A NEW TECHNIQUE FOR FAULT DETECTION IN MULTI-SENSOR PROBES

L A. Smith, K Godfrey, P Fox† and K Warwick†

University of Warwick, U K † University of Reading, U K

INTRODUCTION

In this presentation we are concerned with fault detection in a multisensor environment. We propose a novel approach to determine the self consistency of the observed values from a collection of sensors. While the original techniques were developed for deterministic chaotic systems, the generalizations presented here apply to both deterministic and stochastic systems. We are interested in the case where a particular process or system is observed by a large number of sensors, each measuring some distinct characteristic of the system. In brief, our approach distinguishes observations which correspond to physically realizable states from those which represent physically nonrealizable states; the latter implying error in the observation (i.e. sensor error).

Before giving the details of the method, we illustrate it with an intuitive example, the simple harmonic oscillator.

$$\frac{dx}{dt} = v \qquad (1)$$

$$\frac{dv}{dt} = -kx \qquad (2)$$

For this "system" consider two sensors which measure the position and velocity (the acceleration would be another *observable* or *measurement function* for this system). Consider a plot of the observed position $x(t)$ against the observed velocity, $v(t)$ as in Figure 1; if these were true observations, the curve would not be continuous, but a collection of arbitrarily close points. Some in the dynamical systems community would call this ellipse an attractor, others would argue that technically this is not a true attractor. As we shall discuss stochastic systems where it is agreed no attractors exist we will avoid the word attractor where possible and, in its stead, consider an *allowed region*. In this case the allowed region would consist of a band of finite width centered on the ellipse in which all the states of the system fall. The remainder of the plane is then the *forbidden region*. Our technique is then simply stated: given an observation of the system, determine whether it is in a forbidden region of the plane. If so, then either the observation is in error (i.e. a sensor is malfunctioning) or the system has changed in some fundamental way In the simple harmonic oscillator, for example, consider the result if one of the sensors, say the position sensor, sticks at a constant value. All observations then fall upon a vertical line as illustrated in Figure 1. As the system evolves, the observation will leave the allowed region along this line thus indicating that a sensor has failed. Our technique is of interest because,as is shown below, it is applicable to systems far more complex than the simple harmonic oscillator and sensor failure far more subtle than returning a constant value. It has been tested on a number of chaotic systems (one example below) where the behavior of the system appears to be quite complex and the individual sensors display no linear statistical relationship.

Chaotic and periodic systems are deterministic in that, given their (exact) position in *phase space*, the future and past state of the system is defined for all time. For chaotic systems, any error in the initial position will grow *on average* exponentially fast, making long term prediction practically impossible and giving these systems a complex appearance. Deterministic systems are contrasted with stochastic systems, where random motion removes this determinism. In stochastic systems, the location in phase space completely specifies the state of the system, but not its future. At the close of this paper, it is shown that the analysis presented below applies to stochastic systems as well.

PHASE SPACE AND SENSOR SPACE

An important property of the systems we shall discuss is the number of *degrees of freedom* they possess. The simple harmonic oscillator has 2; once position and velocity have been given, then the state of the system is completely specified and its evolution is determined by Equations 1 and 2. In general, the state of a system with L degrees of freedom may be represented by a point in an L dimensional phase space. For deterministic systems, as noted above, one point in phase space defines the complete future and past history of the system as well as its present state; in many cases of interest, the observed dynamics of the system lie on a manifold of dimension $d_0 < L$. This is the case for the simple harmonic oscillator where $L = 2$, yet the motion in phase space is confined to a one dimensional closed curve (i.e. $d_0 = 1$). In any case, specification of the L components of the phase space vector $\mathbf{x}(t)$ uniquely defines the present state of the system.

In most instances, direct measurement of all phase space variables is not possible and hence we cannot determine the current state of the system (i.e. its true location in phase space). What are accessible are the values of some measurement functions, which will be called *observations* of the system. Symbolically,

$$s_i(t) = M_i(\mathbf{x}(t)) \qquad i = 1, 2, ..., n_s \qquad (3)$$

where we represent the observed value of the i^{th} sensor at time t by $s_i(t)$. An observation reflects the current state of the system, $\mathbf{x}(t)$, as viewed through the measurement function M_i.

