一组预测方法的比较分析

陈冬冬[1,2],彭其渊[3]
CHEN Dong-dong[1,2],PENG Qi-yuan[3]

1.西南交通大学 物流学院,成都 610031
2.四川农业大学 经济管理学院,四川 雅安 625014
3.西南交通大学 交通运输学院,成都 610031
1.College of Logistics,Southwest Jiaotong University,Chengdu 610031,China
2.Economic Management College,Sichuan Agricultural University,Ya'an,Sichuan 625014,China
3.College of Traffic and Transportation,Southwest Jiaotong University,Chengdu 610031,China
E-mail:dongchen15@126.com

CHEN Dong-dong,PENG Qi-yuan.Group of forecasting methods comparative analysis.Computer Engineering and Applications,2009,45(26):199-203.

Abstract: The forecast is the base of scientific decision-making,how to choose the right forecasting method,is very important for forecasting results.This paper first briefly introduces a group of forecasting methods on popular,then uses the data of total freight in 1978-2006,and using the methods which describe to forecast and determinate the error,and finally gives various methods a simple comments.The conclusion is proved that the complicated method of forecasting is not necessarily better than the simple one once again.The paper has a certain theoretical and practical significance for the further study of combination forecasting methods and laid the foundation for further research.
Key words: forecasting methods;time-series data;total freight;comparative analysis

摘 要:预测是科学决策的基础,如何选择合适的预测方法,对于获得好的预测结果具有重要影响。就现在比较流行的一组预测方法做简单介绍后,利用中国1978年~2006年货运周转量数据,用文中介绍的方法分别做出预测和预测误差测定,并对各种方法给出简单评述,并再次验证了复杂方法预测效果并不一定比简单方法好的结论。具有一定的理论和实践意义,为进一步研究组合预测方法奠定了基础。
关键词:预测方法;时间序列数据;货运周转量;比较分析
DOI:10.3778/j.issn.1002-8331.2009.26.060 **文章编号:**1002-8331(2009)26-0199-05 **文献标识码:**A **中图分类号:**F272.1;TP18

1 引言

在当前的社会经济生活中,所有组织都在不确定的环境中运营,都面临一个共同的问题,就是现在必须做出影响组织将来的决策。对未来做出有根据的预测对于组织管理人员来说比无根据的猜测更有价值。要想得到有效的预测结果,合理的判断必须基于科学的预测方法和严谨的数值结果,从而要求组织管理人员具有能力运用完善的数据分析技术进行预测,并且能理解这些方法的本质,从而避免因为采用的预测方法不当或不准确的预测可能导致的错误决策。

其中,由于每种具体的预测方法适用于不同的预测环境,一种方法适合于所有案例的情况很少见,因而需要在特定的环境下试用若干种方法,以比较和选择预测结果最精确的方法,并且还应考虑预测的及时性和易理解性等因素。

除了单一预测法外,Clemen(1989)在其里程碑性的文章中指出,通过对多个单一预测方法的结果的组合,可以从本质上提高组合预测的精度,同时指出简单的组合方法相对于复杂的组合方法表现更好。Armstrong(1989)也明确指出:组合预测比单一预测好,简单平均和复杂的统计方法预测的表现结果是一致的。

目的是利用1978年~2006年度我国货运周转量时间序列数据,应用简单移动平均法、指数平滑法、线性趋势预测法、二次曲线法、指数曲线、ARIMA法、灰色预测及其改进方法以及神经网络和支持向量机等方法做比较,系统分析各种方法的各种误差,从而对各种预测方法的适用性进行研究,为后续组合预测研究打下基础。

2 各种预测方法的简单介绍

2.1 简单移动平均预测方法

简单移动平均[1-4]利用所有数据的平均值进行预测。时间序列数据 $Y=\{y_1,y_2,\cdots,y_T\}$,T 是样本长度,公式(1)给出了简单移

基金项目:四川省哲社基金项目(No.C07B058)。
作者简介:陈冬冬(1974-),男,博士研究生,讲师,研究方向:物流工程、供应链管理、数量经济;彭其渊,院长,教授,博士生导师。
收稿日期:2008-10-28 **修回日期:**2008-12-03

