Applied Energy 128 (2014) 1–14

Contents lists available at ScienceDirect

Applied Energy

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

A novel data-characteristic-driven modeling methodology for nuclear energy consumption forecasting

Ling Tang [a], Lean Yu [a,*], Kaijian He [a]

[a] School of Economics and Management, Beijing University of Chemical Technology, Beijing 100029, China

HIGHLIGHTS

- A novel data-characteristic-driven modeling methodology is proposed.
- The methodology formulates forecast model based on data's own data characteristics.
- Two steps are involved: data analysis and forecast modeling.
- Relationships between data characteristics and forecasting models are discussed.
- Empirical results statistically verify the effectiveness of our novel methodology.

ARTICLE INFO

Article history:
Received 5 February 2014
Received in revised form 3 April 2014
Accepted 4 April 2014
Available online 8 May 2014

Keywords:
Data characteristics
Nuclear energy consumption
Time series forecasting
Data driven modeling

ABSTRACT

Due to the unique features of nuclear energy market, this paper tries to propose a novel data-characteristic-driven modeling methodology based on the principle of "data-characteristic-driven modeling", aiming at formulating appropriate forecasting model closely in terms of sample data's own data characteristics. In the novel data-characteristic-driven modeling methodology, two steps are mainly involved, i.e., data analysis and forecasting modeling. First, the sample data of nuclear energy consumption are thoroughly investigated in order to capture the main inner rules and hidden patterns driving the data dynamics, in terms of data characteristics. Second, the corresponding forecasting model is accordingly formulated and designed based on these data characteristics. For illustration and verification purposes, the proposed methodology is implemented to predict the nuclear energy consumption of USA and China. The empirical results demonstrate that the novel methodology with the principle of "data-characteristic-driven modeling" strikingly improves prediction performance, since the models elaborately built based on data characteristics statistically outperform all other benchmark models without consideration of data characteristics. This further confirms that the proposed methodology is a very promising tool in both analyzing and forecasting nuclear energy consumption.

1. Introduction

Confronting with the double pressures of resource shortage and environment destruction, nuclear power, as a clean and renewable energy, has been playing an increasingly important role in the world energy system. In the last decades, the fast growth of global economy has accordingly caused rapid growth in energy demand. Furthermore, energy structure is still heavily dependent on fossil energy with limited deposits and high carbon emission, e.g., coal and oil, thus leading to two significant problems of environment damage and energy safety [1,2]. Due to the economic and environmental merits, nuclear energy has become one of the most promising energy options, especially to address the above two issues [3]. For example, since government policy changed in the late 1990s, the USA has undergone a significant growth in nuclear energy consumption and became the world's largest producer of nuclear power, accounting for more than 30% of worldwide nuclear electricity generation. As one of the largest developing countries, China's government also pays much attention to nuclear energy development and has proposed the medium- and long-term nuclear energy development plan (2005–2020), in order to promote nuclear power's production capacity and thence consumption.

Under such background, an accurate prediction for nuclear energy market, particularly the consumption, is quite necessary for both planning makers and producers, to ensure a rapid and

* Corresponding author. Address: School of Economics and Management, Beijing University of Chemical Technology, 15 Beisanhuan East Road, Beijing 100029, China. Tel./fax: +86 10 6443 8793.

E-mail address: yulean@amass.ac.cn (L. Yu).

http://dx.doi.org/10.1016/j.apenergy.2014.04.021
0306-2619/

Nomenclature

AI	artificial intelligence
AIC	akaike information criterion
ANN	artificial neural network
AR	auto regressive
ARIMA	auto-regressive integrated moving average
DGP	data generating process
FNN	feed-forward neural network
FT	fourier transform
ICSS	iterative cumulative sums of squares
LSSVR	least squares support vector regression
MA	moving average
OLS	ordinary least squares
SVM	support vector machine

stable nuclear energy development as expected. However, besides the factors in traditional energy market, a great many of other special ones also exist, significantly impacting the nuclear energy market, which lead to insurmountable difficulties in modeling nuclear energy. For example, the appalling nuclear accident at Fukushima in March 2011 has aroused a worldwide anxious about nuclear security, significantly shaking global nuclear energy market [4]. Besides, due to comparative immaturity, investment supports, technology improvement and political bias are also taking extremely important roles in nuclear energy market [5]. Therefore, the forecasting technologies, although doing well for other energy forms, might lose power in the case of nuclear energy with its special features.

