Causality in Control Systems Based on Data-Driven Oscillation Identification

Michał J. Falkowski 1, Paweł D. Domański 1,∗ and Ewa Pawłuszewicz 2

1 Institute of Control and Computation Engineering, Faculty of Electronics and Information Technology, Warsaw University of Technology, Nowowiejska 15/19, 00-665 Warsaw, Poland; michal.falkowski.dokt@pw.edu.pl
2 Faculty of Mechanical Engineering, Białystok University of Technology, Ul. Wiejska 45C, 15-351 Białystok, Poland; e.pawluszewicz@pb.edu.pl
* Correspondence: pawel.domanski@pw.edu.pl

Abstract: This paper addresses the subject of causality analysis using simulation data and data collected from a real control system. Simulated data includes Gaussian and Cauchy noise signals. Real-time series include various, mostly unknown distortions, like trends, oscillations, and noises. Presented research focuses on the oscillatory component in data and its propagation in multi-loop control systems. Oscillation identification is based on a deep decomposition process for control error time series. Identified periodic signals are used for further causality processing. The analysis uses the Transfer Entropy approach. This method belongs to the group of model-free methods. The determination of information pathways is conducted without any model or a priori process knowledge. The research investigates the impact of the oscillation time-series component on the Transfer Entropy causality analysis. The summary shows the observations obtained for given simulated datasets and those collected from real processes. The obtained results show that simulated analysis works properly. On the contrary, the direct application of the oscillation decomposition in real industrial cases may be misleading. Large datasets demand modification in the methodology. Different variants are tested. They show that oscillation propagation is biased in real systems and, therefore, the decomposition should be applied with caution. Furthermore, it is important to remember that the algorithm transition from simulated data to real industrial ones is demanding and should be done with the utmost care.

Keywords: causality analysis; Transfer Entropy; faults propagation; oscillations identification; MEEMD; statistical analysis; large-scale industrial processes; median filter

1. Introduction

Causality analysis requires a specific approach and designated methods that allow one to determine the internal relations between different variables [1]. One may find various approaches, which have been developed in different domains. Actually, there are four leading approaches to detecting causality from historical process data: Granger Causality, Transfer Entropy, Cross-correlation, and Partial Directed Coherence [2]. Cross-correlation is simple, but it assumes linearity [3]. Some of them are based on the process model, like the coherence method [4], the Granger causality [5] and its extension Partial Directed Coherence. Spectral Granger causality uses Fourier Transform, which limits the method to the linear and stationary cases. Others belong to the group of data-driven and model-free methods, like the Transfer Entropy approach [6]. The analysis of complex and varying control systems, for which the exact model is unknown or its estimation is highly uncertain, encourages the use of the data-driven method. It is a great simplification in the field of causality determination as well as fault propagation research. Faults can be observed as system oscillations, among others.
The industrial demands for the causality analysis have been defined quite early [7], but the research on the causality root-cause analysis in the area of process control is relatively new and only few works are published. The introductory paper [2] has been followed by some examples of the Granger causality and Transfer Entropy comparison in [8,9]. Transfer Entropy research has been followed in [10] with the hybrid combination of TE and the modified conditional mutual information (CMI) approach, which uses not data time series but multi-valued alarm series. Authors of [11] used an opposite approach. They are assuming that periodical elements are not bringing any value so they use denoising and periodicity-removing time-delayed convergent cross mapping (TD-CCM), which is an extension of the original algorithm [12] and its time-delayed version [13].

The Transfer Entropy approach has been addressed by a few works. The Transfer Entropy systematic workflow for oscillation diagnosis is proposed in [14], while the outliers impact is addressed in [15]. Finally, the TE approach is improved by information granulation in [16].

AI-driven methods, like neural or Bayesian networks, are uncommon in the process industry applications. They are many times harder to train because of the increasing size and complexity of the used data sets [11]. Moreover, we would need to know the causality examples to train it. The authors decided not to use AI and develop the TE approach.

Our approach assumes that periodicity contains important information and propagates through the system. Therefore, the Transfer Entropy approach uses decomposed oscillation elements.

An example of the oscillatory behavior propagation through the control system is discussed in [17]. Oscillation diagnosis may allow one to trace the root cause of these faults. It allows answering the question of which control loop of a given system introduces disturbances into the system causing its improper performance. As the result, too aggressive or too sluggish operation [18] is observed. Thus, assuming that the oscillations in a multi-loop control system carry out valuable information, it is imperative to conduct the research in such a direction. In the considered analysis, the innovative application of the Transfer Entropy approach is combined with the oscillation decomposition leading toward time-frequency causality analysis.

The algorithm should work not only in the simulation or laboratory environment but also in industrial reality, which introduces various disturbances and limitations. Non-stationarity and non-Gaussian properties are common. They often turn out in the form of the variables distribution tails and outliers. One has to be aware that outlying observations are very frequent in industrial practice [19]. They may deteriorate statistical estimation and regression analysis [20] and also oscillation decomposition [21]. This research takes this aspect into account as well trying to omit the problem through the subsets decomposition and median filtering. Our contribution consists of two parts. First, we show how the performance of decomposition methods can be extended for very long time series. Second, real-world examples have shown the inadequacy of existing oscillation decomposition methods for industrial data. This observation opens a new direction for further research.

Thus, assuming that the oscillations in a multi-loop control system carry out valuable information, it is imperative to conduct the research in such a direction. In the considered analysis, the innovative application of the Transfer Entropy approach is combined with the oscillation decomposition leading toward time-frequency causality analysis. The combination of the Transfer Entropy method with frequency analysis has not been used either with simulation or real data, however, it may be a great simplification in the field of diagnosis of fault propagation methods based on a data-driven model-free approach.

Once the causality graph is identified the tracing of the misbehavior is allowed in the multi-loop configurations. Frequently the poor tuning of a single loop causes low performance for the rest of the process. Thus, the identification of such loops may help control engineers in the process of multi-loop system tuning and fault diagnosis. Moreover, the causality graphs may help to properly organize the intelligent alarming system protecting from multiple alarming making the installation operator’s job easier.
The paper is organized as follows: Section 2 contains an overview of the applied methods and algorithms, while Section 3 describes the used simulation and industrial case study. Section 4 presents obtained results including causality diagrams reflecting relationships between variables based on oscillation signals. Section 5 concludes the paper and identifies open issues for further research.

2. Proposed Methodology

It is shown that the diagnosis of large-scale industrial control systems, especially the creation of its good representation by a mathematical model, is a tedious process and its accuracy depends on many factors [22]. Depending on the adopted methodology, analysis can be based on various types of data, ranging from data measured directly from the given control loops (raw datasets), through the noise, to oscillations occurring in almost every system [15]. The last-mentioned type of data is strongly related to Time-Frequency (T-F) analysis, which is a powerful tool and can be successfully used to determine interconnections between variables and units (causality). These relations represent the information flow (caused by control loops) in the process and can be “measured” thanks to the Transfer Entropy (TE) approach. The combination of T-F analysis with TE has not been previously observed in the literature, in particular in industry-related applications.

