CAUSALITY IN TIME SERIES:
DYNAMIC TIME WARPING VERSUS GRANGER CAUSALITY

by

Leyla Zeynep Yallak

B.S., Chemical Engineering, Boğaziçi University, 2010

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemical Engineering

Boğaziçi University

2012

CAUSALITY IN TIME SERIES:
DYNAMIC TIME WARPING VERSUS GRANGER CAUSALITY

APPROVED BY:

Prof. Uğur AKMAN …….
(Thesis Supervisor)

Prof. Türkan HALİLOĞLU …….

Prof. Nesrin OKAY …….

DATE OF APPROVAL:

ACKNOWLEDGEMENTS

I would like to thank my thesis supervisor Prof. Uğur Akman for his always positive and encouraging feedback and support.

Many thanks to Prof. Türkan Haliloğlu and Prof. Nesrin Okay for their time in reading and commenting on my thesis along with all professors and supporting personnel in the chemical engineering department for everything I have learnt from them during my six years at the department.

I am grateful to all my friends, especially to Ceyda Önol, with whom we have met many deadlines together, for her kindness and support.

I would also like to acknowledge the two-year financial support of TÜBİTAK BİDEB 2210 Graduate Programme and the financial support of Türk Eğitim Vakfı during my undergraduate study.

Finally, greatest thanks to my family, who are always with me wherever they may be, for their patience and understanding.

ABSTRACT

CAUSALITY IN TIME SERIES: DYNAMIC TIME WARPING VERSUS GRANGER CAUSALITY

Causality is a concept studied in various areas such as economics, and engineering. Identifying the cause and effect relations among variables is important as it enables the control of the effected variable by the variation of the cause or helps predict the future behavior of the effected variable based on the behavior of the cause. Granger Causality (GC) test is a statistical test mainly used for causality detection in economics and recently in bioinformatics. The GC test determines whether one series Granger causes the other or not, or if there exists a feedback relation. However, results of the GC tests do not elucidate how these relations change with time. Dynamic Time Warping (DTW) is a method employed for similarity measurement in classification and clustering applications, in areas such as speech recognition and batch trajectory synchronization. In the DTW method, principles of dynamic programing are utilized and the series are aligned nonlinearly in the time axis. In this thesis work, it is proposed that DTW can help determine the temporal order and the lead/lag relations of the series, therefore, the causal relations. The DTW method is tested on selected synthetic data sets and on data from chemical and biochemical processes, and engineering related economic indicators. The DTW-based causality results are compared with those of the GC tests and cross-correlation analyses. The DTW-based results were found to be as expected and in accordance with the GC test only for the simple examples, for multivariable sets and nonlinearly-related variables, the method was unsuccessful.

ÖZET

ZAMAN SERİLERİNDE NEDENSELLİK: DİNAMİK ZAMAN BÜKMESİ VE GRANGER NEDENSELLİĞİ

Nedensellik, ekonometri ve mühendislik gibi birçok farklı alanda üzerinde çalışılan bir kavramdır. Değişkenler arasındaki neden-sonuç ilişkisinin belirlenmesi, nedenlere sonucu yönlendirme amacıyla müdahale edilmesi ya da nedenlerin incelenerek sonuç hakkında tahminlerde bulunulabilmesi açısından önemlidir. Granger Nedensellik (GN) sınaması ağırlıklı olarak ekonometri alanında, son zamanlarda ise biyoenformatik alanında uygulamaları olan istatistiksel bir testtir. Bu testle, değişkenler arasında tek ya da çift yönlü nedensellik ilişkilerinin varlığı incelenebilir ancak, bu ilişkilerde zamanla meydana gelen değişimler hakkında fikir sahibi olunamamaktadır. Dinamik Zaman Bükmesi (DZB), seriler arası benzerlik ölçümü amacıyla, konuşma tanıma ve kesikli proses yörüngelerinin senkronizasyonu gibi alanlarda sınıflandırma ve kümeleme uygulamalarında kullanılan bir yöntemdir. DZB yöntemde, dinamik programlama ilkeleri kullanılarak, seriler zaman ekseninde doğrusal olmayan bir şekilde eşleştirilirler. Bu tez çalışmasında, DZB yönteminin seriler arasındaki zamansal sıralamanın ve bu sayede nedenselliğin belirlenmesinde kullanılabilirliği araştırılmıştır. Bu amaçla, DZB yöntemi özel olarak tasarlanmış veri kümeleri, kimyasal ve biyokimyasal işlemler ve mühendislik ile ilgili ekonomik gösterge verileri üzerinde denenmiştir. Uygulamalarda DZB'ye dayalı sonuçlar GN sınaması ve çapraz korelasyon uygulamalarıyla elde edilen sonuçlarla karşılaştırılmıştır. DZB yöntemi basit uygulamalarda beklenen ve GN sınamasına uygun sonuçlar vermiş olsa da çok değişkenli kümelerde ve doğrusal olmayan ilişkilerlerin bulunduğu verilerde yöntem başarılı olmamıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

a	Equation parameter
b	Equation parameter
c	Equation parameter
C_A	Concentrations of CNBr and compound A
C_B	Concentrations of CH_3NH_2 and compound B
C_C	Concentrations of CH_3Br and compound C
C_D	Concentration of $NCNH_2$
$c_{\mathbf{XY}}(k)$	Cross-covariance between **X** and **Y** at lag k
d	Equation/model parameter
d	Local distance matrix
D	Cumulative distance matrix
d_i	Squared local distance between any two points
Dist, $Dist_{DTW}$	Cumulative warping distance
$Dist_E$	Euclidean distance
F	Value of the F-statistic
f	Fox population
f_0	Initial amount of the fox population
F_c	Critical value of the F-statistic
H_0	Null hypothesis
K	Length of the warping path
k	Optimal lag
k_1, k_2, k_3	Reaction rate constant
k_s	Biochemical system model parameter
L	Biochemical system model parameter
l	Optimal lag
m	Optimal lag
n	Number of data points
p	Optimal lag
q	Biochemical system model parameter
r	Rabbit population

r_0	Initial amount of the rabbit population
$\mathbf{r_M}$	Time series of length M
r_s	Biochemical system model parameter
$r_{\mathbf{XY}}(k)$	Cross-correlation between **X** and **Y** at lag k
s**x**	Standard deviation of **X**
s**y**	Standard deviation of **Y**
t	Time
$\mathbf{t_N}$	Time series of length N
u	Time series
v_0	Volumetric flow rate
V	Reactor volume
V_0	Initial volume of the reactor
W	Warping path
w_k	kth element of the warping path
x_i	Time series
X	Time series
X_t	Value of the time series **X** at time t
$\overline{X}$	Sample mean of **X**
y	Time series
Y	Time series
Y_t	Value of the time series **Y** at time t
$\overline{Y}$	Sample mean of **Y**

α	Growth rate of rabbits when there is no interaction with foxes
α	Significance level for the Granger causality test
α	Substrate concentration
α_0	Initial substrate concentration
β	Coefficient of the predator-prey interactions decreasing the rabbit population
γ	Death/emigration rate of foxes in the absence of interactions with rabbits
γ	Product concentration
γ_0	Initial product concentration

δ	Growth of the fox population in case of interaction with rabbit population
ε_i	Prediction error, noise
η_i	Prediction error, noise
μ	Biochemical system model parameter
ν	Biochemical system flow parameter
σ	Biochemical system model parameter
φ	Biochemical system model parameter
$\leftarrow, \rightarrow, \leftrightarrow$	Granger causation directions
$\Leftarrow, \Rightarrow, \Leftrightarrow$	Lead/lag, precedence directions

LIST OF ACRONYMS/ABBREVIATIONS

ADF	Augmented Dickey-Fuller
BIC	Bayesian information criterion
CEPCI	Chemical engineering plant cost index
CorOrg	Correlation between the unwarped series
CorDTW	Correlation between the series after warping
CS	Consumption
DTW	Dynamic time warping
DZB	Dinamik zaman bükmesi
GC	Granger causality
GDP	Gross domestic product
GN	Granger nedenselliği
IP	Industrial production
KPSS	Kwiatkowski–Phillips–Schmidt–Shin
MaxLag	Maximum number of lags considered in the model
Min	Minimize, minimum
O	Order of magnitude
OP	Oil price
RSS	Residual sum of squares
TimeShift	The mode of the time shifts between the series in the warping path

1. INTRODUCTION

There are a number of different approaches to causality in various fields such as philosophy, physics, and economics. A review on temporality and causality can be found in Karimi (2010). Granger causality (GC) is one of these methods that is based on the improvement in the predictability of a variable by using the past values of another variable (Granger, 1969). Conditional Granger causality, which searches for causality between variables by taking the effect of additional variables into consideration, was introduced by Geweke (1984). Guo *et al*. (2008) presented a method called 'partial Granger causality' that can discover the causal relations in linear and nonlinear systems where exogenous variables are present. Moreover, a multivariate nonlinear Granger causality test was developed by Bai *et al*. (2010). For a reliable test of GC, the variables in the model should be chosen with care and sample sizes should be large enough (Stern, 2011). The GC test is mainly used in economics, with recent applications in bioinformatics. To name a few of these studies, Nagarajan and Upreti (2010) used GC for the determination of the functional relationships between gene expression profiles of the human cell-cycle. In a study by Mukhopadhyay and Chatterjee (2007), GC is used in the formation of a causality network of microarray time series data. Li *et al*. (2006) applied GC on the detection of the causality between intermediate phenotypes and Krishna *et al*. (2011) utilized partial GC for process monitoring and fault detection in bioreactors.

Mullen (2010) criticized the fallacy that correlation implies causality with a picturesque example as follows:

> "… To illustrate: suppose we are observing two ants independently following a pheromone trail towards some tasty morsel. Ant 1 started the journey two minutes before Ant 2 and so he appears to be 'leading' Ant 2. If we compute the cross-correlation between the two ants' trajectories for a range of time lags we would find a high correlation between their trajectories and, furthermore, we would find the correlation was peaked at a non-zero lag with Ant 1 leading Ant 2 by a lag of two minutes. But it would be foolish to say that Ant 1 was 'causing' the behavior of Ant 2. In fact, not only is there no causal relationship whatsoever between the two, but there is not even any information being transmitted between the two ants. They are conditionally independent of each other, given their own past history and given the fact that each is independently following the pheromone trail (this is the 'common (exogenous) cause' that synchronizes their behavior). If we were to intervene and remove Ant 2 (Ant 1), Ant 1 (Ant 2) would continue on his way, oblivious to the fact that his comrade is no longer in lock-step with him. Consequently, if we calculate the Granger-causality between the two trajectories we will find that the GC is zero in both directions: there is no information in the history of either ant that can help predict the future of the other ant above and beyond the information already contained in each ant's respective past…".

Mullen stated that the cross-correlation method does not tell if one of the variables contains useful information for the prediction of the other variable that cannot be provided by any other variable. The GC test does this by using regression. On the other hand, a maximum value of cross-correlation at a non-zero lag means that one of the variables has temporal precedence over the other variable and this satisfies a requirement of causality that is "causes must precede effects in time" (Mullen, 2010). However, a third (hidden/unobservable) variable manipulating both of the variables at different times may be the reason for the correlation at lagged values (Sornette and Zhou, 2005). Moreover, in a study by Kispersky *et al*. (2011), correlated noises were generated with correlation varying between −1 and 1, and one of the series was lagged behind the other. The direction of causality was examined by the application of the GC test. GC detected the causation in all cases but the uncorrelated case. It was found that when the series were highly correlated/anti-correlated, they were able to capture the lag between the series and the direction of causality by using cross-correlation. However, as the correlation between the generated series decreased, the relation became less clear. The disadvantage of GC is that it does not provide information on the lag (Kispersky *et al*., 2011).

A brief review of DTW can be started with the mention of Sakoe and Chiba (1978) who used dynamic programing for the nonlinear warping of the time axis in speech-pattern recognition. They also compared symmetric and asymmetric forms of the method and introduced slope constraints. Kassidas *et al*. (1998) performed DTW on batch trajectory synchronization, required for the comparison and statistical analysis of the batches. In their application, they used multivariate batch data. Keogh and Ratanamahatana (2005) showed that by using a modified piecewise aggregate approximation, DTW can be utilized in the indexing of time series. The use of DTW as a classifier can be seen in Legrand *et al*. (2008) who applied DTW on classification of human chromosomes.

In the literature, the only study found that uses DTW in the same context with GC was carried out by Peng *et al*. (2008). In the study, a semi-automatic system was presented in which DTW and GC methods were utilized together in leading indicator discovery. DTW was considered more convenient than Euclidean distance or cross-correlation coefficients in measurements of correlation among the indicators. For this purpose, the cumulative DTW distance was used. In addition, the time shifts in the warping path were

considered and they were used alongside the GC test in the determination of the time order and the causal direction of the highly correlated series (Peng *et al*., 2008). However, they have not regarded the use of DTW as a means to determine the variations in the relations with time, which could be happening in their system, and examining only the overall time shifts may lead to a failure to notice. There is also a US Patent (Peng *et al*., 2011) which claims the discovery of a "semi-automatic system with an iterative learning method for uncovering the leading indicators in business processes".

Another study with a method and aim similar to the present study was conducted by Sornette and Zhou (2005). In this study, "the optimal thermal causal path method" was applied on time series. This method, resembling DTW, also used distance measures between series, and in order to uncover the lead/lag relationships matched the series dynamically in time dimension by minimizing the cumulative distance. The robustness of the method was increased by the addition of a non-zero temperature. Synthetic series were used as test cases for the demonstration of the performance with changing lags and predictability. Better results were obtained compared with the cross-correlation method. The method was then applied on several economic variables. The advantage of the method is that it can reflect the changes in the behavior of the series with time. In addition, it is stated that a large amount of data is not needed and nonlinear relations can be handled as well (Sornette and Zhou, 2005). These are mainly what have been aimed with the application of DTW as an alternative to the GC test.

By comparing the magnitudes of the series with respect to the time dimension, DTW matches similar trends; for example, an increasing trend in a series is matched with a similar trend in the second series. Therefore, the observation of how the series are matched can provide information on the time order and it can be assumed that the series in which the trend is observed earlier leads the other series. As stated previously, this property satisfies the "causes must precede the effect" requirement of causality. Although it helps determine temporal precedence and whether it changes over time, concluding that a variable causes the other one only by using DTW may not be correct. As it is with the cross-correlation method, the series can actually be causally related or a third (hidden/unobservable) series may be the common cause. DTW is not expected to

differentiate between these two cases. Likewise, the GC test applied in this work does not consider additional (exogenous) variables; therefore, spurious relations may be obtained.

In this study, the aim is to see whether DTW can be used as an alternative method to the bivariate GC test in detection of causality. Unlike GC, DTW does not require large sample sizes and statistical tests, moreover it can enable the examination of relations that are changing over time, which cannot be observed with the GC test.

Chapter 2 of this thesis work provides information on the DTW method and algorithm. Chapter 3 summarizes the GC test and briefly mentions the cross-correlation method. In Chapter 4, the results of the applications of DTW on data sets of wide scope are given. The data sets are also investigated with the GC test and cross-correlation analyses and the results are compared with those of DTW. Chapter 5 provides the conclusions drawn from the applications, and recommendations on future work are presented as well. In the Appendix, MATLAB codes for the DTW method and GC test are provided for selected examples.

2. DYNAMIC TIME WARPING METHOD

This chapter provides theoretical information on DTW and DTW algorithm. DTW method and its use in this thesis work are explained in Section 2.1. In Section 2.2, the algorithm used in the applications is provided and it is demonstrated on a simple example.

2.1. DTW Problem Formulation

Measure of similarity between time series is useful for classification and clustering purposes. Given two time series **t** and **r** of lengths N and M respectively, $\mathbf{t} = t_1, t_2, \dots, t_N$ and $\mathbf{r} = r_1, r_2, \dots, r_M$, similarity of the series can be measured by using Euclidean distance, $Dist_E$ which is the sum of the squared distances (d_i) between the corresponding points of each series, $(t_1,r_1),(t_2,r_2),\dots,(t_N,r_N)$. Equation 2.1 shows the calculation of the cumulative Euclidean distance.

$$Dist_E = \sum d_i = \sum_{i=1}^{i=N} (t_i - r_i)^2 \tag{2.1}$$

The cumulative distance indicates dissimilarity, so the smaller the distance between the series, the more similar they are. However, in some cases Euclidean distance is not a very reliable choice for similarity measure because of the one-to-one matching between the points. The series can be in fact very similar but slightly shifted in time dimension, resulting in a greater distance when the Euclidean distance measure is used. DTW, on the other hand, solves this problem by warping the time axis dynamically, stretching/compressing the axis where necessary, therefore allowing time-wise different points of the series be matched. There are a number of constraints on this nonlinear matching of the points. The boundary condition forces the first and the last elements of the series to be matched, $(t_1,r_1),\dots,(t_i,r_j),\dots,(t_N,r_M)$. The monotonicity condition requires the matching to proceed with monotonically increasing indexes, thus preventing going back in the time dimension, $i_{k+1} \geq i_k$ and $j_{k+1} \geq j_k$. The continuity condition requires movement only towards adjacent indexes, ensuring that each point is used and excessive compression or expansion of the time axis is avoided, $i_{k+1} - i_k \leq 1$ and $j_{k+1} - j_k \leq 1$ (Pravdova *et al.*, 2002; Keogh and Ratanamahatana, 2005).

The complete matching of the two series, starting at the beginning and finishing at the end points is called the warping path. An example of a warping path is depicted in Figure 2.1.

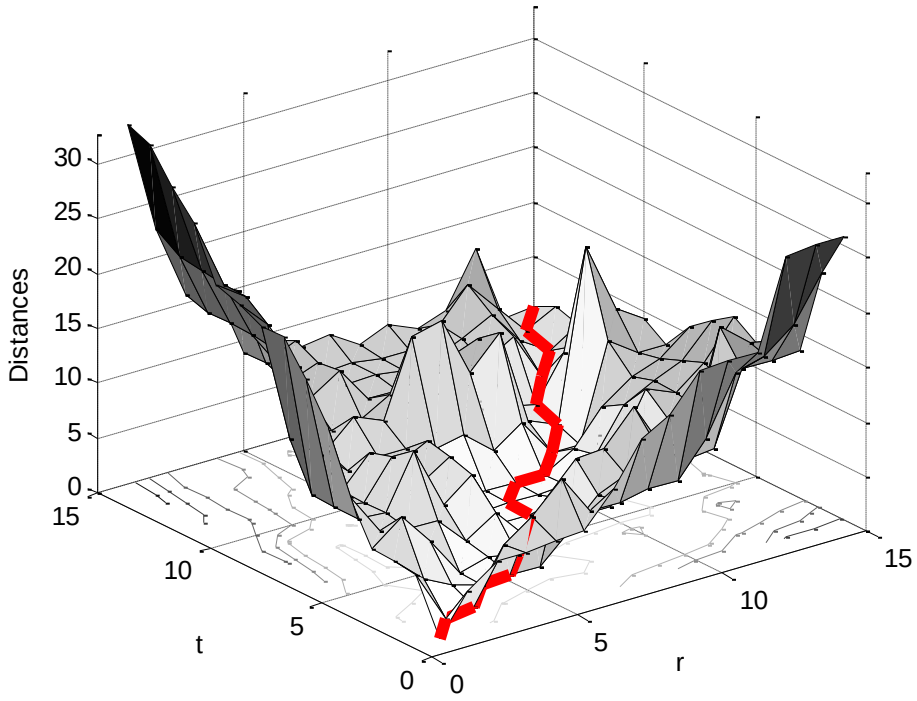

Figure 2.1. 3D representation of a distance matrix and the optimal warping path.

The aforementioned constraints restrict the predecessors of a point (i,j) on the path to three points (i,j−1), (i−1,j) or (i−1,j−1). However, there are still many warping paths satisfying this restriction. The one of interest is the warping path that results in the smallest cumulative distance between the series; this path is called the optimal warping path. Dynamic programing is used in the optimal warping path determination. Instead of calculating the cumulative distance for each possible path and then finding the path with minimum distance, this method breaks the optimization problem in Equation 2.2 of minimizing the cumulative distance, to smaller steps and at each step by moving to a point on the path from the minimum of the three points preceding it, the overall minimum distance is obtained (Ramaker *et al.*, 2003). The indexes showing how the series are matched can then be recovered by backtracking. In Equation 2.2, $Dist_{DTW}$ is the cumulative distance, the warping path is $\mathbf{W}=w_1,w_2,\ldots,w_K$, $w_k=(i,j)_k$ and K is the length of the path.

$$Dist_{DTW}=\min_{W}\left[\sum_{k=1}^{K} d(w_k)\right]=\min_{W}\left[\sum_{k=1}^{K} d(i_k,j_k)\right] \quad (2.2)$$

The two series that are examined with DTW can have different number of data points, thus series of equal length are not required for this method. However, DTW has a time and space complexity of O(NM) and this makes it more suitable for series containing less than a few thousand data points (Keogh and Ratanamahatana, 2005; Salvador and Chan, 2004).

In this thesis work, DTW is not utilized as a similarity measure for clustering or classification. It is used to warp the series to their most similar shape by making the distance between them minimum and then to determine the lead/lag relation between the series by examining the position of the optimal warping path over the cumulative distance matrix relative to the diagonal line. The diagonal line represents a special case of DTW where the identical indexes of the series are matched and Euclidean distance is obtained, $w_k=(i, j)_k$, i=j=k. In an analysis of DTW(**t**,**r**), the path lies above/right of this line when more recent points of the second series (series **r**) are matched with the past points of the first series (series **t**), $w_k=(i, j)_k$, i < j and as a result it is concluded that the series **t** precedes the series **r**. The reverse is said when the path lies below/left of the diagonal line. For this kind of a lead/lag examination of the series, one needs to use a nonlinear time measure since using one-to-one measures do not allow the matching of similar trends occurring at different times and thus uncovering the order of their occurrence by the investigation of the path.

2.2. DTW Algorithm Applied to an Example Problem

In this section, DTW algorithm is explained on a simple example. The sample series are generated in Excel by using RAND() function. The series **t** is generated by RAND()×100 and the series **r** by RAND()×80. The lengths of the series are 15.

t={36, 93, 44, 27, 86, 97, 83, 78, 32, 0, 53, 93, 73, 79, 53}

r={6, 56, 45, 6, 26, 12, 54, 57, 42, 55, 16, 3, 79, 24, 11}

Z-score normalization is applied on the series and the following series with mean zero and standard deviation one are obtained. From this point on, **t** and **r** represent the normalized series and all applications are carried out on them.

t={-0.89, 1.07, -0.62,-1.19, 0.83, 1.20, 0.74, 0.55, -1.02, -2.11,-0.29, 1.06, 0.38, 0.59,-0.29}

r={-1.12, 0.97, 0.52, -1.12, -0.30, -0.86, 0.90, 1.02, 0.36, 0.91, -0.69, -1.22, 1.92, -0.37, -0.92}

The local distance between any two points of the series is given by Equation 2.3 and the local distance matrix $\mathbf{d_{NxM}}$ is constructed by applying this formula for each pair. Table 2.1 shows the distances between t and r.

$$d_{n,m}=(t_n - r_m)^2 \tag{2.3}$$

Table 2.1. Local distance matrix **d**.

0.1	3.5	2.0	0.1	0.3	0.0	3.2	3.6	1.6	3.2	0.0	0.1	7.9	0.3	0.0
4.8	0.0	0.3	4.8	1.9	3.7	0.0	0.0	0.5	0.0	3.1	5.3	0.7	2.1	4.0
0.2	2.5	1.3	0.2	0.1	0.1	2.3	2.7	1.0	2.3	0.0	0.4	6.5	0.1	0.1
0.0	4.7	2.9	0.0	0.8	0.1	4.4	4.9	2.4	4.4	0.2	0.0	9.6	0.7	0.1
3.8	0.0	0.1	3.8	1.3	2.9	0.0	0.0	0.2	0.0	2.3	4.2	1.2	1.4	3.1
5.4	0.0	0.5	5.4	2.2	4.2	0.1	0.0	0.7	0.1	3.6	5.9	0.5	2.5	4.5
3.5	0.1	0.0	3.5	1.1	2.6	0.0	0.1	0.1	0.0	2.0	3.9	1.4	1.2	2.8
2.8	0.2	0.0	2.8	0.7	2.0	0.1	0.2	0.0	0.1	1.5	3.2	1.9	0.9	2.2
0.0	4.0	2.4	0.0	0.5	0.0	3.7	4.2	1.9	3.7	0.1	0.0	8.7	0.4	0.0
1.0	9.5	7.0	1.0	3.3	1.6	9.1	9.8	6.1	9.1	2.0	0.8	16.3	3.0	1.4
0.7	1.6	0.7	0.7	0.0	0.3	1.4	1.7	0.4	1.4	0.2	0.9	4.9	0.0	0.4
4.8	0.0	0.3	4.8	1.9	3.7	0.0	0.0	0.5	0.0	3.1	5.2	0.7	2.1	3.9
2.2	0.4	0.0	2.2	0.5	1.5	0.3	0.4	0.0	0.3	1.1	2.6	2.4	0.6	1.7
2.9	0.1	0.0	2.9	0.8	2.1	0.1	0.2	0.1	0.1	1.6	3.3	1.8	0.9	2.3
0.7	1.6	0.7	0.7	0.0	0.3	1.4	1.7	0.4	1.4	0.2	0.9	4.9	0.0	0.4

The construction of the cumulative distance matrix $\mathbf{D_{NxM}}$ is as follows.

$$D_{n,m}=d_{n,m}+\min\left(D_{n-1,m},D_{n-1,m-1},D_{n,m-1}\right) \qquad 1\leq m\leq M,\ 1\leq n\leq N \tag{2.4}$$

Dynamic programing requires each element of **D** to be calculated by the addition of the local distance between the elements at any point to the minimum of the cumulative distances of the three preceding points as shown in Equation 2.4. Therefore, the first element $D_{1,1}$ becomes the first element of the local distance matrix $d_{1,1}$ since there are no preceding points. $D_{1,2}$ is calculated by using Equation 2.5 and the first row $D_{1,m}$ is filled in the same manner, the only available point being the previous element of the same row. Then, the first column $D_{n,1}$ is calculated by using Equation 2.6 and starting from $D_{2,1}$. Figure 2.2 shows the construction of **D** and the optimal warping path **W**. In the figure, → represents the route to a specific point from the minimum preceding point. **W** is obtained by backtracking from the end point $D_{N,M}$ and ●—● represents the optimal warping path.

$$D_{1,m}=d_{1,m}+D_{1,m-1} \qquad 1\leq m\leq M \tag{2.5}$$

$$D_{n,1}=d_{n,1}+D_{n-1,1} \qquad 1\leq n\leq N \tag{2.6}$$

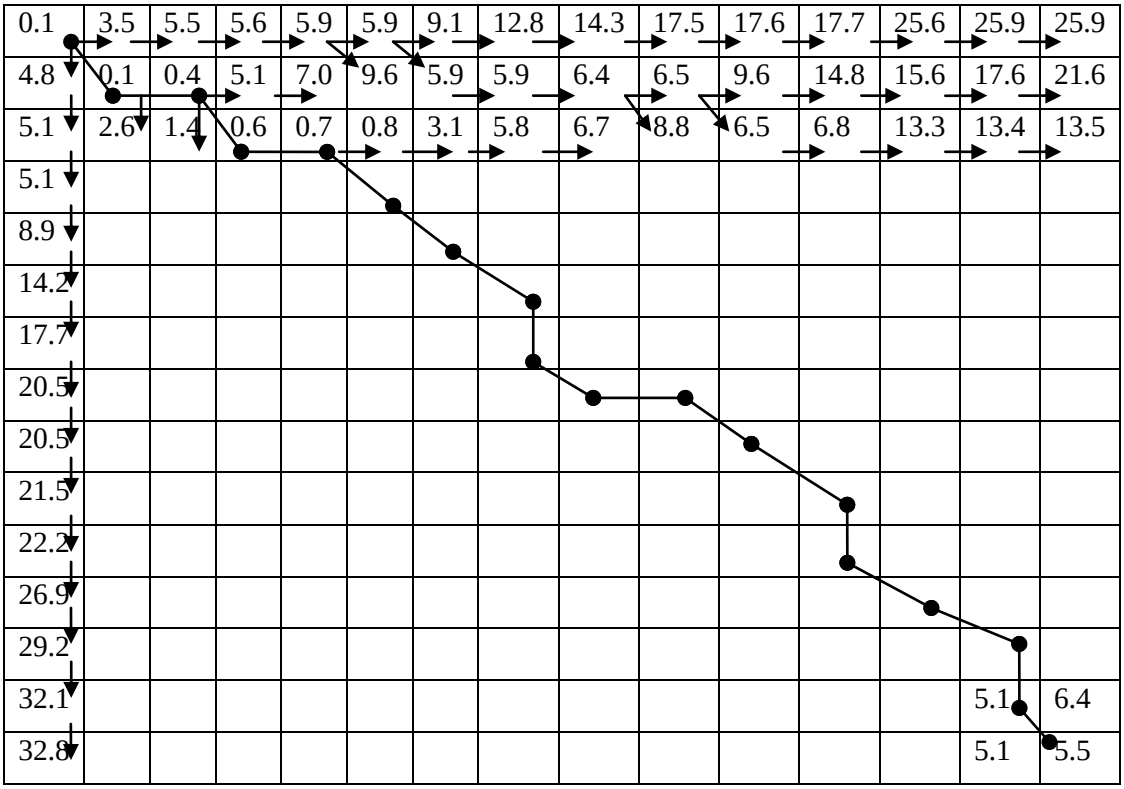

Figure 2.2. Schematic representation of the construction of **D**.

The DTW algorithm (pseudo code) is summarized in Figure 2.3.

Require t and **r** are normalized
for n = 1 to N
  for m = 1 to M **do**
    Compute the local distance vector **d** using Equation 2.3;
  end for
end for
Let D_{NxM} be the cumulative distance matrix
$D_{1,1} \Leftarrow d_{1,1}$;
for n = 2 to N
  Compute the first column $D_{n,1}$ using Equation 2.4;
end
for m = 2 to M
  Compute the first row $D_{1,m}$ using Equation 2.5;
end
for n = 2 to N
  for m = 2 to M **do**
    Compute the cumulative distance matrix **D** using Equation 2.4;
  end for
end for
while n + m ≠ 2
  number ⇐ min([D(n-1,m),D(n,m-1),D(n-1,m-1)])
  switch number
    case 1: n ⇐ n – 1
    case 2: m ⇐ m – 1
    case 3: n ⇐ n – 1
  end switch
  w ⇐ cat(1,w,[n,m])
end while

Figure 2.3. Dynamic time warping algorithm.

The following figures are the results of the DTW application. Figure 2.4 shows the normalized series **t** and **r** before and after warping. The local expansion of the series is clearly observed.

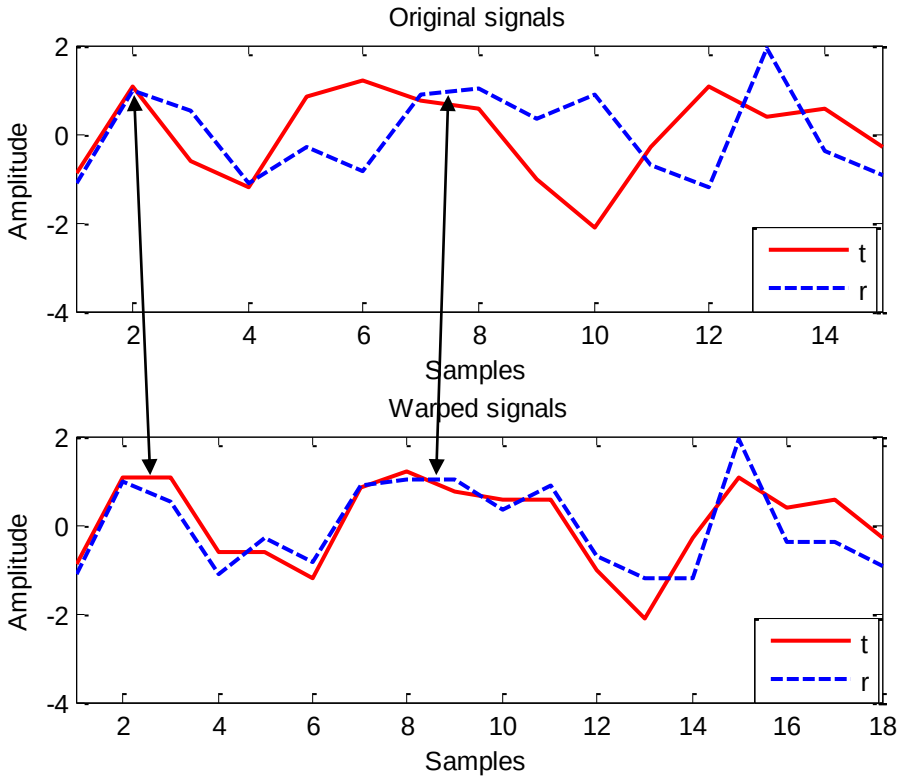

Figure 2.4. Original signals and warped signals after DTW.

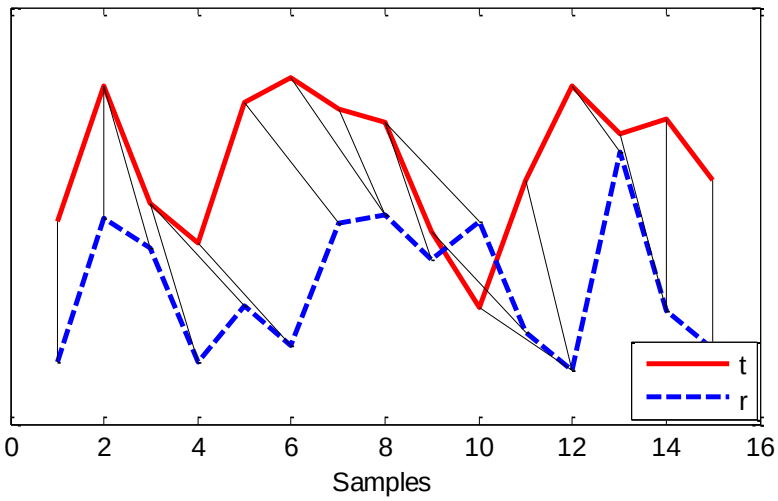

Figure 2.5. Alignment of the series.

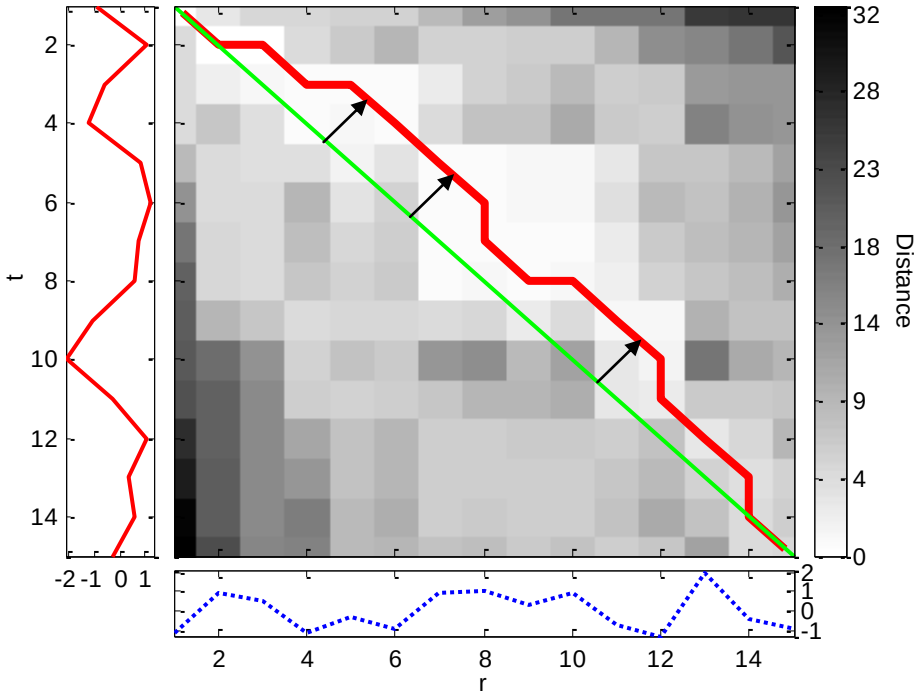

Figure 2.6. DTW distance matrix and the optimal warping path.

Figure 2.5 shows how the points are matched. Except for the first two and the last two points, all points of **r** are matched with previous points of **t**. The cumulative distance matrix **D**, the diagonal line and the optimal warping path **W** are pictured in Figure 2.6. It is clearly seen that the optimal warping path lies above the diagonal line meaning that the series **t** precedes the series **r**.

The mode of the time shifts between the series in the warping path is −2 and the mean of the time shifts is −1, indicating that the series **t** leads the series **r** (**t** ⇒ **r**). The total distance is 5.479 and the length of the optimal path is 18. Moreover, the correlation between the unwarped series is measured as 0.026, however after warping the correlation becomes 0.836.

In this thesis, ⇒, ⇐ and ⇔ are used as the symbolic representations of temporal precedence directions from the left hand side to the right hand side, from the right hand side to the left hand side, and a two-way varying relation, respectively.

In this example, the direction of the lead/lag relation between the series does not change consequently the warping path does not cross the diagonal line. However, this is also a possibility and in such cases examination of the warping path can indicate the regions of different relations separately. Although in some cases with large number of data points or for very similar series, drawing conclusions from the figures may not be easy; therefore the mode and mean of the time shifts are also important indications. Negative values of time shifts suggest that the first series leads the second series and vice versa. It should be noted that, these values do not provide insight on the regional behavior of the series.

3. GRANGER CAUSALITY

In Section 3.1 of this chapter, bivariate Granger causality (GC) method is explained. In Section 3.2, the algorithm used in the applications in this work is provided and it is demonstrated on a simple example in Section 3.3. Even though not directly related to GC, Section 3.4 briefly presents the cross-correlation method also used in this thesis work.

3.1. GC Problem Formulation and Algorithm

GC investigates whether there is a causal relationship between two time series and, if so, the direction of the causation. The GC test is a linear test and it assumes linear causal relationships between the series (Bai *et al*., 2010). The test consists of forming regression models and testing the significance of the coefficients of the variables in these models. By Granger's definition, if the inclusion of the past values of a time series **Y** in the regression model for the time series **X** improves the prediction of **X**, then the series **Y** is said to be Granger causing **X** (Granger, 1969). Then, the test can be conducted for the reverse direction as well. The outcomes can be such that **X** Granger causes **Y** and **Y** Granger causes **X**, meaning that there is a feedback relation between **X** and **Y**. The second possible outcome is the case where **X** Granger causes **Y** however, **Y** does not Granger cause **X**. The reverse of this result is the third outcome namely **Y** Granger causes **X** and not the other way around. Lastly, there may be no causal relationships between the series.

The series are required to be stationary for the application of the GC test (Granger, 1969). For this purpose, unit root and stationarity tests such as the augmented Dickey-Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests can be used to determine the order of integration and check the stationarity of the series (Lim and Yoo, 2012; Giles, 2011). In case of non-stationarity, differences (derivatives; i.e., $X_t - X_{t-1}$ and $Y_t - Y_{t-1}$) of both series are taken before the test. Moreover, Engle-Granger cointegration test (Engle and Granger, 1987) is used to check the cointegration.

For a bivariate test, Equation 3.1 and Equation 3.2 describe the auto-regressive models for **X** and **Y** respectively. X_{t-j}, and Y_{t-j} are the past values of the series, m and p are

the optimal lags used in the regression models, and $\varepsilon_{1,t}$ and $\eta_{1,t}$ are the prediction errors (preferably white noise). These equations are also called the *restricted models*.

$$X_t = \sum_{j=1}^{m} a_{1,j} X_{t-j} + \varepsilon_{1,t} \tag{3.1}$$

$$Y_t = \sum_{j=1}^{p} c_{1,j} Y_{t-j} + \eta_{1,t} \tag{3.2}$$

In order to check Granger causality, cross-dependence variables (past values of **Y** in **X** equation and past values of **X** in **Y** equation) are included in these models. These alternative models are called the *unrestricted models* for **X** and **Y**, and are depicted in Equation 3.3 and Equation 3.4 respectively.

$$X_t = \sum_{j=1}^{m} a_{2,j} X_{t-j} + \sum_{j=1}^{k} b_{2,j} Y_{t-j} + \varepsilon_{2,t} \tag{3.3}$$

$$Y_t = \sum_{j=1}^{p} c_{2,j} Y_{t-j} + \sum_{j=1}^{l} d_{2,j} X_{t-j} + \eta_{2,t} \tag{3.4}$$

The null hypotheses of 'no causality' are expressed as follows:

$$H_0^1: b_{2,1} = \ldots = b_{2,k} = 0$$

$$H_0^2: d_{2,1} = \ldots = d_{2,l} = 0$$

The hypotheses are tested by conducting an F-test for the restricted and unrestricted models. A value of the F-statistic greater than the critical value F_c means that the inclusion of another variable, **Y** to the regression model of **X** decreases the variance of the prediction error, thus improves the prediction. For such a case, H_0 is rejected and it is concluded that **Y** Granger causes **X** (Krishna *et al*., 2011). Equation 3.5 shows the F-statistic for the test of H_0^1.

$$F = \frac{\left(\sum_{t=1}^{n} \varepsilon_{1,t}^2 - \sum_{t=1}^{n} \varepsilon_{2,t}^2\right)/m}{\sum_{t=1}^{n} \varepsilon_{2,t}^2 / \left(n-(m+k+1)\right)} \tag{3.5}$$

where $\sum_{t=1}^{n} \varepsilon_{1,t}^2$ and $\sum_{t=1}^{n} \varepsilon_{2,t}^2$ are the residual sums of squares for the restricted and unrestricted models respectively. m is the number of lags of **X**, k is the number of lags of **Y**

and n is the number of data points. The critical value $F_{\alpha,m,n-(m+k+1)}$ can be looked up from tabulated values (Peng *et al*., 2008).

