

中国民用汽车保有量的组合预测

古继宝 荣 敏 吴剑琳
(中国科学技术大学, 合肥 230026)

摘 要 本文对中国民用汽车保有量进行了中短期的预测。文章基于组合预测误差平方和最小的条件, 将指数平滑法和 BP 神经网络有机结合, 从而提高了预测的精度和稳定性。并通过对中国民用汽车保有量的组合预测结果的分析, 提出有利于中国汽车产业发展的政策建议。

关键词 民用汽车 组合预测 BP 神经网络
中图分类号 F407.471 **文献标识码** A

引 言

中国民用汽车的消费在近几年发展迅速, 在汽车整体消费中所占比例逐年上升, 对汽车产业及其关联产业起着重要影响, 汽车工业产业是现代工业发展的一个主要推动力, 其产业关联度很高, 可以带动很多其他的产业部门的发展, 如钢铁、橡胶、玻璃、电子和化工工业。目前对汽车市场的研究主要集中于整体需求或保有量预测, 对民用和私人汽车购买的针对性研究很少。

汽车产业作为我国经济发展中的支柱产业, 对国民经济走势的影响越来越大, 而民用汽车消费在汽车市场所占比重逐年增大。汽车的私人消费也成为汽车市场发展最主要的拉动因素, 对于民用汽车保有量的研究预测有利于我们对国家宏观经济走势的判断, 同时, 汽车保有量预测也是我国道路发展规划和确定石油发展战略及产业结构调整的必要要求。可以说, 对中国民用汽车保有量的预测研究有利于制定合理的汽车产业政策和国民经济政策。

1 文献回顾

近年来, 对于汽车行业的宏观方向的预测技术很多, 下面介绍的是一些常用的具体方法。

1.1 时间序列预测法

利用过去的资料预测未来状态。时间序列模型的突出特点在于它需要的数据量很小, 并且特别适合短期预测, 它的缺陷是无法考虑诸如人口统计因素的影响, 单一的用时间序列方法进行预测的文章较少。韩亮, 冯小明 (1999) 用弹性系数法和指数平滑法预测轿车市场生产量。

1.2 人工神经网络法

神经网络方法预测的精度较高, 但神经网络方法的稳定性稍差, 隐层层次和节点数只能由试算决定。何明, 过秀成 (2007) 采用主成分分析的方法提炼出较少的与

线性无关的主要因素, 并根据这些因素, 利用 BP 神经网络方法对汽车保有量进行了预测, 最后通过实例, 将 BP 神经网络主成分分析法计算结果和非线性模拟与全要素 BP 神经网络模拟结果进行比较, 得知 BP 神经网络主成分分析法在运算效率、运算精度上较优。吴义虎, 侯志祥利用改进的 BP 神经网络对中国 2000 年, 2005 年, 2010 年的汽车保有量进行预测, 并与灰色预测的结果进行对比, 结果表明, 由于神经网络具有很强的学习与泛化能力, 因此在处理具有一定程度不确定性的非线性系统的建模与预测方面, 神经网络有很好的应用价值。

1.3 组合预测法

考虑到单一预测方法的局限, 组合预测把不同的预测模型组合起来, 综合利用各种预测方法所提供的信息, 以适当的加权平均形式得到组合预测模型。潘志刚、韩颖 (2006) 在我国汽车市场的需求预测中将指数平滑法、灰色系统法和多元回归法进行组合, 提高了预测的精度。王玉梅, 尚金城 (2007) 均运用组合预测法, 将指数平滑法和灰色系统法等方法预测出的结果进行加权组合, 同样提高了预测精度, 避免了单一预测的局部性和不稳定性。

2 模型建立

2.1 组合预测模型

组合预测就是要充分利用每一种预测方法中所包含的独立信息, 确定每种单一预测方法的权重是组合预测的核心内容。在假定各种单一预测方法的结果是无偏的情况下, 在已知各种预测方法的预测结果后, 本文通过最小方差法来确定各种预测方法的权重。赵韩, 许辉, 梁平等 (2008) 建立了关于组合预测的最小方差模型, 其具体的实现方法如下:

对同一预测问题有 m 种模型, 对 n 个时段进行预测, 组合预测的目的就是要找到满足方差和最小时各预

收稿日期: 2009—10—24

测方法的权重，其中：

y_t ——第 t 时段的实际值
 $\bar{y}_t$ ——组合预测方法对第 t 时段的预测值
 $\bar{y}_{it}$ ——第 i 种预测方法对第 t 时段的预测值
 e_t ——组合预测方法对第 t 时段的预测误差
 e_{it} ——第 i 种预测方法对第 t 时段预测误差
 w_i ——第 i 种预测方法的权数
 $e_t^2 = e_t' e_t$ (1)
 $Z = \sum e_t^2$ (2)
 $W = (w_1, w_2, \dots, w_m)^T$

$$E = \begin{bmatrix} \sum_{t=1}^n e_{1t}^2 & \sum_{t=1}^n e_{1t} e_{2t} & \cdots & \sum_{t=1}^n e_{1t} e_{mt} \\ \sum_{t=1}^n e_{2t} e_{1t} & \sum_{t=1}^n e_{2t}^2 & \cdots & \sum_{t=1}^n e_{2t} e_{mt} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{t=1}^n e_{mt} e_{1t} & \sum_{t=1}^n e_{mt} e_{2t} & \cdots & \sum_{t=1}^n e_{mt}^2 \end{bmatrix}$$

则 $Z = W^T E W$

最优的权数为 $W = \frac{E^{-1} R}{R^T E^{-1} R}$ ，其中 $R = (1, 1, \dots, 1)^T$ (3)

本文通过指数平滑法和 BP 神经网络法的组合预测对中国民用汽车保有量进行研究，指数平滑法充分考虑了中国民用汽车保有量本身的时间序列的对未来的影响；BP 神经网络则通过建立网络模型考虑到中国民用汽车保有量的各影响因素对未来的影响，对以上两种方法进行组合就可以全面考虑到中国民用汽车保有量本身及其各影响因素，使预测包含了尽可能多的信息。在得到单一方法的预测结果后，通过以上组合预测的最小方差法可以得到这两种方法的权重。

2.2 指数平滑模型

指数平滑法一种时间序列分析方法，除了比较直观，其最大的优点是在进行预测时，只需要知道实际数值和本期预测值就可以预测下一期的数值，而不需要知道以往的观测值。本文要预测研究的是中国民用汽车保有量，指数平滑法只需要考虑到其本身的时间序列值就可以进行分析预测。

2.3 BP 神经网络模型

BP 网络是人工神经网络其中应用得最广泛的一种非线性动力系统。BP 神经网络是一种误差后向传播网络，一般采用 3 层网络型式，即由输入层、隐含层和输出层 3 层神经元组成，各层神经元的作用都是不同的，其学习过程是由正向传播和误差反向传播所组成。

BP 神经网络的预测模型必须考虑到影响中国民用汽车保有量的因素，从以往的研究中可以看出，中国民用汽车保有量的解释变量有很多，通过变量的分析，选取实际人均可支配收入、实际社会消费品零售总额、相对公路总运输路线的长度、实际人均 GDP、从业人员作为影响使中国民用汽车市场宏观消费量的自变量，把汽车的保有量作为因变量，建立 BP 神经网络的预测模型。

本研究将实际人均可支配收入、实际社会消费品零售总额、相对公路总运输路线的长度、实际人均 GDP、从业人员这 5 个变量作为 BP 神经网络的输入单元，相应年份的汽车保有量作为网络的输出单元。这样神经网络的输入、输出之间就形成一种对应关系，它形成的训练样本中有 5 个值为输入量，即该预测网络输入层有 5 个神经元；1 个值为输出量，即输出神经元为 1 个。对于隐层神经元个数的确定要根据经验和设计者多次试验来确定，不存在一个理想的解析式，通过参考相关资料，确定隐层神经元个数为 12。再者，由于 BP 系统是非线性的，初始值对于学习能否达到局部最小和是否能够收敛的结果关系很大。一般会要求初始权值在输入累加时使每个神经元的状态值接近于零，权值一般取随机数，而且比较小。输入样本也同样要求进行归一化处理，使那些比较大的输入仍然落在传递函数梯度大的地方。

3 数据处理与分析

本文是对中国民用汽车保有量的宏观预测，相关数据可以从中国统计年鉴得到，以下是 1990~2007 年中国民用汽车拥有量及其影响因素具体包括城镇居民年人均可支配收入、社会消费品零售总额、实际人均国内生产总值、从业人员、相对公路运输线路长度的数据，如表 1 所示。