Proposed a median ensembled version of EEMD (MEEMD) that belongs to a class of noise-assisted EMD methods, helps to decompose signals into oscillations required for further analysis. This approach is used together with specific statistical data handling. The analysis is applied to two case studies: the known and simulated environment and the unknown industrial case.

2.1. Transfer Entropy Approach

Transfer Entropy (TE) is an information-theoretic interpretation of Wiener’s causality definition. In practice, this is a measure of information transfer from one variable to another by measuring the reduction of uncertainty while assuming predictability. TE is given by the Equation (1).

\[ T_{x \rightarrow y} = \sum_{y_{i+h},X_j} p(y_{i+h},Y_i,X_j) \cdot \log \frac{p(y_{i+h}|Y_i,X_j)}{p(y_{i+h}|Y_i)} \] (1)

where \( p \) means the complete or conditional Probability Density Function (PDF), variables \( Y_i \) and \( X_i \) respectively as \( Y_i = [y_i, y_{i−τ}, . . . , y_{i−(l−1)τ}] \), \( X_i = [x_i, x_{i−τ}, . . . , x_{i−(k−1)τ}] \), \( τ \) is a sampling, \( h \) is a prediction horizon. Concluding, it is the difference between an information about a future observation of \( x \) obtained from the simultaneous observation of past values of both \( x \) and \( y \), and the information about the future of \( x \) obtained from the past values of \( x \) alone. Phenomenon of entropy in both directions is highly probable that is why a measure described as \( T_{x \rightarrow y} = T_{y|x} - T_{y|y} \) is decisive due to quantity and direction, what is causality.

The practical implementation of the Transfer Entropy approach between pair of variables according to Equation (1) requires its simplification to the form presented in Equation (2):

\[ T_{x \rightarrow y} = \sum_{y_{i+t},X_j} p(y_{i+t},X_{i−τ}) \cdot \log \frac{p(y_{i+t},X_{i−τ})}{p(y_{i−τ},X_{i−τ})} \] (2)

where \( p \) means the conditional (PDF), \( τ \) and \( t \) are the time lags in \( x \) and \( y \), respectively. If the time-series length is short, the \( t \) is set to 1 under the assumption that the maximum auto-transfer of information occurs from the data point immediately before the target value in \( y \).

2.2. Median Ensemble Empirical Mode Decomposition

Median Ensemble Empirical Mode Decomposition (MEEMD) is a variation of the EEMD algorithm that uses the median operator instead of the mean operator to ensemble noisy Intrinsic Mode Function (IMF) trials [23]. The use of this algorithm is a practical exten-

---

**Appl. Sci. 2022, 12, 3784**  
3 of 21
sion of the classic EMD and justified choice with real-world applications. The EMD method was developed so that data can be examined in an adaptive time–frequency–amplitude space for nonlinear and non-stationary signals [24]. It decomposes the input signal into a few Intrinsic Mode Functions and a residue. The given equation will be as follows:

\[ I(n) = \sum_{m=1}^{M} IMF_m(n) + Res_M(n), \]  

where \( I(n) \) is the multi-component signal, \( IMF_m(n) \) is the \( m \)th intrinsic mode function, and \( Res_M(n) \) represents the residue corresponding to \( M \) intrinsic modes. The proposed median EEMD (MEEMD) algorithm, defines median operator as:

\[
\text{median} = \begin{cases} 
\text{IMF}(t)\left(\frac{N+1}{2}\right) & \text{when } k \text{ is odd}, \\
\frac{\text{IMF}(t)\left(\frac{N}{2}\right) + \text{IMF}(t)\left(\frac{N+1}{2}\right)}{2} & \text{when } k \text{ is even}, 
\end{cases}
\]

where \( \text{IMF}(t) \) denotes the ordered IMF list at time instant \( t \), which is obtained from \( N \) independent noise realizations. Consider a real-valued signal \( x(t) \) and a predefined noise amplitude \( \epsilon \), the used MEEMD is outlined in Algorithm 1.

**Algorithm 1** Algorithm of MEEMD

1. Generate the ensemble \( y_n(t) = x(t) + \epsilon \omega_n(t) \) for \( n = 1, \ldots, N \), where \( \omega_n(t) \sim N(0,1) \);
2. Decompose every member of \( y_n(t) \) into \( M_n \) IMFs using the standard EMD, to yield the set \( \{d_{m}^{n}(t)\}_{m=1}^{M_n} \);
3. Assemble same-index IMFs across the ensemble using the median operator to obtain the final IMFs within MEEMD; for instance, the \( m \)th IMF is computed as \( d_{m} = \text{median}(d_{1}^{m}(t), d_{2}^{m}(t), \ldots, d_{N}^{m}(t)) \).

Nevertheless, EMD, and thus its derivative MEEMD has limitations. In particular the occurrence of dependencies between the length of analyzed data and the number of received IMFs resulting from such a decomposition process.

2.3. Statistical Analysis

Actually one should remember that industrial datasets can include many observations and therefore the constraint of the oscillation detection algorithm (datasets cannot be too long) might limit the analysis. Splitting of the long datasets into smaller subsets is only one of the solutions to overcome a limitation of the MEEMD method described in Section 2.2. Statistical analysis suggests that it is natural to use an arithmetic mean. Analyzing datasets/signal values based on the arithmetic mean is not always producing accurate results. It is known that using the arithmetic mean, adequate (true for a given sample) information is obtained only when the sample comes from a population with a normal distribution or close to it. On the other hand, if the population distribution is not close to a normal distribution, the median is used to infer the sample mean value. If industrial data is considered as a time series, the question arises of whether the generated noise can be smoothed using multi-period smoothing (averaging filters). However, with their construction, some information is lost, as a result of which the extremes of the smoothed signal may appear in inappropriate/illogical places, contrary to real conditions. As a result, the information obtained would be highly incorrect or may even be distorted. Previous papers show that raw datasets collected from real-scale industrial processes do not have normal distributions. Considering the above, the use of a median filter is justified.

Several consecutive samples are taken and their median is determined in the classical median filter. A modified sample is obtained as a result of this operation. Working with datasets of a given amount is still an insufficient solution for further analysis and inference. Therefore, such an \( n \)-element dataset is divided into \( i \)-th windows with a length corresponding to natural divisors of number \( n \) (in this case 10, 25, 50, and 100, respectively).
In each window, we have \( k \) samples, which are arranged in a non-decreasing manner, and for each already ordered window we determine the median as follows [25]:

\[
Me = \begin{cases} 
\frac{x_{(k+1)/2}}{2} & \text{when } k \text{ is odd}, \\
\frac{x_{(k/2)} + x_{(k/2)+1}}{2} & \text{when } k \text{ is even}, 
\end{cases}
\]

where \( x_k \) is the value of the \( i \)-th dataset. The position of the median (i.e., the middle unit) in each of the windows is set at half the size of the window under consideration:

\[
N_{Me} = \frac{k}{2}.
\]

In this way, based on the medians from each of the windows, a shorter, filtered sample is obtained.

