Physica A 461 (2016) 629–646

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A study of industrial electricity consumption based on partial Granger causality network

Can-Zhong Yao[a,b,*], Qing-Wen Lin[a], Ji-Nan Lin[c]

[a] School of Economics and Commerce, South China University of Technology, Guangzhou, 510006, China

[b] Industrial and Urban Development Research Center, South China University of Technology, Guangzhou, 510006, China

[c] Department of Economics, University of California-Santa Barbara, Santa Barbara, 93106, USA

HIGHLIGHTS

- The energy transferring path could be studied by partial Granger network.
- The industries having higher degree tend to have stronger causality with others.
- Partial Granger networks have consistent hub industries with Granger network.
- The energy transferring paths in GD are more multidimensional and robust.
- Causal relationship could be found more reliable with bootstrap method.

ARTICLE INFO

Article history:
Received 27 January 2016
Received in revised form 30 April 2016
Available online 11 June 2016

Keywords:
Granger causality network
Partial Granger causality network
Bootstrap
Complementary network
Industrial influence

ABSTRACT

The paper studies the industrial energy transferring paths among the industries of China by distinguishing direct causality from the indirect. With complementary graphs, we propose that industrial causal relationship can be heterogeneous, and provide insights for refining robust industrial causality framework.

First, by analyzing the in-weight and out-weight of the industries in Granger causality networks we find that public utilities have significant causality with other industries, and the industries with higher degree value tend to have stronger causality with others.

Further, we eliminate the exogenous links by partial Granger causality model and find both Granger and partial Granger networks have consistent hub industries while some outliers emerge in partial Granger causality networks. Besides, compared with GX, GZ, HN and YN, the correlation between the volume of electricity consumption and the weight of each industry is more significant in the networks of GD and NF. By studying the characteristics of complementary graphs, we show that the industrial energy transferring paths in GD are more multidimensional, and the corresponding interdependent relationship among industries is more robust.

Finally, using bootstrap method we verify the reliability of each industrial relationship network. Results exhibit that GD, GX and NF have more reliable causal relationship networks than other provinces, revealing their industrial structure to be more stable.

* Corresponding author at: Industrial and Urban Development Research Center, South China University of Technology, Guangzhou, 510006, China.
E-mail address: ycz20120911@gmail.com (C.-Z. Yao).

http://dx.doi.org/10.1016/j.physa.2016.06.072

1. Introduction

Understanding how the industries interact is one of the most important issues in the analysis of industrial chain including value chains, supply chains and so on, which could also provide a theoretical insight and practical use for the researches of macro industry development and policy choices. Yao et al. [1] initially studied the evolution of industrial electricity-consumption sequences by minimal spanning trees in pre-crisis and after-crisis periods, the work also intuitively revealed the influence of financial crisis on industrial upgrading and transformation.

Frameworks of causal relationship network analysis originated from entity network analysis. A network can be defined as a random matrix, which is determined by the relationship of each node. There are some simple introductions to the networking of psychophysics, psychology, and neurophysiology [2]. Recently, network analysis has been involved in various studies, such as data mining, knowledge management, data visualization, statistical analysis, social capital, information dissemination and so on. For example, Getchell et al. [3] applied network analysis to official accounts of Twitter during the West Virginia water crisis. McGuire et al. [4] used network model to reveal multi-scale controls on stream-water chemistry.

Method of partial Granger causality stemmed from the Granger causality. In 1969, Granger [5] proposed causality of "prediction" (Granger causality), initially inspired by the thought of causality analysis proposed by Wiener [6], initially.

After the development done by Simmons, Granger causality test has been widely accepted and used by economists as a measurement method. It is defined in the way of hypothesis testing, to conclude whether there is a causal relationship between two time series. Since then, Granger causality analysis has been adopted in many fields. Yu et al. [7] used linear and nonlinear Granger causality to find out the relationship between carbon market and crude oil market. It turned out that for long time-scale, the long-term trends of two markets exhibited an obvious linear relationship. These methods are also actively applied in the stock price–volume Relation [8], neuroscience analysis [9–13] and climatology [14,15].

Granger causality characterizes the causal relationship between two time series. However, whether the causal relationship is direct or indirect is not obvious. For only two groups of time series, Granger causality analysis shows direct relationship. As for complex network analysis with multiple time series, Granger causality could hardly detects direct causal relationship. Therefore, we refer to the partial correlation causality method to distinguish between direct and indirect relation.

Partial Granger causality is commonly used to analyze the causal relationship among three or more variables in the model. In the causality analysis of two variable defined as X and Z, if adding a new variable Y to the models and removing the influence of the variable Z at the same time could improve the accuracy to predict the variable X, we can conclude that Y is the cause of X. Based on this method we extract the direct causal relationship among the variables of a certain model.

In aspect of interactive tempo-spatial time series detection, partial Granger causality network is becoming a competitive model. Relevant studies have been carried out by many scholars. For example, Krishna et al. [16] applied partial Granger to the analysis of chemical or biological systems to study the penicillin production process under several operating conditions. Guo et al. [17] eliminated exogenous input and latent variables by partial Granger causality networks, and further revealed relationship among time series with effective variables. Moreover, Roelstraete et al. [18] further validated partial Granger causality and found an answer why negative partial Granger causality estimates were reported. It was finally demonstrated that partial Granger causality would better eliminate latent variables and exogenous inputs than conditional Granger causality [19] which was also used to eliminate the influence of latent variables and exogenous inputs.

Motivated by these previous studies, this paper mainly applies Granger causality and partial Grainger causality to the analysis of industrial causality. Based on the electricity-consumption sequences including the industries of GD, GX, YN, GZ HN and NF respectively, we apply the partial Granger causality network to study the interactions among different industries. Moreover, by comparing the partial Granger causality with Granger causality model, we provide new perspectives for the industries' indirect relationship.

Bootstrap technique is widely used for estimating the confidence level in the fields of biological DNA sequence analysis [20–22], climatology [23,24], *etc.* We adopt bootstrap technique to test reliability among sectors, and to reveal the existence of incidental levels in the interactive influence of the industries.

2. Methodology statement

2.1. Granger causal model

In practice [25], the Granger causality is based on the following regressive models.

$$X_t = \sum_{i=1}^{p} a_i X_{t-i} + u_t \tag{1}$$

$$X_t = \sum_{i=1}^{p} b_i X_{t-i} + \sum_{i=1}^{p} c_i Y_{t-i} + v_t. \tag{2}$$

The variables are defined as follows:

X—the data of a certain industry wanted to find the cause;

Y—the data of other industries except X used to test whether it is the cause of X;