International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 5 Issue: 7 297 – 305

Reconstruction of Time Series using Optimal Ordering of ICA Components

Amr Goneid and Abear Kamel
Department of Computer Science & Engineering,
The American University in Cairo,
Cairo, Egypt
e-mail: goneid@aucegypt.edu

***Abstract* —** We investigate the application of Independent Component Analysis (ICA) and the process of optimal ordering of independent components in reconstructing time series generated by mixed independent sources. We use a modified fast neural learning ICA algorithm with a non-linearity dependent on the statistical properties of the observed time series to obtain independent components (IC's). Experimental results are presented on the reconstruction of both artificial time series and actual time series of currency exchange rates using different error measures. The area of the error profile is introduced as a minimizing parameter to obtain optimal ordered lists of IC's for the different series. We compare different error measures and different algorithms for determining optimal ordering lists. Our results support the use of an Euclidean error measure for evaluating reconstruction errors and are in favor of a method for obtaining optimal ordering lists based on minimizing the error profile between contributions of independent components in the lists and the observed time series. For the majority of the series considered, we find that quite acceptable reconstructions can be obtained with only the first few dominant IC's in the lists.

Keywords: *Independent component analysis; Independent component ordering; Data Reconstruction*

*****

I. INTRODUCTION

Independent Component Analysis (ICA) has been considered as a powerful tool in generating *independent* features in problems where sets of observed time series may be considered as results of linearly mixed instantaneous source signals [1,2]. It has been successfully applied in a wide variety of problems (see survey in [3]). In the application of ICA in time series analysis, there is great value in the reconstruction of observed data for trend discovery and forecasting [4,5]. The process requires the optimal ordering of IC's for the reconstruction process. Although this problem had some attention in the literature, the methods suggested vary considerably. For example, the components are sorted according to their non-gaussianity [6], or by selecting a subset of the components based on the mutual information between the observations and the individual components [7]. Also, in the work [8], the L_∞ norm of each individual component is used to decide on the component ordering where the order is determined based on each individual component only. More recently, component ordering is suggested to be based on component power [9]. On the other hand, the works in [4] and [10] consider the joint contributions of IC's in data reconstruction, which naturally leads the component ordering to a typical combinatorial optimization problem.

In the present paper, we investigate the process of reconstruction of observed time series using ICA. We use a modified fast neural learning ICA algorithm with a non-linearity dependent on the statistical properties of the observed time series. Experimental results are presented on the reconstruction of both artificial time series and actual time series of currency exchange rates using different error measures. The area of the error profile is introduced as a minimizing parameter to obtain optimal ordered lists of IC's for the different series. We compare different error measures and different algorithms for determining optimal ordering lists.

The paper is organized as follows: section II introduces the ICA of time series and the modified fast ICA algorithm used in the present work; section III describes the process of reconstruction of observed time series; section IV presents the different methods for the determination of optimal ordering lists; sections V and VI give results of experimentation; and finally section VII presents the summary and conclusion of our work.

II. ICA OF TIME SERIES

A. The ICA Model

Consider the observed k time series $\boldsymbol{X} = x(t) = [x_1(t), \ldots x_k(t)]^T$, $1 \le t \le N$ to be the instantaneous linear mixture of unknown statistically independent components $\boldsymbol{Y} = y(t) = [y_1(t), \ldots . y_k(t)]^T$, i.e., $\boldsymbol{X} = \boldsymbol{A}\ \boldsymbol{Y}$, where $\boldsymbol{A}$ is an unknown $k \times k$ nonsingular mixing matrix. Given $\boldsymbol{X}$, the basic ICA problem is to find an estimate $\hat{\boldsymbol{y}}$ of $\boldsymbol{Y}$ and the mixing matrix $\boldsymbol{A}$ such that $\hat{\boldsymbol{y}} = \boldsymbol{W}\ \boldsymbol{X} = \boldsymbol{W}\ \boldsymbol{A}\ \boldsymbol{Y} = \boldsymbol{G}\ \boldsymbol{Y} \approx \boldsymbol{Y}$, where $\boldsymbol{W} = \boldsymbol{A}^{-1}$ is the *unmixing* matrix, and $\boldsymbol{G} = \boldsymbol{W}\ \boldsymbol{A}$ is usually called the Global Transfer Function or Global Separating-Mixing (GSM) Matrix. The linear mapping $\boldsymbol{W}$ is such that the unmixed signals $\hat{\boldsymbol{y}}$ are statistically independent. However, the sources are recovered only up to scaling and permutation. In practice, the estimate of the unmixing matrix $\boldsymbol{W}$ is not exactly the inverse of the mixing matrix $\boldsymbol{A}$. Hence, the departure of $\boldsymbol{G}$ from the

identity matrix ***I*** can be a measure of the error in achieving complete separation of sources.

B. The Neural Learning ICA Algorithm

For computing the independent components (IC's) from the observed time series, we have adopted the modified algorithm used before in [10], which is based on the Fast ICA algorithm originally given by [11]. Basically, the algorithm uses a fixed-point iteration method to maximize the negentropy using a Newton iteration method. We assume that the observation matrix ***X*** of *k* time series and *N* samples has been preprocessed by centering followed by whitening or sphering to remove correlations. Centering removes means via the transformation $X \leftarrow X - E\{X\}$ and whitening is done using a linear transform (PCA like) $Z = VX$, where ***V*** is a whitening matrix. A popular whitening matrix is $V = D^{-1/2} E^T$, where ***E*** and ***D*** are the eigenvector and eigenvalue matrices of the covariance matrix of ***X***, respectively. The resulting new matrix ***Z*** is therefore characterized by $E\{ZZ^T\} = I$ and $E\{Z\} = 0$. After obtaining the unmixing matrix ***W*** from whitened data, the total unmixing matrix is then $W \leftarrow W V$. The algorithm estimates several or all components in parallel using symmetric orthogonalization by setting $W \leftarrow (W W^T)^{-1/2} W$ in each iteration.

