

经济新常态下基于偏最小二乘回归的中长期负荷预测模型

王雁凌, 吴梦凯

(华北电力大学 电气与电子工程学院, 北京 102206)

摘要:经济新常态下,中长期负荷预测面临负荷趋势呈S形、可依赖数据样本少等问题,传统预测方法精度受限,因此提出一种基于偏最小二乘回归和情景分析法的中长期负荷预测模型。从经济新常态的速度、结构、动力3个宏观方面入手,建立一个宏观与微观相结合、层次化的电力负荷影响因素指标体系;采用偏最小二乘回归法获取电力负荷与各影响因素的关系方程;根据“十二五”及“十三五”期间经济和电力发展的特点,采用情景分析法设定多个情景及相应参数,获得不同情景下有区别的负荷预测结果,以降低预测风险。利用所提模型对某省“十三五”期间的逐年用电量进行预测,通过与现有成熟方法预测结果及官方公布数据的对比,验证了所提模型的有效性。

关键词:经济新常态;中长期负荷预测;偏最小二乘回归;情景分析法;模型

中图分类号:TM 715

文献标识码:A

DOI:10.16081/j.issn.1006-6047.2018.03.018

0 引言

中长期负荷预测,一般指时间尺度为5~10 a的负荷预测,是电力系统的一个重要研究领域,对于保障电源电网规划、运行的安全性和经济性具有重要意义^[1-2]。

2012年以来,随着我国经济发展进入新常态,经济环境对于电力负荷的影响愈发显著,也给中长期负荷预测带来新的挑战。一方面,全国电力消费告别高速增长阶段,呈现出明显的“S”型特征,负荷发展的不确定性增强,这使得一些比较适用于负荷稳定增长阶段的传统预测方法精度受限;另一方面,经济新常态阶段的负荷预测能够依赖的基础数据主要是“十二五”中后期的数据,样本数量少,为建模带来很大的困难。因此,需要提出负荷增长新阶段下的中长期负荷预测模型。

国内外关于中长期负荷预测的模型和方法主要有:传统的单耗法、趋势外推法、弹性系数法,计量经济学中的时间序列法、回归分析法,现代预测技术中的神经网络法、模糊预测法和灰色预测法等^[3],以及博采各方法之众长的组合预测法^[4]。近年来,考虑要素更全面、建模更精细的预测模型逐渐涌现。文献[5]针对城市化形成的复杂数据环境,采用层次分析法确定各要素对电力负荷影响的权重,再采用模糊聚类分析法预测中长期负荷;文献[6]借鉴向量误差修正理论,从短期扰动和历史存续2个方面深入分析不同行业的电力需求关系,提出一种新型的负荷预测方法;文献[7]从负荷数据的增长率无后效性出发,采用马尔可夫链划分区间,通过方差-协方差方法分配权重,提出一种组合预测模型;文献[8]利用经验模态分解模型从时间和空间上对负荷进行分解,再采用支持向量机算法进行负荷预测。

综上,中长期负荷预测发展迅速,预测效果不断改善,但是在以下2个方面仍然有一定改进的空间。其一,多数方法的预测思路本质上都是采用时序外推法,即利用负荷及其影响因素的历史数据建立负荷变化规律模型,再外推预测未来的负荷。这种方法比较适合用于负荷稳定增长、存在大量数据样本的情况。而在负荷增长的新阶段,真正较有价值的数据仅为近几年的数据,因此,需要探索在样本数量少情况下的负荷预测模型建模方法。其二,很多模型往往仅能提供单一的预测结果,或通过简单的上下浮动得到结果区间。而中长期负荷预测具有不确定性,在新常态下,未来的经济环境更是难以准确估计,单一的预测结果存在很大的风险与挑战。

本文提出基于偏最小二乘回归和情景分析法的中长期负荷预测模型。首先,从经济新常态的速度、结构、动力3个宏观方面入手,构建宏观与微观相结合、层次化的电力负荷影响因素指标体系。其次,利用偏最小二乘回归PLS(Partial Least-Square regression),解决样本数量少和变量多重相关性的问题,获取新常态下电力负荷与各影响因素的关系方程;利用情景分析法,设定未来多个经济发展情景及相应参数,以获得针对不同情景的负荷预测结果。最后,利用所提出的模型对浙江省“十三五”期间逐年的用电量进行预测。

1 经济新常态下的负荷影响因素指标

1.1 经济新常态对电力负荷的影响分析

我国自2012年开始,宏观经济由高速发展进入中高速发展阶段。2014年5月,习近平总书记在考察河南的行程中,首次提及“新常态”一词。同年11月9日,在亚太经合组织工商领导人峰会中,“经济新常态”得到系统阐述,主要有以下3个方面:一是

