

基于分位数回归的电力负荷预测影响因素研究及应用

传统回归模型只对电力负荷进行均值预测,无法详细地分析在不同时期各因素对电力负荷规模影响程度的差异。针对这一问题,建立负荷预测的分位数回归模型,保留历史数据的统计信息,采用电力需求已进入稳定阶段的日本用电量数据进行仿真,定量分析了不同社会发展水平下各社会经济因素对电力需求的影响大小,得出了电力需求量与分位点选取呈正相关且存在周期波动的规律,并将这一结论应用于我国华东某城市饱和负荷预测,为电力负荷预测研究提供了一种新的方法。

薛承荣 顾洁/上海交通大学电气工程系 孙金华/国网上海节能服务有限公司 赵建平 侯兴政/上海交通大学继续教育学院



薛承荣/硕士研究生

关键词/Keywords

电力系统·
负荷预测·
相关性·
分位数回归·

负荷预测是电力生产的基础^[1],其方法可以概括为传统预测(如回归、时间序列法等)及现代人工智能方法(如人工神经网络、支持向量机和灰色理论等)两大类^[2-5]。对于影响电力消费的因素,文献[6]提出了基于M-估计的稳健回归预测算法,具有较好的抗干扰性。

传统的回归分析,如最小二乘回归(Ordinary Least Square, OLS)只关注变量的平均取值(称为“均值回归”),对于影响因素的分析不够全面,尤其在电力负荷预测的过程中,由于预测所涉及的时间跨度长,不确定性因素多,在发展的过程中影响因素及因素的影响权重都在发生变化,采用传统的回归方法可能会对预测结果产生不利影响。Koenker等人提出的分位数回归可以较好地克服均值回归的不足,同时能够完整地考察响应变量的整个条件分

布,在经济和金融等很多领域获得了广泛运用,但目前在电力负荷预测方面的应用成果则较少。

鉴于在不同时期电力需求可能受不同因素的影响,同时影响程度也会有所不同,本文引入分位数回归理论,对电力需求发展已进入饱和阶段的日本用电量历史数据进行分位数回归分析,预测其用电量,分析不同时期各因素的影响权重,结合其经济发展情况分析得到最优分位点选取和电力需求之间的规律,并将其应用于我国华东某城市饱和负荷预测中,为我国的电力负荷预测工作提供有益的借鉴。

分位数回归基本理论

通常的回归方法为均值回归,Koenker和Bassett于1978年提出了分位数回归的思想,保留了变量的统计信息,在一定程度上消除了异方差问题,是对以古典条件均值模型为基础的最小二乘法的延伸,具体描述如下:

考虑受到 k 个因素 $X_1, X_2, \dots, X_k$ 影响的随机变量 Y ,其分布函数为

$$F(y) = P(Y \leq y) \quad (1)$$

则 Y 的第 τ 分位数为满足 $F(y) \geq \tau$ 的最小 τ 值,即

$$Q(\tau) = \inf\{y: F(y) \geq \tau\} \quad 0 < \tau < 1 \quad (2)$$

Koenker等人提出了如下的线性分位数回归模型

$$Q_Y(\tau|X) = \beta_0(\tau) + \beta_1(\tau)X_1 + \beta_2(\tau)X_2 + \dots + \beta_k(\tau)X_k = X^T \beta(\tau) \quad (3)$$

式中, $Q_Y(\tau|X)$ 为响应变量 Y 在解释变量

上海交通大学文理交叉专项基金项目(11JCY11)。

$X = [X_1, X_2, \dots, X_k]^T$ 给定下的条件 τ 分位数; τ 为分位点, $\tau \in (0, 1)$; $\beta(\tau)$ 为回归系数向量, $\beta(\tau) = [\beta_0(\tau), \beta_1(\tau), \beta_2(\tau), \dots, \beta_k(\tau)]^T$ 。

式(3)中 $\beta(\tau)$ 的求解, 可以转化为求解以下优化问题

$$\min_{\beta} \sum_{i|Y_i \geq X_i' \beta} \tau |Y_i - X_i' \beta| + \sum_{i|Y_i < X_i' \beta} (1-\tau) |Y_i - X_i' \beta| \quad (4)$$

