
摘 要

影响原油期货价格的因素错综复杂，原油期货价格有很大的波动性和不确定性，因此传统的时间序列预测模型较难描述原油期货价格时间序列的这一特征。人工神经网络依据数据本身的内在联系建模，具有良好的自组织性、自适应性，有很强的学习能力、抗干扰能力。它能自动从历史数据中提取有关知识，因而神经网络模型的出现给这类预测问题提供了新的理论和方法。

本文深入分析了Brent原油期货价格时间序列的趋势性和波动性的特点，提出了基于概率统计的组合算法，将线性模型与神经网络模型进行组合，建立了组合预测模型。通过用传统的统计学预测模型、时间序列预测模型，以及若干个神经网络预测模型与之进行对比。结果表明：本文提出的组合预测模型在预测精度和预测稳定性方面有所提高。

关键字：时间序列，神经网络，组合预测

Abstract

There are complicated factors influencing the futures prices of crude oil futures which are fairly fluctuant and dubious. For the characteristic that traditional time series forecasting model is difficult to describe time series of crude oil futures prices, artificial neural network model accord to the inherent relations of the history data, which provides good self-organizing, self-flexibility, powerful study ability and anti-jamming ability, and it can automatically extract corresponding knowledge from history data. Therefore, artificial neural network model offers new theories and approaches for this type of forecasting issues.

This article has deeply analyzed the trend and fluctuation characteristics of time series of Brent crude oil futures price, adopted the combination approach of linear models and artificial neural network, and proposed assembled algorithms based on probability & statistics. Meanwhile, traditional statistics forecasting model, time series forecasting model and some artificial neural network forecasting model for Brent crude oil futures price closing quotation price have been established. In conclusion: the combined forecasting model of linear model and artificial neural network model could improve the precision and stability of forecasting.

Keywords: Time Series, Artificial Neural Network, Combined Forecasting

目 录

1 绪论	1
1.1 研究意义.....	1
1.2 国内外研究现状.....	2
1.2.1 传统的统计学模型.....	2
1.2.2 常规的时间序列预测模型.....	3
1.2.3 神经网络预测模型.....	4
1.2.4 组合预测模型.....	6
1.3 本文的主要研究内容.....	7
2 运用常规预测模型进行预测	9
2.1 基础工作.....	9
2.1.1 历史数据的收集整理.....	9
2.1.2 Brent 原油期货价格时间序列分析.....	11
2.1.3 数据指标相关分析.....	12
2.2 多元线性回归模型.....	13
2.2.1 多元线性回归算法.....	14
2.2.2 应用多元线性回归模型预测 Brent 原油期货价格.....	14
2.3 灰色模型.....	15
2.3.1 灰色预测算法.....	16
2.3.2 应用灰色模型预测 Brent 原油期货价格.....	17
2.4 ARIMA 模型.....	17
2.4.1 ARIMA 模型算法.....	18
2.4.2 应用 ARIMA 模型预测 Brent 原油期货价格.....	19
3 人工神经网络预测理论	21
3.1 人工神经网络模型概述.....	21
3.1.1 人工神经网络的发展.....	21
3.1.2 人工神经网络的结构.....	24
3.2 BP 神经网络预测模型.....	25
3.2.1 BP 神经网络概述.....	25
3.2.2 BP 神经网络算法.....	26
3.2.3 BP 神经网络结构.....	27
3.2.4 BP 神经网络预测模型.....	28
3.2.5 BP 神经网络的不足和改进.....	31
3.3 径向基神经网络预测模型.....	32
3.3.1 径向基神经网络的基本原理.....	32
3.3.2 径向基神经网络结构.....	33
3.3.3 径向基神经网络预测算法.....	33
4 组合预测理论	35
4.1 组合预测简介.....	35