动平均的预测。第 k 阶 $MA(k)$ 移动平均值可通过以下公式计算：

$$\hat{y}_{t+1}=\frac{y_t+y_{t-1}+\cdots+y_{t-k+1}}{k} \tag{1}$$

其中：$\hat{y}_{t+1}$ 表示下一期的预测值；y_t 表示第 t 期的实际值，k 表示移动平均期数。

时期 t 的移动平均值为最近 k 期观测结果的算术平均值。移动平均法只处理已知的最近 k 期数据，计算平均值数据点的数量并不随着时间提前而变化，该法对趋势和季节性数据的处理并不出色，但是简单易用和易理解。

二重移动平均可以处理线性趋势的时间序列数据，即以一次移动平均值的基础上，再做一次移动平均。计算公式为：

$$M_t=\hat{y}_{t+1}=\frac{y_t+y_{t-1}+\cdots+y_{t-k+1}}{k} \qquad M_t'=\frac{M_t+M_{t-1}+\cdots+M_{t-k+1}}{k}$$

$$a_t=M_t+(M_t-M_t')=2M_t-M_t' \qquad b_t=\frac{2(M_t-M_t')}{k-1}$$

预测未来 p 期的值为公式：

$$\hat{y}_{t+p}=a_t+b_t p \tag{2}$$

其中：k 表示移动平均值期数；p 表示预测值之前的期数。

2.2 指数平滑法预测

该方法根据最近的经验对估算值进行不断修正，以指数递减方式平均（平滑）[1-4]序列的历史值为基础。最近的观测结果赋予的权重 $\alpha(0<\alpha<1)$ 最大，下一个最近的观测值权重较小，为 $\alpha(1-\alpha)$，过去两期的权重为 $\alpha(1-\alpha)^2$，以此类推。指数平滑法的公式为：

