

# Energy Consumption Prediction in Yunnan Province Based on Mixed Time Series Model

Xiaoxiao Yin<sup>1</sup>, Wen Gan<sup>2</sup>

<sup>1</sup>School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming

<sup>2</sup>International Business School, Yunnan University of Finance and Economics, Kunming

Email: [yinxiaoxiao2011@163.com](mailto:yinxiaoxiao2011@163.com)

Received: Oct. 10<sup>th</sup>, 2014; revised: Nov. 12<sup>th</sup>, 2014; accepted: Nov. 21<sup>st</sup>, 2014

Copyright © 2014 by authors 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

---

## Abstract

Energy is an important material security for human survival, economic development and social progress. Based on the historical data of energy consumption in Yunnan Province, we first establish a trend extrapolation model to estimate the total energy consumption of Yunnan Province in the future. Second, we combine the trend extrapolation model and ARIMA model to give a mixed time series model. The prediction results of these two models are analysed and compared. The results show that the mixed time series model is performed better. This indicates that the mixed time series model is a useful theoretical tool for energy prediction in Yunnan Province.

## Keywords

Trend Extrapolation Model, ARIMA Model, Mixed Time Series Model, Energy Consumption, Predictions

---

# 基于混合时间序列模型的云南省能源消费预测研究

尹潇潇<sup>1</sup>, 干文<sup>2</sup>

<sup>1</sup>云南财经大学统计与数学学院, 昆明

<sup>2</sup>云南财经大学国际工商学院, 昆明

Email: [yinxiaoxiao2011@163.com](mailto:yinxiaoxiao2011@163.com)

收稿日期：2014年10月10日；修回日期：2014年11月12日；录用日期：2014年11月21日

## 摘要

能源是人类生存、经济发展和社会进步的重要物质保障。本文结合云南省能源消费总量的历史数据，首先建立趋势外推模型，预测云南省未来能源消费总量；然后将趋势外推模型和ARIMA模型相结合，利用混合时间序列模型进行预测和分析；通过比较上述预测结果，发现混合时间序列模型具有更好的预测效果，说明该模型对云南省能源消费总量的预测有重要的理论与现实意义。

## 关键词

趋势外推，ARIMA模型，混合时间序列模型，能源消费，预测

## 1. 引言

能源是人类生存和发展的重要物质保障，随着经济的高速发展，社会的快速进步以及人们生活水平的不断提高，中国能源消耗量在不断的增加，人类对能源的需求量也大幅度的增加，从而导致能源供给日趋紧张，能源问题也将成为我国未来经济发展和社会进步的制约因素。能源在国民经济中占有重要的地位，做好能源预测的相关工作，准确预测能源未来消费的发展趋势，可以给有关部门制定科学的能源战略规划，合理的能源消费政策提供依据，同时有利于维护我国国民经济健康、持续、稳定的发展，建设节约型和谐社会，这具有非常重要的现实意义和战略意义。

国内外很多学者采用不同的预测方法对我国能源消费总量进行研究，当然不同的方法其预测精度也不同。国外学者[1][2]分别建立ARIMA模型和GM(1,1)进行预测。国内学者如韩君[3]运用趋势外推中的三次曲线模型对我国1990~2003年能源消费总量进行预测；程静[4]以广东省1979~2006年能源消费总量为基础，运用Eviews软件，建立ARMA模型预测广东省2007~2010年的能源消费总量；徐明德[5]采用灰色预测法进行预测分析；孙文生[6]使用BP神经网络法预测河南省煤炭消费总量。目前，组合模型预测法备受欢迎，如柴元春[7]以我国1989~2009年能源生产总量数据为基础，建立组合模型并证明其预测效果比单一模型要好；王惠婷[8]将趋势外推模型与ARIMA模型组合，对河南省许昌市粮食产量进行预测，发现混合模型预测占优势；韩君也采用趋势外推与ARIMA组合模型分析预测了我国能源需求情况。