In the modified version [10] of the algorithm, the performance during the iterative learning process is measured using the matrix $G = W A$, which is supposed to converge to a permutation of the scaled identity matrix at complete separation of the IC's. This is done by decomposing $G = Q P$, where ***P*** is a positive definite *stretching* matrix and ***Q*** is an orthogonal *rotational* matrix. The cosine of the rotation angle is to be found on the diagonal of ***Q*** so that a convergence criterion is taken as $\Delta\,|diag(Q)|_{min} < \varepsilon$, where ε is a threshold value. This leads to the normalized performance (error) measure, *E3* introduced in [10] as:

$$E3 = \frac{e3}{2k(k-1)} \qquad \{0,1\} \tag{1}$$

with,

$$e3 = \sum_{i=1}^{k}\{\sum_{j=1}^{k}|g_{ij}| - M_i + |M_i - 1|\} + \sum_{j=1}^{k}\{\sum_{i=1}^{k}|g_{ij}| - M_j + |M_j - 1|\} \tag{2}$$

where g_{ij} is the ij^{th} element of the matrix ***G*** of dimensions *k* x *k*, $M_i = max_k\,|\,g_{ik}\,|$ is the absolute value of the maximum element in row (*i*) and $M_j = max_k\,|\,g_{kj}\,|$ is the corresponding quantity for column (*j*).

The algorithm is summarized in the following steps:

- Preprocess observation matrix ***X*** to get ***Z***
- Choose random initial orthonormal vectors w_i to form intial ***W*** and random ***A***
- Set $W_{old} \leftarrow W$

Iterate:

1. Do Symmetric orthogonalization of ***W*** by setting $W \leftarrow (W W^T)^{-1/2} W$
2. Compute dewhitened matrix ***A*** and new $G = W A$ and do polar decomposition of $G = Q P$
3. Compute error *E3*
4. If not the first iteration, test for convergence: $\Delta\,|\,diag(Q)\,|_{min} < \varepsilon$
5. If converged, break.
6. Set $W_{old} \leftarrow W$
7. For each component w_i of ***W,*** update using learning rule $w_i \leftarrow E\{z\, f\,(w_i\, z)\} - E\{f'\,(w_i\, z)\}\, w_i$

- After convergence, dewhiten using $W \leftarrow W V$
- Compute independent components $\hat{y} = W X$

In the update step, ***z*** is a column vector of one sample from ***Z,*** and the expectation is the average over the *N* samples in ***Z.***

C. Choice of Non-linearity

In the above neural learning algorithm, a non-linearity *f(y)* (and its derivative *f '(y))* is essential in the optimization process and for the learning rule that updates the estimates of the unmixing matrix ***W*** and, overall, it is important for the stability and robustness of the convergence process. It is common to use non-linearities *f (x)* that are derived from assumed source models such that $f(x) = -(\partial p(x)/\partial x) / p(x)$, where *p(x)* is the PDF of the source. For super-gaussian sources, the PDF *p(x) = 1/cosh(x)* leads to the general purpose non-linearity $\psi_{\pm}(x,a,b,\mu) = C\exp\{-\rho\,|x \mp \mu|^b\} = C\,\phi_{\pm}\ ,\ \rho = a^{-b}$ *f (x) = tanh(x).*

For sub-gaussian sources, it was shown in [12] that the PDF can be represented by a bimodal Exponential Power Distibution (EPD) symmetric mixture density

$$p(x,a,b,\mu) = (1/2)\{\psi_{+}(x,a,b,\mu) + \psi_{-}(x,a,b,\mu)\},\ \ \mu > 0 \tag{3}$$

where,

(4)

The function (4) is the single EPD with scale parameter (*a*), a shape parameter (*b*) and location parameter μ. The above symmetric density leads to a non-linearity

$$f(x) = -\frac{\partial p(x)/\partial x}{p(x)} = \frac{\eta_{+} + \eta_{-}}{\phi_{+} + \phi_{-}} \tag{5}$$

where,

$$\eta_{\pm} = \rho\, b\,|\,x \mp \mu\,|^{b-1}\,\phi_{\pm}\ sign\,(x \mp \mu),\quad b > 1 \tag{6}$$

Considering the stability profiles of such mixture model, the work in [12] has given for sub-gaussian sources the following set of parameters for the non-linearity given by equation (3):

$a = 3, b = 2, \mu = 3$

III. RECONSTRUCTION OF OBSERVED TIME SERIES USING INDEPENDENT COMPONENTS

A. Contribution of Independent Components to Observed Time Series

Assume we have a list L_i of independent components indices expressing a specific component ordering. The process of reconstructing time series x_i from the estimated independent components Y_j, $j = 1...k$ can be done by summing their contributions in the order given by the list L_i. Following [4], [10], the contribution may be expressed by the 3-D space:

$$u(i, j, t) = W^{-1}(i, j)Y_j(t), \quad 1 \le j \le k \tag{7}$$

where $W^{-1}(i,j)$ is the $(i,j)th$ element in the inverse of the W matrix. The reconstructed series of x_i by the first m independent components in the list L_i is given by the sum of the contributions of the individual component, i.e.,

(8)

where (s) denotes the s[th] element of L_i.

B. Reconstruction Error Profile

The reconstruction error for x_i can be computed under a certain error measure (e.g. Euclidean Distance or MSE, Average Mutual Information (AMI), etc). At a given time (t) in the time series x_i, this error may be denoted by the quantity $q(x_i(t), \hat{x}_i^m(L_i, t))$. For the whole series, the average is given by:

$$Q(m) = Q(x_i, \hat{x}_i^m(L_i)) = \underset{1 \le t \le N}{aver} \left\{ q(x_i(t), \hat{x}_i^m(L_i, t)) \right\} \tag{9}$$

The dependence of the above quantity on m represents the error profile for the reconstruction of series x_i using list L_i. Naturally, it will also depend on the error measure used.

