# Nearest Neighbors Methods for Root Cause Analysis of Plant-Wide Disturbances

Margret Bauer[+], John W. Cox[†], Michelle H. Caveness[†], James J. Downs[†], Nina F. Thornhill[*]

[+]Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, UK

[†]Eastman Chemical Company, Advanced Controls Technology Group, Kingsport, TN 37662-5280, USA

[*]Department of Chemical Engineering, Imperial College London, London SW7 2AZ, UK (author for correspondence).

Abstract: In continuous chemical processes, disturbances travel along propagation paths in the direction of the control path or process flow. This article applies a method based on the nearest neighbors of embedded vectors to historical process data for the purpose of identifying the direction of propagation of disturbances. The resulting measure is sensitive to directionality even in the absence of an observable time delay. Its performance is studied in two industrial case studies and default settings for the parameters in the algorithm are derived so that it can be applied in a large scale setting.

Keywords: Nearest neighbors; control loop performance; fault diagnosis; oscillation; whole-plant disturbance; process history; time series analysis.

## 1    Introduction

A fault occurring in a continuous process often results in variations in the local process measurements, and the disturbance can also travel and affect other more remote process measurements.

1

It is then called a plant-wide or whole-plant disturbance and will have a negative impact on the economic performance of the process. In the best case, the plant cannot be run at the desired set points and in the worst case it may have to be shut down. In their study on industrial processes, Desborough and Miller [1] identified common root causes of plant-wide disturbances as

- Process and constraint problems;

- Controller tuning problems;

- Control valve problems.

While the analysis of valve and tuning problems has been widely addressed for single loops [2, 3, 4, 5], few systematic approaches to analyze a whole process are available [6, 7]. The contribution of this article is a data-driven method to find the root cause of a plant-wide disturbance by investigating relationships between pairs of process measurements and thus tracing the fault propagation path.

The work presented here focuses on dynamic plant-wide disturbances that persist over a time horizon of hours or days. The onset of the disturbance is often hard to find, however, especially if it started some time earlier or if it grew gradually. The dynamic nature of these disturbances motivates the treatment of the measurements as time series rather than the use of static methods such as principal component analysis and contribution charts which have had success in on-line fault detection and diagnosis of slowly-developing faults such as catalyst degradation or fouling of a heat exchanger. Papers from industrial commentators [1, 8] emphasise that dynamic plant-wide disturbances are common and very problematical in continuous processes.

When a disturbance propagates along the process flow passing through process equipment, its nature changes. For instance, deviations with sharp spiky features tend to become smoother because of low-pass filtering due to the process dynamics, and there may also be measurement delays. A method that is sensitive to these features in the time series can therefore give information about the propagation path. A number of recently developed statistical measures [9, 10, 11, 12, 13, 14, 15] exploit this information using a method referred to as generalized synchronization [16, 17]. Of two

2

related variables, one is identified as the driver and the other as the response variable [18]. The relationship is described through a mutual phase space representation which is constructed from historical data and the nearest neighbors of the mutual variables. A recent review [19, 20] discussed the construction of digraphs from process models. By contrast, directional information obtained from the analysis proposed in this paper can be used directly without the need for a process model.

The application of a nearest neighbors method to process data gives a *directionality measure* for retracing the fault propagation path in the chemical process. It was proposed by the authors in a conference publication [21] and the contribution of this article is to present the method in depth and to suggest default parameter settings. The methods are tested on two industrial case studies with plant-wide disturbances. One of the disturbances investigated is oscillatory in nature while the other disturbance shows an irregular pattern. The directionality measure is applied to both kinds of disturbances. The results are compared to those from a related and established fault diagnostic tool which uses nonlinearity [22] to show that the directionality measure is able to resolve ambiguous nonlinearity results.

## 2    Nearest Neighbors Methods

Nearest neighbour methods are data-driven and operate on the process measurements stored in a data historian. This section describes the concept of embedded vectors and shows how the nearest neighbors of embedded vectors can be used to detect interdependence and directionality.

### 2.1    Embedded Vectors

For each variable $X$ measurements $x_i$ at time instances $i = 1 \ldots N$ are taken where $N$ is the number of samples. In this context, $X$ refers to the stochastic variable while $x_i$ refers to a sample value of that stochastic variable. The objective is to find information about the state of the process and the relationship between process measurements. A time series method based on the nearest

neighbors of a phase state representation called an *embedding* is investigated to provide a measure of interdependence between two variables.

Embedded vectors give a generic, high dimensional representation of a dynamic system [23]. They are constructed from the measurements by stacking $m$ values in sequential order. At time instance $i$, the embedded vector is defined as $\mathbf{x}_i = \left[ x_i, x_{i-\kappa}, \ldots x_{i-(m-1)\kappa} \right]$ where the embedding delay $\kappa$ can be viewed as a sub-sampling interval. The embedded vectors of a measurement $X$ with $N$ samples can be arranged in the following matrix:

$$
\mathbf{X} = \begin{pmatrix} \mathbf{x}_{(m-1)\kappa+1} \\ \mathbf{x}_{(m-1)\kappa+2} \\ \vdots \\ \mathbf{x}_N \end{pmatrix} = \begin{pmatrix} x_{(m-1)\kappa+1} & \cdots & x_1 \\ x_{(m-1)\kappa+2} & \cdots & x_2 \\ \vdots & & \vdots \\ x_N & \cdots & x_{\tilde{N}} \end{pmatrix} \tag{1}
$$

where $\tilde{N} = N - (m-1)\kappa$ is the number of embedded vectors. The Euclidean norm measuring the distance between embedded vectors $i$ and $j$ is denoted as $d_{i,j} = ||\mathbf{x}_i - \mathbf{x}_j||$. Suggested default values for the embedding parameters $m$ and $\kappa$ as well as the prediction parameter $h$ are given in Section 4. All measurements are treated equivalently regardless of their mean value and engineering units by scaling them to zero mean and unit variance.

The nearest neighbors of $\mathbf{x}_i$ are defined as the embedded vectors $\mathbf{x}_j$ that have the smallest values of $d_{i,j}$. Altogether, $K$ nearest neighbors are sought and the indices of the $K$ nearest neighbors are denoted as $r_{i,j}$, where $j = 1 \ldots K$, and $r_{i,j}$ is the index for the $j$th nearest neighbor of the $i$th embedded vector $\mathbf{x}_i$. Embedded vectors are used for prediction as follows. A future value called the *prediction value* is assigned to each embedded vector. For instance, in the case of self-prediction, the vector $\mathbf{x}_i$ is assigned the prediction value $x_{i+h}$ where $h$ is the prediction horizon.
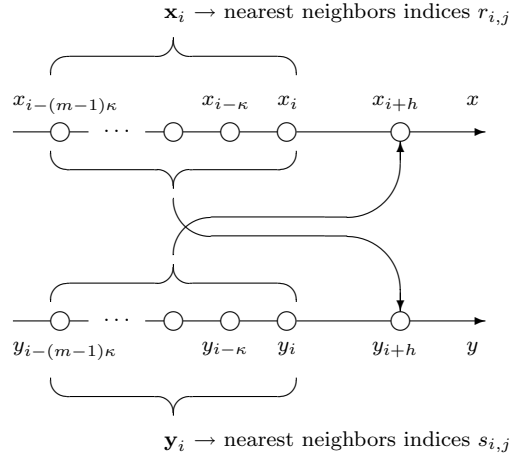
$$\mathbf{x}_i \rightarrow \text{nearest neighbors indices } r_{i,j}$$

$$\mathbf{y}_i \rightarrow \text{nearest neighbors indices } s_{i,j}$$

Figure 1: Find the nearest neighbors of $\mathbf{x}_i$ and $\mathbf{y}_i$ to predict future values $x_{i+h}$ and $y_{i+h}$.