If multiple, independent sensors are available, one may construct a *sensor vector*

$$\mathbf{s}(t) = (s_1, s_2, ..., s_{n_s}) \qquad (4)$$

Just as every point in phase space will map onto a particular value for an individual sensor $s_i(t)$, so each state of the system maps into a single sensor vector $\mathbf{s}(t)$. (Note that this map is not invertible in general: a given sensor vector may represent several states of the system. Such a vector is said to have multiple *preimages* in phase space.) Similarly, the trajectory produced by the evolution of the system in phase space is projected into a trajectory in this *sensor space*. In fact, the entire allowed region is mapped into the sensor space. If $d_0 > n_s$ then this map must be many to one; nevertheless this image need not be dense throughout sensor space. Points in sensor space which have no allowed preimages will form forbidden regions, while the image of the allowed region in phase space will define an allowed region in sensor space.

It is crucial to maintain the distinction between the state of the system in phase space and the characterization of this state via the sensors. The former we shall refer to as the state of the system, the latter as an observation of the system. Implicit is the understanding that the observation depends not only on the state of the system, but also on the sensors. Thus while the state of the system is unique and well defined, an observation may be in error. We also emphasize that even in the case of stochastic dynamics, forbidden regions may still exist. An observation within a forbidden region indicates sensor failure.

Two approaches to error detection are presented below. The first is a simple geometric way of determining whether a given observation is in a forbidden region described above. The second, a predictive method, applies when the number of sensors exceeds the number of degrees of freedom being observed(i.e. $n_s >> L$, this condition is sufficient, but not necessary); in this case there is redundant information in the sensor vector (the map from $\mathbf{x}$ to $\mathbf{s}$ becomes one to one) and we may determine if a given sensor is working by seeing if we can "predict" its value from the others.

In both approaches, the behavior of the system is compared to the expected behavior of the correctly functioning system. Rather than impose some *a priori* model of the system, however, the expected behavior is determined from a *learning data set* taken over a period during which all n_s sensors are operating appropriately. Thus **no explicit model of the system** is required. The learning set provides a collection of sensor vectors, denoted $\mathbf{s}(t^i_{learn}), i = 1, 2, ..., N$, which form the backbone of the allowed region. For clarity, the superscript i shall be suppressed hereafter.

METHOD I : THE GEOMETRIC APPROACH

To determine the reliability of an observation, $\mathbf{s}(t_{test})$ of unknown quality, one tests to see if it lies in the allowed region. The simplest method to do this is determine its nearest neighbor in the learning set, let this distance be δ_{nn}. If this separation is "small", the observation will be considered to be in the allowed region, while the maximum acceptable distance, δ_{max}, will depend on the observational accuracy with which the system is viewed. Thus if $\delta_{nn} < \delta_{max}$, the observation is considered valid. In other words, all points with a distance, δ_{max} of any point in the learning set form the allowed region. Explicitly, let

$$\delta_{nn} = min(|\mathbf{s}(t_{test}) - \mathbf{s}(t_{learn})|) \qquad (5)$$

where the minimum is taken over all vectors in the learning set, $\mathbf{s}(t_{learn})$. This will determine the nearest neighbor to this observation from the learning set. Note that the components of the vector separating the test point from its nearest neighbor may be of use in determining a reasonable weighting for the reliability of each sensor (once they exceed some statistically determined lower cutoff).

Recall the allowed region of the simple harmonic oscillator. It is clear from Figure 1 that the collection of points which pass this test in this system form an elliptical ring of width $2\delta_{max}$ about the ellipse shown in the figure, assuming, of course, that the number of points in the learning set is large enough to cover the ellipse (i.e. so that the nearest neighbor distance for points *within the learning set itself* is less than δ_{max}.) Data requirements are a serious consideration in implementing this formalism.