经济增速由高速增长转为中高速增长(速度变化);二是经济结构正不断优化升级,第三产业、消费需求逐步成为主体,城乡区域差距逐步缩小,居民收入占比上升(结构优化);三是从要素驱动、投资驱动转向创新驱动(动力转换)^[9]。

以下就从上述 3 个方面入手,具体分析我国经济新常态对电力负荷的影响,并提炼量化指标。

a. 速度变化方面。

经济增长减速是新常态的逻辑起点和事实起点,同时对电力负荷也具有全局性、根本性的影响。由于经济发展水平在很大程度上决定了全社会电力消费水平和结构的变化,因此经济增长减速对电力负荷的直接影响就是用电量的明显下滑。我国 2012 至 2014 年的 GDP (Gross Domestic Product) 增速分别为 7.65%、7.67%、7.40%,明显低于过去 30 a 平均两位数的增速;与此同时,这 3 年的全社会用电量增长率分别为 5.5%、7.5%、3.8%,相比 2011 年的增长率 11.7%、2010 年的增长率 14.6%,下滑更为明显^[10]。

因此,衡量经济新常态速度变化方面对电力负荷的影响,直观可取 GDP 作为量化指标。但由于通常电力负荷相比 GDP 波动幅度更大,因此还可以再取 CPI (Consumer Price Index)、PPI (Producer Price Index) 等作为量化指标。

b. 结构优化方面。

经济结构的优化是新常态的重要特点,既包括传统的一二三产比重的调整,又涵盖重振内需、提高居民收入、区域结构调整等新的结构调整措施。受此影响,用电结构也有了明显的改善。二产用电占比从 2011 年的 75% 下降到 2015 年的 72%;三产和居民用电占比则从 23% 上升到 26%^[10],对全社会用电量增长的贡献超过七成。

因此,衡量经济新常态结构优化方面对电力负荷的影响,直观可取二产、三产及居民用电量占比作为量化指标。另外,还可取人均可支配收入作为重振内需等对负荷影响的量化指标。

c. 动力转换方面。

经济新常态的本质是增长动力的转换,即从要素驱动、投资驱动转向创新驱动。具体到电力负荷,则是要逐步摆脱过于依赖高耗能行业的局面,实现电力的低碳化、清洁化、创新式发展。

2012 至 2015 年,高耗能行业产能相对过剩,电力消费呈现低迷态势,增速同比全线下滑。2012 年,四大高耗能行业合计用电量同比增长 2.7%,增速比同期的全社会用电量平均水平低 2.8%。至 2015 年,四大高耗能行业用电量比重同比下降 3.4%,造成全社会用电量增速大幅下降、第二产业及其工业用电量出现负增长^[10]。不过与此同时,我

国出台了关于促进可再生能源消纳、大气污染防治和电能替代等一系列政策,将在促进节能减排的同时缓解用电量下滑的压力。

经济新常态下,经济、电力的发展将转换为依靠创新驱动,而创新投入则必然是从原来的高耗能、高污染转为低耗能、清洁化。因此,可取万元 GDP 能耗和废气排放总量作为衡量经济新常态动力转换方面对电力负荷影响的量化指标。

1.2 电力负荷影响因素指标体系的构建

由第 1.2 节,本文从经济新常态的速度、结构、动力 3 个宏观方面入手,选取 GDP (x_1)、CPI (x_2)、PPI (x_3)、二产用电占比 (x_4)、三产及居民用电占比 (x_5)、人均可支配收入 (x_6)、万元 GDP 能耗 (x_7)、废气排放总量 (x_8) 8 个微观指标来建立电力负荷影响因素指标体系,如图 1 所示。

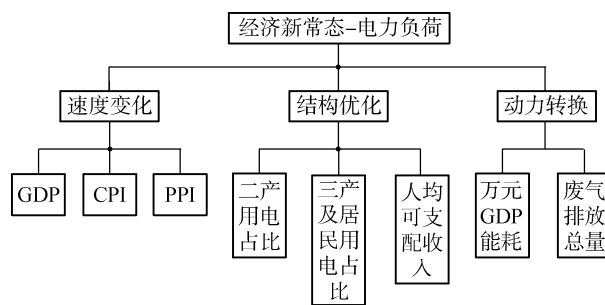


图 1 经济新常态下电力负荷影响因素指标体系

Fig.1 Index system of factors influencing power load under new economic normality

2 PLS 模型

PLS^[11-13]是一种新型的多元统计分析方法,有机结合了主成分分析、典型相关性分析及多元线性回归分析,能在样本个数较少以及自变量存在严重多重相关性的条件下建模,且模型对实际有较强的解释力。

考虑到负荷增长新阶段的负荷及其影响因素数据样本较少,并且各个因素变量之间可能存在多重相关性,因此 PLS 是探索新阶段下负荷与其影响因素之间关系的较好方法。考虑到中长期负荷预测的实际需要,下面仅给出一种单因变量 PLS 模型的建模方法。