式(4)表明, 对回归系数的求解, 实质是求解关于 y 与其拟合值之间加权残差最小的问题。

基于分位数回归的电力需求影响因素分析

近十多年来, 分位数回归在国外得到了迅猛的发展及应用, 其应用领域包括经济、医学、环境科学、生存分析以及动植物学等方面。本文将分位数回归技术应用于求解电力负荷回归预测模型, 深入挖掘各个分位数下解释变量与被解释变量的关系, 便于对影响用电需求的各种因素进行研究。

1. 分析基本思路

日本作为亚洲发达国家, 其经济社会发展历程及相应电力需求的变化规律较为成熟完整, 能够为我国的电力需求分析等提供重要的参考, 为此本文针对日本的社会经济发展及电力需求等数据建立回归预测模型, 同时利用分位数计算变量参数对各因素进行分析, 为我国未来的电力需求预测规律分析提供参考。

首先以日本1965~1995年31年的用电量数据为基础建立基于经济和社会发展指标的用电量回归模型, 应用分位数回归技术得到在不同分位点下回归模型中的参数; 然后应用在不同分位点下计算得到的模型分别预测日本1996~2005年10年间的用电量, 并同实际值相比较, 分析在不同用电规模和社会经济发展水平下分位数选取的规律, 进而分析不同影响因素在各分位点的影响权重, 从而得到模型中各因素对电力负荷需求的影响大小; 最后对日本1980~2005年用电量以相对误差最小为目标进行分位数回归分析, 得出分位点选取同年用电量增长率呈正相关且具有周期波动的规律, 利用这一规律预测得到我国华东某城市饱和负荷规模。

2. 用电量分位数回归预测模型

根据式(3)建立用电量的线性回归模型, X^T 即为影响用电量的各影响因素向量。在应用时需要结合各个地区实际社会经济发展状况选择模型变量, 并利用分位数回归计算模型中的参数。根据参考文献[7], 日本的用电量与宏观经济、人口规模以及第二、三产业比重的相关度均非常高, 即日本的电力需求主要是由规模效应(如GDP与人口)与结构效应(如三产业比重)决定, 为此建立以下用电量回归预测模型

$$ELE(\tau) = \beta_0(\tau) + \beta_1(\tau) \times GDP + \beta_2(\tau) \times POP + \beta_4(\tau) \times SEC + \beta_5(\tau) \times THIR \quad (5)$$

式中, ELE为用电量; GDP为国民生产总值; POP为总人口数; SEC为第二产业比重; THIR为第三产业比重。在实际计算时, 为消除异方差, 对各变量取对数。

算例仿真

1. 用电量回归预测结果

本文收集了日本1965~2005年41年间的用电量数据, 利用1965~1995年31年间的的历史数据, 采用分位数回归, 得到不同分位点对同一模型的不同组系数的估计值。为简化计算量, 同时又不失代表性, 取5个不同分位点得到式(5)参数, 预测日本1996~2005年10年的年用电量, 并同最小二乘回归预测结果相比较, 年用电量回归预测值如图1所示, 用电量预测结果如表1所示。

对比不同分位点下用电量拟合值可以看出, 高分位点下的模型参数对应得到的拟合值一般较

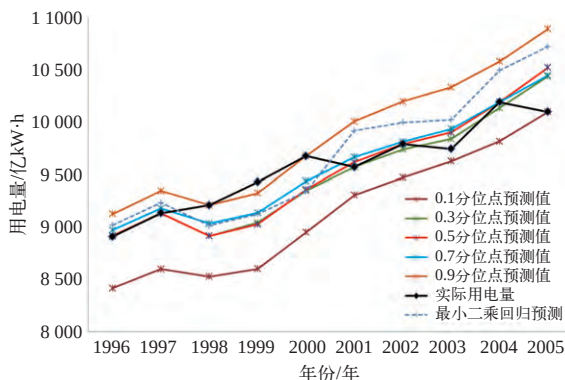


图1 不同分位点下日本年用电量回归预测值

表1 不同分位点下日本用电量预测结果

(单位: 亿kW·h)

