70E-04	1.79E-04
LCC	67	67	67	9.25E-20	2.59E-19	6.48E-19
CMLCC	5	0	0	2.36E-15	7.48E-10	2.44E-08
RMPE	0	0	0	1.60E-04	4.99E-04	4.09E-04
SMPE	64	64	64	3.71E-21	5.37E-20	3.05E-21
MMLT	66	66	66	1.81E-20	1.29E-19	1.15E-20

The root mean power error (RMPE) and scaled mean power error (SMPE) operators contain a variable parameter known as the Minkowski parameter. This parameter influences the proportional effect of each component's difference upon the entire vector distance. At low

parameter values, all of the component errors affect the distance. As the Minkowsky parameter is allowed to increase, the distance decreases asymptotically towards the largest component error. As displayed in Figure 2, this parameter was varied to observe the effect upon test prediction error.

Figure 2. NSET Test Prediction Error Using RMPE Operator and Varying Minkowski Parameter

As sum-squared error is a useful comparison metric between prediction methods, but not a very informative standalone statistic, Figures 3-12 display measured and predicted individual signals over the entire test set - first with no measurement error, then with measurement error, and finally, with the diagnosis of the calibration verification system. Figure 3 displays both the measured and estimated cooling tower inlet temperature. The NSET estimate is obtained using candidate operator #3, the Euclidean distance operator. This operator was found to provide the best combination of estimation accuracy and insensitivity to measurement error.

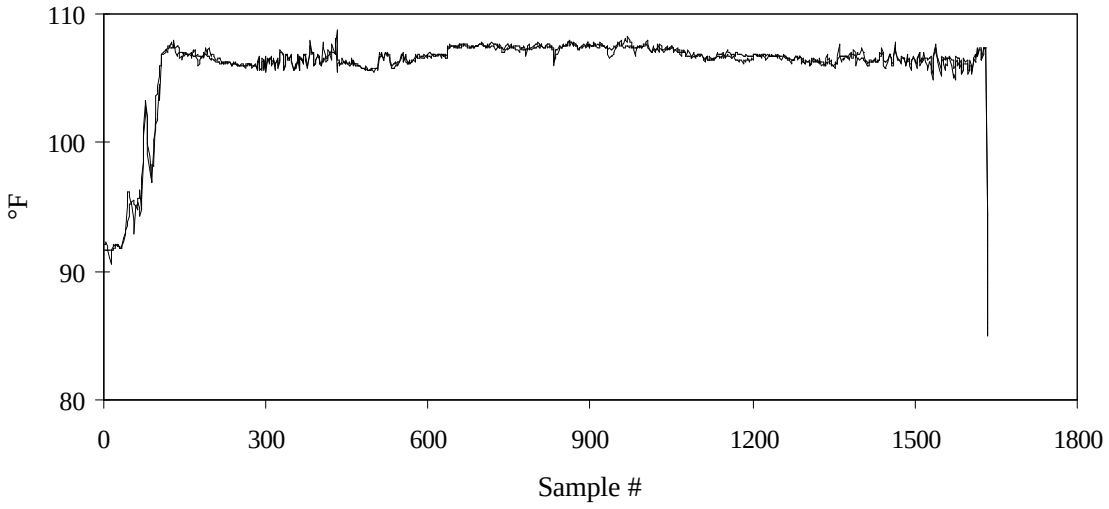


Figure 3. Estimated and Measured Cooling Tower Inlet Temperature (DIST – no Scaling)

Figure 4 contains the measured cooling tower inlet temperature, this time incorporating a synthetic linear drift of 0.002% full-scale (%FS) per sample, as well as the 'DIST operator and no scaling' version of the NSET estimate.