As a more dynamically interesting system than the simple harmonic oscillator, consider the Moore-Spiegel system. Originally proposed as a model of the motion of a parcel of ionized gas in the atmosphere of a star, we will use the equations to produce complex motions of a system in a 3-dimensional phase space. The equations are (Moore and Spiegel, 1966):

$$\frac{dz}{dt} = v \qquad (6)$$

$$\frac{dv}{dt} = a \qquad (7)$$

$$\frac{da}{dt} + a + (T - R + Rz^2)v + Tz = 0. \qquad (8)$$

where z, v and a are the position, velocity and acceleration of the parcel. We shall consider the case $R = 100.0$ and $T = 36.0$ for which this system displays chaotic behavior. It is important to note that when employing the analysis below we are **not** assuming a knowledge of the form of these equations; they are used only to produce the time series which are then analyzed without reference to the underlying analytical model.

A typical time series of these three variables is shown in the top three panels of Figure 2. It is far from obvious, upon observation of this series, that a simple deterministic system is producing them; standard statistical tests (e.g. the power spectra) suggest that such series are complicated (see Moore and Spiegel, 1966 and Smith, 1987). None the less, the motion of the system in phase space is restricted to a very small region. (In this case, it is an attractor and actually has zero volume in 3-dimensional space!) Figure 3a shows the same $v(t)$ series which is corrupted by pseudo-random noise (with a maximum value of 10 % the range of the signal) added from the midpoint of the graph onwards. The existence of this "error" is detected in Figure 3b, a graph of δ_{nn} with time. Similarly, Figure 3c is again the $v(t)$ series, this time corrupted with a sinusoidal error with a frequency comparable to that of the signal; Figure 3d, the corresponding δ_{nn} graph, shows that this type of error is detected as well.

One of the restrictions of this approach is the amount of data required to outline the allowed region, and this is a system dependent problem. A less data hungry method is available when the number of independent sensors is greater than the number of active degrees of freedom. Before giving the details of this method, we note that for deterministic systems, time delays of a single sensor value may be used in place of additional sensors. That is, $s_j(t) = s_i(t - \tau)$ may be used in addition to $s_i(t)$. This result has a firm mathematical footing (Takens, 1981) and has contributed to the advance of analysis of nonlinear dynamical systems (Packard, et al, 1980). A description of this approach and discussion of the selection of τ are beyond the scope of this contributions, but may be found in Eubank (1990) along with an introduction to chaotic systems. We shall find this approach useful in demonstrating method II, as it avoids the need to construct artificial measurement functions.

METHOD II : THE PREDICTIVE APPROACH

Here we restrict attention to the case where $n_s > L$, that is where there is a unique relation between the location of the system in phase space and its image in sensor space and this property is maintained if any one sensor is turned off. In this case there is redundant information in the sensor vector; exploiting this nonlinear relationship, in particular observing when it fails, allows an evaluation of the quality of the observation. Thus we require a method of detecting deviations from a complex surface of arbitrary dimension defined on a limited, unevenly spaced data set. To accomplish this we apply an interpolation of the surface using radial basis functions. This approach has proven quite powerful in predicting the dynamics of chaotic systems in both numerical systems (Casdagli, 1989,

Smith 1990) and laboratory systems(Smith and King,1990). Here we are presented with the simpler problem of predicting the values of different measurement functions at the same time.

Consider each observation to give rise to a ***basis vector*** consisting of all sensor values save the one to be tested. Let the sensor space have dimension M_s. Denote the basis vector as $\mathbf{x}_j$ and the remaining sensor as $y_j = s_{i_{test}}$. Thus $\mathbf{x}_j$ is a projection of $\mathbf{s}_j$ containing all its components except the component due to the i_{test} sensor.

The predictor is constructed from a learning set consisting of n distinct base vectors ($\mathbf{x}_j, j = 1, 2, ..., n$) and the corresponding y_j taken when all the sensors are functioning properly.

The goal is to determine a function $f(\mathbf{x}_i) : R^{M_s - 1} \rightarrow R^1$ such that

$$f(\mathbf{x}_j) = y_j. \tag{9}$$

We will consider predictors of the form

$$f(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i \phi(||\mathbf{x} - \mathbf{x}_i||) \tag{10}$$

where $\phi(r)$ are radial basis functions (Powell, 1985) and the λ_i are constants which are uniquely determined by equations (9) provided the matrix

$$A_{ij} = \phi(||\mathbf{x}_i - \mathbf{x}_j||) \tag{11}$$

is nonsingular. This is always the case when the $\mathbf{x}_i$ are distinct and the $\phi(r)$ are radial basis functions (Powell, 1985, Michelli 1986). To date, our examination of this approach has been restricted to functions of the form

$$\phi(r) = (r^\alpha + c^2)^{-\beta} \tag{12}$$

for $\beta > -1$ and $\alpha\beta \neq 0$.