2.1 PLS 方程的建立

a. 设已知因变量 y 和 m 个自变量 $x_1, x_2, \dots, x_m$, 样本数为 n , 形成因变量矩阵 $Y = [y]_{n \times 1}$ 和自变量矩阵 $X = [x_1, x_2, \dots, x_m]_{n \times m}$ 。将 X 和 Y 分别标准化处理^[14]为 E_0 和 F_0 。

b. 从 E_0 中提取第 1 个主成分 t_1 。 t_1 是 E_0 各列的线性组合,同时要求 t_1 尽可能多地携带 E_0 中的变异信息,且与 F_0 的相关程度最大。这样,就尽可能好地代表了自变量 $x_1, x_2, \dots, x_m$,同时对因变量 y 有最

强的解释能力。

具体地,取 ω_1 为 E_0 的第 1 主轴,即 $\|\omega_1\| = 1$ 。由于 E_0 和 F_0 均为标准化矩阵,则有:

$$\omega_1 = \frac{E_0^T F_0}{\|E_0^T F_0\|} = \frac{1}{\sqrt{\sum_{j=1}^m r^2(x_j, y)}} \begin{bmatrix} r(x_1, y) \\ \vdots \\ r(x_m, y) \end{bmatrix} \quad (1)$$

$$t_1 = E_0 \omega_1 = \frac{1}{\sqrt{\sum_{j=1}^m r^2(x_j, y)}} \times [r(x_1, y) E_{01} + r(x_2, y) E_{02} + \cdots + r(x_m, y) E_{0m}] \quad (2)$$

其中, $E_{0i} (i = 1, 2, \cdots, m)$ 为 E_0 的第 i 列; $r(x_j, y)$ ($j = 1, 2, \cdots, m$) 为 x_j 与 y 的相关系数。在简化计算中,可不进行 F_0 对 t_1 的回归,只需求出 E_0 对 t_1 的回归系数 p_1 即可。

$$E_0 = t_1 p_1^T + E_1 \quad (3)$$

$$p_1 = \frac{E_0^T t_1}{\|t_1\|^2} \quad (4)$$

其中, E_1 为回归方程的残差矩阵。

c. 检验交叉有效性。若满足停止条件,则只提取一个成分 t_1 ; 否则,重复步骤 b,并以 E_1 取代 E_0 ,类似得到 ω_2 。须注意,由于 E_1 不再是标准化矩阵,则:

$$\omega_2 = \frac{E_1^T F_0}{\|E_1^T F_0\|} = \frac{1}{\sqrt{\sum_{j=1}^m \text{cov}^2(E_{1j}, y)}} \begin{bmatrix} \text{cov}(E_{11}, y) \\ \text{cov}(E_{12}, y) \\ \vdots \\ \text{cov}(E_{1m}, y) \end{bmatrix} \quad (5)$$

$$t_2 = E_1 \omega_2 \quad (6)$$

$$p_2 = \frac{E_1^T t_2}{\|t_2\|^2} \quad (7)$$

$$E_2 = E_1 - t_2 p_2^T \quad (8)$$

其中, $\text{cov}(E_{1j}, y)$ 为 E_{1j} 与 y 的协方差。

d. 根据所提取的 k 个成分 $t_1, t_2, \cdots, t_k$, 实施 F_0 在 $t_1, t_2, \cdots, t_k$ 上的回归,得到:

$$F_0 = r_1 t_1 + r_2 t_2 + \cdots + r_k t_k \quad (9)$$

其中, $r_u (u = 1, 2, \cdots, k)$ 为 F_0 关于 $t_1, t_2, \cdots, t_k$ 的回归系数。

由于 $t_g (g = 1, 2, \cdots, k)$ 均为 E_0 中各列的线性组合,故合并同类项后通过标准化的逆过程,可得到 y 关于 $x_1, x_2, \cdots, x_m$ 的回归方程为:

$$\hat{y} = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m \quad (10)$$

其中, $\beta_f (f = 1, 2, \cdots, m)$ 为 y 关于 $x_1, x_2, \cdots, x_m$ 的回归系数。

2.2 交叉有效性检验

根据上述建模过程,PLS 方程一般不需要选用全部成分进行回归建模。究竟选择多少个成分为宜,可通过增加一个新的成分后能否对模型的预测功能有明显的改进来确定。

设已选择 $h-1$ 个成分,现确定第 h 个成分是否需要选取。首先,将第 i 个样本 y_i 从所有 n 个样本中提取出来,用剩下的 $n-1$ 个样本提取 h 个成分得到一个 PLS 方程,并将第 i 个样本点代入,得到预测值 $\hat{y}_{h(-i)}$ 。循环选取 $i = 1, 2, \cdots, n$, 得到预测误差的平方和指标为:

$$P_h = \sum_{i=1}^n (y_i - \hat{y}_{h(-i)})^2 \quad (11)$$

然后,用全部 n 个样本提取 $h-1$ 个成分得到回归方程,该方程对所有样本预测误差的平方和记为:

$$S_{h-1} = \sum_{i=1}^n (y_i - \hat{y}_{(h-1)i})^2 \quad (12)$$

定义交叉有效性为:

$$Q_h^2 = 1 - \frac{P_h}{S_{h-1}} \quad (13)$$

一般认为,当 $Q_h^2 \geq 0.0975$ 时,第 h 个成分的边际贡献率是显著的,增加它是有益的;否则,可以认为不应再增加该成分。

3 基于 PLS 和情景分析法的负荷预测模型

3.1 情景分析法

情景分析法^[15-18]在能源资源规划、政策分析以及决策管理等领域应用广泛,该方法的实质是构建一套中长期战略框架,基于现实情况和可预见趋势,预测未来发展过程的一系列方案。情景分析法应用于电力负荷预测时,通常首先设定经济、政策等参数获得未来的不同情景,从而模拟分析各个情景下有区别的电力需求。

经济新常态下的中长期负荷预测面对的不确定性较大,应用情景分析法可使决策者有效避免过高或过低的决策,从而选择相对可靠的预测方案。

3.2 负荷预测情景设定原则

根据第 1.1 节中经济新常态速度、结构、动力 3 个方面的变化特点,同时对比国家“十二五”与“十三五”规划可知,“十三五”期间,我国经济发展环境总体上形势更加严峻,创新驱动、民生保障和环境保护放在了更加突出的位置,三产比重规划进一步提高等。由此,可设定 3 个未来经济社会发展的不同情景,相关指标参数设定原则如表 1 所示。

表 1 未来经济社会发展指标参数设定

Table 1 Parameters setting for economic and social development indicators

参数	变化趋势	情景 1	情景 2	情景 3
GDP	降速趋缓	√		
	历史趋势		√	
CPI	降速加大			√
	历史趋势	√	√	
PPI	增速趋缓			√
	历史趋势	√	√	√
二产用电占比	降速加大	√		
	历史趋势		√	√
三产及居民用电占比	增速加大	√		
	历史趋势		√	√
人均可支配收入	增速加大	√		
	历史趋势		√	
万元 GDP 能耗	降速加大	√		
	历史趋势		√	√
废气排放总量	降速加大	√		
	历史趋势		√	√

注：“√”表示对应情景下所对应的变化趋势。

a. 情景 1。在该情景下,经济减速得到了有效控制,经济结构优化、动力转换加快。具体而言,GDP、CPI、PPI 相比 2012 至 2015 年的减速趋势有所趋缓;二产用电占比下降更为明显,三产和居民用电增长势头更为强劲,人均可支配收入较大幅度提高,消费需求逐步成为主体;节能减排工作深入推进,发展模型更显低碳、清洁和创新。

b. 情景 2。按照进入经济发展新常态的近几年(2012 至 2015 年)惯有的经济增长速度、用电结构、居民收入和节能减排情况,保持惯性发展带来的电力负荷需求情景。

c. 情景 3。在该情景下,经济发展不容乐观,减速更为明显,且呈现通货紧缩形势;人均可支配收入增速趋缓,三产和消费的带动作用不强;用电结构和节能减排基本维持历史趋势。

3.3 中长期负荷预测模型框架

本文在预测中长期电力负荷时,综合应用 PLS 和情景分析法,模型以定量分析为主,并结合了定性分析的优势。首先,基于地区电力负荷及其影响因素指标的历史数据(主要采用“十二五”期间负荷增长新阶段下的数据),利用 PLS 建立负荷与影响因素的关系模型,来反映负荷发展的客观规律;其次,基于决策者对未来趋势的研判,以相对易于把握的经济增长、用电结构以及节能减排目标为基础,设定不同的预测情景及参数;最后,将参数代入 PLS 方程,求得不同情景下的中长期电力负荷需求。模型框架如图 2 所示。

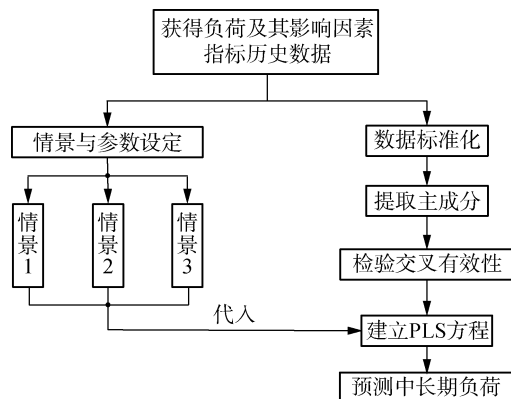


图 2 基于 PLS 和情景分析法的中长期负荷预测模型

Fig.2 Medium and long term load forecasting model based on PLS and scenario analysis method