C. Error Profile Area

The error profile $Q(m)$ depends on the list L_i and the used error measure. For comparison purposes, we introduce the error profile area as a collective error measure for a given profile. From observations, we find that typical error profiles exhibit an approximate exponential decrease with increasing m. If we express the error as $Q(m) = C \exp(-\lambda m)$, $m > 0$, then the area under the error profile will be

$$A(L_i) = \int_{m=1}^{\infty} Q(m)\, dm = (C/\lambda_i)\exp(-\lambda_i) \tag{10}$$

In practice, we compute the area as

$$A(L_i) = \sum_{m=1}^{k-1} \{ \frac{1}{2} | Q(m) - Q(m+1) | + \min(Q(m), Q(m+1)) \} \tag{11}$$

The error profile area is therefore a sensitive measure of IC's order given by the list L_i. Using such area, an optimal ordering list may be obtained as:

$$L_i^{opt} = \arg\min_{L_i}(A(L_i)) \tag{12}$$

IV. DETERMINATION OF OPTIMAL ORDERING LISTS

We investigate five different methods for determining the optimal list for the reconstruction of an observed series x_i. These methods are summarized as follows:

A. Exhaustive Search Method (ES)

A list L_i of k indices of IC's has $k!$ permutations. Exhaustive search will examine the error profile for each of these permutations for a given observed series x_i. Based on a given error measure, the resulting $k!$ error profile areas $A(L_i)$ are then used to find L_i^{opt} as given by (12). This method involves $k!$ reconstruction steps.

B. The L∞ Norm Method (LN)

The contribution of independent component (j) to the reconstruction of observed series x_i is given by the quantity $u(i,j,t)$. In this method, the $L\infty$ norm of each of these individual components is used to decide on the component ordering. The optimal list represents the component indices sorted in descending order of their $L\infty$ norm.

C. Exclude the Least Contributing IC First (EL)

This is basically the Testing-and Acceptance (TnA) method given by [4]. The method first selects from the set of k IC's the component that when excluded from the list will minimize the reconstruction error. This component is then removed from the set of IC's and its index becomes the last in the order list L_i. The process is repeated on what remains in the component set to select the second-last in the order list, and so on. The algorithm for this method is given in the following:

Algorithm Exclude-Least-First

L_i = { } is an empty list; Z = set of indices 1: k;
r = 1 is the reverse order of the least contributing IC;
While (r ≤ k)
  For all indices j remaining in Z
    Find the sum of contributions excluding component (j),

$$C_{ij}(t) = \sum_{m \in Z, m \ne j} u(i, m, t)$$

    Find the error ε_{ij} of reconstructing x_i using C_{ij};
  End For
  α = index (j) for which ε_{ij} is minimum = index of IC of least contribution;
  L_i (r) ← α;
  Remove α from set Z;
  r ← r +1;
End While

$L_i \leftarrow$ Reverse (L_i);

This algorithm will involve $k(k+1)/2-1$ reconstruction steps.

D. Maximizing the Average Mutual Information (AMI)

Average mutual information (AMI) measures the dependence between pairs of random variables. It can be used in selecting a subset of the components based on the mutual information between the observations and the individual components [7]. In the present work, an optimal list may be obtained by sorting in ascending order the AMI between individual contributions *u(i,j,t)* and the observed series x_i. This method involves only *k* reconstruction steps.

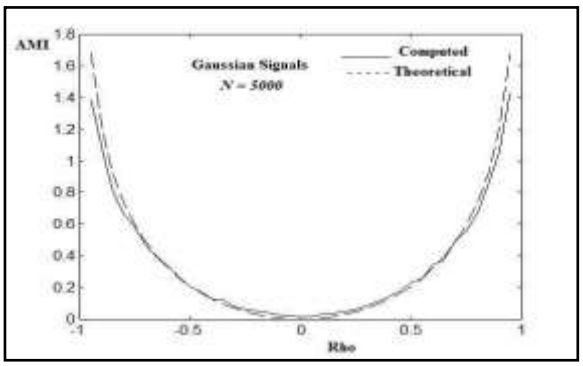

Figure 1. Dependence of AMI on ρ for bivariate Gaussian series (x,y)

Several algorithms exist for computing the AMI between two series. [13]. A fast implementation given in [14] will be used here. In order to test the AMI algorithm, it is well known [e.g. 15] that if (x,y) are bivariate normal, then the AMI between x and y depends only on the correlation coefficient ρ between them. Specifically in this case,

$$AMI(x,y) = -(1/2)\log_2(1-\rho^2)$$

We have generated two random Gaussian series *x* and *y* with a correlation coefficient between them in the range -0.95 ≤ ρ ≤ +0.95 in steps of 0.05. The AMI obtained from the algorithm is compared with the theoretical values as shown in Fig. 1 for series length of N = 5000. The validity of the algorithm is evident from the comparison of computed and theoretical AMI values.

E. Minimizing the Reconstruction Error for Individual IC's Contribution (ME)

Given a certain error measure, the list is obtained by sorting in ascending order the error between individual contributions *u(i,j,t)* and the observed series x_i. This method involves only *k* reconstruction steps.

