

文章编号: 2095-2163(2019)04-0033-07

中图分类号: TM715;F426.61;F224

文献标志码: A

基于 Shapley 组合模型的上海市电力需求趋势预测

王涛

(上海理工大学 管理学院, 上海 200093)

摘要: 现代社会中,家家户户都离不开电力的供应,稳定可靠的电力供应对于经济、社会的发展以及人民生活水平的提高起着重要作用。上海市作为中国重要的经济和贸易中心,电力需求一直保持着高速增长,对于电力供应单位而言,需要对用户的电力需求进行科学预测,以满足用户的用电需求,稳定电力供应,助力经济社会发展。本文采用 ARIMA 模型、BP 神经网络模型和 Holt 指数平滑模型分别对 1995~2017 年上海市电力需求数据进行建模分析,并运用合作对策理论中的 Shapley 组合模型方法进行组合建模,求解了上述三种单一预测方法的权重,构建了组合模型,并依据组合模型预测了未来五年的上海市电力需求量趋势。根据实例结果,Shapley 组合模型具备更高的预测精度,预测平均相对误差仅为 2.33%,拟合效果较好,有利于电力供应单位稳定电力供应,促进经济社会发展。

关键词: ARIMA 模型; BP 神经网络; Holt 指数平滑法; Shapley 组合模型; 电力需求预测

Research on Shanghai electricity demand trend forecast based on Shapley combination model

WANG Tao

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

[Abstract] In modern society, people cannot live without electricity supply. Stable and reliable electricity supply plays an important role in economic and social development and helps to improve people's living standards. As an important economic and trade center in China, Shanghai's power demand has been growing at a high speed. For power supply companies, it is necessary to forecast the power demand of users scientifically in order to meet the power demand of users, stabilize the power supply and help economic and social development. In this paper, ARIMA model, BP neural network model and Holt exponential smoothing model are used to model and analyze the power demand data of Shanghai from 1995 to 2017 respectively, and Shapley combination model method in cooperative countermeasure theory is used to carry out combination modeling. In this paper, the weights of the above three single prediction methods are solved, the combination model is constructed, and the power demand trend of Shanghai in the next five years is predicted according to the combination model. According to the results, Shapley combination model has higher prediction accuracy, the average relative error of prediction is only 2.33%, and the fitting effect is good, which is conducive to the stable power supply of power supply companies and the promotion of economic and social development.

[Key words] ARIMA model; BP neural network; Holt exponential smoothing method; Shapley combination model; electricity demand forecast

0 引言

随着现代社会的高速发展,电力资源作为必备的重要能源,对于促进社会发展、提高人民物质生活水平起着至关重要的作用。现代社会中,家家户户都离不开电力的供应,稳定可靠、价格合适的电力供应成为了人们关注的焦点。因此,电力供应单位需要准确地制定供电计划,确保对用电量的准确预测以及电力的稳定供应。上海市作为中国重要的经济和贸易中心,自改革开放以来经济一直保持着良好增长态势,人民生活水平显著提高,物质生活获得了极大的丰富,由此也带来了用电量的激增,这对供电单位而言同样也是极大的考验。因而选择科学、有

效的预测模型,准确地对市场中的用电需求进行预测,有助于供电单位制定供电计划,稳定电力供应,助力社会经济的高速平稳运行,不断提高人民生活质量。

近年来,国内外学者对于预测方法的研究有很多,也取得了可观成果^[1-17]。目前,常用的预测方法主要包括系统动力学方法^[1]、随机森林法^[2-3]、支持向量机法^[4-5]、神经网络法^[6-9]、灰色预测法^[10-13]、回归分析法^[14-17]等。文献[1]从系统动力学的角度研究了城市长期电力需求的预测方法,在经济新常态的背景下,城市支柱产业面临着由要素驱动向创新驱动的转变,因而传统的电力需求预测模型逐渐不再适用,文章结合新的影响因素进行了定量分析,