目前，云南省能源消费量相关预测的研究很少，基于能源对人类的重要性以及云南省是一个能源大省的事实，做好云南省未来能源消费总量预测工作意义重大，可以为云南省政府制定合理的能源战略及规划提供科学依据，保障云南省能源的合理利用以及健康稳定的发展。本文在一些学者研究的基础上，建立了混合时间序列模型进行预测分析，该模型弥补了单一趋势外推模型不能解释非趋势分量的缺陷，而且预测效果优于趋势外推模型。

## 2. 模型及方法

通过阅读大量的文献，我们了解到国内外学者对能源预测模型进行了大量的研究，目前主要的预测方法有趋势外推法、时间序列分析法、神经网络模型法、灰色预测法、能源弹性系数法、投入产出法和组合模型预测等方法。对同一数据使用不同的方法建立模型，得到的预测结果会存在一定的差异，它们的预测精度也各不相同。本文基于所选数据的特征，选择建立趋势外推模型和混合时间序列模型对

云南省能源消费总量进行分析预测。

## 2.1. 趋势外推模型

### 2.1.1. 趋势外推法基本介绍

趋势外推法[9]，又称趋势外插法，是根据过去和现在的发展趋势推断未来的一类方法的总称。其基本依据是预测的连续性原理，根据预测对象发展具有规律性的特点，通过正确把握预测对象过去和现在的发展状况，来预测未来的发展趋势。

随着时间的变化，事物的发展呈现一定的规律性，而且所要预测的对象也会具有一定的趋势，这时我们可以寻找合适的函数曲线来反映这种变化趋势，以时间  $t$  为自变量，时间序列值  $y$  为因变量，建立趋势外推模型：

$$y = f(t) \quad (1)$$

而以顺延的时间作为已知条件，根据拟合的模型可以得到趋势值即预测值。由此看来，模型拟合的好坏将直接影响到预测的准确程度，我们常用最小二乘法拟合趋势模型，因为它拟合出的模型的预测标准误差最小。

### 2.1.2. 趋势外推模型的识别与选择

预测结果的准确与否与所建立的趋势模型有关，因此，曲线模型的选择至关重要，下面给出几种常见曲线模型的识别与选择方法：

1) 直观判断法。绘制已知时间序列数据的散点或趋势图，与常见的曲线模型比较，选择分布比较接近的曲线模型作为趋势外推模型。

2) 特征分析法。通过分析已知数据所具有的特征，从我们所熟悉的曲线模型中选择出与此特征相符的模型。

3) 预测精度比较法。当我们不能确定究竟选择哪一种模型来拟合历史数据时，我们可以把所有可能的曲线模型都分别来拟合一次，然后比较这几种模型的拟合优度，从而选择出最好的预测模型。

### 2.1.3. 曲线模型估计的步骤

当我们在解决实际问题时，不能确定哪一种曲线最能反映该问题的趋势情况，这时我们可以运用曲线估计。在估计过程中有很多曲线可供我们选择，如：线性、二次、三次、指数、增长等曲线，只要这些曲线可以描述变量之间的大概关系，我们就可以进行曲线估计分析。分析步骤如下：

- 1) 绘制已知时间序列数据的趋势图；
- 2) 根据趋势图，选择几种可能的曲线模型进行拟合；
- 3) 采用最小二乘法得到曲线模型的参数估计值，以及统计量  $R$  方、 $F$  值和  $P$  值；
- 4) 对参数估计的相关统计量进行检验，看其是否通过显著性检验；
- 5) 通常选择  $R$  方较大的模型作为预测模型，当然也可以进一步比较它们的预测精度进行模型筛选。

## 2.2. 时间序列分析法

时间序列分析法是以预测对象时间序列的历史数据为基础，运用一定的数学方法使其向外延伸，并对未来的发展变化趋势进行预测。

### 2.2.1. ARIMA 模型基本介绍

ARIMA 模型的全称为差分自回归移动平均模型，是由博克思(Box)和詹金斯(Jenkins)于 70 年代初提出的著名时间序列预测方法，所以又称为 Box-Jenkins 模型、博克思 - 詹金斯法。所谓 ARIMA 模型，是

