Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-1 (2016) 077–082

Estimation of network connectivity strengths in linear causal dynamic systems

Sudhakar Kathari * **Arun K. Tangirala** *

* *Indian Institute of Technology Madras, Chennai-600 036, INDIA (e-mail: arunkt@iitm.ac.in)*

Abstract: Identification of network structure and quantifying the connectivity strengths in multivariate systems is an important problem in many scientific areas. Data-driven approach to network reconstruction based on causality measures is an emerging field of research in this respect. Among several recently introduced data-driven causality measures, the partial directed coherence (PDC) and direct power transfer (DPT) have been shown to be very effective for linear systems. While the PDC is useful in reconstructing the network, DPT has been proved to be effective in both identifying the network structure as well as quantifying the strength of connectivity. In this work, we study the problem of obtaining efficient estimates of network connectivity strengths, which has hitherto not been addressed in the literature. To this end, we study two different estimation methods for network connectivity strengths and demonstrate that the goodness of estimates depends on nature of the data generating process (DGP). In order to characterize the multivariate DGP, we introduce two statistics, namely, the vector-valued autocorrelation function (VACF) and the vector-valued partial autocorrelation function (VPACF), and estimators of the same. Our studies show that the parametric models used in estimating connectivity strengths should be commensurate with the dynamics of the process as characterized by the newly introduced VACF and VPACF. Simulation studies are presented under different scenarios to support our findings and the newly introduced measures.

Keywords: VACF, VPACF, causality, direct power transfer, strength of connectivity

1. INTRODUCTION

Identification of network structures in multivariate systems from measurements is an important problem in many areas such as engineering, systems biology, econometrics, statistics, sociology and climatology etc (Granger, 1969; Baccala and Sameshima, 2001; Gigi and Tangirala, 2012). The main objective in network reconstruction is to identify causal interactions between the variables from given time series data. The problem of identifying the causal relationships from the measurements was initially addressed by Wiener (1956). Granger (1969) proposed the definition of causality, based on Wiener's ideas, which is known today as Granger causality. The concept of Granger causality is based on prediction. A variety of data-driven causality measures that work either in the time- or frequency-domain have been proposed for over four decades now (Hlavackova-Schindler et al., 2007; Winterhalder et al., 2005). A majority of these measures rely on the concept of Granger causality. Among the *frequency-domain measures*, the partial directed coherence (PDC) and direct power transfer (DPT) are well-suited for structure determination as they measure direct influences between the variables (Baccala and Sameshima, 2001; Gigi and Tangirala, 2010). Both the methods use parametric time-series models, namely, the *vector auto-regressive* (VAR) models (see Appendix A) as the primary vehicles, regardless of the underlying process. The choice of this structure is motivated primarily by the ease of estimation. VAR models give rise to linear predictors thereby admitting least squares estimators to provide unique solutions. On the other hand, vector moving average (VMA) models yield predictors that are non-linear functions of parameters and are therefore, more complicated to estimate (Lutkepohl, 2005).

An additional problem of interest in network reconstruction is the determination of connectivity strengths. The knowledge of network connectivity strengths is valuable in several applications. A common use of the connectivity strength is in determining the strongest and weakest links, which finds applications in fault diagnosis, control of networks, etc. Despite its importance, relatively little effort has gone into defining and estimating strengths of connectivities. An ad hoc definition and computation is provided by Baccala and Sameshima (2001) based on the PDC. The definition therein lacks a transparent connection with any statistical relationship between the variables. On the other hand, the DPT-based definition introduced by Gigi and Tangirala (2010) directly quantifies the "amount" of transfer of power (or variability) from the source to the sink variable. The key step is the decomposition of the total power into direct, indirect and interference terms at each frequency (Gigi and Tangirala, 2010). Subsequently, the connectivity strength is derived as the normalized DPT between two variables (see §2.2). However, obtaining efficient estimates of this strength of connectivity has neither been addressed nor studied in the literature.

The objective of this work is to develop / identify a suitable method for obtaining unbiased and efficient estimates of connectivity strengths in linear, causal dynamical systems,

10.1016/j.ifacol.2016.03.032

or jointly stationary linear processes, based on the definition given in Gigi and Tangirala (2010). The specific questions of interest are:

(1) Does blindly adapting a VAR structure, without paying attention to the process characteristics, lead to inefficient estimates of connectivity strengths?
(2) Is it possible to devise practically useful measures that provide insights into the underlying multivariate data generating process?
(3) Does estimation of strength of connectivity by fitting a model that is commensurate with the process characteristics always give rise to efficient estimates in all the cases?