作者简介: 王涛(1993-),男,硕士研究生,主要研究方向:电力需求预测分析。

收稿日期: 2019-05-13

提出了一种适用于经济新常态的长期电力需求预测模型的动力学方法,并取得了较好的效果。文献[9]基于BP神经网络算法,研究了北京市夏季日最大电力负荷的预测方法。文献[16]基于2009~2013年北京市电力负荷和气象资料,通过多元回归的思想构建了北京市日最大电力负荷的预测模型,并取得了较好的预测准确率,展现了良好的预报能力。文献[17]将自回归积分滑动平均模型(ARIMA模型)和支持向量机法(SVM)相结合,基于ARIMA和SVM混合模型对短期电力负荷进行了预测。然而目前大多数的预测文献中采用了单一的预测方法,往往偏重于数据的某一类特征,无法充分利用数据的各种有效信息,从而导致预测的误差较大,未能得到令人满意的预测精度和效果。但是研究可知组合预测模型可以克服单一模型的局限性,最大程度地利用数据的各种有效信息,因而组合预测模型方法更加有利于实现科学的预测,广泛适用于信息量不完备的复杂经济系统。一般而言,电力需求量数据序列为非平稳时间序列,可以通过差分的方法转换为平稳时间序列,进而采用ARIMA模型进行建模预测分析。同时,BP神经网络也是一种常用的预测方法,具有高度的信息处理能力,在处理数据的过程中表现出良好的自学习、自适应能力,因而也尤为突显出该方法独特的优越性;指数平滑法则是一种较简单的预测方法,常适用于光滑、且显出一定趋势的数据,在计算的复杂程度和适应性上占据着优势。

在上述分析基础上,本文拟将采用组合ARIMA模型、BP神经网络模型和指数平滑模型,基于Shapley组合预测模型对上海市电力需求趋势展开预测分析,并给出未来五年上海市电力需求的预测结果。对此可得研究论述如下。

1 ARIMA 模型建模

ARIMA模型,又称自回归积分滑动平均模型,是一种经典的时间序列计量模型,主要用于对时间序列的宏观趋势预测研究^[18]。ARIMA模型通过把非平稳时间序列转化为平稳时间序列,将因变量对其滞后和随机误差项的当前值与滞后进行回归所建立的模型,可以有效地描述时间序列数据的动态连续特性,适用于解决非平稳时间序列趋势预测问题。在ARIMA(p, d, q)模型中,AR称为自回归, p 表示自回归的阶数,MA表示移动平均, q 为移动平均的阶数, d 为将非平稳时间序列转化为平稳时间序列的差分阶数。ARIMA模型的基本结构如下:

$$\begin{cases} \Phi(B) \tilde{N}^d X_t = \Theta(B) \varepsilon_t; \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t; \\ E(X_s \varepsilon_t) = 0, s < t. \end{cases} \quad (1)$$

其中, $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ 表示平稳可逆ARMA(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ARMA(p, q)模型的滑动系数多项式; $\{\varepsilon_t\}$ 为白噪声过程。

对于任意非平稳时间序列,均可以通过差分运算,转化为平稳时间序列,继而进行ARMA模型的拟合。对于时间序列建立ARIMA模型的步骤可分述如下。

(1)时间序列的平稳性检验。根据获得的时间序列,进行ADF单位根检验,依据结果判断是否为平稳时间序列。

(2)将非平稳时间序列转化为平稳序列。对于非平稳时间序列,运用差分运算的方法,转化为平稳时间序列。

(3)识别模型并判断模型阶数。对于平稳时间序列,依据自相关与偏自相关图初步分析模型的阶数 p, q ,并依据AIC或BIC准则最终确定其数值。

(4)模型参数检验。对模型残差进行白噪声检验。

(5)模型拟合预测。依据模型,对时间序列进行预测。

根据上述步骤,对上海市电力需求量时间序列进行ARIMA模型建模。对此,可推得研发阐述如下。

1.1 时间序列的平稳性检验

根据1995~2017年上海市电力需求量历史数据(文中所用数据来自国家统计局官方数据库),绘制1995~2017年上海市电力需求量时序图如图1所示。

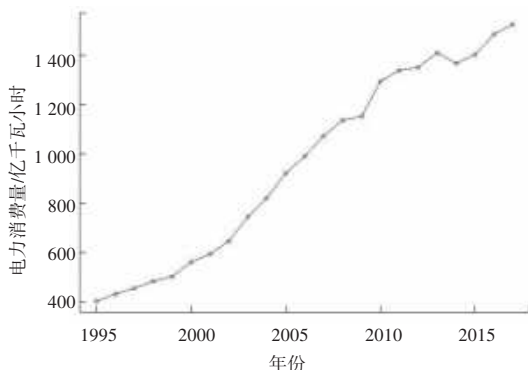


图1 上海市电力需求量趋势曲线图

Fig. 1 Shanghai power demand trend curve

观察图 1 时序可知,上海市电力需求量呈现明显的上升趋势,分析可知这不是一个平稳序列。为了进一步确定上海市电力需求序列的平稳性,运用 ADF 单位根检验的方法进行判断,运算结果见表 1。