表 1 1990~2007 年中国民用汽车保有量及其各影响因素数据

年份	城镇居民年人 均可支配收入 (元)	社会消费品零售 总额 (亿元)	实际人均国内生产总值 (元)	从业人员 (万人)	相对公路运输 线路长度 (万公里)	民用汽车拥有量 (万辆)
1990	1510.2	8300.1	1644.0	64749	102.83	551.36
1991	1700.6	9415.6	1892.8	65491	104.11	606.11
1992	2026.6	10993.7	2311.1	66152	105.67	691.74
1993	2577.4	14270.4	2998.4	66808	108.35	817.58



续 表

年份	城镇居民年人 均可支配收入 (元)	社会消费品零售 总额 (亿元)	实际人均国内生产总值 (元)	从业人员 (万人)	相对公路运输 线路长度 (万公里)	民用汽车拥有量 (万辆)
1994	3496. 2	18622. 9	4044. 0	67455	111. 78	941. 95
1995	4283. 0	23613. 8	5045. 7	68065	115. 7	1040. 01
1996	4838. 9	28360. 2	5845. 9	68950	118. 58	1100. 08
1997	5160. 3	31252. 9	6420. 2	69820	122. 64	1219. 09
1998	5425. 1	33378. 1	6796. 0	70637	127. 85	1319. 30
1999	5854. 0	35647. 9	7158. 5	71394	135. 17	1452. 94
2000	6280. 0	39105. 7	7857. 7	72085	140. 27	1608. 91
2001	6859. 6	43055. 4	8621. 7	73025	169. 8	1802. 04
2002	7702. 8	48135. 9	9398. 1	73740	176. 52	2053. 17
2003	8472. 2	52516. 3	10542. 0	74432	180. 98	2382. 93
2004	9421. 6	59501. 0	12335. 6	75200	187. 07	2693. 71
2005	10493. 0	67176. 6	14053. 0	75825	334. 52	3159. 66
2006	11759. 5	76410. 0	16165. 0	76400	345. 70	3697. 35
2007	13785. 8	89210. 0	18934. 0	76990	358. 37	4358. 36

资料来源：1991～2008 年中国统计年鉴

在 spss 软件的辅助下对 1990～2007 年中国民用汽车保有量进行指数平滑建模和分析，可以得到其指数平滑的预测值。

在 matlab 软件中建立 BP 神经网络模型，以 1990～1999 年的数据作为网络的训练样本，2000～2007 年的数据作为网络的预测效果检验样本，通过 BP 网络训练实现最优化。为了提高 BP 神经网络的效果和精度，在网络训练以前需对各个变量的数据进行归一化处理，本文的归一化公式为 $x_i = 0.1 + 0.8 * (X_i - X_{min}) / X_{max} - X_{min}$ ，由此得到相应的数据代入到神经网络模型中进行训练，

通过训练得到的 BP 神经网络拟合效果很好，在误差精度为 0.001 时，通过 143 次训练 BP 神经网络达到最优拟合，通过训练好的 BP 神经网络，导入相应各个变量值之后可以得到中国民用汽车保有量的预测值。

为了对未来中短期的民用汽车保有量进行预测，我们也需要未来的各个变量的输入值，因此，在对民用汽车保有量进行预测前，我们先用灰色预测的方法对各个民用汽车保有量影响因素的变量进行了未来的短期预测，并把该短期预测值作为 BP 神经网络的输入值，相关预测数值如表 2 所示。

表 2 2008～2012 年中国民用汽车保有量各影响因素灰色预测值

年份	城镇居民人均可 支配收入 (元)	社会消费品零售总额 (亿元)	实际国内人均生产 总值 (元)	从业人员 (万人)	相对公路运输线路长度 (万公里)
2008	15153. 0	99480. 0	21585. 7	77733. 3	452. 99
2009	17042. 0	112943. 0	24873. 4	78406. 4	534. 63
2010	19166. 5	128228. 1	28661. 9	79085. 4	630. 98
2011	21555. 8	145581. 7	33027. 5	79770. 2	744. 70
2012	24243. 0	165283. 8	38058. 0	80461. 0	878. 92