指将非平稳时间序列转化为平稳时间序列，然后将因变量仅对它的滞后值以及随机误差项的现值和滞后值进行回归所建立的模型。其基本思想是：将预测对象随时间推移而形成的数据序列视为一个随机序列，用一定的数学模型来近似描述这个序列。

ARIMA 模型的结构[10]为：

$$\begin{cases} \Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, \quad s \neq t \\ E x_s \varepsilon_t = 0, \quad \forall s < t \end{cases} \quad (2)$$

式(2)中， $\nabla^d = (1-B)^d$ ； $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ ，为平稳可逆 ARIMA( $p, q$ )模型的自回归系数多项式； $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ ，为平稳可逆 ARIMA( $p, q$ )模型的移动平滑系数多项式。

式(2)可以简记为：

$$\nabla^d x_t = \frac{\Theta(B)}{\Phi(B)} \varepsilon_t \quad (3)$$

式(3)中， $\{\varepsilon_t\}$  为零均值白噪声序列。

由(3)式显而易见，ARIMA 模型的实质就是差分运算与 ARMA 模型的组合。这说明任何非平稳序列只要通过适当阶数的差分实现差分后平稳，就可以对差分后序列进行 ARMA 模型拟合了。

### 2.2.2. ARIMA 模型建立的基本步骤

- 1) 根据所收集的时间序列数据，利用 SPSS 或 Eviews 软件作出该时间序列的散点图；
- 2) 通过 ADF 单位根检验，判断该序列的平稳性；
- 3) 对非平稳序列进行平稳化处理，通常对其进行差分运算使其变为平稳序列；
- 4) 绘制差分序列的自相关和偏自相关图，并得到自相关系数(ACF)和偏自相关系数(PACF)表；
- 5) 确定 ARIMA( $p, d, q$ )模型的阶数  $p$  和  $q$ ，其中  $d$  为差分的次数；
- 6) 估计模型的参数，并检验其是否具有统计意义；
- 7) 对残差序列进行白噪声检验；
- 8) 运用已通过检验的模型进行预测分析。

### 2.3. 混合时间序列模型

本文所谓的混合时间序列模型就是趋势外推法和 ARIMA 模型的结合，即先对时间序列配合趋势模型  $f(t)$ ，用来描述序列中的趋势分量，对于非趋势性分量，即趋势值与实际值的离差序列，再用 ARIMA 模型进行描述，最后把两个模型叠加起来进行预测研究[11]。

混合时间序列模型可表示为：

$$\begin{cases} y_t = f(t) + \mu_t \\ \varphi(B)\mu_t = \theta(B)\varepsilon_t \end{cases} \quad (4)$$

其中  $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p$ ， $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ ， $B$  为滞后算子， $\varepsilon_t$  是服从正态分布的白噪声误差项。

## 3. 实证分析

### 3.1. 数据来源及说明

根据《云南省统计年鉴 - 2011》和《云南省统计年鉴 - 2012》，本文收集了云南省 1978 到 2011 年

的能源消费总量数据，将 1978~2007 年的数据[12]作为训练集，建立预测模型，而 2008~2011 年的数据作为测试集，用于检验模型的预测效果，进而对云南省未来能源消费总量作出合理的预测。

将 1978~2007 年云南省能源消费总量记为  $y_t$  序列，其中  $t = 1, 2, \dots, 30$  (1978 年对应  $t = 1$ )，时间  $t$  为自变量， $y_t$  为因变量。

### 3.2. 云南省能源消费总量的预测模型建立

#### 3.2.1. 建立趋势外推模型

1) 绘制时间序列趋势图

运用 Eviews 软件绘制 1978~2007 年云南省能源消费总量的散点图，见图 1:

2) 模型选择及曲线估计

观察散点图，发现  $y_t$  随时间的变化呈现逐渐上升的趋势，我们可尝试建立以时间  $t$  为自变量， $y_t$  为因变量的二次曲线、三次曲线和指数曲线模型，曲线估计结果见表 1。