This method is demonstrated in Figure 4. In this case we have used a sensor vector:

$$s = (z(t), z(t-\tau), z(t-2\tau), z(t-3\tau))$$

where $\tau = 0.100$. (Note that τ is large compared to the dynamical time of the system, this is *not* an approximation of the derivative.) The solid line shows the observed evolution of the system, while the boxes show the time τ prediction of the signal. In Figure 4a, the uncorrupted signal is demonstrated, while the error is shown in 4b; similar error graphs for additive white noise (Figure 4c) and sinusoidal (Figure 4d) errors are also shown. In the later two, the sensor error in introduced at the midpoint of the graph. All error figures are to the same scale. In this example we have used 256 points in the learning set.

FAULT LOCATION

The methods outlined above are sufficient for the detection of sensor failure but the identification of which sensor is faulty requires more work. For Method I, the analysis may be repeated omitting one sensor at a time from the sensor vector; if the magnitude of δ_{nn} is dominated by contributions from a single component, then that sensor is the one in error. For Method II, we propose the following approach for identifying the failed sensor. It is based on the observation that if predictions are made for a pair of sensors at a time (deleting both from the input basis vector) then, when one sensor fails, the only good predictions will be for its partner. This follows from the observation that, as the failed sensor cannot itself be predicted and as its use will introduce error into the basis vector when estimating the other s_i's, the only case of clean prediction occurs when the failed sensor's output is omitted from both the image and basis vector. This will narrow the choice to two sensors, a situation in which a good prediction is realized only when the valid partner is predicted thus identifying the failed sensor.

THE APPLICATION TO STOCHASITC SYSTEMS

Consider now the applicability of this approach to stochastic systems. To keep things simple, consider a stochastic system with two degrees of freedom; a single sensor directly measuring either of these variables (assuming this could be done) would reveal smooth random variations within some bounds. Now consider this process as observed by a collection of $n_s >> 2$ measurement functions whose outputs are linearly independent. The motion in this n_s-dimensional sensor space could be quite complex, exploring macroscopically all n_s directions; nevertheless, there are still only two variables which determine all the observables of this system. In principle, a knowledge of these two variables is sufficient to determine ("predict") the simultaneous value of any measurement function of this system. This demonstrates the applicability of Method II. When considering the application of Method I to this system, we are faced with the initially paradoxical observation of a stochastic system whose motion is on a low dimensional surface. The paradox is easily resolved with the realization that a multiple sensor probe considers a constant time slice of data; the motion *in time* is stochastic - hence the method of time delays is not as useful for this system- but the set of points accessible to the system in sensor space (at any given time) is very restricted, in fact, low dimensional. This is sufficient to define a nontrivial allowed region and justify the approach of Method I. (It also creates difficulties in interpreting dimension calculations in multi-sensor reconstructions.)

CONCLUSIONS

We conclude with a few observations

a) As noted above, the dynamics of a system is often restricted to a manifold of dimension less than L. Although it is not crucial that this be the case, the lower the volume of the allowed region, the more computationally efficient our method will be, in general.

b) The data requirements for Method II are significantly less than those of Method I which effectively reconstructs geometry from a time series (see Smith, 1988), the reasons why this is the case are discussed in Smith (1990).

c) When applied to observations of deterministic systems, the approach may be viewed as determining whether or not the observation corresponds to a point on the attractor upon which the system's dynamics lie. In this case, this information may be used to make predictions of the future behavior of the system. Obviously, such predictions may provide a valuable aid to both the evaluation and control of such systems.

d) It is important to note that the role played by the sensors in the above analysis need not physically correspond to true sensors; the numerical values used (or estimated) could be input rates, control instructions, and so on. For example, in a single input multiple output system, the learning set might be constructed from a data set where both the input and outputs were monitored, the technique itself could then be used to estimate the unknown input from later measurements of the outputs. An approach similar to this has been discussed by Casdagli (1990).

