

Prediction of Total Energy Consumption in China and Analysis of Its Influencing Factors

Yue Li, Yujie Cui

College of Sciences, North China University of Technology, Beijing
Email: 1059273523@qq.com

Received: Nov. 16th, 2019; accepted: Dec. 4th, 2019; published: Dec. 11th, 2019

Abstract

With the development of society, our demand for energy is growing. By 2018, China's total energy consumption has reached 4640 million tons of standard coal. Energy is very important for a country's economic development. In this paper, multiple linear regression model and ARMA model are used to predict the total energy consumption. It is of practical significance for China's energy sector to formulate effective and reasonable energy consumption policies to predict the future development trend of China's total energy consumption through reasonable methods.

Keywords

Total Energy Consumption, Multiple Linear Regression, ARCH Model

中国能源消费总量的预测及影响因素分析

李悦, 崔玉杰

北方工业大学理学院, 北京
Email: 1059273523@qq.com

收稿日期: 2019年11月16日; 录用日期: 2019年12月4日; 发布日期: 2019年12月11日

摘要

随着社会的发展, 我们对能源的需求变得越来越大。截止到2018年, 我国的能源消费总量达到464,000万吨标准煤。能源对于一个国家的经济发展非常重要。本文分别用多元线性回归模型和ARMA模型对能源消费总量进行预测。通过合理的方法来预测我国能源消费总量的未来发展趋势, 对于我国能源部门制定有效合理的能源消费政策具有一定的现实意义。

关键词

能源消费总量, 多元线性回归, ARCH模型

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

能源是人类赖以生存的物质基础。随着时代的发展和社会的进步,我们对能源的需求量变得越来越大。1999年我国的能源消费总量为140,569万吨标准煤,经过20年的发展,2018年我国的能源消费总量高达464,000万吨标准煤,是1999年的2.3倍。由此导致了我国能源供给紧张的形势。能源在一个国家的经济中占据很重要的地位。因此,采用合理的方法准确地预测能源消费的未来发展趋势,能够为有关能源部门制定合理化的能源消费政策提供建议,同时可以促进我国经济健康发展,具有一定的现实意义。

王倩(2018)通过对我国能源消费现状进行分析,发现我国在过去几十年里能源消费达到了很高的程度,需要对我国的能源消费进行优化策略[1]。方圆,张万益等(2018)通过将中国的能源消费情况与世界各国的能源消费情况对比分析,发现我国过度依赖于煤炭、石油和天然气,核能发展相对落后[2]。汪琳,尹传凯(2016)认为对我国的消费现状和能源结构进行优化调整,是解决我国能源问题的根本之路[3]。杨杰,宋马林等(2009)基于协整理论、Granger因果关系检验等方法对我国的能源消费进行了实证分析[4]。张子荣(2018)运用VAR模型研究我国能源消费与GDP之间的关系,结果显示经济增长与能源消费之间是一种因果关系[5]。

2. 我国能源状况分析

2.1. 我国能源消费总量状况分析

由图1我们可以看到从1999年到2018年我国的能源消费总量是逐年递增的,2018年我国的能源消费总量已经达到了464,000万吨标准煤。在这期间,2004年的增速最快,为16.84%;2015年的增速最慢,为0.96%。

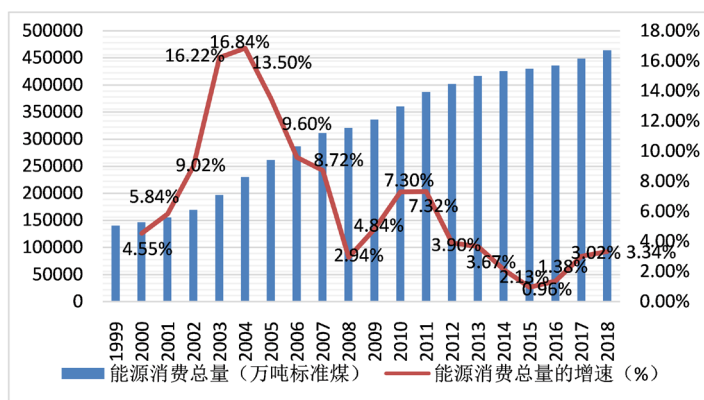


Figure 1. Change in total energy consumption in China from 1999 to 2018

图 1. 1999 年~2018 年中国能源消费总量变化图

2.2. 我国主要能源品种消费情况分析

由图 2 可以看到我国主要能源品种有煤炭、石油、天然气、水电、核电、风电。其中煤炭和石油这两种能源品种所占的比重相对较大, 同时我们看到随着科技的不断发展, 考虑到保护环境的重要性, 水电、核电、风电所占比重逐渐变大。