$$\hat{y}_{t+1}=\alpha y_t+(1-\alpha)\hat{y}_t \tag{3}$$

该方法的关键是如何设定 α 值。当时间序列比较平稳和平滑的时候，该值设定较小；反之，当时间序列有较大随机波动时，就需要设置的较大一些。

2.3 线形趋势预测

线性趋势[1-4]指现象随时间的推移而呈现稳定增长或下降的线性变化规律。可以用线性趋势方程来描述：

$$\hat{y}_t=\alpha+\beta t \tag{4}$$

式中：$\hat{y}_t$ 代表时间序列 y_t 的趋势值，参数 α 和 β 分别为方程的截距和斜率，可用最小二乘法进行估计。

2.4 二次曲线

若时间序列数据呈现二次非线性趋势，可以使用二次曲线[1-4]拟合和预测。其一般方程为：

$$\hat{y}_t=\alpha+\beta t+\gamma t^2 \tag{5}$$

式中参数和公式（4）相同，参数的估计也可以用最小二乘法估计。

2.5 指数曲线

指数曲线[1-4]用于描述以几何级数变化的数据。指数曲线的一般形式为：

$$\hat{y}_t=\alpha+\beta^t \tag{6}$$

要确定参数 α 和 β，需要做线性化处理，两端取对数，公式（6）变换为：

$$\ln(\hat{y}_t)=\ln\alpha+t\ln\beta \tag{7}$$

利用最小二乘法估计 $\ln\alpha$ 和 $\ln\beta$，再取反对数即可得参数 α 和 β。

2.6 修正指数曲线

修正指数曲线[2-4]是在一般指数曲线的技术上增加一个常数 K，即：

$$\hat{y}_t=K+\alpha\beta^t \tag{8}$$

式中：K、α 和 β 为待估计参数，$K>0$，$\alpha\neq0$，$0<\beta\neq1$。可以采用三和法求解参数。

2.7 ARIMA 预测

由于时间序列数据一般都是非平稳的，在应用 ARMA 方法[3-6]时，首先需要对数据做平稳化处理，一般常用的方法是差分的方法。其步骤为：

（1）模型识别确定模型的使用形式：检查数据是否平稳，若是非平稳的，则做差分平稳化。非平稳序列模型称为自回归累积移动平均模型，表示为 $ARIMA(p,d,q)$，其中，p 代表自回归部分的阶数，d 表示差分阶数，q 代表移动平均部分的阶数。如 $ARIMA(1,1,1)$ 可表示为：

$$y_t-y_{t-1}=\phi_1(y_t-y_{t-1})+\varepsilon_t-\theta_1\varepsilon_{t-1} \tag{9}$$

然后根据序列的自相关系数和偏自相关系数确定 AR 和 MA 的阶数，并利用简约的原则最终确定模型的形式。

（2）模型估计：ARIMA 模型的参数是通过最小拟合误差平方和来估算的，主要使用方法是非线性最小二乘法。

（3）模型检验：主要是对残差进行检验，看是否是白噪声，是则说明模型可行。

（4）用模型进行预测。

2.8 灰色预测 GM(1,1)模型

GM(1,1)模型[7-8]是最常用的一种灰色动态预测模型，该模型由一个单变量的一阶微分方程构成。

其步骤大致可以分为：

（1）级比检验和建模可行性判断：对原始序列 $X^{(0)}=\{x^{(0)}(t)|t=1,2,\cdots,n\}$，计算级比：$\sigma(k)=\dfrac{x(k-1)}{x(k)}$，$k=2,3,\cdots,n$ 并判断级比值是否落入可容覆盖中，若是，则可进行建模。

（2）数据变换处理：对级比检验不合格的序列，需要做数据变换处理，主要有：平移变换、对数变换、方根变换。

（3）GM(1,1)建模：对符合建模的数据做一次累加变换，生成新的序列：$X^{(1)}=\{x^{(1)}(t)|t=1,2,\cdots,n\}$，其中：$x^{(1)}(t)=\sum_{k=1}^{t}x^{(0)}(k)$。计算 $x^{(1)}(t)$ 的均值序列 $z^{(1)}(k)=0.5(x^{(1)}(k)+x^{(1)}(k-1))$，计算中间参数：

$$C=\sum_{k=2}^{n}z^{(1)}(k),D=\sum_{k=2}^{n}x^{0}(k),E=\sum_{k=2}^{n}z^{(1)}(k)x^{0}(k),F=\sum_{k=2}^{n}z^{(1)}(k)^2$$

利用中间参数计算模型参数：

$$a=\frac{CD-(n-1)E}{(n-1)F-C^2} \qquad b=\frac{DF-CE}{(n-1)F-C^2}$$

最终可得 GM(1,1)模型：

$$\hat{x}^{(1)}(k+1)=(x^{(0)}(1)-\frac{b}{a})e^{-ak}+\frac{b}{a} \tag{10}$$

经累减还原的原序列数据为：

$$\hat{x}^{(0)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k)$$

（4）残差检验和级比检验。

（5）预测。

2.9 改进灰色预测

对于上述灰色预测[9]可以知道，GM（1，1）模型预测，其预测精度和被预测对象的递变规律密切相关，若原始数据序列以指数规律变化，则预测精度较高。基于这样的原理，这里通过用 a^{-x^m} 对原始序列做变换，是变换后的序列 $a^{-x(k)^m}$ 更接近指数规律，从而提高预测精度。其余部分和传统的灰色预测建模都是相同的，这里不再详述。