V. EXPERIMENTS WITH ARTIFICIALLY GENERATED DATA

A. ICA of Artificially Generated Time Series

In order to validate the present methodology and to compare between different methods for determining reconstruction lists, we have used *k* = 6 artificially generated time series, each of length N = 2000. The generated series ***S*** were mixed by a random mixing matrix ***A*** to obtain the simulated observed series ***X*** = ***A S***. Fig. 2 shows portions of the source signals **S**, the mixed signals ***X*** and the IC's obtained from the fast ICA algorithm given in Section II.B. For choosing the appropriate non-linearity, we have computed the normalized kurtosis *K* for the mixed series ***X*** as given by $K = E\{x^4\} / (E\{x^2\})^2 - 3$. with the results shown in Table (1).

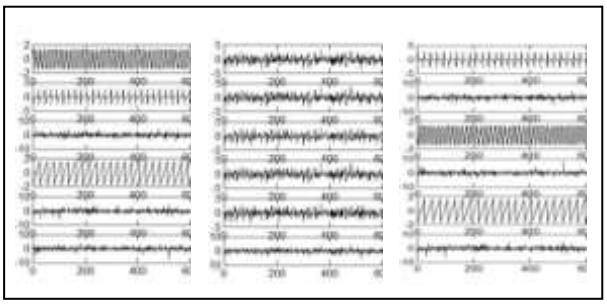

Figure 2. Artificial source signals, mixed series and obtained IC's

TABLE 1. Kurtosis of mixed series

X Series	1	2	3	4	5	6
Kurtosis	0.156	0.275	0.441	0.182	0.365	1.262

Table (1) shows that the mixed signals are all super-gaussian ($K > 0$) and a suitable non-linearity would be the general purpose one $f(x) = tanh(x)$.

B. Error Measures

For evaluating reconstruction errors, we have selected 3 error measures as follows:

1. Relative Hamming Distance (RHD)

The Q-measure using RHD is given as:

$$Q(x_i, \hat{x}_{L_i}^m) = RHD(x_i, \hat{x}_{L_i}^m) = \\ = \frac{1}{N-1}\sum_{t=1}^{N-1}[R_i(t) - \hat{R}_{L_i}^m(t)]^2 \qquad (13)$$

where,

$$R_i(t) = sign[x_i(t+1) - x_i(t)]\,,$$

$$\hat{R}_{L_i}^m(t) = sign[\hat{x}_{L_i}^m(t+1) - \hat{x}_{L_i}^m(t)]$$

2. Euclidean Distance or Mean Square Error (MSE)
The Q-measure using MSE is given as:

$$Q(x_i, \hat{x}_{L_i}^m) = MSE(x_i, \hat{x}_{L_i}^m) = \frac{1}{N}\sum_{t=1}^{N}[\hat{x}_{L_i}^m(t) - x_i(t)]^2 \quad (14)$$

3. AMI Distance (AMID)
The Average Mutual Information (AMI) may be used as an error measure equivalent to the distance between AMI (max) and AMI (x,y). In the present work, we use AMI Q-measure in the form:

$$Q(x_i, \hat{x}_{L_i}^m) = AMID(x_i, \hat{x}_{L_i}^m) = AMI(x_i, x_i) - AMI((x_i, \hat{x}_{L_i}^m) \quad (15)$$

C. Results for Series Reconstruction

In order to reconstruct observed series ***X*** from the obtained IC's, we have to first determine the error measure most appropriate for such process. We have derived trial ordered lists of IC indices using the method of minimizing the MSE and computed the error profiles for the 3 different error measures (RHD, MSE, and AMID). Fig. 3a shows an example for x_1 of such profiles representing the dependence of the reconstruction error *Q(m)* on the number of contributing IC's (*m*) taken in order from the ordered list.

It can be seen from the figure that the MSE measure gives the desired lowest profile. Such result is also found in the reconstruction of all other series ***X***. This result is also evident from computations of the error profile area as shown in Fig. 3b where *A(L)* is plotted against the indices of the observed series ***X***. From these results, it is possible to conclude that the MSE error measure is superior to RHD and AMID for the evaluation of reconstruction error.

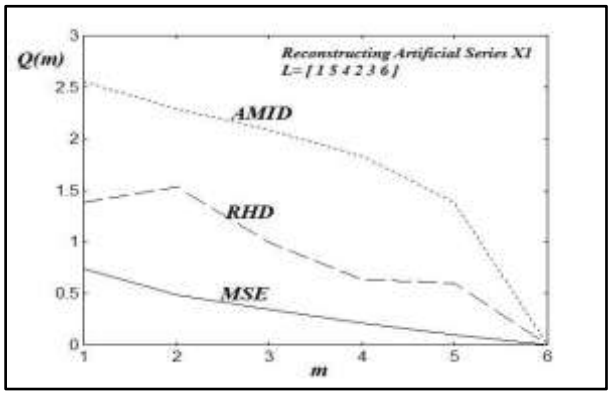

Figure 3a. Example error profiles

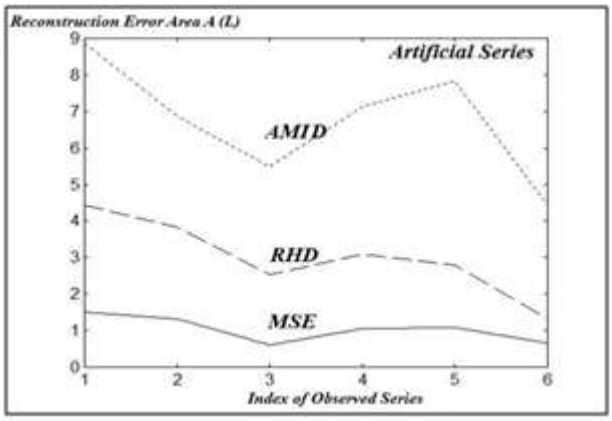

Figure 3b. Reconstruction error area

Turning now to the problem of determining optimal lists for reconstruction, we have made calculations of the error profiles and error profile areas for the 5 methods given in Section IV using MSE as the error measure. Fig. 4 shows a comparison of the different methods for the observed artificial series ***X***.

