

基于支持向量机模型的北京市能源需求预测研究

张跃军¹, 周彬², 王丽³

(1. 北京理工大学 管理与经济学院 能源与环境政策研究中心, 北京 100081;
2. 北京理工大学 光电学院, 北京 100081; 3. 中国人民银行 营业管理部, 北京 100045)

摘要:北京市能源需求系统具有非线性、历史数据较少而影响因素众多等复杂特征,而支持向量机模型在解决小样本、非线性及高维模式识别问题方面具有突出优势。为此,引入支持向量机模型对北京市1978—2010年能源需求进行建模,并据此对2012—2020年能源需求量进行预测。结果表明:支持向量机模型能有效拟合北京市能源需求系统的复杂变化趋势,比其他传统方法有更高的预测精度。研究发现,2012—2020年北京市能源需求量逐年增加,年均增速2.75%;另外,北京市能源需求的增速在“十三五”期间会比“十二五”期间略有趋缓。

关键词: 能源需求; 支持向量机; 预测模型

中图分类号: F426, C931

文献标志码: A

文章编号: 1009-3370(2013)03-0008-05

能源是经济社会发展的重要物质基础。“十一五”期间,北京市经济平稳较快发展,年均增速达到11.4%,而能源消费量也不断提高,年均增速为4.7%。2011年末,北京市常住人口突破2 000万,GDP增长8.1%,能源消费量再攀新高,接近7 000万吨标准煤。但是,北京市能源资源非常匮乏,90%的能源消费量依赖于外省调入。同时,北京作为中国首都,以能源的持续发展支撑经济的持续发展,具有重要的政治和社会意义。因此,了解影响北京市能源需求变化的主要因素,引入合适的模型方法预测北京市未来能源需求的变化趋势,对于首都能源运行监测和预警、经济稳定及安全发展都具有重要意义。

过去几十年,国内外许多学者或研究机构都对能源需求预测进行过研究。现有文献对能源需求的预测不仅包括能源需求总量预测,包括电力、煤炭、石油、可再生能源等能源品种需求预测^[1-2],还包括一些部门的能源需求预测,如交通部门能源需求预测等^[3]。在研究过程中,现有文献形成了一些常用的预测方法,如消费弹性法、灰色模型、神经网络、粒子群优化模型、LEAP模型、时间序列ARIMA模型等^[4-7]。这些传统能源需求预测方法为能源系统建模和能源发展规划等提供了重要的科学支持。此外,近些年鉴于现有研究的不足,部分学者尝试引入前沿的软计算方法讨论能源需求预测问题,如模糊逻辑方法^[8]、微粒群优化算法与蚁群算法、神经网络或

遗传算法等相结合的混合算法^[9-11]、灰色马尔科夫模型^[12]等。

实际上,能源需求系统是一个典型的非线性系统,各种影响因素相互渗透,客观上增加了能源需求预测的复杂性,而部分传统预测方法(如消费弹性法)对样本数据没有学习过程,难以准确刻画能源系统中的非线性关系,导致预测精度可能不够;一些前沿的软计算预测方法又过于复杂,同时也存在一定主观性,可能减弱能源需求预测的性能。另外,北京市能源需求可用的历史数据较少,导致常见的时序模型的预测可靠性不够;采用神经网络模型来预测会由于样本得不到充分训练而导致预测性能不稳定,并且还存在着推广能力不强和“过学习”等缺点^[13];而灰色模型预测的实质是指数预测,原数据具有指数递增特性时预测效果较好,否则效果可能不佳。因此,预测北京市未来能源需求需要引入新的模型方法。

实际上,已有部分文献开始定量研究北京市能源需求问题。例如,张兴平^[14]对北京市17个行业2005—2009年间的能源消费变动进行因素分解分析,发现生产效应是导致北京市能源消费增长的主要因素,结构效应和强度效应则延缓了能源消费的增长,但生产效应明显大于结构效应和强度效应之和,导致能源消费在样本区间内不断增加。胡军峰^[15]引入面板协整模型分析了北京市能源消费和经济增长的关系,发现短期内存在能源消费到经济增长