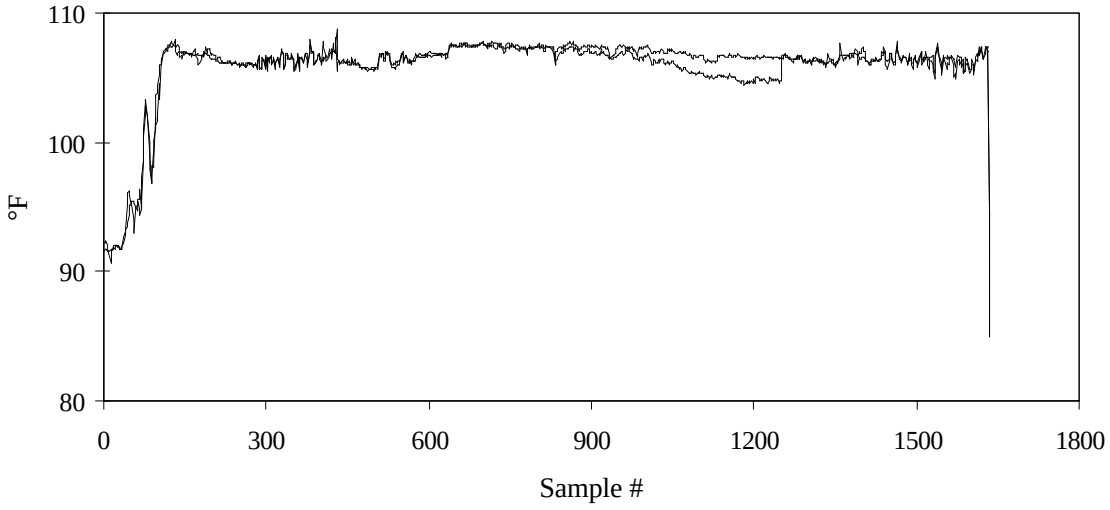


Figure 4. Drifted Measurement and Estimate of Cooling Tower Inlet Temperature (DIST – no scaling)

Note that the synthetic measurement drift present from sample 751 to sample 1250 is only negligibly present in the NSET estimate. This immunity to measurement error makes the NSET an appropriate modeling engine in a calibration monitoring system.

Figure 5 redisplays the measurement and estimate in increased detail. Figure 6 displays the diagnosis of an NSET-based ISCV system on the same scale.

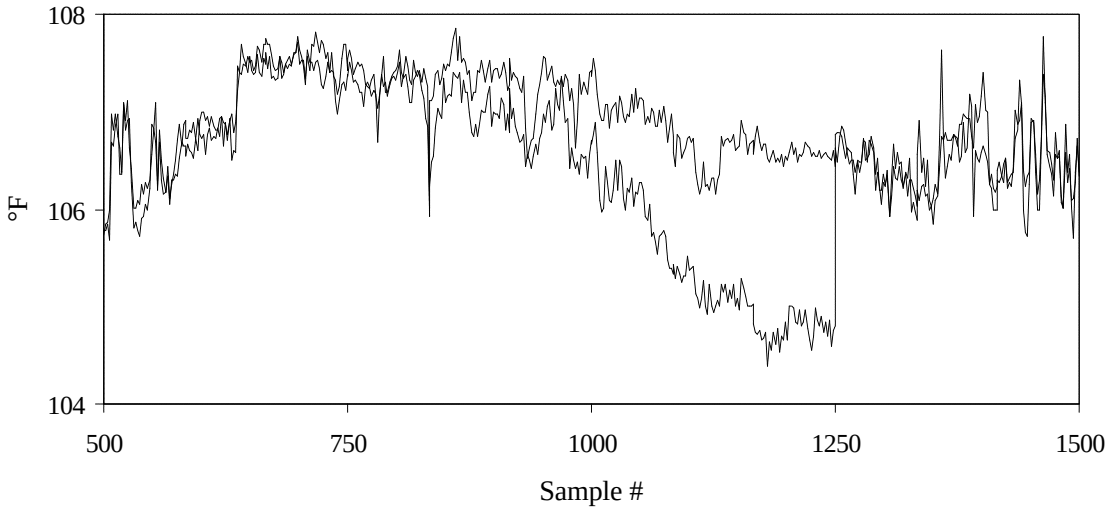


Figure 5. Drifted Measurement and NSET Estimate (Drift commencing at Sample 251)

The output of the SPRT module, contained in Figure 6, corresponds to the drift diagnosis. The diagnosis assumes 5 discrete values that correspond to different states of measurement error. Gross drift, indicated by +/-1.0, has been arbitrarily defined as 30 %FS. Fine drift, indicated by +/0.5, has arbitrarily been set to be 0.5 %FS. No drift is indicated by an SPRT output of 0.0.