The same formulations are made for a second variable $Y$ such that

$$\mathbf{Y} = \begin{pmatrix} \mathbf{y}_{(m-1)\kappa+1} \\ \mathbf{y}_{(m-1)\kappa+2} \\ \vdots \\ \mathbf{y}_N \end{pmatrix} = \begin{pmatrix} y_{(m-1)\kappa+1} & \cdots & y_1 \\ y_{(m-1)\kappa+2} & \cdots & y_2 \\ \vdots & & \vdots \\ y_N & \cdots & y_{\tilde{N}} \end{pmatrix}. \tag{2}$$

To each embedded row vector $\mathbf{y}_i$, the nearest neighbors indices $s_{i,j}$ and the prediction value $y_{i+h}$ are assigned accordingly.

## 2.2 Directionality Measures

Figure 1 shows the concept of a directionality measure using cross-prediction based on nearest neighbors. The method evaluates how well the embedded vectors in $X$ anticipate the prediction values $y_{i+h}$, and vice versa. The prediction value $y_{i+h}$ is now assigned to the embedded vector of the first sequence $\mathbf{x}_i$ instead of being assigned to $\mathbf{y}_i$ . The prediction value $y_{i+h}$ is then compared to the predicted values $y_{r_{i,j}+h}$ of the nearest neighbors $\mathbf{x}_{r_{i,j}}$. If $y_{i+h}$ and all $y_{r_{i,j}+h}$ have similar values this means that $X$ is a good predictor of $Y$. The prediction from $Y$ to $X$ is measured similarly by

assigning to each embedded vector $\mathbf{y}_i$ a predicted value $x_{i+h}$ and comparing the predicted value with the prediction of the nearest neighbors. If one variable is a better predictor of the other variable then a directional relationship is inferred.

For clarification, an example is considered. Assume two time sequences $x_i$ and $y_i$ are available with $i = 1 \ldots N$ and that $m = 4$ and $\kappa = 1$. Suppose that the nearest neighbor of row 1 of the embedded matrix $\mathbf{X}$ in Equation 1 is row 50. That is, the nearest neighbor of $\mathbf{x}_4 = [x_4, x_3, x_2, x_1]$ is $\mathbf{x}_{53} = [x_{53}, x_{52}, x_{51}, x_{50}]$. If the difference between the prediction values $y_{4+h}$ and $y_{53+h}$ are small then $X$ is a good predictor of $Y$. Similarly, suppose that the nearest neighbor of row 1 of the embedded matrix $\mathbf{Y}$ from Equation 2 is row 50. If the difference between the prediction values $x_{4+h}$ and $x_{53+h}$ is large then $Y$ is a poor predictor of $X$.

Directionality measures based on nearest neighbors have been developed only recently, with the first methods being applied to long data sequences in the mid 1990s. The reason for this is that it is not an analytical method and its computation requires a large number of numerical calculations that were only available with the advancement of modern PC technology. Among the first to propose measuring interdependency in this way were Le Van Quyen et al. [9]. Schiff et al. [11] presented a similar method that measured generalized synchrony by assigning to each embedded vector $\mathbf{x}_i$ the corresponding vector $\mathbf{y}_i$ and finding the nearest neighbors. Variations of the methods proposed in [9] and [11] were compared by Arnhold et al. [15]. Two recent methods [10, 12] successfully investigated the scaling to take account of self-predictability. Self-predictability assesses whether the time series is predictable from its own past history by assigning a prediction value $x_{i+h}$ to each embedded vector $\mathbf{x}_i$.

## 2.3   Algorithm

The directionality measure as adapted from [13] determines the differences between the prediction value of each time instance $i$ and the prediction values of the nearest neighbors $s_{i,j}$ of $\mathbf{y}_i$. The

algorithm incorporates self-predictability and is based on a distance measure:

$$D_i(X|Y) = \frac{1}{K} \sum_{j=1}^{K} |x_{i+h} - x_{s_{i,j}+h}|. \tag{3}$$

where $j$ is the index of the $j$th nearest neighbor. In Equation 3, $x_{i+h}$ is the prediction value of $x$ assigned to $\mathbf{y}_i$ and $x_{s_{i,j}+h}$ is the prediction value assigned to the $j$'th nearest neighbor of $\mathbf{y}_i$. For robustness in the presence of outliers, the measure is summed over $K$ nearest neighbors instead of choosing just one single nearest neighbor. Outliers can distort the result by accidentally being a particularly good or poor predictor and without averaging over the $K$ nearest neighbors, the impact of an outlier would be significant. If $D_i(X|Y)$ is small then $Y$ is a good predictor of $X$ and if $D_i(Y|X)$ is large then $Y$ is a poor predictor. The distance measure is scaled by a self-predictability factor:

$$D_i(X) = \frac{1}{K} \sum_{j=1}^{K} |x_{i+h} - x_{r_{i,j}+h}| \tag{4}$$

where $x_{i+h}$ is the prediction value of $x$ assigned to $\mathbf{x}_i$ and $x_{r_{i,j}+h}$ is the prediction value assigned to the $j$'th nearest neighbor of $\mathbf{x}_i$. The quantities $D_i(X|Y)$ and $D_i(X)$ are computed for all embedded vectors and time instances $i$. The result are then scaled and averaged to give the following accumulated interdependence measure:

$$H(X|Y) = \frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} \frac{D_i(X|Y)}{D_i(X)}. \tag{5}$$

The complementary measure $H(Y|X)$ is defined similarly by exchanging $X$ and $Y$ in the formulation. To establish whether $X$ influences $Y$ more than $Y$ influences $X$, these two measures are compared to give a directionality measure $H_{X \to Y}$:

$$H_{X \to Y} = H(X|Y) - H(Y|X). \tag{6}$$

A positive value of $H_{X \to Y}$ is taken to mean that $X$ influences $Y$, if it is negative then $Y$ influences $X$. The formulation given above in Equations 3 to 6 broadly follows that of [13]. The main difference is that the algorithm of [13] sums the square of the differences in Equations 3 and 4 while in this
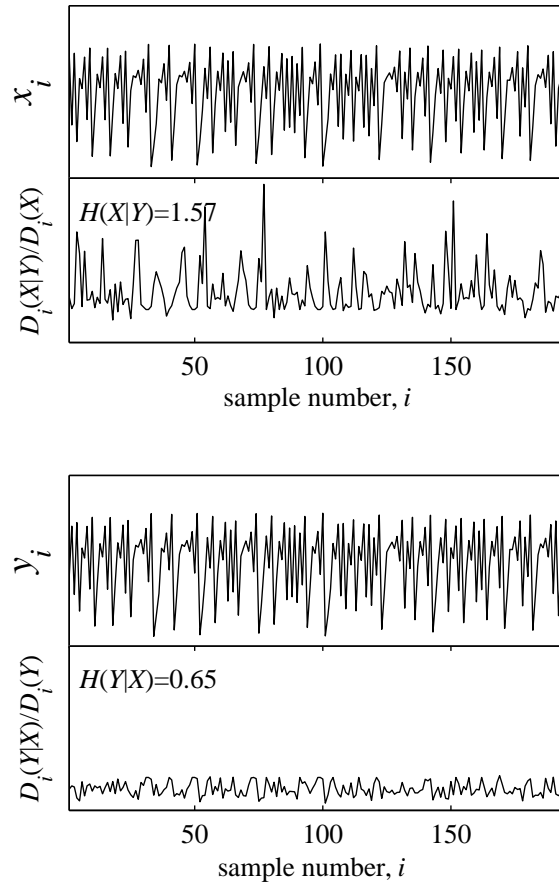
Figure 2: Construction of the directionality measure from $D_i(X|Y)/D_i(X)$ and $D_i(Y|X)/D_i(Y)$ for the Hénon map, $a = 1.4$, $b = 0.3$. ($m = 5$, $\kappa = h = 1$, $K = 10$, $N = 200$).

paper, the absolute value is summed. This modification was introduced since the absolute value is more intuitive as a measure of difference.

The construction of the causality measure is illustrated in Figure 2 for the Hénon map. The Hénon map is a dynamic system that generates two related time series which are often used as a reference [15]. The equations of the dynamic system are given by $x_{i+1} = y_i + 1 - ax_i^2$ and $y_{i+1} = bx_i$. The resulting time series are chaotic with $X$ influencing $Y$ to a greater extent than $Y$ influences $X$, as discussed in [15], [24] and elsewhere. The normalized time series are plotted in Figure 2.

The distance measure $D_i$ is computed and $D_i(X|Y)/D_i(X)$ as well as $D_i(Y|X)/D_i(Y)$ are plotted for each time instance $i$. Figure 2 shows that $D_i(X|Y)/D_i(X)$ is significantly larger, especially at some time instances which result in large peaks. The values of the interdependence measures are $H(X|Y) = 1.57$ indicating $Y$ is a poor predictor of $X$, and $H(Y|X) = 0.65$ which means that $X$ is a better predictor of $Y$. The directionality measure $H_{X \to Y}$ is positive, showing directionality from $X$ to $Y$, as expected.

## 2.4  Significance Level

A threshold is needed to determine whether a value of $H_{X \to Y}$ is significant. The approach used in this paper is to test whether the asymmetry of the interdependency measures is significant. If it is not significant, then a directional relationship can be ruled out. If the asymmetry is significant then a directional relationship cannot be ruled out. Pereda et. al. [30] explain that it can be difficult to be sure that asymmetry in the interdependence measures $H(X|Y)$ and $H(Y|X)$ is attributable to a directional dependence. They warn that other factors such as differences in the dynamics of the time series may contribute to the asymmetry. In the plant-wide case, a dynamic disturbance arises in one location and spreads through the process. The dynamics of the time series measured at different locations are therefore likely to be similar, suggesting that asymmetry can in practice reasonably be attributed to directionality.

A basic approach would be to create a set of random sequences two at a time and to compute their $H_{X \to Y}$ values. If the $H_{X \to Y}$ value for the $X$ and $Y$ time series under test is significantly different from those of the random sequences, then one cannot rule out the possibility that directional dependence might be present between $X$ and $Y$. While random sequences do not have any interdependence, their disadvantage is that they do not generally resemble the test time series. There is thus no logical basis for choosing the settings for the parameters in the algorithm because these depend on the nature of the time series. A set of surrogate time series may be used to overcome this shortcoming [25], [26].

An surrogate time series has a power spectrum, autocorrelation function and probability density function that are well matched to those of the time series under test, but with the phases of the frequency components randomized. The amplitude adjusted Fourier transform method [25] provides surrogates of this type. They key steps are: (i) nonlinear scaling to convert the amplitude distribution to Gaussian, (ii) adding random phases to the arguments of the discrete Fourier transform (DFT), and (iii) re-inversion via the inverse DFT and rescaling to restore the original amplitude distribution. Such surrogate time series derived from $X$ and $Y$ have no interdependence because of phase randomization, but they retain the key aspects of the original time series that influence the choice of embedding parameters. Computational aspects such as dealing with the aliased components of the DFT and end-matching are explained in [27], while practical examples using process data are demonstrated in [22].

The null hypothesis is that $X$ does *not* influence $Y$ and is denoted by $\lambda_0 = H_{X \to Y}$. Suitable alternative numbers are values of the directionality measure generated by computing the measure from surrogate time series of $x_i$ and $y_i$, $i = 1 \ldots N$: $\lambda_k = H^k_{x_{surr} \to y_{surr}}$ for $k = 1 \ldots N_{surr}$. The null hypothesis is rejected if $\lambda_0$ varies considerably from all surrogates values $\lambda_k$. The significance level is estimated to measure the deviation of the directionality measure of the original from the surrogate as follows:

$$\Omega_{X \to Y} = \frac{\lambda_0 - \mu_\lambda}{\sigma_\lambda} \tag{7}$$

with mean value

$$\mu_\lambda = \frac{1}{N_{surr}} \sum_{k=1}^{N_{surr}} \lambda_k \tag{8}$$

and variance

$$\sigma_\lambda^2 = \frac{1}{N_{surr} - 1} \sum_{k=1}^{N_{surr}} (\lambda_k - \mu_\lambda)^2. \tag{9}$$

If $\Omega_{X \to Y}$ is large then the null hypothesis that $X$ does not influence $Y$ can be rejected. The threshold for $\Omega_{X \to Y}$ in Theiler et al. is set to three which indicates that the distance of the directionality measure to the mean of the measure for the surrogates is three times the standard deviation of the

10

surrogate distribution. In the following, the threshold is set accordingly to $\Omega_{X \to Y} \geq 3$.

Detailed discussion of surrogates for interdependence testing can be found in [28], [29] and [30]. The test outlined above uses univariate surrogates and therefore tests only the null hypothesis that the $X$ and $Y$ time series are independent linear stochastic processes. It concludes that interdependence exists if the null hypothesis is rejected but would require the use of bivariate surrogates to determine the nature of the interdependence. For instance, as explained in [29], the significance test described here does not distinguish between the nonlinear interdependence that is present in the dynamic equations for the Hénon map and linear interdependence such as when two sources are combined in a linear manner to give the coupled $X$ and $Y$ time series.