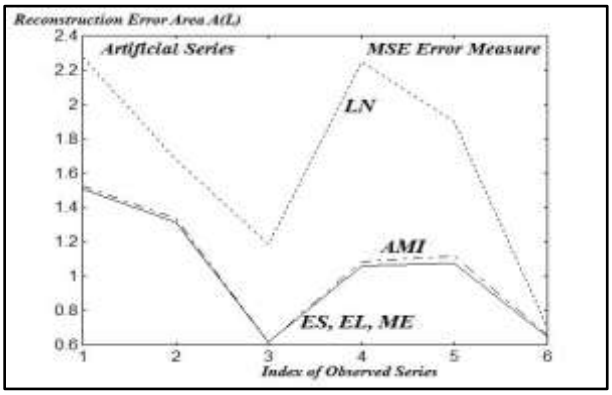

Figure 4. Reconstruction error area for observed series

It can be seen from this figure that the $L\infty$ norm method (LN**)** cannot compete with the other methods. It is also clear that the three methods of Exhaustive Search (ES), Exclude Least (EL), and Minimizing MSE (ME) give exactly the same results and are superior to all other methods, including maximizing AMI.

Table (2) gives an example of the optimal lists obtained for the reconstruction of observed series x_1 again highlighting the conclusions drawn from Fig. 4. In particular, the ES, EL, and ME (with MSE measure) methods give exactly similar lists.

TABLE 2. Optimal IC lists for observed series x_1

Method	List					
ES	1	5	4	2	3	6
LN	2	4	6	1	5	3
EL	1	5	4	2	3	6
AMI	1	5	4	3	2	6
ME (MSE)	1	5	4	2	3	6

On the other hand, the AMI method gives slightly different lists resulting in slightly higher reconstruction error, while the lists obtained by the LN method cannot be judged to be optimal and therefore produce much higher reconstruction errors.

Although tables like Table (2) for the other 5 series are not given here, it can be shown that similar conclusions apply to all the 6 observed series ***X***.

D. Obtained Reconstruction Ordered Lists

There is an exact match between the ordered lists obtained by the 3 methods of ES, EL and ME using the MSE error measure and all give the same minimum reconstruction error. Recall that the number of reconstruction steps required by these three methods are: $k!$, $k(k+1)/2 - 1$, and k respectively, where k is the number of observed time series. For our case of $k = 6$, the number of reconstruction steps are 720, 20 and 6, respectively. Hence, we may consider the ME (MSE) method to be of least complexity. Using this last method, we obtain the following set of ordered lists for the 6 artificially generated time series:

$L_1 = [1\ 5\ 4\ 2\ 3\ 6]$ $L_2 = [1\ 5\ 2\ 3\ 4\ 6]$
$L_3 = [1\ 3\ 6\ 5\ 2\ 4]$ $L_4 = [3\ 2\ 1\ 5\ 4\ 6]$
$L_5 = [1\ 6\ 3\ 4\ 2\ 5]$ $L_6 = [2\ 4\ 1\ 5\ 6\ 3]$

Note that exactly the same sets are also obtained by the Exhaustive Search (ES) and the Exclude Least (EL) methods.

VI. EXPERIMENTS WITH FINANCIAL DATA

A. Exchange Rate Time Series Dataset

As a second set of experiments, actual time series of 6 foreign exchange rate series were selected representing USD versus Brazilian Real (BRL), Canadian Dollar (CAD), Danish Krone (DKK), Japanese Yen (JPY), Swedish Krona (SEK), and Swiss Franc (CHF) in the period from January 4, 2010 till December 31, 2015. The dataset size was 6 time series over 1504 days collected from different historical exchange rates data sources such as [16, 17, 18]. Fig. 5 shows these time series.

In order to choose the appropriate non-linearity for the fast ICA algorithm given in Section II.B., we have computed the normalized kurtosis K for the above series ***X*** with the results shown in Table (3). The observed series represent a mixture between super-gaussian ($K > 0$) and sub-gaussian signals ($K < 0$). We found that our fast ICA algorithm would converge in fewer number of iterations if we use the mixture non-linearity given by (5) rather than the general purpose non-linearity $f(x) = tanh(x)$.

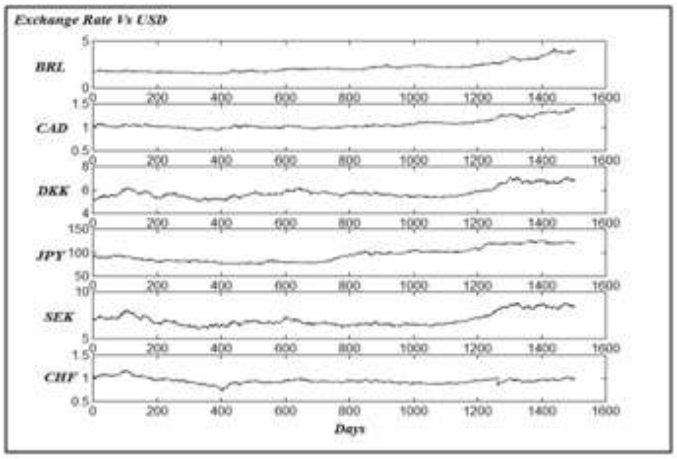

Figure 5. Exchange rate time series

TABLE 3. Kurtosis of observed series

X Series	1	2	3	4	5	6
Kurtosis	1.491	0.867	0.330	- 1.004	- 0.082	1.628

B. Results for Series Reconstruction

Using the IC's obtained from the fast ICA algorithm, we have derived trial ordered lists of IC indices using the method of minimizing the MSE and computed the error profiles for the 3 different error measures (RHD, MSE, and AMID). Fig.6a shows an example for $x_5(SEK)$ of such profiles representing the dependence of the reconstruction error $Q(m)$ on the number of contributing IC's (m) taken in order from the ordered list.