These sensitivity settings generally result in an appropriate alarm when the measurement has drifted halfway to one of the setpoints. An alarm at this halfway point denotes that the SPRT has decided that it is more likely than not that the system has drifted. As illustrated in Figure 6, the SPRT catches the drift after 132 samples, corresponding to a drift of 0.26 %FS.

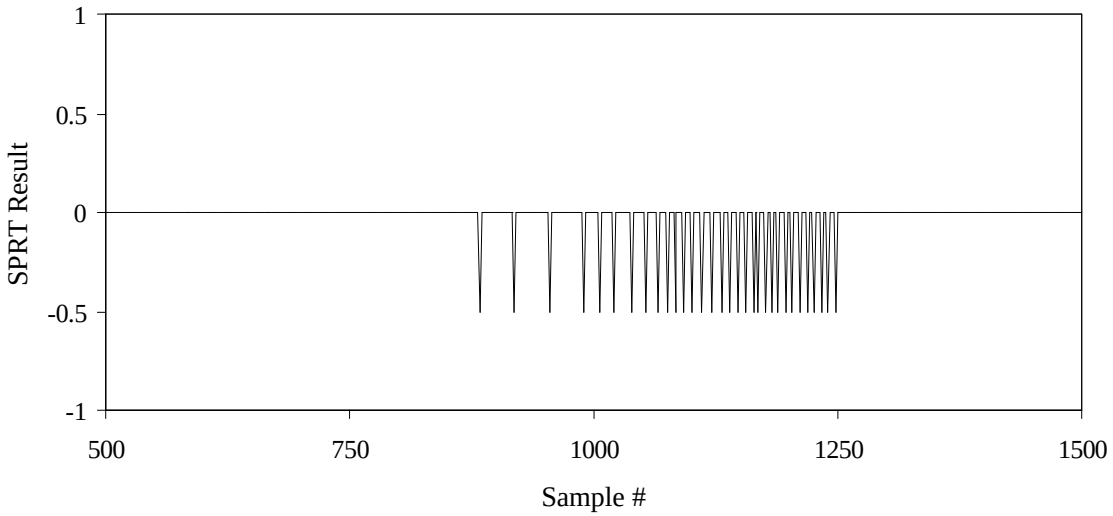


Figure 6. SPRT Drift Diagnosis from NSET (DIST – no scaling) ISCV System

One behavior of the SPRT detection system should be noted: after the initial detection, the SPRT redetects the drift with a diminishingly smaller required number of samples. For example, the second detection only takes another 28 samples. Eventually, as the drift approaches 1.0 %FS, it is redetected on every other sample. Fewer and fewer samples are required as the drift magnitude gets larger with respect to the nominal variance or random error associated with the residual between the measurement and the NSET estimate.

Figures 7-10 display the modeling performance of the NSET using the Euclidean distance operator for the outlet temperature 1 (servo) signal. These figures profile the performance with another typical signal of the same NSET as before, using the same 100 measurement vector prototype matrix, scaling (none), SPRT sensitivities, and the preferred comparison operator.

In Figure 7, the measurement and NSET estimate are displayed with no measurement error. Figure 8 and 9 demonstrate the behavior of the estimate in the presence of measurement error. A linear drift of 0.002 %FS begins at sample 751 and continues until measurement bias of 1 %FS has been reached at sample 1250.