## 2.5    Nonlinearity Measure

A property that can assist in finding the root cause of a plant-wide disturbance is the nonlinearity of a time series [22]. As established in [22] and demonstrated in several industrial case studies, the closer a measurement is to a nonlinear root cause, the more nonlinear is its resulting time series. The basic reason for the observation is that nonlinear features of the time series such as phase coupling and harmonics are reduced as the disturbance propagates through the process. Thus, the variable which shows the highest nonlinearity is most likely to be closest to the root cause. A nonlinearity measure can be constructed as follows based on [26]:

$$\Gamma_x = \frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} \left( x_{i+h} - \frac{1}{K} \sum_{j=1}^{K} x_{r_{i,j}+h} \right)^2 \tag{10}$$

where $r_{i,j}$ is the index of the $j$th nearest neighbor of $\mathbf{x}_i$. A significance level $\Omega_X$ is computed for each time series in the same fashion as the computation of $\Omega_{X \to Y}$ through Equations 7 to 9 only with a scaling factor of three, that is, $\Omega_X = \frac{\lambda_0 - \mu_\lambda}{3\sigma_\lambda}$. Since the nonlinearity index is an established method and based on the same dynamical features as the directionality measure, it will be used to give insights into the results in the case studies. The $3\sigma_\lambda$ term in the denominator scales the index
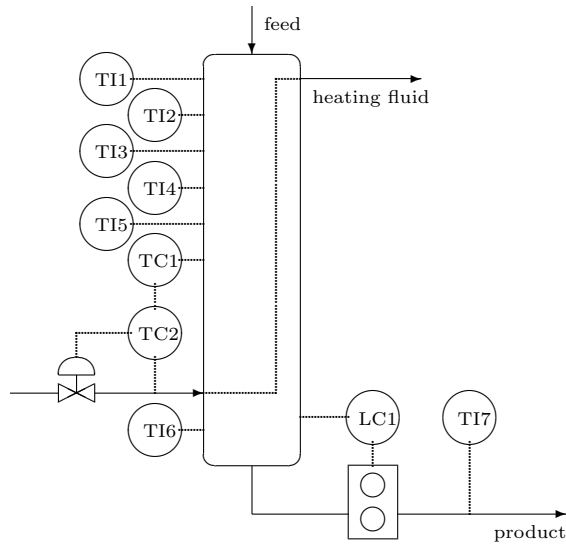
11

Figure 3: Process schematic of Case Study I.

so that $\Omega_X \geq 1$ indicates nonlinearity.

# 3 Applications

In this section, two case studies of industrial processes with plant-wide disturbances are introduced. The root cause of each disturbance as well as the propagation path are known in both cases.

## 3.1 Case Study I

Case Study I is a distillation unit within a larger production plant at Eastman Chemical Company. Figure 3 shows the process schematic. Altogether nine temperature measurements, two of them controlled variables, and one controlled level measurement were available. A feed enters the top of the column and is separated into the desired product that exits the column at the bottom and a by-product that exits the column at a side draw not shown in Figure 3. Heating fluid is pumped through a piping system along the length of the column and exits at the top. The heating fluid
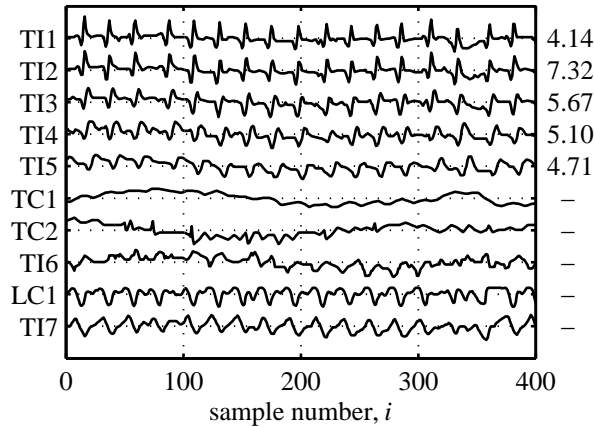
Figure 4: Segment of the process measurements of Case Study I, with nonlinearity index.

flow is controlled by the heating fluid temperature because the heating fluid is a shared facility with a varying temperature. The temperature in the column is controlled by a cascade loop for which the master controller (TC1) measures the temperature in the middle of the column and the slave controller (TC2) uses the temperature of the heating fluid. The flow out of the column is the manipulated variable for the bottom tray level (LC1) and is adjusted through a pump. Temperatures are additionally measured along the upper part of the column (TI1 to TI5), at the bottom tray (TI6) and downstream (TI7).

A periodic disturbance affected all process measurements, as can be seen in Figure 4 which shows a segment of the process measurements that were analyzed. The disturbance affects all measurements, most distinctively TI1 to TI5 and LC1 as well as TI7. The sampling interval is 20 seconds and the period of oscillation is around 61 samples, roughly 20 minutes. Although all measurements share the same period of oscillation, suggesting that it is indeed the same plant-wide disturbance, the shape of the waveform is different for each variable. TI1 and TI2 show spiky peaks while TI3, TI4 and TI5 exhibit a smoother trend and TI7 shows an oscillation with an almost triangular shape.

For the nearest neighbors analysis, the time sequences are sub-sampled by a factor of three. A time frame of 500 samples is considered comprising 25 full cycles with a period of oscillation of

13

20 samples. Reducing the number of samples to be analyzed keeps the computation time within bounds, and sub-sampling enables more cycles to be analyzed using a limited data ensemble.

The root cause of the disturbance was a process problem in an upstream reactor. A pressure set point was set too high and was suspected to lead to flow surges. The uneven flow affected the heating of the product along the top part of the column and thus the temperature. The level was upset by flow surges as well, and once the upstream pressure set point was corrected the disturbance went away. The data set shown in Figure 4 was captured before the root cause was eliminated.

The expectation for Case Study I is that the nearest neighbors method should point to temperature measurements TI1 and TI2 at the top of the distillation column as closest to the root cause because these measurements are closest to the upstream disturbance. This is also supported by the nonlinearity index that is indicated on the right hand side of Figure 4. All the temperature indicators in the top half of the column (TI1 to TI5) show high nonlinearity. While the nonlinearity index $\Omega_X$ points to the top of the column, the directionality measure can additionally point to the origin and propagation of the disturbance. Results of the nearest neighbors analysis for this case study are presented in Section 5.1.

## 3.2 Case Study II

While the first case study showed a periodic disturbance, the disturbance of the second case study is non-oscillatory. The unit is a reactive distillation column at Eastman Chemical Company which is shown in Figure 5. The head outflow of the column is recycled via the condenser and reflux tank and the pressure in the column is controlled through the inert gas flow via PC1. To prevent the reflux tank from overflowing an exit line outflow is supervised by LC1. The temperature in the column is critical for the chemical reaction to take place and therefore not only controlled by TC1 and TC2 but also observed by TI1 and TI2. A pressure and temperature gradient can be observed along the column. PI1 and PI2 measure steam and inert gas pressure into the reboiler and flash pot.
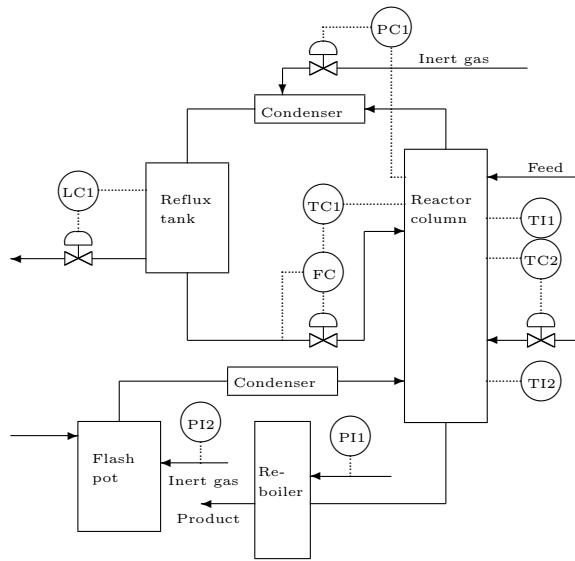
14

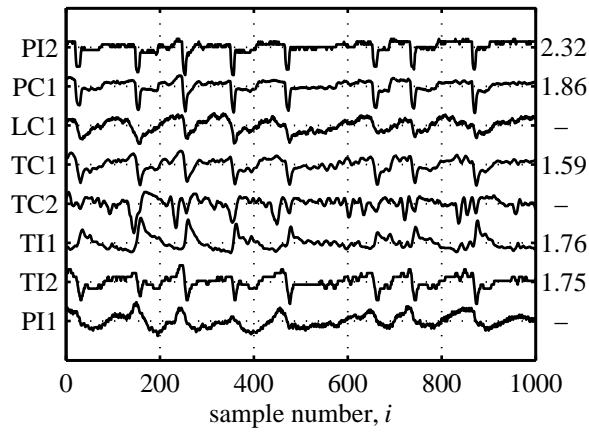Figure 5: Process schematic of Case Study II.