Bayesian Information Criterion (BIC) is used in the determination of the optimum number of lags to be included in the regression models. The lag that minimizes BIC is chosen for each variable. Equation 3.6 shows the calculation of BIC.

$$BIC=n \ln\left(\frac{RSS}{n}\right)+k \ln(n) \tag{3.6}$$

where RSS is the residual sum of squares, k is the number of regressors and n is the number of data points.

The algorithm (pseudo code) is summarized in Figure 3.1.

Require X and **Y** are normalized
Check stationarity and cointegration;
Let MaxLag be the maximum number of lags to be considered
for i=1 to MaxLag **do**
  Compute the coefficient estimates, residuals of Equation 3.1 by linear regression;
end
Determine the optimal lag length of **X** with BIC using Equation 3.6;
for i=1 to MaxLag **do**
  Compute the coefficient estimates, residuals of Equation 3.3 by linear regression;
end for
Determine the optimal lag length of **Y** with BIC using Equation 3.6;
Compute F using Equation 3.5;
if $F > F_c$ **then**
  Reject the null hypothesis;
else
  Cannot reject the null hypothesis;
end if

Figure 3.1. The GC test algorithm.

3.2. A Simple Example of GC Application

In this section, the causal relationship between Gross Domestic Product (GDP) and Consumption (CS) is investigated by the help of the GC algorithm explained in the previous section. The series are taken from EViews (a widely used econometrics software) data file "Chow_var" (IHS, Inc., 2011). Each series contains 193 data points. The data are illustrated in Figure 3.2.

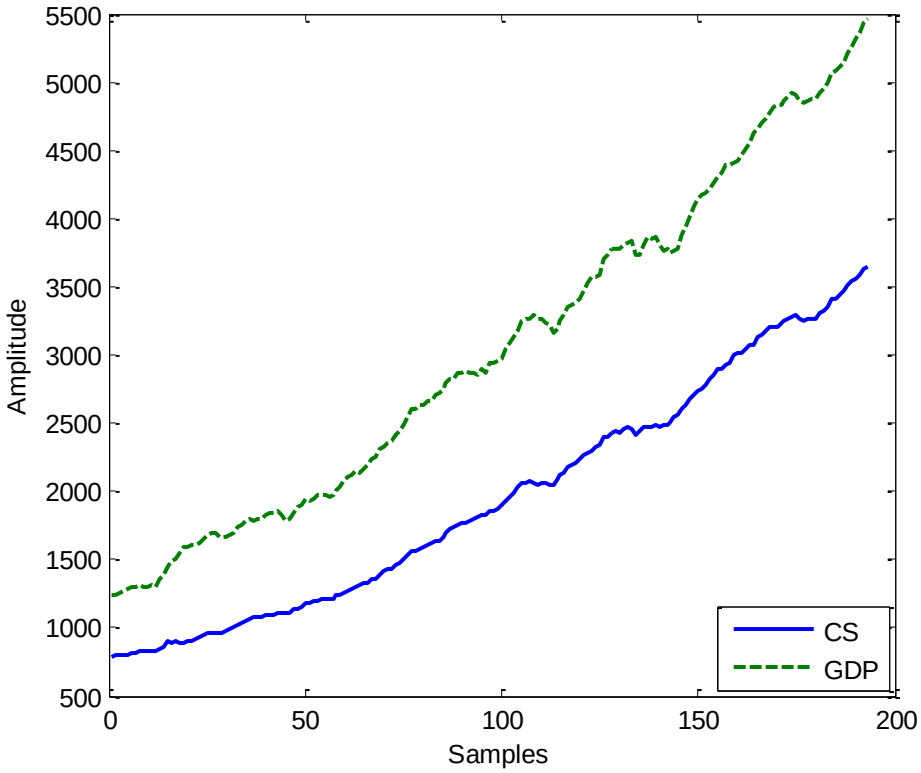

Figure 3.2. Consumption and GDP data.

Z-score normalization is applied before the GC test. Series are non-stationary and not cointegrated, first order difference is taken for the GC test to make the series stationary with respect to ADF test. The normalized stationary series are depicted in Figure 3.3. The stationarity and cointegration tests were performed using the Econometrics Toolbox of MATLAB using `adftest` and `egcitest` functions. Significance level for the GC test is 0.95.

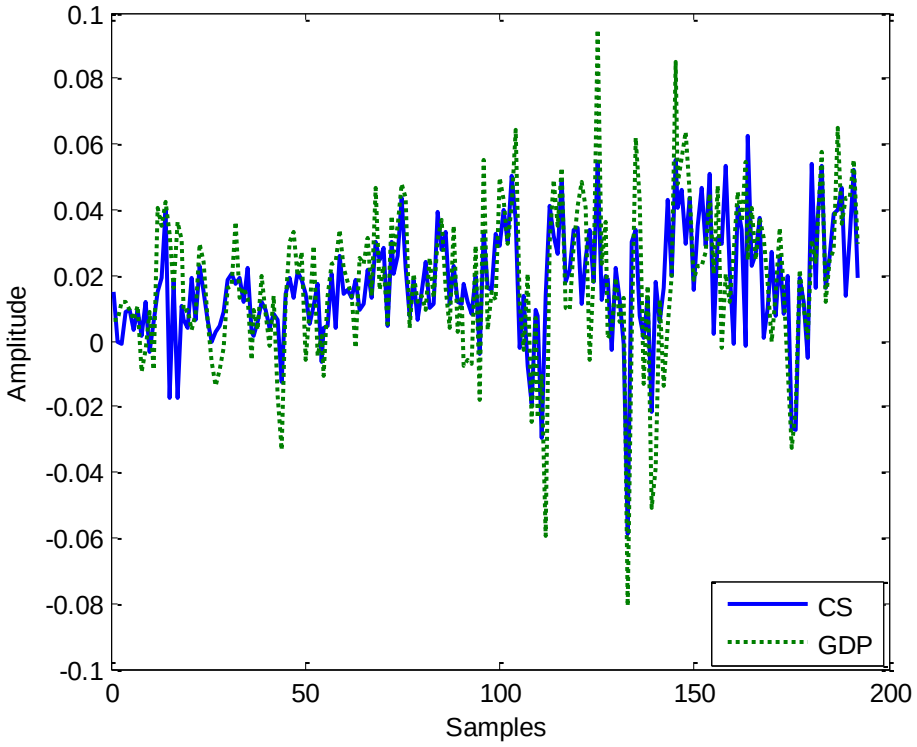

Figure 3.3. Consumption and GDP data after normalization and differentiation.

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on the Bayesian Information Criterion showed the optimum number of lags to be used in the GC tests as three, one and one, two for CS–GDP and GDP–CS pairs, respectively. The results of the GC test are as follows:

Table 3.1. GC test results for the CS and GDP data.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
GDP does not Granger cause CS	3.892	0.275	Do not reject H_0
CS does not Granger cause GDP	3.044	10.762	Reject H_0

From the results in Table 3.1 it is seen that since F is > F_c consumption causes GDP (CS → GDP) and not the other way around. This is meaningful and as expected since it is generally known that public consumption increases country's production which in turn increases country's GDP.

In this thesis, →, ← and ↔ are used as the symbolic representations of causality directions from the left hand side to the right hand side, from the right hand side to the left hand side, and a two-way relation, respectively.

3.3. Cross-Correlation Calculation

Sample cross-correlation r between two series **X** and **Y** is estimated by using Equation 3.7.

$$r_{\mathbf{XY}}(k)=\frac{c_{\mathbf{XY}}(k)}{s_{\mathbf{X}}s_{\mathbf{Y}}} \qquad k=0,\pm1,\pm2,\ldots \tag{3.7}$$

where $c_{\mathbf{XY}}(k)$ is the cross-covariance between **X** and **Y** at lag k and $s_{\mathbf{X}}$ and $s_{\mathbf{Y}}$ are the sample standard deviations of **X** and **Y** respectively. Computation of the cross-covariance can be performed in either of the two ways shown in Equation 3.8,

$$c_{\mathbf{XY}}(k)=\begin{cases}\frac{1}{n}\sum_{t=1}^{n-k}(X_t-\overline{X})\,(Y_{t+k}-\overline{Y}) & k=0,1,2,\ldots\\ \frac{1}{n}\sum_{t=1}^{n+k}(Y_t-\overline{Y})\,(X_{t-k}-\overline{X}) & k=0,-1,-2,\ldots\end{cases} \tag{3.8}$$

where n is the number of samples, and $\overline{X}$ and $\overline{Y}$ are the sample means of **X** and **Y**. The sample standard deviations are the square roots of auto-covariances at zero lag as depicted in Equation 3.9 and Equation 3.10 (The MathWorks, Inc., 2012).

$$s_{\mathbf{X}}=\sqrt{c_{\mathbf{XX}}(0)} \tag{3.9}$$

$$s_{\mathbf{Y}}=\sqrt{c_{\mathbf{YY}}(0)} \tag{3.10}$$

In this thesis cross-correlations between series at varying lags are computed using MATLAB's `crosscorr` function.

4. APPLICATIONS AND COMPARISON OF DTW AND GC

In this chapter, the Dynamic Time Warping (DTW) and the Granger Causality (GC) methods will be applied on relatively wide spectrum of data sets and the methods will be compared in terms of the causality between the series and the direction of the leads or lags. The data sets used are as follows:

- Time series with interval-dependent varying lags
- Multidimensional linear time series models
- Multidimensional nonlinear time series models
- Hair dryer input/output model
- Economic indexes and the Chemical Engineering Plant Cost Index (CEPCI)
- Predator-prey model
- Biochemical system models
- Chemical reaction models
- Shifted series

4.1. Illustrative Examples of Time Series with Interval-Dependent Varying Lags

The two examples in this section demonstrate the role of DTW in uncovering the changes in the lead/lag structures of the time series. Two time series are used and they are constructed as follows: one of the series is a first order AR-process with Gaussian noise, the second series is created from the first one and it either simultaneously occurs, lags behind, or leads the first series (Sornette and Zhou, 2005). According to the interval-dependent data generation models (Equations 4.1 and 4.2), the expected results are that, for the first and last intervals there are no lead/lag relations between the series, for the second and third intervals the series x leads the series y, and for the fourth interval the series y leads the series x. The series are created by using MATLAB.

4.1.1. First Example with Interval-Dependent Varying Lags

In this example, the coefficient a in Equation 4.1 is taken as 0.7 and b in Equation 4.2 is taken as 0.8. ε_1 and ε_2 are noises with zero mean and with standard deviations 0.5 and 0.1, respectively. The lag between the series y and x is changed in every 100 time steps. Some additional parameters and properties of the data and tests are as follows: Each series is made of 500 data points. Z-score normalization is applied before DTW, cross-correlation calculations, and the GC test. Series are stationary with respect to ADF test and cointegrated. Significance level for the GC test is 0.95.

$$x(n)=ax(n-1)+\varepsilon_1 \tag{4.1}$$

$$y(n)=\begin{cases} bx(n)+\varepsilon_2, & 1\leq n\leq 100, \\ bx(n-10)+\varepsilon_2, & 101\leq n\leq 200, \\ bx(n-5)+\varepsilon_2, & 201\leq n\leq 300, \\ bx(n+5)+\varepsilon_2, & 301\leq n\leq 400, \\ bx(n)+\varepsilon_2, & 401\leq n\leq 500. \end{cases} \tag{4.2}$$

The results of the tests for the entire data range, interval $1 \leq n \leq 500$, are as follows:

Table 4.1. GC test results for the first data set with interval-dependent lags: the entire data range.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause x	2.232	6.513	Reject H_0
x does not Granger cause y	1.850	9.036	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 10, and the automated analysis based on the Bayesian Information Criterion (BIC) showed the optimum number of lags to be used in the GC tests as one, five and one, 10 for x – y and y – x pairs, respectively.

The mode of the time shifts between the series in the warping path is zero, indicating that there is no lead/lag relation between the series and the mean of the time shifts is −2, indicating that the series x leads y ($x \Rightarrow y$) in the entire data range.

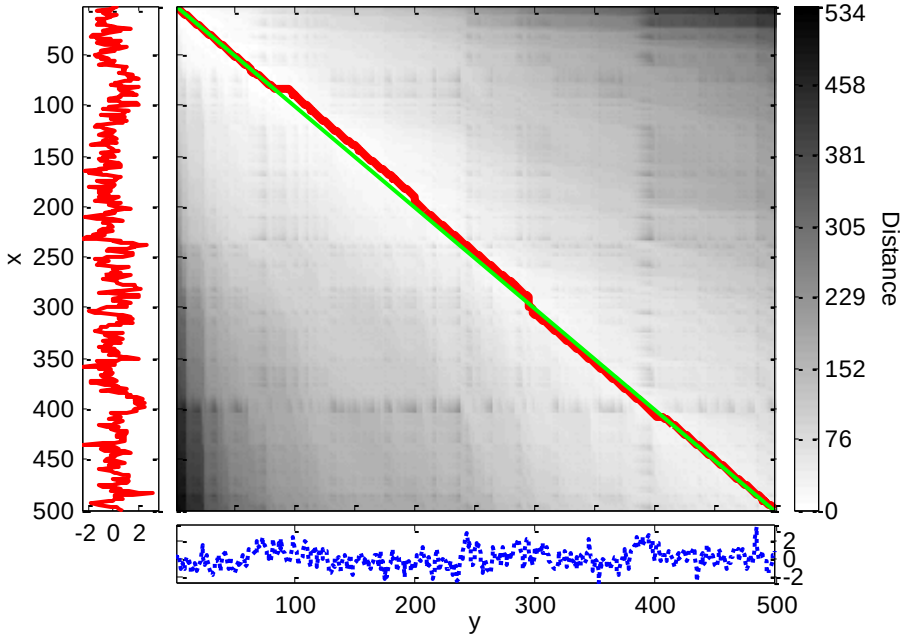

Figure 4.1. DTW distance matrix and the optimal warping path for the first data set with interval-dependent lags: the entire data range.

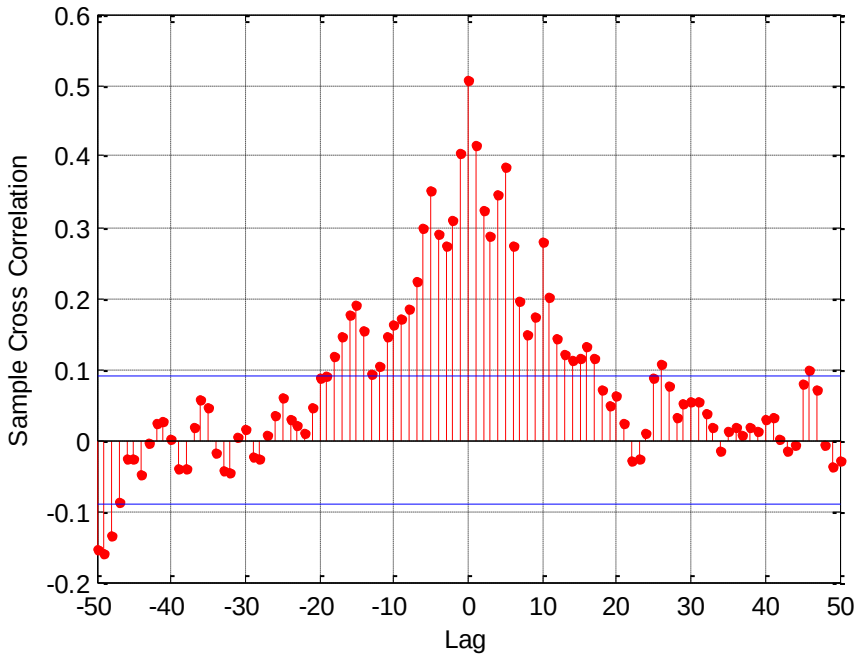

Figure 4.2. Sample cross-correlation function for the first data set with interval-dependent lags: the entire data range.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relationship between the series x and the series y.

The results of the tests for the interval $1 \leq n \leq 100$ are as follows:

Table 4.2. GC test results for the first data set with interval-dependent lags: $1 \leq n \leq 100$.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause x	3.939	0.264	Do not reject H_0
x does not Granger cause y	3.939	5.728	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 10, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as one, one and one, one for x – y and y – x pairs, respectively.

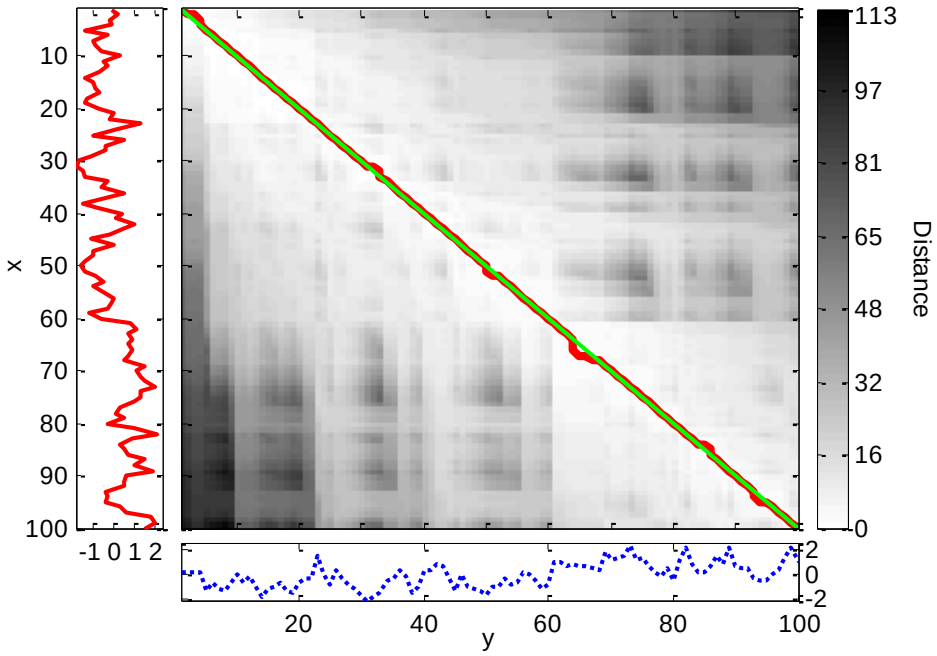

Figure 4.3. DTW distance matrix and the optimal warping path for the first data set with interval-dependent lags: $1 \leq n \leq 100$.

The mode and the mean of the time shifts between the series in the warping path are zero, indicating that there is no lead/lag relation between the series x and the series y in the interval $1 \leq n \leq 100$.

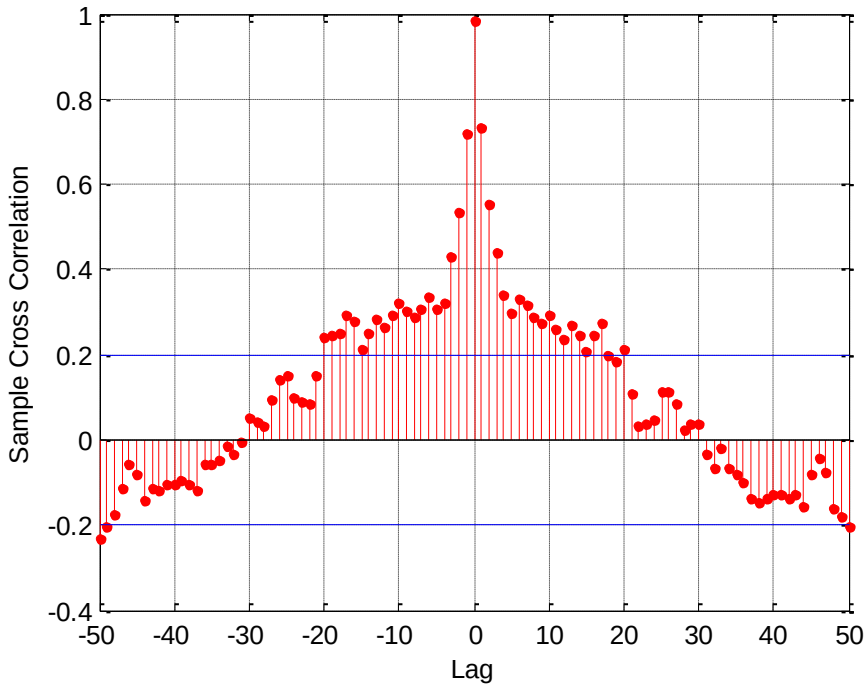

Figure 4.4. Sample cross-correlation function for the first data set with interval-dependent lags: $1 \leq n \leq 100$.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relationship between the series x and the series y in the interval $1 \leq n \leq 100$.

The results of the tests for the interval $101 \leq n \leq 200$ are as follows:

Table 4.3. GC test results for the first data set with interval-dependent lags: $101 \leq n \leq 200$.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause x	3.939	0.357	Do not reject H_0
x does not Granger cause y	1.940	160.646	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 10, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as one, one and one, 10 for x – y and y – x pairs, respectively.

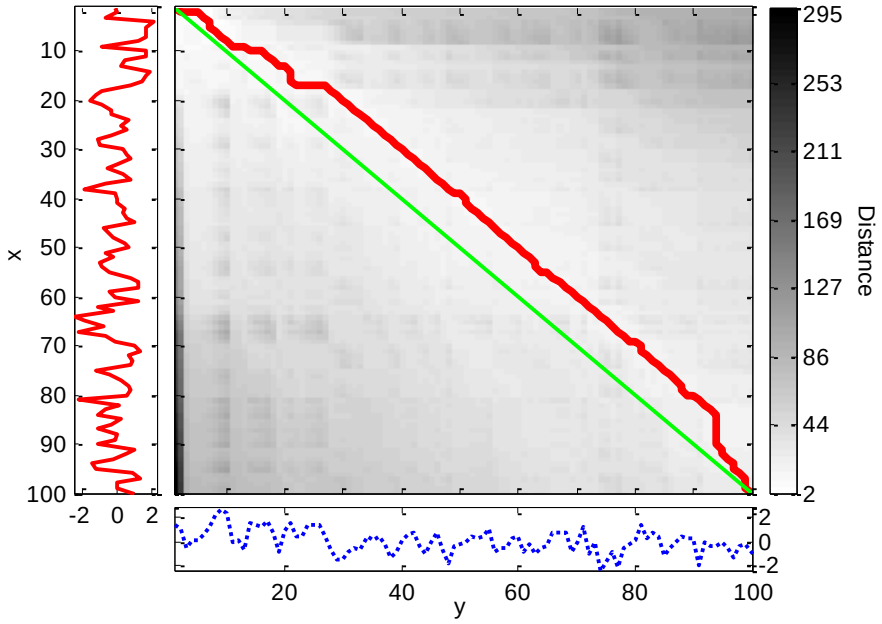

Figure 4.5. DTW distance matrix and the optimal warping path for the first data set with interval-dependent lags: $101 \leq n \leq 200$.

The mode of the time shifts between the series in the warping path is −10 and the mean of the time shifts is −8, indicating that the series x leads the series y ($x \Rightarrow y$) in the interval $101 \leq n \leq 200$.

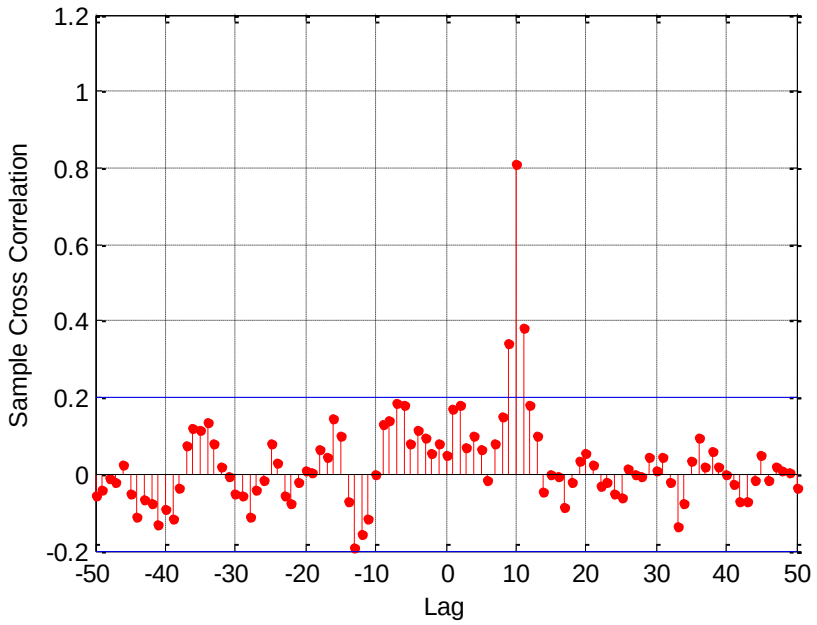

Figure 4.6. Sample cross-correlation function for the first data set with interval-dependent lags: $101 \leq n \leq 200$.

The maximum value of cross-correlation is observed at lag 10, indicating that the series x leads the series y (x ⇒ y) in the interval $101 \leq n \leq 200$.

The results of the tests for the interval $201 \leq n \leq 300$ are as follows:

Table 4.4. GC test results for the first data set with interval-dependent lags: $201 \leq n \leq 300$.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause x	3.939	0.098	Do not reject H_0
x does not Granger cause y	2.312	368.503	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 10, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as one, one and one, five for x – y and y – x pairs, respectively.

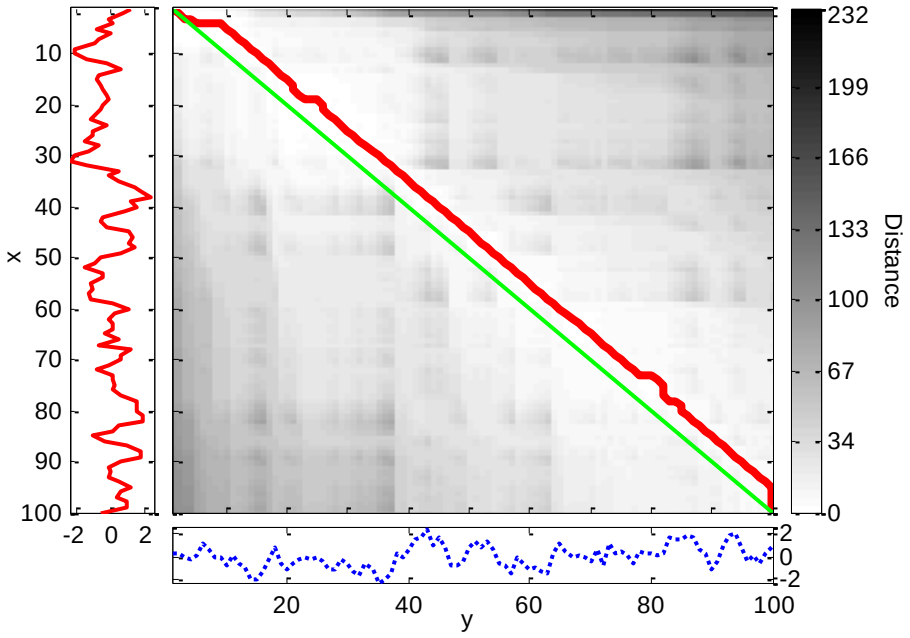

Figure 4.7. DTW distance matrix and the optimal warping path for the first data set with interval-dependent lags: $201 \leq n \leq 300$.

The mode and the mean of the time shifts between the series in the warping path are −5, indicating that the series x leads the series y (x ⇒ y) in the interval $201 \leq n \leq 300$.

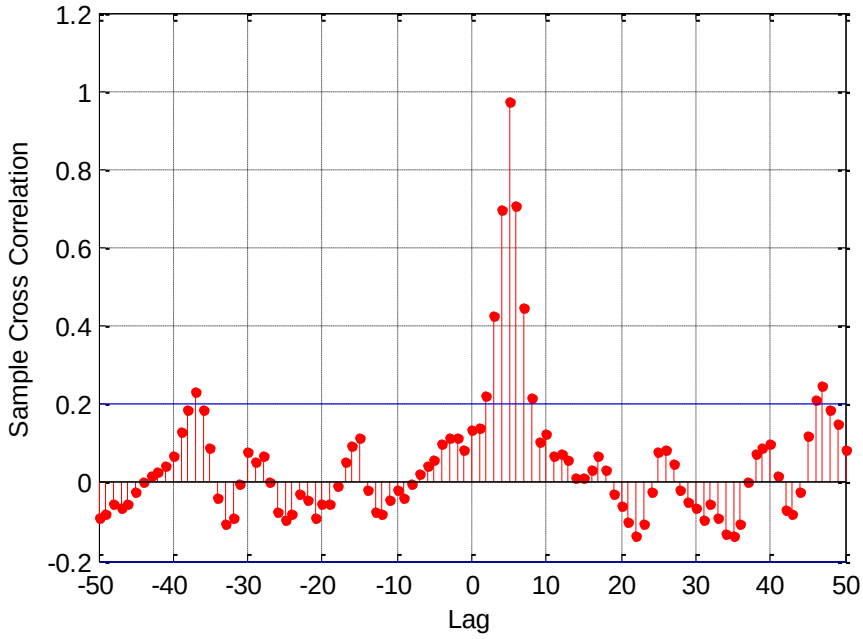

Figure 4.8. Sample cross-correlation function for the first data set with interval-dependent lags: $201 \leq n \leq 300$.

The maximum value of cross-correlation is observed at lag five, indicating that the series x leads the series y ($x \Rightarrow y$) in the interval $201 \leq n \leq 300$.

The results of the tests for the interval $301 \leq n \leq 400$ are as follows:

Table 4.5. GC test results for the first data set with interval-dependent lags: $301 \leq n \leq 400$.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause x	2.312	485.372	Reject H_0
x does not Granger cause y	3.939	0.902	Do not reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 10, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as one, five and one, one for x – y and y – x pairs, respectively.

The mode and the mean of the time shifts between the series in the warping path are five, indicating that the series x lags behind the series y ($x \Leftarrow y$) in the interval $301 \leq n \leq 400$.

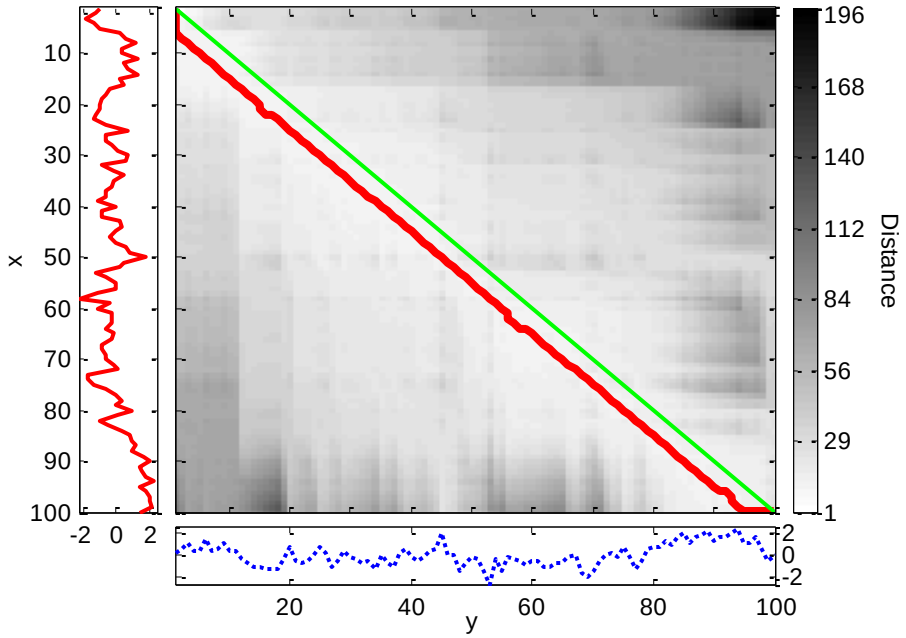

Figure 4.9. DTW distance matrix and the optimal warping path for the first data set with interval-dependent lags: $301 \leq n \leq 400$.

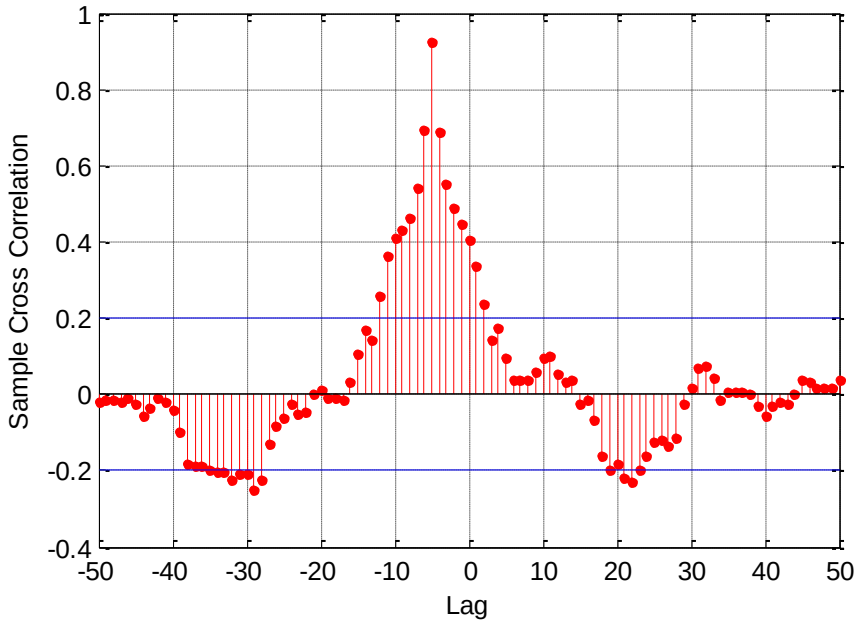

Figure 4.10. Sample cross-correlation function for the first data set with interval-dependent lags: $301 \leq n \leq 400$.

The maximum value of cross-correlation is observed at lag −5, indicating that the series x lags behind the series y ($x \Leftarrow y$) in the interval $301 \leq n \leq 400$.

The results of the tests for the interval $401 \leq n \leq 500$ are as follows:

Table 4.6. GC test results for the first data set with interval-dependent lags: $401 \leq n \leq 500$.

Null Hypothesis (H_0)	Critical Value	F-Statistics	Result
y does not Granger cause x	3.940	1.216	Do not reject H_0
x does not Granger cause y	3.940	1.535	Do not reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 10, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as two, one and two, one for x – y and y – x pairs, respectively.

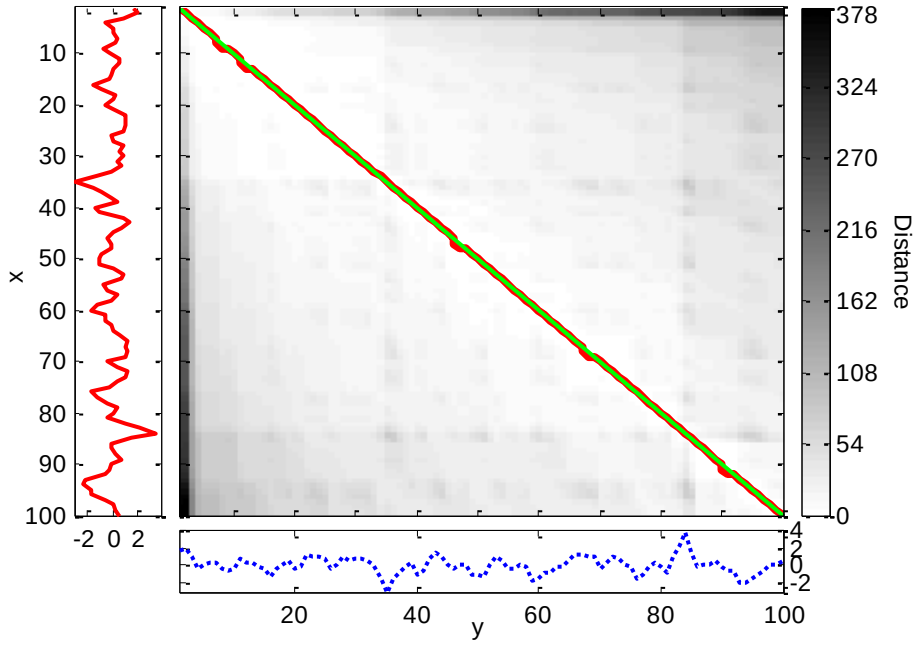

Figure 4.11. DTW distance matrix and the optimal warping path for the first data set with interval-dependent lags: $401 \leq n \leq 500$.

The mode and the mean of the time shifts between the series in the warping path are zero, indicating that there is no lead/lag relation between the series x and the series y in the interval $401 \leq n \leq 500$.

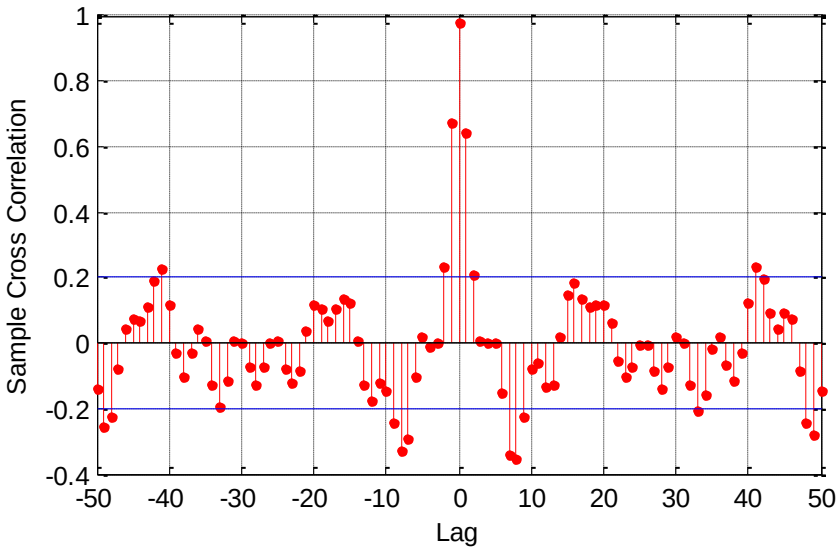

Figure 4.12. Sample cross-correlation function for the first data set with interval-dependent lags: $401 \leq n \leq 500$.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relation between the series x and the series y in the interval $401 \leq n \leq 500$.

Table 4.7 summarizes the results of the GC test, DTW, and cross-correlation analyses, stated by Table 4.1 through Table 4.6 and Figure 4.1 through Figure 4.12.

Table 4.7. Summary of the results for the first data set with interval-dependent lags.

Interval	**GC**	**DTW**	**Cross-correlation**
$1 \leq n \leq 100$	$x \rightarrow y$	No precedence	No precedence
$101 \leq n \leq 200$	$x \rightarrow y$	$x \Rightarrow y$	$x \Rightarrow y$
$201 \leq n \leq 300$	$x \rightarrow y$	$x \Rightarrow y$	$x \Rightarrow y$
$301 \leq n \leq 400$	$x \leftarrow y$	$x \Leftarrow y$	$x \Leftarrow y$
$401 \leq n \leq 500$	No causality	No precedence	No precedence
$1 \leq n \leq 500$	$x \leftrightarrow y$	Partially no precedence, $x \Rightarrow y$ and $x \Leftarrow y$	No precedence

In this example, DTW method finds all relations as expected according to the interval-dependent data generation models (Equations 4.1 and 4.2). These relations between the series x and the series y can be observed by the examination of the overall data with DTW as well as each interval individually where the lead/lag relation remains

constant. From this example it is seen that DTW can discover the lead/lag relations between series even when the direction changes over time. Cross-correlation analysis also provides the correct results however, as opposed to DTW, when the data is examined as a whole, individual zones where one series leads the other cannot be detected. The GC test shows similar results in general, the only different result is obtained in the interval $1 \leq n \leq 100$ where it falsely finds that $x \rightarrow y$, whereas DTW correctly points to no precedence between the series.

4.1.2. Second Example with Interval-Dependent Varying Lags

In this example, the coefficients a and b in Equation 4.1 and Equation 4.2 are kept the same as before, however, the standard deviations of noises ε_1 and ε_2 are changed to one and 0.5, respectively. The lag between the series y and x is again varied in every 100 time steps. On the other hand, the magnitude of the lags is increased as given by Equation 4.3 and Equation 4.4. Some additional parameters and properties of the data and tests are as follows: Each series is made of 500 data points. Z-score normalization is applied before DTW, cross-correlation calculations, and the GC test. Series are stationary with respect to ADF test and cointegrated. Significance level for the GC test is 0.95.

$$x(n)=ax(n-1)+\varepsilon_1 \tag{4.3}$$

$$y(n)=\begin{cases} bx(n)+\varepsilon_2, & 1\leq n\leq 100, \\ bx(n-20)+\varepsilon_2, & 101\leq n\leq 200, \\ bx(n-10)+\varepsilon_2, & 201\leq n\leq 300, \\ bx(n+10)+\varepsilon_2, & 301\leq n\leq 400, \\ bx(n)+\varepsilon_2, & 401\leq n\leq 500. \end{cases} \tag{4.4}$$

The results of the tests for the entire data range, interval $1 \leq n \leq 500$, are as follows:

Table 4.8. GC test results for the second data set with interval-dependent lags: the entire data range.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause x	1.808	5.959	Reject H_0
x does not Granger cause y	3.860	3.983	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 20, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as three, 11 and two, one for x – y and y – x pairs, respectively.

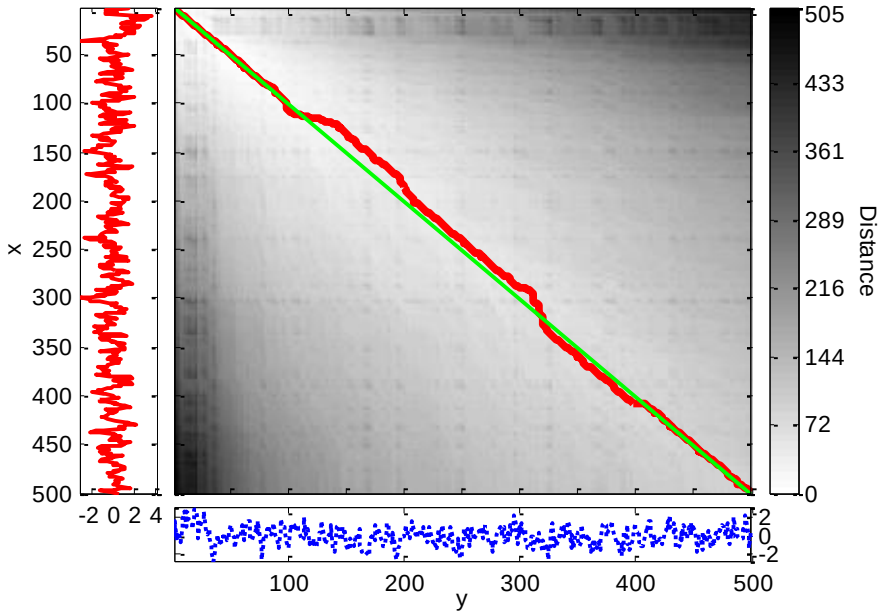

Figure 4.13. DTW distance matrix and the optimal warping path for the second data set with interval-dependent lags: the entire data range.