4.2 简单平均组合预测模型	35
4.3 最优组合预测方法	36
4.4 变权重预测方法	37
5 基于概率统计的组合预测模型	38
5.1 模型算法.....	38
5.1.1 组合算法	39
5.1.2 组合权系数的确定.....	40
5.1.3 单项模型的选择标准.....	42
5.2 单项预测模型.....	43
5.2.1 模型 L_bp_all.....	44
5.2.2 模型 L_bp4.....	44
5.2.3 模型 L_bp5s	45
5.2.4 模型 L_bp10s_3z.....	46
5.2.5 模型 L_rbf_all	47
5.2.6 模型 L_rbf4	48
5.2.7 模型 L_rbf5s.....	49
5.2.8 模型 L_rbf10b_3z.....	50
5.2.9 单项预测模型预测性能比较	51
5.3 组合预测模型.....	52
5.3.1 组合模型（一）	52
5.3.2 组合模型（二）	53
5.3.3 与简单平均组合模型比较	54
6 结论与展望	56
6.1 总结	56
6.2 进一步研究.....	57
致谢.....	58
攻读硕士期间发表的学术论文	59
参考文献.....	60

图 目 录

图 1.1 研究框架图	8
图 2.1 BRENT 原油期货价格时间序列图	11
图 2.2 一阶差分序列图	11
图 2.3 各数据指标相关系数图	12
图 2.4 多元线性回归模型预测效果图	15
图 2.5 灰色模型预测效果图	17
图 2.6 数据样本时间序列图	19
图 2.7 一次差分后时间序列图	20
图 2.8 ARIMA 模型预测效果图	20
图 3.1 人工神经元结构图	24
图 3.2 三层 BP 神经网络结构图	28
图 5.1 序列趋势与波动分离图	39
图 5.2 BRENT 原油期货价格随机波动序列图	39
图 5.3 波动序列全体样本直方图	40
图 5.4 测试集附近波动序列直方图	41
图 5.5 L_BP_ALL 模型预测效果图	44
图 5.6 L_BP4 模型预测效果图	45
图 5.7 L_BP5S 模型预测效果图	46
图 5.8 L_BP10S_3Z 模型预测效果图	47
图 5.9 L_RBF_ALL 模型预测效果图	48
图 5.10 L_RBF4 模型预测效果图	49
图 5.11 L_RBF5S 模型预测效果图	50
图 5.12 L_RBF10B_3Z 模型预测效果图	51
图 5.13 组合模型（一）预测效果图	53
图 5.14 组合模型（二）预测效果图	54

1 绪论

1.1 研究意义

期货市场在稳定与促进市场经济发展方面具有许多功能。首先，它具有回避价格风险的功能，此外，还具有发现价格的功能、有利于市场供求和价格的稳定，期货交易还是一种重要的投资工具，有助于合理利用社会闲置资金。因而，期货交易在市场经济里是必不可少的一分子。

原油期货是最重要的石油期货品种，目前世界上重要的原油期货合约有4个：纽约商业交易所（NYMEX）的轻质低硫原油（WTI）即“西德克萨斯中质油”期货合约、高硫原油期货合约，伦敦国际石油交易所（IPE）的布伦特(Brent)原油期货合约，新加坡国际金融交易所（SIMEX）的迪拜酸性原油期货合约^[13]。

伦敦国际石油交易所（IPE）是欧洲最重要的能源期货和期权的交易场所。它成立于1980年，是非营利性机构。1988年6月23日，IPE推出国际四种基准原油之一的布伦特原油期货合约。IPE布伦特原油期货合约特别设计用以满足石油工业对于国际原油期货合约的需求，是一个高度灵活的规避风险及进行交易的工具。IPE的布伦特原油期货合约上市后取得了巨大成功，迅速成为国际原油期货交易中心之一，而布伦特原油期货价格也成为了国际油价的基准之一。现在，布伦特原油期货合约是布伦特原油定价体系的一部分，包括现货及远期合约市场。该价格体系涵盖了世界原油交易量的65%^[12]。