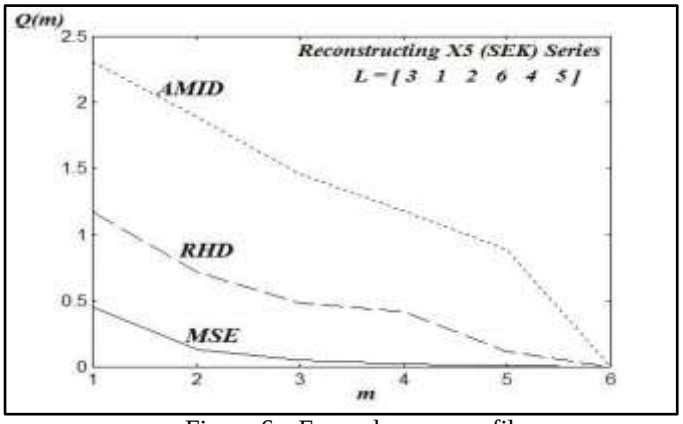

Figure 6a. Example error profiles

This result is also evident from computations of the error profile area as shown in Fig.6b where $A(L)$ is plotted against the indices of the observed series X. It can be seen from that figure that the MSE measure also gives the desired lowest profile for the financial series. Such result is also found in the reconstruction of all other series ***X***.

 ISSN: 2321-8169
297 – 305

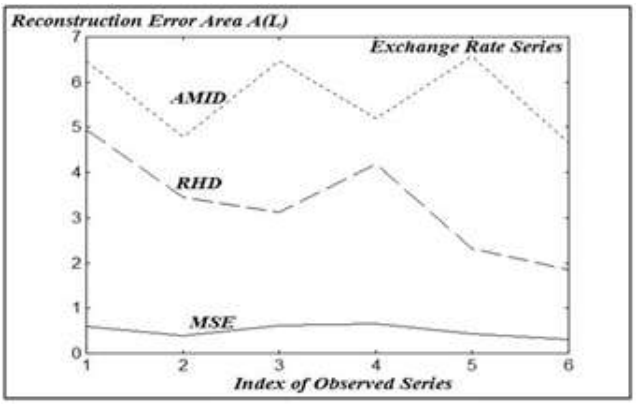

Figure 6b. Reconstruction error area

For determining optimal lists for reconstruction, we have made calculations of the error profiles and error profile areas for the 5 methods given in Section IV using MSE as the error measure. Fig. 7 shows a comparison of the different methods for the observed artificial series ***X***

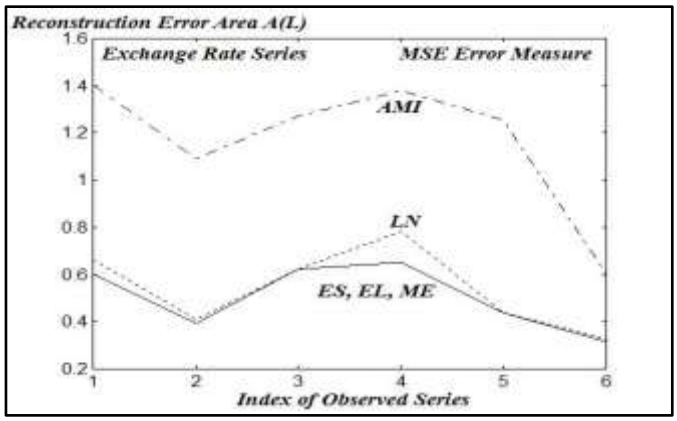

Figure 7. Reconstruction error area for observed Exchange rate series ***X:*** Comparison of Methods for obtaining optimal lists

It can be seen from Fig. 7 that the three methods of Exhaustive Search (ES), Exclude Least (EL), and Minimizing MSE (ME) give exactly the same results and are superior to all other methods, a result that is in accordance with our findings for artificially generated series.

Table (4) gives an example of the optimal lists obtained for the reconstruction of observed series x_2 (CAD) again highlighting the conclusions drawn from Fig. 7. In particular, the ES, EL, and ME (with MSE measure) methods give exactly similar lists.

It can also be seen from Table (4) that the LN method gives slightly different lists resulting in slightly higher reconstruction error, while the lists obtained by the AMI method produce much higher reconstruction errors. Although tables like Table (4) for the other 5 series are not given here, it can be shown that similar conclusions apply to all the 6 exchange rate series.

TABLE 4. Optimal IC lists for exchange rate series x_2 *(CAD)*

Method	List					
ES	1	3	5	6	2	4
LN	1	3	6	5	2	4
EL	1	3	5	6	2	4
AMI	5	3	1	2	6	4
ME	1	3	5	6	2	4

C. Results for Ordered Lists and Contributions of IC's

Although the ES, EL, and ME methods give exactly similar lists, the ME method is superior since it involves only $k = 6$ reconstruction steps. Using this method, we obtain the set of ordered lists for the 6 exchange rate time series ***X***. Table (5) shows the obtained ordered lists.

TABLE 5. Obtained ordered lists

Series	Ordered List					
BRL	1	3	5	4	6	2
CAD	1	3	5	6	2	4
DKK	3	1	6	4	5	2
JPY	3	5	1	6	4	2
SEK	3	1	2	6	4	5
CHF	2	6	5	1	4	3

Table (6) gives the percentage cumulative contribution of the first m IC's from the lists to the reconstruction of the exchange rate time series. In this table, the contributions are calculated as 1 -$MSE\left(x_i, \hat{x}_{L_i}^m\right)$ / $variance(x_i)$ = 1- $Q(m)$, since the series have zero mean and unit variance.