收稿日期: 2012-12-05

基金项目: 国家自然科学基金资助项目(71001008,71273028,71020107026); 高等学校博士学科点专项科研基金资助项目(20101101120041); 教育部人文社会科学研究基金资助项目(09YJC630011); 北京市优秀人才培养资助计划(2011D009011000008)

作者简介: 张跃军(1980—),男,副教授,管理学博士,E-mail: zyjmis@126.com

的单向因果关系,长期来看能源消费和经济增长之间存在双向因果关系。张峰和刘伟^[6]采用时间序列ARIMA模型预测了北京市2007—2015年的能源消费量,结果认为,北京市能源消费量将呈加快增长态势,2015年达到1.1亿吨标准煤。但近几年的实际数据表明,他们的预测结果明显偏大。

总体而言,讨论北京市能源需求变化特征,预测北京市未来能源需求的文献还很少,特别是现有常用预测模型方法的可靠性还值得商榷,而支持向量机(SVM)模型是一种可取的选择。它具有可靠的统计学习理论基础和优越的学习性能,是一种专门研究有限样本情况下非参数估计问题的机器学习新方法,在解决小样本、非线性及高维模式识别问题方面与传统方法相比优势突出,并已成功应用于文本自动分类、多维函数预测等领域^[7]。

鉴于此,本文在分析影响北京市能源需求主要因素的基础上,针对其能源需求系统具有非线性、历史数据较少以及影响因素众多等复杂特征,引入支持向量机模型,预测北京市2012—2020年能源需求变化趋势,以期有关部门的决策提供依据。

一、研究方法 with 数据说明

(一)影响能源需求的主要因素

能源需求系统是一个非线性复杂系统,能源需求量受人口、经济、社会等众多因素影响。预测北京市能源需求变化趋势,需要根据其消费特征,归纳主要影响因素,并考虑它们之间错综复杂的关系。本文认为,影响北京市能源需求变化的主要因素可以归纳如下几个方面:

1. 经济发展水平。能源是经济发展的动力,以能源的可持续发展支持经济的可持续发展是政府的工作目标。未来几年,北京市经济发展“稳中求进”,还会拉动能源需求不断增加。因此,北京市经济发展水平是影响能源消费的关键因素。

2. 人口规模和城镇化。人口总量的多少直接影响能源消费量;而且,由于城镇居民和农村居民的能源利用方式和消费水平有很大差别,城镇居民的能耗水平一般比农村居民的能耗水平高。因此,随着北京市城镇化进程的加快,全市能源需求还会因此而不断增长。

3. 第二产业比重。由于第二产业的能耗指数远高于第一、第三产业,2011年北京市第二产业比重为23.4%,远高于发达国家的大城市,未来随着北京市产业结构的逐步调整,特别是第二产业比重的调整,能源消费量将受到影响。

4. 技术进步。技术进步对能源需求量的影响主要包括两个方面:一是由于科技水平提高特别是节能技术进一步发展,节能率提高,从而降低能源消费;二是技术进步使得新能源开发利用成为可能,导致能源消费结构改变,从而影响北京市能源需求量的变化。

5. 居民生活消费水平。随着人们生活水平不断提高,消费观念和行出现很大变化,人们不再满足于基本的消费需求,而且要求高级的消费享受,从而增加了对能源直接消费的需求,特别是对电力、液体和气体燃料等优质能源的需求。

此外,还有其他因素也会对能源需求产生影响,如北京市节能减排政策、宏观经济政策、汽车限购政策等,但这些因素难以量化或没有相关的统计数据,因此暂时不予考虑。

(二)预测能源需求的支持向量机模型

由于影响能源需求变化的各种因素与能源需求量的关系非常复杂,很难用一个具体的解析模型描述清楚,因此,本文采用SVM模型来构建能源需求系统中输入(能源需求影响因素)与输出(能源需求量)的关系;即基于1978—2010年间北京市能源需求量以及各种影响因素的历史数据,构造一个多输入、单输出的支持向量机预测模型。