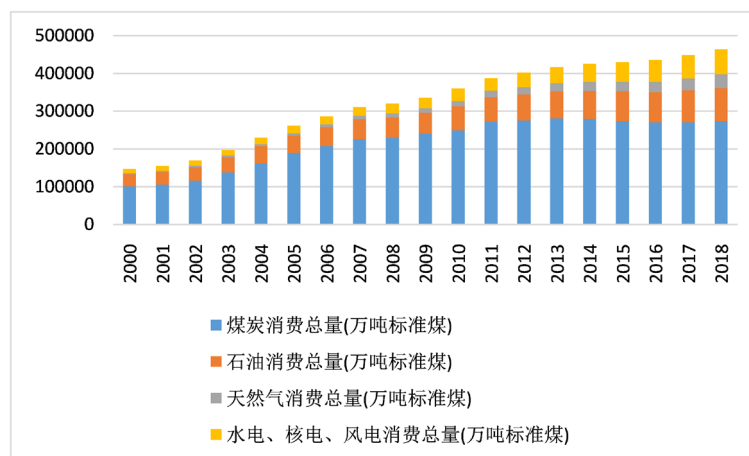


Figure 2. Structure of energy consumption of major varieties in China from 2000 to 2018

图 2. 2000 年~2018 年中国主要品种能源消费结构图

3. 我国能源消费总量影响因素分析

3.1. 数据的选择与处理

本次研究共收集了 1999 年到 2018 年共 20 年的能源消费总量相关数据, 为了去量纲化, 本文采用全序列法, 全序列功效系数的形式如下:

$$x_{ij}^*(t_k) = \phi + \frac{x_{ij}(t_k) - \min_{i,k} \{x_{ij}(t_k)\}}{\max_{i,k} \{x_{ij}(t_k)\} - \min_{i,k} \{x_{ij}(t_k)\}} \times \phi$$

$x_{ij}^*(t_k)$ 为第 i 个对象第 j 项指标在 t_k 时刻的无量纲数据。本文取 $\phi = 0, \varphi = 1$ 。

3.2. 基于多元回归分析的能源消费总量预测

3.2.1. 多元线性回归模型的基本描述

多元线性回归模型是一元线性回归模型的推广。用回归方程定量刻画一个因变量 y 的均值与多个自变量 x_1, x_2, \dots, x_p 存在线性关系, 多元线性回归模型的形式如下:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon$$

y 被称为因变量或者被解释变量, x_1, x_2, \dots, x_p 被称为自变量或者解释变量, $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ 被称为回归系数, ε 服从均值为 0, 方差为 σ^2 的正态随机误差项。

3.2.2. 基于逐步回归的方法建立能源消费总量预测模型

逐步回归分析方法通过将所有自变量对因变量的影响程度从大到小依次逐个引入方程中, 同时引入一个变量就对模型中的所有变量进行检验, 不显著的变量剔除, 保证模型中的含有自变量对因变量都是

显著的。

将能源消费总量(万吨标准煤) (CON)作为因变量。同时找到了影响我国能源消费总量的经济因素和社会因素, 分别为: 城镇化率(%) (URB)、固定资产投资(亿元) (INV)、国内生产总值(亿元) (GDP)、研究与试验经费发展支出(亿元) (EXP)、第三产业所占比重(%) (TER)。运用 SPSS 软件, 采用逐步回归的方法建立回归方程, 输出结果如下。

Table 1. Regression coefficient table

表 1. 回归系数表

Model	Unstandardized Coefficients		t	Sig.
	B	Std. Error		
(Constant)	-0.037	0.009	-4.220	0.001
URB	1.493	0.042	35.969	0.000
TER	-0.457	0.044	-10.380	0.000

由表 1 我们可以看出城镇化率(URB)、第三产业所占比重(TER)这两个变量进入模型, 其他变量都未进入, 两个变量 t 检验的 P 值都是小于 0.05 的, 说明回归方程线性显著。得到的回归方程为:

$$\hat{C\hat{O}N} = -0.037 + 1.493URB - 0.457TER$$

由表 2 可以看到复相关系数、判定系数和调整的判定系数, 三者均大于 0.9, 说明模型的拟合效果很好。

Table 2. Model summary table

表 2. 模型摘要表

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
2	0.999 ^b	0.997	0.997	0.0193502