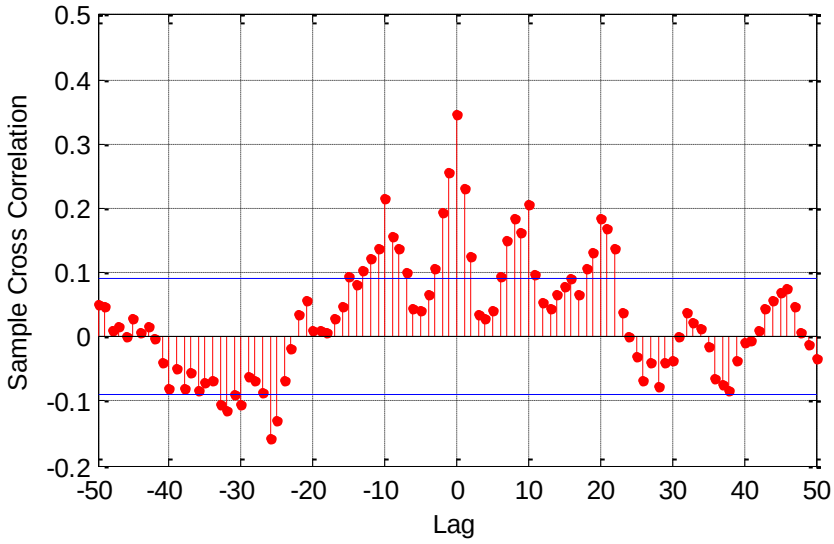

Figure 4.14. Sample cross-correlation function for the second data set with interval-dependent lags: the entire data range.

The mode of the time shifts between the series in the warping path is zero, indicating that there is no lead/lag relation between the series and the mean of the time shifts is −4, indicating that the series x leads y (x ⇒ y) in the entire data range.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relation between the series x and the series y in the entire data range.

The results of the tests for the interval $1 \leq n \leq 100$ are as follows:

Table 4.9. GC test results for the second data set with interval-dependent lags: $1 \leq n \leq 100$.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause x	3.091	6.765	Reject H_0
x does not Granger cause y	3.941	2.622	Do not reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 20, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as one, two and three, one for x – y and y – x pairs, respectively.

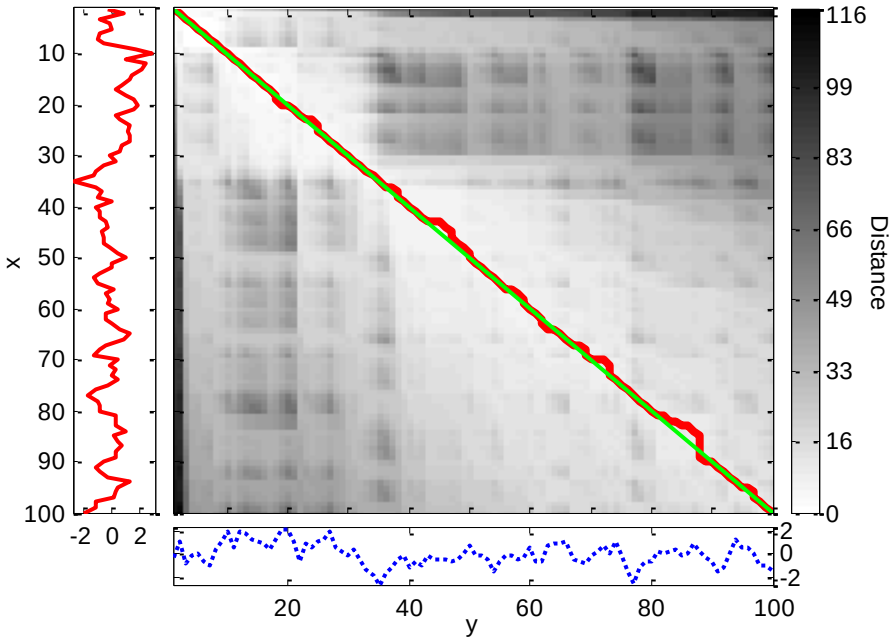

Figure 4.15. DTW distance matrix and the optimal warping path for the second data set with interval-dependent lags: $1 \leq n \leq 100$.

The mode and the mean of the time shifts between the series in the warping path are zero, indicating that there is no lead/lag relation between the series x and the series y in the interval $1 \leq n \leq 100$.

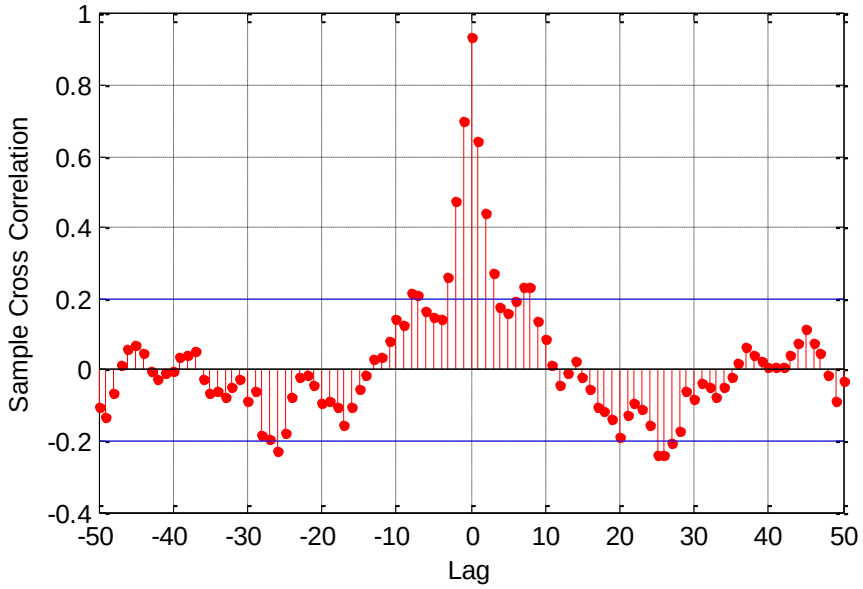

Figure 4.16. Sample cross-correlation function for the second data set with interval-dependent lags: $1 \leq n \leq 100$.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relation between the series x and the series y in the interval $1 \leq n \leq 100$.

The results of the tests for the interval $101 \leq n \leq 200$ are as follows:

Table 4.10. GC test results for the second data set with interval-dependent lags: $101 \leq n \leq 200$.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause x	3.939	1.489	Do not reject H_0
x does not Granger cause y	1.707	22.484	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 20, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as one, one and one, 20 for x – y and y – x pairs, respectively.

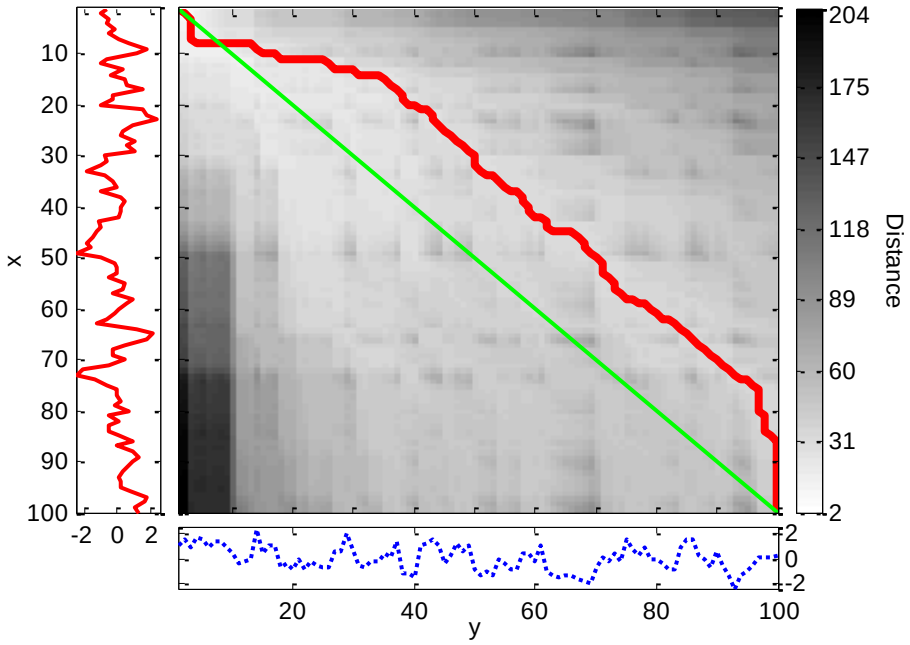

Figure 4.17. DTW distance matrix and the optimal warping path for the second data set with interval-dependent lags: $101 \leq n \leq 200$.

The mode of the time shifts between the series in the warping path is −20 and the mean of the time shifts is −14, indicating that the series x leads the series y ($x \Rightarrow y$) in the interval $101 \leq n \leq 200$.

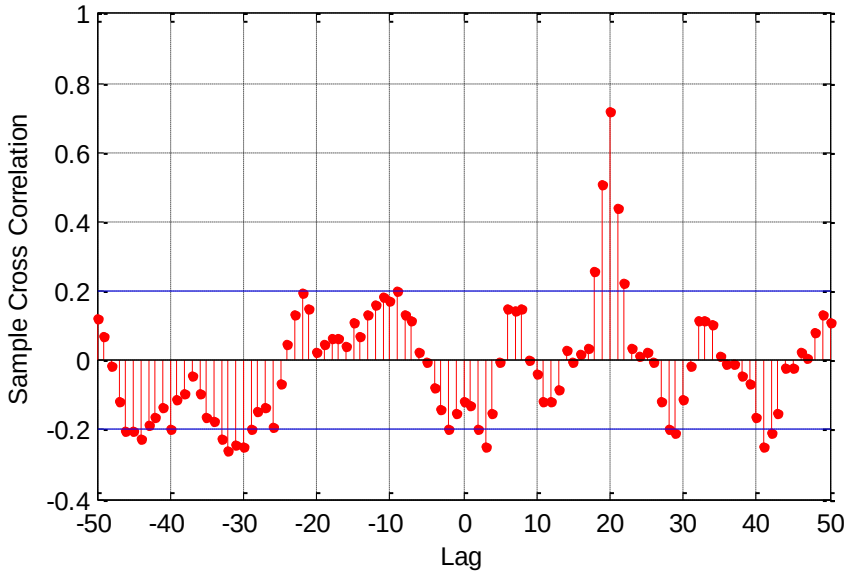

Figure 4.18. Sample cross-correlation function for the second data set with interval-dependent lags: $101 \leq n \leq 200$.

The maximum value of cross-correlation is observed at lag 20, indicating that the series x leads the series y (x ⇒ y) in the interval 101 ≤ n ≤ 200.

The results of the tests for the interval 201 ≤ n ≤ 300 are as follows:

Table 4.11. GC test results for the second data set with interval-dependent lags: 201 ≤ n ≤ 300.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause x	3.939	1.415	Do not reject H_0
x does not Granger cause y	1.940	22.386	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 20, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as one, one and one, 10 for x – y and y – x pairs, respectively.

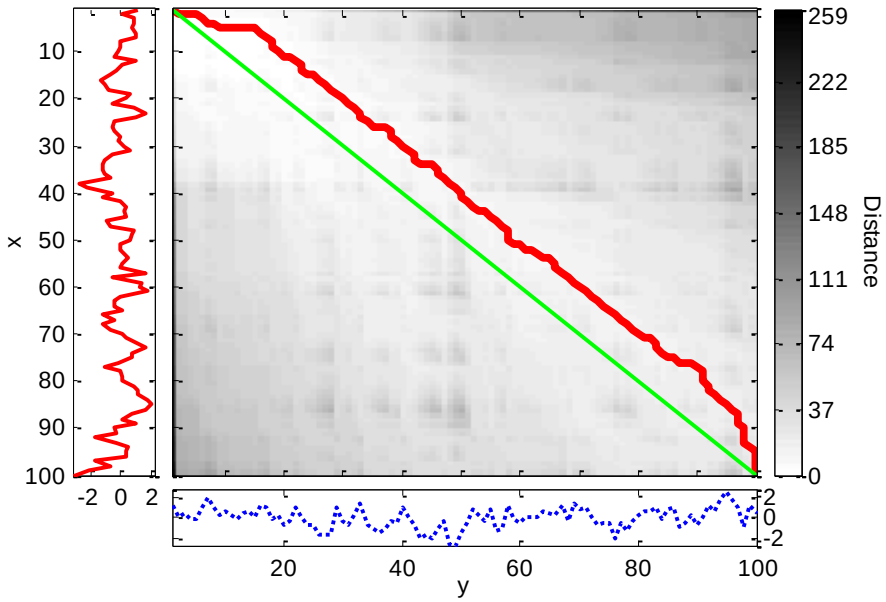

Figure 4.19. DTW distance matrix and the optimal warping path for the second data set with interval-dependent lags: 201 ≤ n ≤ 300.

The mode of the time shifts between the series in the warping path is −10 and the mean of the time shifts is −9, indicating that the series x leads the series y (x ⇒ y) in the interval 201 ≤ n ≤ 300.

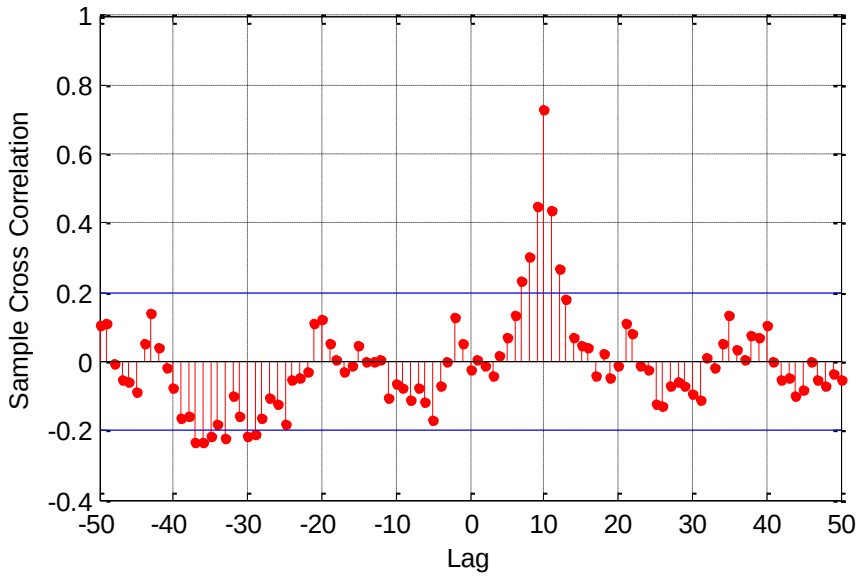

Figure 4.20. Sample cross-correlation function for the second data set with interval-dependent lags: $201 \leq n \leq 300$.

The maximum value of cross-correlation is observed at lag 10, indicating that the series x leads the series y ($x \Rightarrow y$) in the interval $201 \leq n \leq 300$.

The results of the tests for the interval $301 \leq n \leq 400$ are as follows:

Table 4.12. GC test results for the second data set with interval-dependent lags: $301 \leq n \leq 400$.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause x	1.940	32.741	Reject H_0
x does not Granger cause y	3.939	1.628	Do not reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 20, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as one, 10 and one, one for x – y and y – x pairs, respectively.

The mode of the time shifts between the series in the warping path is 10 and the mean of the time shifts is nine, indicating that the series x lags behind the series y ($x \Leftarrow y$) in the interval $301 \leq n \leq 400$.

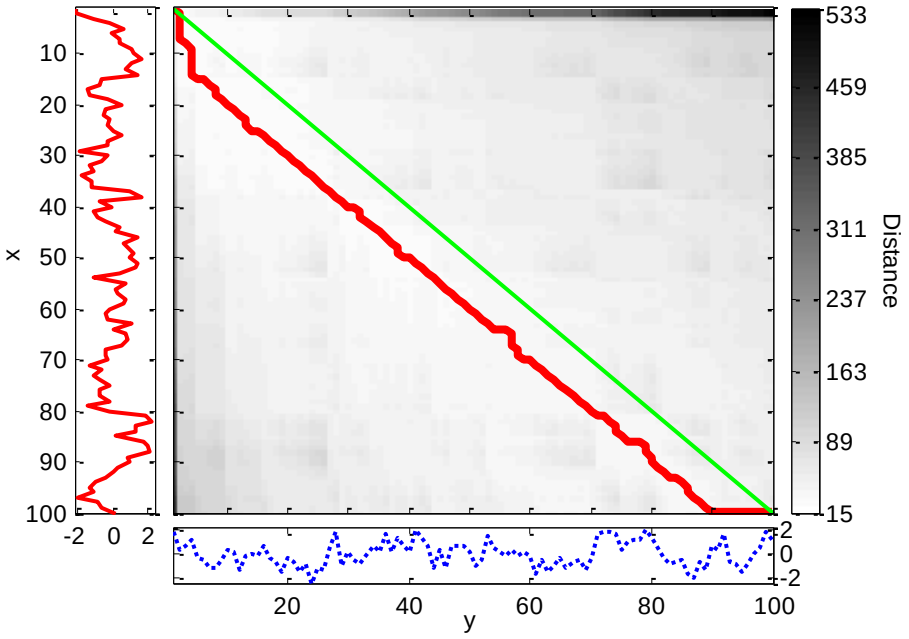

Figure 4.21. DTW distance matrix and the optimal warping path for the second data set with interval-dependent lags: $301 \leq n \leq 400$.

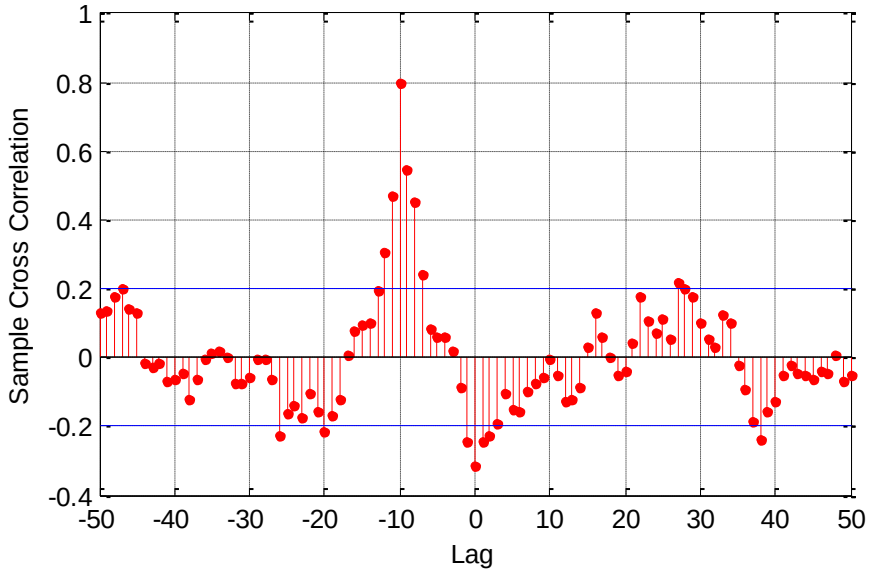

Figure 4.22. Sample cross-correlation function for the second data set with interval-dependent lags: $301 \leq n \leq 400$.

The maximum value of cross-correlation is observed at lag −10, indicating that the series x lags behind the series y (x ⇐ y) in the interval $301 \leq n \leq 400$.

The results of the tests for the interval $401 \leq n \leq 500$ are as follows:

Table 4.13. GC test results for the second data set with interval-dependent lags: $401 \leq n \leq 500$.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause x	3.941	3.507	Do not reject H_0
x does not Granger cause y	2.700	8.737	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is 20, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as three, one and one, three for x – y and y – x pairs, respectively.

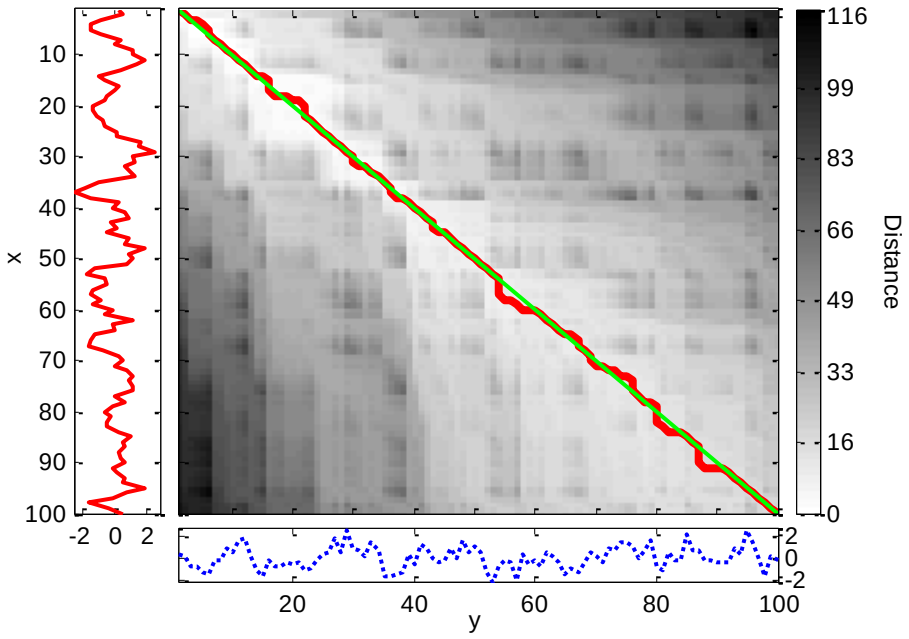

Figure 4.23. DTW distance matrix and the optimal warping path for the second data set with interval-dependent lags: $401 \leq n \leq 500$.

The mode and the mean of the time shifts between the series in the warping path are zero, indicating that there is no lead/lag relation between the series x and the series y in the interval $401 \leq n \leq 500$.

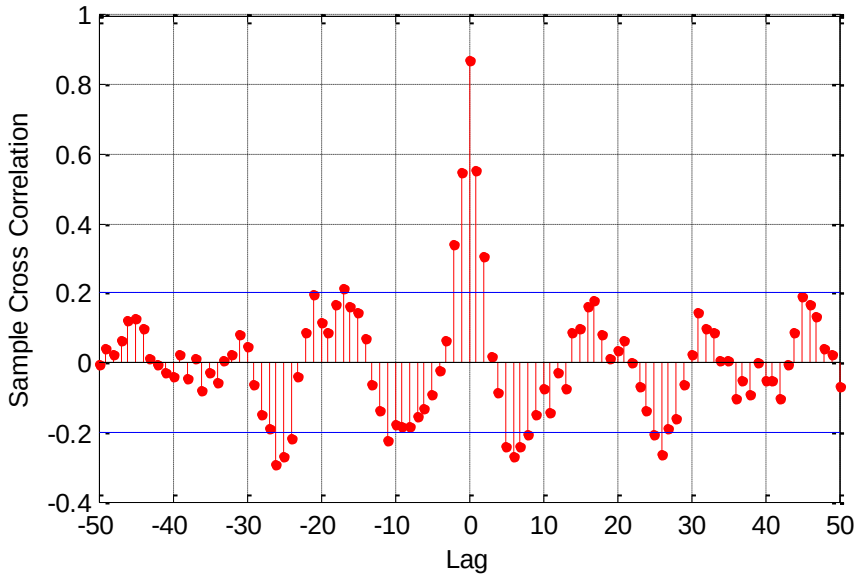

Figure 4.24. Sample cross-correlation function for the second data set with interval-dependent lags: $401 \leq n \leq 500$.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relation between the series x and the series y in the interval $401 \leq n \leq 500$.

Table 4.14 summarizes the results of the GC test, DTW, and cross-correlation analyses, stated by Table 4.8 through Table 4.13 and Figure 4.13 through Figure 4.24.

Table 4.14. Summary of the results for the time series with interval-dependent lags.

Interval	**GC**	**DTW**	**Cross-correlation**
$1 \leq n \leq 100$	$x \leftarrow y$	No precedence	No precedence
$101 \leq n \leq 200$	$x \rightarrow y$	$x \Rightarrow y$	$x \Rightarrow y$
$201 \leq n \leq 300$	$x \rightarrow y$	$x \Rightarrow y$	$x \Rightarrow y$
$301 \leq n \leq 400$	$x \leftarrow y$	$x \Leftarrow y$	$x \Leftarrow y$
$401 \leq n \leq 500$	$x \rightarrow y$	No precedence	No precedence
$1 \leq n \leq 500$	$x \leftrightarrow y$	Partially no precedence, $x \Rightarrow y$ and $x \Leftarrow y$	No precedence

The results of DTW presented in Table 4.14 are again as expected according to the interval-dependent data generation models (Equations 4.3 and 4.4). In this second example, since the magnitude of the lags is increased in the creation of the data points, the lead/lag

relation in Figure 4.13 is even clearer from the shape of the path than it is in the first example. Cross-correlation analysis results are in accordance with DTW results when the intervals are examined individually, however, when the complete data is used, the varying lead/lag relations cannot be recovered. The GC test shows similar results in general, different results are obtained only in the intervals $1 \leq n \leq 100$ and $401 \leq n \leq 500$ where it is found that $x \leftarrow y$ and $x \rightarrow y$, respectively. For these intervals DTW points to no precedence between the series.

4.2. Multidimensional Linear Time Series Models

The two examples in this section examine data sets generated from two separate multidimensional linear time series models where the direction of leads/lags between the series remains constant. The series are created by using MATLAB.

4.2.1. Five-Dimensional Linear Time Series Model

The model consists of five time-series simultaneously produced from the linear equations, Equation 4.5 through Equation 4.9. x_1 is independent of other series; x_2, x_3 and x_4 are directly caused by x_1; and there is a feedback relation between x_4 and x_5 as depicted in Figure 4.25. There are also several indirect relations in the model. ε_1, ε_2, ε_3, ε_4 and ε_5 are noises with zero mean and variance one. The coefficient of the noises, c is taken as one (Krishna and Guo, 2008). Some additional parameters and properties of the data and tests are as follows: Each series is made of 500 data points. Z-score normalization is applied before DTW, cross-correlation calculations, and the GC test. Series are stationary with respect to ADF test and cointegrated. Significance level for the GC test is 0.95.

$$x_1(n)=0.95\sqrt{2}x_1(n-1) - 0.9025x_1(n-2)+c\varepsilon_1(n) \tag{4.5}$$

$$x_2(n)=0.5x_1(n-2)+c\varepsilon_2(n) \tag{4.6}$$

$$x_3(n)=-0.4x_1(n-3)+c\varepsilon_3(n) \tag{4.7}$$

$$x_4(n)=-0.5x_1(n-2)+0.25\sqrt{2}x_4(n-1)+0.25\sqrt{2}x_5(n-1)+c\varepsilon_4(n) \tag{4.8}$$

$$x_5(n)=-0.25\sqrt{2}x_4(n-1)+0.25\sqrt{2}x_5(n-1)+c\varepsilon_5(n) \tag{4.9}$$

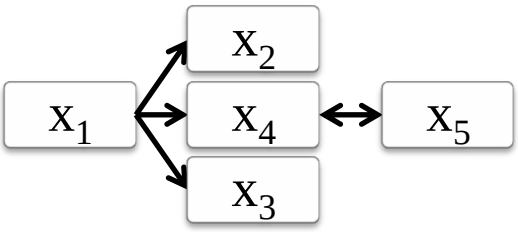

Figure 4.25. The schematic representation of the causal relations among the series in Equation 4.5 through Equation 4.9.

The results of the tests for the pair $x_1 - x_2$ are as follows:

Table 4.15. GC test results for the $x_1 - x_2$ pair of the five-dimensional example.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_2 does not Granger cause x_1	3.860	0.910	Do not reject H_0
x_1 does not Granger cause x_2	3.014	211.510	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as two, one and five, two for $x_1 - x_2$ and $x_2 - x_1$ pairs, respectively.

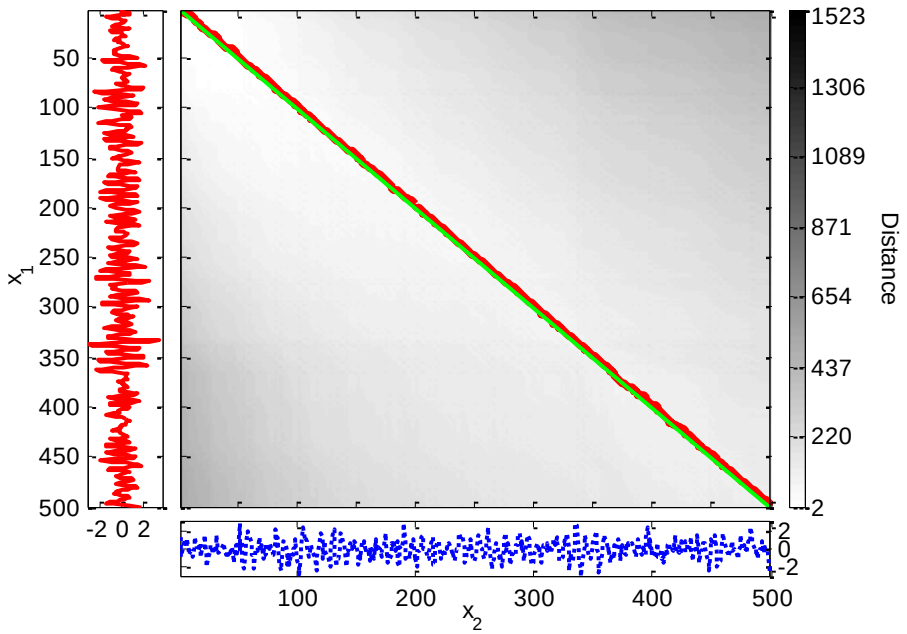

Figure 4.26. DTW distance matrix and the optimal warping path for the $x_1 - x_2$ pair of the five-dimensional example.

The mode and the mean of the time shifts between the series in the warping path are −2, indicating that the series x_1 leads the series x_2 ($x_1 \Rightarrow x_2$).

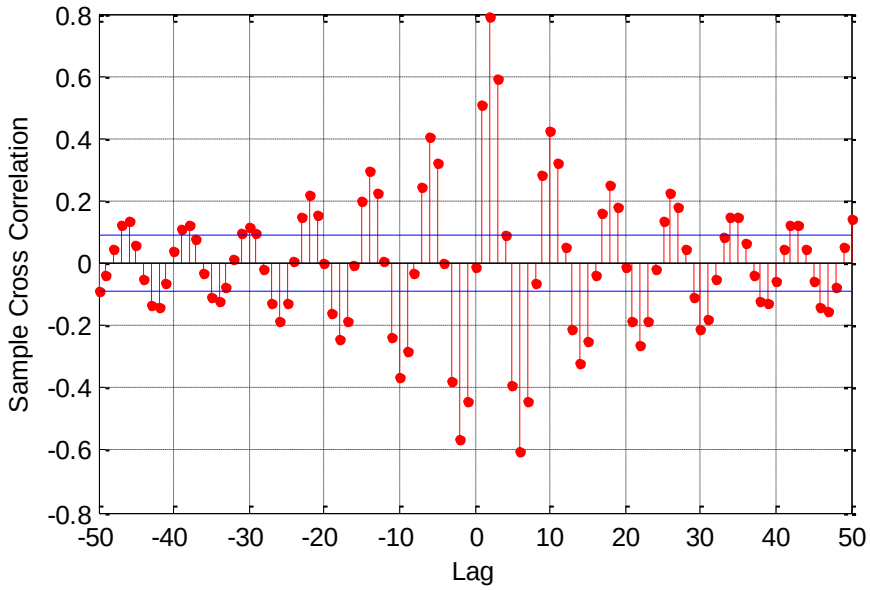

Figure 4.27. Sample cross-correlation function for the $x_1 - x_2$ pair of the five-dimensional example.

The maximum value of cross-correlation is observed at lag two, indicating that the series x_1 leads the series x_2 ($x_1 \Rightarrow x_2$).

The results of the tests for the pair $x_1 - x_3$ are as follows:

Table 4.16. GC test results for the $x_1 - x_3$ pair of the five-dimensional example.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_3 does not Granger cause x_1	3.860	1.136	Do not reject H_0
x_1 does not Granger cause x_3	2.623	99.108	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as two, one and five, three for $x_1 - x_3$ and $x_3 - x_1$ pairs, respectively.

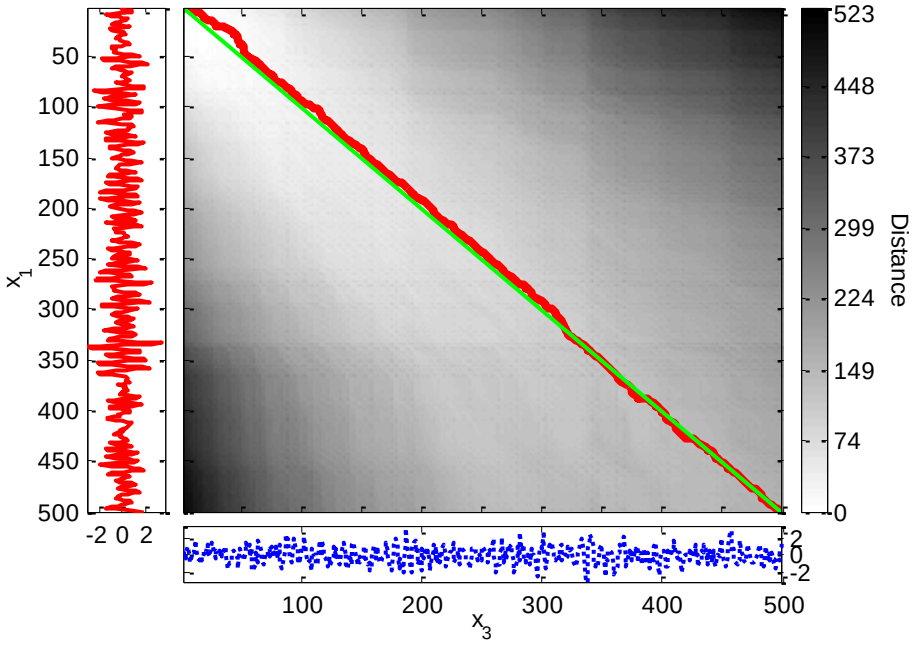

Figure 4.28. DTW distance matrix and the optimal warping path for the $x_1 - x_3$ pair of the five-dimensional example.

The mode of the time shifts between the series in the warping path is −7 and the mean of the time shifts is −4, indicating that the series x_1 leads the series x_3 ($x_1 \Rightarrow x_3$).

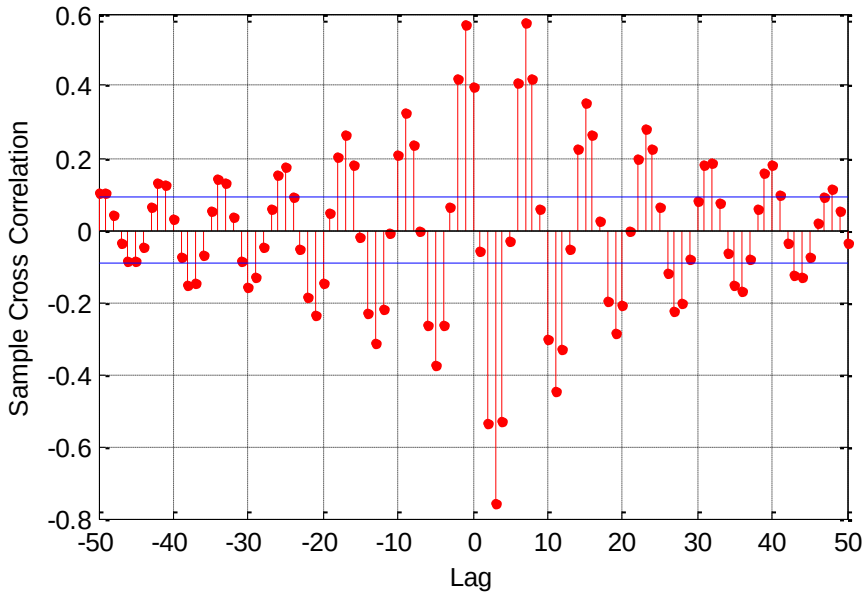

Figure 4.29. Sample cross-correlation function for the $x_1 - x_3$ pair of the five-dimensional example.

The maximum value of cross-correlation is observed at lag seven, indicating that the series x_1 leads the series x_3 ($x_1 \Rightarrow x_3$).

The results of the tests for the pair $x_1 - x_4$ are as follows:

Table 4.17. GC test results for the $x_1 - x_4$ pair of the five-dimensional example.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_4 does not Granger cause x_1	3.860	0.644	Do not reject H_0
x_1 does not Granger cause x_4	3.014	172.925	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as two, one and four, two for $x_1 - x_4$ and $x_4 - x_1$ pairs, respectively.

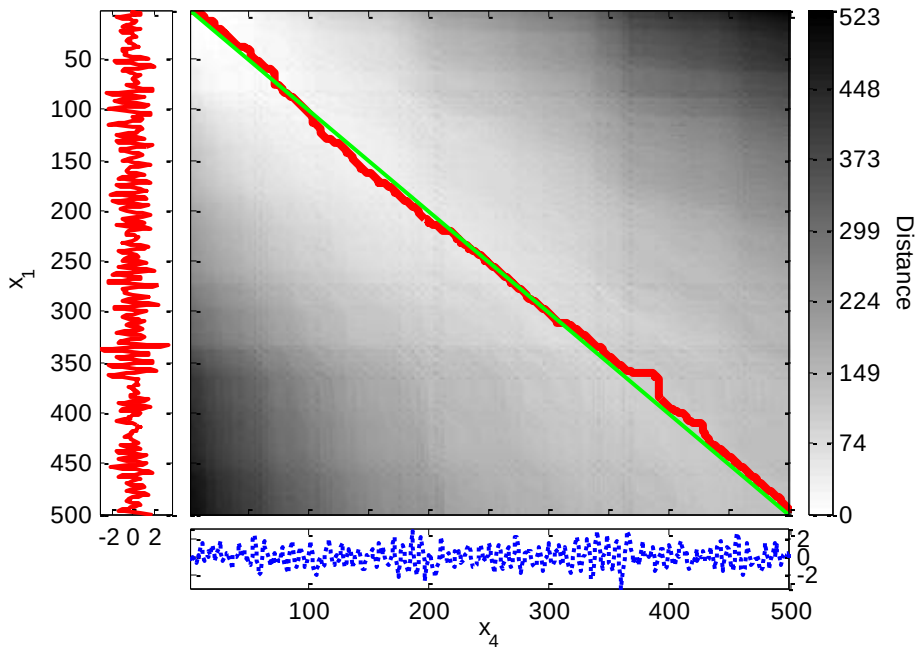

Figure 4.30. DTW distance matrix and the optimal warping path for the $x_1 - x_4$ pair of the five-dimensional example.

The mode of the time shifts between the series in the warping path is −6 and the mean of the time shifts is −2, indicating that the series x_1 leads the series x_4 ($x_1 \Rightarrow x_4$).

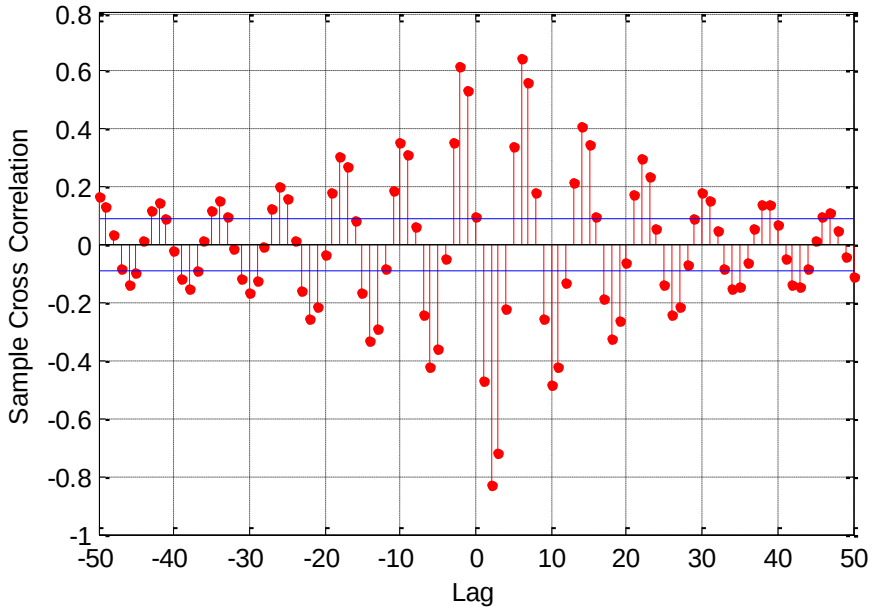

Figure 4.31. Sample cross-correlation function for the $x_1 - x_4$ pair of the five-dimensional example.

The maximum value of cross-correlation is observed at lag six, indicating that the series x_1 leads the series x_4 ($x_1 \Rightarrow x_4$).

The results of the tests for the pair $x_3 - x_4$ are as follows:

Table 4.18. GC test results for the $x_3 - x_4$ pair of the five-dimensional example.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_4 does not Granger cause x_3	3.860	102.604	Reject H_0
x_3 does not Granger cause x_4	3.014	8.274	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as five, one and four, two for $x_3 - x_4$ and $x_4 - x_3$ pairs, respectively.

The mode of the time shifts between the series in the warping path is one and the mean of the time shifts is two, indicating that the series x_3 lags behind the series x_4 ($x_3 \Leftarrow x_4$).