表 1 ADF 单位根检验结果
Tab. 1 ADF unit root test results

ADF 统计量	临界值			
	1%置信水平	5%置信水平	10%置信水平	
T 统计量	0.163 505	- 4.440 739	- 3.632 896	- 3.254 671
P 值	0.995 900	-	-	-

根据 ADF 单位根检验结果, P 值为 0.995 900, 不能拒绝原假设,上海市电力需求量序列不是一个平稳的时间序列。因此需要将数据进行差分平稳性处理。

1.2 差分运算处理

对上海市电力需求量时间序列进行一阶差分处理。为了判断一阶差分后该序列是否平稳,绘制一阶差分后序列的自相关图和偏自相关图如图 2 所示。

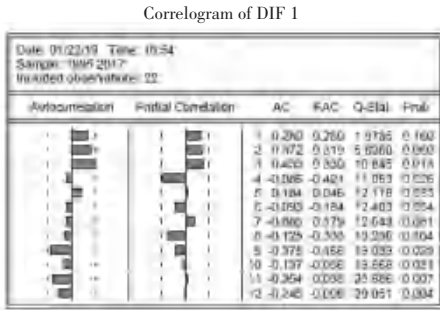


图 2 一阶差分后序列自相关图和偏自相关图

Fig. 2 First-order differential sequence autocorrelation graph and partial autocorrelation graph

由图 2 可知,初步判断一阶差分后序列为平稳序列。为了保证判断的正确性,同理,可以通过 ADF 单位根检验的方法对差分后的时间序列进行判断。预算结果见表 2。

表 2 一阶差分序列 ADF 单位根检验结果

Tab. 2 First-order differential sequence ADF unit root test result

ADF 统计量	临界值			
	1%置信水平	5%置信水平	10%置信水平	
T 统计量	- 3.865 812	- 4.467 895	- 3.644 963	- 3.261 452
P 值	0.032 900	-	-	-

由 ADF 单位根检验结果, P 值 0.032 900 < 0.05, 拒绝原假设,可以认为一阶差分之后为平稳序列。因此,可以考虑采用 ARIMA(p, 1, q) 模型。

基于准确性原则和数据最优化考虑,对一阶差

分后的时间序列再做一次差分运算。为了判断二阶差分后该序列是否平稳,绘制二阶差分后序列的自相关图和偏自相关图如图 3 所示。

根据图 3 可初步判断二阶差分后序列为平稳序列。为了保证判断的正确性,同理通过 ADF 单位根检验的方法对差分后的时间序列进行判断。结果见表 3。

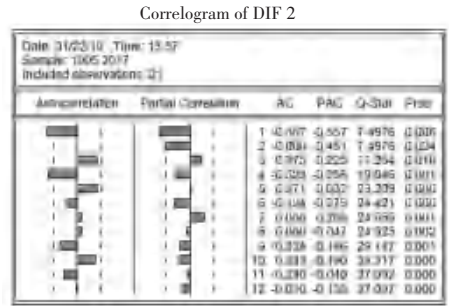


图 3 二阶差分后序列自相关图和偏自相关图

Fig. 3 Second-order differential sequence autocorrelation graph and partial autocorrelation graph

表 3 二阶差分序列 ADF 单位根检验结果

Tab. 3 Second-order differential sequence ADF unit root test result

ADF 统计量	临界值			
	1%置信水平	5%置信水平	10%置信水平	
T 统计量	- 5.905 598	- 2.692 358	- 1.960 171	- 1.607 051
P 值	0.000 000	-	-	-

由表 3 分析可知,二阶差分之后序列即为平稳序列。因此,可以考虑 ARIMA(p, 2, q) 模型。

1.3 识别模型,判断模型阶数

根据上述 ADF 单位根检验的结果,可以看出上海市电力需求数据在一阶差分和二阶差分下 ADF 检验 P 值均小于 0.05,即差分后的时间序列均为平稳序列,参数 d 的选取需要考虑 1 与 2 两个值。

在 ARIMA 模型中除了考虑差分阶数 d 的选择,还需要考虑参数 p 与 q 的选择,时间序列的自相关系数(ACF)与偏自相关系数(PACF)可以提供对于参数 p, q 的判断方法。若自相关函数在 q 阶截尾和偏自相关系数在 p 阶截尾,则可初步判定模型为 ARIMA(p, d, q)。