Figure 6: Process measurements for Case Study II, with nonlinearity index.

15

A disturbance affected the measurements around the acid column. Measurements participating in the disturbance are shown in Figure 6 with a sampling interval of 10 seconds. A common pattern can be observed in all time trends. Sharp spikes occur at around the same time in level, pressure and temperatures throughout the process. The spikes are followed by periods of more steady operation of varying length so that the disturbance exhibits an irregular pattern. Visual inspection shows that the spikes are part of the same disturbance since the peaks line up in the various measurements. Level measurement LC1 was distorted by random noise and has been filtered with a sixth order moving-average low pass filter prior to analysis and plotting. The acausal nature of the filter did not introduce a time delay that would interfere with the interdependence analysis.

The disturbance was due to fluctuations of the inert gas pressure caused by a split valve controller further upstream in the inert gas inlet. Investigations showed that the pressure controller had oversized split settings which caused the inflow to be upset thus affecting temperatures and pressures in the process. The process schematic in Figure 5 shows that PI2 and PC1 are two pressures connected to the inert gas flow. The expectation for the result of the nearest neighbors method for this case study is that it should show the two pressures PI2 and PC1 are closest to the root cause. This expectation is supported by the nonlinearity index as shown in Figure 6. The nonlinearity index has the highest value ($\Omega_X = 2.32$) for PI2 and the second highest ($\Omega_X = 1.86$) for PC1. Thus it is expected that PI2 and PC1 cause the other measurements of Case Study II. Again, the nearest neighbors method will enhance the root cause analysis by showing the propagation path. Results of the nearest neighbors analysis for this case study are presented in Section 5.2.

# 4    Parameter Settings

The values selected for the parameters in the algorithm are those which give high significance levels $\Omega_{X \to Y}$ for the directionality measure. When constructing the embedded vectors, the following parameters have to be selected:

16

- Embedding dimension $m$;

- Embedding delay $\kappa$;

- Prediction horizon $h$;

- Number of samples $N$;

- Number of nearest neighbors $K$;

- Number of surrogates $N_{surr}$.

The first three parameters capture the dynamics of the time series. The number of samples $N$ has an impact on the magnitude of the results but not on the direction since more data should give a better and thus more pronounced directionality. The number of nearest neighbors $K$ and the number of surrogates $N_{surr}$ are expected to have only a limited impact on the measure above a certain threshold since these variables are statistical averages. The obvious choice of $K$ is to set it equal to the number of cycles that are analyzed in case of an oscillatory disturbance. In case of a non-oscillatory disturbance, as in Case Study II, the number of repetitive peaks is an equally straightforward choice. The number of surrogates $N_{surr}$ is set to 20 as this is expected to give statistically relevant results while keeping the computational effort within a feasible limit. The number of samples is chosen as 500, however, since the number of samples has the strongest impact on the computation time, the minimum number of samples is investigated later on in this section.

The selection of default parameters is carried out in three steps. First, the embedding dimension $m$ is varied while the embedding delay $\kappa$ and the prediction horizon $h$ are kept at a fixed value. The value of embedding dimension $m$ for which the significance level has its maximum for the majority of relationships is selected. In the second step, the best value of embedding delay $\kappa$ is found by varying $\kappa$ while the embedding dimension $m$ is kept at the specified best value. The prediction horizon $h$ is varied in the last step with the same procedure. The minimum number of samples is
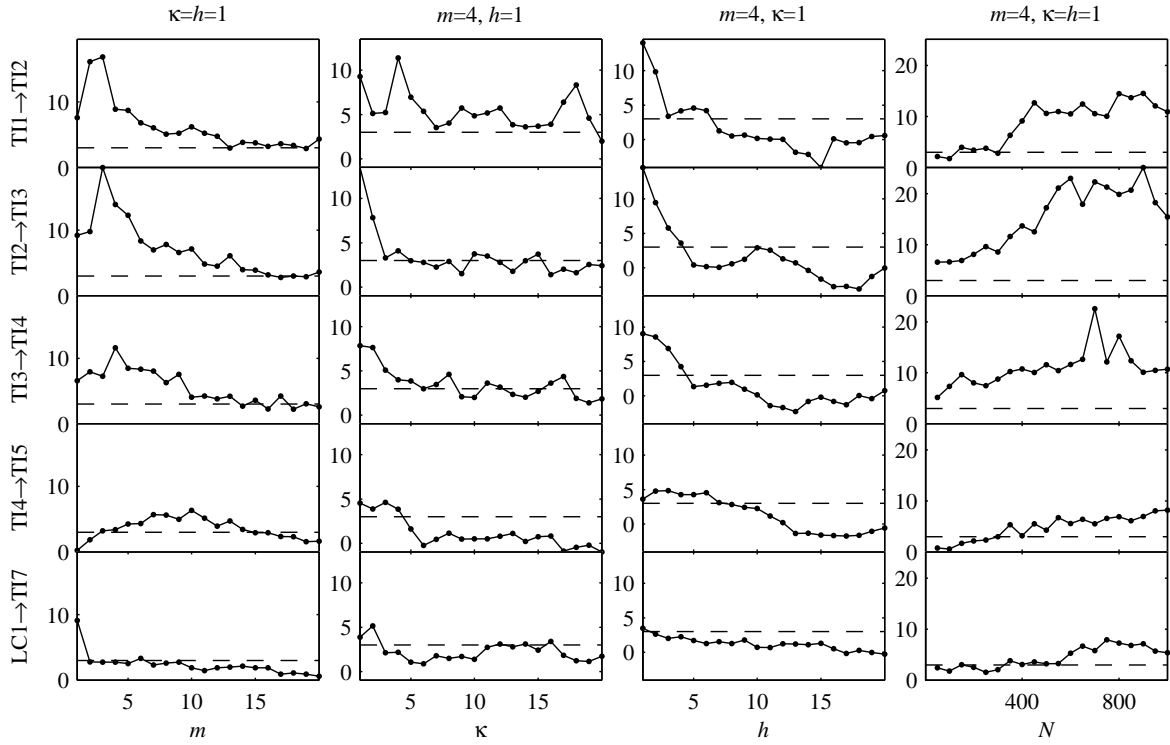
Figure 7: Parameter setting: maximizing significance level $\Omega_{X \to Y}$ of nearest neighbors algorithm for the directionality measures of Case Study I ($N = 500$, $K = 25$, $N_{\text{surr}} = 20$).

then investigated using the best values of $m$, $\kappa$ and $h$. The results of this procedure are shown in Figures 7 and 8 which show the significance levels $\Omega_{X \to Y}$ for Case Studies I and II.