这些原油期货市场对稳定该地区市场经济和回避企业风险起到了非常重要的作用。并且其期货价格成为了现货交易价格的重要参考依据。

我国石油需求量增长迅速，自1993年开始，成为石油净进口国，目前年进口原油7000多万吨，花费近200亿美元，前年由于国际油价上涨多支付了数十亿美元。目前我国石油供求和价格对国外资源的依赖程度越来越高，承受的风险越来越大，国内企业对恢复原油期货交易的呼声很高。

对原油期货价格及其趋势进行预测，将对石油及相关企业的决策起到重要参考作用，也对人们进行期货投资具有重要指导和参考意义。

通常对时间序列数据的预测可以分为长期预测、中期预测和短期预测。长期预测，是指对5年或5年以上的价格变化前景的预测；中期预测，是指对一个季度以上两年以下的价格变化的预测；短期预测，是指以日、周、旬、月为单位，对一个季度以下的价格变化的预测。

由于影响原油期货价格的因素错综复杂，原油期货价格有很大的波动性和不确定性，因而对原油期货价格的中长期预测非常困难，主要预测方法是由石油行业专家凭自己的实践经验和理论、业务水平，对原油期货行情及其发展前景的性质、方向和程度作出判断进行定性预测，也称为判断预测或调研预测。

本文研究内容是用定量的方法对伦敦国际石油交易所的布伦特原油期货合约的下一天收盘价进行预测。

1.2 国内外研究现状

时间序列预测法，就是将预测目标的历史数据按照时间的顺序排列成为时间序列，然后分析它随时间的变化趋势，外推预测目标的未来值。这样，就把影响预测目标变化的一切因素由“时间”综合起来描述了。时间序列预测方法是预测方法体系中的重要组成部分。时间序列分析的最基本的理论基础是40年代分别由Norbert Wiener和Andrei Kolmogorov独立给出的，他们对发展时间序列的参数模型拟合和推断过程作出了贡献，提供了与此有关的重要的文献，促进了时间序列分析方法在工程领域上的应用。

1.2.1 传统的统计学模型

在国民生产和日常生活中，传统的统计学预测模型应用非常广泛，这些方法包括简单平均法、移动平均法、指数平滑法和线性回归法。

移动平均是预测技术中的一种古老方法。它对一组给定的历史数据，计算其平均值，并将这一平均值作为下一时期的预测值。移动平均分为移动算术平均与移动几何平均，以及移动加权平均。这种方法非常简单但其预测精度比较低。

指数平滑方法是一种较为实用的方法。最具代表性的是Holt-Winters的线性趋势平滑模型。它分为一次指数平滑、二次指数平滑与多次指数平滑。指数平滑方法是一种低费用、简单和不精致的方法，但其预测精度却可以与许多精致的、更具统计基础的方法相比拟。

上述方法主要是通过对时间序列进行平滑以消除其随机波动，描述序列趋势，也可以用于对平稳时间序列进行短期预测。主要应用在工厂年度产量预测、企业季度销售额预测等等。这些方法的优点是算法简单、花费时间成本小。

1.2.2 常规的时间序列预测模型

20世纪80年代初，我国学者邓聚龙教授提出了灰色系统理论和方法，在许多领域得到了广泛关注，它能够有效处理不确定性显著和数据样本较少的系统，因此越来越多地得到了应用。

文献[1]中，运用灰色预测模型对呼和浩特市的人口规模进行了预测。文献[2]中，把灰色系统理论应用于城市年用电量预测。文献[3]将灰色系统理论首次引入证券价格趋势分析;利用灰色系统理论和关联度分析，建立能进一步准确反映证券价格动态过程发展态势的量化模型;通过模型，对证券价格变动趋势进行预测研究，建立灰色GM模型，并针对上海证券交易所某个时间段进行实证模型建立、分析及精度检验，取得了满意的结果^{[1][2][3]}。