2.4. The Method

The methodology of the analysis depends on the type of data, especially on the length of the time series. Simulation data presented in Figure 1a may require a simpler approach as no additional data preprocessing is required. In our case, they are disturbed by randomly generated Gaussian measurement noise and optionally by the Cauchy disturbance. Oscillations are introduced using added sinusoidal signal of frequency equal to 60 Hz, not exceeding 10% of the mean value of the data. Real process data consists of longer time series. Therefore, they require preprocessing that is proposed to be done in two separate variants, as in Figure 1b.

![Flow charts presenting the analysis procedure for (a) simulation and (b) industrial data.](image)

Both approaches can be summarized in a single algorithm. The methodology is described by the algorithm Algorithm 2. It uses acquired process data representing control errors for \( M \) loops of length equal to \( N \) observations \( X_k \in \mathbb{R}^{M \times N} \). In the proposed analysis the suggested value for the maximum time-series length to work without decomposition is set to \( N_{\text{max}} = 2000 \).
Algorithm 2 TE algorithm with oscillation decomposition

Acquire data time series $X_k$, $k = 1, \ldots, N$

if $N < N_{\text{max}}$ then
    Decompose $M$ loops with MEEMD to IMF functions $M_k$ \(\triangleright\) Use Algorithm 1
    Calculate TE coefficients $d_i$ for each pair of IMFs $M_k$ \(\triangleright\) Use Equation (2)
else
    Detrend data using polynomial/spline interpolation
    Select splitting variant
    Split detrended data into $n$ datasets of length $N < N_{\text{max}}$
    Decompose each subset data with MEEMD to $M_k = \text{round}(\log(N))$ \(\triangleright\) Use Algorithm 1
    Calculate TE coefficients $d_i$ for corresponding pairs of IMFs $M_k$ \(\triangleright\) Use Equation (2)
end if

Draw causality diagram using $d_i$ coefficients

3. Research Case Studies

Both simulation environment and industrial data are summarized below.

3.1. Simulation Data

To generate data for analysis, an artificial multi-loop simulation model is created in MATLAB Simulink® environment. The model scheme is presented in Figure 2. This model consists of five control loops—four PID controllers ($R_1$, $R_2$, $R_3$, and $R_4$) and one PI ($R_5$) controller. The control errors from the loops, marked in the scheme as $\epsilon_1$, $\epsilon_2$, $\epsilon_3$, $\epsilon_4$, and $\epsilon_5$, respectively, are the basis for the problem analysis. The model incorporates the possibility to apply noises. In this case, two variants are given—first that simulation data is distorted by Gaussian noise (represented by variable $\sigma_i$) and second that simulation data is distorted by both Gaussian and Cauchy (represented by variable $\gamma_i$) noises. A sinusoidal signal is added to the Cauchy disturbance to simulate potential loop oscillations of known frequency and their propagation.

![Simulation environment presenting multi-loop PID-based control layout.](image)

Simulated process transfer functions are in form of the following linear models:

$$
G_1(s) = \frac{1}{0.15s + 1}, \quad G_2(s) = \frac{0.25(-s + 1)}{s(2s + 1)}, \quad G_3(s) = \frac{1}{(s + 1)(0.04s^2 + 0.04s + 1)},
$$
$$
G_4(s) = \frac{-0.25(-s + 1)}{s(s + 1)(2s + 1)}, \quad G_5(s) = \frac{1}{(0.2s + 1)^2 e^{-s}}. \quad (7)
$$

Feedforward filters are defined as

$$
F_3(s) = \frac{1}{4s^2 + 3.2s + 1}, \quad F_4(s) = \frac{1}{s^2 + 1.2s + 1}, \quad F_5(s) = \frac{1}{4s^2 + 2s + 1},
$$
$$
F_{d3}(s) = \frac{1}{2s + 1}, \quad F_{d4}(s) = \frac{0.1}{s + 1}, \quad F_{d5}(s) = \frac{4}{2s + 1}. \quad (8)
$$
Controllers are roughly tuned with their parameters sketched in Table 1. Disturbance decoupling is done using industrial design [26] with a block diagram sketched in Figure 3.

**Table 1.** Tuning parameters of simulated controllers $R_i$—Matlab PID implementation.

| $i$ | P  | I  | D  | N  |
|-----|----|----|----|----|
| 1   | 0.2| 0.01| 0.02| 10 |
| 2   | 2  | 0.2| –  | 10 |
| 3   | 0.13| 0.5| 0.08| 100|
| 4   | 0.27| 0.61| 0.21| –  |
| 5   | 0.01| 0.2| –  | –  |

**Figure 3.** Layout of the industrial realization of the feedforward disturbance decoupling.

The generated simulation data, including Gaussian noise, is presented in Figure 4. For each $\sigma_i$ variable, the noise parameters are selected randomly. In the case of the second variant, both Gauss and Cauchy disturbances are randomly selected, and the frequency of the added sinusoidal signal equals 30 Hz. The result of the above assumptions is presented in the form of trends in Figure 5. Next, the control error histograms are evaluated (see Figure 6) to check the data’s statistical properties. It is shown that even in the case of only Gaussian noise the loop statistical properties are not Gaussian and they exhibit fat tails. Therefore, robust statistical estimators should be used to protect against misfitting.

Due to the fact that the relationships between the signals from the control loop are known, Figure 7 shows the causality diagram for the simulation model. The solid line means direct relationship, the dashed line means indirect relationship.

**Figure 4.** Control errors time series for the simulation data impacted by the Gaussian noise.
Figure 5. Control errors time series for the simulation data impacted by the Gaussian noise and Cauchy disturbance.

Figure 6. Control errors histograms for the simulation data impacted by the Gaussian noise with PDF fitting. (a) Control error $\epsilon_1$, (b) Control error $\epsilon_2$, (c) Control error $\epsilon_3$, (d) Control error $\epsilon_4$, (e) Control error $\epsilon_5$. 
3.2. Industrial Data Overview

Industrial data used in this part of the research originates from the ammonia production plant, specifically from the ammonia synthesis loop. The process is using the catalytic reaction of the hydrogen, which is produced in earlier stages of production, with nitrogen being obtained from ambient air to produce anhydrous liquid ammonia. The process runs in the so-called synthesis loop, where the synthesis gas circulates and mixes with the fresh gas input. The reaction takes place on the ferric catalyst. The heat, which is obtained, heats up the steam and preheats the gas entering the installation. Included time series of the control errors describe five PID loops controlling the final stage of the ammonia productions, i.e., the synthesis loop. Loops denoted as $u, w, x, y$ represent level control in the separators while the loop $z$ is the temperature control.