## 4.1 Embedding Dimension $m$

In the first step of the parameter setting procedure, the embedding dimension is varied from 1 to 20 while the embedding delay and prediction horizon are fixed to $\kappa = h = 1$. The result for the two case studies can be seen in the left hand columns of Figures 7 and 8. The result of a directionality calculation is accepted if the significance level $\Omega_{X \to Y}$ is above the threshold of 3, as discussed in section 2.3 and indicated by the dashed lines. For Case Study I, an intuitive choice of $m$ would be to
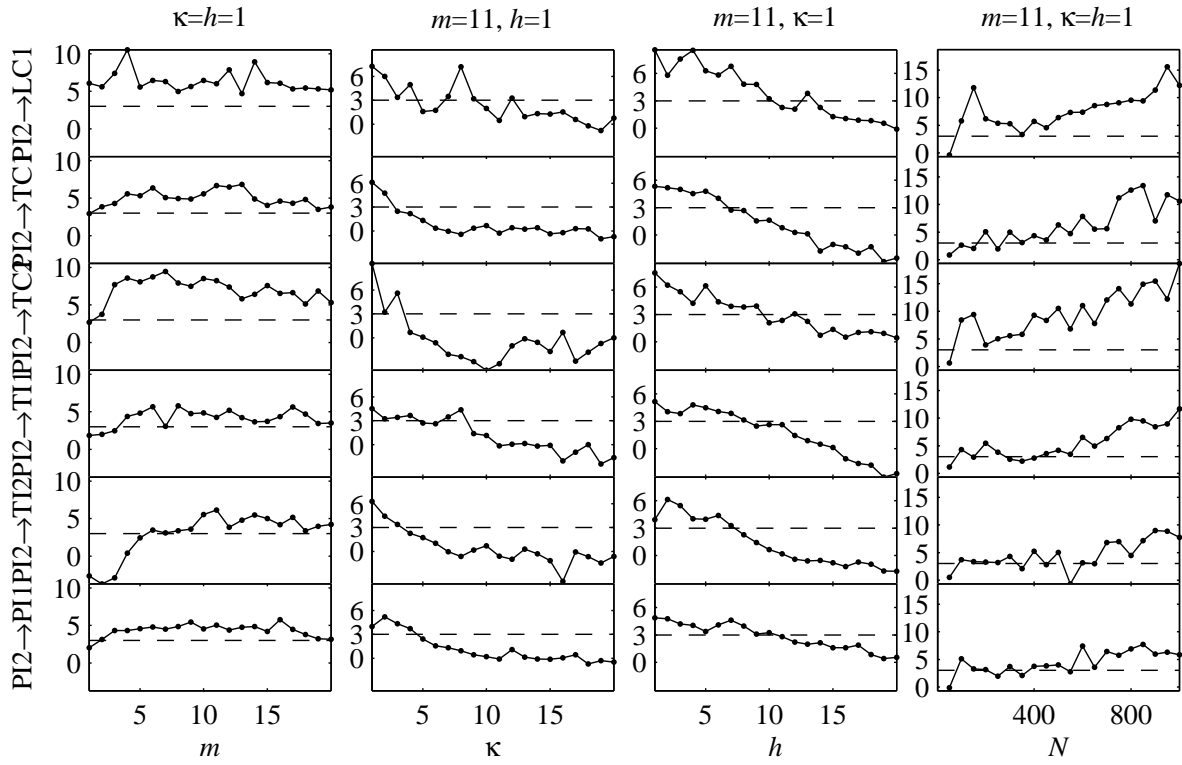
Figure 8: Parameter setting: maximizing significance level $\Omega_{X \to Y}$ of nearest neighbors algorithm for the directionality measures of Case Study II ($N = 500$, $K = 8$, $N_{\text{surr}} = 20$).

set it to the oscillation period which is 20 samples ($m = 20$). The experimental results, however, do not confirm this choice. Instead, the significance level is highest for smaller embedding dimensions at around $m = 4$ indicating that the directionality result is most robust when predicting small stretches of the oscillation. Case Study II gives a different result. The duration of the irregular sharp spikes of the disturbance is around 11 samples and the intuitive choice of $m$ would be the length of the duration of the irregular feature, that is, $m = 11$. The plots in the left hand column of Figure 8 show that the significance level is consistently large if $m$ is between 4 and 14. The maximum value of the average of significance levels is at $m = 11$. The recommended choice of $m$ for the irregular disturbances is therefore the length of the irregular features.

## 4.2    Embedding Delay $\kappa$

Once the embedding dimension is fixed, the embedding delay $\kappa$ is varied in the second step of the parameter selection procedure. The results are shown in the second columns from the left of Figures 7 and 8 in which $\kappa$ is varied from 1 to 20. For Case Study I, the embedding dimension is set to $m = 4$ while the prediction horizon is still kept at $h = 1$. The threshold of $\Omega_{X \to Y}$ is only exceeded for all directionality relationships if $\kappa$ equals to 1 or 2 samples. Best results are obtained if $\kappa = 1$. This suggests that short term dynamics are most relevant for contributing to the directionality result. For Case Study II, $m$ is set to 11 samples while $h$ is kept at 1 sample. The results look similar to Case Study I, that is, the significance only exceeds the threshold for small value of $\kappa$. The best value is again at $\kappa = 1$.

## 4.3    Prediction Horizon $h$

After finding values for embedding dimension and embedding delay, the prediction horizon $h$ is varied in the third step of the parameter selection procedure. The results are shown in the second column from the right of Figures 7 and 8. The significance level shows a similar relationship to the prediction horizon as it did to the embedding delay $\kappa$. As a general trend for both case

studies, $\Omega_{X \to Y}$ decreases with prediction horizon $h$. The largest value of the significance level for all relationships in both case studies, with the exception of PI2→TI2 in Case Study II, is at $h = 1$. Thus, the prediction horizon is set to $h = 1$.

## 4.4   Minimum Number of Samples $N$

As a general rule, the directionality measure $H_{X \to Y}$ and its significance $\Omega_{X \to Y}$ increase with increasing number of samples used for the computation. The reason for this is that a good predictor can find better prediction values because the number of similar near neighbors grows while the poor predictor remains unchanged. The number of samples $N$ was varied in order to find the value $N_{min}$ above which the significance level is robustly above the 3 threshold. For this purpose, the significance level is computed for a number of data sub-sets constructed from the original sequence, and averaged over these sub-sets. The result is shown in the right hand columns of Figures 7 and 8 for the two case studies. The results indicate the threshold value is reached at about $N_{min} = 400$. $N > 400$ is therefore the recommended value. As the computational effort increases significantly with the number of samples, a compromise must be made depending on the application. If the number of samples is too low, however, not all directional dependencies in the investigated process might be detected. For instance, the significance levels in the lowest panel of the right hand column of Figure 7 is on the threshold if $N = 400$.