通过观察自相关函数与偏自相关函数图, p 值可以选取 0, 1, 2, q 值可以选取 1, 2。根据上述结论,初步设定的模型为 ARIMA(0, 1, 1)、ARIMA(0, 2, 1)、ARIMA(1, 2, 1)、ARIMA(2, 2, 1)、ARIMA(0, 2, 2)、ARIMA(1, 2, 2)、ARIMA(2, 2, 2)。由于上述模型的确定主要是根据自相关图和偏自相关图定性判

断得出,为了精准确定最优模型参数,可以使用 *AIC* 和 *BIC* 准则进行判断。一般认为,*AIC* 和 *BIC* 函数值较小的模型参数有着较高的准确性,更加接近真实模型的阶数。使用 R 语言对上述七种模型进行比较,各模型的 *AIC* 和 *BIC* 准则值见表 4。

表 4 各模型参数对比表

Tab. 4 Comparison table of models parameter

模型	Likelihood	AIC	BIC
ARIMA(0,1,1)	50.75	-97.50	-95.31
ARIMA(0,2,1)	55.91	-107.82	-105.73
ARIMA(1,2,1)	56.30	-106.60	-103.46
ARIMA(2,2,1)	59.84	-111.68	-107.50
ARIMA(0,2,2)	56.41	-106.81	-103.68
ARIMA(1,2,2)	56.43	-104.86	-100.67
ARIMA(2,2,2)	60.08	-110.15	-104.93

根据 *AIC* 和 *BIC* 准则对上述模型进行比较,发现模型 ARIMA(2,2,1) 结果较好。选取 ARIMA(2,2,1) 模型为最优模型,序列 $\{x_1, x_2, \dots, x_n\}$ 的拟合结果为:

$$x_t = -1.322 2x_{t-1} - 0.734 8x_{t-2} + \varepsilon_t + 0.891 0\varepsilon_{t-1}. \quad (2)$$

1.4 模型参数检验

对 ARIMA(2,2,1) 模型残差序列进行白噪声检验,得到 p 值为 0.959 $6 > 0.05$,表明残差序列是纯随机的,为白噪声序列,说明模型显著,拟合效果较好,ARIMA(2,2,1) 模型对该时间序列拟合成功。

1.5 模型预测

应用 ARIMA(2,2,1) 模型对上海市 2018~2022 年电力需求量进行短期预测,结果见表 5。

表 5 ARIMA 模型预测结果

Tab. 5 ARIMA model prediction results

年份/年	2018	2019	2020	2021	2022
预测值/ 亿千瓦小时	1 602.94	1 670.21	1 730.1	1 816.2	1 881.37

2 BP 神经网络

根据 1995~2017 年上海市电力需求量的历史数据,选择每前三年的数据来预测第四年的数据。即在 BP 神经网络中,输入层选择 3 个神经元,输出层选择 1 个神经元,根据文献[9],在隐含层的节点应为输入层节点的 2~3 倍左右搜索,经过多次实验模拟,测试发现在输入层节点为 3,输出层节点为 1 的情况下,隐含层节点数为 10 时,BP 神经网络的拟合效果较好。因此可按照如下参数来训练 BP 神经

网络建模,即:输入层节点数为 3,隐含层节点数为 10,输出层节点数为 1,训练次数为 3 000 次,神经网络训练误差设为 0.000 1,网络的初始权值/学习速度和阈值等由 Matlab 软件自动选取。

研究中,在保证训练次数和训练误差的前提下,应用 BP 神经网络模型对上海市 2018~2022 年电力需求量进行短期预测,结果见表 6。

表 6 BP 神经网络模型预测结果

Tab. 6 Prediction results of BP neural network model

年份/年	2018	2019	2020	2021	2022
预测值/ 亿千瓦小时	1 617.24	1 709.48	1 768.06	1 795.53	1 803.72

3 Holt 指数平滑法

指数平滑法,是一种基于预测对象历史数据利用加权平均的方法进行数据的拟合与预测的方法^[19]。根据上述关于 1995~2017 年上海市电力需求量的变化趋势图可以看出,上海市电力需求呈现出一定程度的线性变化趋势,具有水平截距和线性趋势斜率,无季节因素影响。基于指数平滑理论,Holt 指数平滑法可以对有水平项和趋势项的时间序列实现拟合预测,本文采取 Holt 指数平滑的方法对上海市电力需求量变化趋势进行预测。这里,研究推得 Holt 指数平滑法的计算模型的数学公式可顺次表述如下。