We have introduced a new method for error detection when multiple sensors are used to observe a dynamical system. Simply put, the multiple sensor output is considered as a vector in a reconstructed phase space. The deviation of this vector from allowed regions of this space indicates either sensor error or a change in the basic properties of the system. This approach appears promising as it directly senses the nonlinear relationships between the different sensors. This means that sensor faults may be detected even in cases where the individual sensor outputs maintain the correct statistical properties. Additional applications to both numerical and physical systems are currently being investigated.

ACKNOWLEDGEMENTS

This research was funded in part by the Science and Engineering Research Council and the United States Office of Naval Research.

REFERENCES

Casdagli M., 1989, Nonlinear prediction of chaotic time series, *Physica D* **35**, 335-356.

Casdagli M., 1990, in Modeling Randomly Forced Nonlinear Systems, presented to Sante Fe Institute workshop on "Nonlinear Modeling and Forecasting." Sante Fe, NM.

Eubank, S. and D. Farmer, 1990, An introduction to chaos and randomness, to appear in *1989 Lectures in Complex Systems, SFI Studies in Complexity*, Sante Fe Institute, Sante Fe.

Moore, D.W. and E.A. Spiegel, 1966, A Thermally Excited Nonlinear Oscillator, *Astrophys. J.*, **143**:871.

Michelli, C. A., 1986, Interpolation of Scattered Data *Constr. Approx.*, **2**:11.

Packard, N.H., J.P.Cruchfield, J.D.Farmer, and R.S. Shaw, 1980, Geometry from a Time Series, *Phys. Rev. Lett.*, **45**:712.

Powell, M.J.D., 1987, Radial basis functions for multivariate interpolation: a review, Preprint, D.A.M.T.P., Univ of Cambridge.

Smith, Leonard A., 1987, Lacunarity and Chaos in Nature, Ph.D. Thesis, Columbia University, New York.

Smith, Leonard A., 1988, Intrinsic Limits on Dimension Calculations, *Phys. Lett. A*, **133**:283.

Smith, Leonard A., 1990, Quantifying Chaos with Predictive Flows and Maps: Locating Unstable Periodic Orbits, in **Measures of Complexity and Chaos**, ed. N.B.Abraham,et al. NATO ASI Series, Plenum Press.

Smith, L.A. and Ĝ. King, 1990, Unstable Periodic Orbits in an Experimental Chaotic Attractor, preprint, Mathematics Institute, University of Warwick.

Takens, F. 1981, Detecting Strange Attractors in Fluid Turbulence. In *Dynamical Systems and Turbulence*,ed D. Rand and L.-S. Young. Berlin: Springer-Verlag.