4 算例分析

本文以浙江省为例,运用所提基于 PLS 和情景分析法的中长期负荷预测模型,对“十三五”期间的年用电量需求进行预测。

4.1 取用历史数据及多重相关性诊断

浙江省“十二五”期间的年用电量及其影响因素指标历史数据^[19],如表 2 所示。

表 2 浙江省“十二五”期间经济社会指标

Table 2 Economic and social indicators of Zhejiang province during 12th five-year plan

年份	x_1 /亿元	x_2	x_3	x_4	x_5
2011	32 363.38	100.00	100.00	0.779	0.215
2012	34 739.13	102.20	97.30	0.763	0.231
2013	37 756.58	104.55	95.55	0.752	0.241
2014	40 173.03	106.75	94.40	0.756	0.237
2015	42 886.49	108.24	91.00	0.742	0.251

年份	x_6 /元	x_7 /吨	x_8 /亿 m^3	电量/(亿 kW·h)	电量增速/%
2011	30 971	0.551	24 940	3 116.91	10.49
2012	34 550	0.520	23 967	3 210.55	3.00
2013	37 080	0.494	24 565	3 453.05	7.55
2014	40 393	0.469	26 958	3 506.39	1.54
2015	43 714	0.457	26 841	3 553.90	1.35

注:CPI、PPI 以 2011 年为基期,取同基指数;由于统计口径的变化,人均可支配收入取城镇人均可支配收入数据; x_1 — x_5 列数据为其相应元素值, x_7 、 x_8 的计量标准分别为标准煤、标准状态,下同。

分析表 2 易知,“十二五”期间,浙江省经济发展与全国保持大体相同的态势,自 2012 年进入新常态。新常态下,浙江电力负荷增速下降明显,用电结构更加优化,节能减排工作稳中有进。

自变量系统的多重相关性可通过计算方差膨胀因子 VIF 进行判断^[14]。

$$(VIF)_i = (1 - R^2)^{-1} \quad (14)$$

其中, R^2 为以变量 x_i 为因变量时对其他自变量回归的复测定系数。所有 $(VIF)_i$ 中最大的一个用来作为诊

断多重相关性的指标 VIF。如果 $VIF > 10$, 通常认为多重相关性将严重影响最小二乘回归的估计值。

对表 2 中的自变量 $x_1, x_2, \dots, x_8$ 进行方差膨胀因子诊断, 结果如表 3 所示。由表 3 可知, 最大方差膨胀因子为 42 597。由此说明, 自变量系统中存在严重的多重相关性。

表 3 自变量方差膨胀因子

Table 3 Variance inflation factors of independent variables

自变量	VIF	自变量	VIF	自变量	VIF
x_1	42 597	x_4	38	x_7	2 347
x_2	6 328	x_5	112	x_8	3 021
x_3	912	x_6	3 989		

4.2 情景与参数设定

以表 2 数据反映出的浙江省经济社会各指标趋势为基准, 并根据第 3.2 节的负荷预测情景设定原则, 参考相关规划, 设定 2016 至 2020 年间的经济社会发展情景 1—3 的各指标参数, 如表 4 所示(由于篇幅所限, 仅给出情景 1、2 的指标参数)。

表 4 浙江省“十三五”期间经济社会指标设定

Table 4 Economic and social indicator setting of Zhejiang Province during 13th five-year plan

年份	情景	x_1 /亿元	x_2	x_3	x_4
2016	1	45 888.54	109.50	89.10	0.726
	2	45 674.11	109.43	88.91	0.734
2017	1	49 100.74	111.08	87.30	0.710
	2	48 642.92	110.53	86.87	0.725
2018	1	52 537.79	112.41	85.55	0.702
	2	51 804.71	111.52	84.87	0.716
2019	1	56 215.43	113.54	83.74	0.694
	2	55 172.02	112.41	82.92	0.708
2020	1	60 150.52	114.33	81.82	0.679
	2	58 758.20	113.20	81.01	0.700
年份	情景	x_5	x_6 /元	x_7 /吨	x_8 /亿 m^3
2016	1	0.266	47 887	0.432	26 199
	2	0.261	47 648	0.437	26 734
2017	1	0.277	52 196	0.413	26 094
	2	0.271	51 937	0.417	26 627
2018	1	0.288	56 894	0.386	25 990
	2	0.282	56 611	0.398	26 520
2019	1	0.299	62 014	0.361	25 886
	2	0.294	61 706	0.380	26 414
2020	1	0.311	67 596	0.345	25 782
	2	0.305	67 259	0.363	26 308