表中的 R 方为可决系数，是衡量自变量与因变量之间密切关系的指标，其范围介于 0 到 1 之间，R 方的值越大，表明方程的解释能力越好，对数据的拟合优度越高。由表 1 中的数据可知，二次模型、三次模型以及指数模型的判定系数都较高，但相对而言，三次曲线模型的可决系数最大。我们选择三次曲线模型对云南省能源消费总量建模较好，其判定系数为 0.986，拟合优度较好；而且模型中的 F 统计量为 615.804，相伴概率 Sig 为 0.000，小于 0.05，方程和系数都通过了检验，三次曲线模型具有显著的统计学意义。因此，我们选择三次曲线模型对云南省能源消费总量进行预测。

由表 1 中的数据我们可以得到以时间  $t$  为自变量，云南省能源消费总量  $y_t$  为因变量的三次曲线模型表达式为：

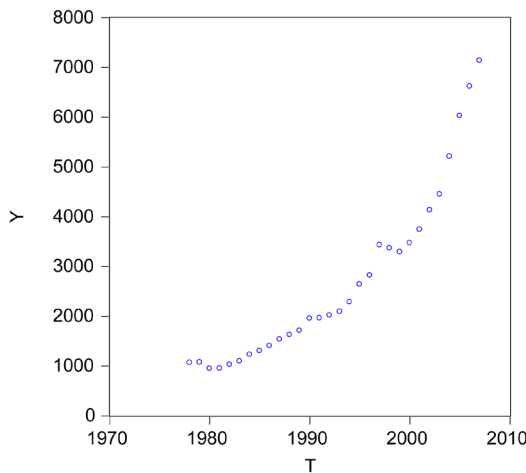


Figure 1. Scatter graph of energy consumption  
图 1. 能源消费总量的散点图

Table 1. The estimation of curve about energy consumption

表 1. 能源消费总量的曲线估计结果

方程	R 方	F	df1	df2	Sig.	常数	b1	b2	b3
二次	0.974	515.661	2	27	0	1293.945	-84.407	8.672	
三次	0.986	615.804	3	26	0	694.035	130.407	-8.371	0.367
指数	0.978	1247.631	1	28	0	770.441	0.069		

$$y = 0.367t^3 - 8.371t^2 + 130.407t + 694.035 \quad (5)$$

1978~2007年云南省能源消费总量的实际值、三次曲线模型的拟合值以及它们的残差趋势图,见图2。

3) 进一步说明三次曲线模型的拟合效果

由表1还可得二次模型和指数模型的表达式,分别为:

二次模型:

$$g(t) = 8.672t^2 - 84.407t + 1293.945 \quad (6)$$

指数模型:

$$h(t) = 770.441e^{0.069t} \quad (7)$$

基于模型(5)、(6)、(7)以及时间 $t$ 可预测1978~2007年云南省能源消费总量值,然后结合历史数据,根据公式:

$$\text{相对误差绝对值} = \frac{|\text{预测值} - \text{真实值}|}{\text{真实值}} \times 100\% \quad (8)$$

计算出各个模型每年的相对误差绝对值,既而可以求得各个模型的平均相对误差绝对值,结果见表2。

由表2明显发现三次曲线模型的平均相对误差绝对值 $MAPE1 = 6.55\%$ ,是三个模型中最小的,即 $MAPE1 < MAPE3 < MAPE2$ 。综上分析知:三次曲线模型对云南省能源消费总量的拟合预测效果优于二次和指数模型。因此,我们选择三次曲线模型作为云南省能源消费总量的预测模型。

### 3.2.2. 建立混合时间序列模型

趋势外推模型产生预测误差的主要原因是它只反映了能源消费总量随时间变化的规律,并没有解释其趋势值与实际值的残差序列。所以为了使云南省能源消费总量的预测精度提高,下面采用ARIMA模