文献[4]中，用灰色预测模型对沪深股市的多支股票价格的未来走势、波动方向及波动幅度进行了预测。预测平均相对误差均小于1%。灰色预测法是一种对既含有已知信息又含有不确定因素的系统进行预测的方法，它的特点是所需信息量少，不仅能够将无序离散的原始序列转化为有序序列，而且预测精度较高，能够保持原系统的特征，较好地反映系统的实际情况^[4]。

文献[5]肯定了灰色预测模型在经济学、地震学、城市规划等学科和领域实践中取得的巨大成果，同时剖析了灰色预测模型的预测原理，证明了灰色预测模型的非完备性。认为灰色预测模型并不是万能的预测模型，而是有一定的适用条件。各种系统特别是社会系统、经济系统的复杂性、多变性，其发展虽然有一定的规律，但是决不是简单的函数关系，因此应当根据具体数据的特点，选择合适的预测模型^[5]。

1970年, 美国学者George Box和英国统计学家Gwilym Jenkins 提出了Box-Jenkins模型, 简记为ARIMA(p,d,q)。基本思想是: 对于非平稳的时间序列, 用若干次差分使其成平稳序列, 再将此序列表示成关于序列直到过去某一点的自回归和关于白噪声移动平均的组合。该模型可看作是经过d次差分后的AR(P)模型和MA(q)模型结合而成的, 在国际上被称为最复杂、最高级的时间序列预测模型^[9]。

国内外学者对线性时间序列进行了大量的研究, ARIMA模型的建模方法、非平稳时序模型的识别、模型阶的确定、模态参数的估计等已基本趋于成熟。ARIMA模型使时间序列分析广泛地运用成为可能, 并很快成为预测领域的主要方法这一。

文献[7]中, 建立了ARIAM模型对上海证券交易所股票大盘走势进行预测, 预测显示, 短期预测平均相对误差在4%左右。较长期(第5期以后)的预测相对误差在8%左右。ARIMA模型由于只考虑时间序列本身的特性来预测, 没有考虑到股市本身受许多不可预测的复杂因素影响, 因此这些突然变化的因素在ARIMA模型在预测值中是无法表现出来的, 但ARIMA模型在短期内的预测结果是可以接受的^[7]。

文献[9]中, 应用ARIMA预测模型对我国2003年出口贸易额进行预测, 预测相对误差为6.99%。结果表明: 第一、ARIMA模型有较高的预测精度, 预测相对误差小于10%; 第二、该模型只适合于短期预测, 随着预测延长, 预测误差增大非常明显; 第三、该模型只考虑了时间序列本身的特性, 而没有考虑其他一些不确定因素对出口的影响, 因此这些因素在模型中是以随机项来反映, 在预测的期望值中是无法反映出来的, 从而使预测结果产生了一定的偏差。

ARIMA预测法适合于对时间序列的典型特征难以作出判断的时间序列的预测, 而且它也无需像回归分析中必须花很多时间寻找解释变量。它只要事先假设一个可能适用的模型, 然后按照一定的程序反复识别改进, 以求得一个较为满意的预测模型。ARIMA预测法是一种精度较高的短期预测方法, 但其计算过程非常复杂, 必须由统计软件来完成。

1.2.3 神经网络预测模型

上述的时间序列预测方法, 如移动平均法、指数平滑方法、ARIMA模型、线性回归法等等。应用这些传统的预测方法可以预测一段时间内预测对象变化的大致趋势,

但只适合于预测对象变化比较平稳的情况，需要事先知道各种参数，以及参数在什么情况下应做怎样的修正^[18]。而原油期货价格、汇率、股价等时间序列的变化是许多非常复杂因素共同影响的结果，其中广泛存在着非线性、时变性和不确定作用关系。传统的预测方法在发挥巨大作用的同时，也逐渐显露出它的缺陷，即很难把握时间序列变化的非线性现象，因而必然造成预测的误差性加大。