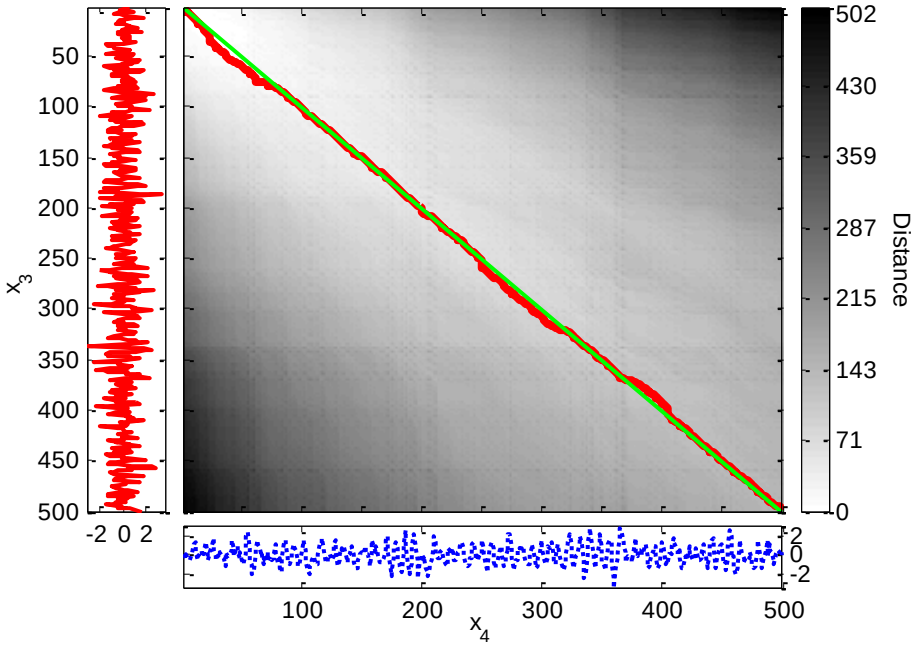

Figure 4.32. DTW distance matrix and the optimal warping path for the $x_3 - x_4$ pair of the five-dimensional example.

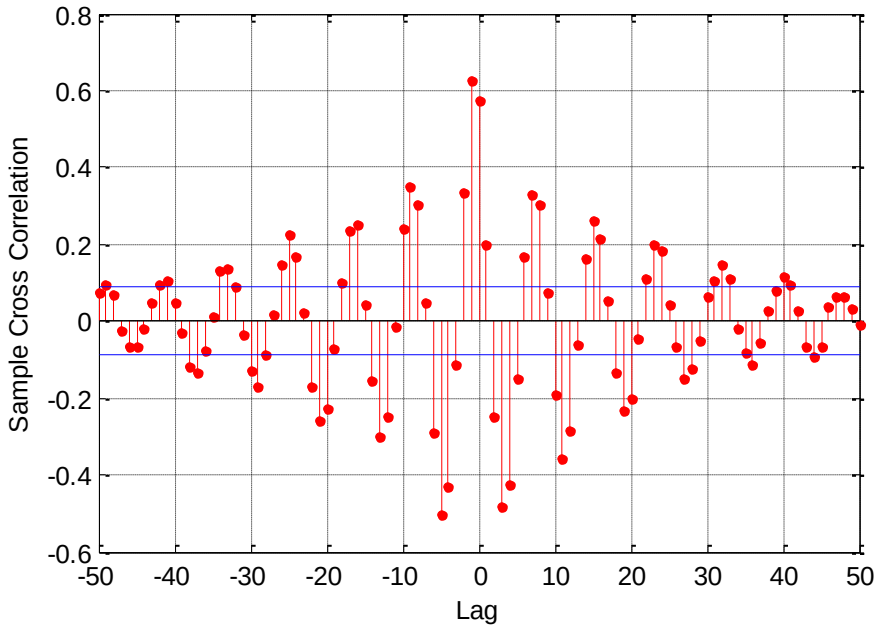

Figure 4.33. Sample cross-correlation function for the $x_3 - x_4$ pair of the five-dimensional example.

The maximum value of cross-correlation is observed at lag −2, indicating that the series x_3 lags behind the series x_4 ($x_3 \Leftarrow x_4$).

The results of the tests for the pair $x_4 – x_5$ are as follows:

Table 4.19. GC test results for the $x_4 – x_5$ pair of the five-dimensional example.

Null Hypothesis (H_0)	Critical Value	F-Statistics	Result
x_5 does not Granger cause x_4	3.860	28.360	Reject H_0
x_4 does not Granger cause x_5	3.860	165.818	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as four, one and four, one for $x_4 – x_5$ and $x_5 – x_4$ pairs, respectively.

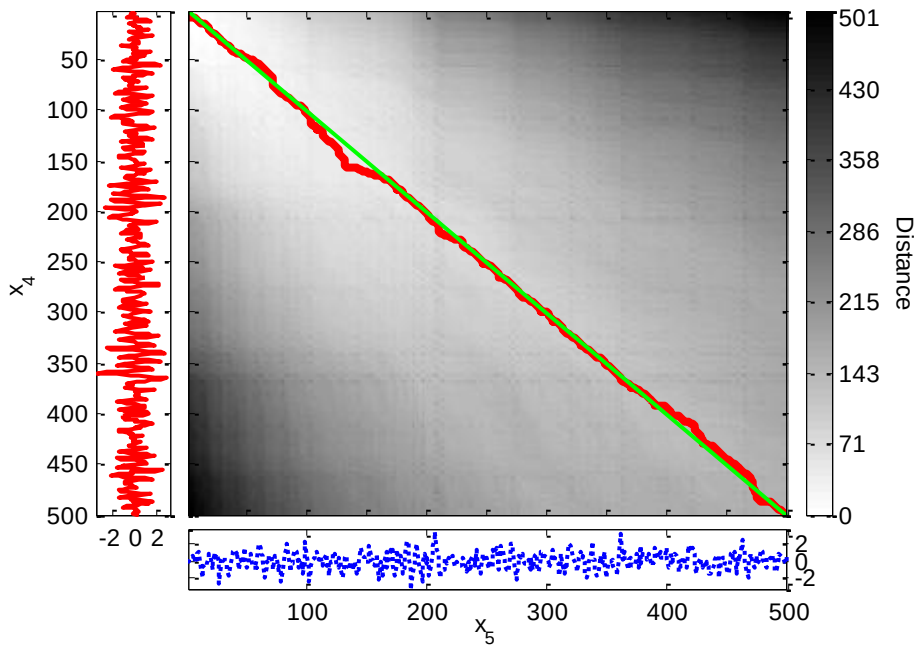

Figure 4.34. DTW distance matrix and the optimal warping path for the $x_4 – x_5$ pair of the five-dimensional example.

The mode and the mean of the time shifts between the series in the warping path are two, indicating that the series x_4 lags behind the series x_5 ($x_4 \Leftarrow x_5$).

The maximum value of cross-correlation is observed at lag −2, indicating that the series x_4 lags behind the series x_5 ($x_4 \Leftarrow x_5$).

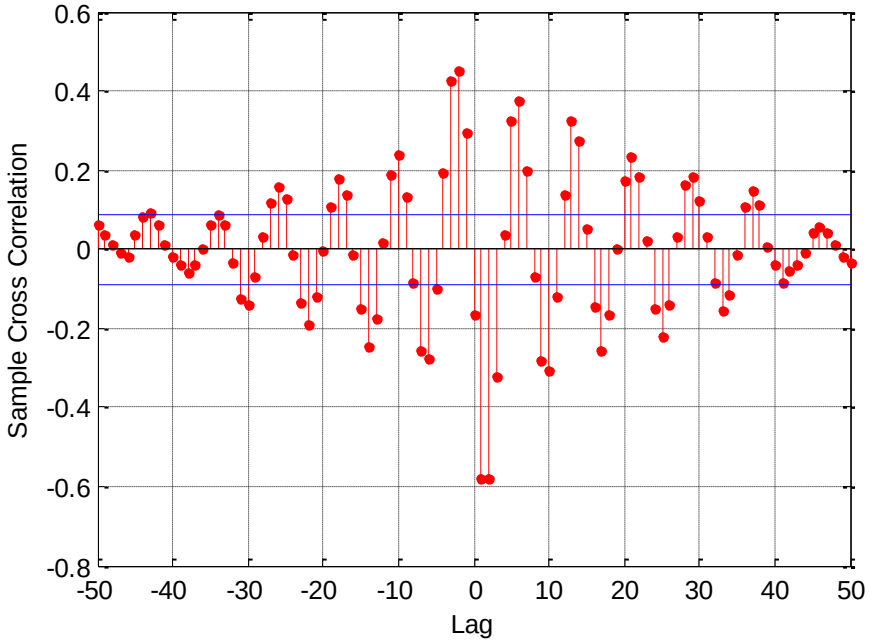

Figure 4.35. Sample cross-correlation function for the $x_4 - x_5$ pair of the five-dimensional example.

Table 4.20 summarizes the results of the GC test, DTW, and cross-correlation analyses, stated by Table 4.15 through Table 4.19 and Figure 4.26 through Figure 4.35.

Table 4.20. Summary of the results for the five-dimensional time series.

Pair	**GC**	**DTW**	**Cross-correlation**
x_1 & x_2	$x_1 \rightarrow x_2$	$x_1 \Rightarrow x_2$	$x_1 \Rightarrow x_2$
x_1 & x_3	$x_1 \rightarrow x_3$	$x_1 \Rightarrow x_3$	$x_1 \Rightarrow x_3$
x_1 & x_4	$x_1 \rightarrow x_4$	$x_1 \Rightarrow x_4$ (Figure 4.30: $x_1 \Leftrightarrow x_4$)	$x_1 \Rightarrow x_4$
x_3 & x_4	$x_3 \leftrightarrow x_4$	$x_3 \Leftarrow x_4$ (Figure 4.32: no precedence)	$x_3 \Leftarrow x_4$
x_4 & x_5	$x_4 \leftrightarrow x_5$	$x_4 \Leftarrow x_5$ (Figure 4.34: no precedence)	$x_4 \Leftarrow x_5$

The relations between the pairs $x_1 - x_2$, $x_1 - x_3$ and $x_1 - x_4$ are correctly acquired by the GC test and the DTW. The indirect relation between $x_3 - x_4$ ($x_3 \rightarrow x_4$, $x_3 \leftarrow x_4$) found by the GC test is due to the common cause x_1. DTW fails to reflect the feedback relation

between $x_4 - x_5$ by showing mainly no lead/lag in Figure 4.34 and $x_4 \Leftarrow x_5$ by examining the time shifts. Cross-correlation analysis results are identical with the DTW results.

4.2.2. Three-Dimensional Linear Time Series Model

In the second model there are three time series generated with the equations Equation 4.10 through Equation 4.12. Of these series x_2 directly causes x_3 and x_3 causes x_1 as depicted in Figure 4.36. ε_1, ε_2 and ε_3 are noises with zero mean and variances 0.3, one and 0.2 respectively. Data are tested pairwise with the GC test, DTW and cross-correlation analysis (Ding *et al*., 2006). Some additional parameters and properties of the data and tests are as follows: Each series is made of 100 data points. Z-score normalization is applied before DTW, cross-correlation calculations, and the GC test. Series are stationary with respect to ADF test and cointegrated. Significance level for the GC test is 0.95.

$$x_1(n)=0.8x_1(n-1) - 0.5x_1(n-2)+0.4x_3(n-1)+\varepsilon_1 \quad (4.10)$$

$$x_2(n)=0.9x_2(n-1) - 0.8x_2(n-2)+\varepsilon_2 \quad (4.11)$$

$$x_3(n)=0.5x_3(n-1) - 0.2x_3(n-2)+0.5x_2(n-1)+\varepsilon_3 \quad (4.12)$$

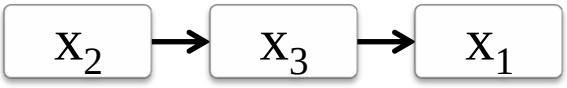

Figure 4.36. The schematic representation of the causal relations among the series in Equation 4.10 through Equation 4.12.

The results of the tests for the pair $x_1 - x_2$ are as follows:

Table 4.21. GC test results for the $x_1 - x_2$ pair of the three-dimensional example.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_2 does not Granger cause x_1	3.092	31.138	Reject H_0
x_1 does not Granger cause x_2	3.941	0.461	Do not reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as two, two and three, one for $x_1 - x_2$ and $x_2 - x_1$ pairs, respectively.

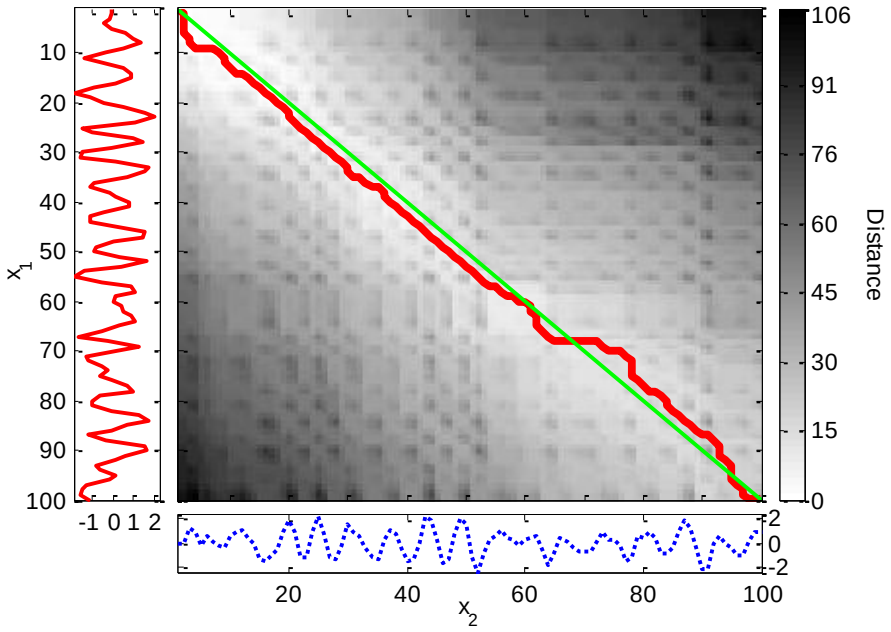

Figure 4.37. DTW distance matrix and the optimal warping path for the $x_1 - x_2$ pair of the three-dimensional example.

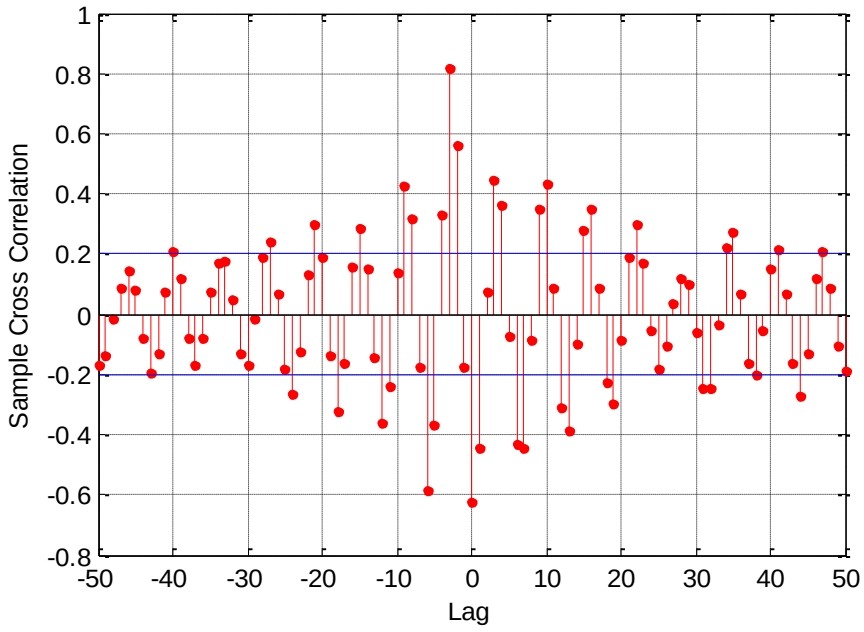

Figure 4.38. Sample cross-correlation function for the $x_1 - x_2$ pair of the three-dimensional example.

The mode of the time shifts between the series in the warping path is three and the mean of the time shifts is one, indicating that the series x_1 lags behind the series x_2 ($x_1 \Leftarrow x_2$).

The maximum value of cross-correlation is observed at lag −2, indicating that the series x_1 lags behind the series x_2 ($x_1 \Leftarrow x_2$).

The results of the tests for the pair $x_1 - x_3$ are as follows:

Table 4.22. GC test results for the $x_1 - x_3$ pair of the three-dimensional example.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_3 does not Granger cause x_1	3.940	99.579	Reject H_0
x_1 does not Granger cause x_3	3.942	0.620	Do not reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as two, one and four, one for $x_1 - x_3$ and $x_3 - x_1$ pairs, respectively.

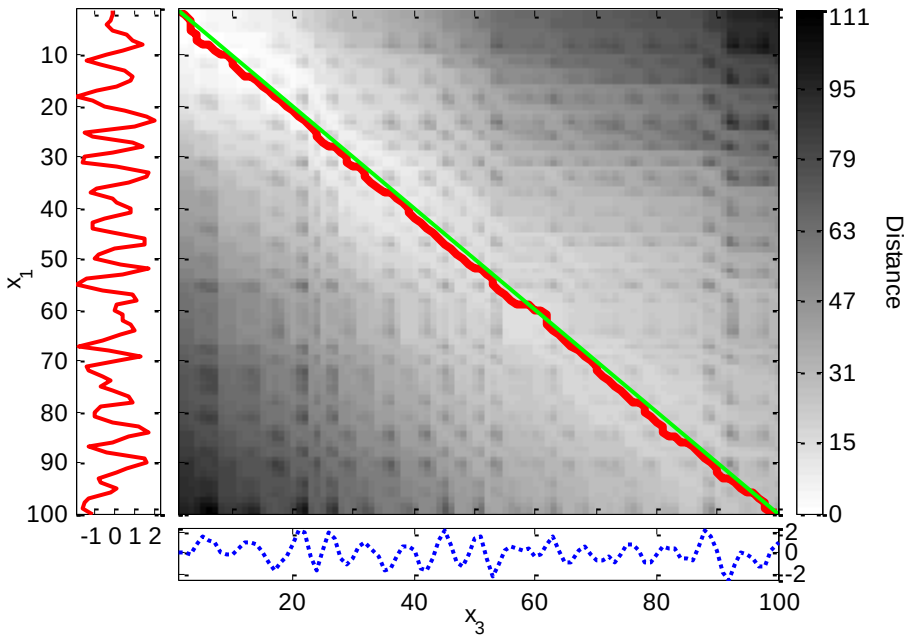

Figure 4.39. DTW distance matrix and the optimal warping path for the $x_1 - x_3$ pair of the three-dimensional example.

The mode and the mean of the time shifts between the series in the warping path are one, indicating that the series x_1 lags behind the series x_3 ($x_1 \Leftarrow x_3$).

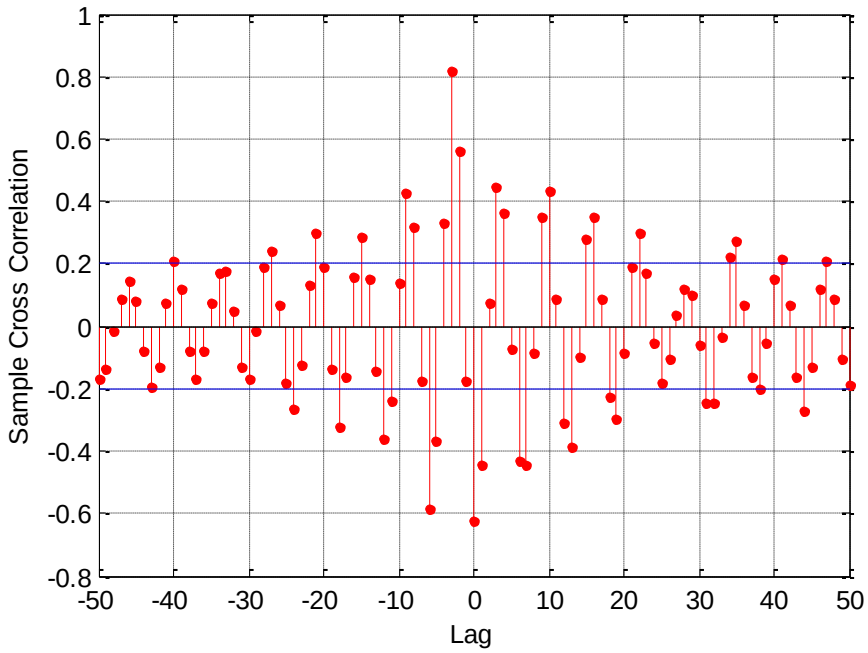

Figure 4.40. Sample cross-correlation function for the $x_1 - x_3$ pair of the three-dimensional example.

The maximum value of cross-correlation is observed at lag −2, indicating that the series x_1 lags behind the series x_3 ($x_1 \Leftarrow x_3$).

The results of the tests for the pair $x_2 - x_3$ are as follows:

Table 4.23. GC test results for the $x_2 - x_3$ pair of the three-dimensional example.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_3 does not Granger cause x_2	3.941	0.414	Do not reject H_0
x_2 does not Granger cause x_3	3.094	108.221	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as three, one and four, two for $x_2 - x_3$ and $x_3 - x_2$ pairs, respectively.

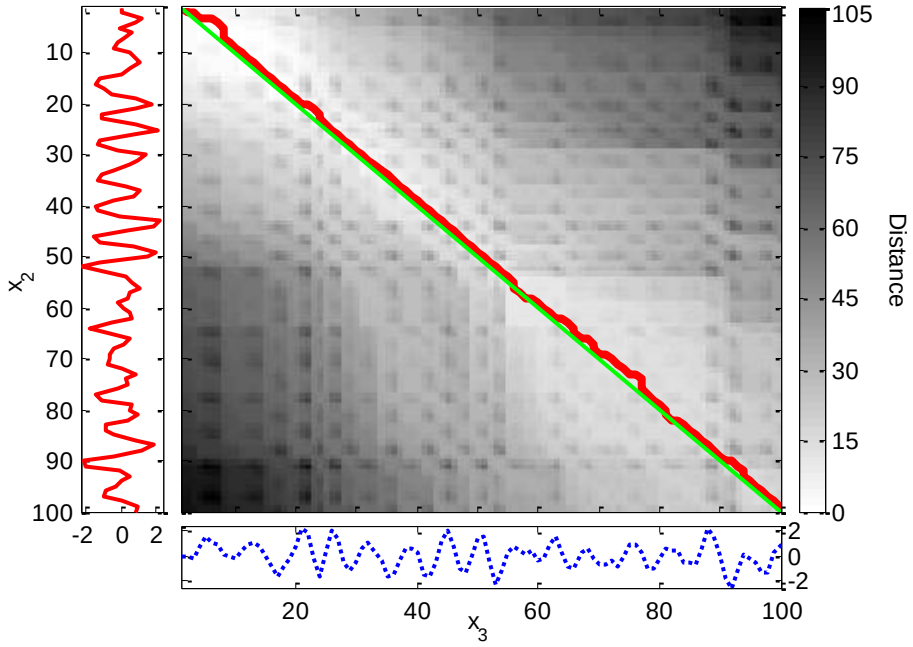

Figure 4.41. DTW distance matrix and the optimal warping path for the $x_2 - x_3$ pair of the three-dimensional example.

The mode and the mean of the time shifts between the series in the warping path are −1, indicating that the series x_2 leads the series x_3 ($x_2 \Rightarrow x_3$).

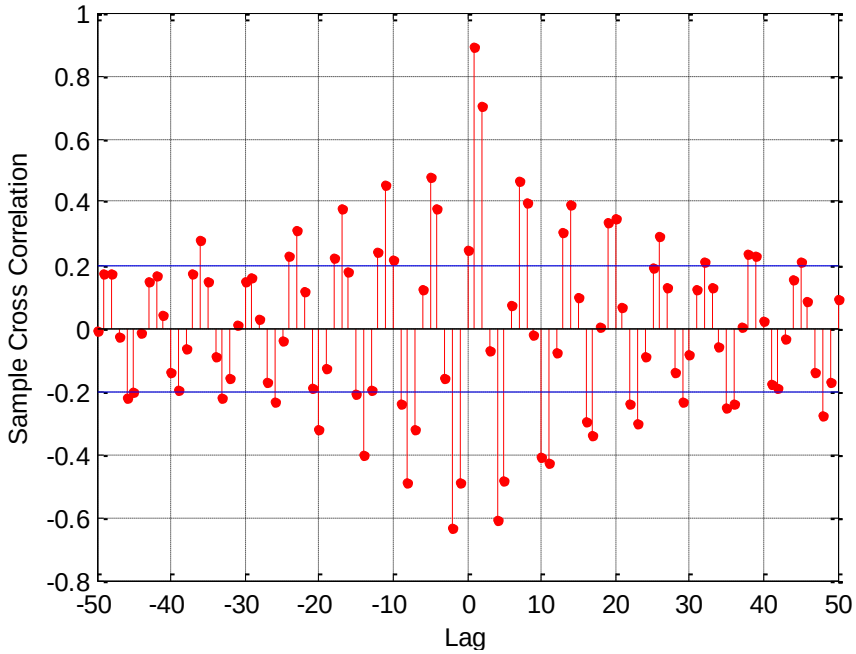

Figure 4.42. Sample cross-correlation function for the $x_2 - x_3$ pair of the three-dimensional example.

The maximum value of cross-correlation is observed at lag one, indicating that the series x_2 leads the series x_3 ($x_2 \Rightarrow x_3$).

Table 4.24 summarizes the results of the GC test, DTW, and cross-correlation analyses, stated by Table 4.21 through Table 4.23 and Figure 4.37 through Figure 4.42.

Table 4.24. Summary of the results for the three-dimensional time series.

Pair	GC	DTW	Cross-correlation
x_1 & x_2	$x_1 \leftarrow x_2$	$x_1 \Leftrightarrow x_2$	$x_1 \Leftarrow x_2$
x_1 & x_3	$x_1 \leftarrow x_3$	$x_1 \Leftarrow x_3$	$x_1 \Leftarrow x_3$
x_2 & x_3	$x_2 \rightarrow x_3$	$x_2 \Rightarrow x_3$	$x_2 \Rightarrow x_3$

The relations between $x_1 - x_3$ and $x_2 - x_3$ are captured correctly by the GC test and the DTW. Even though there is no direct relation between $x_1 - x_2$, the GC test and cross-correlation methods indicate a causal influence from x_2 to x_1, $x_1 \Leftarrow x_2$. DTW shows regional changes in the direction of causality, however it is mainly from x_2 to x_1 and the time shifts indicate this as well. This spurious relation between $x_1 - x_2$ is due to the influence of x_2 on x_3 ($x_2 \Rightarrow x_3$) and x_3's further influence on x_1 ($x_1 \Leftarrow x_3$).

4.3. Multidimensional Nonlinear Time Series Models

Two examples given in this section examine data sets generated from two different nonlinear series where the GC test fails to function correctly. The series are created by using MATLAB.

4.3.1. Two-Dimensional Nonlinear Time Series Model

In this example there are two time series created by using the nonlinear equations Equation 4.13 and Equation 4.14. ε_1 and ε_2 are noises with zero mean and variance one. c is the nonlinear coupling strength, and can take values between zero and one. In this example c is taken as 0.5. The direction of causality is from x_1 to x_2 as depicted in Figure 4.43 (Seth and Príncipe, 2010). Some other parameters and properties of the data and tests are as follows: Each series is made of 500 data points. Z-score normalization is applied

before DTW, cross-correlation calculations, and the GC test. Series are stationary with respect to ADF test and cointegrated. Significance level for the GC test is 0.95.

$$x_1(n)=3.4x_1(n-1)\left(1-x_1^2(n-1)\right)\exp\left(-x_1^2(n-1)\right) + 0.8x_1(n-2)+\varepsilon_1 \quad (4.13)$$

$$x_2(n)=3.4x_2(n-1)\left(1-x_2^2(n-1)\right)\exp\left(-x_2^2(n-1)\right) + 0.5x_2(n-2)+cx_1^2(n-2)+\varepsilon_2 \quad (4.14)$$

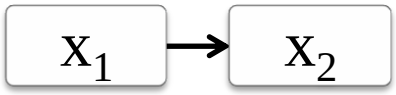

Figure 4.43. The schematic representation of the causal relations among the two-dimensional nonlinear time series in Equation 4.13 and Equation 4.14.

The results of the tests are as follows:

Table 4.25. GC test results for the two-dimensional nonlinear series.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_2 does not Granger cause x_1	3.860	3.160	Do not reject H_0
x_1 does not Granger cause x_2	3.014	9.862	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as two, one and four, two for $x_1 - x_2$ and $x_2 - x_1$ pairs, respectively.

The mode of the time shifts between the series in the warping path is 14 and the mean of the time shifts is 28, indicating that the series x_1 lags behind the series x_2 ($x_1 \Leftarrow x_2$).

The maximum value of cross-correlation is observed at lag 30, indicating that the series x_1 leads the series x_2 ($x_1 \Rightarrow x_2$).

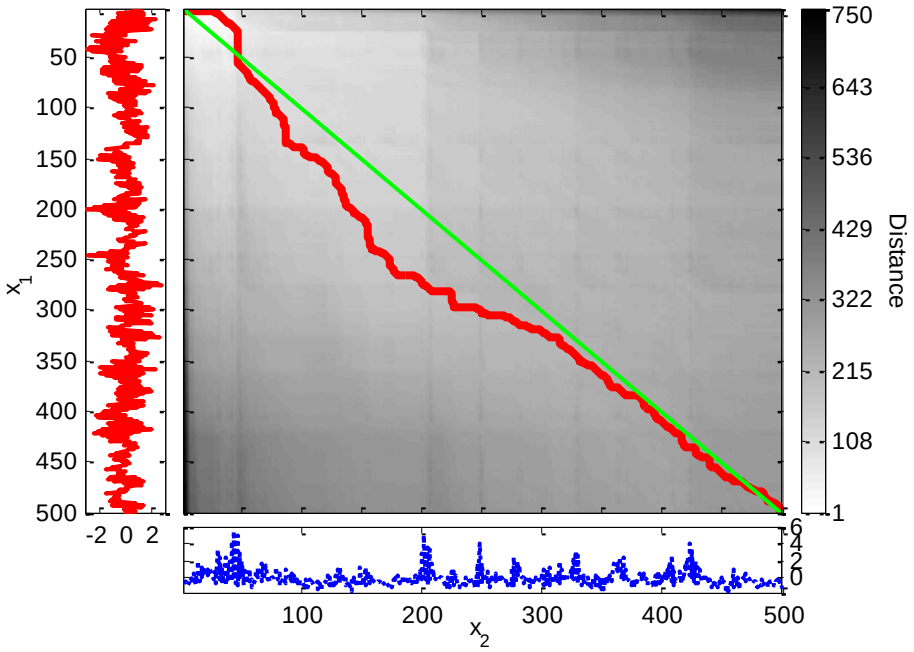

Figure 4.44. DTW distance matrix and the optimal warping path for the two-dimensional nonlinear series.

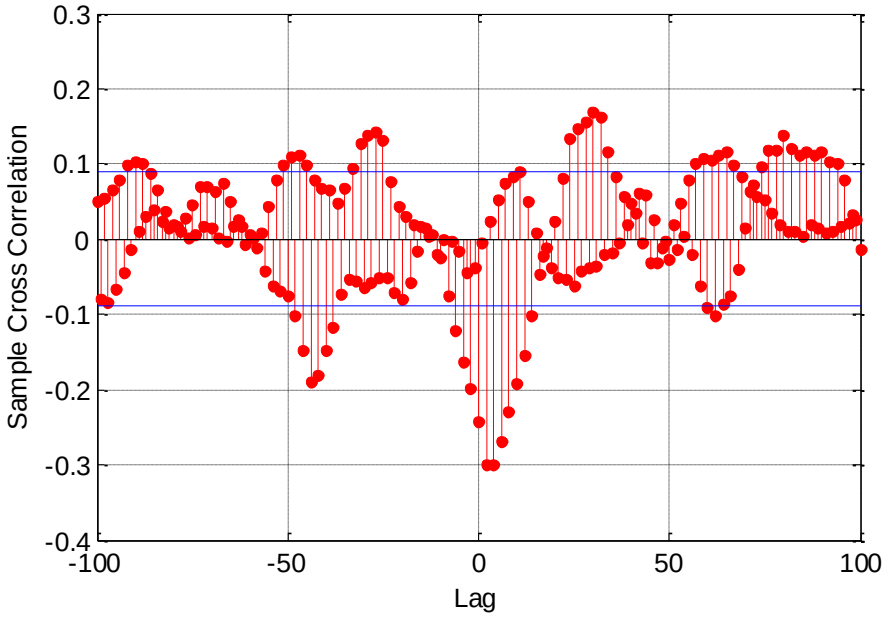

Figure 4.45. Sample cross-correlation function for the two-dimensional nonlinear series.

Table 4.26 summarizes the results of the GC test, DTW, and cross-correlation analyses, stated by Table 4.25, Figure 4.44 and Figure 4.45.

Table 4.26. Summary of the results for the two-dimensional nonlinear series.

Pairs	GC	DTW	Cross-correlation
x_1 & x_2	$x_1 \rightarrow x_2$	$x_1 \Leftarrow x_2$	$x_1 \Rightarrow x_2$

Examining the DTW results, it is seen that the optimal warping path in Figure 4.44 is essentially below the diagonal line except the beginning and end parts, and the mode and mean of the time shifts are positive meaning that x_2 precedes x_1, $x_1 \Leftarrow x_2$. So, in this example the DTW gives incorrect results. Even though the GC test is a linear method and not expected to work correctly in nonlinear problems, in this particular case, correct result is obtained as it is found that $x_1 \rightarrow x_2$ and not the other way around. Cross-correlation analysis results are consistent with the GC test results.

4.3.2. Four-Dimensional Nonlinear Time Series Model

Second example consists of four series created by using the nonlinear equations, Equation 4.15 through Equation 4.17, influencing one another in a complex way. Time series x_1 and x_2 have a feedback relation, and x_2 is also caused by x_4. x_3 does not cause any of the series however x_1 and x_2 have an influence on x_3. x_4 is Gaussian white noise of variance one. The causal relations among the series are depicted in Figure 4.46 (Seth and Príncipe, 2010). Some other parameters and properties of the data and tests are as follows: Each series is made of 500 data points. Z-score normalization is applied before DTW, cross-correlation calculations, and the GC test. Series are stationary with respect to ADF test and cointegrated. Significance level for the GC test is 0.95.

$$x_1(n+1) = \left(1 + \frac{x_1(n)}{1+x_1^2(n)}\right) \sin x_2(n) \tag{4.15}$$

$$x_2(n+1) = x_1(n) \exp\left(-\frac{x_1^2(n)+x_2^2(n)}{8}\right) + x_2(n) \cos x_2(n) + \frac{x_4^3(n)}{\left(1+x_4(n)\right)^2 + 0.5\cos\left(x_1(n)+x_2(n)\right)} \tag{4.16}$$

$$x_3(n+1) = \frac{x_1(n)}{1+0.5\sin x_2(n)} + \frac{x_2(n)}{1+0.5\sin x_1(n)} \tag{4.17}$$

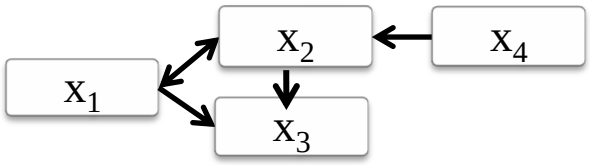

Figure 4.46. The schematic representation of the causal relations among the four-dimensional nonlinear time series in Equation 4.15 through Equation 4.17.

The results of the tests for the pair $x_1 - x_2$ are as follows:

Table 4.27. GC test results for the $x_1 - x_2$ pair of four-dimensional nonlinear time series.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_2 does not Granger cause x_1	2.623	195.210	Reject H_0
x_1 does not Granger cause x_2	2.623	18.240	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as three, three and three, three for $x_1 - x_2$ and $x_2 - x_1$ pairs, respectively.

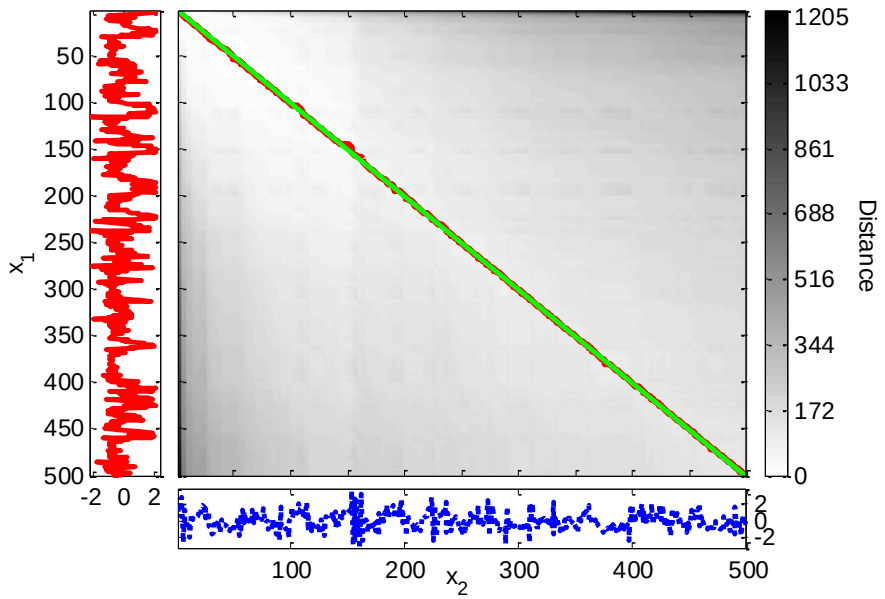

Figure 4.47. DTW distance matrix and the optimal warping path for the $x_1 - x_2$ pair of four-dimensional nonlinear time series.

The mode and the mean of the time shifts between the series in the warping path are one, indicating that the series x_1 lags behind the series x_2 ($x_1 \Leftarrow x_2$).

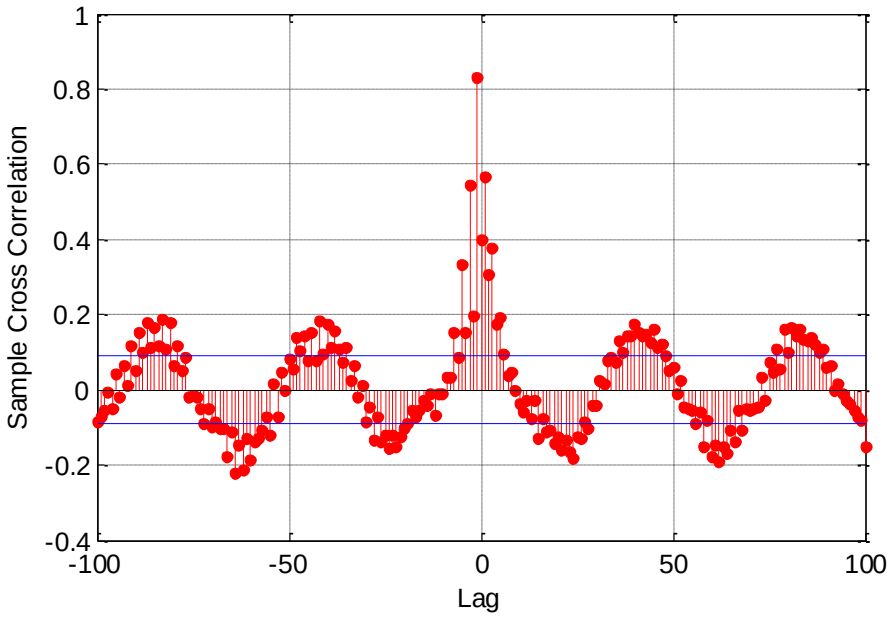

Figure 4.48. Sample cross-correlation function for the $x_1 - x_2$ pair of four-dimensional nonlinear time series.

The maximum value of cross-correlation is observed at lag −1, indicating that the series x_1 lags behind the series x_2 ($x_1 \Leftarrow x_2$).

The results of the tests for the pair $x_1 - x_3$ are as follows:

Table 4.28. GC test results for the $x_1 - x_3$ pair of four-dimensional nonlinear time series.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_3 does not Granger cause x_1	3.014	33.178	Reject H_0
x_1 does not Granger cause x_3	3.014	109.028	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as three, two and two, two for $x_1 - x_3$ and $x_3 - x_1$ pairs, respectively.

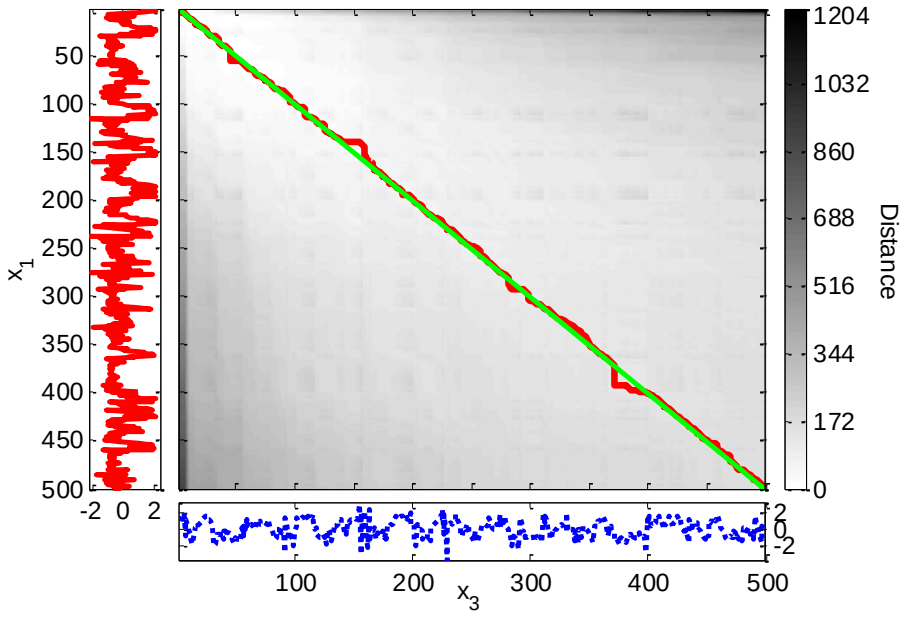

Figure 4.49. DTW distance matrix and the optimal warping path for the $x_1 - x_3$ pair of four-dimensional nonlinear time series.