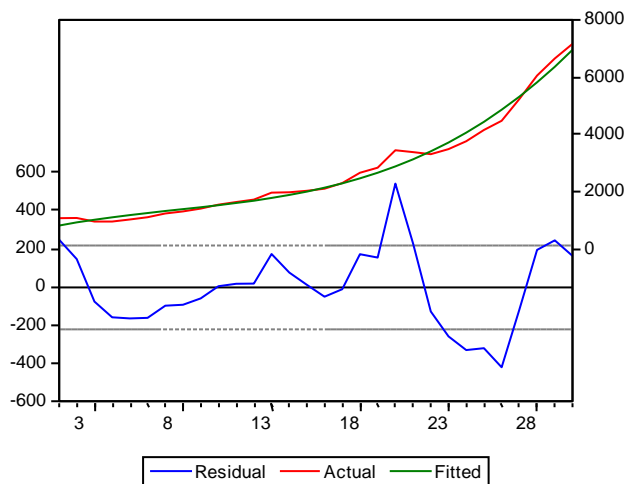


Figure 2. Actual, fitted, residual graph  
图2. 实际值、三次曲线模型拟合值及残差趋势图

Table 2. The absolute value of the average relative error  
表2. 平均相对误差绝对值

方程	三次	二次	指数
平均相对误差绝对值(%)	6.55	8.60	7.11

型对残差项进行预测，最后把趋势外推模型与 ARIMA 模型的预测值叠加起来作为能源消费总量的预测结果。

1) 对残差序列建立 ARIMA 模型

①运用 Eviews 软件画出 1978-2007 年云南省能源消费总量的实际值与拟合值之间的残差图，如图 3 所示。

根据图 3，我们可初步判断残差序列是平稳的，因为图形表现出该残差序列围绕其均值不断波动。这种直观的图示也常产生误导，因此需要进行进一步的判别，下面通过 ADF 单位根检验法，进一步判断其平稳性。

②对残差序列做 ADF 单位根检验，确定其是否平稳

使用 Eviews 软件对筛选出的趋势外推模型即三次曲线模型的残差序列做 ADF 单位根检验，检验结果见表 3。

根据上表，可以发现：在  $\alpha = 0.01$  的水平下，临界值为  $-2.65$ ； $\alpha = 0.05$  的水平下，临界值为  $-1.95$ ； $\alpha = 0.1$  的水平下，临界值为  $-1.61$ ；而统计量的值为  $-3.28$ 。无论在何种水平下，统计量的值都是小于临界值的，因此根据 ADF 单位根检验的判断准则可知，拒绝零假设“存在一个单位根”，故该残差序列不存在单位根，是平稳的时间序列，则可以直接对该残差序列建立 ARMA 模型并作预测。

③建立 ARMA 模型

应用 Eviews 软件可得到残差序列的自相关和偏自相关图以及 ACF 和 PACF 表。如图 4 所示。

根据上图自相关图和偏自相关图的特点，不能确定 ARMA( $p, q$ ) 中  $p$  和  $q$  究竟取几阶所建立的模型拟合效果最好，因此我们尝试建立几个可能的 ARMA( $p, q$ ) 模型，比如：ARMA(2,2)、ARMA(2,1)、ARMA(1,2)、ARMA(1,1)。利用 Eviews 软件分别对上面四种可能的 ARMA 模型进行估计，所得结果整理见表 4。

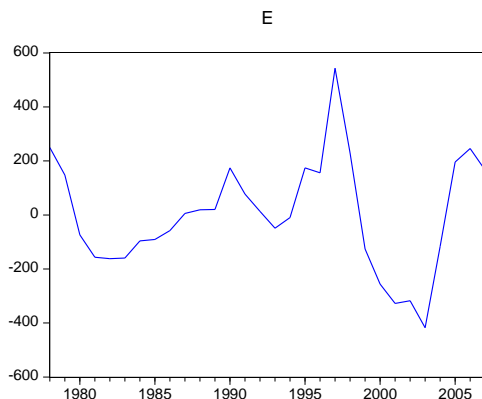


Figure 3. Residual graph  
图 3. 残差图

Table 3. ADF unit root test  
表 3. ADF 单位根检验结果

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.282076	0.0019
Test critical values:		
1% level	-2.650145	
5% level	-1.953381	
10% level	-1.609798	