随着人工神经网络理论的不断发展和完善为时间序列预测提供了新的方法。人工神经网络依据数据本身的内在联系建模，具有良好的自组织性、自适应性，有很强的学习能力、抗干扰能力。它能自动从历史数据中提取有关知识，可以克服传统预测方法的许多局限以及面临的困难，同时也能避免许多人为因素的影响，因而神经网络模型是在非线性时间序列预测领域应用最广泛的模型。

国内外对原油期货价格预测研究还非常少，伊朗德黑兰的Saeed Moshiri在2004年用多层BP人工神经网络对纽约商业期货交易所的原油期货价格进行了预测，并用线性和非线性预测模型与之作比较，结果显示用神经网络模型得到的预测结果更加精确^[33]。

Greg Tkacz 和 Sarah Hu 分别用人工神经网络方法和线性的预测方法对加拿大的GDP增长率进行了预测，人工神经网络模型有效地吸收了金融、经济数据中隐藏的引导信息，预测误差比线性的预测模型预测误差小15%~19%^[38]。

刘新勇等人在预测股价时，不但考虑是看盘价、收盘价、最高价、最低价等指标，而且加入了前几天的威廉指标、相对强弱指标等技术指标。实证分析结果表明，人工神经网络具有非常强的非线性逼近能力^[8]。

金席卷等人把模糊理论与人工神经网络技术相结合，解决了BP神经网络训练收敛慢，容易陷入局部最小的问题。引入气温、日照、降雨量等因素，对电力短期负荷进行预测，预测相对误差达到7%左右，达到工程准度5%的要求^[10]。

此外，BP神经网络、径向基神经网络及其改进算法模型还广泛应用于车站车流量预测、网络业务量预测、铁路客运量预测、股票指数预测等等领域。

1.2.4 组合预测模型

然而，每个不同的单项预测模型的预测效果不一样，有些模型拟合能力较强，有些模型推广能力较强。并且，同一个预测模型，在不同的时间预测效果有好有坏。时间序列组合预测模型理论的提出，为解决这一问题提供了新的途径。

1969年J M Bates 和C W J Granger 首次提出了组合预测的理论和方法^[29]。组合预测定理指出：即使一个很差的单项预测方法，如果它含有系统的独立信息，当它与一种较好的预测方法进行组合后，同样可以提高系统的预测性能力。由于它能有效地提高预测精度，因此受到国内外预测工作者的重视。为此，人们一直在寻求如何有效利用各单项预测方法优点的组合预测方法，以提高预测精度和预测稳定性。

1989年，国际预测领域的权威学术刊物《Journal of Forecasting》还出版了组合预测的专辑，充分说明了组合预测在预测中的重要地位。进入90年代，组合预测的研究更处于一个热潮之中。

2001年，Ajoy Kumar Palit 和 D. Popovic 用组合遗传神经网络对时间序列问题进行预测，结果表明组合神经网络模型优于任何一个单独的神经网络模型^[30]。

近年来，我国在组合预测研究方面也取得了一系列的研究成果。电子科技大学唐小我教授及重庆大学曹长修教授等人的研究成果尤为突出，发表了一系列关于组合预测的学术论文，为促进我国组合预测的理论研究与应用做出了重要的贡献^[28]。

冯春山等人对ARIMA模型和神经网络模型进行了组合研究，充分发挥ARIMA模型对平稳线性时间序列预测优势，又结合了神经网络预测模型善于描述复杂的非线性动力系统的特点。通过对Brent石油价格预测进行实证研究，结果表明，ARIMA模型和神经网络模型的组合预测模型预测效果好于其中任意一个单项预测模型。预测精度分别比两个单项预测模型提高9.11%和10.88%^[24]。

中国科学院管理决策与信息系统重点实验室在对时间序列预测方面也有很多研究成果。汪寿阳教授及其学生用ARIMA模型、指数平滑模型、简单移动平均模型等时间序列预测模型与误差反传网络(BPNN)、径向基函数网络(RBFNN)、广义回归神经网络(GRNN)和概率神经网络(PNN)等神经网络模型组合，并成功地应用于外汇汇率预测、房产销售预测和国际原油价格预测^[31]。