基于支持向量机(SVM)模型预测的基本思想^[7],将输入的各种能源需求影响因素 x^1, x^2, \dots, x^d 映射到一个高维特征空间 $(\varphi(x_1), \varphi(x_2), \dots, \varphi(x_n))$ 。由统计学习理论可知,可将原非线性模型转化为特征空间的线性回归模型如式(1)所示

$$f(x_j) = w^T \varphi(x_j) + b \quad (1)$$

其中, w, b 是模型中需要进行辨识的参数。基于结构风险最小化原则,对式(1)中需要辨识的参数进行处理,如式(2)所示

$$R_{\text{emp}}(f) = \sum_{i=1}^s C(e_i) + \lambda \|w\|^2 \quad (2)$$

其中, $R_{\text{emp}}(f)$ 为经验风险; $\|w\|^2$ 为置信风险; $C(e_i)$ 为损失函数。进一步根据SVM原则,求解式(2)等价于求解式(3)中的优化问题

$$\begin{aligned} \min L &= \frac{1}{2} w^T w + C \sum_{i=1}^s \varepsilon_i^* + \varepsilon_i \\ \text{s.t.} &\begin{cases} y - (w, \varphi(X_i)) - b \leq \varepsilon + \varepsilon_i^* \\ (w, \varphi(X_i)) + b - y \leq \varepsilon + \varepsilon_i \\ \varepsilon_i^*, \varepsilon_i \geq 0 \end{cases} \end{aligned} \quad (3)$$

为了便于求解,往往把式(3)转化为对偶问题,则可得非线性函数 $f(x)$

$$f(x) = \sum_{i=1}^l (a_i - a_i^*) K(X_i, X) + b \quad (4)$$

其中, a_i 和 a_i^* 是支持向量参数; $K(X_t, X)$ 是内积函数。根据 Mercer 条件, 定义核函数, 选择径向基核函数(RBF)

$$K(X_t, X) = \exp\left\{-\frac{\|x_j - x_v\|^2}{\sigma^2}\right\} \quad (5)$$

将式(5)代入式(4)中, 经过等价变换可得到式(6)

$$f(x) = \sum_{i=1}^l a_i \exp\left\{-\frac{\|x_j - x_v\|^2}{\sigma^2}\right\} + b \quad (6)$$

其中, a_j 是支持向量对应的参数值; x_j 是训练年份的

输入数据向量; x_v 是预测年份的输入数据向量; $f(x)$ 为输出向量集合。通过式(6)得到能源需求预测参数 a_j 和 b , 并估计出能源需求预测模型。本文的计算都是通过 Matlab 软件编程实现的。

(三) 数据说明

本文以 1978—2010 年北京市能源需求相关数据为样本, 其中, 采用 1978—2003 年的数据构建模型, 并以 2004—2010 年的数据对模型的预测性能进行测试。样本数据来自历年《北京市统计年鉴》。研究过程中用到的变量及其含义如表 1 所示。

表 1 变量及其含义说明

变量名称	变量含义
居民消费水平	北京市居民在物质产品和劳务的消费过程中, 对满足人们生存、发展和享受需要方面所达到的程度
人口总数	北京市常住人口总数
城镇人口比例	北京市城镇化水平, 采用北京市城镇常住人口与总常住人口之比
第二产业比重	北京市第二产业增加值在 GDP 中所占比重
技术进步	北京市单位 GDP 的能耗水平
实际 GDP	北京市经济发展水平, 按 1978 年不变价计算得到
能源消费总量	北京市用于生产、生活所消费的各种能源数量之和