年份/年	实际值	分位数回归										最小二乘回归	
		0.1	误差	0.3	误差	0.5	误差	0.7	误差	0.9	误差	预测值	误差
1996	8 911	8 417	-5.544	8 911	0.000	8 920	0.101	8 973	0.696	9 128	2.435	9 020	1.223
1997	9 135	8 599	-5.868	9 133	-0.022	9 135	0.000	9 181	0.504	9 347	2.321	9 235	1.095
1998	9 211	8 528	-7.415	8 918	-3.181	8 917	-3.192	9 036	-1.900	9 211	0.000	9 017	-2.106
1999	9 431	8 602	-8.790	9 043	-4.114	9 028	-4.273	9 137	-3.120	9 327	-1.100	9 128	-3.213
2000	9 678	8 953	-7.491	9 349	-3.399	9 361	-3.275	9 439	-2.470	9 678	0.000	9 341	-3.482
2001	9 573	9 307	-2.779	9 573	0.000	9 620	0.491	9 668	0.992	10 010	4.565	9 920	3.625
2002	9 794	9 474	-3.267	9 742	-0.531	9 794	0.000	9 815	0.214	10 200	4.145	9 999	2.093
2003	9 747	9 630	-1.200	9 842	0.975	9 905	1.621	9 938	1.960	10 336	6.043	10 025	2.852
2004	10 194	9 819	-3.679	10 138	-0.549	10 194	0.000	10 194	0.000	10 585	3.836	10 499	2.992
2005	10 101	10 101	0.000	10 441	3.366	10 526	4.208	10 450	3.455	10 896	7.871	10 726	6.188

高, 而低分位点的拟合值则相对较小。因此可以预测出未来不同分位点情况下的电力需求值。在社会经济增长迅速的年份, 电力需求增长旺盛, 可以采用高分位点对应的拟合值进行预测; 反之在经济社会发展速度相对减缓的年份, 电力需求增长也相应减速, 可以考虑选取低分位点对应的拟合值进行预测。

同时可以发现, 最小二乘回归只能得到平均的回归结果, 无法获得不同经济社会水平下的用电量特征; 而分位数回归提供了完整的统计分布信息, 因此可以为不同时期各种电力需求因素的分析提供依据。

2. 负荷预测影响因素及分位点变化规律分析

选取2005年数据, 分析不同分位点下各因素对于日本年用电量的影响大小, 如图2所示。

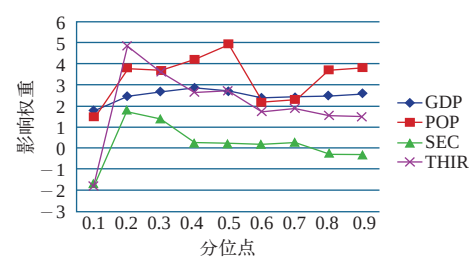


图2 不同分位点下各影响因素变化情况

图2表明了在不同分位点下, 电力需求受GDP、人口数量和产业结构(如第二产业、第三产业占比)的不同影响。通过分析可以得出以下结论:

1) 人口数量、GDP是影响用电量的最主要因素, 尤其是在0.3~0.5中分位段上, 影响权重

最大, 在中高分位点附近影响权重变化则比较平缓。这说明在经济发展的高速阶段, 电力消费增加较快, 人口数量、经济规模成为影响用电量的主要因素, 经济发展趋于稳定之后, 两者对于用电量的影响也保持在一个稳定的水平。对于发展成熟的地区, 高GDP的增长并不是电力需求高速增长的决定因素, 其中对电价、能源效率、产业结构和用电习惯等可能有较大的影响。

2) 产业结构中, 第三产业占比对用电量的影响大于第二产业占比所产生的影响。原因是本文以2005年数据点为分析对象, 此时日本已经进入后工业时期, 相应的日本用电结构(第三产业用电量占比超过第二产业用电量占比)与其工业发展阶段特征相吻合。在高分位点处, 第二产业占比影响系数甚至变为负值, 这说明了在发展处于波峰的市场环境下, 第三产业的用电量将成为总用电量主要的推动力量。