The first question assumes prominence in light of the fact that VAR models are widely used in the network reconstruction of linear causal dynamical systems. While this choice of model structure might serve the purpose of reconstruction, it may not be appropriate for obtaining efficient estimates of connectivity strengths, particularly when the underlying mechanism of the DGP is in significant deviation from the VAR structure. The second issue of interest is a natural follow-up of the first one. Since it is suspected that VAR model structures can lead to biased and / or inefficient estimates of strengths, it is necessary to be equipped with measures that provide insights into the nature of the DGP. The measures that exist in the multivariate time-series literature, namely, the matrix auto-correlation function (ACF) and its partial counterpart (Wei, 2005) do not, unfortunately, present sufficient insights into the collective characteristics of the multivariate process (see Appendix B). They are, however, ideally suited for obtaining a knowledge of the individual channel characteristics. Finally, the third issue of interest, perhaps being the practically most relevant one, is also a question that demands broader and comprehensive study. Therefore, for present, we restrict our scope of study to the first two issues.

The connectivity strength, as defined in (3), is a non-linear rational function of the model coefficients. Therefore, they do not lend themselves easily to a theoretical analysis of the bias and variance, i.e., it is difficult to theoretically derive the distributional characteristics of the estimates. Therefore, we adopt a Monte-Carlo simulation approach in this work. In addition, we introduce two new statistics, namely, the vector-valued ACF and partial ACF, for characterizing the correlation structure in multivariate stationary processes. These measures are useful in providing a collective picture of the correlation and can be thought of as multivariate analogues of the ACF and partial ACF (PACF) for the univariate process (Shumway and Stoffer, 2000; Tangirala, 2014). They possess similar properties as that of the univariate versions and aid in determining the order of VMA and VAR models, respectively. Further, we provide expressions for estimating these functions and study their distributional characteristics through Monte-Carlo simulations.

One of the main findings of this work is that VAR models are only suited for efficient estimation of connectivity strengths when the DGP is also of the VAR type. Any deviation, for e.g., when the DGP is VMA type, the use of VAR models result in biased and / or inefficient estimates, even when the VAR model order is chosen appropriately, i.e., the resulting model passes all necessary diagnostic tests. This finding, while prima facie, may be unsurprising, is also interesting since the connectivity strengths are defined in terms of the VAR model. On the other hand, when the model structure is chosen in accordance with the process characteristics, as determined by the vector-valued ACF / PACF, one obtains unbiased and relatively efficient estimates. A broader message of this work is that it is advisable to choose a model that is commensurate with the data generating mechanism rather than always choosing a VAR model for network reconstruction, which is the general practice. Further, the vector-valued ACF and PACF is potentially useful in developing multivariate time-series models for other applications as well.

The paper is organized as follows. Section 2 reviews vector time series models and quantification of connectivity strengths along with the estimation techniques. In Section 3, we introduce the vector-valued ACF and PACF and their sample versions with three illustrative examples. Simulation case studies are presented in Section 4. The paper ends with a few concluding remarks in Section 5.

2. PRELIMINARIES

This section reviews the quantification of connectivity strengths and their estimation methods based on parametric vector time-series modeling, a brief overview of which is provided in Appendix A.

2.1 Quantification of power transfers

The quantification of direct and indirect influences in terms of power transfer is obtained by developing direct and indirect transfer functions based on the signal flow graph representation of the process (Gigi and Tangirala, 2010). The mathematical expression for direct power transfer function ($h_{D,ij}(\omega)$)from source x_j to sink x_i is given as,

$$h_{D,ij}(\omega) = \begin{cases} \frac{-\bar{a}_{ij}(\omega)\det(\bar{\mathbf{M}}_{ij})}{\det(\bar{\mathbf{A}}(\omega))}, & i \neq j \\ \frac{\det(\bar{\mathbf{M}}_{ij})}{\det(\bar{\mathbf{A}}(\omega))}, & i = j \end{cases} \quad (1)$$

where $\bar{\mathbf{M}}_{ij}(\omega)$ is the minor matrix of the matrix $\bar{\mathbf{A}}(\omega)$, which is obtained from $\bar{\mathbf{A}}(\omega)$ by eliminating both i^{th} and j^{th} row and column.