Therefore, it is not difficult to understand: compared with other energy forms, the existing studies on nuclear energy forecasting are extremely inadequate and insufficient. First, according to extant literature, there are quite fewer studies on nuclear energy forecasting than those on other energy forms [5]. Furthermore, even in the relative few researches, nuclear energy has often been treated as one non-special energy form and simply mentioned amid various energies analyses, without targeted processing (e.g., [6,7]). Finally, the forecasting technologies usually applied in nuclear energy are mainly constrained to traditional models (e.g., [6–8]) and simple scenario analysis (e.g., [9]), while a variety of forecasting technologies are elaborately designed for other energy forms, involving artificial intelligences (AIs) and hybrid learning paradigms (e.g., [10]).

Due to the unique features of nuclear energy market, an interesting concept of "data-characteristic-driven modeling" can be introduced to partly address the difficulty in modeling the special factors in nuclear energy. That is, based on "data-characteristic-driven modeling", a promisingly effective forecasting method for nuclear energy consumption should be formulated closely based on the sample's own feature in terms of data characteristics. Actually, the principle of "data-characteristic-driven modeling" has already cast helpful lights on prediction model formulation. In particular, to capture nonlinearity characteristics of crude oil data, Wang et al. put forward a hybrid methodology, i.e., TEI@I, using different powerful AIs for nonlinear hidden patterns and econometric models for linear ones [11]. Exactly due to the data characteristics of complexity, irregularly and nonlinearity in nuclear market, Tang et al. proposed a hybrid ensemble learning paradigm with strategy of decomposition in order to address the difficulty in modeling nuclear energy consumption by decomposing the complex data into relative simple modes [5]. Similarly, considering the seasonality characteristic of hydropower consumption, Wang et al. introduced seasonal processing and put forward a seasonal decomposition based ensemble learning approach [10]. All results demonstrated that the forecasting models based on the data characteristics of sample data significantly outperformed their benchmark models without considering data characteristics.

However, the above studies just considered certain given data characteristics of the energy data, while other important data characteristics were ignored in modeling. Energy data, especially for nuclear energy data, always hold a series of coexisting and interactive data characteristics [12]. Furthermore, necessary numerical tests for data characteristics were always absent in existing forecasting researches, not to mention comprehensive analyses. However, according to the principle of "data-characteristic-driven modeling", a comprehensive analysis for data characteristics of nuclear energy consumption is indispensable. In particular, a powerful forecasting model should be carefully designed based on data characteristics, not only strictly following the inner rules of data dynamics but also effectively reflecting the main hidden factors [12,13].

Therefore, in terms of "data-characteristic-driven modeling", this paper tends to propose a novel "data-characteristic-driven modeling methodology" in order to formulate appropriate forecasting model based on sample data's own data characteristics, enhancing forecasting accuracy for nuclear energy consumption. Generally, there are two main steps in the proposed modeling methodology, i.e., data analysis and forecasting modeling. In the first step, the main data characteristics of nuclear energy consumption data are thoroughly and systematically tested and investigated. In the second step, closely depending on the data characteristics, forecasting model is accordingly designed to conduct prediction for nuclear energy consumption, by following the relationship between data characteristics and forecasting technologies.

The main innovation of this study is to propose a data-characteristic-driven modeling methodology for nuclear consumption forecasting, based on the principle of "data-characteristic-driven modeling". Two key steps are involved in the proposed method, i.e., data analysis and forecasting modeling, in order to formulate a compatible and appropriate forecasting model for nuclear energy consumption based on the data's own data characteristics. The rest of the paper is organized as follows. Section 2 describes the proposed data-characteristic-driven modeling methodology in detail, where the interrelationships between data characteristics and forecasting technologies are thoroughly discussed. Taking nuclear energy consumption of USA and China as sample data, the empirical study is performed in Section 3 for illustration and verification purposes. Finally, Section 4 concludes the paper and points out the direction of further research.

2. Methodology formulation

Based on the principle of "data-characteristic-driven modeling", this section tends to formulate the novel data-characteristic-driven modeling methodology to construct appropriate and powerful forecasting models depending on data characteristics of sample data. First, the data characteristics of energy time series data, together with their relationships, are thoroughly analyzed, as shown in Section 2.1. Second, Section 2.2 further discusses popularly used prediction technologies, in terms of function design and parameter estimation. Finally, the novel data-characteristic-