基于分位数回归, 对日本1976~2005年30年的用电量数据进行拟合, 针对不同阶段选取使拟合值相对误差最小的分位点, 得到相应的最优分位点选取情况, 如表2所示。

对分位点序列与用电量增长率序列进行相关性分析, 结果显示两者具有高度正相关性, 相关系数为0.914 3, 如图3所示。从图中可以看出, 地区用电量回归分析中最优分位点的选取呈现出周期波动变化的规律。因此可以通过对某地区用电量历史数据进行分位数回归分析, 以拟合值相对

表2 日本用电量分位数回归预测最优分位点选取结果

年份	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985
分位点	0.2	0.5	0.6	0.9	0.8	0.6	0.2	0.4	0.8	0.6
年份	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
分位点	0.1	0.2	0.2	0.5	0.9	0.6	0.3	0.1	0.7	0.6
年份	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
分位点	0.3	0.5	0.8	0.8	0.9	0.6	0.5	0.2	0.6	0.1

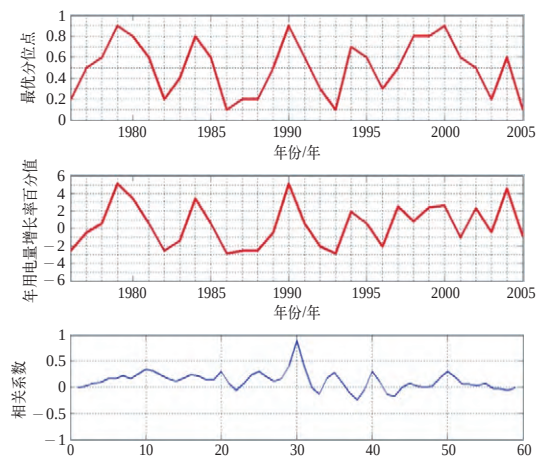


图3 最优分位点同年用电量增长率相关性

误差最小为目标确定每年的最优分位点，进而得到该区分位点选取的周期值，结合对于该地区电力需求饱和时间的研究，可以预测得出该地区电力需求达到饱和时电力需求规模。

3. 饱和负荷预测

对我国华东某市应用前文所建立的模型，在不同阶段取不同分位点，使负荷预测值残差最小，得到结果如表3所示。

表3 华东某市电力最高负荷分位数回归模型拟合结果

年份/年	最高负荷实际值/MW	最高负荷拟合值/MW	相对误差 (%)	分位点
1990	4 310	4 310	0.00	8
1991	4 500	4 500	0.00	7
1992	5 150	5 150	0.00	9
1993	5 550	5 550	0.00	5
1994	5 300	5 778	9.01	1
1995	7 100	7 076	-0.34	8
1996	7 600	7 592	-0.10	4
1997	8 250	8 250	0.00	2
1998	9 200	9 197	-0.03	6
1999	9 200	9 200	0.00	1
2000	10 600	10 634	0.32	3

(续)

年份/年	最高负荷实际值/MW	最高负荷拟合值/MW	相对误差 (%)	分位点
2001	11 300	11 300	0.00	2
2002	12 600	12 543	-0.45	4
2003	13 800	13 761	-0.28	4
2004	15 300	15 300	0.00	4
2005	16 900	16 900	0.00	5
2006	19 430	19 034	-2.04	9
2007	21 190	21 190	0.00	9
2008	22 450	22 450	0.00	7
2009	23 790	23 790	0.00	6
2010	26 090	26 090	0.00	8

由表3可知，除1994年相对误差为9.01%外，其他误差均小于3%，拟合效果较好；并且观察不同阶段分位点选取可以发现呈现一定周期规律，周期约为14年。取2006年为分位数波峰，由文献[7]得该市电力需求饱和时间段约为2025~2027年，经推算饱和年份电力需求处于波谷期，则饱和年分位点为0.3~0.1，饱和最高负荷范围约为39 846.56~41 591.7 MW，与报告预测结果相近。与饱和负荷预测中广泛采用的多变量协整模型预测结果相对比，结果如表4所示。与多变量协整模型拟合结果比较可知，本文提出的方法预测结果较好。