2.10 BP 神经网络预测

BP 网络[10-11]模型处理信息的基本原理是：输入信号 X_i 通过中间节点（隐层点）作用于输出节点，经过非线形变换，产生输出信号 Y_k，网络训练的每个样本包括输入向量 $\boldsymbol{X}_i$ 和期望输出量 t，网络输出值 Y 与期望输出值 t 之间的偏差，通过调整输入节点与隐层节点的联接强度取值 $\boldsymbol{W}_{ij}$ 和隐层节点与输出节点之间的联接强度 T_{jK} 以及阈值，使误差沿梯度方向下降，经过反复学习训练，确定与最小误差相对应的网络参数（权值和阈值），训练即告停止。

BP 网络模型包括输入输出模型、作用函数模型、误差计算模型和自学习模型。

（1）节点输出模型

隐节点输出模型：$O_j=f(\sum_{i=1}\boldsymbol{W}_{ij}\times\boldsymbol{X}_i-\theta_j)$

输出节点输出模型：$Y_k=f(\sum_{i=1}T_{jK}\times O_j-\theta_k)$

其中：f 为非线形作用函数；θ 为神经单元阈值。

（2）作用函数模型

作用函数是反映下层输入对上层节点刺激脉冲强度的函数又称刺激函数，一般取为（0，1）内连续取值 Sigmoid 函数：$f(x)=1/(1+e^{-x})$。

（3）误差计算模型

误差计算模型是反映神经网络期望输出与计算输出之间误差大小的函数。

$$E_p=\frac{1}{2}\sum_{i=1}(t_{pj}-O_{pi})^2 \tag{11}$$

其中 t_{pj} 是 i 节点的期望输出值：O_{pi} 是 i 节点计算输出值。

（4）自学习模型

神经网络的学习过程，即连接下层节点和上层节点之间的权重矩阵 $\boldsymbol{W}_{ij}$ 的设定和误差修正过程。BP 网络有师学习方式（需要设定期望值）和无师学习方式（只需输入模式）之分。自学习模型：

$$\Delta W_{ij}(n+1)=\eta\times\phi_i\times O_j+\alpha\times\Delta W_{ij}(n) \tag{12}$$

其中：η 是学习因子；ϕ_i 是输出节点 i 的计算误差；O_j 是输出节点 j 的计算输出；α 是动量因子。

2.11 支持向量机预测

支持向量机（Support Vector Machine，SVM）[12-14]由 Vapnik 及其合作者以统计学习理论为基础率先提出，在 20 世纪 90 年代中后期得到了全面深入的发展，现已经成为机器学习和数据挖掘领域的主要工具。下面介绍解决回归问题的支持向量回归机。

支持向量回归机的基本思想是通过非线性映射将数据映射到高维特征空间 Ω 中，并在该特征空间进行线性回归：

$$f(x)=(w\cdot\Phi(x))+b \tag{13}$$

式中：$(\cdot)$ 表示在高维特征空间 Ω 中的内积，$\Phi:x\to\Omega$，b 是阈值。即在高维特征空间 Ω 的线性回归对应低维输入空间的非线性回归。

考虑 l 个独立同分布的学习样本 $T=\{(x_1,y_1),\cdots,(x_l,y_l)\}\in(X\times Y)^l$，其中 $x_l\in X\in R^n$，$y_l\in Y\in R$，$i=1,2,\cdots,l$，在高维特征空间中构造回归超平面。

回归超平面的对应优化问题是：

$$\begin{aligned}&\min_{w,b,\xi^{(*)}}\frac{1}{2}\|w\|^2+C\sum_1^l(\xi_i+\xi_i^*)\\&\text{s.t. }(w\cdot\Phi(x))+b-y_i\leqslant\varepsilon+\xi_i,i=1,2,\cdots,l\\&\quad y_i-(w\cdot\Phi(x))-b\leqslant\varepsilon+\xi_i^*,i=1,2,\cdots,l\\&\quad \xi_i,\xi_i^*\geqslant0,i=1,2,\cdots,l\end{aligned} \tag{14}$$