1.3 本文的主要研究内容

本文预测对象是伦敦国际石油交易所（IPE）的布伦特（Brent）原油期货最近月份合约的收盘价，首先对Brent原油期货价格历史数据时间序列进行分析，分析其趋势性、波动性，根据其线性或非线性的特点，分别建立线性和非线性预测模型，比较和总结各个预测模型的优劣及其特性，然后选择代表不同信息片断的单项预测模型进行组合。

本文的主要研究内容包括下面几个部分：

（1）分析时间序列特点。观察Brent原油期货价格历史数据时间序列波动特性，采用统计方法分析时间序列可能包含趋势成分、季节成分、周期成分和随机成分。作为选择预测模型的依据。

（2）建立多个线性和非线性预测模型。根据时间序列特点，建立移动平均、指数平滑、灰色预测模型、ARIMA等预测模型，以及BP、RBF神经网络模型，并对比、分析每个模型的优劣及其特性。

（3）建立简单平均加权组合预测模型。选择多个线性预测模型和神经网络模型，用简单平均加权组合法进行组合预测。

（4）根据价格波动的统计特征，提出了一个基于概率统计的变权重组合算法，建立了组合预测模型。实验表明，该模型提高了预测稳定性和预测精度，具有一定的实用价值。本文论文框架如图1.1所示：