The squared magnitudes of direct power transfer (DPT) from source x_j to sink x_i is,

$$|\psi_{ij}(\omega)|^2 = |h_{D,ij}(\omega)|^2 \quad (2)$$

The DPT gives both the structural information and the strength of connectivities of the process.

2.2 Strength of connectivities

The strength of connectivity is quantified based on the direct power transfer between the variables (Gigi and Tangirala, 2010). For a link connecting source x_j to sink x_i, the normalized connectivity strength is defined as (Gigi and Tangirala, 2012),

$$\beta_{ij} = \frac{\int_0^\pi |\psi_{ij}(\omega)|^2 d\omega}{\int_0^\pi |\psi_{jj}(\omega)|^2 d\omega} \quad (3)$$

It is also possible to use other normalization factors (for example, see Garg and Tangirala (2014)).

2.3 Estimation techniques

In this work, the strength of connectivity is obtained using parametric models based on the expression given in (3). Here, we study the use of three different approaches (i) VAR approximation, (ii) VMA approximation, and (iii) fitting a model that is commensurate with the process characteristics.

VAR approximation: The primary way of obtaining the network connectivity strengths is fitting a VAR model without paying attention to the process characteristics.

VMA approximation: Another alternative way to model VAR/VARMA process is by fitting a pure VMA model. The VMA model is estimated based on the method of maximum likelihood (Gómez, 2015).

Fitting a model that is commensurate with the process characteristics: In this technique, the network connectivity strengths are obtained by fitting the model that is commensurate with the process structure.

The estimated strength is always positive and it is zero when the connection does not exist between the variables. A statistical test of significance limit on connectivity strengths is beyond the scope of this work.

3. VACF AND VPACF

In this section, we introduce VACF and VPACF, which are useful for determining the characteristics of the multivariate process. Theoretical background on ACF and PACF for multivariate process is given in Appendix B.

3.1 Vector correlation functions

The autocorrelation and partial correlation matrices for multivariate processes give the nature of the DGP. We use graphical representation of auto- and cross-correlations to get the characteristics of the process. For example, Fig. 1 shows the sample auto- and cross-correlations for the data simulated from a 3-dimensional VMA(1) process. From the plots, it is observed that both auto- and cross-correlations are zero after the lags $l > 1$, which gives the nature of the process as VMA(1). However, the graphical representation of auto- and cross-correlations become increasingly cumbersome as the dimension of the process increases. Further, it is an admissibly difficult task to get the nature of the DGP from the listing of correlation matrices at different lags, particularly when the order of the process is high.

To overcome this difficulty and also to get the collective picture of the correlation, we introduce an equivalent scalar-valued representation of correlation matrices at each lag which we call as vector-valued ACF (VACF) and vector-valued PACF (VPACF).

The VACF is obtained by solving the following optimization problem

$$J = \min_{c} \sum_{i=1}^{M} \sum_{j=1}^{M} (\rho_{ij}[l] - c)^2 \tag{4}$$

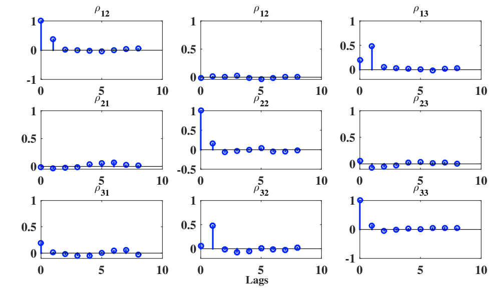
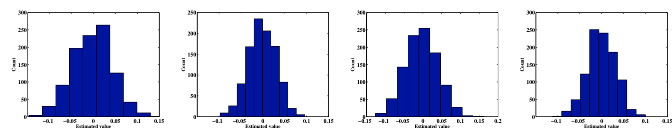
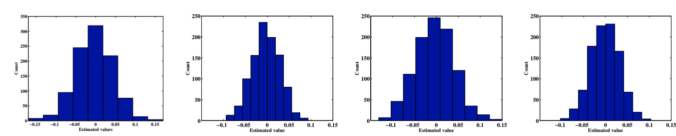


Fig. 1. Sample auto- and cross-correlations for 3-D VMA(1) process



(a) $l = 4, N = 500$ (b) $l = 4, N = 1000$ (c) $l = 9, N = 500$ (d) $l = 9, N = 1000$