式中：$\xi^{(*)}=(\xi_1,\xi_1^*,\cdots,\xi_l,\xi_l^*)$，而 $\varepsilon>0$ 为不敏感损失函数参数，C 为正则化参数。引入 Lagrange 乘子构造 Lagrange 泛函，得到原问题的对偶问题：

$$\begin{aligned}&\max_{a^{(*)}\in R^n}\sum_1^l y_i(a_i^*-a_i)-\frac{1}{2}\sum_{i,j=1}^l(a_i^*-a_i)(a_j^*-a_j)\cdot K(x_i,x_j)-\\&\quad\varepsilon\sum_{i=1}^l(a_i^*+a_i)\\&\text{s.t. }\sum_{i=1}^l(a_i^*-a_i)=0,0\leqslant a_i,a_i^*\leqslant C,i=1,2,\cdots,l\end{aligned} \tag{15}$$

所要求的最优回归超平面可表示为：

$$\begin{aligned}f(x)&=(w\cdot\Phi(x))+b=\sum_{i=1}^{n_{sv}}(a_i^*-a_i)(\Phi(x_i)\cdot\Phi(x))+b=\\&\sum_{i=1}^{n_{sv}}(a_i^*-a_i)K(x_i,x)+b\end{aligned} \tag{16}$$

式中：n_{sv} 为支持向量数，$K(x_i,x)=(\Phi(x_i)\cdot\Phi(x))$ 是满足 Mercer 条件的核函数 $K(x_i,x)$，利用核函数代替高维特征空间中的内积，避免了在高维空间中的复杂的点运算。常用的核函数有多项式核、RBF 核、Sigmoid 核和 Fourier 核及 B 样条核等。

2.12 灰色支持向量机预测

对于时间序列数据进行预测，因为不同的时间序列前后关联的程度不同，选取的历史数据长度对其预测结果的影响也不同，对训练样本选取不同的长度用 SVR 算法进行预测，最终根据均方误差 MSE 最小原则，选取预测步长。

灰色支持向量回归机[15]算法首先利用灰色预测方法把原始数据序列进行一次累加生成，然后利用支持向量回归机拟合非线性数据能力强的特点，利用不同的参数和核函数的组合对测试样本进行预测，通过预测结果的最小均方误差 MSE 选择最优的参数和核函数，对新数据建立预测模型，最后将预测结果进行“累减还原”，得到预测值。具体算法如下：

（1）对原始数据序列

$$X^{(0)}=\{X^{(0)}(1),\cdots,X^{(0)}(n)\}(X^{(0)}(i)>0,i=1,2,\cdots,n)$$

进行一次累加生成，得到新的序列：

$$X^{(1)}=\{X^{(1)}(1),\cdots,X^{(1)}(n)\}(X^{(1)}(k)=\sum_{i=1}^k X^{(0)}(i),k=1,2,\cdots,n)$$