Determining causality between variables of any large-scale industrial process requires proper preparation of data for analysis. As processes are mostly complex and non-linear, the data series obtained from such plants are characterized by a non-Gaussian distribution. The data is affected by noise as well as by the outliers appearing during measurements or resulting from external actions, for example, human activity. Basic preprocessing of such data consists primarily of removing the trend that appears in them, as well as the separation of other components important for the analysis of given signals.

Causality detection is performed on oscillations that are obtained through the decomposition process from the initially pre-processed dataset. It contains five separate variables designed as $u, w, x, y,$ and $z$. Initial data pre-processing consists of trends removal based on the most common polynomial interpolation (shifting the signals by a constant value—0 order polynomial) for $u$ to $y$ and 2nd order spline interpolation for $z$. The detrended data for the analysis is presented in Figure 8. Moreover, signals are characterized by quite high stationarity with a clear presence of outliers. Their impact on the quality of results using the Transfer Entropy approach has been tested [27] and may be ignored for the purposes of this paper. Each given variable is a sequence of 59,900 samples, which is crucial information in further calculations.
Similar to the simulation case, Figure 9 presents control error histograms, which are not Gaussian as well, justifying the use of robust estimators.

![Histograms for time series of the industrial data with PDF fitting.](image)

**Figure 9.** Histograms for time series of the industrial data with PDF fitting. (a) Control error $u(k)$, (b) Control error $w(k)$, (c) Control error $x(k)$, (d) Control error $y(k)$, (e) Control error $z(k)$.

### 4. Results

Results presentation is grouped into two parts, i.e., for simulated and real data.

#### 4.1. Results for Simulation Data

Simulation data is decomposed into 7 IMFs for both variants (data length set to 2000 samples for each $\epsilon_i$). In Figure 10a there are results presented for variable $\epsilon_3$ with Gaussian noise added. Oscillations are barely visible in IMFs $d_4, d_5, d_6,$ and $d_7$. IMFs designated as $d_1, d_2,$ and $d_3$ are considered as noise. The situation is different in the case of the second variant—the oscillation phenomenon is much more visible for the variant with Cauchy noise and sinusoidal signal (of known frequency) added (Figure 10b). Due to the above, research is conducted on the last four IMF functions.
Figure 10. Oscillation decomposition sample using MEEMD algorithm for the variable $\epsilon_3$ in simulated data. The first subplot shows original data (Data) and the consecutive plots show decomposed frequencies and the residuum in the last subplot. (a) Gaussian noise case, (b) Gaussian noise, and Cauchy disturbance case.

Tables 2 and 3 show the values of the Transfer Entropy coefficients for simulation data with Gaussian noise and Gaussian and Cauchy noises respectively. An analysis of the determined TE coefficients values for simulation data with Gaussian noise shows that after the decomposition process, mostly NaN or Inf values type appear (considered as 0). According to the calculated results, in this case, it is impossible to determine the causality graph. There can be many reasons for such poor behavior. The most probably is that Gaussian disturbance is too weak and does not stimulate the simulation system sufficiently. The oscillations detected in the decomposition process are negligible, as shown in Figure 10a. Therefore, the Transfer Entropy approach is ineffective and the determined values of the coefficients should be completely rejected.
According to the analysis of TE coefficients for oscillations with Gauss and Cauchy noises, the effects of the calculations are completely different and the causality graph can be successfully determined. It is presented in Figure 11. In the case of all IMF functions for individual $\epsilon_i$, the largest values of the Transfer Entropy coefficients appear for the same

Table 2. Transfer Entropy coefficients for simulated data (Gaussian noise).

| $T_{row\rightarrow column}$ | $\epsilon_1$ | $\epsilon_2$ | $\epsilon_3$ | $\epsilon_4$ | $\epsilon_5$ |
|-----------------------------|--------------|--------------|--------------|--------------|--------------|
| $\epsilon_1$                |              |              |              |              |              |
| $d_4$                       | NA           | 0            | 0            | 0            | 0            |
| $d_5$                       | NA           | 0            | 0            | 0            | 0.0553       |
| $d_6$                       | NA           | 0            | 0            | 0.0508       | 0.0561       |
| $d_7$                       | NA           | 0            | 0            | 0.0779       | 0.0801       |
| $\epsilon_2$               |              |              |              |              |              |
| $d_4$                       | 0            | NA           | 0            | 0            | 0            |
| $d_5$                       | 0            | NA           | 0            | 0            | 0            |
| $d_6$                       | 0            | NA           | 0            | 0.0420       | 0.0607       |
| $d_7$                       | 0.0651       | NA           | 0            | 0.492        | 0.0561       |
| $\epsilon_3$               |              |              |              |              |              |
| $d_4$                       | 0            | 0            | NA           | 0            | 0            |
| $d_5$                       | 0            | 0            | NA           | 0            | 0            |
| $d_6$                       | 0            | 0            | NA           | 0.0390       | 0.0599       |
| $d_7$                       | 0.0697       | 0            | NA           | 0.0528       | 0.0480       |
| $\epsilon_4$               |              |              |              |              |              |
| $d_4$                       | 0            | 0            | 0            | NA           | 0            |
| $d_5$                       | 0            | 0            | 0            | NA           | 0            |
| $d_6$                       | 0.0515       | 0            | 0            | NA           | 0.0563       |
| $d_7$                       | 0.0807       | 0            | 0            | NA           | 0.0597       |
| $\epsilon_5$               |              |              |              |              |              |
| $d_4$                       | 0            | 0            | 0            | 0            | NA           |
| $d_5$                       | 0.0340       | 0            | 0            | 0            | NA           |
| $d_6$                       | 0            | 0            | 0            | 0.0626       | NA           |
| $d_7$                       | 0.0644       | 0            | 0            | 0.0452       | NA           |

Table 3. Transfer Entropy coefficients for simulated data (Gaussian noise and Cauchy disturbance).