对于时间序列 $\{x_1, x_2, \dots, x_n\}$, Holt 指数平滑后的序列 \hat{x}_t 的计算公式为:

$$\hat{x}_{t+k} = a(t) + k \cdot b(t), \quad (3)$$

其中, a 为水平截距, b 为趋势或斜率。这 2 个参数由下列递推公式计算得到:

$$a_{t+1} = \alpha x_{t+1} + (1 - \alpha)(a_t + b_t), \quad (4)$$

$$b_{t+1} = \beta(a_{t+1} - a_t) + (1 - \beta)b_t.$$

其中, α 和 β 在 (0,1) 之间,均为模型参数。

根据前两期的数据作为初值,即:

$$a_1 = x_1, b_1 = x_2 - x_1, \quad (5)$$

研究可知,R 语言的 forecast 包主要用于时间序列的拟合预测。其中,函数 ets() 主要用于关于指数预测模型的数据拟合,而且还会依据平均相对误差最小准则,自动选取对原始数据拟合优度最高的模型,以上海市电力需求量历史数据的指数模型拟合为例,可得拟合优度最高的 Holt 指数平滑模型。对应模型参数有:

$$\begin{cases} \alpha = 0.846 7, \\ \beta = 0.377 5. \end{cases} \quad (6)$$

应用 Holt 指数平滑模型对上海市 2018~2022 年电力需求量进行短期预测,运算结果见表 7。

表 7 Holt 指数平滑模型预测结果

Tab. 7 Holt exponential smoothing model prediction results

年份/年	2018	2019	2020	2021	2022
预测值/ 亿千瓦时	1 568.72	1 612.12	1 655.51	1 698.91	1 742.31

4 基于 Shapley 组合模型法的上海市电力需求预测

4.1 Shapley 方法基本思路

Shapley 组合法是由 Shapley 在 1953 年提出的用于解决多参与者合作对策问题的一种方法。该方法针对各个参与者在合作中的不同贡献,着眼于每个参与者对合作的平均贡献,突出了合作中每个参与者的的重要性,因而易于为合作的各个参与者所接受,并获得了较好的应用^[20-21]。

在 Shapley 组合法中,假设存在 s 种预测方法,集合为 $I = \{1, 2, \dots, s\}$, 对其中任意子集 $p, q, E(p), E(q)$ 分别为该预测子集的误差。则有:

(1) 对于子集 p, q , 有 $E(p) + E(q) \geq E(p \cup q)$ 。

(2) 设 y_i 为第 i 种预测方法在合作中最终分摊的误差值, 有 $y_i \leq E(i)$, 其中 $E(i)$ 表示第 i 种预测方法产生的误差。

(3) 对于 s 种预测方法基于 Shapley 法组合模型产生的总误差 $E(s)$, 有 $E(s) = \sum_{i \in s} y_i$ 。

假设第 i 种预测方法产生的相对误差的均值为 E_i , 组合模型的总误差为 E , 则有:

$$E_i = \frac{1}{m} \sum_{j=1}^m |e_{ij}|, (i = 1, 2, \dots, s), E = \frac{1}{s} \sum_{i=1}^s E_i, \quad (7)$$

其中, m 表示样本的个数, $|e_{ij}|$ 表示第 i 种预测方法中第 j 个样本的相对误差。

在 Shapley 组合法中, 各个合作参与者的权重分配可参考如下计算公式:

$$E_i = \sum_{i \in p} \omega(|p|) [E(p) - E(p \setminus i)], i = 1, 2, \dots, s, \quad (8)$$

$$\omega(|p|) = \frac{(s - |p|)! (|p| - 1)!}{s!}, \quad (9)$$

其中, $\omega(|p|)$ 为加权系数, 表示在 Shapley 组合中参与者 i 承担的组合同边际贡献; i 表示参与到 Shapley 组合中的某个预测模型; $(p \setminus i)$ 表示在 Shapley 组合中除去模型 i ; E_i 表示参与的预测模型

i 分配的误差, 即 Shapley 值; p 表示在集合 I 中的任意子集; $|p|$ 表示子集中预测模型的个数。

根据上式, 关于在 Shapley 组合预测中, 参与其中的各个预测方法的权重计算将用到如下公式:

$$\omega_i = \frac{1}{s - 1} \cdot \frac{E - E_i}{E}, i = 1, 2, \dots, s. \quad (10)$$