(e) $l = 4, N = 500$ (f) $l = 4, N = 1000$ (g) $l = 9, N = 500$ (h) $l = 9, N = 1000$

Fig. 2. Histograms of VACF$(a-d)$ and VPACF$(e-h)$ estimates for VWN process

the solution of which is

$$c_l^* = \frac{1}{M^2} \sum_{i=1}^{M} \sum_{j=1}^{M} \rho_{ij} \tag{5}$$

where M is the dimension of the process. Therefore, the VACF at any lag l is defined as

$$\rho[l] = \frac{c_l^*}{c_0^*} \tag{6}$$

The VPACF is obtained by a similar procedure. Both these measures have the similar properties as that of univariate versions. Hence, VACF and VPACF are useful in determining the order of VMA and VAR models, respectively.

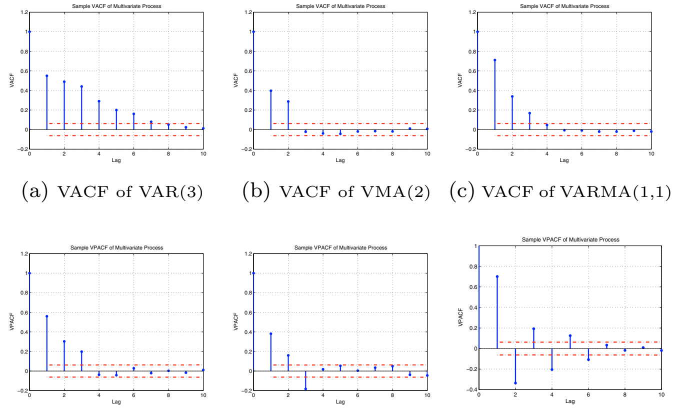
Remark: Ideally it is required to show theoretically that the measure is bounded. However, the theoretical proof is not given here and the work on the same is in progress.

3.2 Sample VACF and VPACF

The sample autocorrelation matrix function at lag l is

$$\hat{\Gamma}_l = \frac{\sum(\mathbf{x}[k] - \hat{\mu})(\mathbf{x}[k-l] - \hat{\mu})'}{\sum(\mathbf{x}[k] - \hat{\mu})(\mathbf{x}[k] - \hat{\mu})'} = \hat{\rho}_{ij}[l], \quad i,j = 1,2,....M \tag{7}$$

where $\hat{\mu}$ is the vector of the sample mean and $\hat{\rho}_{ij}[l]$ is the sample cross correlations when $i \neq j$, and sample autocorrelations when $i = j$. Sample partial correlation matrices are obtained using the recursive procedure (Wei, 2005) by estimation of Σ_l, i.e. using $\hat{\Sigma_l}$ in place of Σ_l. Sample VACF and sample VPACF are obtained, as defined in (6), from the sample autocorrelation and sample partial correlation matrices, respectively. Due to the effect of noise, these sampled versions will have non-zero values even when the true values are zero. It is therefore necessary



(a) VACF of VAR(3) (b) VACF of VMA(2) (c) VACF of VARMA(1,1)

(d) VPACF of VAR(3) (e) VPACF of VMA(2) (f) VPACF of VARMA(1,1)

Fig. 3. Sample VACF and VPACF for various multivariate processes

to determine significance levels. These are obtained using Monte Carlo simulations. The histograms of VACF and VPACF estimates for VWN process at different lags are shown in Fig. 2. It is observed that estimates are normally distributed with mean 0 and the variance $1/N$ (N = Number of samples). Therefore, the 95% significance levels for the estimates are

$$\pm 1.96\sigma_{\hat{\rho}[l]} \tag{8}$$

where $\sigma_{\hat{\rho}[l]}$ is the standard deviation of the estimates at l.

For example, Fig. 3(a) and Fig. 3(d) shows the sample VACF and VPACF obtained for the data simulated from a 3-dimensional VAR(3) process. From the plots it is observed that the VPACF is zero after the lags $l > 3$. Similarly the sample VACF and VPACF for a 4-dimensional VMA (2) are shown in Fig. 3(b) and Fig. 3(e), respectively. Fig. 3(c) and Fig. 3(f) shows the sample VACF and VPACF of VARMA(1,1) process. Therefore, the VACF and VPACF are useful for determining the overall characteristics of multivariate process.

4. SIMULATION RESULTS

Three case studies pertaining to different scenarios are presented to demonstrate the determination of network structure of the causal dynamical processes and quantification of strengths.