| $T_{row\rightarrow column}$ | $\epsilon_1$ | $\epsilon_2$ | $\epsilon_3$ | $\epsilon_4$ | $\epsilon_5$ |
|-----------------------------|--------------|--------------|--------------|--------------|--------------|
| $\epsilon_1$                |              |              |              |              |              |
| $d_4$                       | NA           | 0.1146       | 0.0344       | 0.0242       | 0.0297       |
| $d_5$                       | NA           | 0.1438       | 0.0496       | 0.0495       | 0.0468       |
| $d_6$                       | NA           | 0.1515       | 0.0750       | 0.0643       | 0.0908       |
| $d_7$                       | NA           | 0.1299       | 0.0826       | 0.0730       | 0.1071       |
| $\epsilon_2$               |              |              |              |              |              |
| $d_4$                       | 0.0834       | NA           | 0.0300       | 0.0227       | 0.0251       |
| $d_5$                       | 0.1504       | NA           | 0.0587       | 0.0429       | 0.0536       |
| $d_6$                       | 0.0753       | NA           | 0.0560       | 0.0549       | 0.0646       |
| $d_7$                       | 0.0953       | NA           | 0.0612       | 0.0755       | 0.0940       |
| $\epsilon_3$               |              |              |              |              |              |
| $d_4$                       | 0            | 0            | NA           | 0.0358       | 0.0244       |
| $d_5$                       | 0.0467       | 0.0520       | NA           | 0.0358       | 0.0411       |
| $d_6$                       | 0.0734       | 0.0914       | NA           | 0.0508       | 0.0698       |
| $d_7$                       | 0.0788       | 0.1058       | NA           | 0.0665       | 0.0767       |
| $\epsilon_4$               |              |              |              |              |              |
| $d_4$                       | 0            | 0            | NA           | 0.0233       | 0.0219       |
| $d_5$                       | 0.0429       | 0.0362       | 0.0350       | NA           | 0.0402       |
| $d_6$                       | 0.0701       | 0.0600       | 0.0637       | NA           | 0.0625       |
| $d_7$                       | 0.0958       | 0.0947       | 0.0641       | NA           | 0.0752       |
| $\epsilon_5$               |              |              |              |              |              |
| $d_4$                       | 0            | 0            | 0.0223       | 0.0275       | NA           |
| $d_5$                       | 0.0432       | 0.0518       | 0.0410       | 0.0447       | NA           |
| $d_6$                       | 0.0611       | 0.0709       | 0.0573       | 0.0656       | NA           |
| $d_7$                       | 0.1442       | 0.1657       | 0.0633       | 0.0612       | NA           |
pairs of variables. The expected causality diagram for the simulation model (Figure 7) assumes direct and indirect relations, but focusing on causality for the highest value of TE in a row, such a complicated graph cannot be determined. However, when comparing the two graphs, there are some subtle differences. There is the lack of dependence between the variables $\epsilon_3$ and $\epsilon_1$, and the direction of causality of the variables $\epsilon_5$ and $\epsilon_2$. Considering the second largest value of TE coefficient for the pair $\epsilon_3$ with the other variables, the causality graph would show the relationship between $\epsilon_3$ and $\epsilon_1$, as originally assumed. The inverse, presumed causality between $\epsilon_5$ and $\epsilon_2$ should be considered in further research. Correctly obtained results are the effect of stimulating the simulation system not only with additional Cauchy noise but also with a sinusoidal signal with a known frequency. The causality graph obtained as a result of calculations should be considered correct and sufficient.

![Figure 11. Causality diagram for simulated data—Gaussian noise and Cauchy disturbance.](image)

4.2. Results for Industrial Data

In the case of industrial data, due to the length of time series, each signal from the detrended dataset can be decomposed into multiple IMFs. However, it is unacceptable. In the literature, it can be found that only the first five IMFs are selected for analysis, as the remaining ones being the sum of non-significant IMFs. This assumption is too subjective and unjustified to ignore the fact that there are more IMFs obtained for such a long time series. Whereby an application of the MEEMD algorithm is computationally demanding and the following assumptions for industrial data are stated: first analyzed variant is that the whole dataset is split into 50 (and respectively 49) sub-datasets of equal lengths in two ways ($S_1=\{s_{11}, s_{12}, \ldots\}$ and $S_2=\{s_{21}, s_{22}, \ldots\}$), as it is shown in the Figure 12. Such overlapping of the obtained sub-datasets will allow for checking the potentially emerging frequency bands that are key in given considerations. The second variant is to shorten the dataset by using Median Filter (MF) due to the mentioned enormous number of IMFs and savings resulting from computation time. This operation is widely described in Section 2.3. Such action is aimed not only at comparing variants of oscillations with different frequencies on the same data but also at confronting the statement that the use of a Median Filter will delete (filter) higher frequencies so they are not visible. For simulation data, the above does not apply because it is possible to generate time series of a required length.

![Figure 12. Decomposition of signal $u$ into sub-datasets for the oscillation identification process.](image)
As a result, industrial datasets are divided into 50 (and 49) overlapping sub-datasets successively. Next the MEEMD algorithm is applied to decompose individual signals from the control loops. Sample data for 1st variant is presented in Figure 13. It shows the MEEMD results for a randomly selected pair of overlapping sub-datasets (in this case for signal $y$). It can be noticed that there are six IMFs functions ($d_1$ to $d_6$) for both given sub-dataset lengths in variant 1st. The first three IMFs are considered as noise. The most interesting ones are the IMFs denoted as $d_4$, $d_5$, and $d_6$ for which actual oscillation occurs. Through analysis of Figure 13 it is confirmed that due to this operation emerging frequency bands are not skipped. According to the multiple numbers of sub-datasets both for 1st and 2nd variants, it is impossible to show all the MEEMD algorithm operation results.

![Figure 13](image_url)  
\((a)\) Results for 32 of 50 subdatasets ($d_{132}$), \((b)\) Results for 32 of 49 subdatasets ($d_{233}$).
The use of the median operator is justified through the fact that industrial control errors time series are in the majority non-Gaussian fat-tailed [28]. Therefore, a robust shift estimator is required. Median is the simplest robust estimator. It holds a breakdown point of 50% [20], meaning that half the points must be outlying observations before the median can be moved outside the inlying range and thus it is used.

Due to the large amount of data, it is impossible to illustrate the causality in the form of Transfer Entropy coefficients for each dataset. For both variants, the relationship between every single signal (variable) \((u, w, x, y, z)\) with the others is determined. It is noted that no universal threshold method for the Transfer Entropy approach has been developed. Reminding that each signal is decomposed into 6 IMF functions (the most interesting are \(d_3\) to \(d_6\) oscillations) for 1st variant and from 4 to 7 IMF functions for 2nd variant, it is rational to make the following assumptions: if the value of TE is determined in the form of \(NaN\) or \(Inf\) it equals 0; in other cases, the relationship between the variables is determined based on the highest value in a given row. The results presented in Sections 4.2.1 and 4.2.2 are the totals of cases where the value of the TE coefficient in a row for a given variables pair is the highest.

### 4.2.1. The 1st Variant

Tables 4 and 5 show the sum of cases where the value of the Transfer Entropy coefficient in a row for a given variables pair is the highest respectively for 50 and 49 overlapping sub-datasets. An analysis of the determined TE coefficients values shows that after the decomposition process of individual signals \(u\) to \(z\), only in the case of \(d_4\), \(NaN\) or \(Inf\) values type appear. The TE coefficients for pairs from the \(d_4\) are also clearly greater than those obtained for \(d_5\) or \(d_6\). This suggests the rejection of this dataset, which could affect the shape of the resultant causality graph.

Due to the fact that it is decided to average the values of the TE coefficient for \(d_5\) and \(d_6\), and on this basis to check the relationship between signals, including 0 to the mean (as an assumption for \(NaN\) and \(Inf\)) would clearly distort the causality results. Tables 6 and 7 show the results for this operation.