其中, s 表示预测方法的个数。

4.2 基于 Shapley 方法的上海市电力需求量预测

使用均方误差 (MAPE) 准则对上述三种单一预测方法, 即: ARIMA 模型、BP 神经网络模型和 Holt 指数平滑模型的预测误差进行计算分析。根据上述三种单一预测方法的预测误差, 可求得组合预测的总误差值为 $(2.49 + 2.99 + 2.79) / 3 = 2.76$ 。由 Shapley 方法, 参与组合预测模型总误差分配的“合作关系”成员为: $I = \{1, 2, 3\}$, I 所有子集的组合同误差分别为: $E\{1\}$ 、 $E\{2\}$ 、 $E\{3\}$ 、 $E\{1, 2\}$ 、 $E\{1, 3\}$ 、 $E\{2, 3\}$ 、 $E\{1, 2, 3\}$, 其数值的大小为该子集所包括的向量的误差平均值, 计算结果见表 8。

表 8 预测误差分配结果

Tab. 8 Prediction error distribution result

组合 $E\{1\}$	$E\{2\}$	$E\{3\}$	$E\{1, 2\}$	$E\{1, 3\}$	$E\{2, 3\}$	$E\{1, 2, 3\}$
数值 2.488 5	2.991 2	2.793 9	2.739 9	2.641 2	2.892 5	2.757 9

根据 Shapley 组合法给出的计算公式, 可以计算出组合中对应的成员 1 的 Shapley 值为:

$$\begin{aligned} E_1 &= \frac{0! \times 2!}{3!} [E\{1\} - E(\{1\} - \{1\})] + \\ &\frac{1! \times 1!}{3!} [E\{1, 2\} - E(\{1, 2\} - \{1\})] + \\ &\frac{1! \times 1!}{3!} [E\{1, 3\} - E(\{1, 3\} - \{1\})] + \\ &\frac{2! \times 0!}{3!} [E\{1, 2, 3\} - E(\{1, 2, 3\} - \{1\})] = \\ &\frac{1}{3} \times 2.488 5 + \frac{1}{6} \times (2.739 9 - 2.991 2) + \\ &\frac{1}{6} \times (2.641 2 - 2.793 9) + \frac{1}{3} \times \\ &(2.757 9 - 2.892 5) = 0.717 3, \quad (11) \end{aligned}$$

同理, $E_2 = 1.094 3$, $E_3 = 0.946 3$ 。3 个成员 Shapley 值之和为: $E_1 + E_2 + E_3 = 0.717 3 + 1.094 3 + 0.946 3 = 2.757 9$, 等于组合预测的总误差。

由权重公式(10), 可以计算出各个预测方法的权重为 $\omega = (0.370 0 \ 0.301 6 \ 0.328 4)$, 则 Shapley 组合预测模型为:

$$\hat{y}_t = 0.370 0 \hat{y}_{1t} + 0.301 6 \hat{y}_{2t} + 0.328 4 \hat{y}_{3t}. \quad (12)$$

其中, \hat{y}_t 表示上海市电力需求的 Shapley 组合模型预测值; \hat{y}_{1t} 表示 ARIMA 模型预测值; \hat{y}_{2t} 表示 BP 神经网络模型预测值; \hat{y}_{3t} 表示 Holt 指数平滑模型预测值。

由上述对于各个预测模型的权重公式, 可以对

表 9 基于 Shapley 组合模型的上海市电力需求拟合结果

Tab. 9 Shanghai electric power demand fitting results based on Shapley combined model