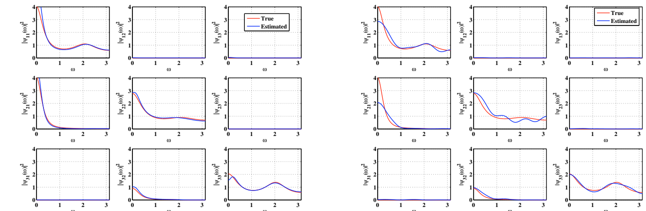
4.1 Case study 1

Consider a 3-dimensional VAR(3) process as

$$\mathbf{x}[k] = \begin{pmatrix} 0.2 & 0 & 0 \\ 0.3 & 0.2 & 0 \\ 0 & 0.1 & 0.1 \end{pmatrix}\mathbf{x}[k-1] + \begin{pmatrix} 0.1 & 0 & 0 \\ 0.2 & 0.1 & 0 \\ 0 & 0.2 & 0 \end{pmatrix}\mathbf{x}[k-2] + \begin{pmatrix} 0.2 & 0 & 0 \\ 0.1 & 0.1 & 0 \\ 0 & 0.1 & 0.2 \end{pmatrix}\mathbf{x}[k-3] + \mathbf{e}[k] \tag{9}$$

where $\mathbf{e}[k]$ is an uncorrelated VWN process with zero mean and unit variance. The process is simulated to obtain 1000 samples of the vector of $\mathbf{x}$. The true network along with connectivity strengths is shown in Fig. 5.

The sample VACF and VPACF shown in Fig. 3(a) and Fig. 3(d) suggest a VAR(3) DGP. This order is identical to the one suggested by the Akaike Information Criteria(AIC). The connectivity strengths are estimated from the direct power transfers by fitting a VAR(3) model. The plots



(a) No model-DGP mismatch (b) Using VMA approximation

Fig. 4. Estimated magnitude squares of DPT for case study 1.

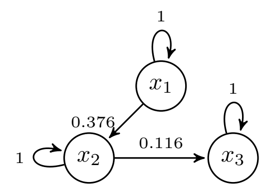
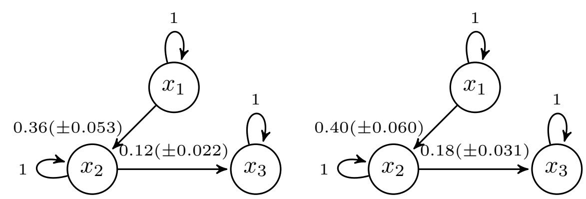


Fig. 5. True network structure for case study 1



(a) No model-DGP mismatch (b) VMA approximation

Fig. 6. Estimated network structure with strength of connectivities using different techniques for case study 1

of direct power transfers and the corresponding reconstructed network with connectivity strengths are shown in Fig.4(a) and Fig.6(a) respectively. The diagonal blocks represent the self-connectivity while the off-diagonal block $(i, j)^{th}$ represents the direct influence from source x_j to sink x_i. In the network connectivity diagram, variables are represented by nodes and the edge with an arrow between two nodes indicates the existence of a directed connection. The cyclic arrow on the node represents the influence of its own variable, if it exists. The value within the parenthesis represents the standard error in the estimates.

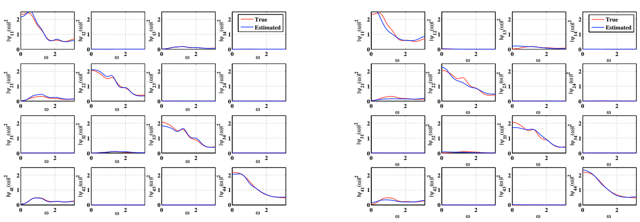
For comparison purpose, the connectivity strengths are also estimated by using VMA approximation. The estimated direct power transfers and reconstructed network along with connectivity strengths are shown in Fig. 4(b) and Fig. 6(b), respectively.