表 3 是回归方程的显著性检验, 由检验的 P 值小于 0.05 可知回归方程线性显著。

Table 3. Analysis of variance table

表 3. 方差分析表

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	2.249	2	1.125	3003.511	0.000 ^b
Residual	0.006	17	0.000		
Total	2.256	19			

表 4 是 SPSS 软件根据建立的回归模型输出的我国能源消费总量各个年份的预测值、实际值以及两者相对误差的绝对值。

4. 基于 ARMA 模型的能源消费总量的预测

4.1. ARMA 模型的基本描述

ARMA 模型(自回归移动平均模型)的基本思想为一个变量现在的取值, 不仅会受到它本身过去值的影响, 也会受到现在和过去各种随机因素的影响。

Table 4. Model fit value and relative error absolute value table**表 4.** 模型拟合值及相对误差绝对值表

年份	实际值	拟合值	相对误差绝对值(%)
1999	0.0000		
2000	0.0198	0.01231	37.83
2001	0.0463	0.05489	18.55
2002	0.0897	0.10948	22.05
2003	0.1747	0.20247	15.90
2004	0.2774	0.30171	8.76
2005	0.3735	0.37261	0.24
2006	0.4511	0.43813	2.88
2007	0.5283	0.49679	5.96
2008	0.5567	0.56301	1.13
2009	0.6046	0.59703	1.25
2010	0.6805	0.70025	2.90
2011	0.7621	0.77657	1.90
2012	0.8087	0.81702	1.03
2013	0.8544	0.84275	1.36
2014	0.8819	0.87071	1.27
2015	0.8946	0.87202	2.52
2016	0.9129	0.90632	0.72
2017	0.9536	0.97361	2.10
2018	1.0000	0.99962	0.04

该模型表示为:

$$X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

其中 $\varphi_j (1 \leq j \leq p)$ 和 $\theta_j (1 \leq j \leq q)$ 为实数, ε_t 为白噪声过程, $\varepsilon_t \sim WN(0, \sigma^2)$ 。

4.2. 平稳性检验

原始序列经过二阶差分处理后的输出结果见表 5。

Table 5. Unit root test after differential processing**表 5.** 差分处理后单位根检验

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.196718	0.0079
Test critical values:		
1% level	-4.057910	
5% level	-3.119910	
10% level	-2.701103	

由表 5 我们可以看到此时的 P 值小于 0.05, 拒绝原假设, 表明该序列是二阶差分平稳序列, 即二阶单整。

4.3. 模型形式的识别

由图 3 可以发现样本 ACF 图拖尾, 样本 PACF 拖尾, 可以初步判断是 ARMA 模型。

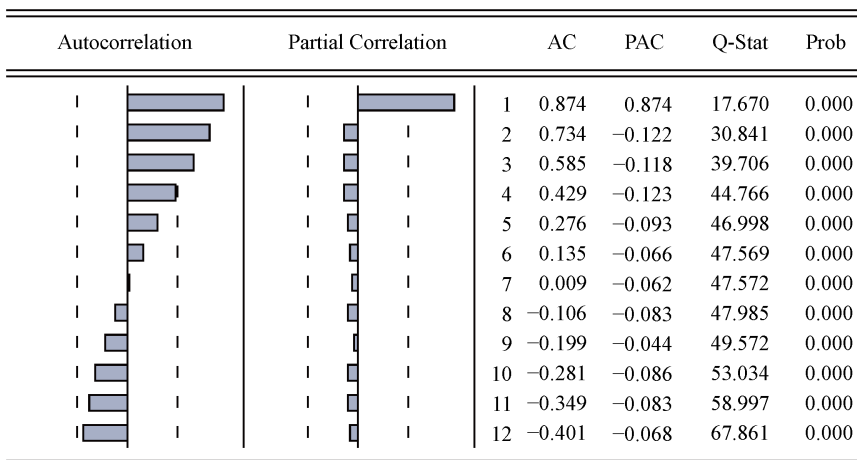


Figure 3. Sample PAF and sample ACF plot
图 3. 样本 PAF 和样本 ACF 图

4.4. 滞后阶数的确定

分别对 ARMA (1,1)、ARMA (1,2)、ARMA (2,1)、ARMA (2,2)模型进行回归, 记下模型中各最大滞后变量对应的 t 统计量值, 模型的 AIC、SC、HQ 值等信息。如表 6 所示。

Table 6. Values for the statistics of the four models
表 6. 四个模型的统计量的值

	AIC	SC	HQ
ARMA (1,1)	-4.6801	-4.5309	-4.6548
ARMA (1,2)	-5.1841	-4.9853	-5.1505
ARMA (2,1)	-4.8346	-4.6367	-4.8073
ARMA (2,2)	-5.1942	-4.9469	-5.1601