（2）对新序列进行不同的核函数和参数组合的计算，利用最小均方误差 MSE 确定核函数和参数。

表 1 预测效果评价表

预测方法	模型参数	年份	实际值/亿吨公里	预测值/亿吨公里	MAD	MSE	MAPE
移动平均法	三期二重移动平均	2005	80 258	69 776.00	2 181.000	13 319 679	0.050 6743
		2006	88 952	85 827.00			
指数平滑法	α=0.9	2005	80 258	79 017.42	3 110.493	22 691 486	0.081 561
		2006	88 952	87 958.97			
线性趋势	α=198.072 9	2005	80 258	62 185.41	5 518.724	56 005 860	0.181 184
	β=2 213.833	2006	88 952	64 399.24			
二次曲线	α=14 607.27 β=-575.042	2005	80 258	71 388.70	3 657.399	22 250 083	0.110 074
	γ=92.962 53	2006	88 952	76 112.52			
指数曲线	α=9 990.885	2005	80 258	69 368.24	2 724.53	19 044 058	0.062 439
	β=1.071 7	2006	88 952	74 338.93			
修正指数曲线	α=33 827.99 β=1.033 51	2005	80 258	59 307.68	2 997.538	47 286 763	0.065 023
	K=-25 884.5	2006	88 952	62 161.88			
ARIMA 法[1]	ARIMA(1,0,0)	2005	80 258	81 312.23	3 121.407	21 854 633	0.085 229
		2006	88 952	89 389.20			
灰色 GM(1,1)	a=-0.073 86	2005	80 258	69 559.00	3 318.781	19 763 917	0.092 772
	b=9 032.038	2006	88 952	74 891.00			
改进灰色 GM(1,1)[2]	a=0.016 489	2005	80 258	59 360.00	5 375.176	67 398 367	0.180 778
	b=1.017 716	2006	88 952	61 240.00			
BP 神经网络	三层 1-5-1BP 网络	2005	80 258	76 925.69	2 271.962	7 707 409	0.074 623
		2006	88 952	87 413.13			
支持向量回归	C=10,μ=0.2	2005	80 258	71 363.19	1 817.028	11 748 906	0.045 420
	2 阶多项式核函数	2006	88 952	87 337.54			
灰色支持向量回归	C=1 000,μ=0.2	2005	80 258	85 918.74	1 588.088	8 701 479	0.033 210
	3 阶多项式核函数	2006	88 952	80 026.56			

[1] 该方法先对原始数据取对数，然后再进行计算，最后把中间结果取反对数得到的结果。

[2] 该方法把原始数据转换为千亿吨做的计算。

（3）构造预测模型：$f(x)=\sum_{i=1}^{n_{sv}}(a_i^*-a_i)K(x_i,x)+b$。

（4）计算累加序列 $X^{(1)}$ 的预测值 $\hat{X}^{(1)}$。

（5）对 $\hat{X}^{(1)}$ 进行"累减还原"，得到原始数据序列 $X^{(0)}$ 的预测模型：

$$\hat{X}^{(0)}(k+1)=\hat{X}^{(1)}(k+1),\cdots,\hat{X}^{(1)}(k),k=n+1,n+2,\cdots \tag{17}$$

3 预测误差的测定

评定预测好坏的方式较多，主要有：绝对平均偏差（MAD）、均方误差（MSE）、绝对平均百分比（MAPE）以及平均百分比误差（MPE）等。若以 $\hat{y}_t$ 表示预测值，y_t 表示实际值，则各误差测定值计算公式分别为：

$$MAD=\frac{1}{n}\sum_{t=1}^{n}\left|y_t-\hat{y}_t\right| \qquad MSE=\frac{1}{n}\sum_{t=1}^{n}(y_t-\hat{y}_t)^2$$

$$MAPE=\frac{1}{n}\sum_{t=1}^{n}\frac{\left|y_t-\hat{y}_t\right|}{y_t} \qquad MPE=\frac{1}{n}\sum_{t=1}^{n}\frac{(y_t-\hat{y}_t)}{y_t}$$

这里使用前三种预测误差测定方法对预测结果进行比较分析。

4 各预测方法的比较分析

4.1 数据选择

物流在国民经济发展中起着举足轻重的作用，它涉及到整个国民经济的运行效率和运行质量，作为介于社会再生产过程中生产与消费之间的重要产业，它的发展与整个国民经济的发展有着十分密切而又复杂的联系。企业的生产组织到产品的加工、储运、分销等，从生产地到消费地、生产者到消费者过程中所形成的包括信息传递在内的一系列计划、执行、管理和控制的过程都离不开物流活动的参与。科学合理的预测物流量，具有一定的现实意义，这里选择我国 1978 年~2006 年货物周转量作为物流发展水平进行预测。所用数据来自《中国统计年鉴》（1978-2007）。