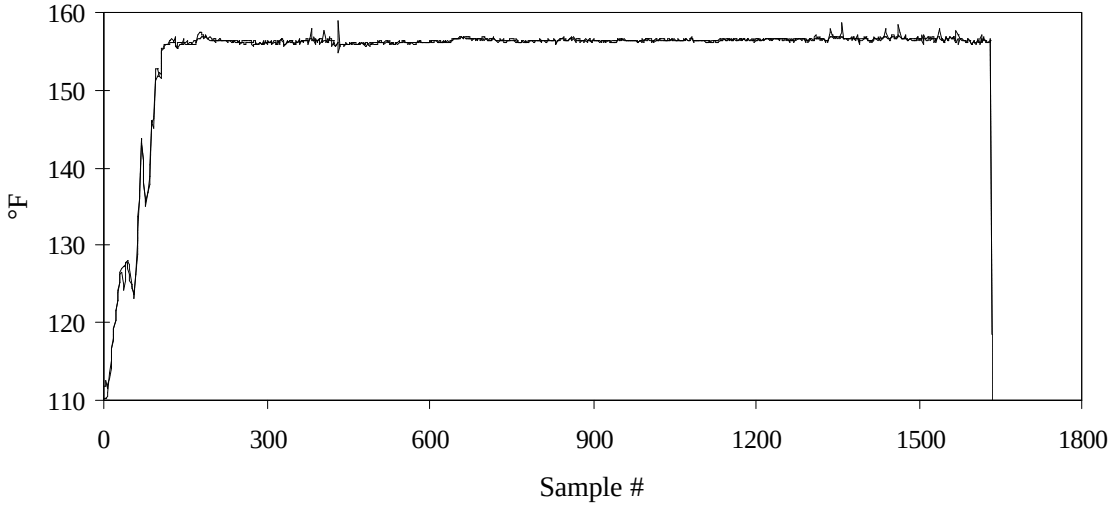


Figure 7. Estimated and Measured Outlet Temperature 1 (Servo) (DIST – no Scaling)

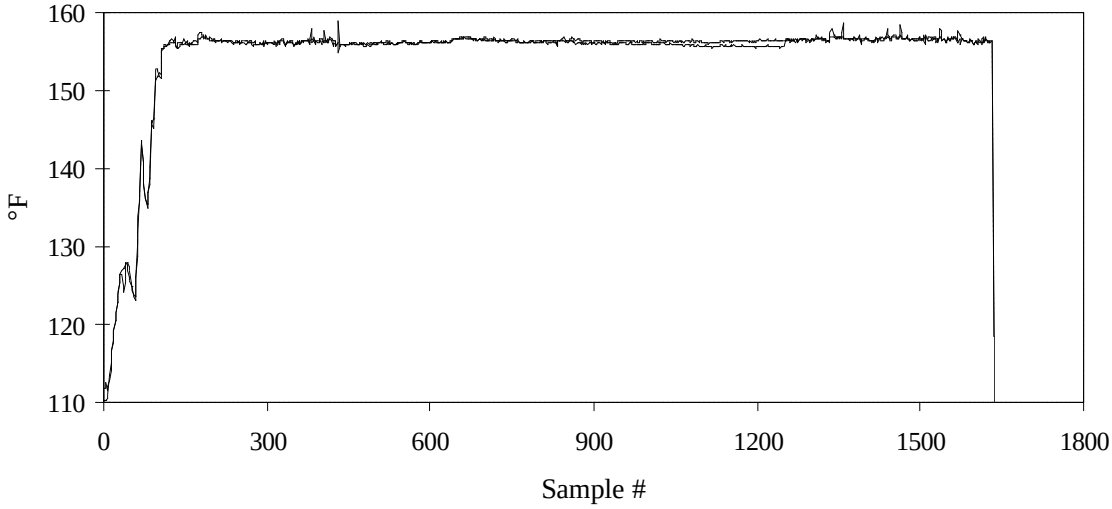


Figure 8. Drifted Measurement and Estimate of Outlet Temperature 1 (Servo) (DIST – no scaling)

In Figure 10, the drift has been detected after 125 samples. The magnitude of the drift at this initial point of detection is 0.25 %FS. This sensitivity, as previously mentioned, was arbitrarily selected to represent a typical application. Most applications allow for some calibration drift, otherwise continuous calibration is required (drift happens). Much finer sensitivity, when desirable (or even practical) is obtained by adjusting a single parameter of the SPRT.