TABLE 6. Cumulative Contributions of IC's (%)

Series	m=1	m=2	m=3	m=4	m=5	m=6
BRL	58.1	81.6	88.5	93.3	97.3	100.0
CAD	49.3	90.3	95.8	99.8	100.0	100.0
DKK	49.5	77.4	88.0	97.7	100.0	100.0
JPY	40.8	73.5	93.4	97.5	100.0	100.0
SEK	54.6	86.8	95.2	97.6	99.2	100.0
CHF	64.1	93.1	96.0	98.0	99.3	100.0

The results given in Table (6) indicate that most of the observed exchange rate series considered here can be reconstructed to an excellent degree using the first 4 IC's in their respective ordered lists (contributions > 97 %). For the majority of series, quite acceptable reconstructions (> 93 %) can also be obtained with only the first 3 IC's in the lists

D. Comparison between Reconstructed and Observed Series

As an example, Fig. 8 compares the observed CAD/USA series with the reconstructed ones using the first one, the first two and first 3 IC's in their ordered lists.

 ISSN: 2321-8169
297 – 305

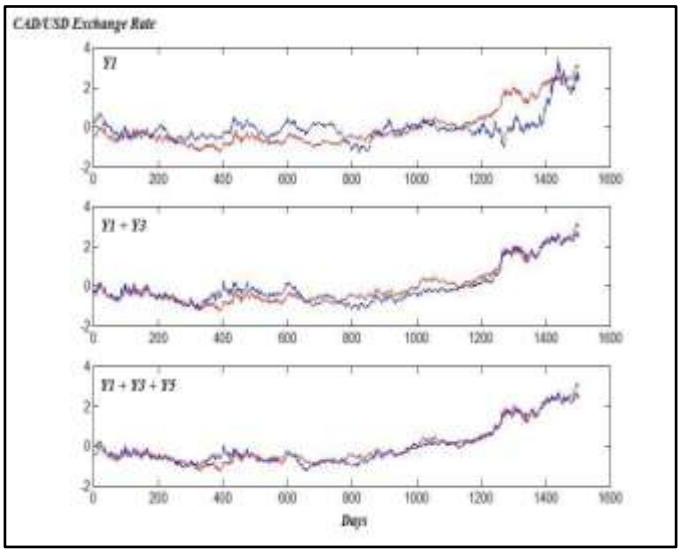

Figure 8. Reconstruction of exchange rate time series X_2 *(*CAD), Red (Observed), Blue (constructed)

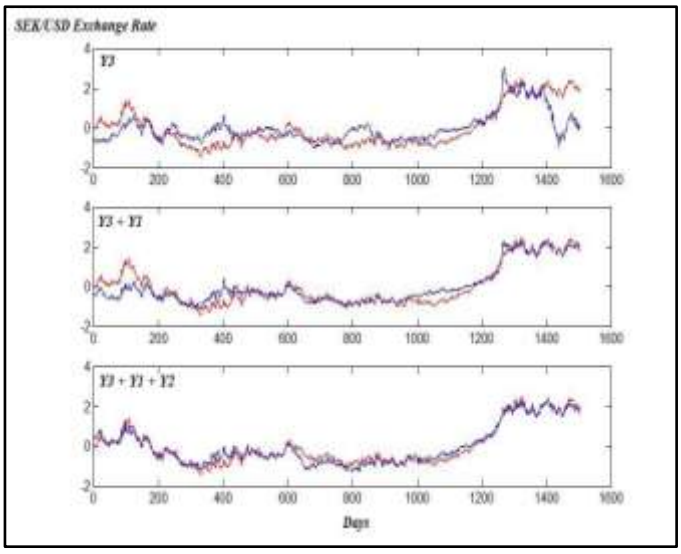

Figure 9. Reconstruction of exchange rate time series X_5 *(*SEK), Red (Observed), Blue (constructed)

A second example is shown in Fig. 9 for the SEK/USD series. Notice that the series compared in these figures have zero mean and unit variance as obtained from the preprocessing of the data for the ICA algorithm. A third example is also shown in Fig. 10 for the CHF/USD series.

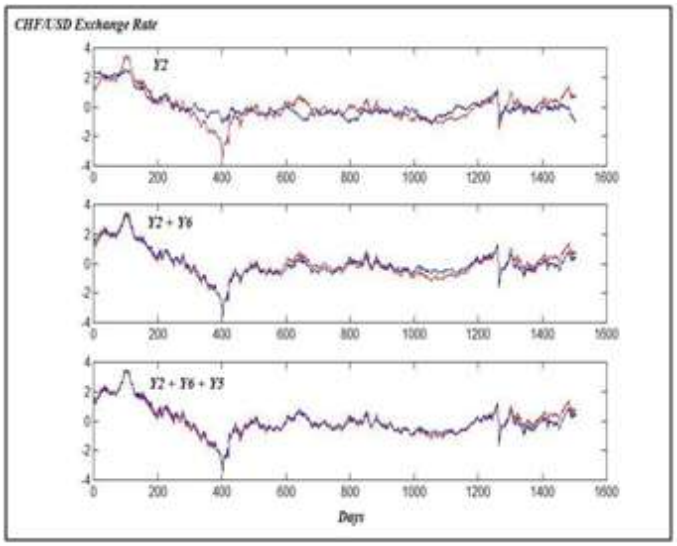

Figure 10. Reconstruction of exchange rate time series X_6 *(*CHF), Red (Observed), Blue (constructed)

It can be seen from these figures that the reconstruction of observed series can preserve the general trends with one or two dominant IC's and that quite acceptable matching can be realized with only the dominant 3 IC's in the lists.