4.2 Case study 2

Consider a 4-dimensional VMA(2) process as

$$\mathbf{x}[k] = \begin{pmatrix} 0.4 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.4 & 0 \\ 0 & 0 & 0 & 0.4 \end{pmatrix}\mathbf{e}[k-1] + \begin{pmatrix} 0.2 & 0 & 0.3 & 0 \\ 0.4 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 0 \\ 0.5 & 0 & 0 & 0.1 \end{pmatrix}\mathbf{e}[k-2] + \mathbf{e}[k] \tag{10}$$

where $\mathbf{e}[k]$ is an VWN process with zero mean and unit variance. The sample VACF and VPACF shown in Fig. 3(b) and Fig. 3(e) indicates that the DGP has the characteristics of VMA(2) process. The true network along with connectivity strengths is shown in Fig. 8. The estimated direct power transfers and the network structure with connectivity strengths obtained by fitting a model with



(a) No model-DGP mismatch (b) Using VAR approximation

Fig. 7. Estimated squared magnitudes of DPT for case study 2

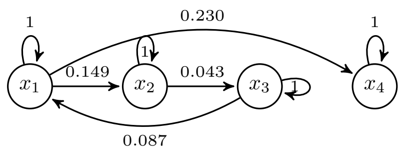
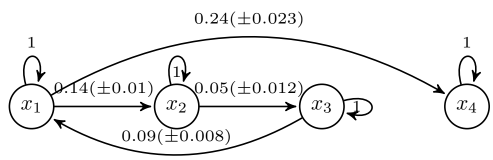
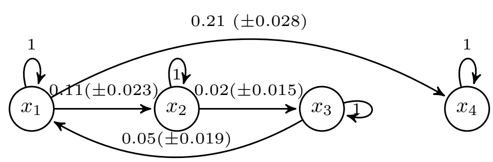


Fig. 8. True network structure for case study 2



(a) No model-DGP mismatch



(b) VAR approximation

Fig. 9. Estimated network structures with strength of connectivities using different techniques for case study 2

identical structure are shown in Fig.7(a) and Fig. 9(a), respectively.

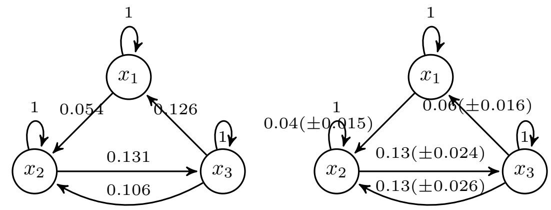
The direct power transfers and connectivity strengths obtained using VAR approximation are shown in Fig. 7(b) and Fig. 9(b), respectively. From these results, it is observed that estimation of connectivity strengths using VAR approximation leads to biased and inefficient estimates when the DGP is VMA.

4.3 Case study 3

Consider a 3-dimensional VARMA(1,1) process as

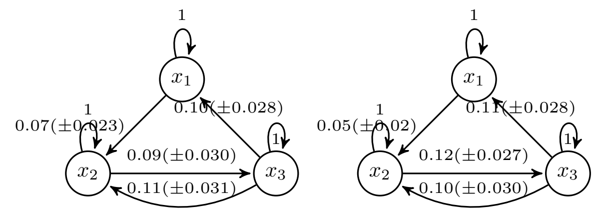
$$\mathbf{x}[k] = \begin{pmatrix} 0.3 & -0.1 & 0.2 \\ 0.1 & 0.2 & 0.1 \\ -0.1 & 0.2 & 0.4 \end{pmatrix} \mathbf{x}[k-1] + \begin{pmatrix} 0.5 & 0.2 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0 & 0.1 & 0.4 \end{pmatrix} \mathbf{e}[k-1] + \mathbf{e}[k] \quad (11)$$

where $\mathbf{e}[k]$ is the VWN with zero mean and unit variance. The sample VPACF in Fig. 3(f) suggests a VAR (6) approximation while the sample VACF shown in Fig. 3(c) indicates a VMA(3) approximation. The network structure along with connectivity strengths obtained by fitting a model with identical structure is shown in 11(b). Also, the estimated network connectivity strengths using VAR and VMA approximation are shown in Fig. 10(b) and Fig. 11(a) respectively. The true network connectivity strengths are shown in Fig. 10(a). From these results, it is observed that fitting a model that is commensurate with the process characteristics results in unbiased estimates when the underlying DGP is VARMA.



(a) True network (b) VAR approximation

Fig. 10. True and esitmated network structures for case study 3



(a) VMA approximation (b) No model-DGP mismatch

Fig. 11. Estimated network structures with connectivity strengths using different techiques for case study 3

A comparison of bias and standard errors of different estimation techniques for network connectivity strengths obtained using 100 Monte-Carlo simulations is shown in Table 1. It is observed that adapting a VAR structure without paying attention to the process characteristics, results in biased and / or inefficient estimates.