### Table 4. Sum of cases where the Transfer Entropy coefficient in a row for a given variables pair is the highest—1st variant (50 subdatasets).

| \(T_{row→column}\) | \(u\) | \(w\) | \(x\) | \(y\) | \(z\) |
|-------------------|-----|-----|-----|-----|-----|
| \(d_4\)          | NA  | 5   | 14  | 10  | 21  |
| \(d_5\)          | NA  | 9   | 9   | 16  | 16  |
| \(d_6\)          | NA  | 8   | 4   | 13  | 25  |
| \(d_4\)          | 11  | NA  | 6   | 15  | 8   |
| \(d_5\)          | 17  | NA  | 11  | 7   | 15  |
| \(d_6\)          | 7   | NA  | 4   | 13  | 26  |
| \(d_4\)          | 20  | 4   | NA  | 4   | 22  |
| \(d_5\)          | 11  | 6   | NA  | 9   | 24  |
| \(d_6\)          | 16  | 6   | NA  | 8   | 20  |
| \(d_4\)          | 20  | 9   | 6   | NA  | 15  |
| \(d_5\)          | 13  | 10  | 9   | NA  | 18  |
| \(d_6\)          | 13  | 11  | 8   | NA  | 18  |
| \(d_4\)          | 24  | 6   | 16  | 4   | NA  |
| \(d_5\)          | 12  | 10  | 16  | 12  | NA  |
| \(d_6\)          | 19  | 9   | 6   | 16  | NA  |
Table 5. Sum of cases where the Transfer Entropy coefficient in a row for a given variables pair is the highest—1st variant (49 subdatasets).

| $T_{row\rightarrow column}$ | $u$ | $w$ | $x$ | $y$ | $z$ |
|-----------------------------|-----|-----|-----|-----|-----|
| $u$                         | $d_4$ | NA  | 2   | 13  | 8   | 26 |
| $d_5$                       | NA   | 16  | 8   | 9   | 16  |
| $d_6$                       | NA   | 10  | 6   | 8   | 26  |
| $w$                         | $d_4$ | 11  | NA  | 8   | 9   | 21 |
| $d_5$                       | 10   | NA  | 15  | 8   | 16  |
| $d_6$                       | 13   | NA  | 4   | 12  | 20  |
| $x$                         | $d_4$ | 12  | 3   | NA  | 8   | 26 |
| $d_5$                       | 11   | 7   | NA  | 9   | 22  |
| $d_6$                       | 8    | 7   | NA  | 8   | 26  |
| $y$                         | $d_4$ | 15  | 5   | 8   | NA  | 21 |
| $d_5$                       | 14   | 6   | 5   | NA  | 24  |
| $d_6$                       | 14   | 3   | 8   | NA  | 24  |
| $z$                         | $d_4$ | 17  | 10  | 14  | 8   | NA |
| $d_5$                       | 13   | 6   | 16  | 14  | NA  |
| $d_6$                       | 16   | 11  | 8   | 14  | NA  |

Table 6. Sum of cases where the mean value of the Transfer Entropy coefficient in a row for a given variables pair is the highest—1st variant (50 subdatasets).

| $T_{row\rightarrow column}$ | $u$ | $w$ | $x$ | $y$ | $z$ |
|-----------------------------|-----|-----|-----|-----|-----|
| $u$                         | NA  | 6   | 4   | 15  | 25  |
| $w$                         | 11  | NA  | 7   | 12  | 20  |
| $x$                         | 13  | 8   | NA  | 4   | 25  |
| $y$                         | 11  | 9   | 7   | NA  | 23  |
| $z$                         | 18  | 10  | 10  | 12  | NA  |

Table 7. Sum of cases where the mean value of the Transfer Entropy coefficient in a row for a given variables pair is the highest—1st variant (49 subdatasets).

| $T_{row\rightarrow column}$ | $u$ | $w$ | $x$ | $y$ | $z$ |
|-----------------------------|-----|-----|-----|-----|-----|
| $u$                         | NA  | 11  | 4   | 9   | 25  |
| $w$                         | 12  | NA  | 5   | 12  | 20  |
| $x$                         | 7   | 8   | NA  | 8   | 26  |
| $y$                         | 12  | 2   | 6   | NA  | 29  |
| $z$                         | 14  | 8   | 13  | 14  | NA  |

Based on the results given in the Tables 4–7, a causality graph can be determined for the most frequently repeated pattern. It is shown in a Figure 14. The analysis of the tables allows one to unequivocally state that the signals $u$, $w$, $x$, and $y$ show a close causal relationship with the signal $z$. There are of course exceptions as for the signal $u$ ($d_5$) (16 casual relationships for both $y$ and $z$ in Table 4 and the same number for $w$ and $z$ in Table 5) or for the signal $w$ (Table 4—$d_5$), but the general rule is maintained. Doubts arise in the case of signal $z$—it can be noticed that a relationship appears between both signals $u$ and $x$ separately. For the sums of the highest Transfer Entropy coefficients determined on the basis of the average TE values of the sets $d_5$ and $d_6$, the results for the 50 sub-datasets unequivocally define the causality of the signals $u$, $w$, $x$, and $y$ with the signal $z$ and the
signal z with the signal u, however, in the case of 49 subdatasets the causal relationship is the same for u and y. Due to the fact that such a relationship does not repeat before, this can be omitted.

![Causality diagrams](image)

**Figure 14.** Causality diagrams for the 1st variant: bilateral causal relationship between (a) u and z and (b) x and z.

### 4.2.2. The 2nd Variant

According to Section 2.3, the median filter is used to shorten the data. Operating with MF where n divisor equals 10 and determining the Transfer Entropy coefficients for the signal pairs of the whole dataset, only NaN or Inf values are obtained. The same situation is repeated for the n divisor equals 25, so these two variants are completely rejected from the analysis. Correct results in the form of real numbers are obtained for the n divisors equal to 50 and 100 (except for causality for signal u—0 for Nan or Inf). Tables 8 and 9 show the measure of causal relationships for the oscillations of individual pairs of variables for n divisor equals 50 and 100, respectively. Due to the lack of consistency and repeatability of the results for the 2nd variant, it cannot be generalized to the form of only two versions of the causality graph, as in Section 4.2.1. Figure 15 presents causality graphs for oscillations from decomposed detrended signals.

Comparing the results obtained as part of the division of datasets according to the assumptions of the 1st variant, especially causality graphs are shown in Figure 14, it is seen that those presented in Figure 15 do not have any common features. The strong relationship with the variable z does not occur as strongly as it does in the 1st variant. The results vary significantly not only in this aspect but also in the repeatability of the TE coefficient values for the corresponding IMFs. The situation is much better in case of the results presented in Table 8, which is reflected in the causality graphs (Figure 15a,b).