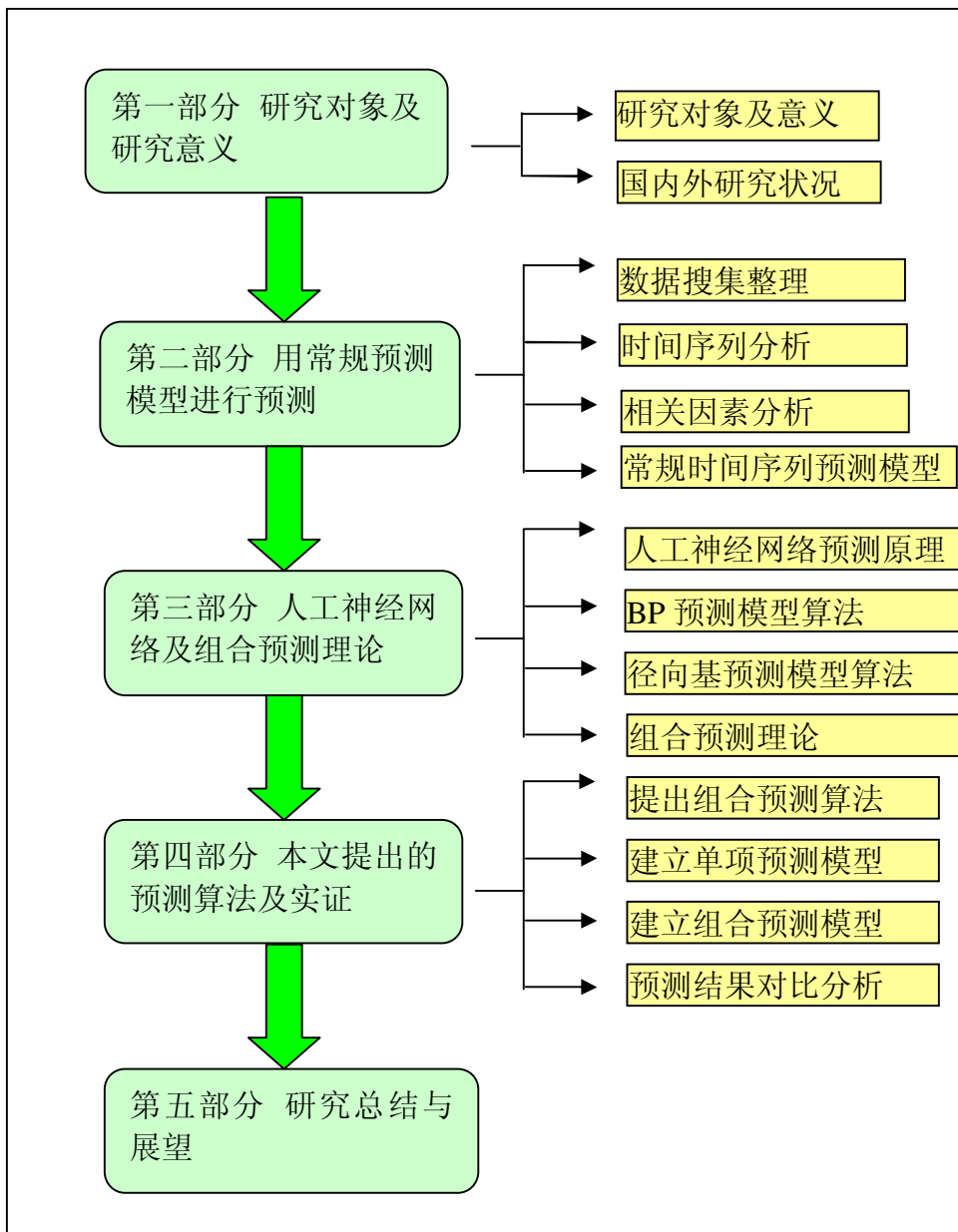


图 1.1 研究框架图

2 运用常规预测模型进行预测

2.1 基础工作

2.1.1 历史数据的收集整理

本文收集了2000年1月至2006年3月的Brent原油期货价格、NYMEX交易所的WTI原油期货价格的开盘价、最高价、最低价、收盘价、成交量及持仓量，道琼斯指数、纳斯达克指数、SP500指数、美元英镑汇率、美元欧元汇率等历史数据。汇率数据来源于美联储网站，原油期货价格数据和其它经济指标数据来源于雅虎财经。

Brent原油期货市场周一至周五开放，人工喊价交易时间为伦敦时间10:02-19:30。本文要预测的就是该市场最近月份合约的原油期货收盘价。

WTI原油期货市场也是周一至周五开放，人工喊价交易时间为美国东部时间时间10:00-14:30。本文收集的是该市场最近月份合约的原油期货价格。

本文收集的英镑对美元汇率和欧元对美元汇率为美国东部时间12:00的数据。

股票价格指数是运用指数方法编制而成的，反映股票市场总体价格或某类股票价格平均水平及其变动趋势的指标。目前世界上比较权威和有影响的股票价格指数主要有：

(1) 道·琼斯股价平均指数(Dow Jones Averages)，由道·琼斯公司编制，选取在纽约证券交易所上市的若干种具有代表性的股票作为样本，以1928年10月1日为基期，并令基期指数为100。道·琼斯股价指数每日都在《华尔街日报》上公布。

(2) 标准普尔指数，由标准普尔公司编制，目前选用在纽约证券交易所上市的500种股票作为样本，其中包括400种工业股票、40种公用事业股票、20种运输业股票和40种金融业股票，以1941年-1943年为基期，并令基期指数为10，以各种股票的发行量为权数，采用加权平均法进行计算。标准普尔指数的最大特点是取样面广、代表性强。本文采用的是标普500(SP500)指数。

(3)纳斯达克(NASDAQ)是美国全国证券交易商协会于1968年着手创建的自动报价系统名称的英文简称。纳斯达克的特点是收集和发布场外交易非上市股票的证券商报价。它现已成为全球最大的证券交易市场。目前的上市公司有5200多家。纳斯达克又是全世界第一个采用电子交易的股市,它在55个国家和地区设有26万多个计算机销售终端。纳斯达克综合指数是反映纳斯达克证券市场行情变化的股票价格平均指数。纳斯达克的上市公司涵盖所有新技术行业,包括软件和计算机、电信、生物技术、零售和批发贸易等。