此外, 本文的能源需求系统考虑的 7 个输入、输出指标的量纲并不一致, 数据在量级上差异很大, 可能导致预测结果变动较大、性能降低。因此, 需要预先对各指标数据进行归一化处理, 即通过式(7)把各种指标数据转化到[0, 1]之间

$$x = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (7)$$

二、研究结果分析

(一) 数据训练及预测模型精度分析

首先, 将北京市能源需求影响因素作为支持向量机(SVM)模型的输入, 把能源消费量作为输出, 对 1978—2003 年的历史数据(训练样本)进行模拟与仿真。

在进行样本训练与预测时, 需要确定 SVM 正规化参数 C 和 RBF 核函数参数 σ^2 。将最终预测结果所产生的误差作为评价标准, 根据验证集上的性能表现, 来确定合适的取值。本文经过反复试验, 最后确定当 $C=6000$, $\sigma^2=100$ 时, 预测结果比较理想; 其中, 通过训练可获得 SVM 模型参数的 a_j 值, 当 a_j 不为零时对应的输入样本是支持向量(训练样本点); b 值为 1.195 6。这时将所得的参数 a_j 和 b 代入式(6)中, 估计出能源需求预测模型。

为了验证所建模型的有效性, 采用模型对北京市 2004—2010 年能源消费量(测试样本)进行预测, 并将预测结果与实际能源消费量进行比较, 结果如表 2 所示。

表 2 基于支持向量机的北京市能源需求预测结果

年份	实际值/万吨标准煤	预测值/万吨标准煤	预测误差/%
2004	5 139.60	5 135.94	0.07
2005	5 521.90	5 468.55	0.97
2006	5 904.10	5 775.43	2.18
2007	6 285.00	6 128.68	2.49
2008	6 327.10	6 379.58	-0.83
2009	6 570.30	6 720.01	-2.28
2010	6 954.10	7 168.25	-3.08

由表 2 可知, 采用支持向量机模型预测北京市 2004—2010 年能源需求量, 预测误差都在 5% 以内, 平均误差 1.70%, 实际值与预测值的相关系数达到

0.987, 说明预测结果非常接近能源消费的真实值, 即支持向量机模型对北京市能源需求的预测效果优异, 预测能力或推广能力突出, 可以作为预测北

京市未来能源需求变化的有效方法。

此外,为了进一步验证 SVM 模型在北京市能源需求预测方面的突出优势,本文也引入常用的经济预测模型 GM(1,1)和 BP 神经网络,并按照类似的

思路,对 1978—2003 年的能源需求相关指标的历史数据进行建模,然后基于模型对 2004—2010 年北京市能源需求量进行预测,以检验模型的预测性能。预测误差如表 3 所示。

表 3 基于 GM(1,1)模型和 BP 神经网络模型的北京市能源需求预测结果

年份	实际值/万吨标准煤	GM 模型预测值/ 万吨标准煤	误差/%	BP 神经网络模型预测值/ 万吨标准煤	误差/%
2004	5 139.60	4 832.00	5.98	5 022.30	2.28
2005	5 521.90	5 025.80	8.98	5 213.10	5.59
2006	5 904.10	5 227.40	11.46	5 532.50	6.29
2007	6 285.00	5 437.10	13.49	5 943.20	5.44
2008	6 327.10	5 655.20	10.62	6 132.90	3.07
2009	6 570.30	5 882.10	10.47	6 342.70	3.46
2010	6 954.10	6 118.00	12.02	6 632.60	4.63

比较结果显示,GM(1,1)模型的预测误差明显超过 5%,大部分年份的预测误差超过 10%,平均预测误差 10.43%,说明预测效果并不可取,可见,虽然 GM(1,1)模型在小样本预测方面具有一定优势,但对北京市能源需求预测的结果并不如意,预测精度较差。另外,BP 神经网络的预测性能虽然要比 GM(1,1)好,预测误差都在 5%附近,平均预测误差 4.39%,但与支持向量机模型的预测误差(平均预测误差 1.70%)相比,其预测性能明显逊色。

