Energy Policy 88 (2016) 74–85

Contents lists available at ScienceDirect

Energy Policy

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

A comparative study on the influential factors of China's provincial energy intensity

Guangfei Yang [a], Wenli Li [a], Jianliang Wang [b], Dongqing Zhang [c,*]

[a] School of Management Science and Engineering, Dalian University of Technology, Dalian, China
[b] School of Business Administration, China University of Petroleum, Beijing, China
[c] College of Engineering, Nanjing Agricultural University, Nanjing, China

H I G H L I G H T S

- Identify the important factors of China's energy intensity by symbolic regression.
- Analyze China's energy intensity using provincial-level panel data from 1985 to 2012.
- Intelligently investigate nonlinear models and the emergence of important factors.
- The Total Population is discovered to be the most important influential factor.
- Provinces are naturally classified into four categories by the influential factors.

A R T I C L E I N F O

Article history:
Received 16 May 2015
Received in revised form
8 October 2015
Accepted 14 October 2015

Keywords:
Energy intensity
Symbolic regression
Energy consumption
Population
China

A B S T R A C T

China has become the largest energy consumer worldwide, and it is important to study the energy intensity to realize the sustainable development goal of China. This paper focuses on investigating the influential factors of China's energy intensity using provincial-level panel data from 1985 to 2012. More specifically, we try to identify which factor is relatively more important to pay attention to. A novel approach based on evolutionary computation is proposed to intelligently mine the intrinsic relations between observed phenomena and to let the important factors automatically emerge from the discovered nonlinear models. However, due to China's vast territory and significant heterogeneities, this approach may fail to examine some detailed or hidden information when analyzing the country as a whole. Instead, we concentrate on the provincial level because the provinces play vital roles in reducing energy intensity in China. From our analytical results, the main findings are as follows: (1) the Total Population is the most important influential factor across China's provinces, while the Energy Price Index has the least impact; and (2) the provinces could be naturally classified into four categories based on the primary factors emerged from data, and such classification could reveal more about the true underlying features of each area.

1. Introduction

The energy issue has been a key factor constraining the survival of human society and economic development. Since the opening and reform policy was initiated in 1978, China's economy has been growing at a high speed. Meanwhile, its energy demand has also been increasing for decades. With an economy that is expected to maintain a growth rate of 7–8% for decades, China's role in the world energy market has become increasingly influential (Crompton and Wu, 2005). In 2009, China surpassed the USA to become the largest energy consumer in the world (IEA, 2010). The total energy consumption of China in 2012 reached 3.617 billion TCE, including 66.6% of coal, 18.8% of crude oil, 5.2% of natural gas and 9.4% of hydropower, nuclear power and wind power (NSB, 2012). During the Twelfth Five-year period, China's energy demand gap will amount to 1.5 billion TCE (Qi, 2011).

In recent years, there has been much literature studying China's energy intensity. Energy intensity (EI) is an important indicator to measure a country's energy efficiency, indicating human development and progress, economic structure, fuel mix, and the technological level of a country (Sun, 2002). Generally speaking, the energy intensity in China has fallen since 1978 (Fisher-Vanden et al., 2004). During 2011–2015, China's energy intensity has been reduced by 16% (Seligsohn and Hsu, 2011). In the year 2014, China focused on controlling total energy consumption to reduce its

* Corresponding author.

http://dx.doi.org/10.1016/j.enpol.2015.10.011
0301-4215/© 2015 Published by Elsevier Ltd.

Table 1
Related research on China's energy intensity.

Authors	Period	Methodology	Focused influential factors
The decomposition method			
Huang (1993)	1980–1988	Decomposition method (Divisia index)	Intensity effect, structural effect
Sinton and Levine (1994)	1980s	Decomposition method (Laspeyres index)	Structural shift, intensity change
Garbaccio et al. (1999)	1987 and 1992	I–O tables/Decomposition method (SDA)	Technical change, structural change
Fisher-Vanden et al. (2004)	1997–1999	Decomposition method (IDA)	Efficiency change, structural change
Qi and Luo (2007)	1995–2005	Decomposition method (LMDI)	Efficiency change, structural change
Liao et al. (2007)	1997–2002	Decomposition method (TÖrnqvist index and Sato-Vartia index)	Sectoral structural effects and efficiency effects
Ma and Stern (2008)	1980–2003	Decomposition method (LMDI)	Technical change, structural change
Zhao et al. (2010)	1998–2006	Decomposition method (LMDI)	Production effect (or output effect), structural effect and efficiency effect (or intensity effect)
Zeng et al. (2014)	1997–2007	I–O tables (EIOA & MIOTs)/decomposition method (SDA)	Energy mix, sectoral energy efficiency, production structure, final demand structure among sectors, final demand composition
The regression method			
Karl and Chen (2010)	1996–2006	Regression method	Government expenditure, ratio of the tertiary sector to GDP, productivity, and energy prices
Zheng et al. (2011)	1999–2007	Multiple regression analysis	Exports, input in technological innovations, Foreign Direct Investment (FDI) intensity
Wu (2012)	1981–2007	Regression method/decomposition method (IDA)	Efficiency effect, structural effect & income, energy price
Herrerias et al. (2013)	1985–2008	Regression method	Investment ownership, imports, share of industry, sectoral composition, energy prices