4.3 建立 PLS 方程与负荷预测结果

根据第 2 节的 PLS 模型建模步骤, 以浙江省 2011 至 2015 年的用电量及其影响因素指标数据为基础, 建立 PLS 方程如下:

$$y = 2\,988.7 + 19.5 x_2 + 17.8 x_3 - 3\,741.4 x_4 + 2\,710.7 x_5 - 2\,351.3 x_7 \quad (15)$$

其中, y 为年用电量(亿 $kW \cdot h$)。

运用式(15)拟合 2011 至 2015 年的用电量, 校验模型拟合精度, 如表 5 所示。

表 5 PLS 方程拟合精度校验

Table 5 Verification of fitting accuracy for PLS equation

年份	用电量实际值/ (亿 $kW \cdot h$)	用电量拟合值/ (亿 $kW \cdot h$)	拟合误差/%
2011	3 116.91	3 093.22	-0.76
2012	3 210.55	3 262.04	1.60
2013	3 453.05	3 404.61	-1.40
2014	3 506.39	3 460.62	-1.31
2015	3 553.90	3 547.38	-0.18

由表 5 可知, 式(15)的 PLS 方程用于拟合 2011 至 2015 年的用电量, 相对误差均小于 2%, 取得了很好的拟合效果。

为验证 PLS 模型可以较好地解决变量多重相关性问题, 采用传统的最小二乘法建立回归方程如式(16)所示。

$$y = -0.311\,9 x_1 + 600.908\,3 x_2 - 383.010\,7 x_3 - 0.387\,7 x_6 + 0.137\,5 x_8 \quad (16)$$

将表 4 中情景 1 的各指标参数代入式(16), 得到基于传统最小二乘回归法的用电量预测值如表 6 所示。

表 6 最小二乘回归法的用电量预测结果

Table 6 Results of electricity consumption forecasting based on least-square regression method

年份	用电量预测值/ (亿 $kW \cdot h$)	年份	用电量预测值/ (亿 $kW \cdot h$)
2016	2 397	2019	-1 863
2017	1 349	2020	-4 059
2018	-89		

由表 6 可知, 虽然运用传统的最小二乘回归法也能够得到负荷与其影响因素的关系方程, 但是用其进行负荷预测时, 得出的结果十分不可靠, 甚至出现了负值。这主要是因为传统最小二乘回归法不能有效解决变量的多重相关性问题, 不能同时保证既剔除相关性变量, 又尽可能多地将变量计入模型中, 所以在预测过程中可能导致较大误差。而 PLS 则通过提取主成分, 较好地解决了这一问题。

将第 4.2 节中的各情景参数设定值代入 PLS 方程式(15), 得到浙江省“十三五”期间逐年的用电量如表 7 所示。

表 7 浙江省“十三五”期间逐年用电量预测值

Table 7 Forecasting value of electricity consumption of Zhejiang Province during 13th five-year plan

年份	情景 1		情景 2		情景 3	
	预测值/ (亿 $kW \cdot h$)	增速/ %	预测值/ (亿 $kW \cdot h$)	增速/ %	预测值/ (亿 $kW \cdot h$)	增速/ %
2016	3 699.11	4.28	3 641.71	2.66	3 577.69	0.85
2017	3 830.81	3.56	3 733.54	2.52	3 647.33	1.95
2018	3 949.41	3.10	3 822.93	2.39	3 714.20	1.83
2019	4 060.24	2.81	3 909.96	2.28	3 778.44	1.73
2020	4 168.95	2.68	3 994.69	2.17	3 840.17	1.63

为对比说明本文所提模型的预测精度,选用目前较为成熟且应用广泛的趋势外推法、时间序列法、灰色 GM(1,1) 模型 3 种中长期负荷预测方法进行预测,得到结果如表 8 所示。

表 8 不同方法的用电量预测结果

Table 8 Results of electricity consumption forecasting for different methods

年份	趋势外推法		时间序列法		灰色 GM(1,1) 模型	
	预测值/ (亿 kW·h)	增速/ %	预测值/ (亿 kW·h)	增速/ %	预测值/ (亿 kW·h)	增速/ %
2016	3 736.69	5.14	3 618.79	1.83	3 706.92	4.31
2017	3 870.00	3.57	3 681.69	1.74	3 824.46	3.17
2018	4 008.07	3.57	3 744.59	1.71	3 945.73	3.17
2019	4 151.07	3.57	3 807.49	1.68	4 070.84	3.17
2020	4 299.17	3.57	3 870.39	1.65	4 199.92	3.17