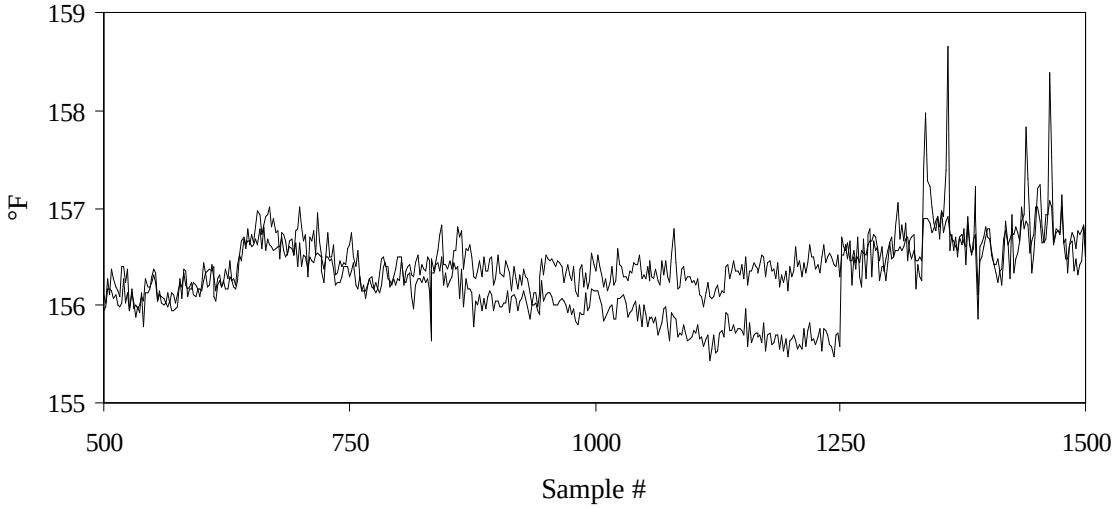


Figure 9. Drifted Measurement and NSET Estimate (Drift commencing at Sample 251)

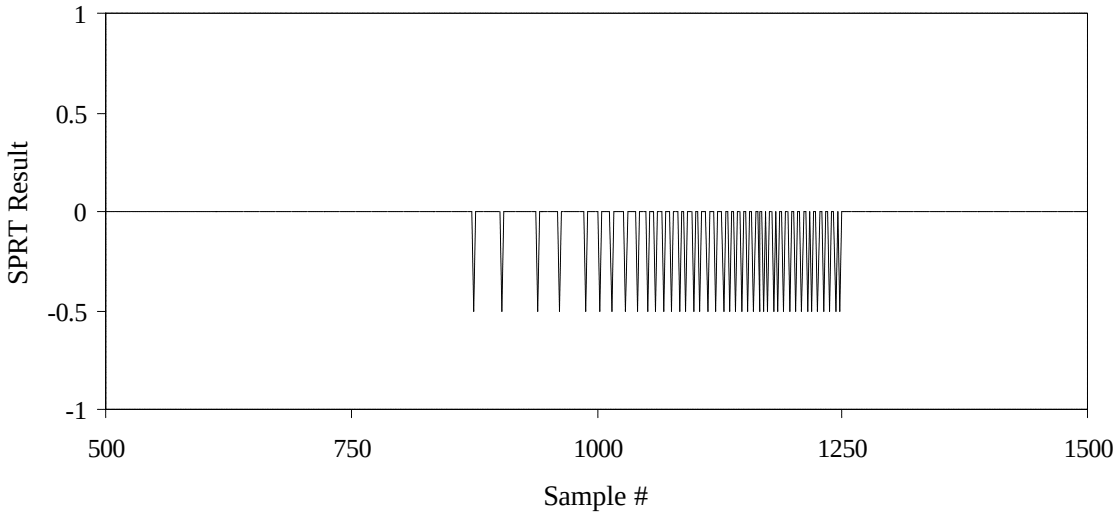


Figure 10. SPRT Drift Diagnosis from NSET (DIST – no scaling) ISCV System

As is apparent in Table 1, several operators produced acceptable prediction errors on the natural test data (containing no drift or measurement error). Some of these operators were found to be highly susceptible to measurement error and thereby less suitable than the chosen DIST/RMPE operators.

Figures 11 and 12 demonstrate respectively the high accuracy clean data estimation and the deficiency typical of NSET estimation with some of these operators, in this case the common mean linear correlation coefficient (CMLCC). First, estimation is demonstrated with no drift or measurement error in Figure 11. Figure 12 then demonstrates the model performance with 0.002 %FS drift per sample. As seen previously, the CMLCC seems to provide very accurate prediction,

as long as the measurements are relatively consistent and free of measurement error. However, as soon as the measurement in Figure 12 begin to drift, an undesirable behavior of the system incorporating the CMLCC operator comes to light: the estimate follows the measurement closely, even as the measurement drifts! The NSET with the CMLCC, as well as with other unsuitable operator choices, cannot predict drift and other forms of measurement error.