VII. SUMMARY AND CONCLUSIONS

We have used Independent Component Analysis (ICA) in the process of reconstructing observed time series that might have been generated by mixed independent sources. A modified fast ICA algorithm is adopted with enhancements achieved by using non-linearities dependent on the normalized kurtosis of the observed series.

We have experimented with both artificially generated series and actual observed series of currency exchange rates. The artificial data were all super-gaussian for which we have used a general purpose non-linearity. The exchange rate data were a mixture between super-gaussian and sub-gaussian signals. For this latter dataset, we found that our fast ICA algorithm would converge in fewer iterations if we use a mixture non-linearity derived from a bimodal Exponential Power Distribution (EPD) symmetric mixture density.

For the reconstruction of the observed series from the obtained IC's, one has to obtain ordered lists of such components using an error measure that would minimize the error between observed and reconstructed series.

We have experimented with three different error measures namely, Relative Hamming Distance (RHD), Mean Square Error (MSE) and Average Mutual Information Distance (AMID). The area of the error profile is introduced as a sensitive measure of IC's order given by a list and to compare between the performances of the different error measures for a given ordered list. We find that the MSE measure is superior to the other two methods for all of the series concerned.

We have also compared five different methods for determining the optimal lists. These utilize Exhaustive Search (ES), the $L\infty$ norm (LN), Exclude Least contribution (EL), Maximizing Average Mutual Information (AMI) and Minimizing Mean square Error (MSE). Naturally ES gives optimal ordered lists but it requires $k!$ reconstruction steps where k is the number of IC's used.

From experiments with artificial and exchange rate datasets, we find that both EL and MSE algorithms offer excellent results since they give exactly the same ordered lists as the ES method. However, we consider the MSE to be the most efficient since it requires only k reconstruction steps compared to $k\ (k+1)/2-1$ steps for the EL algorithm and $k!$ steps for the ES algorithm.

On the other hand, the LN and AMI algorithms sometimes produce lists that are not optimal and therefore cannot compete with the other three algorithms. Here we might argue that the LN method does not necessarily consider the whole contribution of a given IC to the reconstruction process. As for the AMI algorithm, it is observed that AMI computations converge to correct results only for long series length [13] and that AMI values are not very sensitive to correlations in the pairs of series when they are characterized by low correlation values.

For comparing observed exchange rate series with the reconstructed ones, we have produced results for the percentage contribution of cumulative IC's from the ordered lists obtained using the MSE algorithm. For the majority of the series considered, we find that quite acceptable reconstructions (> 93 %) can be obtained with only the first 3 dominant IC's in the lists. Even the reconstruction with only the first, or with the first and the second IC's, the trends in the observed series are acceptably preserved.

REFERENCES

[1] A. Hyvärinen, J. Karhunen, and E. Oja, "Independent Component Analysis", John Wiley and sons , New York , NT, 2001

[2] A. Hyvärinen, "Independent component analysis: recent advances", Phil. Trans. R. Soc. A 371: 20110534, http://dx.doi.org/10.1098/rsta.2011.0534, 2013

[3] A. Mansour, and M. Kawamoto, "ICA papers classified according to their applications and performances", IEICE Trans. Fundamentals, vol. E86-A, No. 3, pp. 620-633, 2003

[4] Cheung, Yiu-ming, and Lei Xu. "Independent component ordering in ICA time series analysis", *Neurocomputing*, vol. 41.1, pp. 145-152, 2001

[5] Chi-Jie Lu, Tian-Shyug Lee and Chih-Chou Chiu, "Time series forecasting using independent component analysis and support vector regression", Decision Support Systems, vol. 47-2, pp. 115-125, 2009

[6] A. Hyvärinen, "Survey on independent component analysis", *Neural Computing Surveys 2*, pp. 94-128, 1999

[7] A. D. Back, and T.P. Trappenberg, "Input variable selection using independent component analysis" , *Proceedings of International Joint Conference on Neural Networks*, vol. 2, pp. 989-992, 1999

[8] A.D. Back, and A.S. Weigend, "A first application of independent component analysis to extracting structure from stock returns", Int. J. Neural Systems, vol. 8(4), pp. 473-484, 1997

[9] A. J. Hendrikse, R.N.J. Veldhuis, and L. J. Spreeuwers, "Component ordering in independent component analysis based on data power", *28th Symposium on Information Theory in the Benelux*, pp. 211-218, 2007

[10] A. Kamel, A. Goneid, and D. Mokhtar, "Ordering of dominant independent components in time series Analysis using fast ICA algorithm", Egyptian Computer Science Journal (ISSN-1110-2586), vol. 41–2, pp. 1-10, 2017 http://ecsjournal.org/Archive/Volume41_Issue2.aspx

[11] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis", *IEEE Trans. on Neural Networks*, vol. 10(3), pp. 626-634, 1999

[12] A. Goneid, A. Kamel, and I. Farag, "Generalized Mixture Models for Blind Source Separation", Egyptian Computer Science Journal, (ISSN-1110-2586), vol. 34-1, pp. 1-14, 2010

[13] R. Thomas, N. Moses, E. Semple, and A. Strang, "An efficient algorithm for the computation of average mutual information: Validation and implementation in Matlab", Journal of Mathematical Psychology, vol. 61, pp. 45-59, 2014

[14] Aslak Grinsted, https://www.mathworks.com/matlabcentral/fileexchange/10040-average-mutual-information, 2006

[15] C. J. Cellucci, A. M. Albano, and P.E. Rapp, "Statistical validation of mutual information calculations: Comparison of alternative numerical algorithms", Phys. Rev. E, Vol. 71, pp. 066208-1-066208-14, 2005

[16] www.bankofengland.co.uk/boeapps/iadb/

[17] www.oanda.com/solutions-for-business/historical-rates

[18] www.xe.com/currencytables/