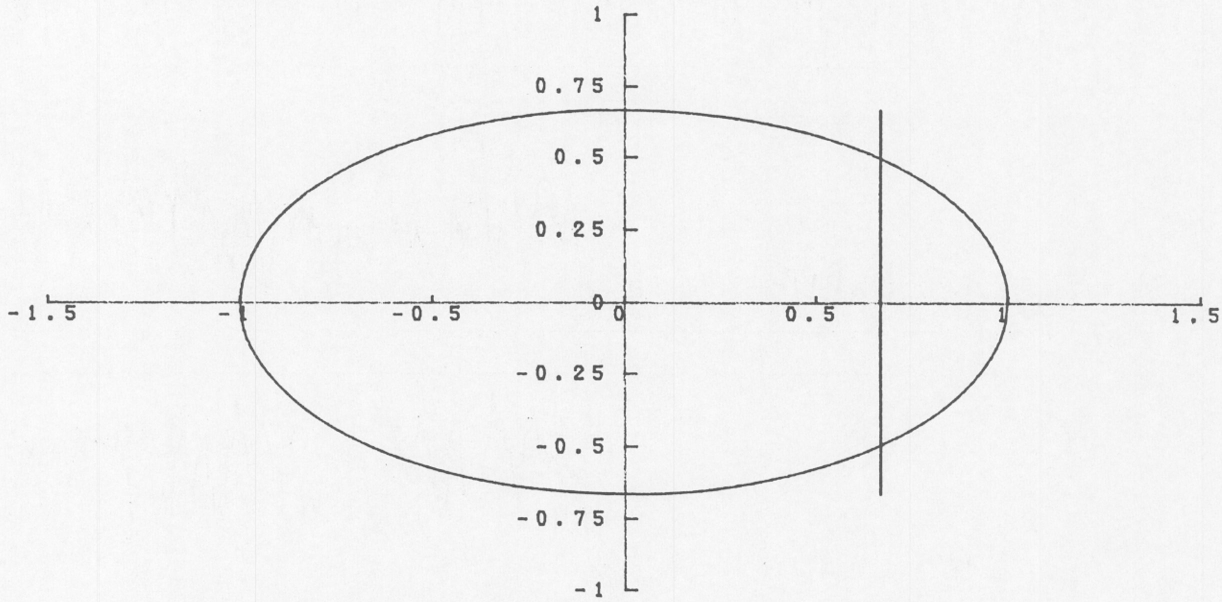

Figure 1. The phase space of the simple harmonic oscillator. Position is given along the x axis and velocity along the y-axis.

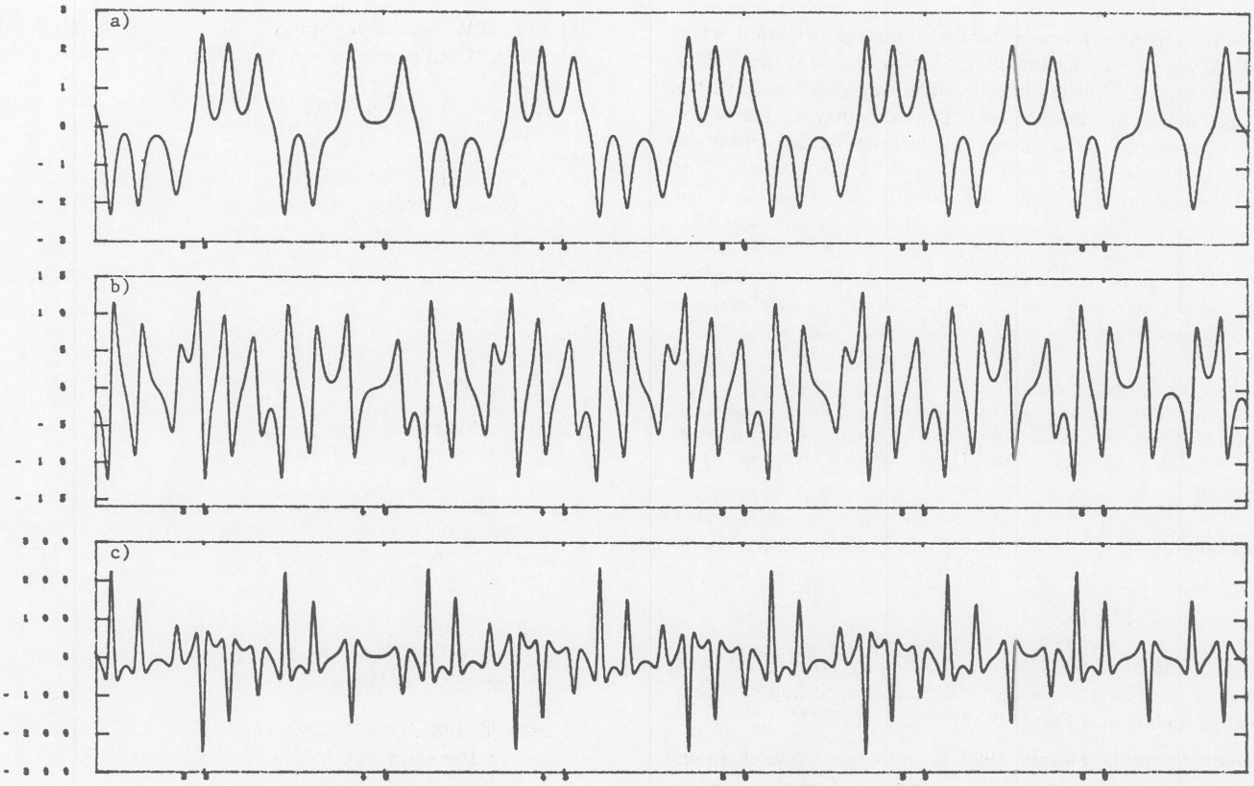

Figure 2. A sample time series for the Moore-Spiegel system with $R = 100$ and $T = 26$. (a) $z(t)$, (b) $v(t)$, and (c) $a(t)$.