4.5. 模型回归与检验

表 7 是残差序列相关 LM 检验选取滞后 1 阶进行残差相关检验, LM 检验统计量 obs*R-squared 对应的 P 值 0.523, 在 5%的显著性水平下, 不拒绝原假设, 检验结果表明残差无序列相关。表明变量无遗漏变量, 即滞后阶数的选取是合理的。

Table 7. Breusch-Godfrey serial correlation LM test
表 7. 序列相关 LM 检验表

F-statistic	1.845363	Prob.F (1,14)	0.1958
Obs*R-squared	2.049095	Prob.Chi-Square (1)	0.1523

由表 ARCH (1,2)可以得到模型的回归方程为:

$$CONS_t = 2.1826 * (1 - 0.9626) + 0.9626CONS_{t-1} + \hat{v}_t + 1.4666\hat{v}_{t-1} + 0.9880\hat{v}_{t-2}$$

采用静态预测的方法得到的预测图, 见图 4。

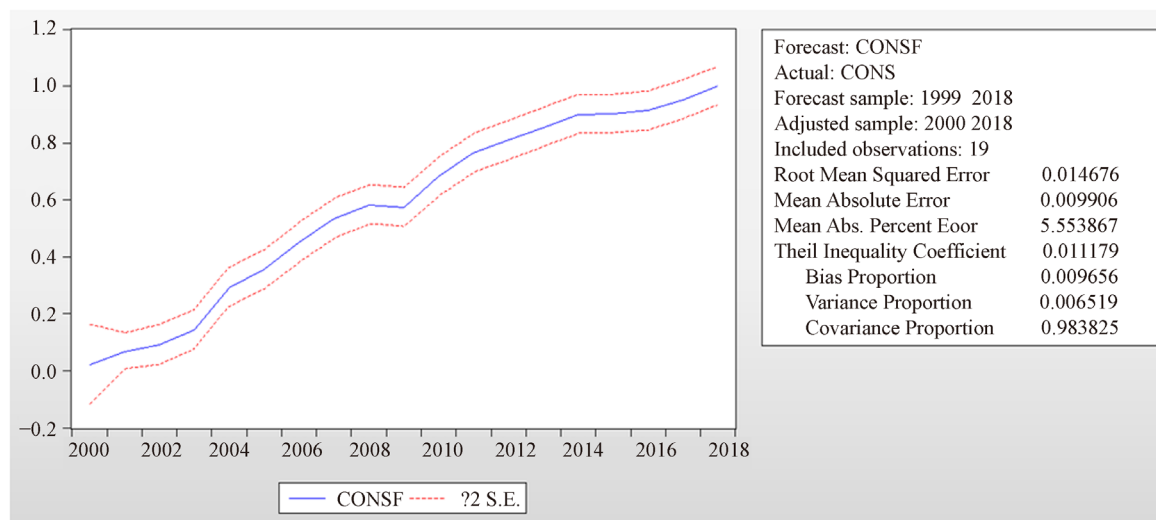


Figure 4. Forecast graph

图 4. 预测图

将实际值与运用 ARCH 模型预测得到的拟合值分别放到同一张表, 并计算对应的相对误差绝对值, 如表 8 所示。

Table 8. Absolute value of relative error

表 8. 相对误差绝对值

年份	实际值	拟合值	相对误差绝对值(%)
1999	0.0000		
2000	0.0198	0.022310	12.68
2001	0.0463	0.068394	47.72
2002	0.0897	0.091317	1.80
2003	0.1747	0.143760	17.71
2004	0.2774	0.293648	5.86
2005	0.3735	0.355373	4.85
2006	0.4511	0.451661	0.12
2007	0.5283	0.532931	0.88
2008	0.5567	0.582855	4.70
2009	0.6046	0.574499	4.98
2010	0.6805	0.681967	0.22
2011	0.7621	0.764184	0.27
2012	0.8087	0.810579	0.23
2013	0.8544	0.855310	0.11
2014	0.8819	0.900952	2.16
2015	0.8946	0.901746	0.80
2016	0.9129	0.913439	0.06
2017	0.9536	0.952446	0.12
2018	1.0000	1.000750	0.08