The mode and the mean of the time shifts between the series in the warping path are zero, indicating that there is no lead/lag relation between the series x_1 and the series x_3.

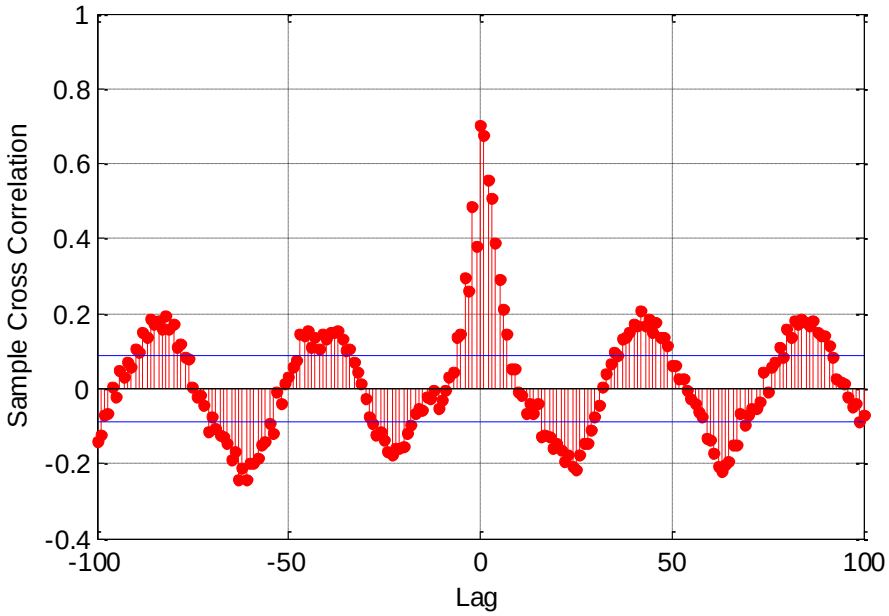

Figure 4.50. Sample cross-correlation function for the $x_1 - x_3$ pair of four-dimensional nonlinear time series.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relation between the series x_1 and the series x_3.

The results of the tests for the pair $x_2 - x_3$ are as follows:

Table 4.29. GC test results for the $x_2 - x_3$ pair of four-dimensional nonlinear time series.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_3 does not Granger cause x_2	2.623	6.632	Reject H_0
x_2 does not Granger cause x_3	2.390	690.468	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as three, three and two, four for $x_2 - x_3$ and $x_3 - x_2$ pairs, respectively.

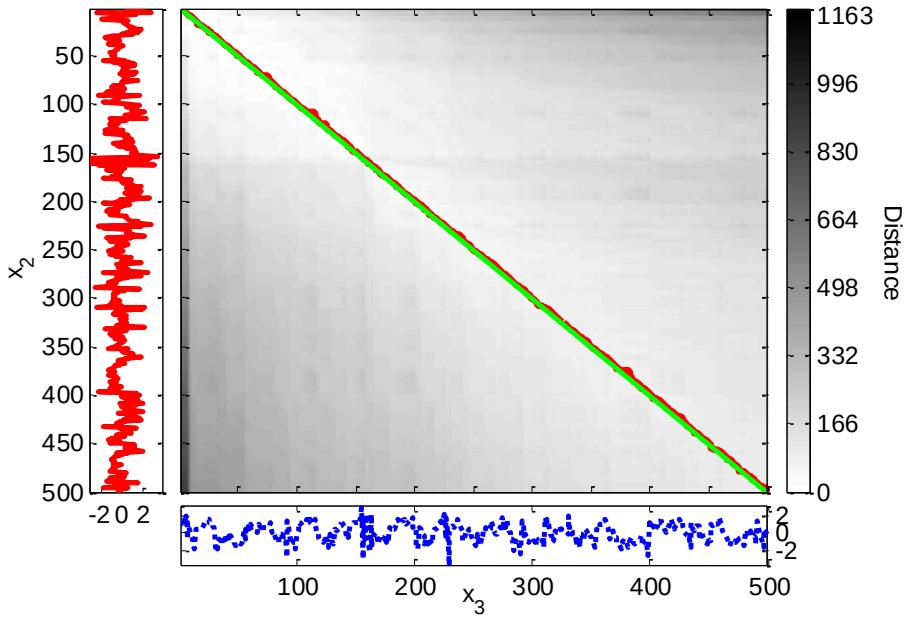

Figure 4.51. DTW distance matrix and the optimal warping path for the $x_2 - x_3$ pair of four-dimensional nonlinear time series.

The mode and the mean of the time shifts between the series in the warping path are −1, indicating that the series x_2 leads the series x_3 ($x_2 \Rightarrow x_3$).

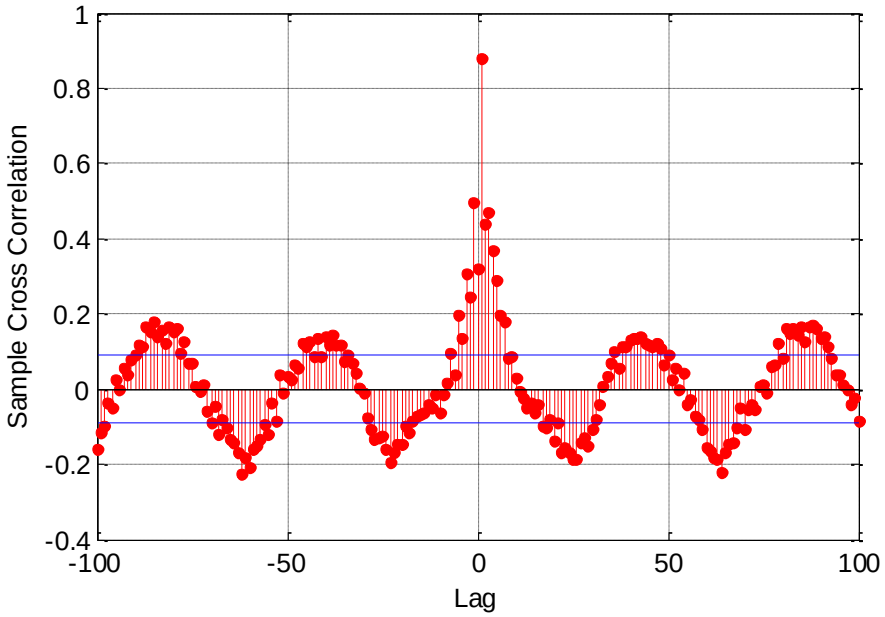

Figure 4.52. Sample cross-correlation function for the $x_2 - x_3$ pair of four-dimensional nonlinear time series.

The maximum value of cross-correlation is observed at lag one, indicating that the series x_2 leads the series x_3 ($x_2 \Rightarrow x_3$).

The results of the tests for the pair $x_2 - x_4$ are as follows:

Table 4.30. GC test results for the $x_2 - x_4$ pair of four-dimensional nonlinear time series.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
x_4 does not Granger cause x_2	2.623	146.107	Reject H_0
x_2 does not Granger cause x_4	3.860	4.152	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as three, three and one, one for $x_2 - x_4$ and $x_4 - x_2$ pairs, respectively.

The mode of the time shifts between the series in the warping path is one indicating that the series x_2 lags behind the series x_4 ($x_2 \Leftarrow x_4$), and the mean of the time shifts is −6, indicating that the series x_2 leads the series x_4 ($x_2 \Rightarrow x_4$).

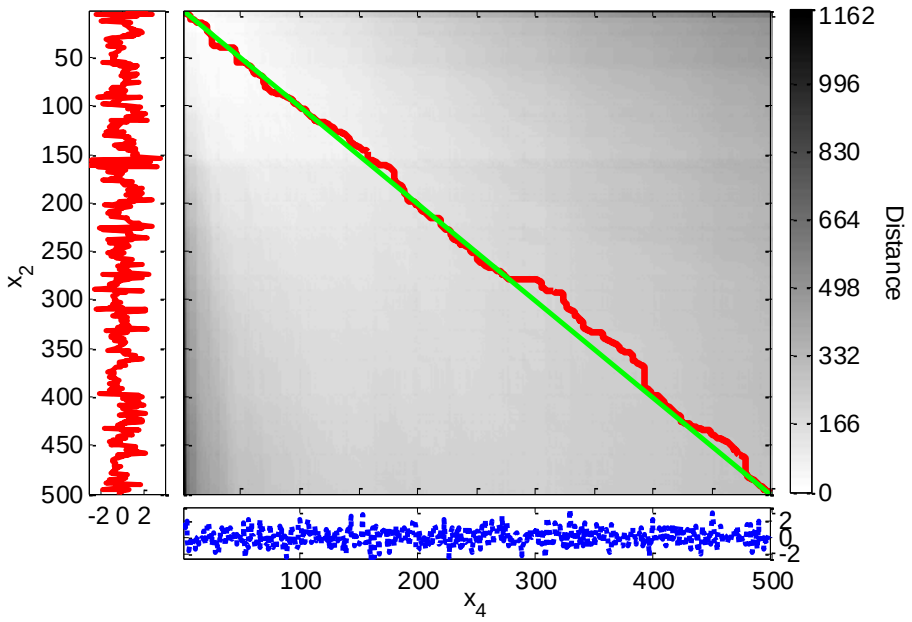

Figure 4.53. DTW distance matrix and the optimal warping path for the $x_2 - x_4$ pair of four-dimensional nonlinear time series.

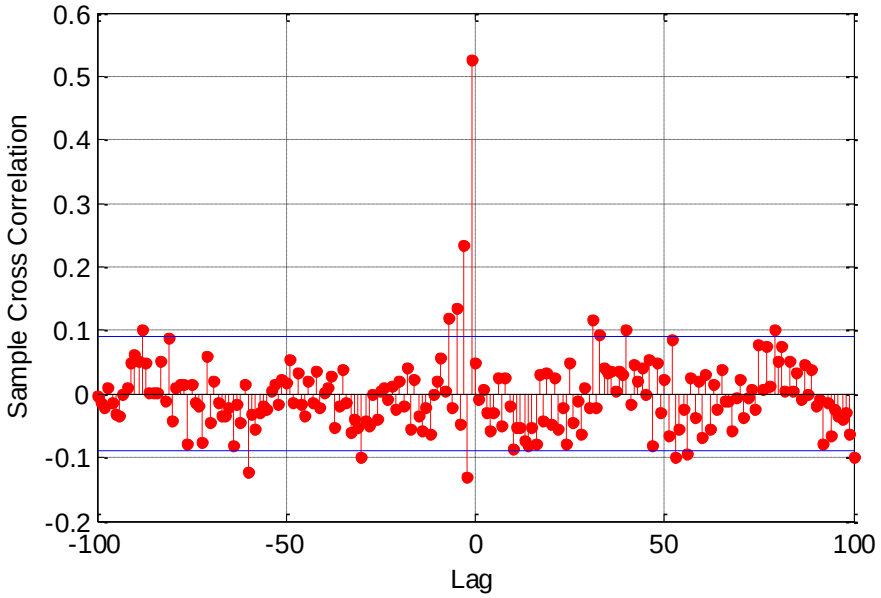

Figure 4.54. Sample cross-correlation function for the $x_2 - x_4$ pair of four-dimensional nonlinear time series.

The maximum value of cross-correlation is observed at lag −1, that the series x_2 lags behind the series x_4 ($x_2 \Leftarrow x_4$).

Table 4.31 summarizes the results of the GC test, DTW, and cross-correlation analyses, stated by Table 4.27 through Table 4.30 and Figure 4.47 through Figure 4.54.

Table 4.31. Summary of the results for the four-dimensional nonlinear time series.

Pair	**GC**	**DTW**	**Cross-correlation**
x_1 & x_2	$x_1 \leftrightarrow x_2$	$x_1 \Leftarrow x_2$	$x_1 \Leftarrow x_2$
x_1 & x_3	$x_1 \leftrightarrow x_3$	No precedence	No precedence
x_2 & x_3	$x_2 \leftrightarrow x_3$	$x_2 \Rightarrow x_3$	$x_2 \Rightarrow x_3$
x_2 & x_4	$x_2 \leftrightarrow x_4$	$x_2 \Leftrightarrow x_4$	$x_2 \Leftarrow x_4$

The GC test incorrectly indicates that x_3 causes x_1 and x_2 ($x_1 \leftarrow x_3$, $x_2 \leftarrow x_3$), and x_2 causes x_4 ($x_2 \rightarrow x_4$). x_3 causing x_1 can be explained by the fact that x_2 causes both x_1 and x_3, similarly x_1 is a common cause of both x_2 and x_3 so the GC test erroneously finds that x_3 causes x_2. DTW captures only the relation between x_2 and x_3 correctly, failing to uncover the expected causal relation among all series. However, overall it produces similar results with cross-correlation analysis.

4.4. Hair Dryer Input/Output Model

The data set was acquired from a MATLAB demo for nonlinear system identification (The MathWorks, Inc., 2007). It was originally obtained from a laboratory device resembling a hair dryer. In this device air is heated at the inlet of a tube and its temperature is measured at the outlet. Voltage over the resistor wires is the input, u(k) and the air temperature at the outlet is the output, y(k) as seen from the model in Figure 4.55. The data set can be accessed from the System Identification Toolbox of MATLAB. The expected result is that the input (u) leads the output (y).

Figure 4.55. System model for the hair dryer device.

Some parameters and properties of the data and tests are as follows: Each series is made of 1000 data points. Z-score normalization is applied before DTW, cross-correlation calculations, and the GC test. Series are stationary with respect to ADF test and cointegrated. Significance level for the GC test is 0.95.

The results of the tests are as follows:

Table 4.32. GC test results for the hair dryer data.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
y does not Granger cause u	3.851	5.053	Reject H_0
u does not Granger cause y	2.223	1194.081	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as one, one and three, five for u – y and y – u pairs, respectively.

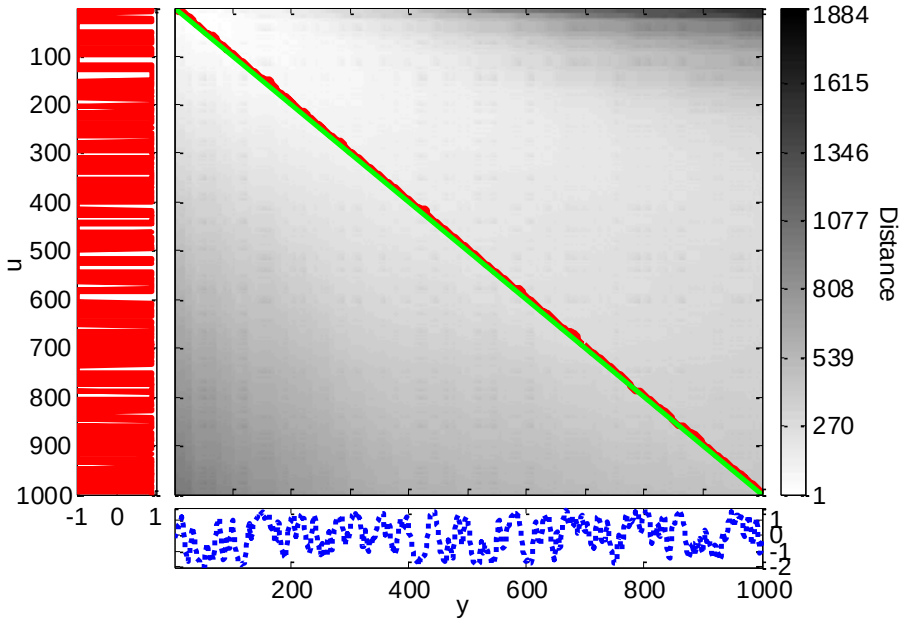

Figure 4.56. DTW distance matrix and the optimal warping path for the hair dryer data.

The mode of the time shifts between the series in the warping path is −5 and the mean of the time shifts is −6, indicating that the series u leads the series y (u ⇒ y).

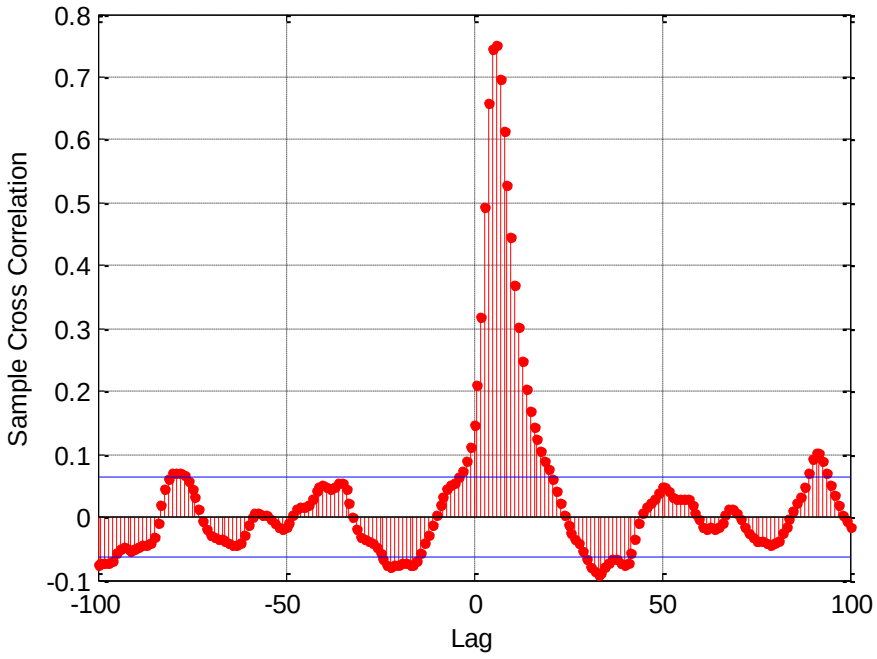

Figure 4.57. Sample cross-correlation function for the hair dryer data.

The maximum value of cross-correlation is observed at lag six, indicating that the series u leads the series y (u ⇒ y).

Table 4.33 summarizes the results of the GC test, DTW and cross-correlation analyses, shown by Table 4.32, Figure 4.56 and Figure 4.57.

Table 4.33. Summary of the results for the hair dryer data.

Pair	**GC**	**DTW**	**Cross-correlation**
u & y	u ↔ y	u ⇒ y	u ⇒ y

The GC test shows a feedback relation between the input (u) and output (y); although in the first test, the value of the F-statistic is very close to the critical value whereas in the second test it is greater than the critical value by a very large amount and causality from the output to the input (y → u), can be considered weak. On the other hand, DTW and cross-correlation results are in accordance with the expected results and indicate that the input leads the output; u ⇒ y.

4.5. Economic Indexes and the Chemical Engineering Plant Cost Index

In this section, the causal relations among various US macroeconomic indicators and chemical engineering related cost index are investigated. For this purpose annual data on Chemical Engineering Plant Cost Index, Industrial Production Index, Spot Oil Prices and Gross Domestic Product between the years 1950 and 2010 are employed.

Plant cost indexes use the construction costs for chemical plants as the basis and reflect the effect of time on the costs. They are utilized in the estimation of process-equipment costs and chemical plant investments (Baasel, 1990). The indexes are based on averages and their accuracy is about 10% for periods up to five years. Chemical Engineering Plant Cost Index (CEPCI) is commonly used and published in the *Chemical Engineering Magazine* and years 1957-1959 are taken as the reference period (Loh *et al*., 2002; Coker, 2007). The CEPCI was designed mainly for the US chemical industry and uses US macroeconomic indicators in its construction and updates (Vatavuk, 2002).

Industrial Production Index (IP) evaluates the amount of output from various industries as well as the manufacturing industry in the USA. 2007 is the reference year where the index is taken as 100. IP is a major economic indicator showing how well the industries are performing. An increase in IP causes an increase in the national currency rate and Gross National Product. Data of the index is made public monthly by the Federal Reserve Board of St. Louis (Industrial Production Index – IPI).

The West Texas Intermediate oil price index (OP) is an index for partially refined oil prices from the US. Data is provided by *the Federal Reserve Bank of St. Louis* and the values are based on *the Spot Oil Price: West Texas Intermediate* in *Wall Street Journal*. (Bottazzi, 2012).

Gross Domestic Product (GDP) is a measure of a country's standard of living and economic wellbeing. It consists of all private consumption in a nation's economy, the sum of government spending, total net exports and the business spending on capital in a country for a specified time period which is usually one year (Gross Domestic Product – GDP).

Some parameters and properties of the data and tests are as follows: All series are made of 61 data points. Z-score normalization is applied before DTW, cross-correlation calculations and the GC test. Series are non-stationary and not cointegrated. First order difference is taken for the CEPCI – IP, CEPCI – OP and IP – OP data sets and second order difference is taken for the IP – GDP, OP – GDP, and CEPCI – GDP data sets before the GC test to make the series stationary with respect to ADF and KPSS tests. Significance level for the GC test is 0.95.

The results of the tests for the CEPCI – IP data sets are as follows:

Table 4.34. GC test results for the CEPCI and IP data.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
IP does not Granger cause CEPCI	4.023	1.265	Do not reject H_0
CEPCI does not Granger cause IP	2.546	9.936	Reject H_0

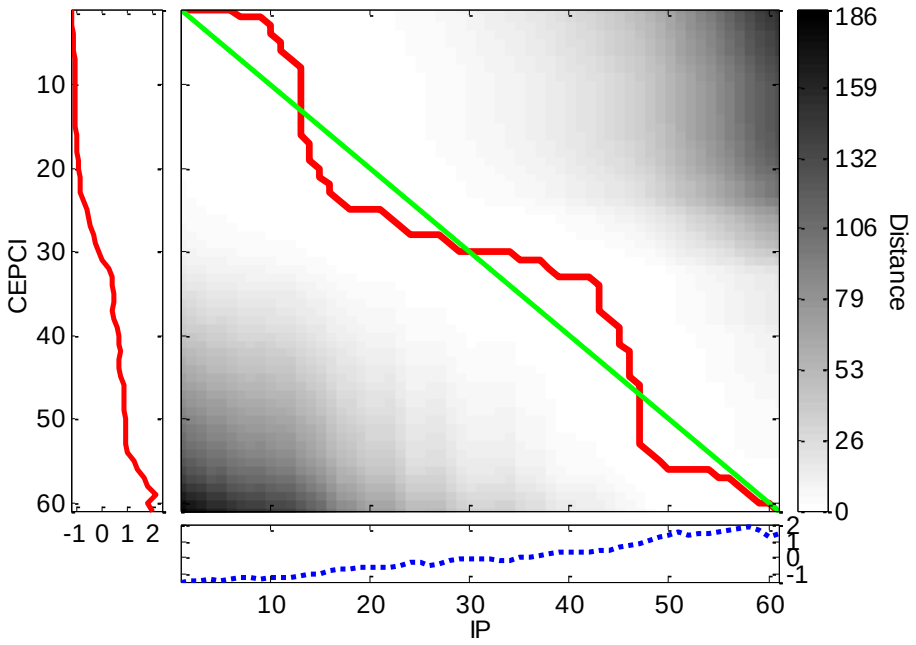

Figure 4.58. DTW distance matrix and the optimal warping path for the CEPCI and IP data.

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be

used in the GC tests as five, one and two, four for CEPCI – IP and IP – CEPCI pairs, respectively.

The mode of the time shifts between the series in the warping path is −6 and the mean of the time shifts is −1, indicating that CEPCI precedes IP.

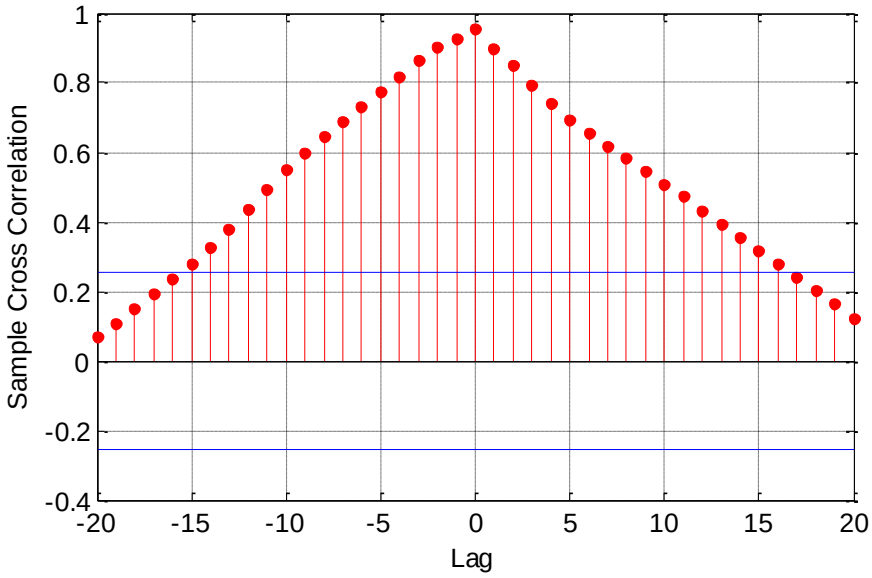

Figure 4.59. Sample cross-correlation function for the CEPCI and IP data.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relation between CEPCI and IP.

The results of the tests for the CEPCI – OP data sets are as follows:

Table 4.35. GC test results for the CEPCI and OP data.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
OP does not Granger cause CEPCI	4.023	1.728	Do not reject H_0
CEPCI does not Granger cause OP	2.389	5.292	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as five, one and one, five for CEPCI – OP and OP – CEPCI pairs, respectively.

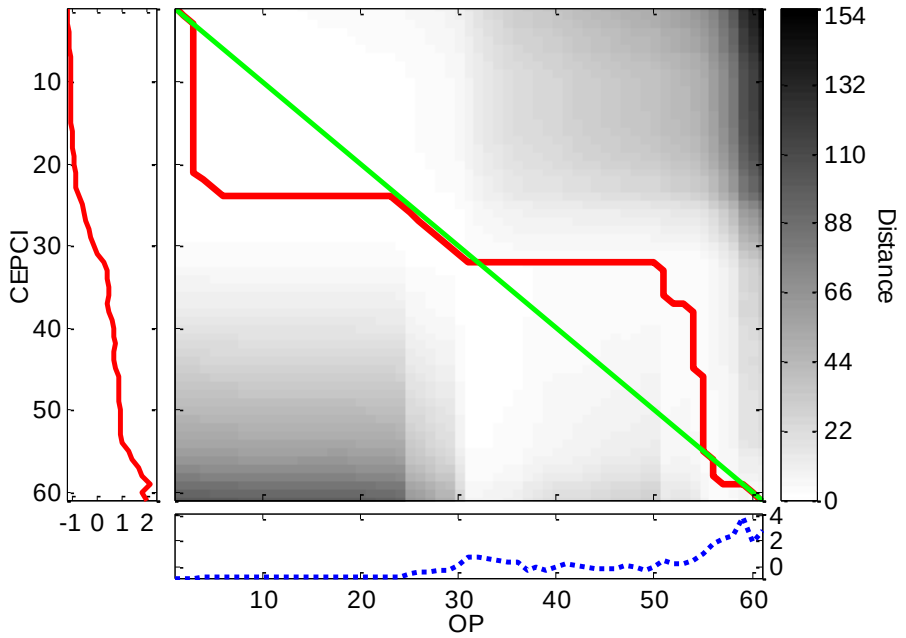

Figure 4.60. DTW distance matrix and the optimal warping path for the CEPCI and OP data.

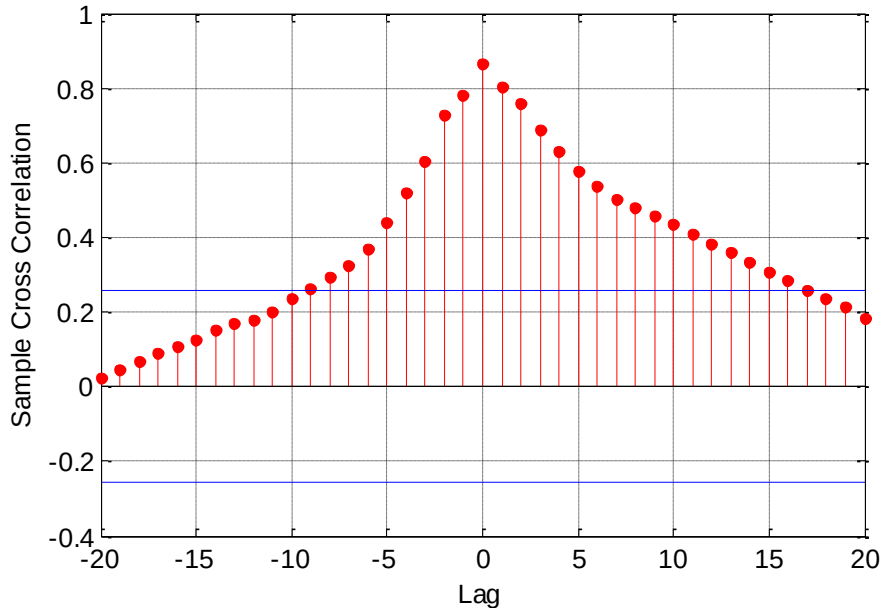

Figure 4.61. Sample cross-correlation function for the CEPCI and OP data.

The mode of the time shifts between the series in the warping path is one, indicating that CEPCI lags behind OP, and the mean of the time shifts is zero, indicating that there is no lead/lag relation between the series.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relation between CEPCI and OP.

The results of the tests for the IP – OP data sets are as follows:

Table 4.36. GC test results for the IP and OP data.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
OP does not Granger cause IP	4.013	26.501	Reject H_0
IP does not Granger cause OP	4.010	2.845	Do not reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as two, one and one, for IP – OP and OP – IP pairs, respectively.

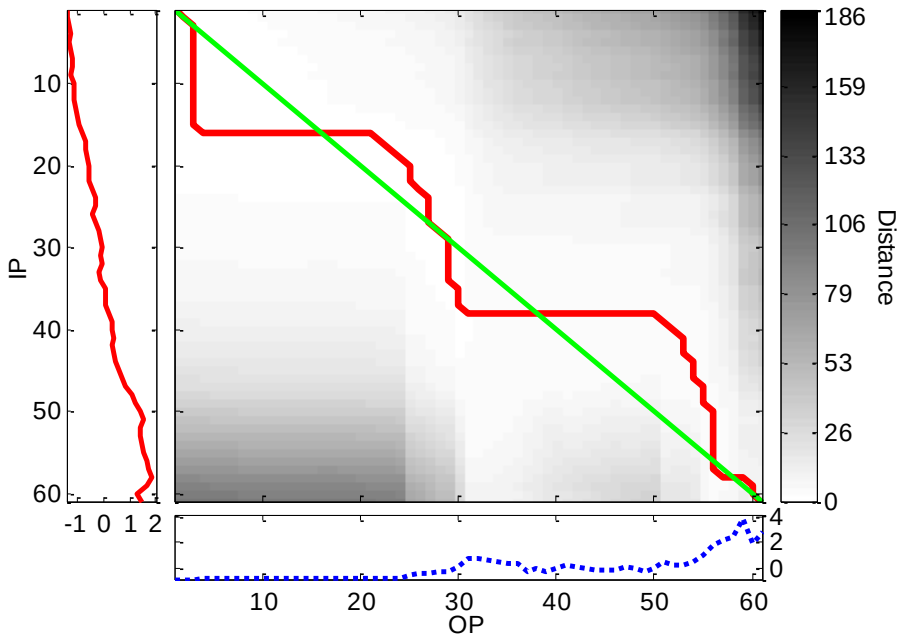

Figure 4.62. DTW distance matrix and the optimal warping path for the IP and OP data.

The mode and the mean of the time shifts between the series in the warping path are zero, indicating that there is no lead/lag relation between IP and OP.

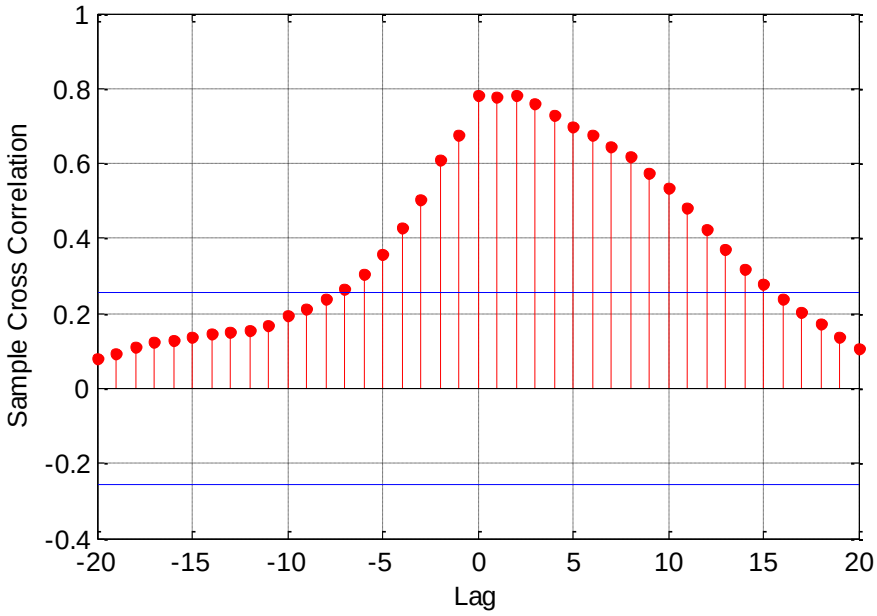

Figure 4.63. Sample cross-correlation function for the IP and OP data.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relation between IP and OP.

The results of the tests for the IP – GDP data sets are as follows:

Table 4.37. GC test results for the IP and GDP data.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
GDP does not Granger cause IP	2.404	5.037	Reject H_0
IP does not Granger cause GDP	4.027	0.283	Do not reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as four, five and five, one for IP – GDP and GDP – IP pairs, respectively.

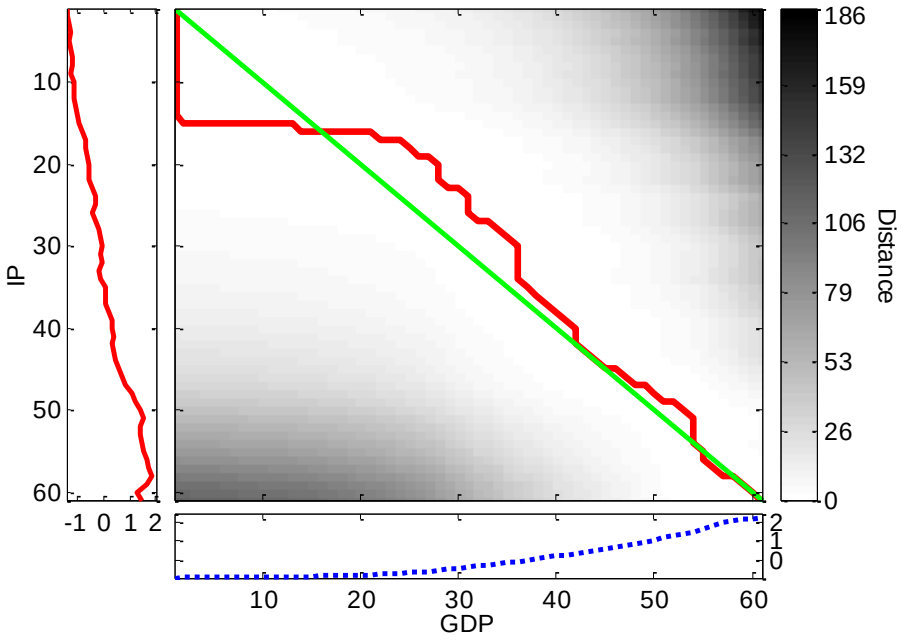

Figure 4.64. DTW distance matrix and the optimal warping path for the IP and GDP data.

The mode of the time shifts between the series in the warping path is −2, indicating that IP precedes GDP, and the mean of the time shifts is zero, indicating that there is no lead/lag relation between the series.

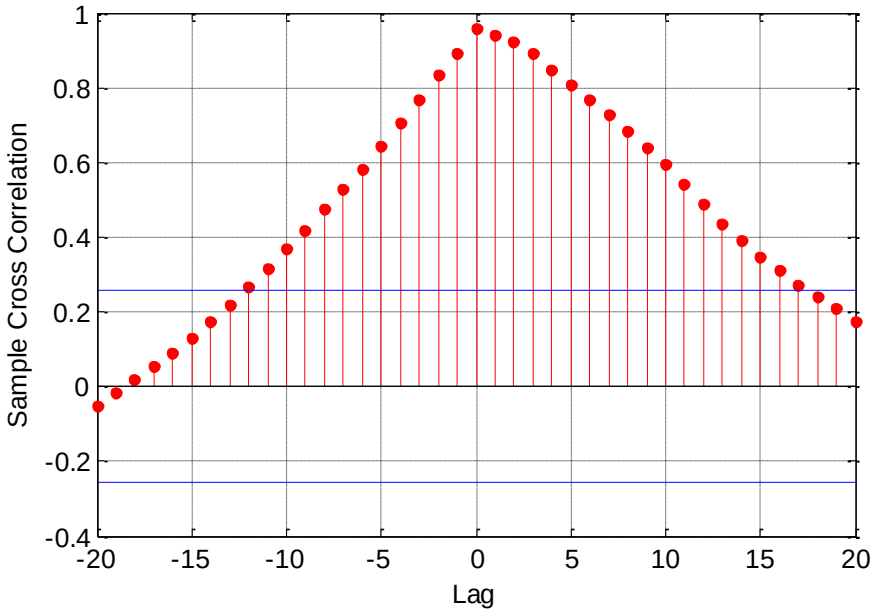

Figure 4.65. Sample cross-correlation function for the IP and GDP data.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relation between IP and GDP.

The results of the tests for the OP – GDP data sets are as follows:

Table 4.38. GC test results for the OP and GDP data.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
GDP does not Granger cause OP	4.016	24.541	Reject H_0
OP does not Granger cause GDP	4.027	12.550	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as two, one and five, one for OP – GDP and GDP – OP pairs, respectively.

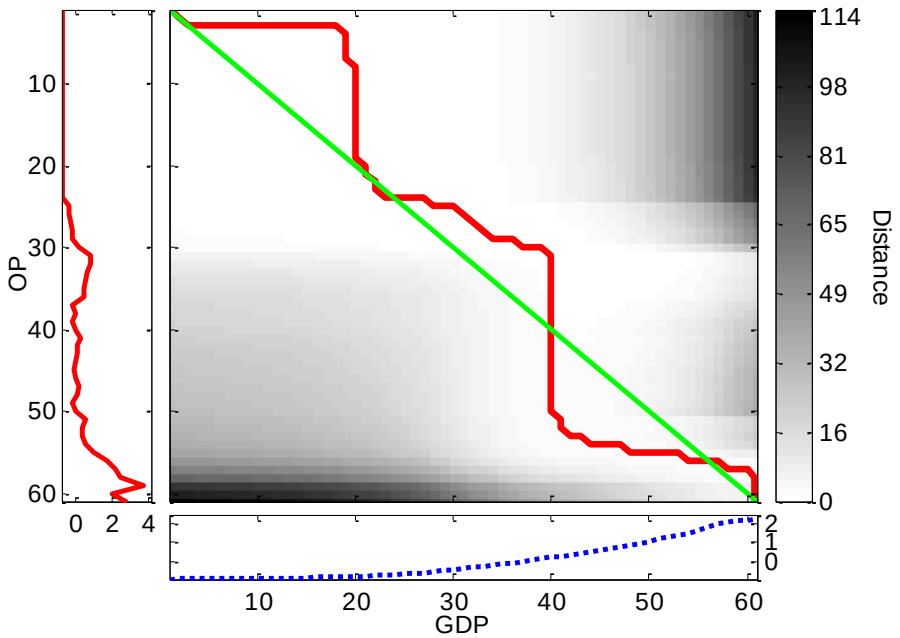

Figure 4.66. DTW distance matrix and the optimal warping path for the OP and GDP data.

The mode of the time shifts between the series in the warping path is zero, indicating that there is no lead/lag relation between the series, and the mean of the time shifts is −2, indicating that OP precedes GDP.

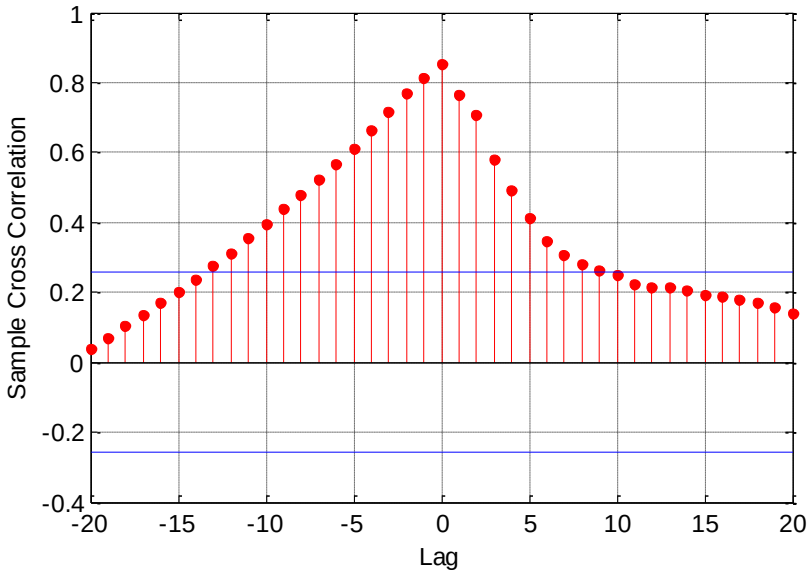

Figure 4.67. Sample cross-correlation function for the OP and GDP data.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relation between OP and GDP.