**Table 8.** Transfer Entropy coefficient for 2nd variant (n divisor equals 50).

| $T_{row \rightarrow column}$ | u   | w   | x   | y    | z       |
|-----------------------------|-----|-----|-----|------|---------|
| u                           | $d_4$ | NA  | 0.0355 | 0.0296 | **0.0779** | 0.0616 |
|                             | $d_5$ | NA  | 0.0432 | 0.0338 | **0.0605** | 0.0470 |
| w                           | $d_4$ | 0   | NA   | 0.0829 | **0.1045** | 0.0844 |
|                             | $d_5$ | 0   | NA   | 0.0438 | **0.0721** | 0.0535 |
| x                           | $d_4$ | 0   | 0.0485 | NA    | 0.1078   | **0.1295** |
|                             | $d_5$ | 0   | 0.0407 | NA    | 0.0479   | **0.0691** |
| y                           | $d_4$ | 0   | 0.0560 | 0.1312 | NA       | **0.1633** |
|                             | $d_5$ | 0   | 0.0462 | **0.0819** | NA   | 0.0754   |
| z                           | $d_4$ | 0   | 0.0547 | 0.1405 | **0.1675** | NA       |
|                             | $d_5$ | 0   | 0.0479 | **0.0859** | 0.0724   | NA       |
Table 9. Transfer Entropy coefficient for 2nd variant (n divisor equals 100).

| $T_{row\rightarrow column}$ | $u$ | $w$ | $x$ | $y$ | $z$ |
|-----------------------------|-----|-----|-----|-----|-----|
| $u$ | $d_3$ | NA | 0.0358 | 0.0395 | **0.1235** | 0.1008 |
| $d_4$ | NA | 0.0784 | 0.0431 | 0.0662 | **0.0899** |
| $w$ | $d_3$ | 0.0461 | NA | **0.0997** | 0.0898 | 0.0906 |
| $d_4$ | **0.0945** | NA | 0.0874 | 0.0675 | 0.0750 |
| $x$ | $d_3$ | 0 | 0.0541 | NA | 0.1097 | **0.1758** |
| $d_4$ | 0.0574 | 0.0625 | NA | 0.0679 | **0.1024** |
| $y$ | $d_3$ | 0 | 0.0741 | 0.0952 | NA | **0.1817** |
| $d_4$ | **0.1508** | 0.0558 | 0.0587 | NA | 0.0926 |
| $z$ | $d_3$ | 0 | 0.0946 | 0.1919 | **0.1961** | NA |
| $d_4$ | 0.0808 | 0.0547 | 0.0888 | **0.0922** | NA |

Figure 15. Causality diagrams for the 2nd variant: (a,b) reflect $n = 50$; (c,d) reflect $n = 100$.

Graphs are similar to each other to a certain extent however, they can not be generalized to a single case. Even worse conclusions are drawn when it comes to the analysis of results presented in Table 9. It can be noticed that the highest TE coefficient value in a row occurs for different pairs of variables within one signal ($d_3$ and $d_4$). Attention is also drawn to the fact that in the case when the divisor of $n$ equals 100, for the IMFs marked as $d_3$, for pairs of the variable $u$ with the variables $x$, $y$, and $z$ there are 0 (Nan or Inf). For $d_4$ such a problem does not occur. Nevertheless, causality graphs can be determined and and their form are presented in Figure 15c,d. Graphs are unfortunately so inconsistent from each other that one can not be accepted as a correct result. Due to the occurrence of Nan’s and Inf’s adopted as 0, it is impossible to perform the averaging of the entropy coefficients values, as it is possible for the 1st variant (Tables 6 and 7).

As one can see, it is hard to find a single explanation for the above-given industrial example. The process knowledge shows that inferring causality from the highest TE coefficient for oscillations is most likely wrong, and the causality graphs do not reflect the correct relationships. The information contained in the oscillations obtained during the decomposition process for 1st variant may be insufficient for the successful Transfer Entropy method application or may be lost. The use of the Median Filter has a positive effect on the computational demands. However, further analysis of the oscillations obtained from the minimized datasets does not bring the desired results.

These observations clearly show that the analysis of the industrial data differs completely from the theoretical analyzes. Simple relations and approaches developed in the sterile simulated environment do not have to extend towards real applications.
5. Conclusions and Further Research

The presented approach extends the classical TE model-free approach, changing the domain of causality detection. Original approaches take the process data as it is. We try to investigate the hypothesis that the identification can be done by taking into account the only oscillatory element of the data. This modification can be simply done with the TE method using common oscillation detection methods, for instance, the MEEMD algorithm. It is difficult to compare the TE method with other approaches as this method is model-free, while the others assume some process models. One has to be aware that such models are rarely known in the process industry and any misfit in modeling may arise the dilemma of what is wrong: the model or the causality detection. Therefore the the methods requiring model assumptions are not considered. In the case of the industrial data used in the analysis, the modeling effort would unnecessarily extend the complexity without any positive add-ons.

The paper focuses on simulated and industrial data processing used to discover the oscillations existing in control-loop variables. Time series decomposition is performed using the MEEMD algorithm. The disadvantages and limitations of this approach force the authors to prepare specific industrial data for analysis. Two variants are checked. The key issue is the use of the Transfer Entropy algorithm on the obtained oscillations from the simulation and industrial data and then the generation of causality graphs.

It is shown that causality analysis is a complex task. There is no one single solution that works. The paper discovers two realities. Simulation example, despite the fact that is is non-Gaussian works perfectly. There is no doubt that the proposed methodology works in that case. Application of the successfully tested method to the normal, not specifically treated industrial data opens the Pandora box. The results become counterintuitive, variable, unrepeatable, and difficult to interpret. It appears that long data files pose a challenge to the methods used. The question of why is hard to answer. There might be several reasons: known and unknown. It might be non-stationarity, non-linearity, outliers, data granulation, fractality, human interventions, uncoupled disturbances (weather?), and many others.

In the case of the simulated data, the data are coherent with the applied methods and algorithms limitations are not violated. However, in the case of industrial data, we are never sure about their actual properties. We do not know if they are stationary. We do not know what type of non-stationarity they exhibit. We should investigate the outliers. We should identify the oscillations. We also need to analyze the residuum and its statistical properties. There are various issues that should be taken into consideration. Once we properly identify the process data, we may select or adopt the proper methodology. However, the research on industrial data internal properties and their compliance with the limitations of the methods used is infrequent.

However, the above observation should not discourage anyone. It represents a research challenge. However, this challenge requires a different approach. Simulations are not sufficient and in-depth, step-by-step comparative analysis must be carried out on several industrial sites; real ones, not laboratory ones. Of course, this is difficult and demanding. But only then it is possible to delve deeper into the issue and perhaps, but not certainly, find a solution.