由上述指数平滑法和 BP 神经网络法对中国民用汽车保有量分别进行预测得到 2001～2007 年的预测值，两

种单一预测方法的结果如表 3 所示。

表 3 2001~2007 年中国民用汽车保有量预测值与实际值比较

年份	实际值	BP 神经网络			指数平滑		
		预测值	绝对误差	相对误差	预测值	绝对误差	相对误差
2000	1608. 91	1690. 6	81. 69	0. 050774	1586. 58	22. 33	0. 013879
2001	1802. 04	1849. 55	47. 51	0. 026365	1764. 88	37. 16	0. 020621
2002	2053. 17	2139. 83	86. 66	0. 042208	1995. 17	58	0. 028249
2003	2382. 93	2364. 92	18. 01	0. 007558	2304. 3	78. 63	0. 032997
2004	2693. 71	2805. 1	111. 39	0. 041352	2712. 69	18. 98	0. 007046
2005	3159. 66	3083. 97	75. 69	0. 023955	3004. 49	155. 17	0. 04911
2006	3697. 35	3573. 17	124. 18	0. 033586	3625. 61	71. 74	0. 019403
2007	4358. 36	4125. 66	232. 7	0. 053392	4235. 04	123. 32	0. 028295

本文通过 BP 神经网络法和指数平滑法组合的最小方差法对民用汽车拥有量进行组合预测,按最优的组合预测方法,由公式 (3) 我们可以得到

$$E = \begin{vmatrix} 104471.443 & 61496.2523 \\ 61496.2523 & 56218.5907 \end{vmatrix}, \text{ 两种预测方法的权重}$$

$W = (0.158507, 0.841493)^T$,按照确定的权重可以容易的得出中国民用汽车保有量的组合预测值,以及组合的结果与其它两种单一预测的效果比较。

表 4 2001~2007 年中国民用汽车保有量组合预测值

年份	实际值	BP 神经网络		指数平滑		组合预测	
		预测值	相对误差	预测值	相对误差	预测值	相对误差
2000	1608. 91	1690. 6	0. 05077	1586. 58	0. 013879	1603. 07	0. 003631
2001	1802. 04	1849. 55	0. 02636	1764. 88	0. 020621	1778. 3	0. 013174
2002	2053. 17	2139. 83	0. 04221	1995. 17	0. 028249	2018. 1	0. 017081
2003	2382. 93	2364. 92	0. 00756	2304. 3	0. 032997	2313. 91	0. 028965
2004	2693. 71	2805. 1	0. 04135	2712. 69	0. 007046	2727. 34	0. 012484
2005	3159. 66	3083. 97	0. 02396	3004. 49	0. 04911	3017. 09	0. 045123
2006	3697. 35	3573. 17	0. 03359	3625. 61	0. 019403	3617. 3	0. 021651
2007	4358. 36	4125. 66	0. 05339	4235. 04	0. 028295	4217. 7	0. 032273

由表 4 可知,通过中国民用汽车保有量 2001~2007 年实际值与各种预测值的比较,组合预测的相对误差最高达到 0. 045123,最低为 0. 003631,预测精度较高,至于具体的各种预测方法效果的比较如表 5 所示。

表 5 各预测方法效果比较

模型	BP 神经网络	指数平滑	组合预测
平均相对误差	0. 034898	0. 024950	0. 021798
误差平方和	104471. 443	56218. 5907	54241. 9774

比较单一模型和组合预测模型的结果,组合预测模型的平均相对误差 0. 021798,低于神经网络和指数平滑预测的平均误差;误差平方和 54241. 9774 同样也小于单一模型预测结果的误差平方和,也就是说,组合预测在总体上起到了提高预测精度。在对已有数据的分析中表明,组合预测方法相对于指数平滑和 BP 神经网络的预测方法是有效和可性的。

表 6 2008~2012 年中国民用汽车保有量的预测值

年份	BP 神经网络预测值	指数平滑预测值	组合预测值	相对增长率
2008	4478.28	5019.37	4933.6	
2009	4615.33	5680.38	5511.56	11.7 %
2010	4748.58	6341.39	6088.92	10.5 %
2011	4962.72	7002.4	6679.1	9.7 %
2012	5557.56	7663.41	7329.62	9.7 %