The results of the tests for the CEPCI – GDP data sets are as follows:

Table 4.39. GC test results for the CEPCI and GDP data.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
GDP does not Granger cause CEPCI	4.027	10.178	Reject H_0
CEPCI does not Granger cause GDP	4.027	20.530	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as five, one and five, one for CEPCI – GDP and GDP – CEPCI pairs, respectively.

The mode of the time shifts between the series in the warping path is zero, indicating that there is no lead/lag relation between the series, and the mean of the time shifts is three, indicating that CEPCI lags behind GDP.

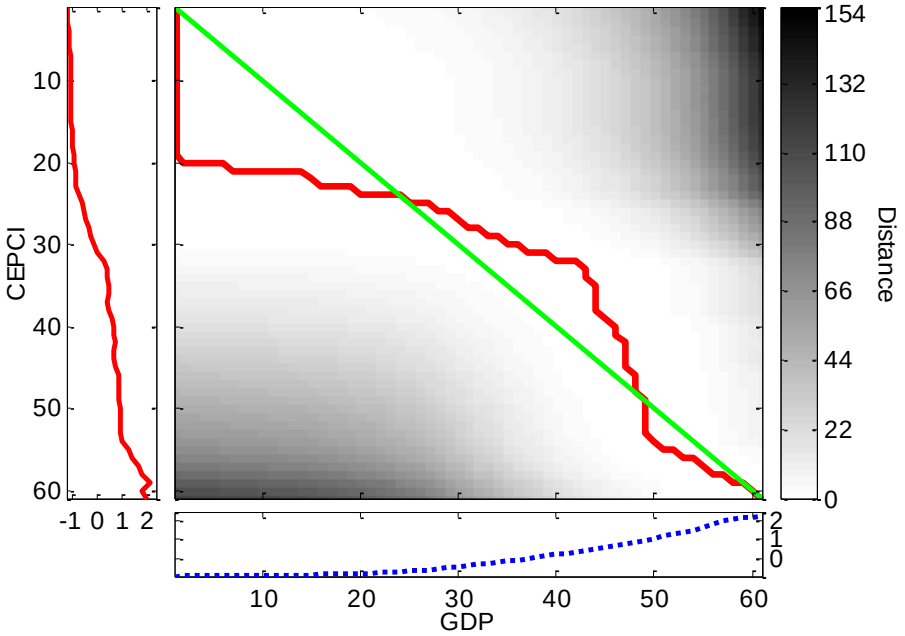

Figure 4.68. DTW distance matrix and the optimal warping path for the CEPCI and GDP data.

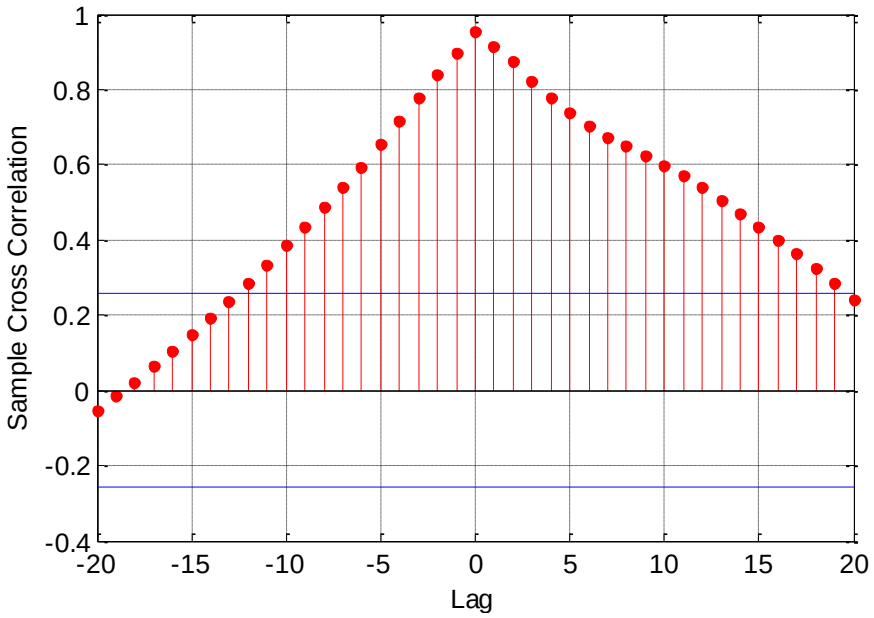

Figure 4.69. Sample cross-correlation function for the CEPCI and GDP data.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relation between CEPCI and GDP.

Table 4.40 summarizes the results of the GC test, DTW, and cross-correlation analyses to the index pairs, stated by Table 4.34 through Table 4.39 and Figure 4.58 through Figure 4.69.

Table 4.40. Summary of the results for the economic indexes.

Pair	GC	DTW	Cross-correlation
CEPCI & IP	CEPCI → IP	CEPCI ⇔ IP	No precedence
CEPCI & OP	CEPCI → OP	CEPCI ⇔ OP	No precedence
IP & OP	IP ← OP	IP ⇔ OP	No precedence
IP & GDP	IP ← GDP	IP ⇔ GDP	No precedence
OP & GDP	OP ↔ GDP	OP ⇔ GDP	No precedence
CEPCI & GDP	CEPCI ↔ GDP	CEPCI ⇔ GDP	No precedence

The GC results, showing a causal relation directed from OP to IP and OP to GDP, are as expected. CEPCI is found to be causing IP, OP and GDP, however this should not be the case since these indicators do not actually depend on CEPCI, on the contrary CEPCI depends mainly on the other economic indicators. In these cases, the GC test could not reflect the true/expected relation. When the optimal warping paths of DTW are studied, regional changes are seen around the diagonal. Therefore, DTW incorrectly suggests that all these pairs precede one another at different times. In these applications, the cross-correlation analysis implies that the series have the highest correlation as they are and no causal relation is discovered.

4.6. Predator-Prey Model

The predator-prey models are nonlinear differential equations of two interacting variables, a predator and a prey. They can also characterize parasite-host models, virus-immune system models and so on. These predator-prey interactions are represented by Lotka-Volterra equations (Hoppensteadt, 2006).

In this example, the prey is rabbits and the predator is foxes. The models for the prey (rabbit) population and the predator (fox) population are described by Equation 4.18 and Equation 4.19 respectively. In Equation 4.18, the coefficient α is the growth rate of rabbits

when there is no interaction with foxes, β is the coefficient of the predator-prey interactions that result in a decrease in the rabbit population. The coefficient γ in Equation 4.19 is the death or emigration rate of foxes in the absence of interactions with rabbits and δ is the coefficient of the growth of the fox population in case of interaction with rabbit population.

$$\frac{dr}{dt}=\alpha r-\beta rf \tag{4.18}$$

$$\frac{df}{dt}=-\gamma f+\delta rf \tag{4.19}$$

The equation parameters are α=2, β=2, γ=1, δ=1, and the initial values are r_0=1, f_0=3. The differential equations are solved with fourth order Runge-Kutta method and h=0.1 is used as the step size (Wang, 1998).

Some additional parameters and properties of the data and tests are as follows: Each series is made of 101 data points. Z-score normalization is applied before DTW, cross-correlation calculations, and the GC test. Series are non-stationary and not cointegrated, third order difference is taken for the GC test to make the series stationary with respect to ADF and KPSS tests. Significance level for the GC test is 0.95.

The results of the tests are as follows:

Table 4.41. GC test results for the predator-prey data.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
Predator does not Granger cause prey	2.319	5647.191	Reject H_0
Prey does not Granger cause predator	2.319	6821.313	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as five, five and five, five for Prey – Predator and Predator – Prey pairs, respectively.

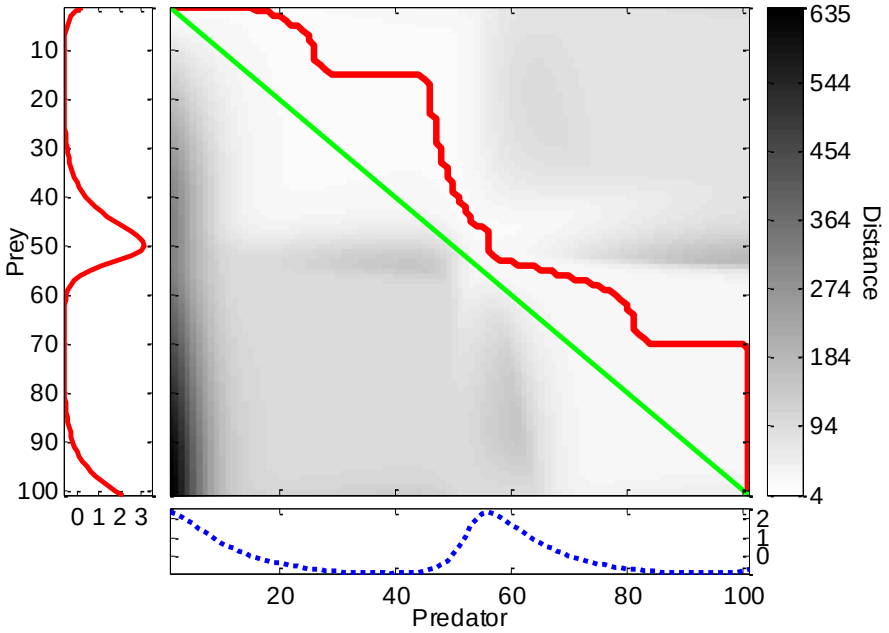

Figure 4.70. DTW distance matrix and the optimal warping path for the predator-prey data.

The mode of the time shifts between the series in the warping path is −14 and the mean of the time shifts is −16, indicating that the prey leads the predator (prey ⇒ predator).

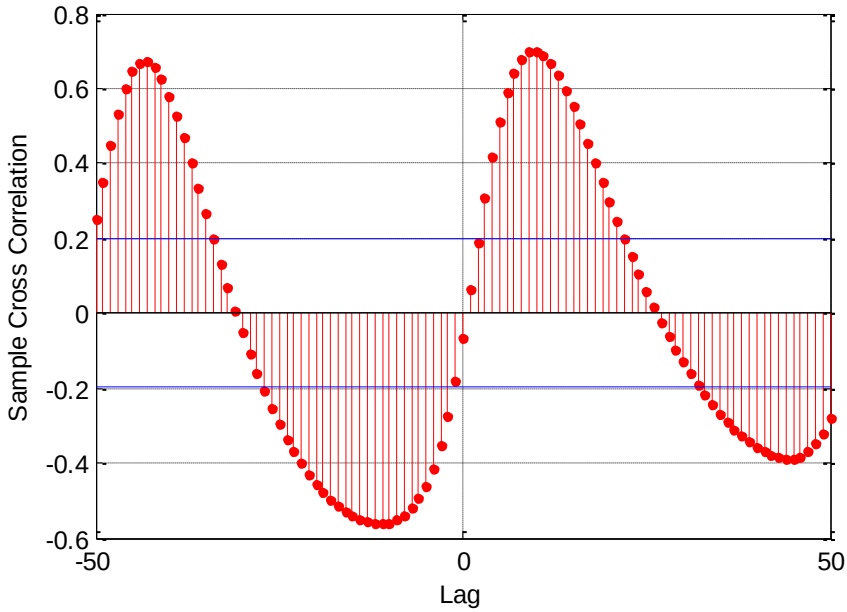

Figure 4.71. Sample cross-correlation function for the predator-prey data.

The maximum value of cross-correlation is observed at lag 10, indicating that the prey leads the predator (prey ⇒ predator).

Table 4.42 summarizes the results of the GC test, DTW, and cross-correlation analyses, stated by Table 4.41, Figure 4.70 and Figure 4.71.

Table 4.42. Summary of the results for the predator-prey data.

Pair	**GC**	**DTW**	**Cross-correlation**
Prey & Predator	Prey ↔ Predator	Prey ⇒ Predator	Prey ⇒ Predator

When the model equations, Equation 4.18 and Equation 4.19 are investigated, it can be expected that the prey population causes the predator population and vice versa. However, DTW warps the series in a way that suggests that only the prey population leads the predator population. Cross-correlation analysis indicates the same relation, as prey ⇒ predator. The GC test results on the other hand, correctly reflect the expected relation and imply that both the prey population causes the predator population (Prey → Predator) and the predator population causes the prey population (Prey ← Predator).

4.7. Biochemical System Models

The two models in this section describe glycolytic oscillations. They were used as test models by Lebiedz and Skanda in their study of dynamic biochemical systems (Skanda and Lebiedz, 2010). Equation 4.20 and Equation 4.21 represent the first and second models respectively. In the first model, positive feedback is observed with linear product sink. In the second model, the product sink is expressed with Michaelis-Menten kinetics. Series are expected to cause one another in both models.

$$\begin{aligned} \frac{d\alpha}{dt} &= \nu - \sigma\varphi(\alpha,\gamma), \\ \frac{d\gamma}{dt} &= q\sigma\varphi(\alpha,\gamma) - k_s\gamma, \\ \varphi(\alpha,\gamma) &= \frac{\alpha(1+\alpha)(1+\gamma)^2}{L+(1+\alpha)^2(1+\gamma)^2}. \end{aligned} \tag{4.20}$$

$$\begin{aligned} \frac{d\alpha}{dt} &= \nu - \sigma\varphi(\alpha,\gamma), \\ \frac{d\gamma}{dt} &= q\varphi(\alpha,\gamma) - \frac{r_s\gamma}{\mu+\gamma}, \\ \varphi(\alpha,\gamma) &= \frac{\alpha(1+\gamma)}{L+(1+\alpha)(1+\gamma)}. \end{aligned} \tag{4.21}$$

In these equations α is the substrate concentration and γ is the product concentration, ν is the flow parameter and it is taken as 0.22. The values of the additional parameters are $\sigma=8.79\times10^{-1}$, q=2.03 , $k_s=1.09\times10^{-1}$, and $L=1.61\times10^{4}$, for the first model and the initial concentrations are $\alpha_0=1.50\times10^{1}$, $\gamma_0=1.92$. For the second model q=6.47, $r_s=3.33$, μ=5.11, and $L=2.44\times10^{2}$ are used, and the initial concentrations are taken as $\alpha_0=1.47\times10^{1}$, $\gamma_0=1.16$.

Some additional parameters and properties of the data and tests are as follows: Each series is made of 201 data points. Z-score normalization is applied before DTW, cross-correlation calculations, and the GC test. Series in the first and second models are non-stationary and not cointegrated, fifth order difference is taken for the first model and fourth order difference is taken for the second model before the GC test to make the series stationary with respect to ADF and KPSS tests. Significance level for the GC test is 0.95. The solution of the differential equations is performed by using MATLAB.

The results of the tests for the first biochemical system data are as follows:

Table 4.43. GC test results for the first biochemical system data.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
Product does not Granger cause substrate	2.263	135.522	Reject H_0
Substrate does not Granger cause product	2.263	68.718	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as five, five and five, five for Substrate – Product and Product – Substrate pairs, respectively.

The mode of the time shifts between the series in the warping path is −12 and the mean of the time shifts is −13, indicating that the substrate precedes the product (substrate ⇒ product).

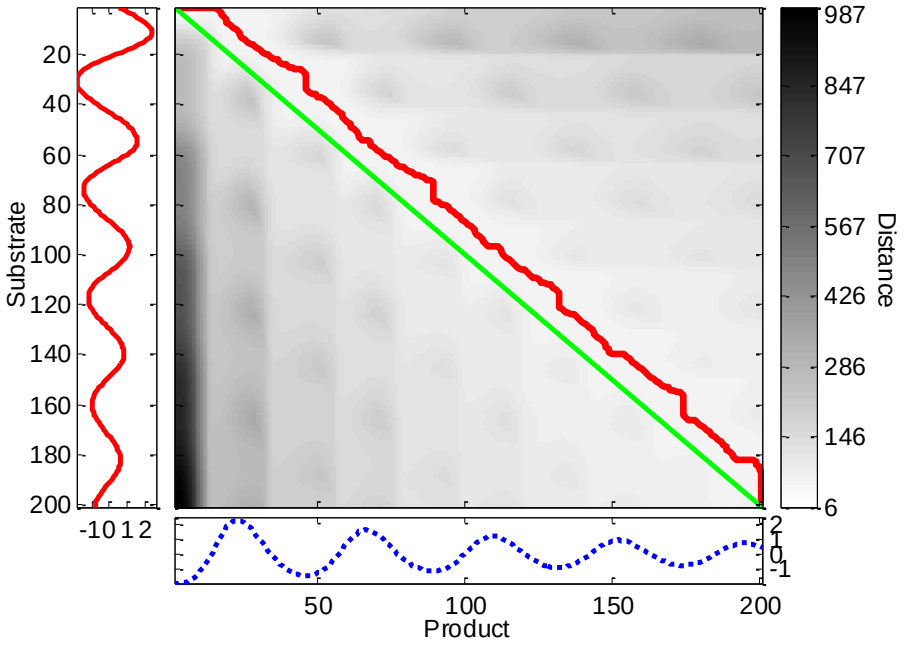

Figure 4.72. DTW distance matrix and the optimal warping path for the first biochemical system data.

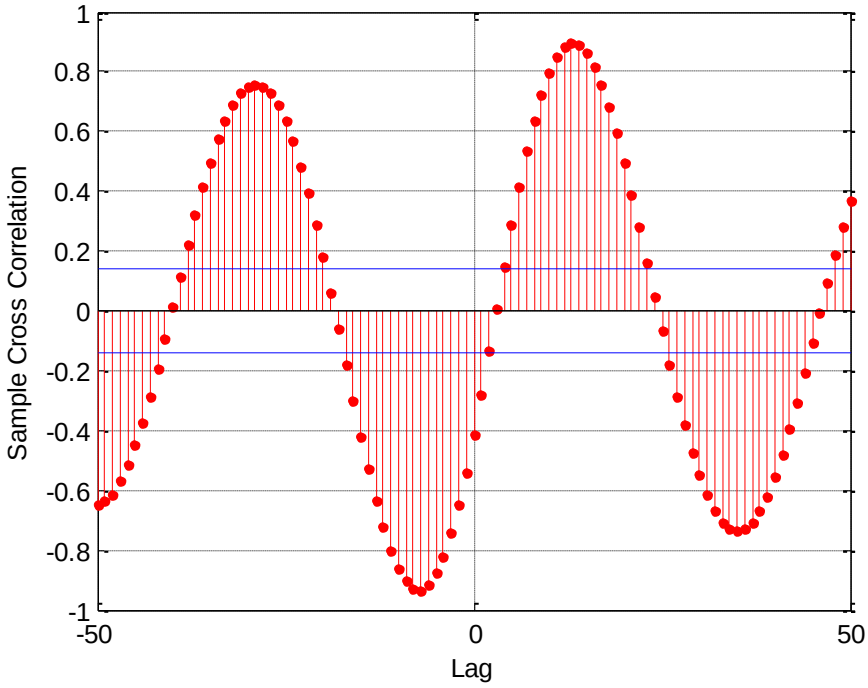

Figure 4.73. Sample cross-correlation function for the first biochemical system data.

The maximum value of cross-correlation is observed at lag 13, indicating that the substrate precedes the product (substrate ⇒ product).

Table 4.44 summarizes the results of the GC test, DTW, and cross-correlation analyses, stated by Table 4.43, Figure 4.72 and Figure 4.73.

Table 4.44. Summary of the results for the first biochemical system data.

Pair	**GC**	**DTW**	**Cross-correlation**
Substrate & Product	Substrate ↔ Product	Substrate ⇒ Product	Substrate ⇒Product

The results of the applications are similar to the results obtained for the predator – prey model. The DTW and cross-correlation analyses show that substrate population leads the product population for all times (Substrate ⇒ Product). However, the GC test results correctly suggest a feedback relation between substrate and product series (Substrate ↔ Product).

The results of the tests for the first biochemical system data are as follows:

Table 4.45. GC test results for the second biochemical system data.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
Product does not Granger cause substrate	2.420	1905.611	Reject H_0
Substrate does not Granger cause product	2.420	400.144	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as five, four and five, four for Substrate – Product and Product – Substrate pairs, respectively.

The mode and the mean of the time shifts between the series in the warping path are −12, indicating that the substrate precedes the product (substrate ⇒ product).

The maximum value of cross-correlation is observed at lag 13, indicating that the substrate precedes the product (substrate ⇒ product).

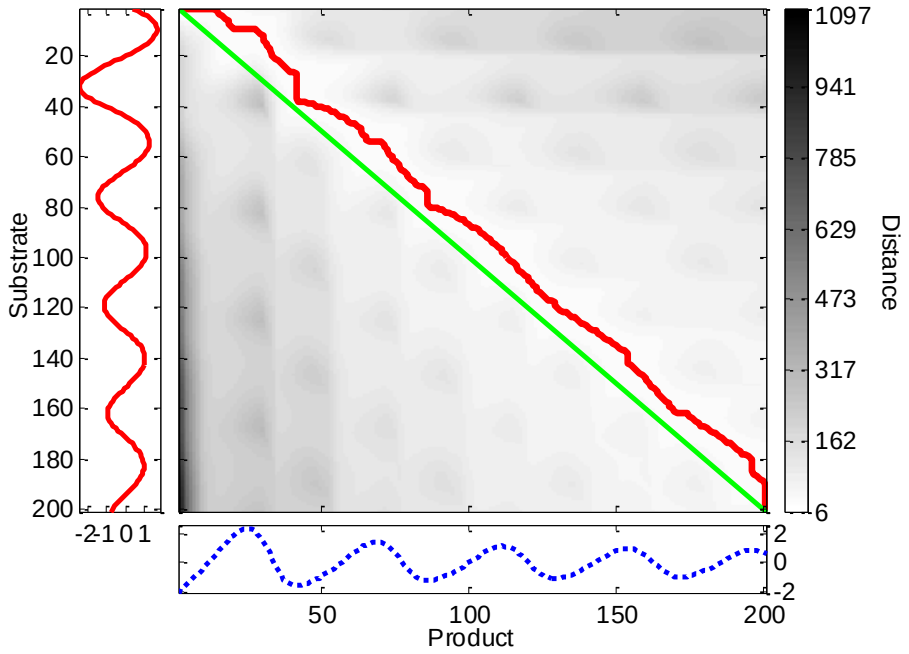

Figure 4.74. DTW distance matrix and the optimal warping path for the second biochemical system data.

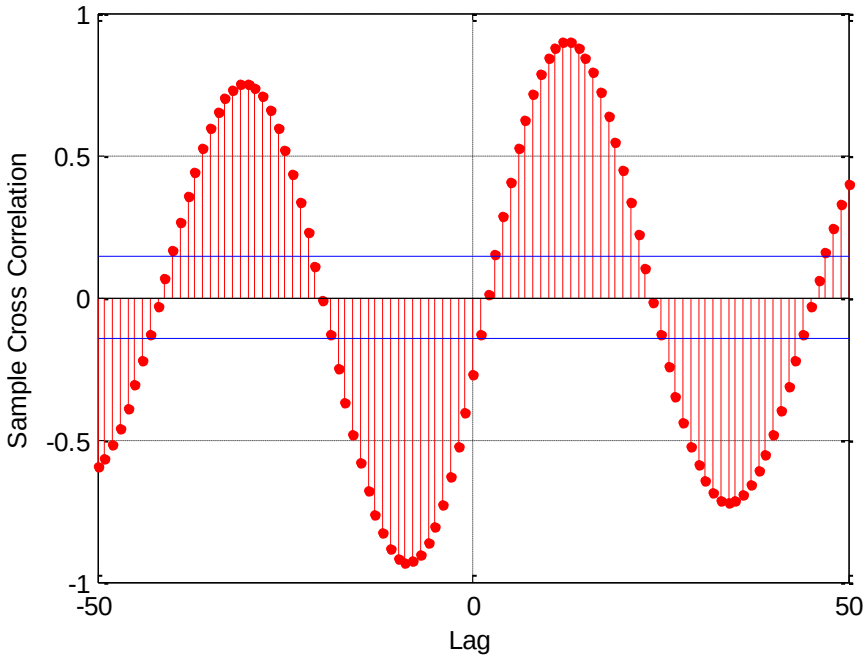

Figure 4.75. Sample cross-correlation function for the second biochemical system data.

Table 4.46 summarizes the results of the GC test, DTW, and cross-correlation analyses, stated by Table 4.45, Figure 4.74 and Figure 4.75.

Table 4.46. Summary of the results for the second biochemical system data.

Pair	GC	DTW	Cross-correlation
Substrate & Product	Substrate ↔ Product	Substrate ⇒ Product	Substrate ⇒Product

The same results are acquired for the second model. The GC test provides the expected results correctly (Substrate ↔ Product) whereas the DTW and cross-correlation analyses show only a one-way relation from substrate to product (Substrate ⇒ Product).

4.8. Chemical Reaction Models

Two example problems of chemical reactions are examined in this section. The tests are applied to the concentration data of the chemical species that are either used or generated in these reactions.

In the analysis of the reactions, the expected result for the DTW is to find no lead/lag relation between the series. The concentration of the products depends on the concentration of the reactants through the balance equations and rate laws however for a lead/lag relation to exist there should be a time difference in the way the series affect each other, one of the series should be affected by the past values of the other series. For this reason, no precedence is expected to be found between the series after warping.

4.8.1. Irreversible Reaction in a Semibatch Reactor

The example problem is taken from Fogler (2008).

"The production of methyl bromide is an irreversible liquid-phase reaction that follows an elementary rate law. The reaction,

$$CNBr+CH_3NH_2 \rightarrow CH_3Br+NCNH_2$$

is carried out isothermally in a semibatch reactor. An aqueous solution of methyl amine (B) at a concentration of 0.025 mol/dm^3 is to be fed at a rate of 0.05 dm^3/s to an aqueous solution of bromine cyanide (A) contained in a glass-lined reactor. The initial volume of

fluid in the vat is to be 5 dm^3 with a bromine cyanide concentration of 0.05 mol/dm^3. The specific reaction rate constant, k is 2.2 $dm^3/s.mol$".

The differential equations, Equation 4.22 through Equation 4.25 are constructed for each compound.

$$\frac{dC_A}{dt} = -kC_AC_B - \frac{v_0C_A}{V} \tag{4.22}$$

$$\frac{dC_B}{dt} = -kC_AC_B + \frac{v_0(C_{B0} - C_B)}{V} \tag{4.23}$$

$$\frac{dC_C}{dt} = kC_AC_B - \frac{v_0C_C}{V} \tag{4.24}$$

$$\frac{dC_D}{dt} = kC_AC_B - \frac{v_0C_D}{V} \tag{4.25}$$

C_A, C_B, C_C and C_D are CNBr, CH_3NH_2, CH_3Br and $NCNH_2$ concentrations respectively, and v_0 is the volumetric flow rate. Additionally, the volume of the reactor, V is expressed as a function of time in Equation 4.26, with V_0 as the initial volume.

$$V = V_0 + v_0t \tag{4.26}$$

Polymath is utilized for the solution of the equations. Some parameters and properties of the data and tests are as follows: Each series is made of 100 data points. Z-score normalization is applied before DTW, cross-correlation calculations, and the GC test. Series are non-stationary and not cointegrated, second order difference is taken for the GC test to make the series stationary with respect to ADF and KPSS tests. Significance level for the GC test is 0.95.

The results of the tests for the series A and C are as follows:

Table 4.47. GC test results for the series A and C of the first reaction example.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
C does not Granger cause A	2.475	18.655	Reject H_0
A does not Granger cause C	2.475	27.610	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as five, four and five, four for A – C and C – A pairs, respectively.

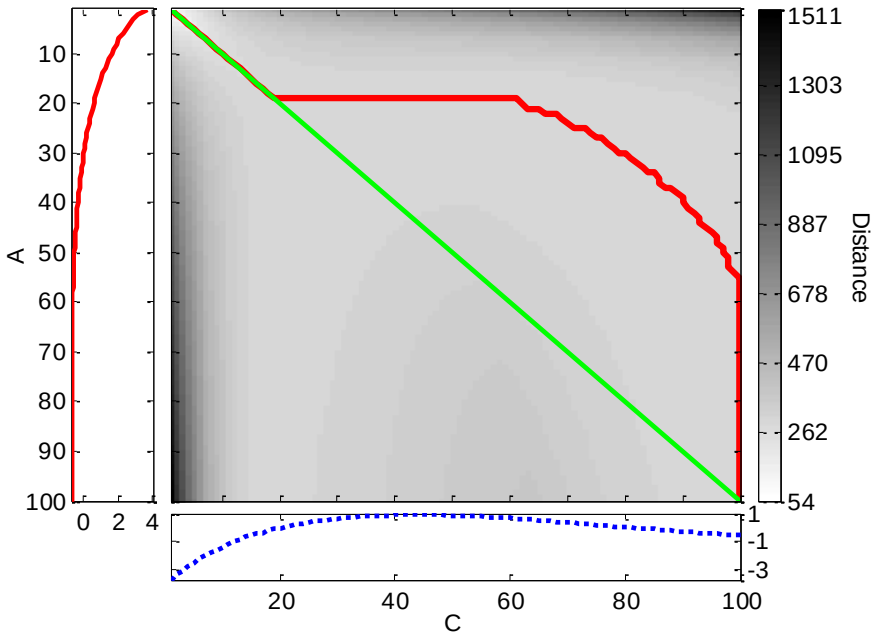

Figure 4.76. DTW distance matrix and the optimal warping path for the series A and C of the first reaction example.

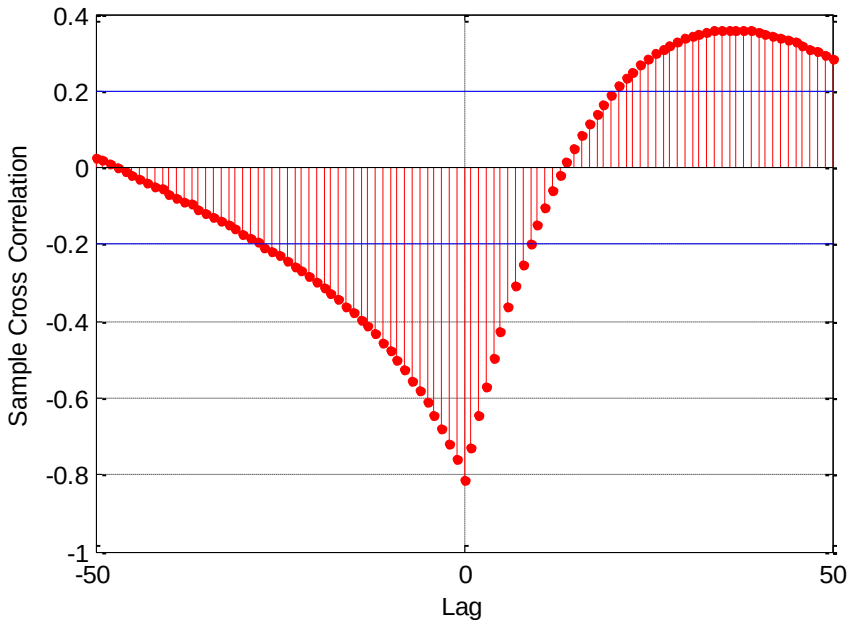

Figure 4.77. Sample cross-correlation function for the series A and C of the first reaction example.

The mode of the time shifts between the series in the warping path is zero, indicating that there is no lead/lag relation between the series and the mean of the time shifts is −27, indicating that the series A leads the series C (A ⇒ C).

The maximum value of cross-correlation is observed at lag 37, indicating that the series A leads the series C (A ⇒ C).

The results of the tests for the series B and D are as follows:

Table 4.48. GC test results for the series B and D of the first reaction example.

Null Hypothesis (H_0)	Critical Value	F-Statistics	Result
D does not Granger cause B	2.319	21.055	Reject H_0
B does not Granger cause D	2.475	27.620	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as five, five and five, four for B – D and D – B pairs, respectively.

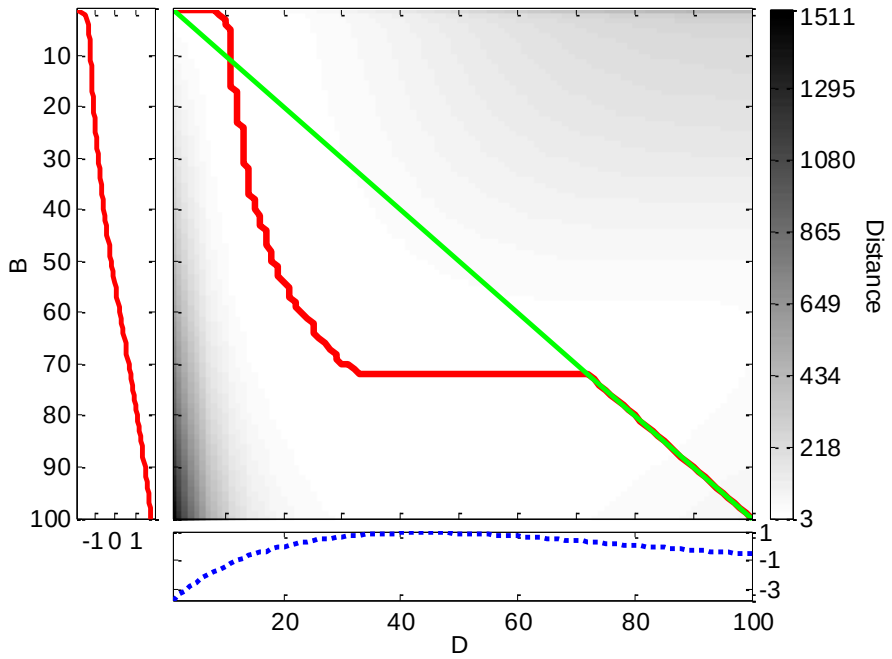

Figure 4.78. DTW distance matrix and the optimal warping path for the series B and D of the first reaction example.

The mode of the time shifts between the series in the warping path is zero, indicating that there is no lead/lag relation between the series and the mean of the time shifts is 15, indicating that the series B lags behind the series D (B ⇐ D).

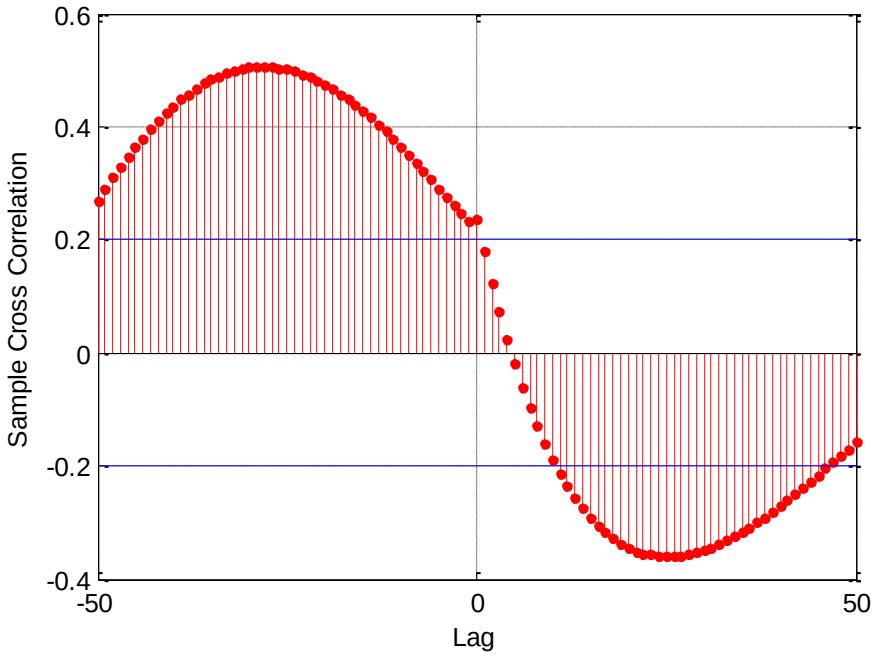

Figure 4.79. Sample cross-correlation function for the series B and D of the first reaction example.

The maximum value of cross-correlation is observed at lag −28, indicating that the series B lags behind the series D (B ⇐ D).

Table 4.49 summarizes the results of the GC test, DTW, and cross-correlation analyses, stated by Table 4.47 and Table 4.48, and Figure 4.76 through Figure 4.79.

DTW results mainly suggest a leading relation from the reactant A to the product C (A ⇒ C) and from the reactant B to the product D (B ⇐ D). These results are in accordance with the cross-correlation analysis results. On the other hand, the GC test shows a feedback relation between the compounds A and C and the compounds B and D, (A ↔ C and B ↔ D).

Table 4.49. Summary of the results for the first chemical reaction example.

Pair	GC	DTW	Cross-correlation
A & C	A ↔ C	A ⇒ C	A ⇒ C
B & D	B ↔ D	B ⇐ D	B ⇐ D

4.8.2. Reversible and Irreversible Reactions in a CSTR

In this example problem, the following reversible and irreversible reactions take place in a CSTR.

$$A \leftrightarrow B \rightarrow C$$

The reaction rate constant for the forward direction of the first reaction, k_1 is two and for the reverse direction, k_2 is 7/8. The rate constant of the second reaction producing C, k_3 is 1/8 (Akman, 2011).

The differential equations written for each compound are

$$\frac{dC_A}{dt} = -k_1 C_A + k_2 C_B \tag{4.27}$$

$$\frac{dC_B}{dt} = -k_1 C_A + (k_2 + k_3) C_B \tag{4.28}$$

$$\frac{dC_C}{dt} = k_3 C_B \tag{4.29}$$

The initial conditions are

$$@\ t=0 \qquad C_A=1, \quad C_B=0, \quad C_C=0$$

The analytical solutions are as follows.

$$C_A = \left(\frac{r_2 - k_1}{r_2 - r_1}\right) e^{-r_1 t} + \left(\frac{k_1 - r_1}{r_2 - r_1}\right) e^{-r_2 t} \tag{4.30}$$

$$C_B = \left(\frac{r_2 - k_1}{r_2 - r_1}\right) \left(\frac{k_1 - r_1}{k_2}\right) e^{-r_1 t} + \left(\frac{k_1 - r_1}{r_2 - r_1}\right) \left(\frac{k_1 - r_2}{k_2}\right) e^{-r_2 t} \tag{4.31}$$

$$C_C = \frac{k_3}{k_2}\left[\left(\frac{k_1 - r_2}{r_2 - r_1}\right)\left(\frac{k_1 - r_1}{r_1}\right)(e^{-r_1 t} - 1) + \left(\frac{r_1 - k_1}{r_2 - r_1}\right)\left(\frac{k_1 - r_2}{r_2}\right)(e^{-r_2 t} - 1)\right] \quad (4.32)$$

with $r_1 = 3/2 - \sqrt{2}$, and $r_2 = 3/2 + \sqrt{2}$.

Some parameters and properties of the data and tests are as follows: Each series is made of 100 data points. Z-score normalization is applied before DTW, cross-correlation calculations, and the GC test. Series are non-stationary but cointegrated. 15th order difference is taken for A – B set and 18th order difference is taken for B – C and A – C sets for the GC test to make the series stationary with respect to ADF and KPSS tests. Significance level for the GC test is 0.95.

The results of the tests for the series A and B are as follows:

Table 4.50. GC test results for the series A and B of the second reaction example.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
B does not Granger cause A	3.963	6.596	Reject H_0
A does not Granger cause B	3.963	0.793	Do not reject H_0

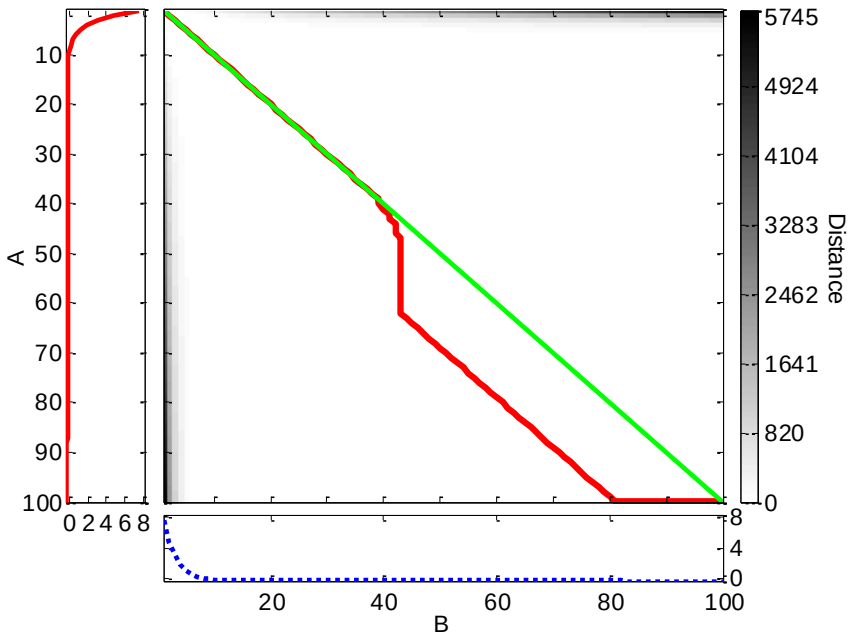

Figure 4.80. DTW distance matrix and the optimal warping path for the series A and B of the second reaction example.

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as five, one and five, one for A – B and B – A pairs, respectively.

The mode of the time shifts between the series in the warping path is zero, indicating that there is no lead/lag relation between the series and the mean of the time shifts is nine, indicating that the series A lags behind the series B (A ⇐ B).

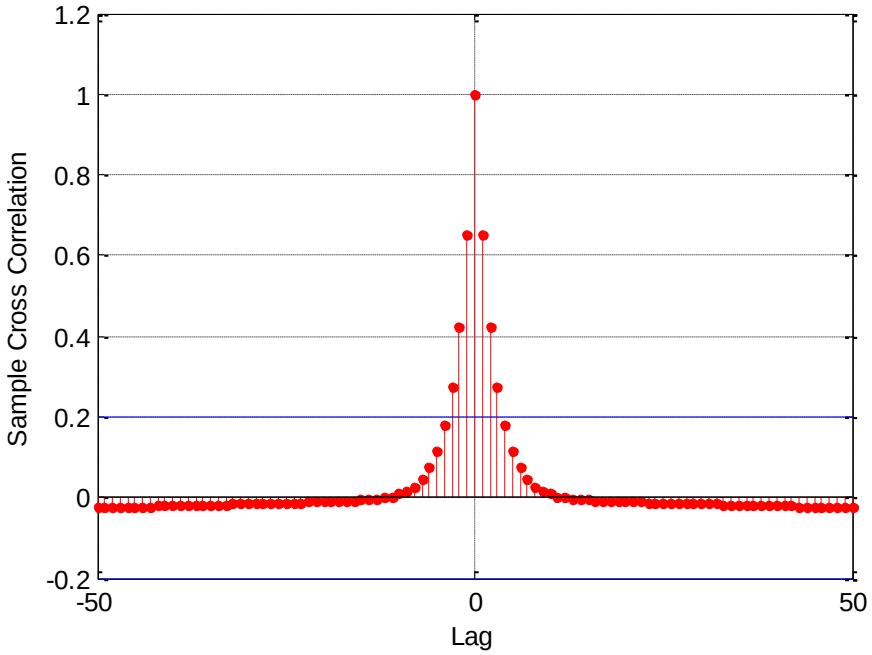

Figure 4.81. Sample cross-correlation function for the series A and B of the second reaction example.