因此,整体上看,可以认为支持向量机模型在北京市能源需求预测方面的性能明显优于其他两种常见预测方法。平均而言,在本文样本区间内,支持向量机模型的预测精度相对于 GM(1,1)和 BP 神经网络分别提高了 84%和 61%。所以,可以采用支持向量机模型对北京市未来能源需求变化进行预测,为有关部门监测能源运行形势,制定能源发展规划等提供科学依据。

(二)北京市未来能源需求预测结果分析

鉴于支持向量机模型突出的预测性能,本文基于训练得到的模型,对北京市 2012—2020 年的能源需求量进行预测,结果如图 1 所示。

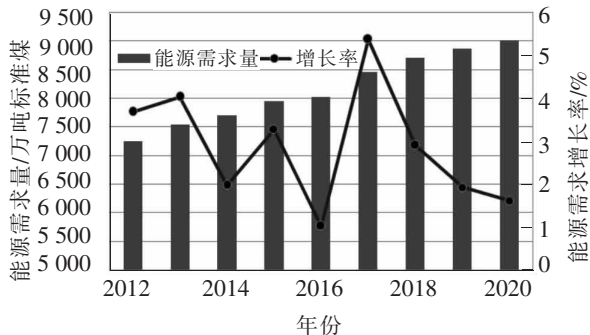


图 1 北京市 2012—2020 年能源需求量预测结果

可以发现,一方面,“十二五”、“十三五”期间

(2012—2020 年),北京市能源需求量不断攀升,年均增速 2.75%,到 2020 年达到 9 008 万吨标准煤。可见,随着北京市工业化、城镇化快速推进,能源消费量较快增长与资源短缺约束的矛盾将更加突出。特别是“十二五”期间,北京市对能源特别是优质能源的需求将保持较快增长,年均增速 2.6%。即便如此,本文预计,“十二五”期末(2015 年)北京市能源需求达到 7 941 万吨标准煤,而《北京市“十二五”时期能源发展建设规划》提出,按照“保发展、促节能、控总量、留余地”的原则,2015 年全市能源消费总量控制在 9 000 万吨标准煤左右,因此,实现北京市政府提出的能源消费量控制目标并不困难。

另一方面,“十三五”期间(2016—2020 年),虽然北京市能源需求量仍然不断增长,但增速逐步下降,年均增速 2.3%,这与北京市经济发展水平和发展模式有关。展望未来,在我国经济发展速度趋缓、节能减排稳步推进的大环境下,北京市经济发展增速逐步放缓,能源效率不断提高,能源结构加快调整,能源体制机制不断完善,能源消费需求的增速逐步趋缓也是自然的结果。

三、结论

通过上述研究,可见使用支持向量机(SVM)模型预测北京市未来能源需求变化趋势,具有方法上简捷、实用等优势,也克服了部分传统预测方法“过学习”、精度较低等缺点,同时还避免了人为主观因素对建模的影响。数据结果表明,支持向量机模型通过综合考虑各种影响能源需求变化的因素,能够有效解决北京市能源需求系统小样本、非线性及高维模式识别等问题,适合北京市能源需求预测,预测精度较高。此外,预测结果表明,2012—2020 年,北京市能源需求不断走高,但“十三五”期间,随着

北京市经济发展增速降低,能源效率提高,能源需求的增速将趋缓。

实际上,北京市能源需求研究还有很多工作可以进一步开展。例如,一些重点耗能部门(如交通部

门)的能源需求评估和预测问题,全市节能潜力评估问题,全市清洁能源的需求管理问题,各区县的能源需求合理配置等。

参考文献:

- [1] Inglesi R. Aggregate electricity demand in South Africa: conditional forecasts to 2030[J]. *Applied Energy*, 2010, 87(1): 197–204.
- [2] Celiktas M S, Kocar G. From potential forecast to foresight of Turkey's renewable energy with Delphi approach [J]. *Energy*, 2010, 35(5): 1973–1980.
- [3] Daly H E, Gallachóir B P. Modelling future private car energy demand in Ireland[J]. *Energy Policy*, 2011, 39(2): 7815–7824.
- [4] Ünler A. Improvement of energy demand forecasts using swarm intelligence: the case of Turkey with projections to 2025[J]. *Energy Policy*, 2008, 36(6): 1937–1944.
- [5] Suganthi L, Samuel A A. Energy models for demand forecasting—a review [J]. *Renewable and Sustainable Energy Reviews*, 2012, 16(2): 1223–1240.
- [6] Huang Y, Bor Y J, Peng C Y. The long-term forecast of Taiwan's energy supply and demand: LEAP model application[J]. *Energy Policy*, 2011, 39(11): 6790–6803.
- [7] Ediger V S, Akar S. ARIMA forecasting of primary energy demand by fuel in Turkey[J]. *Energy Policy*, 2007, 35(3): 1701–1708.
- [8] Kucukali S, Baris K. Turkey's short-term gross annual electricity demand forecast by fuzzy logic approach [J]. *Energy Policy*, 2010, 38(5): 2438–2445.
- [9] Kran M S, Özceylan E, Gündüz M, Paksoy T. A novel hybrid approach based on particle swarm optimization and ant colony algorithm to forecast energy demand of Turkey[J]. *Energy Conversion and Management*, 2012, 53(1): 75–83.
- [10] Yu S, Wei Y M, Wang K. China's primary energy demands in 2020: predictions from an MPSO-RBF estimation model[J]. *Energy Conversion and Management*, 2012, 61: 59–66.
- [11] Yu S, Zhu K, Zhang X. Energy demand projection of China using a path-coefficient analysis and PSO-GA approach [J]. *Energy Conversion and Management*, 2012, 53(1): 142–153.
- [12] Kumar U, Jain V K. Time series models (Grey-markov, grey model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India[J]. *Energy*, 2010, 35(4): 1709–1716.
- [13] Murat Y S, Ceylan H. Use of artificial neural networks for transport energy demand modeling [J]. *Energy Policy*, 2006, 34(17): 3165–3172.
- [14] 张兴平, 汪辰晨, 张帆. 北京市能源消费的因素分解分析[J]. *工业技术经济*, 2012(1): 13–18.
- [15] 胡军峰, 赵晓丽, 欧阳超. 北京市能源消费与经济增长关系研究[J]. *统计研究*, 2011(3): 79–85.
- [16] 张峰, 刘伟. 北京市能源消费预测与政策建议[J]. *中国人口·资源与环境*, 2008, 18(3): 99–102.
- [17] 邓乃扬, 田英杰. 数据挖掘中的新方法: 支持向量机[M]. 北京: 科学出版社, 2004.

Energy Demand Projection Based on Support Vector Machine Model: Evidence from Beijing

ZHANG Yuejun¹, ZHOU Bin², WANG Li³

(1. Center for Energy and Environmental Policy Research, School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China;

2. School of Optoelectronics, Beijing Institute of Technology, Beijing 100081, China;

3. Business Management, The People's Bank of China, Beijing 100045 China)

Abstract: The energy demand system of Beijing has a number of features including nonlinearity, limited historical data and numerous drivers, while the Support Vector Machine (SVM) model owns unique advantage in small samples, nonlinear and high-dimensional pattern recognition. Therefore, the SVM model is employed to fit the related historical data about energy demand in Beijing from 1978–2010 and then project the energy demand during 2012–2020. The results indicate that the projection power of SVM model evidently outweighs that of other traditional and commonly-used models, which can effectively consider the complex features in energy demand system of Beijing. Additionally, the projection results also suggest that Beijing's energy demand may increase year by year during 2012–2020, with the average annual growth rate 2.75%; and its growth rate may appear relatively slower in the 13th Five-Year Plan period than that in the 12th Five Year Plan period.

Key words: energy demand; support vector machine; projection model