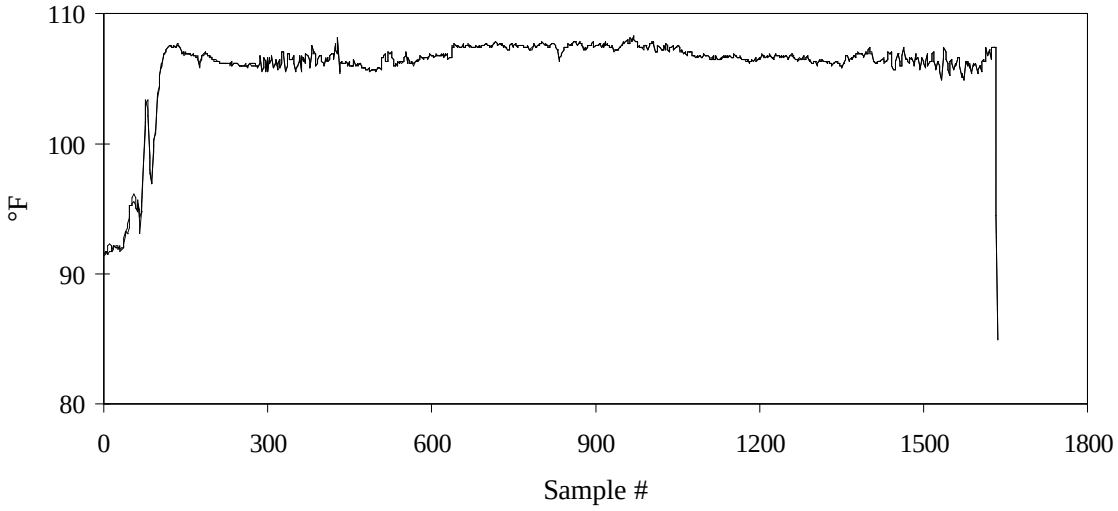


Figure 11. Cooling Tower Inlet Temperature Measurement and CMLCC NSET Estimate

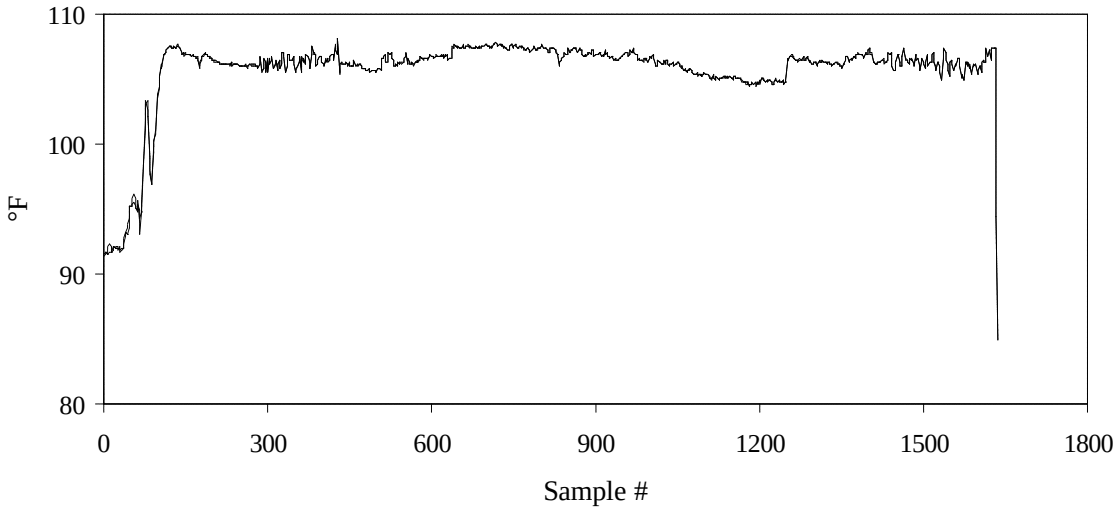


Figure 12. Cooling Tower Inlet Temperature Measurement (Drifted) and CMLCC NSET Estimate

CONCLUSIONS

A nonlinear state estimation technique (NSET) has been developed as a generalization of the multivariate state estimation technique (MSET) [1], which is an extension of the multiple

regression equation. Several nonlinear operators have been employed and explored for the measurement similarity/distance pattern comparison implicit in the NSET. Although several operators provided very good 'clean data' model prediction, only a few operators were found to give robust estimation in the presence of measurement errors.