观察表,我们发现:四种可能的 ARMA 模型中,ARMA(2,2)模型的残差平方和最小,但是其 AIC 值比 ARMA(2,1)的要大,而且 MA(2)的系数不显著;ARMA(2,1)这个模型的 AIC 值是四个可能模型中最小的,而且各变量的系数都通过了显著性检验;另外,ARMA(1,2)和 ARMA(1,1)模型不仅残差平方和与 AIC 值比其它两种模型都要大,个别变量的系数也不显著。根据 AIC 准则可知,当选择模型时,AIC 值越小越好,因此建立 ARMA(2,1)模型相对比较合适。

使用 Eviews 软件对 ARMA(2,1)模型估计参数,所得结果见表 5。

由表中的信息,可得如下 ARMA(2,1)模型:

$$(1-1.54B+0.78B^2)\mu_t = (1-0.95B)\varepsilon_t \tag{9}$$

④模型检验

ARMA(p,q)模型的识别与估计是在随机干扰项  $\varepsilon_t$  是一个白噪声的基础上进行的,因此,如果估计的模型确认正确的话,残差  $\hat{\varepsilon}_t$  应代表一个白噪声序列。ARMA(2,1)模型对残差  $\mu_t$  的拟合见图 5。

从图 5 可以初步判断  $\varepsilon_t$  是平稳的,基本接近白噪声,此外模型的拟合效果整体上还不错。下面运用  $\chi^2$  检验进一步说明  $\varepsilon_t$  是否是白噪声序列。 $Q_{LB}$  统计量[13]:

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.657	0.657	14.274	0.000
		2	0.198	-0.410	15.618	0.000
		3	-0.183	-0.206	16.813	0.001
		4	-0.409	-0.169	23.002	0.000
		5	-0.453	-0.105	30.886	0.000
		6	-0.392	-0.167	37.026	0.000
		7	-0.122	0.218	37.648	0.000
		8	0.046	-0.260	37.742	0.000
		9	0.103	-0.067	38.232	0.000
		10	0.064	-0.141	38.429	0.000
		11	-0.041	-0.138	38.514	0.000
		12	-0.103	-0.112	39.079	0.000
		13	-0.162	-0.121	40.559	0.000
		14	-0.134	-0.168	41.630	0.000
		15	-0.063	-0.143	41.887	0.000
		16	0.030	-0.126	41.950	0.000

Figure 4. The autocorrelation and partial autocorrelation coefficient of residual in the model  
图 4. 模型(5)的残差序列自相关和偏自相关图及系数

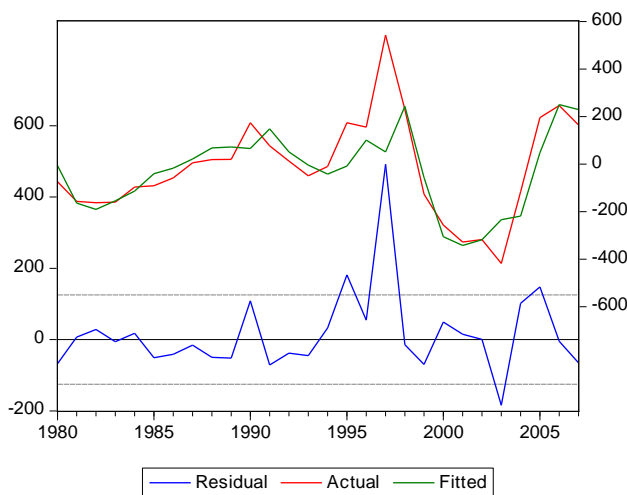
Table 4. The estimation of parameters in the model ARMA  
表 4. 模型相关参数估计