注:趋势外推法采用增长型曲线;时间序列法采用指数平滑法中的布朗趋势模型。

由表 2、7、8 易知,3 种负荷预测对比方法均只能提供单一的预测结果,而本文提出的基于 PLS 和情景分析法的中长期负荷预测模型则可以提供不同经济发展情景下的多套负荷预测结果,以降低预测风险。同时,趋势外推法和灰色 GM(1,1) 模型的负荷预测结果没有体现出负荷增长新阶段电力消费增长减速放缓的特点,而时间序列法的预测结果则又过于保守。

另外,2016 年 9 月,浙江省发展改革委发布《浙江省电力发展“十三五”规划》^[20],其中公布了全省“十三五”期间的电力需求预测值。经对比,利用本文提出的中长期电力负荷预测模型在情景 1 下的预测结果与浙江省官方发布的数据十分接近,进一步验证了模型的有效性。但由于情景 1 是对经济社会发展采取了较为乐观的估计,且预测出的用电量增速高于 2012 至 2015 年的平均水平。因此,本文推荐浙江省“十三五”期间的电力负荷需求应采用介于情景 1 和情景 2 之间的预测结果。

5 结论

在我国经济发展进入新常态的背景下,本文提出一种基于 PLS 和情景分析法的中长期负荷预测模型。主要结论有:

a. 经济新常态的 3 个主要宏观表现(速度变化、结构优化、动力转换)均会对电力负荷产生影响。选取了 8 个微观指标来反映这些宏观表现,建立起一个经济新常态下宏观与微观相结合、层次化的电力负荷影响因素指标体系。

b. 在负荷增长的新阶段,电力负荷及其影响因素数据样本较少,且各因素变量之间可能存在多重相关性,利用 PLS 可较好地解决这些问题,便于得到负荷与其影响因素的定量关系。

c. 本文提出的中长期负荷预测模型结合了情景分析法,融入决策者对未来经济环境的研判,能够提供基于不同经济发展情景下有区别的负荷预测结果,有利于降低预测风险,为电力规划、运行等部门提供更为可靠的电力负荷预测方案。

参考文献:

- [1] KHUNTIA S R, RUEDA J L, MART A M M, VAN DER M. Forecasting the load of electrical power systems in mid- and long-term horizons: a review[J]. IET Generation, Transmission & Distribution, 2016, 10(16): 3971-3977.
- [2] 许梁, 孙涛, 徐箭, 等. 基于函数型非参数回归模型的中长期日负荷曲线预测[J]. 电力自动化设备, 2015, 35(7): 89-94.
XU Liang, SUN Tao, XU Jian, et al. Mid- and long-term daily load curve forecasting based on functional nonparametric regression model[J]. Electric Power Automation Equipment, 2015, 35(7): 89-94.
- [3] 康重庆, 夏清, 张伯明. 电力系统负荷预测研究综述与发展方向的探讨[J]. 电力系统自动化, 2004, 28(17): 1-11.
KANG Chongqing, XIA Qing, ZHANG Boming. Review of power system load forecasting and its development[J]. Automation of Electric Power Systems, 2004, 28(17): 1-11.
- [4] 张栋梁, 严健, 李晓波, 等. 基于马尔可夫链筛选组合预测模型的中长期负荷预测方法[J]. 电力系统保护与控制, 2016, 44(12): 63-67.
ZHANG Dongliang, YAN Jian, LI Xiaobo, et al. Mid-long term load forecasting based on Markov chain screening combination forecasting models[J]. Power System Protection and Control, 2016, 44(12): 63-67.
- [5] 李亦言, 严正, 冯冬涵. 考虑城市化因素的中长期负荷预测模型[J]. 电力自动化设备, 2016, 36(4): 54-61.
LI Yiyan, YAN Zheng, FENG Donghan. Mid/long-term load forecasting model considering urbanization characteristics[J]. Electric Power Automation Equipment, 2016, 36(4): 54-61.
- [6] 王鹏, 陈启鑫, 夏清, 等. 应用向量误差修正模型的行业电力需求关联分析与负荷预测方法[J]. 中国电机工程学报, 2012, 32(4): 100-107.
WANG Peng, CHEN Qixin, XIA Qing, et al. Correlation analysis and forecasting method on industrial electricity demand based on vector error correction model[J]. Proceedings of the CSEE, 2012, 32(4): 100-107.
- [7] ZHANG D L, YAN J, WANG W H, et al. Mid-long term load forecasting of the unstable growth sequence based on Markov chains screening combination forecasting models[C] // 2016 China International Conference on Electricity Distribution (CICED 2016). Xi'an, China: CIRED, 2016: 1-5.
- [8] GHELARDONI L, GHIO A, ANGUITA D. Energy load forecasting using empirical mode decomposition and support vector regression[J]. IEEE Transactions on Smart Grid, 2013, 4(1): 549-556.
- [9] 国家行政学院经济学教研部. 中国经济新常态[M]. 北京: 人民出版社, 2014: 47-72.
- [10] 中华人民共和国国家统计局. 中国统计年鉴: 2016[M]. 北京: 中国统计出版社, 2016: 156-204.
- [11] WU F Y, ASADA H H. Implicit and intuitive grasp posture control for wearable robotic fingers: a data-driven method using partial least squares[J]. IEEE Transactions on Robotics, 2016, 32(1): 176-