## 4.5   Parameter Summary

The guideline parameters used for the directionality measure and which result from the parameter setting procedure in this section are summarized in Table 1.

| | |
|---|---|
| Embedding dimension | $m = 4$ (oscillatory) |
| | $m = N_p$ (non-oscillatory) |
| embedding delay | $\kappa = 1$ |
| Prediction horizon | $h = 1$ |
| Number of samples | $N > 400$ |
| Number of nearest neighbors | $K =$ the number of peaks or cycles |
| Number of surrogates | $N_{surr} = 20$ |

Table 1: Guideline parameters for nearest neighbors one-step-ahead-prediction algorithm. $N_p$ is the duration of the non-oscillatory irregular disturbance.

# 5    Results

The directionality measure $H_{X \to Y}$ is now applied to the relationships for all combinations of process measurements in Case Studies I and II to determine the fault propagation paths in the two processes.

For $p$ variables, $\frac{p(p-1)}{2}$ values of the directionality measure are computed representing the relationship between all combinations of these variables. The results are visualized in a chart that places all potential causes on the vertical axis and all potential effects on the horizontal axis. If a directional relationship between PV1 and PV2 is established then an entry is made at the intersection of PV1 on the vertical axis and PV2 on the horizontal axis. The entries are circles whose sizes represent the relative values of the directionality measure $H_{X \to Y}$ for those $H_{X \to Y}$ values which are statistically significant. The resulting graph is referred to as bubble plot. The root cause is then taken to be the variable that causes most other variables while not being the effect of any other variable. Thus, the column associated in the bubble graph with the root cause variable will have no entry while its row will have one or more entries. Algorithms for the construction and manipulation of the bubble plots are discussed in [31].
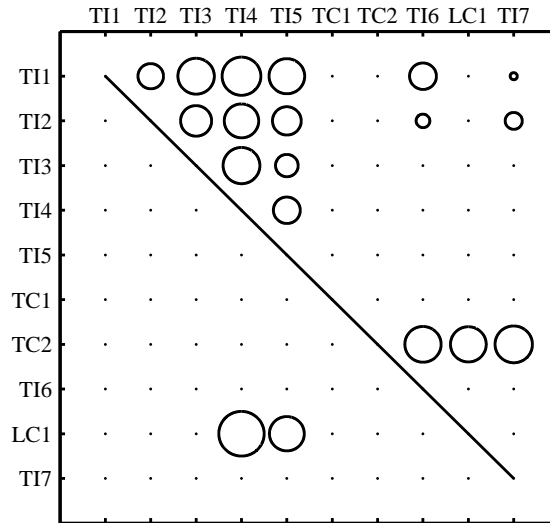
Figure 9: Results of Case Study I, directionality measure $H_{X \to Y}$ with significance level above the threshold for $m = 4$, $\kappa = h = 1$, $N = 1000$, $K = 25$, $N_{surr} = 20$.

## 5.1 Case Study I

Figure 9 shows the bubble plot for all process measurements in Case Study I. Most entries occur above the main diagonal. TI1, the temperature at the top of the distillation column in Figure 3 has no column entry but six row entries and is therefore most likely to be close to the root cause because it influences the highest number of other variables. A further conclusion from the bubble plot is the order of occurrences of the fault. The disturbance happens first in TI1 which then propagates to TI2, TI2 propagates to TI3, TI3 propagates to TI4 and TI4 propagates to TI5. Also, TC2 propagates to the level in the bottom tray of the distillation column, LC1, which means that the level is upset because of the disturbance in the temperature in the column and not vice versa. These findings match the expectations for Case Study I that were outlined at the end of Section 3.1.

The temperature TC2 causes the temperature further downstream in the process (TI7) to be upset. An additional detected directional dependence is the relationship between LC1 and TI4, TI5
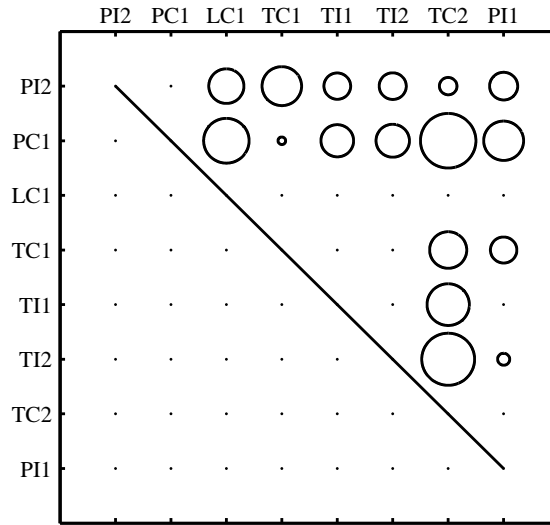
23

Figure 10: Results of Case Study II, directionality measure $H_{X \to Y}$ with significance level above the threshold for $m = 11$, $\kappa = h = 1$, $N = 1000$, $K = 8$, $N_{surr} = 20$.

which can be seen as a pair of bubbles in the lower part of Figure 9. When examining the time series in Figure 4 in detail, flat stretches in LC1 can be observed. This could be explained by a further disturbance such as a valve being stuck to a limit value. The results suggest the directionality measure is capturing the effects of two different disturbances.

## 5.2 Case Study II

The results for Case Study II are shown in Figure 10. For this case study, all detected dependencies are above the main diagonal. Pressure PI2 and PC1 appear to be closest to the root cause since no other variable influences these, and they influence all the other variables. The order of the other events is ambiguous since, for example, no dependency is detected among the group of variables TC1, TI1 and TI2. It is, however, clear that TC2, the controlled temperature at the reactor column and PI1, the pressure at the inflow of the re-boiler, occur last in the chain of events because they are influenced by many other variables.

The known root cause of Case Study II was given in Sections 3.2. The directionality results

presented here point to the known cause, suggesting that the nearest neighbors analysis of measurements from routine operation is capable of determining the source and direction of propagation of the plant-wide disturbance.

# 6  Conclusions

Plant-wide disturbances in a process affect many measurements and thus pose the problem of identifying the source of the disturbance. This paper has shown how the fault propagation path can be traced through the plant from the root cause by detecting interdependency between the time series of measurements from the process and determining the directionality of the relationship. The performance of the method, which is based on nearest neighbors, was investigated using real data from two industrial processes at Eastman Chemical Company. Guidelines for the selection of the parameters were developed through the application to selected variables of the industrial data and the consecutive variation of the parameters. The directionality measure was then applied to the full case studies and shown to be effective in identifying the known root causes and the fault propagation paths.

# References

[1] Desborough, L.; Miller, R. Increasing customer value of industrial control performance monitoring - Honeywell's experience. *AIChE Symposium Series No 326* **2002**, 98, 153-186.

[2] Choudhury, M.A.A.S.; Shah, S.L.; Thornhill, N.F. Diagnosis of poor control loop performance using higher order statistics. *Automatica* **2003**, 40, 1719-1728.

[3] Horch, A. Condition Monitoring of Control Loops. PhD Thesis, Royal Institute of Technology, Stockholm, 2000.

[4] Rengaswamy R.; Hägglund, T.; Venkatasubramanian, V. A qualitative shape analysis formalism for monitoring control loop performances, *Engineering Applications of Artificial Intelligence* **2001**, 14, 23-33.

[5] Harris, T.J. Assessment of control loop performance. *The Canadian Journal of Chemical Engineering* **1989**, 67, 856-861.