Note: I–O tables: Input–output tables. EIOA: Environmental input–output analysis. MIOTs: Monetary input–output tables. SDA: Structural Decomposition Analysis. IDA: Index Decomposition Analysis. LMDI: Logarithmic Mean Divisia Index.

energy intensity that year by more than 3.9%, while in the last year it actually fell 3.7%, which means a reduction of 220 million tons of coal consumption (NPC, 2014).

Many experts and academics are committed to the inquiry about energy intensity as well as its influence factors. This body of research could be roughly classified into two groups according to the methods applied: the decomposition methods and the regression methods. A list of related researches on China's energy intensity is shown in Table 1.

On the one hand, from Table 1, it could be observed that the decomposition analysis has been widely used to study energy intensity. There are two commonly used decomposition methods: the index decomposition analysis (IDA) and the structural decomposition analysis (SDA). There are distinct differences between them: the structure and format of the data and the problem size (Ang and Liu, 2007). SDA uses complex input–output models and data to decompose changes in indicators, whereas IDA uses more aggregate sector level data (Ang and Liu, 2007). By contrast, because of the flexibility in problem formulation and data requirement of IDA, IDA methods have been more widely used than SDA methods in studying the drive forces of energy use and energy-related emissions (Su and Ang, 2012). As for IDA methods, Ang (2004) concluded that the logarithmic mean Divisia index (LMDI) method is the preferred method after compared various index decomposition analysis (IDA) methods. Ang (2005) gave a practical guide to the LMDI decomposition approach which is useful to practitioners interested in adopting this approach. Ang et al. (2009) have studied the properties and linkages of some popular index decomposition analysis (IDA) methods in energy and carbon emission analyses. Up to now, it has become the most popular method and has been frequently employed by international organizations and national statistics departments, as well as to monitor sectoral and economy-wide energy efficiency trends (Ang, 2006; Liu and Ang, 2007; Su and Ang, 2012). The LMDI-based accounting framework has been further advocated (Ang et al., 2010). There are many successful application examples of IDA methods. For instances, Ma and Stern (2008) employed logarithmic mean Divisia index (LMDI) techniques to decompose changes in energy intensity during the period 1980–2003. Zhao et al. (2010) conducted an index decomposition analysis to identify the key forces of the increase trend based on sub-sector data at the two-digit level. Furthermore, this method is being applied to numerous non-traditional areas, even in conjunction with other modeling methods and tools in innovative ways (Ang, 2015).

On the other hand, there are also some researches that employ the regression methods to investigate the energy intensity and its influential factors. Herrerias et al. (2013) investigated whether openness and investment ownership are key factors in explaining the diffusion of energy-saving technologies in China. Zheng et al. (2011) applied regression analysis to investigate the impact of exports on energy intensity in 20 sub-sectors during 1999–2007. Actually, the regression methods are relatively infrequent used compared with the decomposition methods.

Generally speaking, the decomposition method helps us differentiate the efficiency change at the micro-level from the economic structural change at the macro-level. However, just as Metcalf (2008) and Ma et al. (2010) have pointed out, some fundamental factors which greatly affect the energy intensity have not been rigorously examined, especially some aggregative variables, such as GDP, energy consumption, price, etc. As for the general regression method that has a predefined regression structure, different research with different regression structure will result into distinct results. In addition, due to the non-linearity of the relationship between energy intensity and its influential factors, some regression analysis with linear structure may lead to a significant deviation.

One contribution of this paper is that we propose a novel approach to study the influential factors of energy intensity in addition to showing the important factors, different from the researches that mainly apply decomposition analysis and regression analysis. The advantage of symbolic regression is that it is an effective tool to discover the hidden functions or even laws from data (Schmidt and Lipson, 2009), and it could act as a robot scientist to assist domain experts in analyzing complex problems. The core of symbolic regression is genetic programming (Koza, 1992, 1994), an extension of the genetic algorithm inspired from