- 186.
- [12] ZHU Y L, JIA Y F, WANG L Y. Partial discharge pattern recognition method based on variable predictive model-based class discriminate and partial least squares regression[J]. IET Science, Measurement & Technology, 2016, 10(7): 737-744.
- [13] 李建生, 梁军, 负志皓, 等. 基于电气信息评估设备状态渐进变化的概率分析方法[J]. 电工技术学报, 2013, 28(10): 355-363. LI Jiansheng, LIANG Jun, YUN Zhihao, et al. Probability analysis of progressive changes of equipment state based on electrical information[J]. Transactions of China Electrotechnical Society, 2013, 28(10): 355-363.
- [14] 张戈力, 毛安家, 赵岩. 一种 PLS 回归的并网风电项目利润预测方法[J]. 电力系统保护与控制, 2013, 41(8): 87-92. ZHANG Geli, MAO Anjia, ZHAO Yan. A profit forecast method of wind power project based on partial least squares regression[J]. Power System Protection and Control, 2013, 41(8): 87-92.
- [15] 王建, 吴奎华, 刘志珍, 等. 电动汽车充电对配电网负荷的影响及有序控制研究[J]. 电力自动化设备, 2013, 33(8): 47-52. WANG Jian, WU Kuihua, LIU Zhizhen, et al. Impact of electric vehicle charging on distribution network load and coordinated control[J]. Electric Power Automation Equipment, 2013, 33(8): 47-52.
- [16] MEIXELL M J, WU S D. Scenario analysis of demand in a technology market using leading indicators[J]. IEEE Transactions on Semiconductor Manufacturing, 2001, 14(1): 65-75.
- [17] RAHMANI M, ROMERO R A, RIDER M J. Risk/investment-driven transmission expansion planning with multiple scenarios[J]. IET Generation, Transmission & Distribution, 2013, 7(2): 154-165.
- [18] 曾鸣, 田廓, 鄢帆, 等. 基于不同市场主体的输电投资效益计算模型[J]. 电力自动化设备, 2010, 30(12): 7-11. ZENG Ming, TIAN Kuo, YAN Fan, et al. Evaluation model of transmission investment benefit based on different market entities[J]. Electric Power Automation Equipment, 2010, 30(12): 7-11.
- [19] 浙江省统计局, 国家统计局浙江调查总队. 浙江统计年鉴: 2016[M]. 北京: 中国统计出版社, 2016: 78-324.
- [20] 浙江省发展和改革委员会, 浙江省能源局. 浙江省电力发展“十三五”规划[EB/OL]. [2017-01-10]. http://www.zjdpc.gov.cn/art/2016/9/19/art_8_1715438.html.

作者简介:



王雁凌

王雁凌(1970—),女,辽宁沈阳人,副教授,博士,主要研究方向为电力经济、电力系统运行与控制(E-mail: wangyanling@ncepu.edu.cn);

吴梦凯(1993—),男,浙江丽水人,硕士研究生,通信作者,主要研究方向为电力系统负荷预测、电力经济(E-mail: hdu_wmk@163.com)。

Medium and long term load forecasting model based on partial least-square regression under new normal economy

WANG Yanling, WU Mengkai

(School of Electrical and Electronic Engineering, North China Electric Power University, Beijing 102206, China)

Abstract: The medium and long term load forecasting faces problems under the new normal economy, such as S shape load trend, few dependable data sample, etc., and the accuracy of traditional forecasting method is limited, for which, a medium and long term load forecasting method based on partial least-square regression and scenario analysis method is proposed. With the combination of macroscopic view with the microscopic view, a hierarchical index system of factors influencing power load is built from three macroscopic aspects of new normal economy, i.e. speed, structure and driving force. The partial least-square regression method is adopted to obtain the relation equation between power load and its influence factors. According to the characteristics of economy and power development during the 12th and 13th five-year plan, the scenario analysis method is adopted to set up several scenarios and corresponding parameters, and different load forecasting results under each scenarios are obtained for reducing the forecasting risk. The proposed model is applied for forecasting the yearly electricity consumption of a province during the 13th five-year plan, and the comparison with the forecasting results of existing sophisticated methods and official data verify the effectiveness of the proposed model.

Key words: new normal economy; medium and long term load forecasting; partial least-square regression; scenario analysis method; models