年份	实际电力需求量/ 亿千瓦时	单一预测模型			各单一模型相对误差/%			Shapley 法组合模型	
		ARIMA 模型	BP 神经 网络	Holt 指数 平滑模型	ARIMA 模型	BP 神经 网络	Holt 指数 平滑模型	拟合值	相对 误差/%
1998	482.94	482.45	489.39	480.85	0.10	1.33	0.43	484.01	0.22
1999	501.20	511.59	517.07	509.37	2.07	3.17	1.63	512.51	2.26
2000	559.45	528.31	534.29	526.59	5.57	4.50	5.87	529.55	5.34
2001	592.98	604.30	618.19	589.06	1.91	4.25	0.66	603.48	1.77
2002	645.71	628.61	650.88	628.28	2.65	0.80	2.70	635.22	1.63
2003	745.97	721.87	749.70	684.51	3.23	0.50	8.24	717.99	3.75
2004	821.44	804.21	924.66	797.66	2.10	12.57	2.89	838.39	2.06
2005	921.97	939.65	922.56	886.51	1.92	0.06	3.85	917.04	0.53
2006	990.15	1 027.46	989.60	996.58	3.77	0.06	0.65	1 005.90	1.59
2007	1 072.38	1 075.95	1 046.43	1 069.13	0.33	2.42	0.30	1 064.81	0.71
2008	1 138.22	1 182.55	1 135.58	1 150.91	3.89	0.23	1.12	1 157.99	1.74
2009	1 153.38	1 191.94	1 157.10	1 215.14	3.34	0.32	5.35	1 189.05	3.09
2010	1 295.87	1 225.35	1 080.15	1 218.08	5.44	16.65	6.00	1 179.17	9.01
2011	1 339.62	1 380.83	1 339.19	1 364.04	3.08	0.03	1.82	1 362.76	1.73
2012	1 353.40	1 394.94	1 342.44	1 415.66	3.07	0.81	4.60	1 385.91	2.40
2013	1 410.60	1 458.60	1 329.89	1 415.34	3.40	5.72	0.34	1 405.57	0.36
2014	1 369.03	1 392.79	1 405.12	1 462.21	1.74	2.64	6.81	1 419.31	3.67
2015	1 405.55	1 405.34	1 410.54	1 404.41	0.02	0.36	0.08	1 406.60	0.07
2016	1 486.02	1 411.97	1 444.84	1 426.84	4.98	2.77	3.98	1 426.77	3.99
2017	1 526.77	1 517.62	1 516.95	1 517.32	0.60	0.64	0.62	1 517.32	0.62
MAPE	-	-	-	-	2.49	2.99	2.79	-	2.33

根据表 9 中数据的分析比较后可知, 采用 Shapley 组合模型预测法运算得到的 1998~2017 年上海市电力需求量的预测结果的平均相对误差 (MAPE) 为 2.33%, 属于最小。所以, 可以认为基于 Shapley 组合模型预测法得到的预测结果具有较高的预测精度, 拟合效果较好, 相对单一的预测方法较为有效。

在此基础上, 拟对 2018~2022 年五年内上海市电力需求量进行预测。首先采用上述单一模型进行预测, 然后根据 Shapley 组合模型预测法, 基于各单一预测方法的权重值进行组合加权, 最终算出组合预测值。预测结果见表 10。

ARIMA 模型、BP 神经网络模型和 Holt 指数平滑模型进行组合。利用组合预测模型对 1995~2017 年上海市电力需求量进行预测, 该组合模型的预测值与误差详见表 9。

表 10 Shapley 组合模型预测结果

Tab. 10 Shapley combined model prediction results /亿千瓦时

年份	ARIMA 模型 分析 ($\omega_1 = 0.370 0$)	BP 神经 网络分析 ($\omega_2 = 0.301 6$)	Holt 指数 平滑法分析 ($\omega_3 = 0.328 4$)	Shapley 组合 预测模型
2018	1 602.94	1 617.24	1 568.72	1 596.01
2019	1 670.21	1 709.48	1 612.12	1 662.97
2020	1 730.10	1 768.06	1 655.51	1 717.05
2021	1 816.20	1 795.53	1 698.91	1 771.44
2022	1 881.37	1 803.72	1 742.31	1 812.28

5 结束语

本文首先采用 ARIMA 模型、BP 神经网络模型和 Holt 指数平滑模型三种单一预测模型分别对 1995~2017 年上海市电力需求数据进行建模分析,接着运用合作对策理论中的 Shapley 组合模型方法进行组合建模,求解了上述三种单一预测方法的权重,构建了组合模型,并依据组合模型预测了未来五年的上海市电力需求量趋势。根据实例结果,相比于三种单一的预测方法的拟合效果,Shapley 组合模型具备更高的预测精度,预测平均相对误差仅为 2.33%,拟合效果较好。

在实际应用中,根据 Shapley 组合模型的预测结果,未来 5 年内上海市的电力需求仍将呈现较大的增长态势,预计到 2022 年上海市的电力需求量为 1 812.28 亿千瓦时,该预测运算的拟合结果较好,因此可为上海市的供电单位决策提供参考。对于上海市的供电单位而言,需要制定科学合理的供电量增长规划以及供电基础设施的建设工作,稳定上海市电力供应,服务于高速发展的经济社会活动,促进人民生活水平的提高。

此外,本文主要选取了 3 种预测领域应用较多、效果较好的单一预测方法用于 Shapley 组合模型,如何更加全面有效地选择预测方法进行 Shapley 组合预测,提高应用的一般性,获得更高的拟合精度,仍然是一个亟待深入探究和讨论的研究课题。