[6] Chiang, L.H. Fault Detection and Diagnosis for Large-Scale Systems. PhD Thesis, University of Illinois, 2001.

[7] Thornhill, N.F.; Horch, A. Advances and new directions in plant-wide controller performance assessment. *Control Engineering Practice*, **2007**, in press.

[8] Paulonis, M.A.; Cox, J.W. A practical approach for large-scale controller performance assessment, diagnosis, and improvement, *Journal of Process Control* **2003** , 13, 155-168.

[9] Le Van Quyen, M.; Dama, C.; Baulac, M.; Martenierie, J.; Varela, F.J. Nonlinear analysis of interictal EEG map the brain interdependencies in human focal epilepsy. *Brain Research* **1998**, 792, 24-40.

[10] Wiesenfeldt, M.; Parlitz U.; Lauterborn, W. Mixed state analysis of multivariable time series. *International Journal of Bifurcation and Chaos* **2001**, 11, 2217-2226.

[11] Schiff, S.J.; So, P.; Chang, T.; Burke, R.E.; Sauer, T. Detecting dynamical interdependence and generalized synchrony through mutual prediction in a neural ensemble. *Physical Review E* **1996**, 54, 6708-6724.

[12] Feldmann, U.; Bhattacharya, J. Predictability improvement as an asymmetrical measure of interdependence in bivariate time series. *International Journal of Bifurcation and Chaos* **2004**, 14, 505-514.

[13] Bhattacharya, J; Pereda, E.; Petsche, H. Effective detection of coupling in short and noisy bivariate data. *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics* **2003**, 33, 85-95.

[14] Quian Quiroga, R.; Arnhold J.; Grassberger, P. Learning driver-response relationships from synchronization patterns. *Physical Review E* **2000**, 61, 5142-5148.

[15] Arnhold, J.; Grassberger, P.; Lehnertz, K.; Elger, C.E. A robust method for detecting interdependences: Application to intracranially recorded EEG, *Physica D* **1999**, 134, 419-430.

[16] Kocarev, L.; Parlitz, U. Generalized synchronization, predictability and equivalence of unidirectional coupled dynamical systems. *Physical Review Letters* **1996**, 76, 1816-1819.

[17] Bhattacharya J.; Petsche, H.; Pereda, E. Long-range synchrony in the gamma band - role in the music perception. *Journal of Neuroscience* **2001**, 21, 6329-6337.

[18] Rulkov, N.F.; Sushchik, M.M.; Tsimring, L.S.; Arbandel, H.D.I. Generalized synchronization of chaos in directionally coupled chaotic systems. *Physical Review E* **1995**, 51, 980-994.

[19] Maurya, M.R.; Rengaswamy, R.; Venkatasubramanian, V. A systematic framework for the development and analysis of signed digraphs for chemical processes. 1. Algorithms and analysis. *Industrial & Engineering Chemistry Research* **2003**, 42, 4789-4810.

[20] Maurya, M.R.; Rengaswamy, R.; Venkatasubramanian, V. A systematic framework for the development and analysis of signed digraphs for chemical processes. 2. Control loops and flowheet analysis. *Industrial & Engineering Chemistry Research* **2003**, 42, 4811-4827.

[21] Bauer, M.; Cox, J.W.; Thornhill, N.F. Measuring cause and effect of process variables. *IEEE Advanced Process Control Applications for Industry Workshop* **2005**, Vancouver, Canada.

[22] Thornhill, N.F. Finding the source of nonlinearity in a process with plant-wide oscillation, *IEEE Transactions on Control Systems Technology* **2005**, 13, 434-443.

[23] Sauer, T.; Yorke, N.A.; Casdagli, M. Embedology. *Journal of Statistical Physics* **1991**, 65, 579-616.

[24] Schmitz, A. Measuring statistical dependence and coupling of subsystems. *Physical Review E* **2000**, 62, 7508-7511.

[25] Theiler, J.; Eubank, S.; Longtin, A.; Galdrikian, B. Farmer, J.D. Testing for nonlinearity in time series - the method of surrogate data. *Physica D* **1992**, 58, 77-94.

[26] Kantz, H.; Schreiber, T. *Nonlinear Time Series Analysis.* Cambridge University Press: Cambridge, 1997

[27] Hegger, R.; Kantz, H.; Schreiber, T. *TISEAN 2.1 Surrogates Manual, Periodicity Artefacts*, **2000**, Retrieved Apr 4th 2007 from $http : //www.mpipks - dresden.mpg.de/$~tisean/TISEAN$_2.1/index.html$.

[28] Pereda, E.;, Rial, R.; Gamundi, A.; Gonzalez, J.; Assessment of changing interdependencies between human electroencephalograms using nonlinear methods, *Physica D* **2001**, 148, 147-158.

[29] Andrzejak, RG.; Kraskov, A.; Stogbauer, H.; Mormann, F.; Kreuz, T.; Bivariate surrogate techniques: Necessity, strengths, and caveats. *Physical Review E* **2003**, 68, Article 066202, 1-15.

[30] Pereda, E.;, Quian Quiroga, R.; Bhattacharya, J. Nonlinear multivariate analysis of neurophysiological signals. *Progress in Neurobiology* **2005**, 77, 1-37.

[31] Bauer, M.; Cox, J.W.; Caveness, M.H.; Downs, J.J.; Thornhill, N.F. Finding the Direction of Disturbance Propagation in a Chemical Process using Transfer Entropy. *IEEE Transactions on Control Systems Technology* **2007**, 15, 12-21.

# Figure Captions

Figure 1: Find the nearest neighbors of $\mathbf{x}_i$ and $\mathbf{y}_i$ to predict future values $x_{i+h}$ and $y_{i+h}$.

Figure 2: Construction of the directionality measure from $D_i(X|Y)/D_i(X)$ and $D_i(Y|X)/D_i(Y)$ for the Hénon map, $a = 1.4$, $b = 0.3$. ($m = 5$, $\kappa = h = 1$, $K = 10$, $N = 200$).

Figure 3: Process schematic of Case Study I.

Figure 4: Segment of the process measurements of Case Study I, with nonlinearity index.

Figure 5: Process schematic of Case Study II.

Figure 6: Process measurements for Case Study II, with nonlinearity index.

Figure 7: Parameter setting: maximizing significance level $\Omega_{X \rightarrow Y}$ of nearest neighbors algorithm for the directionality measures of Case Study I ($N = 500$, $K = 25$, $N_{\text{surr}} = 20$).

Figure 8: Parameter setting: maximizing significance level $\Omega_{X \rightarrow Y}$ of nearest neighbors algorithm for the directionality measures of Case Study II ($N = 500$, $K = 8$, $N_{\text{surr}} = 20$).

Figure 9: Results of Case Study I, directionality measure $H_{X \rightarrow Y}$ with significance level above the threshold for $m = 4$, $\kappa = h = 1$, $N = 1000$, $K = 25$, $N_{surr} = 20$.

Figure 10: Results of Case Study II, directionality measure $H_{X \rightarrow Y}$ with significance level above the threshold for $m = 11$, $\kappa = h = 1$, $N = 1000$, $K = 8$, $N_{surr} = 20$.

# Table Caption

Table 1: Guideline parameters for nearest neighbors one-step-ahead-prediction algorithm. $N_p$ is the duration of the non-oscillatory irregular disturbance.