Model	R-squared	Sum squared resid	AIC	P			
				AR(1)	AR(2)	MA(1)	MA(2)
ARMA(2,2)	0.66983	377233.1	12.70343	0.0000	0.0001	0.0025	0.8970
ARMA(2,1)	0.656459	392510.4	12.60028	0.0000	0.0000	0.0000	
ARMA(1,2)	0.586374	483078.7	12.83438	0.7526		0.0000	0.0000
ARMA(1,1)	0.523096	556982.1	12.90777	0.0272		0.1187	



**Table 5.** The estimation of parameters in the model ARMA(2,1)  
**表 5.** ARMA(2,1)模型相关参数估计

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	1.535694	0.116532	13.17828	0.0000
AR(2)	-0.777645	0.118049	-6.587454	0.0000
MA(1)	-0.954865	0.073601	-12.97348	0.0000
R-squared	0.656459	Mean dependent var		-14.19170
Adjusted R-squared	0.628976	S.D. dependent var		205.7096
S.E. of regression	125.3013	Akaike info criterion		12.60028
Sum squared resid	392510.4	Schwarz criterion		12.74301
Log likelihood	-173.4039	Hannan-Quinn criter.		12.64391
Durbin-Watson stat	1.854546			



**Figure 5.** The estimation of parameters in the model ARMA (2, 1)  
**图 5.** ARMA(2,1)模型对残差  $\mu_t$  的拟合图

$$Q_{LB} = n(n+2) \sum_{k=1}^m \frac{\gamma_k^2}{n-k} \tag{10}$$

其中  $n$  为样本数， $m$  为滞后长度， $\gamma_k$  为滞后  $k$  阶样本自相关系数。将  $n = 30$ ， $m = 16$  代入公式(10)，经计算得到  $Q_{LB} = 1.40$ ，没有超过 5% 显著性水平的临界值 27.58。因此可以接受所有的自相关系数都为 0 的假设，即  $\varepsilon_t$  为白噪声序列，模型可用于预测。

2) 混合时间序列模型

综合上面的分析，得到能源消费总量的混合时间序列预测模型为：

$$\begin{cases} y_t = 0.367t^3 - 8.371t^2 + 130.407t + 694.035 + \mu_t \\ (1 - 1.54B + 0.78B^2)\mu_t = (1 - 0.95B)\varepsilon_t \end{cases} \tag{11}$$

3.2.3. 趋势外推模型与混合时间序列模型预测效果比较

使用趋势外推模型和混合时间序列模型对云南省 1978~2007 年的能源消费总量分别进行预测，得到

它们的预测值，通过历史真实值和预测值，可以分别计算出两种预测模型的平均相对误差绝对值。经计算得到趋势外推模型的平均相对误差绝对值为  $MAPE_1 = 6.55\%$ ，混合时间序列模型的平均相对误差绝对值为  $MAPE_4 = 5.74\%$ ，从而有  $MAPE_1 > MAPE_4$ ，所以混合时间序列模型的预测效果优于趋势外推模型。

用所建立的趋势外推模型和混合时间序列模型分别对 2008~2011 年云南省能源消费总量进行预测拟合，所得预测结果见表 6。

由表 6 我们发现：运用趋势外推模型进行预测，2008~2010 年模型的相对误差绝对值都比混合时间序列模型的大，趋势外推模型和混合时间序列模型的平均相对误差绝对值分别为 3.30% 和 3.26%，前者大于后者。由此进一步说明，混合时间序列模型的预测精度高于趋势外推模型。

#### 4. 混合时间序列模型对未来云南省能源消费总量的预测应用

混合时间序列模型不仅在训练集(1978~2007 这 30 年的云南省能源消费总量数据集)中的拟合预测结果比趋势外推模型更准确，同时在测试集(2008~2011 这 4 年的数据集[14])中表现也一样好。因此，我们可以利用混合时间序列模型对未来云南省能源消费总量进行准确的预测，2014~2018 年云南省能源消费总量预测结果见表 7。