4.2 各方法的计算及结果

对各方法进行计算的过程中，使用 1978 年~2004 年的数据进行建模或训练，2005 年~2006 年的数据作为预测值，以便比较。预测效果评价见表 1。

4.3 预测结果排序及评述

（1）从预测误差的测定结果可以看出，各方法以预测精度大小排序，见表 2。从表中可以看出，就方法本身而言，支持向量类的方法、神经网络方法、灰色方法以及 ARMA 方法均比较复杂，但是预测的效果则以支持向量类方法较高，而 BP 神经网络方法和灰色预测方法则不是很理想；相反，移动平均和指数平滑两种方法较为简单，但是预测精度却比较高，仅排在支持向量类方法之后，说明简单的方法预测精度不一定低，而复

表 2 预测方法效果排序

序号	方法	MAPE	序号	方法	MAPE
1	灰色支持向量回归	0.033 210	7	指数平滑法	0.081 561
2	支持向量回归	0.045 420	8	ARIMA 法	0.085 229
3	移动平均法	0.050 674	9	灰色 GM(1,1)	0.092 772
4	指数曲线	0.062 439	10	二次曲线	0.110 074
5	修正指数曲线	0.065 023	11	改进灰色 GM(1,1)	0.180 778
6	BP 神经网络	0.074 623	12	线性趋势	0.181 184

杂的方法预测效果不一定好。

（2）就 2005 年和 2006 年两年的具体预测值而言，各方法的表现又有所不同。具体情况见表 3。每种方法的排序情况较表 2 又有所不同。其中唯有支持向量回归一直保持较高的精度，是因为其泛化性能较高的缘故。由于神经网络方法的泛化性能相比支持向量机方法要弱，所以其预测值的相对误差比较大。

表 3　2005 年~2006 年预测值相对误差表

排序	预测方法	2005 年预测相对误差/(%)	2006 年预测相对误差/(%)
1	指数平滑法	0.013 135 513	0.004 915
2	支持向量回归	0.015 457 400	0.011 164
3	二次曲线	0.041 519 973	0.017 300
4	线性趋势	0.070 531 785	0.100 340
5	指数曲线	0.110 509 856	0.144 342
6	改进灰色 GM(1,1)	0.110 827 706	0.018 150
7	灰色支持向量回归	0.130 603 803	0.035 131
8	ARIMA 法	0.133 307 583	0.158 074
9	修正指数曲线	0.135 684 418	0.164 280
10	移动平均法	0.225 181 166	0.276 023
11	灰色 GM(1,1)	0.260 385 258	0.311 539
12	BP 神经网络	0.261 037 155	0.301 175

5　结论

由于选取的数据是 1978 年~2006 年中国货运周转量，样本容量只有 29 个，相对来讲还是比较少。对于一些方法，如 ARMA 法，可能会因为样本容量的问题影响计算和预测精度。另外，由于该时间序列数据具有明显的非线性增长趋势，所以，文中提到的一些方法并不太适合于这种样本数据，故会导致预测精度下降，这并不能说明这种方法预测能力不强，而是应该视具体情况，选择相应的合适的预测方法。而对于灰色预测类方法，则由于其自身的一些缺陷，导致预测精度有时不是很高。

文中通过该时间序列数据，再次验证了复杂的方法其预测精度不一定比简单的方法效果好这一命题。较为系统的对现在比较流行的预测方法给出了实证分析，具有一定的实际应用价值。