表4 华东某市分位数回归模型与多变量协整模型预测结果对比

使用模型	多变量协整模型	分位数回归模型
预测结果/MW	44 576~53 501	39 847~41 592
相对误差 (%)	7.28~28.76	0.1~4.1

结束语

与传统的OLS只得到均值回归结果相比，分位数回归能够更详细地描述变量的统计分布。本文

利用这一特点，将分位数回归理论应用于负荷预测，得到了以下结论：

1) 基于线性回归模型，利用分位数技术计算模型参数，以日本作为典型国家地区，对分位点的选取在对电力需求预测中的特点作出分析，得到了在不同发展阶段各社会经济解释变量对电力需求的不同影响程度。

2) 在回归预测用电量时，以拟合值同实际值相对误差最小为目标，确定最优分位点，并对同年用电量增长率序列进行相关性分析，发现电力需求增长率与最优分位点选取变化的正相关性，且具有周期性波动的规律，并提出利用这一规律预测城市电力需求饱和规模的方法。

实践证明，本文的研究对于分析电力需求影响因素权重、确定回归分析时分位点的选取规律以及基于分位数回归的饱和负荷的预测提供了有益的借鉴。

参考文献

[1] 刘晨晖. 电力系统负荷预报理论与方法 [M]. 哈

尔滨：哈尔滨工业大学出版社，1987.

[2] 廖旋焕，胡智宏，马莹莹，等. 电力系统短期负荷预测方法综述 [J]. 电力系统保护与控制，2011，39 (1)：147-152.

[3] 杜莉，张建军. 神经网络在电力负荷预测中的应用研究 [J]. 计算机仿真，2011，28 (10)：297-300.

[4] 万昆，柳瑞禹. 区间时间序列向量自回归模型在短期电力负荷预测中的应用 [J]. 电网技术，2012，36 (11).

[5] 顾洁，薛季良，杨宗麟，等. 基于扩展索洛长期均衡模型的饱和电力需求预测 [J]. 中国电力，2014 (7)：62-66.

[6] 刘晓娟，方建安. 稳健回归方法在电力消费预测中的应用 [J]. 电力系统及其自动化学报，2013，25 (5)：22-25.

[7] 杨宗麟，程浩忠，顾洁，等. 华东电网四省一市未来饱和负荷水平研究 [R]. 上海：华东电网有限公司，上海交通大学，国家发改委能源研究所，2013.

(收稿日期：2015-12-15) EA

(上接第124页)

和“轴心轨迹图”按钮，将会在页面下方显示结果，图5中右侧可以看出频谱分析，左侧可以看到轴心轨迹。根据前面所介绍的理论知识，特征频率为一倍频，并且幅值很大，根据轴心轨迹为椭圆形，可以判断是转子不平衡故障。

结束语

本文设计了基于Web和OPC技术的煤矿井下水泵机组远程故障诊断系统，详细介绍了系统通信的开发、软件的设计，并且从选取1#水泵机组的小波消噪、信号处理等界面对系统进行了验证，结果表明达到了预期的效果，实现了工作环境和操作环境的分离，有效地提高了煤矿井下水泵机组的运行效率，避免了不必要的安全事故，为煤矿的智能化夯实了基础，使煤矿井下安全水平迈上了一个新台阶。

参考文献

[1] 周晓峰，王结实，阚恒. 我国煤矿安全现状的思考 [J]. 中州煤炭，2009 (5)：110-112.

[2] 王坚，张浩，严隽蔚，等. 基于Internet/Intranet的数字化远程故障诊断系统 [J]. 计算机工程，2000 (10)：28-29.

[3] 蔡俊，朱成杰，郭来功. 矿业水泵房自动排水监控系统设计 [J]. 煤矿机械，2011，32 (4)：30-32.

[4] 李红. 基于B/S模式的机组远程振动监测与故障诊断系统研究 [D]. 北京：华北电力大学，2006.

[5] 杨晓冬. 水泵状态监测与故障诊断系统 [D]. 徐州：中国矿业大学，2008.

[6] 钱东海，李福军. 基于Web和OPC技术的远程监控系统的研究 [J]. 计算机应用与软件，2008 (2)：170-172.

(收稿日期：2015-03-18) EA