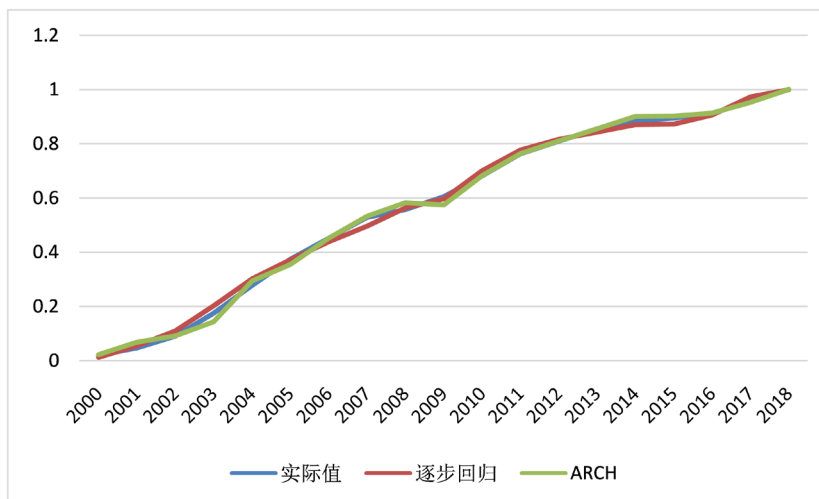


Figure 5. Comparison of model fitting effects
图 5. 模型拟合效果对比图

由上图 5 可以看出, 通过绘制折线图发现逐步回归的方法和 ARCH 模型两种方法预测出的结果与实际值能较好地契合, 说明这两种方法在拟合我国能源消费总量时的效果比较好。

计算两种模型的预测结果与实际结果的相对误差绝对值。

见表 9, 通过计算两种方法的预测结果与实际结果的相对误差绝对值, 由上表可以看到 ARCH 模型计算的相对误差绝对值的平均数要比多元线性回归模型的稍小一些, 然而多元线性回归模型的标准差相对要小一些, 说明我们在对能源消费总量进行预测时可以结合这两种方法的优缺点, 在探索多变量影响因素方面, 可以考虑多元线性回归模型, 而在单一变量进行预测时采用 ARCH 模型。

Table 9. Comparison table of the two methods
表 9. 两种方法的对比表

	平均数	标准差
多元线性回归模型	0.067578209	0.099939741
ARCH 模型	0.05544182	0.112589011

5. 结论与建议

5.1. 结论

1、我国的能源消费总量近 20 年是逐年递增的, 随着科技的不断发展, 水电、核电、风电这些清洁能源所占比重逐渐增大。

2、煤炭在我国主要能源品种消费中占有的比重过大, 石油所占的比重相对较大。

3、我国能源消费总量的增速相比 20 年前已经明显减缓, 说明我国在控制能源消费方面有了显著的成效。

4、城镇化率和第三产业所占比重对我国的能源消费总量影响显著, 其中城镇化率对能源消费总量起到正向的作用, 第三产业所占比重对能源消费总量起到负向的作用。

5、在预测我国能源消费总量时, ARCH 模型和逐步回归方法的拟合效果都比较好, 对于多变量预测可以采用逐步回归的方法, 对于单变量预测可以选用 ARCH 模型。

5.2. 建议

1、推动产业结构优化升级, 大力发展高新技术产业, 可以通过对高新技术及新兴产业发展实行财政补贴和税收优惠。

2、通过技术创新, 提高能源使用效率。政府应当加大对高科技人才的培养与引进, 加强企业对新兴技术的使用, 此外, 我国应当促进与在能源行业有优势的国家的交流与合作。

3、我国应当控制不可再生能源的使用, 可以对煤炭价格进行合理的调整, 限制煤炭资源的开采, 鼓励使用清洁能源。

4、地球上存在着很多可再生能源, 应当鼓励政府科技创新, 通过风力发电, 改善生态环境, 缓解全球变暖。

基金项目

本文受到北方工业大学学生科技活动项目资助, 项目编号: 11005130019XN140/006。

参考文献

- [1] 王倩. 我国能源消费现状及其优化策略[J]. 商业经济研究, 2018(17): 40-42.
- [2] 方圆, 张万益, 曹佳文, 朱龙伟. 我国能源资源现状与发展趋势[J]. 矿产保护与利用, 2018(4): 34-42+47.
- [3] 汪琳, 尹传凯. 中国能源消费现状分析及政策研究[J]. 商场现代化, 2016(6): 237-238.
- [4] 杨杰, 宋马林, 叶小榕. 中国能源消费的实证研究[J]. 粤港澳市场与价格, 2009(2): 17-20.
- [5] 张子荣. 我国经济增长与能源消费关系的实证分析[J]. 商业经济研究, 2018(17): 36-39.