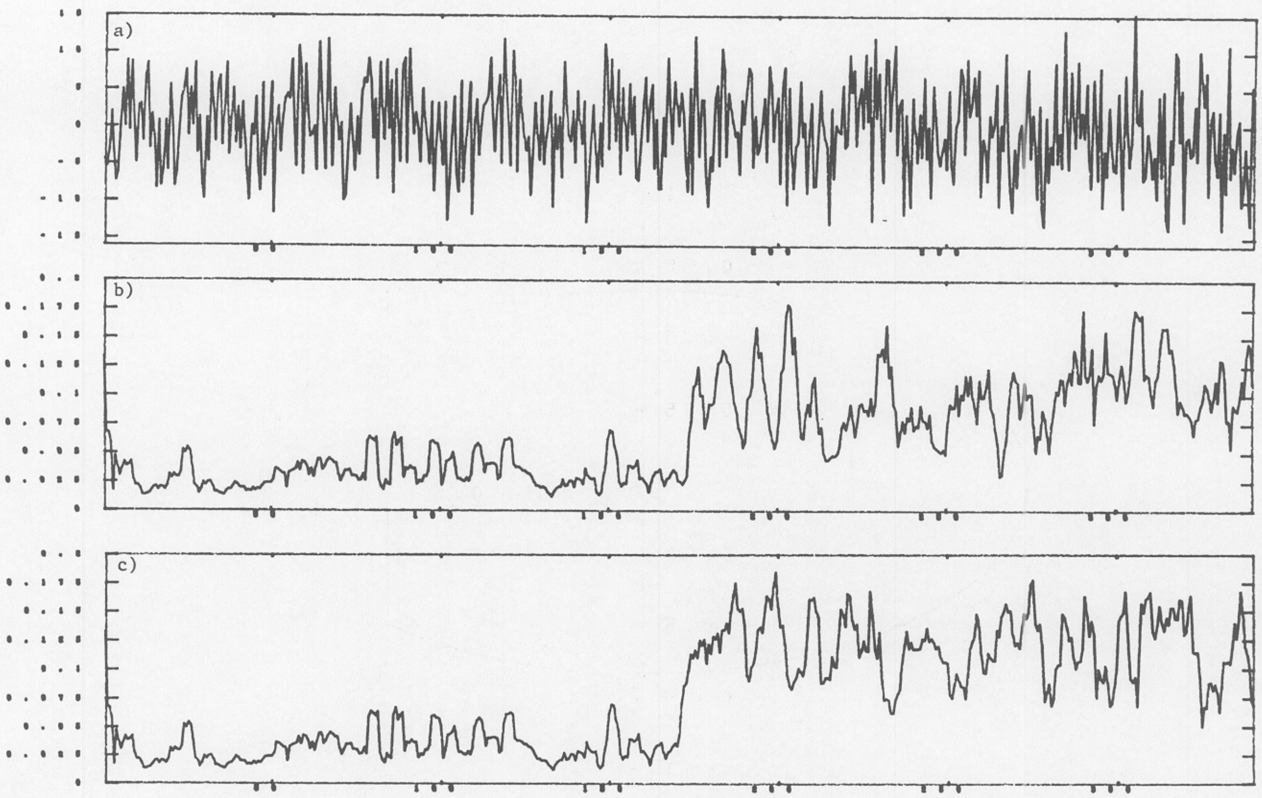

Figure 3. An evaluation of the geometric method as applied to the Moore- Spiegel system. The panels show (a) $v(t)$ and the nearest neighbor distance, $\delta(t)$, for the cases where (b) pseudo-random noise and (c) a sinusoidal term have been added to the observed $v(t)$ values for t in the right half of panels. The presence of this sensor error is clearly refected by the increase in $\delta(t)$. In each case the additional term had a maximum value of 10% the range of the data; a 1 % error was easily detected.