5. CONCLUSIONS

In this work we studied the impact of different estimation techniques on the bias and variance of network connectivity strengths in linear causal dynamical systems. One of the main objectives was to carefully investigate the suitability of VAR models for estimating connectivity strengths given that it is a general practice to use VAR models for network reconstruction. The approach was based on the use of two newly introduced vector measures for characterizing the auto-correlation and partial auto-correlation functions and Monte-Carlo simulations. We have also provided expressions for sample VACF and VPACF, while establishing significance levels for the same through Monte-Carlo simulations.

Simulation results demonstrated a comparison of estimation techniques and determination of network connectivity strengths under different scenarios. The main finding of this work is that fitting a model that is commensurate with the process characteristics results in unbiased and / or, efficient estimates of the connectivity strengths. Use of VAR models should be restricted to the case where DGP is also of VAR type. The statistical properties of the estimates are not investigated in present work and is a subject of future work. FInally, it is also believed that the vector measures introduced for obtaining a consolidated picture of the correlation in the multivariate process can

Table 1. Comparison of bias and standard errors for different estimation techniques

Case study	True strength	Estimated strength using			Bias			Standard error		
		VAR	VMA	VARMA	VAR	VMA	VARMA	VAR	VMA	VARMA
1 (DGP: VAR)	0.376	0.360	0.401	-	0.001	-0.012	-	0.053	0.060	-
	0.116	0.125	0.185	-	0.004	0.015	-	0.022	0.031	-
2 (DGP: VMA)	0.149	0.112	0.147	-	-0.024	0.007	-	0.023	0.010	-
	0.043	0.029	0.053	-	-0.010	0.003	-	0.015	0.012	-
	0.087	0.054	0.092	-	-0.010	0.003	-	0.019	0.008	-
	0.230	0.210	0.241	-	-0.043	0.003	-	0.028	0.023	-
3 (DGP: VARMA)	0.054	0.046	0.075	0.059	-0.012	0.006	0.005	0.015	0.023	0.020
	0.131	0.131	0.098	0.012	-0.012	-0.009	0.001	0.024	0.030	0.025
	0.106	0.138	0.118	0.010	-0.039	0.001	0.001	0.016	0.028	0.027
	0.126	0.062	0.103	0.119	-0.033	0.002	0.005	0.026	0.031	0.030

be used in other applications of multivariate time-series modeling.

REFERENCES

Baccala, L. and Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics*, 84(6), 463–474.

Garg, A. and Tangirala, A.K. (2014). Interaction assessment in multivariable control systems through causality analysis. In *Advances in Control and Optimization of Dynamical Systems*.

Gigi, S. and Tangirala, A.K. (2010). Quantitative analysis of directional strengths in jointly stationary linear multivariate processes. *Biological Cybernetics*, 103(2), 119–133.

Gigi, S. and Tangirala, A.K. (2012). Reconstructing plant connectivity using directed spectral decomposition. In *ADCHEM*, volume 8, 481–486. Singapore.

Gómez, V. (2015). Ssmmatlab: A set of matlab programs for the statistical analysis of state space models. *Journal of Statistical Software*, 66(9), 1–37.

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 424–38.

Hlavackova-Schindler, K., Palus, M., Vejmelka, M., and Bhattacharya, J. (2007). Causality detection based on information-theoretic approaches in time series analysis. *PhysicsReports*, 441.

Lutkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.

Priestley, M.B. (1981). *Spectral analysis and time series*. Academic Press.

Shumway, R.H. and Stoffer, D.S. (2000). *Time Series Analysis and Its Applications*. Springer.

Tangirala, A.K. (2014). *Principles of System Identification: Theory and Practice*. CRC Press.

Tiao, G. and Box, G. (1981). Modeling multiple times series with applications. *Journal of the American Statistical Association*, 76(376), pp. 802–816.

Wei, W. (2005). *Time Series Analysis: Univariate and Multivariate Methods*. Pearson.

Wiener, N. (1956). The theory of prediction. *Modern mathematics for the engineer*, pp. 165–190.

Winterhalder, M., Schelter, B., Hessec, W., Schwab, K., Leistritz, L., Klan, D., Bauer, R., Timmer, J., and Witte, H. (2005). Comparison of linear signal processing techniques to infer directed interactions in multivariate neural systems. *Signal Processing*, 85.