Author Contributions: Conceptualization, P.D.D. and M.J.F.; methodology, M.J.F. and E.P.; software, M.J.F.; validation, M.J.F. and E.P.; formal analysis, P.D.D., M.J.F., and E.P.; investigation, M.J.F.; data curation, M.J.F.; writing—original draft preparation, M.J.F.; writing—review and editing, P.D.D. and E.P.; visualization, M.J.F.; supervision, P.D.D.. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.
Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

- MEEMD: Median Ensemble Empirical Mode Decomposition
- EMD: Empirical Mode Decomposition
- EEMD: Ensemble Empirical Mode Decomposition
- TE: Transfer Entropy
- T-F: Time-Frequency
- IMF: Intrinsic Mode Function

References

1. Yang, F.; Xiao, D. Progress in root cause and fault propagation analysis of large-scale industrial processes. *J. Control Sci. Eng.* 2012, 2012, 478373. [CrossRef]

2. Lindner, B.; Auret, L.; Bauer, M. Investigating the Impact of Perturbations in Chemical Processes on Data-Based Causality Analysis. Part 1: Defining Desired Performance of Causality Analysis Techniques. *IFAC-PapersOnLine* 2017, 50, 3269–3274. [CrossRef]

3. Bauer, M.; Thornhill, N.F. A practical method for identifying the propagation path of plant-wide disturbances. *J. Process Control* 2008, 18, 707–719. [CrossRef]

4. Baccala, L.A.; Sameshima, K. Partial directed coherence: A new concept in neural structure determination. *Biol. Cybern.* 2001, 84, 463–474. [CrossRef] [PubMed]

5. Barnett, L.; Seth, A.K. The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *J. Neurosci. Methods* 2014, 223, 50–68. [CrossRef] [PubMed]

6. Schreiber, T. Measuring information transfer. *Phys. Rev. Lett.* 2000, 85, 461–464. [CrossRef] [PubMed]

7. McCann, M.; Li, Y.; Maguire, L.; Johnston, A. Causality Challenge: Benchmarking relevant signal components for effective monitoring and process control. *J. Mach. Learn. Res.* 2008, 6, 277–288.

8. Lindner, B.; Auret, L.; Bauer, M. Investigating the Impact of Perturbations in Chemical Processes on Data-Based Causality Analysis. Part 2: Testing Granger Causality and Transfer Entropy. *IFAC-PapersOnLine* 2017, 50, 3330–3335. [CrossRef]

9. Lindner, B.; Auret, L.; Bauer, M.; Groenewald, J.W. Comparative analysis of Granger causality and Transfer Entropy to present a decision flow for the application of oscillation diagnosis. *J. Process Control* 2019, 79, 72–84. [CrossRef]

10. Su, J.; Wang, D.; Zhang, Y.; Yang, F.; Zhao, Y.; Pang, X. Capturing Causality for Fault Diagnosis Based on Multi-Valued Alarm Series Using Transfer Entropy. *Entropy* 2017, 19, 663. [CrossRef]

11. Chen, Q.; Fei, X.; Xie, L.; Li, D.; Wang, Q. Causality analysis in process control based on denoising and periodicity-removing CCM. *J. Manuf. Spec. Equip.* 2020, 1, 25–41. [CrossRef]

12. Sugihara, G.; May, R.; Ye, H.; Hsieh, C.H.; Deyle, E.; Fogarty, M.; Munch, S. Detecting causality in complex ecosystems. *Science* 2012, 338, 496–500. [CrossRef] [PubMed]

13. Ye, H.; Deyle, E.R.; Gilarranz, L.J.; Sugihara, G. Distinguishing time-delayed causal interactions using convergent cross mapping. *Sci. Rep.* 2015, 5, 14750. [CrossRef] [PubMed]

14. Lindner, B.; Auret, L.; Bauer, M. A systematic workflow for oscillation diagnosis using Transfer Entropy. *IEEE Trans. Control Syst. Technol.* 2019, 27, 1–2. [CrossRef]

15. Falkowski, M. Causality analysis incorporating outliers information. In *Outliers in Control Engineering. Fractional Calculus Perspective*; Domarński, P.D., Chen, Y.Q., Ławryńczuk, M., Eds.; Walter de Gruyter GmbH: Berlin, Germany; Boston, MA, USA, 2022; pp. 133–148.

16. Zhang, X.; Hu, W.; Yang, F. Detection of Cause-Effect Relations Based on Information Granulation and Transfer Entropy. *Entropy* 2022, 24, 212. [CrossRef] [PubMed]

17. Landman, R.; Kortela, J.; Sun, Q.; Jamasa-Joumela, S.L. Fault propagation analysis of oscillations in control loops using data-driven causality and plant connectivity. *Comput. Chem. Eng.* 2014, 71, 446–456. [CrossRef]

18. Qin, S.J.; Yuan, T. Root cause diagnosis of plant-wide oscillations using Granger causality. *J. Process Control* 2014, 24, 450–459.

19. Domarński, P.D. Study on statistical outlier detection and labelling. *Int. J. Autom. Comput.* 2020, 17, 788–811. [CrossRef]

20. Rousseew, P.J.; Leroy, A.M. *Robust Regression and Outlier Detection*; John Wiley & Sons, Inc.: New York, NY, USA, 1987.

21. Okoński, M.; Domarński, P.D. Study on oscillation detection robustness and outlier filtering impact. In *Outliers in Control Engineering. Fractional Calculus Perspective*; Domarński, P.D., Chen, Y.Q., Ławryńczuk, M., Eds.; Walter de Gruyter GmbH: Berlin, Germany; Boston, MA, USA, 2022; pp. 187–218.

22. Falkowski, M.; Domarński, P.D. Nested NARIMA model of the atmospheric distillation column. In *Advances in Intelligent Systems and Computing Automation*; Szwezyk, R., Zieliński, C., Kaliczyńska, M., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 115–124.

23. Lang, X.; ur Rehman, N.; Zhang, Y.; Xie, L.; Su, H. Median ensemble empirical mode decomposition. *Signal Process.* 2020, 176, 107686. [CrossRef]
24. Mandic, D.P.; ur Rehman, N.; Wu, Z.; Huang, N.E. Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data analysis. *IEEE Signal Process. Mag.* **2013**, *30*, 74–86. [CrossRef]

25. Spiegel, M.R.; Stephen, L.J. *Schaum's Outline of Statistics*, 6th ed.; McGraw-Hill Education: New York, NY, USA 2018.

26. Domanski, P.D. *Control Performance Assessment: Theoretical Analyses and Industrial Practice*; Springer International Publishing: Cham, Switzerland, 2020.

27. Domanski, P.D.; Falkowski, M. Impact of outliers on determining relationships between variables in large-scale industrial processes using Transfer Entropy. In *Proceedings of the 7th International Conference on Control, Decision and Information Technologies (CoDIT)*, Prague, Czech Republic, 29 June–2 July 2020; pp. 807–812.

28. Domanski, P.D. Non-gaussian properties of the real industrial control error in SISO loops. In *Proceedings of the 19th International Conference on System Theory, Control and Computing*, Cheile Gradistei, Romania, 14–16 October 2015; pp. 877–882.