参考文献：

[1] 孙焰.现代物流管理技术[M].上海：同济大学出版社，2004.
[2] 刘国山.数据建模与决策[M].北京：中国人民大学出版社，2004.
[3] 冯文权.经济预测与决策技术[M].武汉：武汉大学出版社，2005.
[4] Hanke J E，Wichern D W.Business forecasting [M].[S.l.]：Pearson Education Inc，2005.
[5] 高铁梅.计量经济分析方法与建模——EVIEWS 应用及实例[M].北京：清华大学出版社，2006.
[6] 王振龙.时间序列分析[M].北京：中国统计出版社，2002.
[7] 刘思峰.灰色系统理论及其应用[M].北京：科学出版社，1999.
[8] 邓聚龙.灰预测与灰决策[M].北京：科学出版社，2004.
[9] 陈洁，许长新.改进的灰色模型在港口吞吐量预测中的应用[J].水运工程，2005(1)：20-23.
[10] Hagan M T，Demuth H B，Beale M H.Neural network design[M].[S.l.]：PWS Publishing Company，1996.
[11] 周开利，康耀红.神经网络模型及其 MATLAB 仿真程序设计[M].北京：清华大学出版社，2005.
[12] 瓦普尼克.统计学习理论[M].许建华，张学工，译.北京：电子工业出版社，2004.
[13] 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，2006.
[14] Cristianini N，Shawe-Taylor J.支持向量机导论[M].李国正，王猛，曾华军，译.北京：电子工业出版社，2004.
[15] 唐万梅.基于灰色支持向量机的新型预测模型[J].系统工程学报，2006(8)：410-413.
[16] Clemen R T.Combining forecasting：A review and annotated bibliography[J].International Journal of Forecasting，1989，5(4)：559-583.
[17] Armstrong J S.Combining forecasts：The end of the beginning or the beginning of the End?[J].International Journal of Forecasting，1989，5(4)：585-588.

（上接 195 页）

通过对比可以看出，提出 ASKO 算法无论在收敛速度还是收敛精度方面都优于其他三种算法，意味着在得到足够精确的优化解的基础上所需要的对目标函数评估次数比其他三种算法小的多，如果将该算法应用于目标特性评估耗时的工程优化设计中时，可以有效地降低计算量，提高优化效率，具有很强的工程应用价值。此外，ASKO 算法是可预测的，属于确定性算法，给定一个函数运行该算法找到近似全局最优解的概率为 1，而且每次运行所需要的迭代次数一定。而 SA 算法、ECTS 算法都属于随机算法，即在运行过程中需要用到随机数导致它们的可靠性不好，运行过程不能预测，因此以上给出的两种算法的目标函数评估次数均为平均值。

4　结论

提出了一种基于 Kriging 代理模型的自适应序贯优化算法。采用 Kriging 模型的预测值与预测标准差的提高概率定义 EI 函数，能有效地避免基于代理模型优化算法中常见的局部收敛问题。数值算例表明，该算法是正确的，其收敛速度，精度都较高，这就能有效地提高优化的效率，而且它是一种确定性的优化算法，运行稳定。可见将提出的算法应用于工程设计领域中，能有效地解决工程优化中耗时较长这个问题。

参考文献：

[1] Oyama A，Liou M S，Obayashi S.Transonic axial flow blade optimization：Evolutionary algorithms/three dimensional Navier-Stokes solver[J].Journal of Propulsion and Power，2004，20(4)：612-619.
[2] Simpson T W，Peplinsk J D，Koch P N.Metamodels for computer-based engineering design：Survey and recommendations[J].Engineering with Computers，2001，17：129-150.
[3] Xu Y，Li G，Wu Z.A novel hybrid genetic algorithm using local optimizer based on heuristic pattern move[J].Applied Artificial Intelligence，2001，15：601-631.
[4] Jones D L.A taxonomy of global optimization methods based on response surfaces[J].Journal of Global Optimization，2001，21：345-383.
[5] Schonlau M.Computer experiments and global optimization[D].Waterloo：University of Waterloo，1997.
[6] Loactelli M.Bayesian algorithms for one-dimensional global optimization[J].Journal of Global Optimization，1997，10：57-76.
[7] Wang G，Dong Z，Aitchison P.Adaptive response surface method：A global optimization scheme for computation-intensive design problems[J].Journal of Engineering Optimization，2001，33(6)：707-734.
[8] Chelouah R，Siarry P.Tuba search applied to global optimization[J].European Journal of Operational Research，2000，123：256-270.