本文对中国民用汽车保有量进行 2008~2012 年的组合预测结果如表 6 显示,中国民用汽车保有量在 2012 年时将达到 7329.62 万辆,在此期间,中国民用汽车保有量将会以 10 %左右的增长速度稳定快速增长。由于 2008 年的部分变量的数据难以搜集,无法以 2008 年的数据作为原始数据进行分析预测,但 2008 年中国民用汽车保有量的实际值为 4975.0 万辆,相对误差为 0.84 %,预测结果精度较高。

4 结 论

本文通过指数平滑法和 BP 神经网络的组合预测法对中国民用汽车保有量进行了中短期的预测研究,显然,组合预测模型通过各单一预测模型的预测误差优化得到其权重,综合了两种预测方法的有效信息,提高了预测的精度,使我们的预测结果具有一定的参考性。

对中国民用汽车保有量 2008~2012 年的预测结果表明,未来几年内我国民用汽车市场处在稳步成长期,民用汽车保有量也大幅增长,至 2012 年中国民用汽车保有量将达到 7329.62 万辆,年平均增长率达到 10 %左右,我国政府在实现汽车工业产业快速发展的同时也必须做好相关的政策准备,具体政策建议如下:

4.1 以产业政策为核心,制定汽车产业促进法和产业促进条例

随着民用汽车市场需求和供给的增加,汽车市场会形成一定程度的混乱,制定产业优化政策、联合重组政策等能使汽车企业的组织结构达到布局合理;实行汽车产品质量纠纷举证责任倒置制度,由汽车制造商和销售商证明汽车产品合格,避免因汽车大规模制造的增加导致的质量下降;加强消费者权益保护立法,侧重处理好汽车企业和消费者之间不平等的地位,保障汽车需求稳步健康的生长。

4.2 制定新能源汽车产业发展政策

我国民用汽车保有量的增长会加剧我国石油消耗速度,为了降低汽车产业对石油的依存度,发展新能源

汽车产业是必然的方向。政府可以在财税政策支持新能源汽车重点产品;在商业化推广方面,启动国家节能和新能源汽车,由中央财政安排资金对示范工程的业主单位给予财政补贴;在促进私人消费方面,研究制订对消费者购买纯电动、充电式混合动力、普通型混合动力汽车等新能源汽车的税收优惠和财政补贴实施办法;进一步制订新能源汽车发展规划,要求县级以上城市人民政府要制定规划,优先在城市公交、出租、公务、环卫、机场领域推广使用新能源汽车。

4.3 加快道路设施建设,缓解交通压力

汽车保有量的快速增长,势必对城市交通设施造成很大压力。因此,要加大城市交通和基础设施投资,加快道路新建、扩建、改建,增加停车泊位,改善用车环境,完善路网结构,构建立体交通枢纽,确保道路基本畅通,提高运行能力。在良好的用车环境得到保障后,汽车需求的增长和汽车产业的壮大会形成一个良性循环。

参 考 文 献

1. Button Kenneth, Nge Ndo, Hine John. Modeling vehicle ownership and use in low income countries [J]. Journal of Transport Economics and Policy, 1993, XXVII (1): 51~69

2. Clemen, R. T. and Winkler, R. L. Combining economic forecasts [J]. Journal of Business and Economic statistics, 1986, 4 (1): 39~46

3. Bunn D W. Combining forecasts [J]. European Journal of Operation Research, 1988, 33 (3): 223~229

4. 唐小我. 最优组合预测方法及其应用 [J]. 数理统计与管理, 1992, 11 (1): 31~35

5. 赵韩, 许辉, 梁平, 等. 最优组合预测方法在家用汽车需求预测中的应用 [J]. 工业工程, 2008, (1): 126~133

6. 王正武, 肖正军. 城市私人小汽车保有量预测 [J]. 重庆交通院学报, 2004, 23 (5): 112~114

7. 黄志刚, 汤洪, 丁胜春. 基于人工神经网络的湖南省汽车保有量预测 [J]. 山西科技, 2005, (2): 106~107

8. 飞思科技产品研发中心. 神经网络理论与 MATLAB7 实现 [M]. 北京: 电子工业出版社, 2005

作者简介 古继宝, 中国科学技术大学管理学院副教授, 博士。研究方向: 战略管理、科研管理。荣敏, 中国科学技术大学管理学院硕士研究生。研究方向: 战略管理。吴剑琳, 中国科学技术大学管理学院讲师, 博士。研究方向: 战略管理。