The maximum value of cross-correlation is observed at zero lag, indicating that there is no lead/lag relationship between the series A and the series B.

The results of the tests for the series B and C are as follows:

Table 4.51. GC test results for the series B and C of the second reaction example.

Null Hypothesis (H_0)	**Critical Value**	**F-Statistics**	**Result**
C does not Granger cause B	2.499	57.940	Reject H_0
B does not Granger cause C	2.499	341.577	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as five, four and five, four for B – C and C – B pairs, respectively.

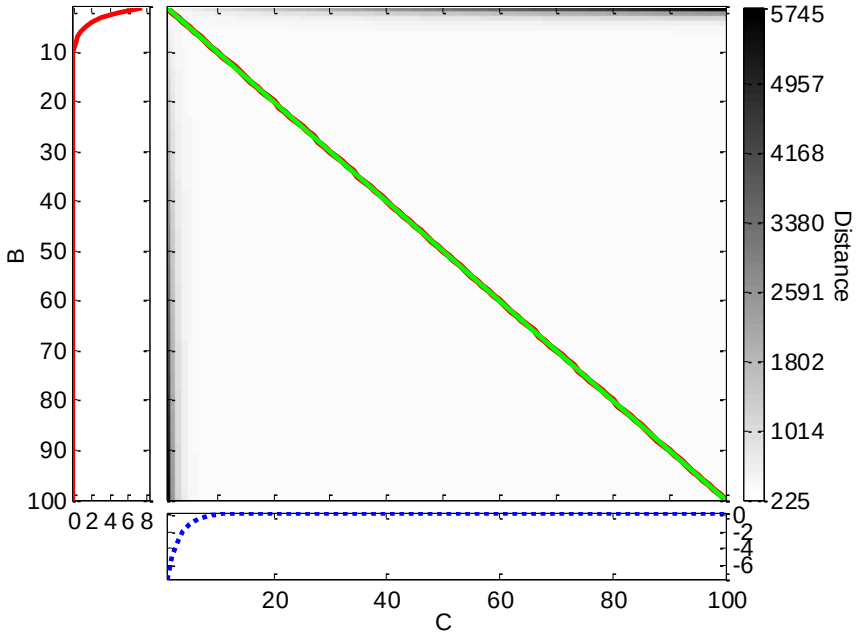

Figure 4.82. DTW distance matrix and the optimal warping path for the series B and C of the second reaction example.

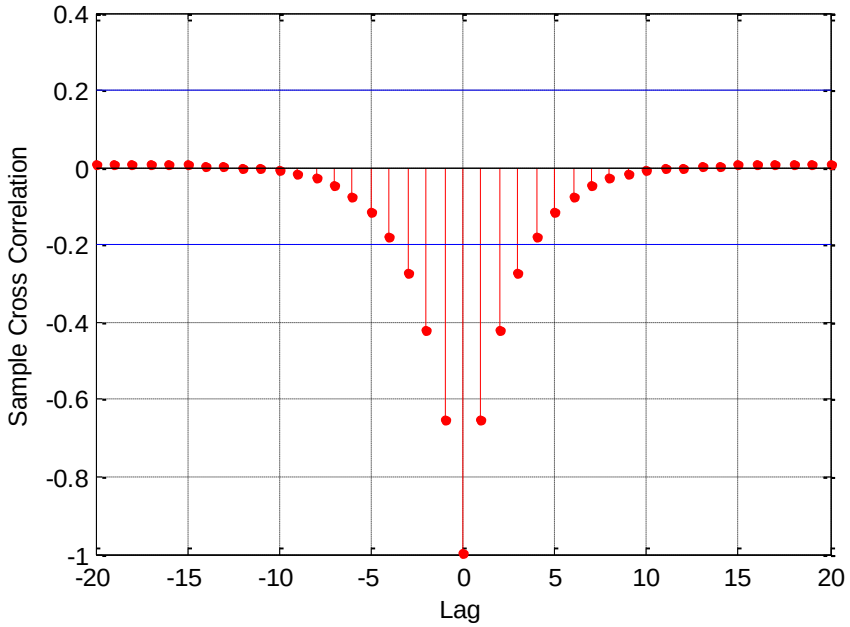

Figure 4.83. Sample cross-correlation function for the series B and C of the second reaction example.

The mode and the mean of the time shifts between the series in the warping path are zero, indicating that there is no lead/lag relation between the series B and the series C.

The maximum value of cross-correlation is observed at lag 20, indicating that the series B leads the series C (B ⇒ C).

The results of the tests for the series A and C are as follows:

Table 4.52. GC test results for the series A and C of the second reaction example.

Null Hypothesis (H_0)	Critical Value	F-Statistics	Result
C does not Granger cause A	2.344	81.195	Reject H_0
A does not Granger cause C	2.344	334.692	Reject H_0

The maximum number of ordered lags to be considered in the GC regression models is five, and the automated analysis based on BIC showed the optimum number of lags to be used in the GC tests as five, five and five, five for A – C and C – A pairs, respectively.

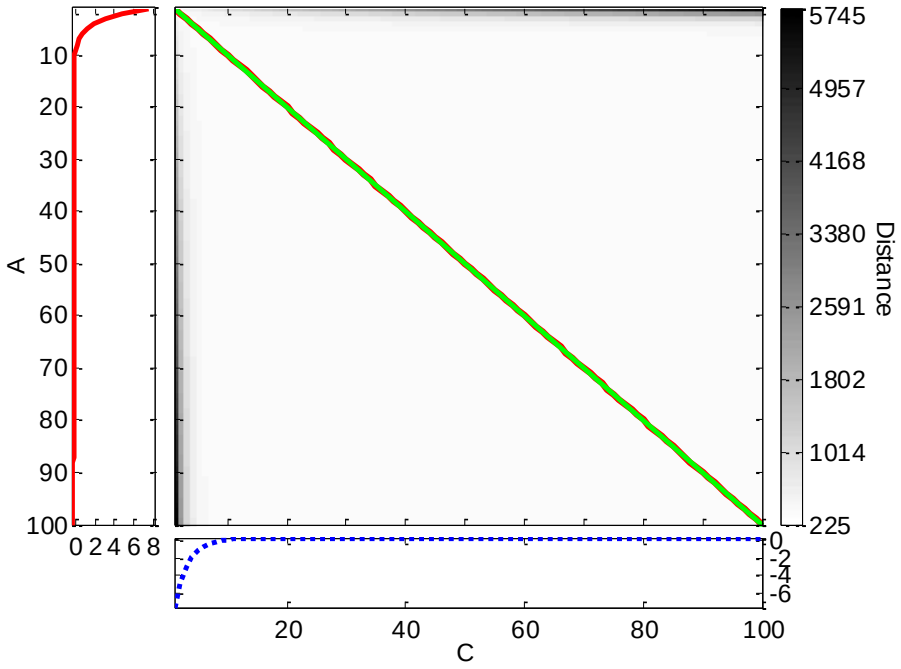

Figure 4.84. DTW distance matrix and the optimal warping path for the series A and C of the second reaction example.

The mode and the mean of the time shifts between the series in the warping path are zero, indicating that there is no lead/lag relation between the series A and the series C.

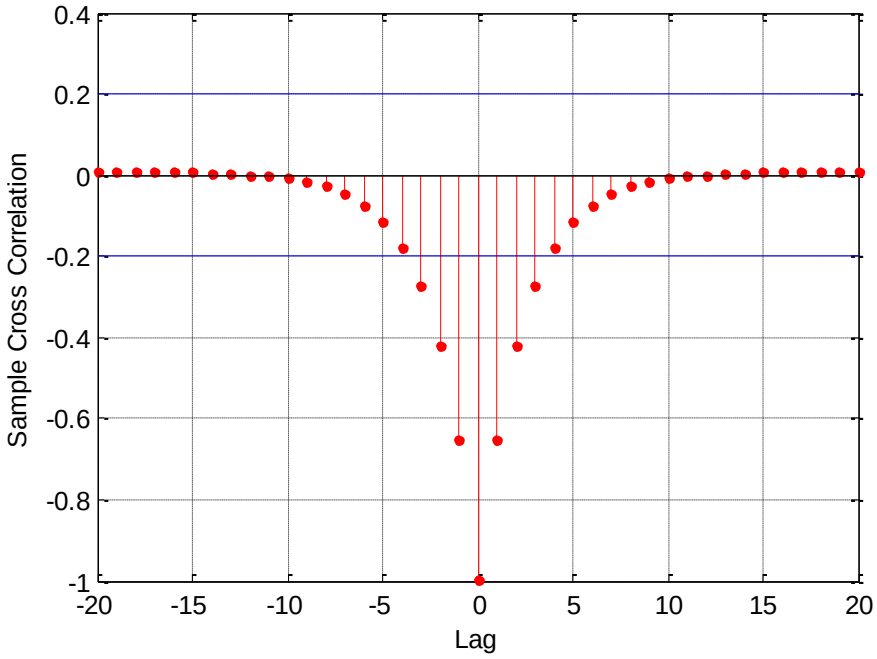

Figure 4.85. Sample cross-correlation function for the series A and C of the second reaction example.

The maximum value of cross-correlation is observed at lag −20, indicating that the series A lags behind the series C (A ⇐ C).

Table 4.53 summarizes the results of the GC test, DTW, and cross-correlation analyses, stated by Table 4.50 through Table 4.52 and Figure 4.80 through Figure 4.85.

Table 4.53. Summary of the results for the second chemical reaction example.

Pair	**GC**	**DTW**	**Cross-correlation**
A & B	A ← B	A ⇐ B	No precedence
B & C	B ↔ C	No precedence	B ⇒ C
A & C	A ↔ C	No precedence	A ⇐ C

DTW analysis suggests that there is no lead/lag relation between the compounds B and C, and A and C. In Figure 4.80 it is seen that, the compound B leads the compound A

for the most part of the series (A ⇐ B); on the other hand, for the beginning section no precedence is found between these series as well. Therefore, DTW results are close to what is expected. Cross-correlation analysis also detects no lead/lag relation for A and B, however it indicates a leading relation from B to C (B ⇒ C) and from C to A (A ⇐ C), and it is observed that at zero lag, the series are negatively correlated. The GC test incorrectly shows that the compound B causes A (A ← B). For the remaining pairs, the GC test shows feedback relations between the compounds (B ↔ C and A ↔ C).

4.9. Shifting the Series

Finally a change is employed on the DTW method by shifting the series before warping them. The reason for this variation in the series is because there is a boundary condition in DTW that requires the alignment of the first and the last elements of the series and this creates inflexibility in the warping around the beginning and end of the series.

In the following application, one of the series is shifted to the left or right with respect to the other depending on the sign of the lag to account for the limitation coming from the boundary condition. The first shifting value is chosen as the value giving the highest cross-correlation. By doing this, the overall time change (overall lag) between the series is aimed to be offset. Moreover, for a range of [–lag, +lag], the series are incrementally shifted and DTW is applied to the shifted series for all incremental lags. The results for each lag are presented on a figure. The investigated parameters are the cumulative distance between the series after warping (Dist), the length of the optimal warping path (K), the ratio of the distance to the path length (Dist/K) which normalizes the cumulative warping distance (warping cost) with the length of the warped path, the correlation between the unwarped series (CorOrg), the correlation between the series after warping (CorDTW) and the mode of the time shifts between the series in the warping path (TimeShift). The lags corresponding to the minimum value of Dist/K and the maximum value of CorDTW are chosen to shift the series by and perform DTW analysis. The tested series in this section are the first and the second examples of the nonlinear series models.

4.9.1. Two-Dimensional Nonlinear Time Series Model with Shifting

The analysis of the model in Section 4.3.1 with DTW presented conflicting results with the GC test and cross-correlation analysis. The following analysis is carried out by shifting the series x_1 and x_2 to the left or right by the lag at which the sample cross-correlation is at the maximum and by the amounts determined from the Figure 4.86. Figure 4.86 shows various results related to the DTW analysis for different lag values.

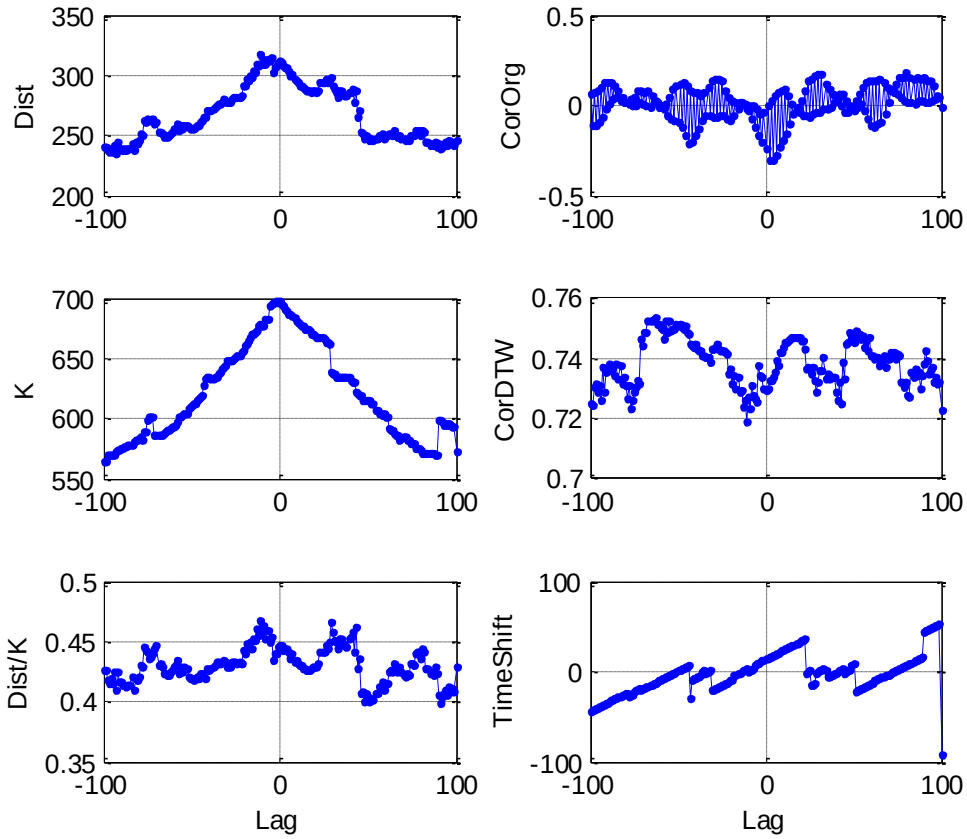

Figure 4.86. DTW results for shifted two-dimensional nonlinear time series.

The maximum value of cross-correlation was observed at lag 30. Therefore the series x_2 is shifted to the left by 30 with respect to x_1 and the results of DTW are as follows.

The mode of the time shifts between the series in the warping path is zero, indicating that there is no lead/lag relation between the series and the mean of the time shifts is −21, indicating that the series x_1 leads the series x_2 ($x_1 \Rightarrow x_2$).

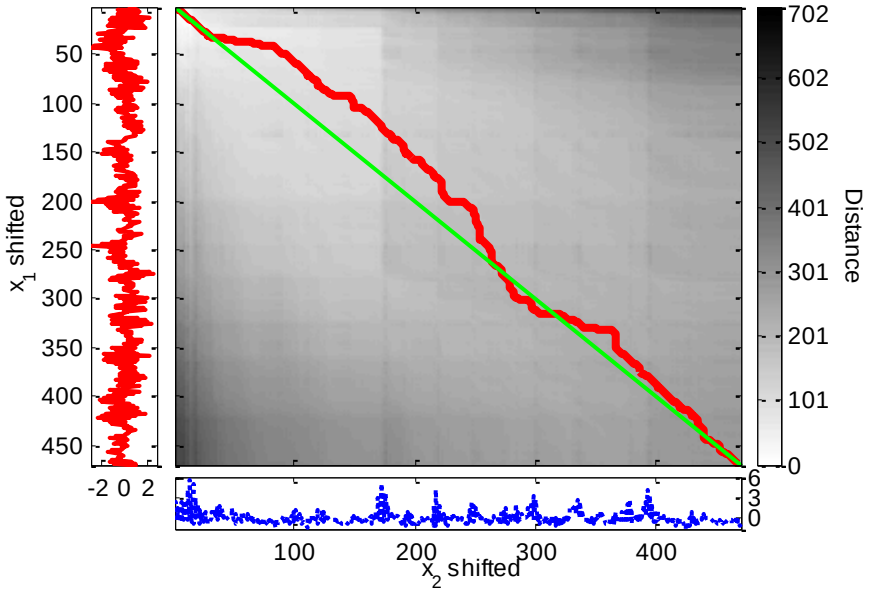

Figure 4.87. DTW distance matrix and the optimal warping path for shifted two-dimensional nonlinear time series (after x_2 is shifted to the left by 30 with respect to x_1).

By inspection of Figure 4.86 for the range of [−30, +30] lags, the minimum value of distance to path length ratio (Dist/K) is found at lag 16 with a value of 0.43. The series x_2 is shifted to the left by 16 with respect to x_1 and the results of DTW are as follows.

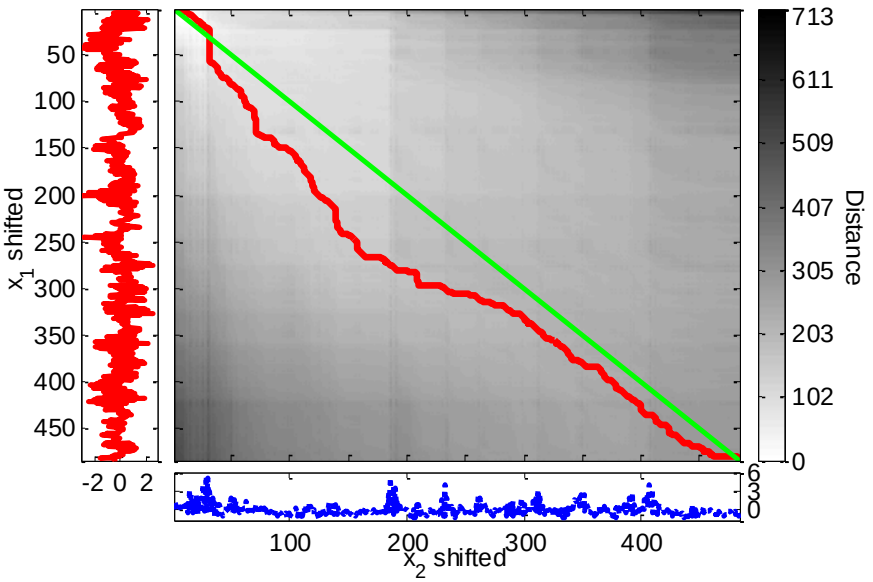

Figure 4.88. DTW distance matrix and the optimal warping path for the shifted two-dimensional nonlinear time series (after x_2 is shifted to the left by 16 with respect to x_1).

The mode of the time shifts between the series in the warping path is 30 and the mean of the time shifts is 44, indicating that the series x_1 lags behind the series x_2 ($x_1 \Leftarrow x_2$). This is the same result obtained in Figure 4.44 before the series are shifted.

For the range of [−30, +30] lags, the maximum value of correlation between the warped series (CorDTW) is obtained at lag 17 with a value of 0.75 as can be seen from Fig. 4.86. The series x_2 is shifted to the left by 17 with respect to x_1 and the results of DTW are as follows.

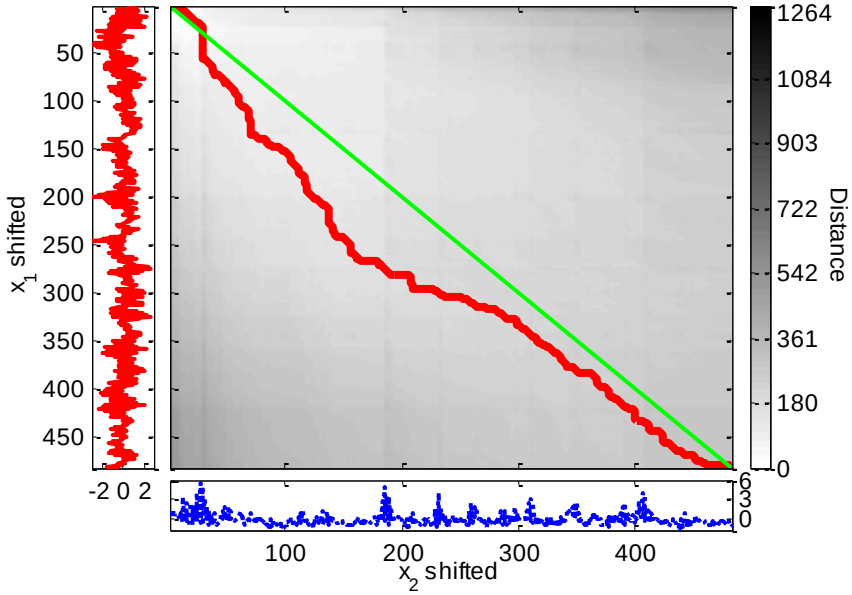

Figure 4.89. DTW distance matrix and the optimal warping path for the shifted two-dimensional nonlinear time series (after x_2 is shifted to the left by 17 with respect to x_1).

The mode of the time shifts between the series in the warping path is 31 and the mean of the time shifts is 45, indicating that the series x_1 lags behind the series x_2 ($x_1 \Leftarrow x_2$). Once more the same result is obtained.

Table 4.54. Summary of the results for the shifted two-dimensional nonlinear time series.

Shift by	$\mathbf{x_1 - x_2}$
Max cross-correlation	$x_1 \Rightarrow x_2$
Min Dist/K	$x_1 \Leftarrow x_2$
Max CorDTW	$x_1 \Leftarrow x_2$

When the results from Table 4.54 are compared with the GC test, DTW and cross-correlation results of the unshifted series from Table 4.26 and with the expected results from Figure 4.43, it is seen that shifting the series by a lag corresponding to the maximum value of cross-correlation produced correct results ($x_1 \Rightarrow x_2$) unlike the results of the unshifted series, or the shifts by the minimum of Dist/K or the maximum of CorDTW ($x_1 \Leftarrow x_2$). The expected result was also obtained by the GC test and cross-correlation analysis on the unshifted series.

4.9.2. Four-Dimensional Nonlinear Time Series Model with Shifting

The analysis of the model in Section 4.3.2 with DTW also showed conflicting results with what was expected. Consequently, the series are shifted by the lag of the maximum cross-correlation and by the lags determined from the Figure 4.90, Figure 4.94, Figure 4.96 and Figure 4.98.

The results of the tests for the pair $x_1 - x_2$ are as follows.

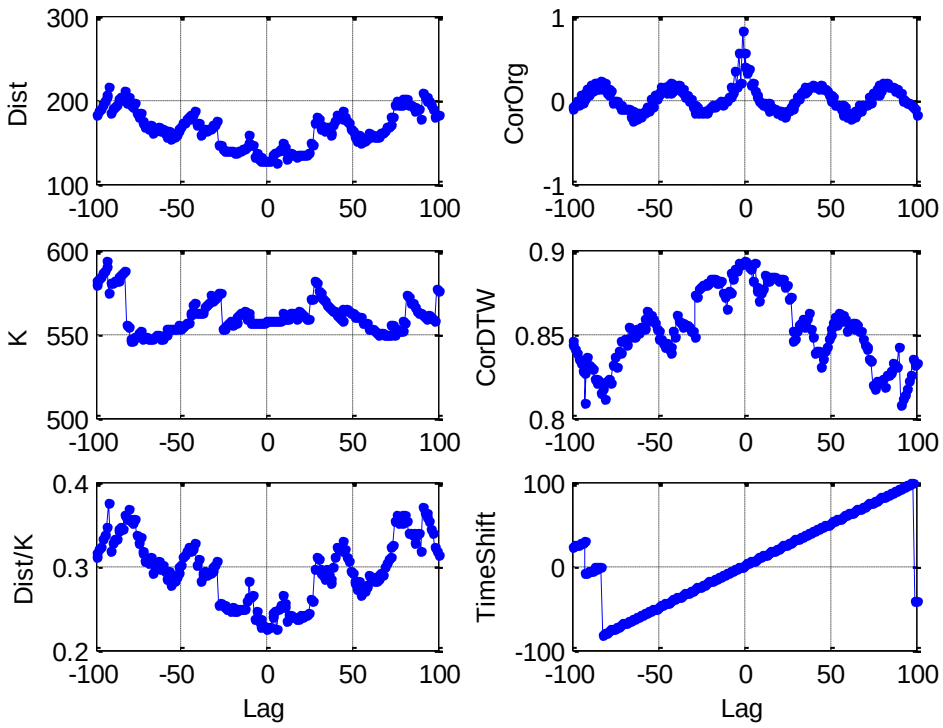

Figure 4.90. DTW results for shifted four-dimensional nonlinear time series for the $x_1 - x_2$ pair.

The maximum value of cross-correlation was observed at lag −1. Therefore the series x_1 is shifted to the left by one with respect to x_2 and the results of DTW are as follows.

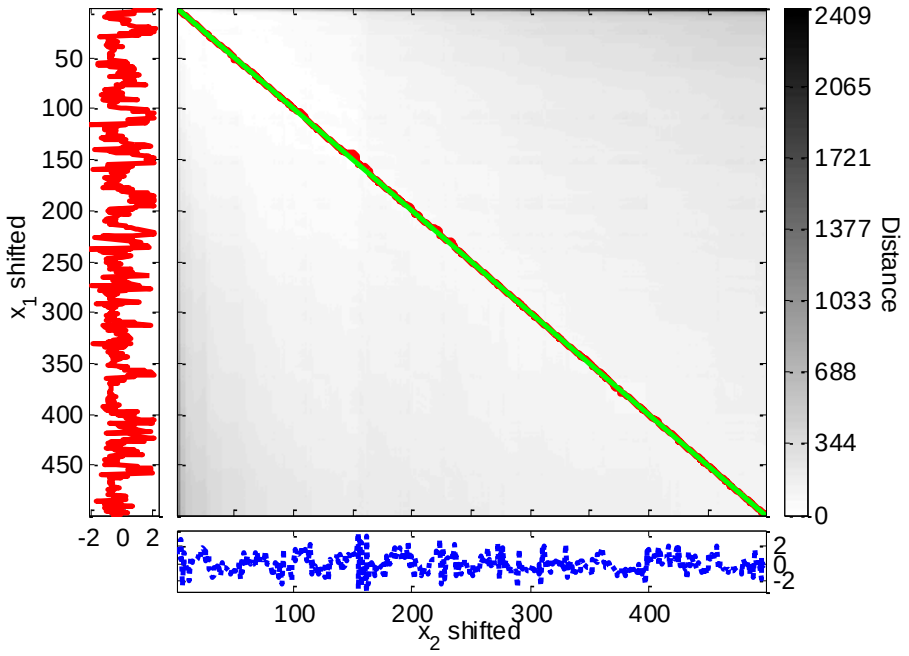

Figure 4.91. DTW distance matrix and the optimal warping path for the $x_1 - x_2$ pair of the shifted four-dimensional nonlinear time series (after x_1 is shifted to the left by one with respect to x_2).

The mode and the mean of the time shifts between the series in the warping path are zero, indicating that there is no lead/lag relation between the series x_1 and the series x_2.

By inspection of Figure 4.90, the minimum value of distance to path length ratio (Dist/K) is found at lag six with a value of 0.22. The series x_2 is shifted to the left by six with respect to x_1 and the results of DTW are as follows.

The mode of the time shifts between the series in the warping path is seven and the mean of the time shifts is six, indicating that the series x_1 lags behind the series x_2 ($x_1 \Leftarrow x_2$). This is the same result as the unshifted series from Figure 4.47 and Table 4.26.

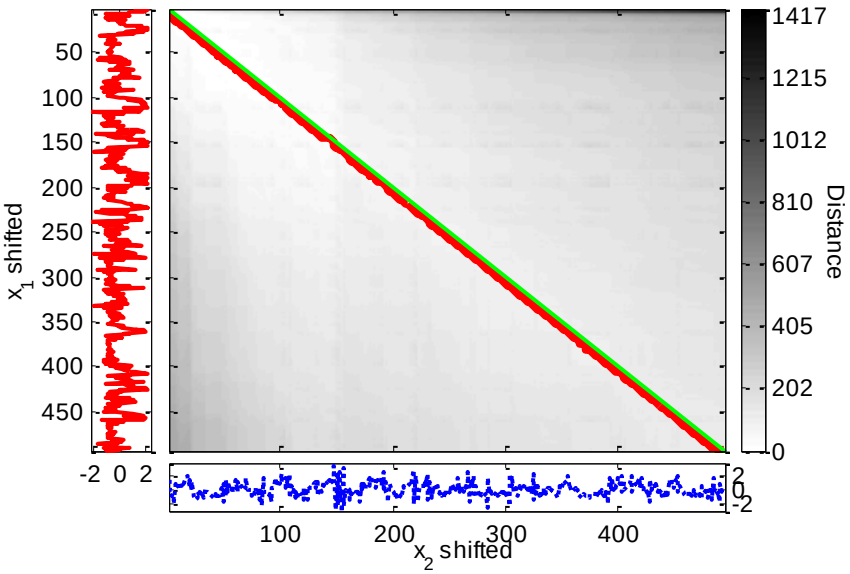

Figure 4.92. DTW distance matrix and the optimal warping path for the $x_1 - x_2$ pair of the shifted four-dimensional nonlinear time series (after x_2 is shifted to the left by six with respect to x_1).

The maximum value of correlation between the warped series (CorDTW) is obtained at lag one with a value of 0.89. The series x_2 is shifted to the left by one with respect to x_1 and the results of DTW are as follows.

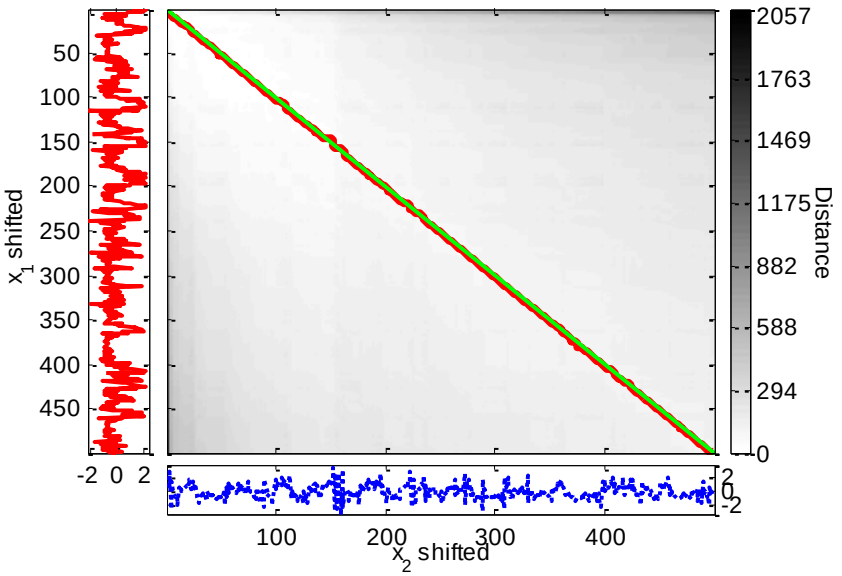

Figure 4.93. DTW distance matrix and the optimal warping path for the $x_1 - x_2$ pair of the shifted four-dimensional nonlinear time series (after x_2 is shifted to the left by one with respect to x_1).

The mode and the mean of the time shifts between the series in the warping path are two, indicating that the series x_1 lags behind the series x_2 ($x_1 \Leftarrow x_2$). The result of the unshifted series is still preserved.

The results of the tests for the pair $x_1 - x_3$ are as follows.

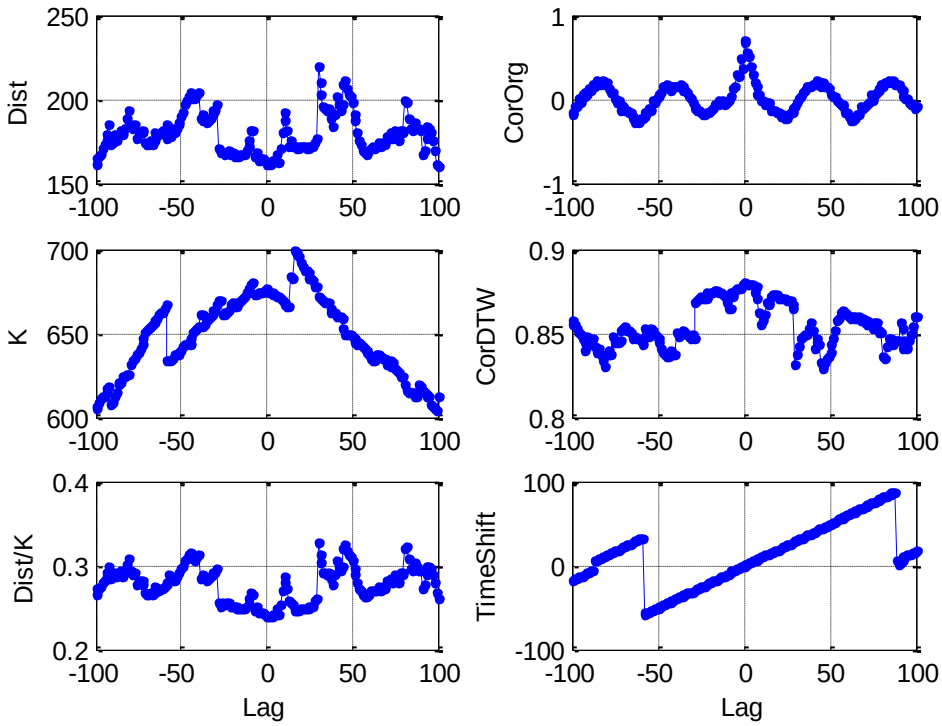

Figure 4.94. DTW results for shifted four-dimensional nonlinear time series for the $x_1 - x_3$ pair.

The maximum value of cross-correlation was observed at zero lag and by inspection of Figure 4.94, the minimum value of distance to path length ratio (Dist/K) is found at lag zero with a value of 0.24. Therefore the series are not shifted.

The maximum value of correlation between the warped series (CorDTW) is obtained at lag one with a value of 0.88. The series x_3 is shifted to the left by one with respect to x_1 and the results of DTW are as follows.

The mode and the mean of the time shifts between the series in the warping path are one, indicating that the series x_1 lags behind the series x_3 ($x_1 \Leftarrow x_3$).

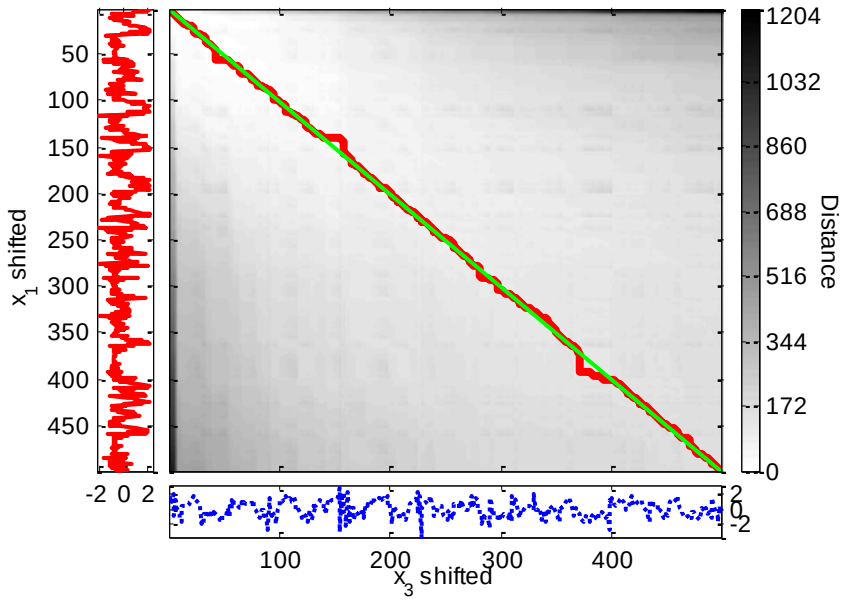

Figure 4.95. DTW distance matrix and the optimal warping path for the $x_1 - x_3$ pair of the shifted four-dimensional nonlinear time series (after x_3 is shifted to the left by one with respect to x_1).

The results of the tests for the pair $x_2 - x_3$ are as follows.

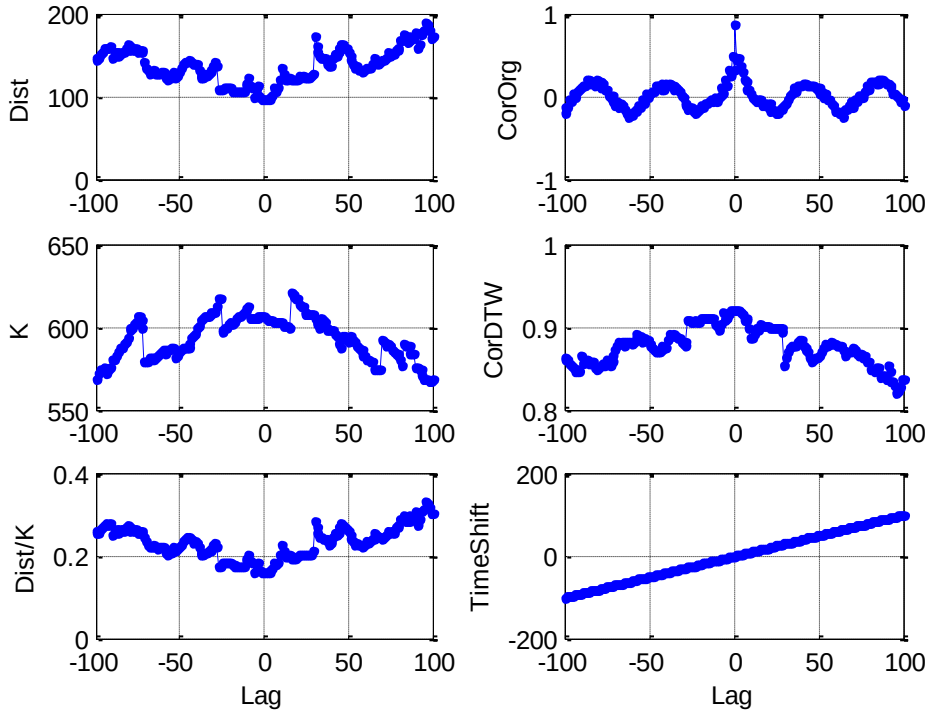

Figure 4.96. DTW results for shifted four-dimensional nonlinear time series for the $x_2 - x_3$ pair.

The maximum value of cross-correlation was observed at lag one. The series x_3 is shifted to the left by one with respect to x_2 and the results of DTW are as follows.

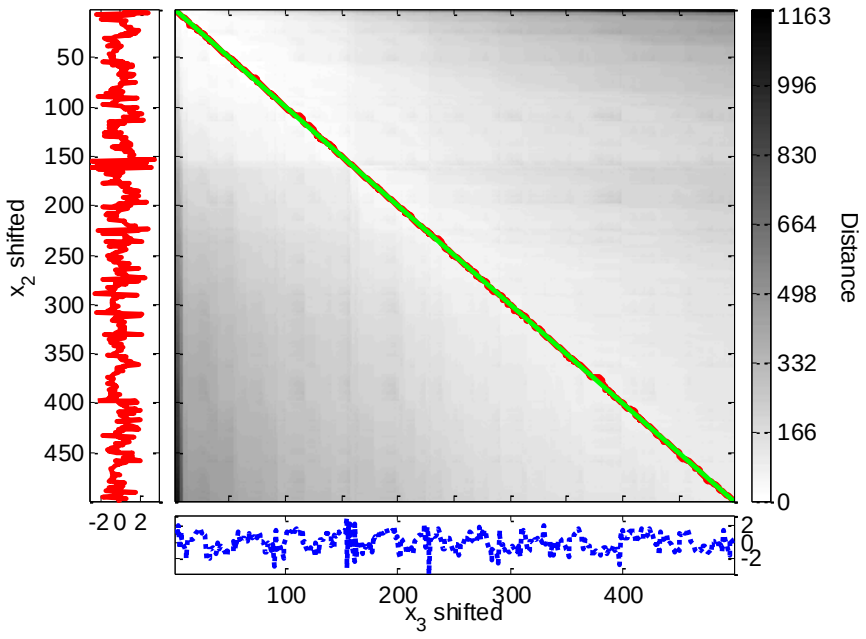

Figure 4.97. DTW distance matrix and the optimal warping path for the $x_2 - x_3$ pair of the shifted four-dimensional nonlinear time series (after x_3 is shifted to the left by one with respect to x_2).

The mode and the mean of the time shifts between the series in the warping path are zero, indicating that there is no lead/lag relation between the series x_2 and the series x_3.

By inspection of Figure 4.96, the minimum value of distance to path length ratio is (Dist/K) found at zero lag with a value of 0.16 and the maximum value of correlation between the warped series (CorDTW) is also obtained at zero lag with a value of 0.92. Therefore the series are not shifted.

The results of the tests for the pair $x_2 - x_4$ are as follows.

The maximum value of cross-correlation was at lag −1. Therefore the series x_2 is shifted to the left by one with respect to x_4 and the results of DTW are as follows.