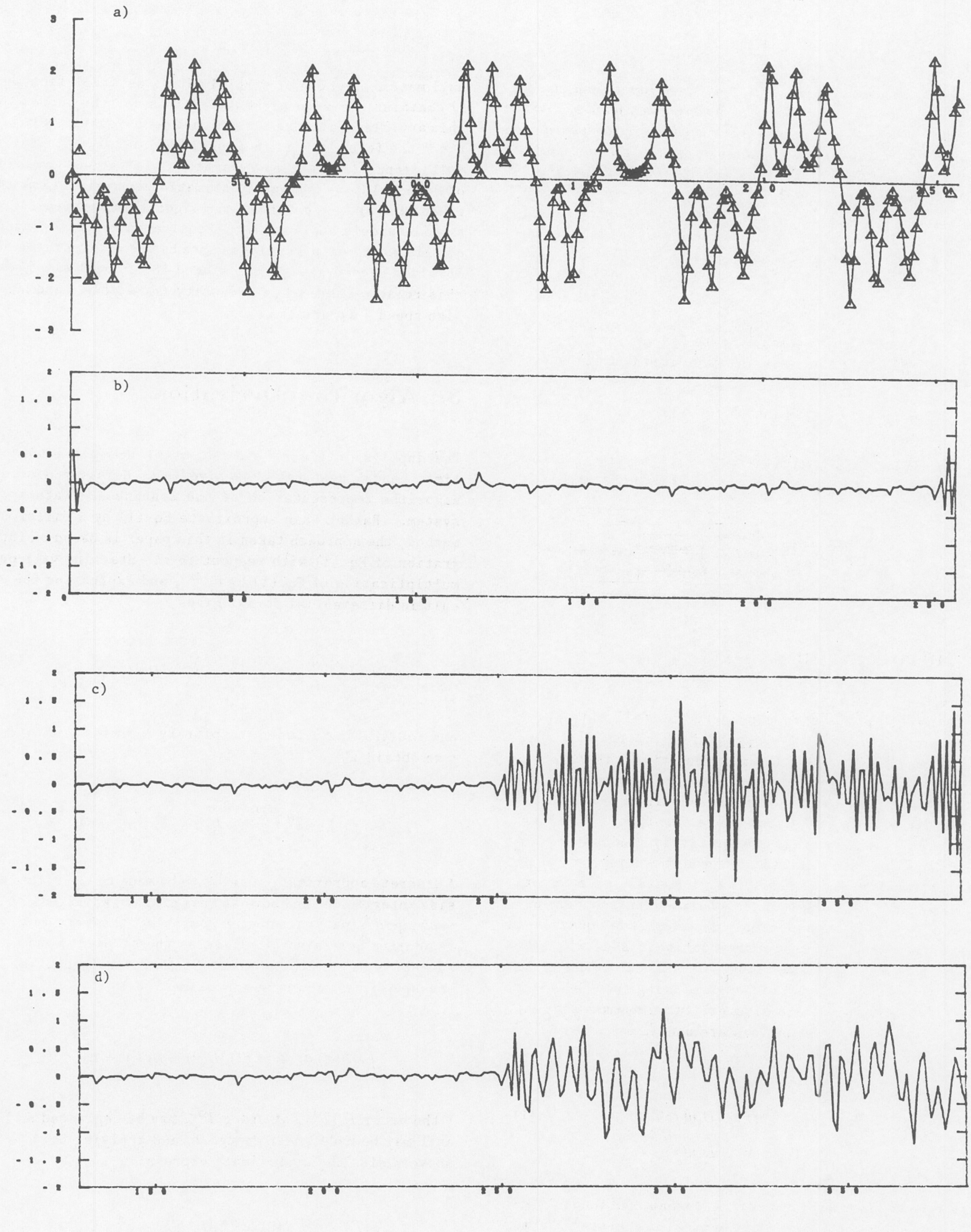

Figure 4. An evaluation of the predictive method as applied to the Moore- Spiegel system. In (a) the system is shown as a solid line, the predictions are boxes. The lower 3 panels show the estimation error for (b) the error free signal, (c) with an additive white pseudo-random noise and (d) with an additive a sinusoidal term. The additive terms were included at the midpoint of each graph as in Figure 3 and they are detected in the increase in the estimation error. In this predictor, 256 points were used with $\phi(r) = r$.