The root mean power error (RMPE) operator was found to provide the most robust prediction, being virtually impervious to measurement error. The RMPE incidentally also provided its most accurate prediction (while retaining its robust character) with a Minkowski parameter of 2.0. This parameter value rendered the RMPE equivalent to its more specific (also independently explored), and conveniently coded (such that it is almost 60% faster) version -- Euclidean distance (DIST).

The NSET outfitted with the RMPE/DIST operator predicted over 140 measurement vectors (each with 34 components) per second in an interpreted environment. Real time or faster than real time prediction and monitoring of thousands of variables would be possible for a compiled standalone ISCV (with detection of fractional %FS errors) or process modeling application incorporating NSET on an dedicated inexpensive personal computer or workstation.

REFERENCES

1. R. M. Singer, K. C. Gross, and R. W. King, "Analytical Enhancement of Automotive Sensory System Reliability," Presented at the **World Congress on Neural Networks**, San Diego, California, June 1994.

2. R. M. Singer, K. C. Gross, J. P. Herzog, R. W. King, and S. Wegerich, "Model-based Nuclear Power Plant Monitoring and Fault Detection: Theoretical Foundations," **9th International Conference on Intelligent Systems Applications to Power Systems**, Seoul, Korea, July 1997.

3. T. Masters, **Advanced Algorithms for Neural Networks: A C++ Sourcebook**, John Wiley and Sons, Inc., New York, NY, pp. 228-229, 1995.

4. W. H. Press, S. A. Teukolsky, et al., **Numerical Recipes in C: The Art of Scientific Computing, Second Edition**, Cambridge University Press, New York, NY, 1992.

5. J.E. Dennis, Jr. and R.B. Schnabel**, Numerical Methods for Unconstrained Optimization and Nonlinear Equations**, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, pp. 64-66, 1983.

6. S. J. Wan, S. K. M. Wong, and P. Prusinkiewicz, "An Algorithm for Multidimensional Data Clustering," **ACM Transactions on Mathematical Software**, Vol.14, No. 2, pp. 153-162, June 1988.

7. C. L. Black, J.W. Hines, and R. E. Uhrig, "Online Implementation of Instrumentation Of Instrument Surveillance and Calibration Verification Using Autoassociative Neural Networks," **Proceedings of Maintenance and Reliability Conference**, Knoxville, TN, 1997.

8. A. Wald, "Sequential Test of Statistical Hypotheses," **Annals of Mathematical Statistics**, Vol. 16, pp. 117-186, 1945.

9. C. L. Black, J.W. Hines, and R. E. Uhrig, "Inferential Neural Networks for Nuclear Power Plant Sensor Channel Drift Monitoring," **ANS Topical Meeting on NPIC & HMIT**, 1996.

10. D. Wrest, J.W. Hines, and R. E. Uhrig, "Autoassociative Neural Networks for Nuclear Power Plant Sensor Channel Drift Monitoring," **ANS Topical Meeting on NPIC & HMIT**, 1996.

11. D. Wrest, J.W. Hines, and R. E. Uhrig, "Instrument Surveillance and Calibration Verification to Improve Nuclear Power Plant Reliability and Safety Using Autoassociative Neural Networks," **Proceedings of the International Atomic Energy Agency Specialist Meeting on Monitoring and Diagnosis Systems to Improve Nuclear Power Plant Reliability and Safety**, Barnwood, Gloucester, United Kingdom, May 14-17, 1996.

12. K. C. Gross, et al., "Application of a Model-based Fault Detection System to Nuclear Plant Signals," **9th International Conference on Intelligent Systems Applications to Power Systems**, Seoul, Korea, July 1997.

13. O. Glockler, "Fault Detection via Sequential Probability Ratio Test of Multivariate Autoregressive Modeling-Based Residual Time Series," **6th Symposium on Nuclear Reactor Surveillance and Diagnostics, SMORN-6**, Gatlinburg, TN, May 19-24, 1991.