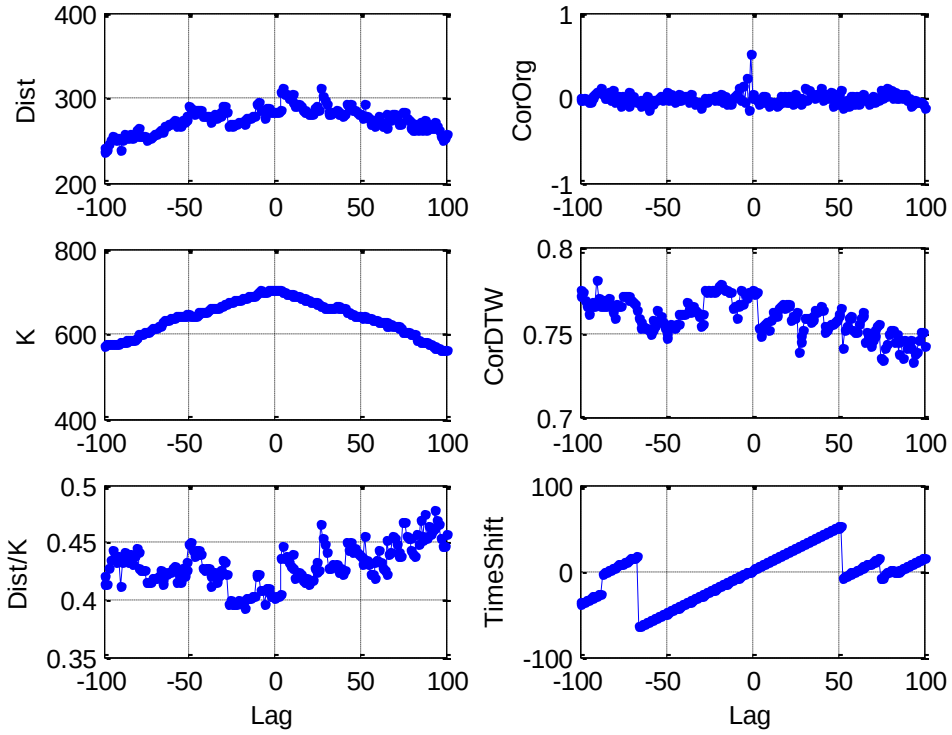

Figure 4.98. DTW results for shifted four-dimensional nonlinear time series for the $x_2 - x_4$ pair.

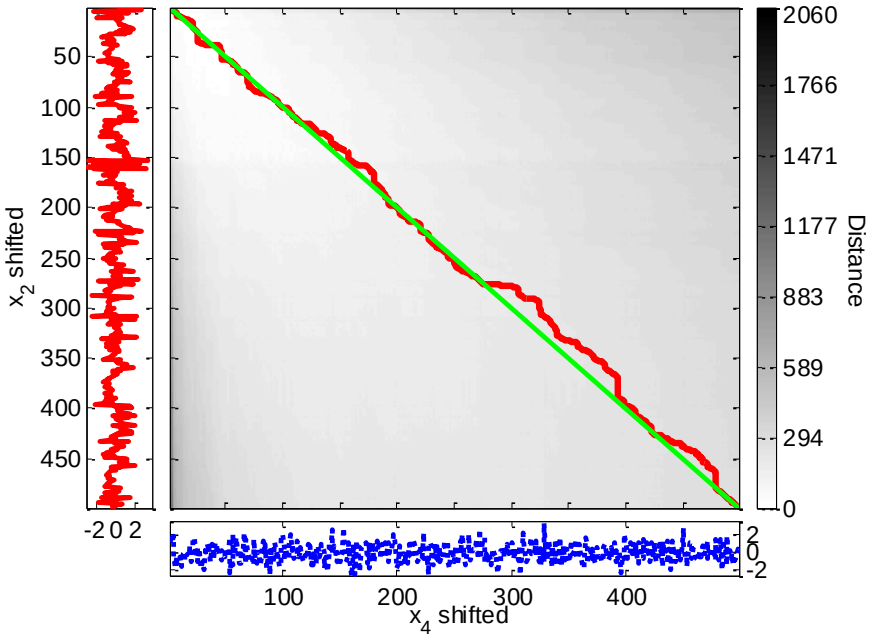

Figure 4.99. DTW distance matrix and the optimal warping path for the $x_2 - x_4$ pair of the shifted four-dimensional nonlinear time series (after x_2 is shifted to the left by one with respect to x_4).

The mode of the time shifts between the series in the warping path is zero, indicating that there is no lead/lag relation between the series and the mean of the time shifts is −7, indicating that the series x_2 leads the series x_4 ($x_2 \Rightarrow x_4$).

By inspection of Figure 4.98, the minimum value of distance to path length ratio (Dist/K) is found at lag −18 with a value of 0.39. The series x_2 is shifted to the left by 18 with respect to x_4 and the results of DTW are as follows.

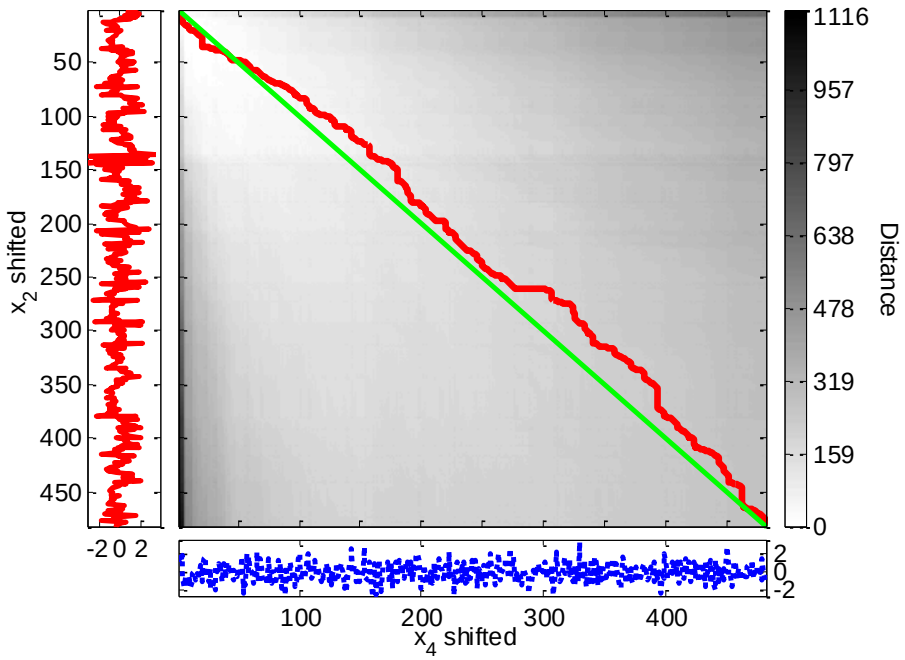

Figure 4.100. DTW distance matrix and the optimal warping path for the $x_2 – x_4$ pair of the shifted four-dimensional nonlinear time series (after x_2 is shifted to the left by 18 with respect to x_4).

The mode of the time shifts between the series in the warping path is −17 and the mean of the time shifts is −20, indicating that the series x_2 leads the series x_4 ($x_2 \Rightarrow x_4$).

For the range of [−30, +30] lags, the maximum value of correlation between the warped series (CorDTW) is again obtained at lag −18 with a value of 0.78.

When the results from Table 4.55 are compared with the GC test, DTW, cross-correlation results of the unshifted series from Table 4.31 and with the expected results from Figure 4.46, it is seen that for $x_1 – x_2$ pair, the expected result of a feedback ($x_1 \leftrightarrow x_2$)

is recovered with a shift corresponding to the maximum value of cross-correlation. Because the series affect each other with the same amount of time lag, DTW finds that none of the series leads the other (no precedence). The same results were obtained by the GC test and cross-correlation analysis on the unshifted series. For $x_1 - x_3$ pair, the only different result is found by a shift of the value corresponding to the maximum CorDTW ($x_1 \Leftarrow x_3$); however it is the opposite of the expected result of x_1 preceding x_3 ($x_1 \Rightarrow x_3$). For the case of the series $x_2 - x_3$, the expected result was obtained by the unshifted series and this also corresponds to the shifts by the minimum of Dist/K and the maximum of CorDTW since they occur at lag zero. Lastly, none of the shifted series provides the correct result that x_4 preceding x_2 ($x_2 \Leftarrow x_4$).

Table 4.55. Summary of the results for the shifted four-dimensional nonlinear time series.

Shift by	$\mathbf{x_1 - x_2}$	$\mathbf{x_1 - x_3}$	$\mathbf{x_2 - x_3}$	$\mathbf{x_2 - x_4}$
Max cross-correlation	No precedence	No precedence	No precedence	$x_2 \Rightarrow x_4$
Min Dist/K	$x_1 \Leftarrow x_2$	No precedence	$x_2 \Rightarrow x_3$	$x_2 \Rightarrow x_4$
Max CorDTW	$x_1 \Leftarrow x_2$	$x_1 \Leftarrow x_3$	$x_2 \Rightarrow x_3$	$x_2 \Rightarrow x_4$

Overall, an appropriate lag to shift the series by could not be determined. Different results are obtained for different series, therefore, they cannot be generalized. In some cases the results of the unshifted series are preserved, in the others the opposite results are obtained and at times these results agreed with what was expected and at other times they did not. However, the number of correct results is higher when the lag value is chosen to be the lag value giving the highest cross-correlation.

5. CONCLUSIONS AND RECOMMENDATIONS

In this thesis work, the aim was to investigate the applicability of the Dynamic Time Warping (DTW) method to explore time-series causality. For this purpose, DTW was first applied on data pairs from linear and nonlinear test sets. The sets were of different dimensions, with constant and varying relations between the variables. In addition, data sets from a hair-dryer device, chemical and biochemical reactions, and chemical engineering plant cost index and macroeconomic indicators were used in the applications.

The DTW results were obtained by investigating the optimal warping paths and time-shift values. The Granger Causality (GC) and cross-correlation methods were also applied on the data sets and the results were presented for a comparison with the DTW method. All methods were implemented in MATLAB.

When the synthetic test sets were examined, it was seen that for the first two data sets where the relationship between the variables changes with time, the DTW method provided the correct results. This conclusion was drawn from the graphical examination of the overall warping path over the cumulative distance matrix. These two examples successfully demonstrated the advantage of DTW in the identification of the time-varying lead/lag relations between two variables. However, these examples only consisted of bivariate models. The performance of the method was then tested on mainly multivariable, linear and nonlinear models with interacting variables, and constant lead/lag relations. In these examples, there existed a number of spurious relations due to the common (hidden) leading variables and the DTW method could not handle additional (exogenous) variables.

Moreover, in the case of feedback relations, existing in several models, DTW did not work properly. DTW was expected to show the direction of temporal precedence correctly, and, for the variables depending on one another through the same value of temporal lagging, DTW should have found no precedence. Majority of the feedback relations were present in nonlinear models and the performance of the method worsened in the applications on the data sets with nonlinear relations.

Another group of data sets was composed of chemical and biological processes and economic indicators. Out of these sets, the application on the input-output data of the hair-dryer process provided the best results: The input was found to be preceding the output as expected. The relations between the economic variables can be changing with time and DTW has found as such. However, the expected relation for the tested variables was mainly one-directional. The reason for the incorrect results may be the existence of additional common factors affecting these series, and isolating the tested pairs from these causes was not possible. Furthermore, the unsuccessful applications on the reactions data may be due to the nonlinearity of the processes. As a variation, the series were shifted, before warping, by values corresponding to several parameters related with the DTW. However, no significant improvement was gained by the application.

In these applications, the GC results were mainly correct apart from certain nonlinear problems. Moreover, spurious relations were mistaken for real causal relations as it has happened in the case of DTW. The cross-correlation results were parallel with the DTW results, except for the problems where the lead/lag relation changed with time. In light of all the applications, it was seen that DTW does not provide the same results as the GC test.

It is concluded that, DTW cannot be used for causality detection with the form used in this work. Nevertheless, the use of DTW in the discovery of temporal precedence can still be investigated.

One possible future approach can be in the way of decreasing the effect of noise. In this regard, the study of Sornette and Zhou (2005) can be taken as an example. However, their method is conceptually different from DTW.

Furthermore, there is a 3D DTW application in Wöllmer et al., (2009), and this can be tested with multidimensional models to see if there is any improvement in the case of common factors. On the other hand, the pairwise projections of the 3D warping path correspond to the 2D results, therefore the 3D view may not provide any better results.

The aim of the work was not the GC application but DTW, therefore a simple bivariate GC test was used in this thesis. However, a partial nonlinear GC test can provide more accurate results.

The GC test is a parametric test, on the other hand the DTW is a non-parametric approach. It may be possible to parameterize the warping path (e.g., using smooth quadratic polynomial parameterization). Although such a parametric warping path cannot yield the global shortest path, suboptimal parametric warping may give causality results that are closer to those of the GC test.

APPENDIX A: MATLAB PROGRAM CODES

A.1. DTW Main Code

```
echo off; clc; clear all; close all; format short g; warning off;

% --- Applications Ch4
% --- First Example with Interval-Dependent Varying Lags
t = xlsread('seriesarticle','long-gc','A2:A501');
r = xlsread('seriesarticle','long-gc','B2:B501');

st = zscore(t);
sr = zscore(r);
tf = st';
rf = sr';

% @ Lag=0 (original)
pflag = 1; Icon = 0;
[Dist,D,k,w] = Dtwfun(tf,rf,pflag,Icon);
% [Dist,D,k,w] = Dtwfun(tf,rf,pflag,Icon,CU,CL);
```

A.2. DTW Function Code

```
function [Dist,D,k,w]=Dtwfun(t,r,pflag,Icon,CU,CL)
[rows,N] = size(t);
[rows,M] = size(r);

switch Icon
    case 0,
        CL = N;
        CU = M;
    case 1,
        %if the lengths of the series are not equal, window constraint
        %should be greater than a certain value.
        a = abs(N-M)+1;
        if (CU<(a))
            disp(['Window constraint should be greater than or equal to:
',  num2str(a)]);
        end
        CL = CU;
    case 2,
        %if the lengths of the series are not equal, upper and lower
constraints
        %should be greater than a certain value.
        l = N-M+1;
        u = M-N+1;
        if (CL<(l))
            disp(['Lower window constraint should be greater than or
equal to: ',  num2str(l)])
        elseif (CU<(u))
```

```
            disp(['Upper window constraint should be greater than or
equal to: ',  num2str(u)])
        end
end

d = NaN(N,M);
D = NaN(size(d));
D = triu(D,CU)+tril(D,-CL);
[I J] = find(D==0);

for k = 1:length(I)
    d(I(k),J(k))=(t(I(k))-r(J(k)))^2;       % Squared Euclidean Distance
end

D(1,1) = d(1,1);

for k = 2:min(N,CL)
    D(k,1) = d(k,1)+D(k-1,1);
end
for k = 2:min(M,CU)
    D(1,k) = d(1,k)+D(1,k-1);
end
for n = 2:N
    for m = 2:M
        if((m-n) <= CU && (n-m) <= CL)
%        D(n,m) = d(n,m)+min([D(n-1,m),D(n-1,m-1),D(n,m-1)]);
         D(n,m) = d(n,m)+min(D(n-1,m),min(D(n-1,m-1),D(n,m-1))); %
Another 10-fold Speed-up can be achived if you use a double-MIN
construction for the Distance matrix
        end
    end
end

Dist = D(N,M);
n = N; m = M; k = 1; w = []; w(1,:) = [N,M];
while ((n+m) ~= 2)
   if (n-1) == 0
      m = m-1;
   elseif (m-1) == 0
      n = n-1;
   else
      [values,number] = min([D(n-1,m),D(n,m-1),D(n-1,m-1)]);
      switch number
      case 1
        n = n-1;
      case 2
        m = m-1;
      case 3
        n = n-1;
        m = m-1;
      end
  end
  k = k+1;
  w = cat(1,w,[n,m]);
end

w = sort(w,'ascend');

TimeShift = mode(w(:,1)-w(:,2))
```

```
TimeShift2 = mean(w(:,1)-w(:,2))

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if pflag


    % --- Warped signals
    figure('Name','Original and warped signals');

    subplot(2,1,1);
    plot(t,'r-'); grid on;  hold on;
    hold on;
    plot(r,'b:'); grid on;  hold on;
    hold off;
    grid;
    legend('signal 1','signal 2');
    title('Original signals');
    xlabel('Samples');
    ylabel('Amplitude');
    xlim([0 max(M,N)]);

    subplot(2,1,2);
    plot(t(w(:,1)),'r-'); grid on;
    hold on;
    plot(r(w(:,2)),'b:'); grid on;
    hold off;
    grid;
    legend('signal 1','signal 2');
    title('Warped signals');
    xlabel('Samples');
    ylabel('Amplitude');
    xlim([0 k]);


    % --- Accumulated distance matrix and optimal path
    figure('Name','Accumulated distance matrix and optimal path');

    main1 = subplot('position',[0.20 0.20 0.65 0.75]);
    imagesc(D);
    colormap(1-gray)
    hold on;
    x = w(:,1); y = w(:,2);
    ind = find(x==1); x(ind) = 1+0.2;
    ind = find(x==N); x(ind) = N-0.2;
    ind = find(y==1); y(ind) = 1+0.2;
    ind = find(y==M); y(ind) = M-0.2;
    plot(y,x,'-r','LineWidth',3);
    line('XData',[w(1,1) w(end,1)],'YData',[w(1,1)
w(end,1)],'Color','g','LineWidth',2)
    hold off;
    axis([1 M 1 N]);
    set(main1, 'FontSize',8, 'XTickLabel','', 'YTickLabel','');

    colorb1 = subplot('position',[0.87 0.20 0.04 0.75]);
    nticks = 8;
    ticks = floor(1:(size(1-gray,1)-1)/(nticks-1):size(1-gray,1));
    mx = max(max(D));
    mn = min(min(D));
```

```
    ticklabels = floor(mn:(mx-mn)/(nticks-1):mx);
    colorbar(colorb1);
    set(colorb1, 'FontSize',9, 'YTick',ticks, 'YTickLabel',ticklabels);
    set(get(colorb1,'YLabel'), 'String','Distance', 'Rotation',-90,
'FontSize',9, 'VerticalAlignment','bottom');

    left1 = subplot('position',[0.08 0.20 0.10 0.75]);
    plot(t,N:-1:1,'r-', 'LineWidth',2);
    set(left1, 'YTick',mod(N,10):10:N, 'YTickLabel',10*floor(N/10):-10:0)
    axis([min(t) 1.1*max(t) 1 N]);
    set(left1, 'FontSize',9);
    set(get(left1,'YLabel'), 'String','Samples', 'FontSize',9,
'Rotation',90, 'VerticalAlignment','cap');
    set(get(left1,'XLabel'), 'String',' ', 'FontSize',6,
'VerticalAlignment','cap');

    bottom1 = subplot('position',[0.20 0.08 0.65 0.10]);
    plot(r,'b:', 'LineWidth',2);
    axis([1 M min(r) 1.1*max(r)]);
    set(bottom1, 'FontSize',9, 'YAxisLocation','right');
    set(get(bottom1,'XLabel'), 'String','Samples', 'FontSize',9,
'VerticalAlignment','middle');
    set(get(bottom1,'YLabel'), 'String',' ', 'Rotation',-90,
'FontSize',6, 'VerticalAlignment','bottom');


    % --- Alignment with lines - series constant
    figure
    ht = ones(N,1);
    hr = zeros(M,1);
    plot(ht,'r','LineWidth',3); grid on;
    hold on;
    plot(hr,'b--','LineWidth',3); grid on;
    xlim([0 max(M,N)]);
    ylim([-0.1 1.1]);
    for kk = 1:k
        line('XData',[w(kk,2) w(kk,1)],'YData',[0
1],'Color','k','LineWidth',1)
    end
    legend('signal 1','signal 2');
    xlabel('Samples');
    set(gca, 'YTickLabel','');


    % --- Alignment with lines-moving series
    DeltaH = 1.8;
    t2 = t+DeltaH;
    figure
    plot(t2,'r-','LineWidth',2); grid on; hold on;
    plot(r,'b--','LineWidth',2);
    xlim([0 max(M,N)]);
    for kk = 1:k
        line('XData',[(w(kk,1)) (w(kk,2))],'YData',[t2(w(kk,1))
r(w(kk,2))],'Color','k','LineWidth',1)
    end
    legend('signal 1','signal 2');
    xlabel('Samples');
    set(gca, 'YTickLabel','');
```

```
    % --- Alignment with lines
    figure
    plot(t,'ro-'); grid on; hold on; plot(r,'b^-');
    for kk = 1:k
        line('XData',[(w(kk,1)) (w(kk,2))],'YData',[t(w(kk,1))
r(w(kk,2))],'Color','k','LineWidth',1)
    end
    legend('signal 1','signal 2');
    xlabel('Samples');
    ylabel('Amplitude');


    % --- Slopes
    for kk = 1:k
        slopew(kk) = 1/(w(kk,1)-w(kk,2));
    end
    figure('Name','Slope');
    plot(slopew,'ro-'); grid on;
    xlabel('Samples');
    ylabel('1/Slope');


    % --- 3D plot of accumulated distance matrix and optimal path
    figure('Name','Optimal path on surface plot');
    surfc(D)
    colormap(1-gray)
    axis([0 M 0 N 0 max(max(D))])
    hold on;
    for m = 1:k
        P(m,:) = [w(m,2) w(m,1) D(w(m,1),w(m,2))];
    end
    plot3(P(:,1),P(:,2),P(:,3),'r','LineWidth',5); grid on;
    xlabel('Signal 2');
    ylabel('Signal 1');
    zlabel('Distances');

    % --- Time Shift Plot
    figure('Name','Time Shift Plot');
    plot((w(:,1)-w(:,2)),'r-','LineWidth',3);
    ylabel('w - w');
    xlabel('Samples');
    xlim([0 k]);

end

% D
% w;
Dist
k

if M == N
CORORG = corr(t',r')
end
CORDTW = corr(t(w(:,1))',r(w(:,2))')

end
```

A.3. GC Test Main Code

```
echo off; clc; clear all; close all; format short g; warning off;

%--- NOTE: The code GrangerCausalityUA implements the equations in
%          SAS_Bivariate Granger Causality Test Example.pdf
% (UA - 12 May 2011)


% Eviews Data: 1st column: CS (consumption) & 2nd column: GDP
load Eviews_Chow_var.txt
DATA = Eviews_Chow_var;

whos DATA
figure
plot(DATA)
legend('signal 1','signal 2');
xlabel('Samples');

t = DATA(:,1);
r = DATA(:,2);
t = zscore(t);
r = zscore(r);
DATA = [t r];


% Pretest for the Order of Integration
% from Cointegration and Pairs Trading with Econometrics Toolbox by
Stuart Kozola
% Augmented Dickey-Fuller test for a unit root &
% Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test for stationarity

% Levels data - t:
fprintf('=== Test t for a unit root ===\n\n')
[h1,pVal1] = adftest(t,'model','ARD') % Left-tail probability

disp(' ')
disp('Null Hypothesis (H0): "Unit root exists"')
if h1 == 1
   disp('h = 1 : Reject the Null Hypothesis')
   disp('Series is stationary')
else
   disp('h = 0 : Do not Reject the Null Hypothesis')
   disp('Series contains a unit root')
end
fprintf('p-Value = %12.6f \n',pVal1);
disp(' ')

fprintf('\n=== Test t for stationarity ===\n\n')
[h0,pVal0] = kpsstest(t,'trend',false) % Right-tail probability

disp(' ')
disp('Null Hypothesis (H0): "Trend stationary"')
if h0 == 1
   disp('h = 1 : Reject the Null Hypothesis')
   disp('Series contains a unit root')
else
   disp('h = 0 : Do not Reject the Null Hypothesis')
```

```
   disp('Series is stationary')
end
fprintf('p-Value = %12.6f \n',pVal0);
disp(' ')


% Levels data - r:
fprintf('=== Test r for a unit root ===\n\n')
[h1,pVal1] = adftest(r,'model','ARD') % Left-tail probability

disp(' ')
disp('Null Hypothesis (H0): "Unit root exists"')
if h1 == 1
   disp('h = 1 : Reject the Null Hypothesis')
   disp('Series is stationary')
else
   disp('h = 0 : Do not Reject the Null Hypothesis')
   disp('Series contains a unit root')
end
fprintf('p-Value = %12.6f \n',pVal1);
disp(' ')

fprintf('\n=== Test r for stationarity ===\n\n')
[h0,pVal0] = kpsstest(r,'trend',false) % Right-tail probability

disp(' ')
disp('Null Hypothesis (H0): "Trend stationary"')
if h0 == 1
   disp('h = 1 : Reject the Null Hypothesis')
   disp('Series contains a unit root')
else
   disp('h = 0 : Do not Reject the Null Hypothesis')
   disp('Series is stationary')
end
fprintf('p-Value = %12.6f \n',pVal0);
disp(' ')

%--- Engle-Granger Cointegration Test
% (Requires Matlab's Version 2.0 (R2011a) Econometrics Toolbox)
%
%    Engle-Granger tests assess the null hypothesis of no
% cointegration among the time series in Y. The test
% regresses Y(:,1) on Y(:,2:end), then tests the residuals
% for a unit root.
%    h: Vector of Boolean decisions for the tests, with length
% equal to the number of tests. Values of h equal to 1 (true)
% indicate rejection of the null in favor of the alternative of
% cointegration. Values of h equal to 0 (false) indicate a
% failure to reject the null.
%   NULL Hypothesis, H0: h=0 No Cointegration
%                        h=1 Cointegration

%[h,pValue,stat,cValue,reg1,reg2] = egcitest(DATA)
figure
plot(DATA)
[h_EGCoint,pValue_EGCoint] = egcitest(DATA);

disp(' ')
disp('Null Hypothesis (H0): "No Cointegration"')
if h_EGCoint == 1
```

```
   disp('h = 1 : Reject the Null Hypothesis')
   disp('Cointegration Among the Time Series')
else
   disp('h = 0 : Do not Reject the Null Hypothesis')
   disp('No Cointegration Among the Time Series')
end
fprintf('p-Value = %12.6f \n',pValue_EGCoint);
disp(' ')

% figure
% plot(diff(DATA,1))
% [h_EGCoint,pValue_EGCoint] = egcitest(diff(DATA,1));
% h_EGCoint,pValue_EGCoint


%-- Perform Bivariate Granger Causality Test
%   USAGE: Granger_Cause(X, Y, alpha, Max_Lag)
%   Below: first X=f(Y), next Y=f(X)
%   where X is the 1st series (1st column of data: CS) and
%         Y is the 2nd series (2nd column of data: GDP)


Alpha = 0.95 % 95% confidence
Max_Lag = 5
[F_Stat,CrVal] = GrangerCausalityUA(DATA(:,1),DATA(:,2),Alpha,Max_Lag);
[F_Stat,CrVal] = GrangerCausalityUA(DATA(:,2),DATA(:,1),Alpha,Max_Lag);

% For this Eviews example: CS Granger Causes GDP
% i.e., Granger Casuality runs one way from CS to GDP
%       and not the other way; i.e., GDP=f(CS) is OK
%       but CS=f(GDP) is not OK. (UA - 12 May 2011)
%
%--- Eviews' Interpretation
% For this example, we cannot reject the hypothesis that GDP does
% not Granger cause CS but we do reject the hypothesis that CS
% does not Granger cause GDP. Therefore it appears that Granger
% causality runs one-way from CS to GDP and not the other way.
```

A.4. GC Test Function Code

```
function [F,c_v] = GrangerCausalityUA(x,y,alpha,max_lag)
% [F,c_v] = granger_cause(x,y,alpha,max_lag)
% Granger Causality test
% Does Y Granger Cause X?
%
% User-Specified Inputs:
%   x -- A column vector of data
%   y -- A column vector of data
%   alpha -- the significance level specified by the user
%   max_lag -- the maximum number of lags to be considered
% User-requested Output:
%   F -- The value of the F-statistic
%   c_v -- The critical value from the F-distribution
%
% The lag length selection is chosen using the Bayesian information
% Criterion
% Note that if F > c_v we reject the null hypothesis that y does not
```

```
% Granger Cause x

% Chandler Lutz, UCR 2009
% Questions/Comments: chandler.lutz@email.ucr.edu
% $Revision: 1.0.0 $  $Date: 09/30/2009 $
% $Revision: 1.0.1 $  $Date: 10/20/2009 $
% $Revision: 1.0.2 $  $Date: 03/18/2009 $

% References:
% [1] Granger, C.W.J., 1969. "Investigating causal relations by
econometric
%     models and cross-spectral methods". Econometrica 37 (3), 424-438.

% Acknowledgements:
%   I would like to thank Mads Dyrholm for his helpful comments and
%   suggestions

% GRANGER_CAUSE is a Granger Causality Test. The null hypothesis is that
% the y does not Granger Cause x. A user specifies the two series, x and
y,
% along with the significance level and the maximum number of lags to be
% considered. The function chooses the optimal lag length for x and y
based
% on the Bayesian Information Criterion. The function produces the
% F-statistic for the Granger Causality Test along with the corresponding
% critical value. We reject the null hypothesis that y does not Granger
% Cause x if the F-statistic is greater than the critical value. Type
help
% granger_cause to learn more.

%Make sure x & y are the same length
if (length(x) ~= length(y))
    error('x and y must be the same length');
end

%Make sure x is a column vector
[a,b] = size(x);
if (b>a)
    %x is a row vector -- fix this
    x = x';
end

%Make sure y is a column vector
[a,b] = size(y);
if (b>a)
    %y is a row vector -- fix this
    y = y';
end

%Make sure max_lag is >= 1
if max_lag < 1
    error('max_lag must be greater than or equal to one');
end


T = length(x);

%First find the proper model specification using the Bayesian Information
%Criterion for the number of lags of x
BIC = zeros(max_lag,1);
```

```
%Specify a matrix for the restricted RSS
RSS_R = zeros(max_lag,1);

i = 1;
while i <= max_lag
    ystar = x(i+1:T,:);
    xstar = [ones(T-i,1) zeros(T-i,i)];
    %Populate the xstar matrix with the corresponding vectors of lags
    j = 1;
    while j <= i
        xstar(:,j+1) = x(i+1-j:T-j);
        j = j+1;
    end
    %Apply the regress function. b = betahat, bint corresponds to the 95%
    %confidence intervals for the regression coefficients and r =
residuals
    [b,bint,r] = regress(ystar,xstar);

    %Find the bayesian information criterion
    BIC(i,:) = T*log(r'*r/T) + (i+1)*log(T);

    %Put the restricted residual sum of squares in the RSS_R vector
    RSS_R(i,:) = r'*r;

    i = i+1;
end
[dummy,x_lag] = min(BIC);


%First find the proper model specification using the Bayesian Information
%Criterion for the number of lags of y
BIC = zeros(max_lag,1);

%Specify a matrix for the unrestricted RSS
RSS_U = zeros(max_lag,1);

i = 1;
while i <= max_lag
    ystar = x(i+x_lag+1:T,:);
    xstar = [ones(T-(i+x_lag),1) zeros(T-(i+x_lag),x_lag+i)];
    %Populate the xstar matrix with the corresponding vectors of lags of
x
    j = 1;
    while j <= x_lag
        xstar(:,j+1) = x(i+x_lag+1-j:T-j,:);
        j = j+1;
    end
    %Populate the xstar matrix with the corresponding vectors of lags of
y
    j = 1;
    while j <= i
        xstar(:,x_lag+j+1) = y(i+x_lag+1-j:T-j,:);
        j = j+1;
    end
    %Apply the regress function. b = betahat, bint corresponds to the 95%
    %confidence intervals for the regression coefficients and r =
residuals
    [b,bint,r] = regress(ystar,xstar);
```

```
    %Find the bayesian information criterion
    BIC(i,:) = T*log(r'*r/T) + (i+1)*log(T);

    RSS_U(i,:) = r'*r;

    i = i+1;
end
[dummy,y_lag] = min(BIC);

%The numerator of the F-statistic
F_num = ((RSS_R(x_lag,:) - RSS_U(y_lag,:))/y_lag);

%The denominator of the F-statistic
F_den = RSS_U(y_lag,:)/(T-(x_lag+y_lag+1));

%The F-Statistic
F = F_num/F_den;


c_v = finv(alpha,y_lag,(T-(x_lag+y_lag+1)));


%--- UA
disp(' ')
disp(' ')
disp('USAGE: Granger_Cause(X, Y, alpha, Max_Lag)')
disp(' ')
fprintf('Number of Lags in the model of X: %3d \n',x_lag);
fprintf('Number of Lags in the model of Y: %3d \n',y_lag);
disp(' ')
disp('Null Hypothesis (H0): "Y does not Granger Cause X"')
   if F>c_v
      disp('F-Statistic > Critical Value: Reject the Null Hypothesis')
      disp('The second argument (Y) Granger Causes the first one (X)')
   else
      disp('F-Statistic <= Critical Value: Do not reject the Null
Hypothesis')
      disp('The second argument (Y) does not Granger Cause the first one
(X)')
   end
disp(' ')
fprintf('F-Statistic    = %12.6f \n',F);
fprintf('Critical Value = %12.6f \n',c_v);
```

REFERENCES

Akman, U., 2011, *ChE 555 Applied Mathematics and Modeling for Chemical Engineers Lecture Notes*.

Baasel, W. D., 1990, *Preliminary Chemical Engineering Plant Design*, 2nd edition, Springer, New York.

Bai, Z., W.-K. Wong, and B. Zhang, 2010, "Multivariate Linear and Nonlinear Causality Tests", *Mathematics and Computers in Simulation*, Vol. 81, No. 1, pp. 5-17.

Bottazzi, G., 2012, *Short Report on Oil Price History*, http://www.cafed.sssup.it/~giulio/other/oil_price/report.html, accessed at February 2012.

Coker A. K., 2007, *Ludwig's Applied Process Design for Chemical and Petrochemical Plants*, Gulf Professional Publishing, Massachusetts.

Ding, M., Y. Chen, and S. L. Bressler, 2006, "Granger Causality: Basic Theory and Application to Neuroscience", *Handbook of Time Series Analysis*, Vol. 57, pp. 451-474.

Engle, R. F., and C. W. J. Granger, 1987, "Co-Integration and Error-Correction: Representation, Estimation, and Testing", *Econometrica*, Vol. 55, No. 2, pp. 251-276.

Fogler, H. S., 2008, *Elements of Chemical Reaction Engineering*, 4th edition, Pearson Education Inc, Massachusetts.

Geweke, J. F., 1984, "Measures of Conditional Linear Dependence and Feedback Between Time Series", *Journal of the American Statistical Association*, Vol. 79, No. 388, pp. 907-915.

Giles, D., 2011, *Testing for Granger Causality*, http://davegiles.blogspot.com/2011/04/testing-for-granger-causality.html, accessed at January 2012.

Granger, C. W. J., 1969, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods", *Econometrica*, Vol. 37, No. 3, pp. 424-438.

Gross Domestic Product – GDP, http://www.investopedia.com/terms/g/gdp.asp#axzz1uJSVsCbO, accessed at February 2012.

Guo, S., A. K. Seth, K. M. Kendrick, C. Zhou, and J. Feng, 2008, "Partial Granger Causality – Eliminating Exogenous Inputs and Latent Variables", *Journal of Neuroscience Methods*, Vol. 172, No. 1, pp. 79-93.

Hoppensteadt F., 2006, *Predator-prey model*, *Scholarpedia*, Vol. 1, No. 10, p. 1563, Revision #91666, http://www.scholarpedia.org/article/Predator-prey_model, accessed at December 2011.

IHS, Inc., 2011, "Chow_var", *Granger Causality*, EViews version 7.

Industrial Production Index – IPI, http://www.investopedia.com/terms/i/ipi.asp#axzz1uJSVsCbO, accessed at February 2012.

Karimi, K., 2010, *A Brief Introduction to Temporality and Causality*, http://arxiv.org/ftp/arxiv/papers/1007/1007.2449.pdf, accessed at December 2011.

Kassidas, A., J. F. MacGregor, and P. A. Taylor, 1998, "Synchronization of Batch Trajectories Using Dynamic Time Warping", *Process Systems Engineering*, Vol. 44, No. 4, pp. 864-875.

Keogh, E., and C. A. Ratanamahatana, 2005, "Exact Indexing of Dynamic Time Warping", *Knowledge and Information Systems*, Vol. 7, No. 3, pp. 358-386.

Kispersky, T., G. J. Gutierrez, and E. Marder, 2011, "Functional Connectivity in a Rhythmic Inhibitory Circuit Using Granger Causality", *Neural Systems & Circuits*, Vol. 1, Article 9.

Krishna, R., and S. Guo, 2008, "A Partial Granger Causality Approach to Explore Causal Networks Derived From Multi-parameter Data", *CMSB '08 Proceedings of the 6th International Conference on Computational Methods in Systems Biology*, pp. 9-27.

Krishna, R., S. Nanda, A. Kulkarni, and S. Patil, 2011, "A Partial Granger Causality Based Method for Analysis of Parameter Interactions in Bioreactors", *Computers and Chemical Engineering*, Vol. 35, No. 1, pp. 121-126.

Legrand, B., C. S. Chang, S. H. Ong, S.-Y. Neo, and N. Palanisamy, 2008, "Chromosome Classification Using Dynamic Time Warping", *Pattern Recognition Letters 29*, Vol. 29, No. 3, pp. 215-222.

Li, W., M. Wang, P. Irigoyen, and P. K. Gregersen, 2006, "Inferring Causal Relationships Among Intermediate Phenotypes and Biomarkers: A Case Study of Rheumatoid Arthritis", *Bioinformatics*, Vol. 22, No. 12, pp. 1503-1507.

Lim, H.-J., and S.-H. Yoo, 2012, "Natural Gas Consumption and Economic Growth in Korea: A Causality Analysis", *Energy Sources, Part B: Economics, Planning, and Policy*, Vol. 7, No. 2, pp. 169-176.

Loh, H. P., J. Lyons, and C. W. White, III, 2002, *Process Equipment Cost Estimation Final Report*, http://www.diquima.upm.es/Evirtual/hda/docs/TQI/process%20equipment%20cost%20estimation.pdf, accessed at March 2012.

Mukhopadhyay, N. D., and S. Chatterjee, 2007, "Causality and Pathway Search in Microarray Time Series Experiment", *Bioinformatics*, Vol. 23, No. 4, pp. 442-449.

Mullen, T., 2010, *Source Information Flow Toolbox (SIFT) An Electrophysiological Information Flow Toolbox for EEGLAB Theoretical Handbook and User Manual*, Release 0.1a, ftp://sccn.ucsd.edu/pub/tim/Advanced_EEGLAB_Workshop/Session%20D%20-%20SIFT/SIFT_manual_0.1a.pdf, accessed at February 2012.

Nagarajan, R., and M. Upreti, 2010, "Granger Causality Analysis of Human Cell-Cycle Gene Expression Profiles", *Statistical Applications in Genetics and Molecular Biology*, Vol. 9, No. 1, Article 31.

Peng W., P. C. Rose and T. Sun, 2011, "Semi-Automatic System with an Iterative Learning Method for Uncovering the Leading Indicators in Business Processes", *U.S. Patent No.: 8,010,589 B2*.

Peng, W., T. Sun, P. Rose, and T. Li, 2008, "Computation and Applications of Industrial Leading Indicators to Business Process Improvement", *International Journal of Intelligent Control and Systems*, Vol. 13, No. 3, pp. 196-207.

Pravdova, V., B. Walczak, and D. L. Massart, 2002, "A comparison of Two Algorithms for Warping of Analytical Signals", *Analytica Chimica Acta*, Vol. 456, No. 1, pp. 77-92.

Ramaker, H.-J., E. N. M. van Sprang, J. A. Westerhuis, and A. K. Smilde, 2003, "Dynamic Time Warping of Spectroscopic Batch Data", *Analytica Chimica Acta*, Vol. 498, No. 1-2, pp. 133-153.

Sakoe, H., and S. Chiba, 1978, "Dynamic Programming Algorithm Optimization for Spoken Word Recognition", *IEEE Transaction on Acoustics, Speech, and Signal Processing*, Vol. 26, No. 1, pp. 43-49.

Salvador, S., and P. Chan, 2004, "FastDTW: Toward Accurate Dynamic Time Warping in Linear Time and Space", *3rd Workshop on Mining Temporal and Sequential Data*.

Seth, S., and J. C. Príncipe, 2010, "A Test of Granger Non-causality Based on Nonparametric Conditional Independence", *Proceedings of the 2010 20th International Conference on Pattern Recognition*, pp. 2620-2623.

Skanda, D., and D. Lebiedz, 2010, "An Optimal Experimental Design Approach to Model Discrimination in Dynamic Biochemical Systems", *Bioinformatics*, Vol. 26, No. 7, pp. 939-945.

Sornette, D., and W.-X. Zhou, 2005, "Non-parametric Determination of Real-time Lag Structure Between Two Time Series: The 'Optimal Thermal Causal Path' Method", *Quantitative Finance*, Vol. 5, No. 6, pp. 577-591.

Stern, D., 2011, *From Correlation to Granger Causality, Crawford School Research Paper No. 13*, http://ssrn.com/abstract=1959624, accessed at February 2011.

The MathWorks, Inc., 2007, "dryer2", *Fuzzy Logic Toolbox Nonlinear System Identification Demo*.

The MathWorks, Inc., 2012, *Econometrics Toolbox Crosscorr*, http://www.mathworks.com/help/toolbox/econ/crosscorr.html, accessed at February 2012.

Vatavuk, W. M., 2002, "Updating the CE Plant Cost Index: Changing ways of building plants are reflected as this widely used index is brought into the 21st century", *Chemical Engineering*, pp. 62-70.

Wang, N. S., 1998, *Predator*, http://terpconnect.umd.edu/~nsw/chbe250/predatorx.pdf, accessed at December 2011.

Wöllmer, M., M. Al-Hames, F. Eyben, B. Schuller, and G. Rigoll, 2009, "A Multidimensional Dynamic Time Warping Algorithm for Efficient Multimodal Fusion of Asynchronous Data Streams", *Neurocomputing*, Vol. 73, No. 1-3, pp. 366-380.