Appendix A. VAR, VMA AND VARMA MODELS

Consider an $M-$dimensional jointly stationary multivariate process is denoted by the vector $\mathbf{x} = (\ x_1 \quad x_2 \quad ...... \quad x_M\ )^T$. Vector autoregressive (VAR) representations are commonly used to express linear relationships among the variables. Mathematically, a VAR model of order P is represented as (Lutkepohl, 2005),

$$\mathbf{x}[k] = \sum_{r=1}^{P} \mathbf{A}_r \mathbf{x}[k-r] + \mathbf{e}[k] \tag{A.1}$$

where $\mathbf{A}_r$ is the AR coefficient matrix at lag r and $\mathbf{e}[k]$ is an M-dimensional vector white noise (VWN) process.

Vector moving average (VMA) models express the variables purely in terms of past and present white noise sequences. Mathematically, VMA (Q) is represented as (Shumway and Stoffer, 2000; Priestley, 1981),

$$\mathbf{x}[k] = \sum_{r=1}^{Q} \mathbf{H}_r \mathbf{e}[k-r] + \mathbf{e}[k] \tag{A.2}$$

where $\mathbf{H}_r$ is the MA coefficient matrix at lag r. The relationship between VAR and VMA models in frequency domain is given by

$$\mathbf{H}(\omega) = \bar{\mathbf{A}}^{-1}(\omega) \tag{A.3}$$

where $\bar{\mathbf{A}}(\omega) = I - \mathbf{A}(\omega)$ and $\mathbf{A}(\omega)$ is the Fourier transform of the AR coefficient matrix, $\mathbf{A}_r$, which is given by

$$\mathbf{A}(\omega) = \sum_{r=1}^{P} \mathbf{A}_r e^{-jwr} \tag{A.4}$$

Inversion of finite order stationary VAR process results in an infinite order VMA process, and vice versa. However, after certain lags the coefficients become insignificant and can be neglected.

The VARMA(P,Q) model has a mix of both VAR and VMA representations. Mathematically (Tiao and Box, 1981),

$$\mathbf{x}[k] = \sum_{r=1}^{P} \mathbf{A}_r \mathbf{x}[k-r] + \sum_{r=1}^{Q} \mathbf{H}_r \mathbf{e}[k-r] + \mathbf{e}[k] \tag{A.5}$$

where $\mathbf{A}_r$ and $\mathbf{H}_r$ are the autoregressive and the moving average coefficient matrices at lag r, respectively.

Appendix B. CORRELATION FUNCTIONS

The matrix ACF of a $M-$dimensional jointly stationary multivariate time series $\{\mathbf{x}[k]\}$ with mean vector μ at lag l is (Lutkepohl, 2005)

$$\Gamma_l = \mathbf{D}^{-1} \Sigma_l \mathbf{D}^{-1} = \rho_{ij}[l], \quad i, j = 1, 2, ..., M \tag{B.1}$$

where $\Sigma_l = E((\mathbf{x}[k] - \mu)(\mathbf{x}[k-l] - \mu)')$ is the autocovariance matrix function at lag l, $\rho_{ij}[l]$ is cross correlation between x_i and x_j when $i \neq j$ and is the autocorrelation when $i = j$, and $\mathbf{D}$ is a diagonal matrix whose elements are the square root of the diagonal elements of Σ_0. The autocorrelation matrices for a stationary VMA(Q) process are zero after the lag $l > Q$. On the other hand, for VAR(P) model these decays exponentially with lag. Hence, the ACF is useful for identifying the VMA models (Tiao and Box, 1981).

The matrix PACF of a jointly stationary process $\mathbf{x}[k]$ at any lag l is the correlation between $\mathbf{x}[k]$ and $\mathbf{x}[k-l]$ after removing their linear dependence on $\mathbf{x}[k-1]$, $\mathbf{x}[k-2]$,...., $\mathbf{x}[k-l+1]$ (Wei, 2005). It is a useful tool for identifying the order of the VAR process. The normalization used in computing the partial correlation is similar to the one used for multivariate ACF in (B.1). This is also termed as the partial lag correlation matrix function. Unlike, the partial autoregression matrix function (Tiao and Box, 1981), this matrix can be interpreted as a correlation matrix since the elements are appropriately normalized. A recursive procedure for computing the PACF is given in Wei (2005). The partial autocorrelation matrices for a stationary VAR(P) model are zero after the lags $l > P$. On the other hand for VMA(Q) model, these matrices exponentially decay with the lag.