#### 5. 结论

通过上面的实证分析，我们发现：趋势外推模型与 ARIMA 模型构成的混合时间序列模型对云南省能源消费总量进行预测，其预测结果比单一趋势外推模型的准确性高，而且预测结果的稳定性也好。因为混合时间序列模型既包含了时间序列中的趋势分量，此分量可由时间变量解释说明，又包含了残差序列，此序列不能被时间变量解释但可由 ARIMA 模型解释，故 ARIMA 模型弥补了单一趋势外推模型的缺陷。所以，混合时间序列模型的预测精度要高于趋势外推模型。

能源是人类赖以生存、人民生活水平不断提高以及社会得以发展、进步的重要物质基础保障，因此，对未来的能源消费情况进行合理、准确、科学的预测意义重大。只有对能源消费总量进行准确的预测，才能为能源问题找到有效的解决办法。运用混合时间序列模型对云南省能源消费总量进行更为准确、合理的预测，这不仅给云南省相关部门制定科学合理的能源战略规划和健全的能源消费政策提供依据和借鉴，同时也有利于维护云南省和我国国民经济健康、持续、稳定的发展，建设节约型和谐社会。所以，研究混合时间序列预测模型，无论在理论上还是实际上都具有重要的意义。

**Table 6.** The prediction, actual, absolute value of the relative error from 2008 to 2011  
**表 6.** 2008~2011 年预测值和真实值以及相对误差绝对值

年份	实际值	趋势外推模型 预测值	相对误差 绝对值(%)	混合时间序列 模型预测值	相对误差绝对值(%)
2008	7510.82	7625.40	1.53%	7617.53	1.42%
2009	8032.06	8321.00	3.60%	8315.95	3.53%
2010	8674.17	9070.30	4.57%	9068.66	4.55%
2011	9540.28	9875.60	3.51%	9877.01	3.53%

**Table 7.** Energy consumption prediction in Yunnan Province from 2014 to 2018  
**表 7.** 2014~2018 年云南省能源消费总量预测结果

年份	2014	2015	2016	2017	2018
能源消费总量预测值(万吨标准煤)	12652.8	13702.51	14818.94	16004.49	17261.48

## 致 谢

本文在云南财经大学统计与数学学院石磊教授的指导和帮助下，得以顺利完成，在此对石磊教授表示衷心的感谢！

## 参考文献 (References)

- [1] Ediger, V.S. and Akar, S. (2007) ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy*, **35**, 1701-1708.
- [2] Kumar, U. and Jain, V.K. (2010) Time series models to forecast energy consumption in India. *Energy*, **35**, 1709-1716.
- [3] 韩君, 梁亚民 (2005) 趋势外推与 ARMA 组合的能源需求预测模型. *兰州商学院学报*, **21**, 92-95.
- [4] 程静, 郑定成, 吴继权 (2010) 基于时间序列 ARMA 模型的广东省能源需求预测. *能源工程*, **1**, 1-5.
- [5] 徐明德 (1995) 灰色 GM(1,1)模型及其在能源预测中的应用. *应用能源技术*, **2**, 5-7.
- [6] 孙文生 (2011) 基于灰色理论与 BP 神经网络的煤炭需求预测. *科技传播*, **8**, 87-88.
- [7] 柴元春 (2011) 我国能源总产量的组合预测. *赤峰学院学报(自然科学版)*, **21**, 71-73.
- [8] 王惠婷 (2013) 基于混合时间序列模型的粮食产量预测. *统计与决策*, **12**, 24.
- [9] 张志启 (2012) 我国能源消费总量预测建模分析. 硕士论文, 内蒙古科技大学, 包头.
- [10] 王燕 (2008) 应用时间序列分析. 第 2 版, 中国人民大学出版社, 北京.
- [11] 郭海明, 李喆 (2006) 混合时间序列模型在经济增长预测中的应用. *西北民族大学学报*, **62**, 89.
- [12] 云南省统计局 (2011) 云南省统计年鉴 - 2011. 中国统计出版社, 北京.
- [13] 李子奈, 潘文卿 (2010) 计量经济学. 第 3 版, 高等教育出版社, 北京.
- [14] 云南省统计局 (2012) 云南省统计年鉴 - 2012. 中国统计出版社, 北京.