参考文献

- [1] HE Yongxiu, JIAO Jie, CHEN Qian, et al. Urban long term electricity demand forecast method based on system dynamics of the new economic normal: The case of Tianjin [J]. *Energy*, 2017, 133(19): 9-22.
- [2] LAHOUAR A, BEN HADJ SLAMA J. Hour-ahead wind power forecast based on random forests[J]. *Renewable Energy*, 2017, 109(5): 529-541.
- [3] LAHOUAR A, BEN HADJ SLAMA J. Day-ahead load forecast using random forest and expert input selection [J]. *Energy Conversion and Management*, 2015, 103(4): 1040-1051.
- [4] WANG Xibin, WEN Junhao, ZHANG Yihao, et al. Real estate price forecasting based on SVM optimized by PSO[J]. *Optik*, 2014, 125(3): 1439-1443.
- [5] MALVONI M, DE GIORGI M G, CONGEDO P M. Data on

- support vector machines (SVM) model to forecast photovoltaic power[J]. *Data in Brief*, 2016, 9(9): 13-16.
- [6] DIMOULKAS I, MAZIDI P, HERRE L. Neural networks for GEFCom2017 probabilistic load forecasting [J]. *International Journal of Forecasting*, 2018(21): 20-24.
- [7] TEALAB A. Time series forecasting using artificial neural networks methodologies: A systematic review [J]. *Future Computing and Informatics Journal*, 2018, 3(2): 334-340.
- [8] WANG Lin, WANG Zhigang, QU Hui, et al. Optimal forecast combination based on neural networks for time series forecasting [J]. *Applied Soft Computing*, 2018, 66(6): 1-17.
- [9] 李琛,郭文利,吴进,等. 基于 BP 神经网络的北京夏季日最大电力负荷预测方法[J]. *气候与环境研究*, 2019, 24(1): 135-142.
- [10] XIE Naiming, YUAN Chaoqing, YANG Yingjie. Forecasting China's energy demand and self-sufficiency rate by grey forecasting model and Markov model [J]. *International Journal of Electrical Power & Energy Systems*, 2015, 66(5): 1-8.
- [11] XU Ning, DANG Yaoguo, GONG Yande. Novel grey prediction model with nonlinear optimized time response method for forecasting of electricity consumption in China [J]. *Energy*, 2017, 118(9): 473-480.
- [12] LI D C, CHANG C J, CHEN C C, et al. Forecasting short-term electricity consumption using the adaptive grey-based approach—An Asian case [J]. *Omega*, 2012, 40(6): 767-773.
- [13] JIN Min, ZHOU Xiang, ZHANG Z M, et al. Short-term power load forecasting using grey correlation contest modeling [J]. *Expert Systems With Applications*, 2012, 39(1): 773-779.
- [14] YUAN Chaoqing, LIU Sifeng, FANG Zhigeng. Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM (1,1) model [J]. *Energy*, 2016, 100(11): 384-390.
- [15] BARAK S, SADEGH S S. Forecasting energy consumption using ensemble ARIMA - ANFIS hybrid algorithm [J]. *International Journal of Electrical Power & Energy Systems*, 2016, 82(8): 92-104.
- [16] 石玉恒,赵娜,王凌,等. 北京地区日最大电力负荷预测模型初探[J/OL]. *中国电力*; 1-5 [2019-01-30]. <http://kns.cnki.net/kcms/detail/11.3265.TM.20190129.2238.004.html>.
- [17] NIE Hongzhan, LIU Guohui, LIU Xiaoman, et al. Hybrid of ARIMA and SVMs for Short-Term load forecasting [J]. *Energy Procedia*, 2012, 16(15): 1455-1460.
- [18] 王燕. 应用时间序列分析 [M]. 4 版. 北京: 中国人民大学出版社, 2015.
- [19] 滑瑶,张辉,刘悦悦,等. 基于 Holt-Winters 无季节指数平滑模型的国内电影票房预测 [J]. *中国传媒大学学报(自然科学版)*, 2017, 24(4): 1-7.
- [20] 王翀. 基于模型组合的我国能源消费需求趋势预测 [J]. *统计与决策*, 2018(20): 86-89.
- [21] 崔杨,曲钰,仲悟之,等. 基于改进 shapley 值的风电汇聚趋势性分状态量化方法 [J/OL]. *电网技术*; 1-10 [2018-11-26]. <https://doi.org/10.13335/j.1000-3673.pst.2018.1451>.