本文收集的道琼斯指数、纳斯达克指数、SP500指数是美国东部时间收盘价格指数。

表 2-1 各价格指数发生时间

价格、指数	数据发生时间(当地)	数据发生时间(格林威治)
Brent原油期货	10:02-19:30	10:02-19:30
WTI原油期货	10:00-14:30	15:00-19:30
道琼斯指数	10:30-17:00(冬) 9:30-16:00(夏)	15:30-22:00(冬) 14:30-21:00(夏)
纳斯达克指数	10:30-17:00(冬) 9:30-16:00(夏)	15:30-22:00(冬) 14:30-21:00(夏)
SP500指数	10:30-17:00(冬) 9:30-16:00(夏)	15:30-22:00(冬) 14:30-21:00(夏)
英镑美元汇率	12:00	17:00
欧元美元汇率	12:00	17:00

从表2-1可以看出,本文收集的道琼斯指数、纳斯达克指数、SP500指数收盘价数据比同天Brent原油期货收盘价晚1.5至3小时;英镑美元汇率、欧元美元汇率数据比同天Brent原油期货收盘价早2.5小时,因此用这些指数来预测Brent原油期货下一天收盘价是合理的。

各数据指数都向Brent原油期货价格数据按日期对齐,如有缺失或者明显错误数据,以近期数据移动平均值表示。

2.1.2 Brent 原油期货价格时间序列分析

取2002年8月30日至2006年4月25日共692个交易日的价格作为数据样本，取前590个数据作为训练样本，剩下的作为测试样本。Brent原油期货价格时间序列，如下图2.1所示：



图 2.1 BRENT 原油期货价格时间序列图

Brent原油期货价格的随机波动非常显著，同时也可以看出明显的上升趋势，季节性、周期性不明显。通过右侧的数据分布直方图可以看出该数据序列不服从正态分布，时间序列不平稳。

图2.2是对原序列进行一次差分后时间序列图，经过一次差分，去除了原时间序列的趋势成分，时间序列在0值附近波动，并服从正态分布，进一步说明该时间序列有明显的趋势性，而季节性、周期性不明显。

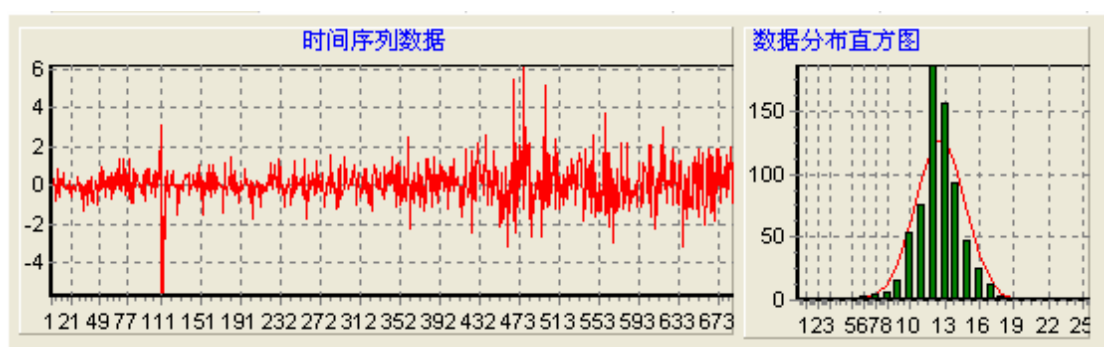


图 2.2 一阶差分序列图

2.1.3 数据指标相关分析

用Brent原油期货价格开盘价、最高价、最低价、成交量、持仓量、WTI原油期货价格的收盘价、纳斯达克指数、SP500指数、道琼斯指数、美元欧元汇率以及美元英镑汇率等指标的前1天至前14天的数据序列分别对当天Brent原油期货价格收盘价数据序列求相关系数，结果如图2.3所示。

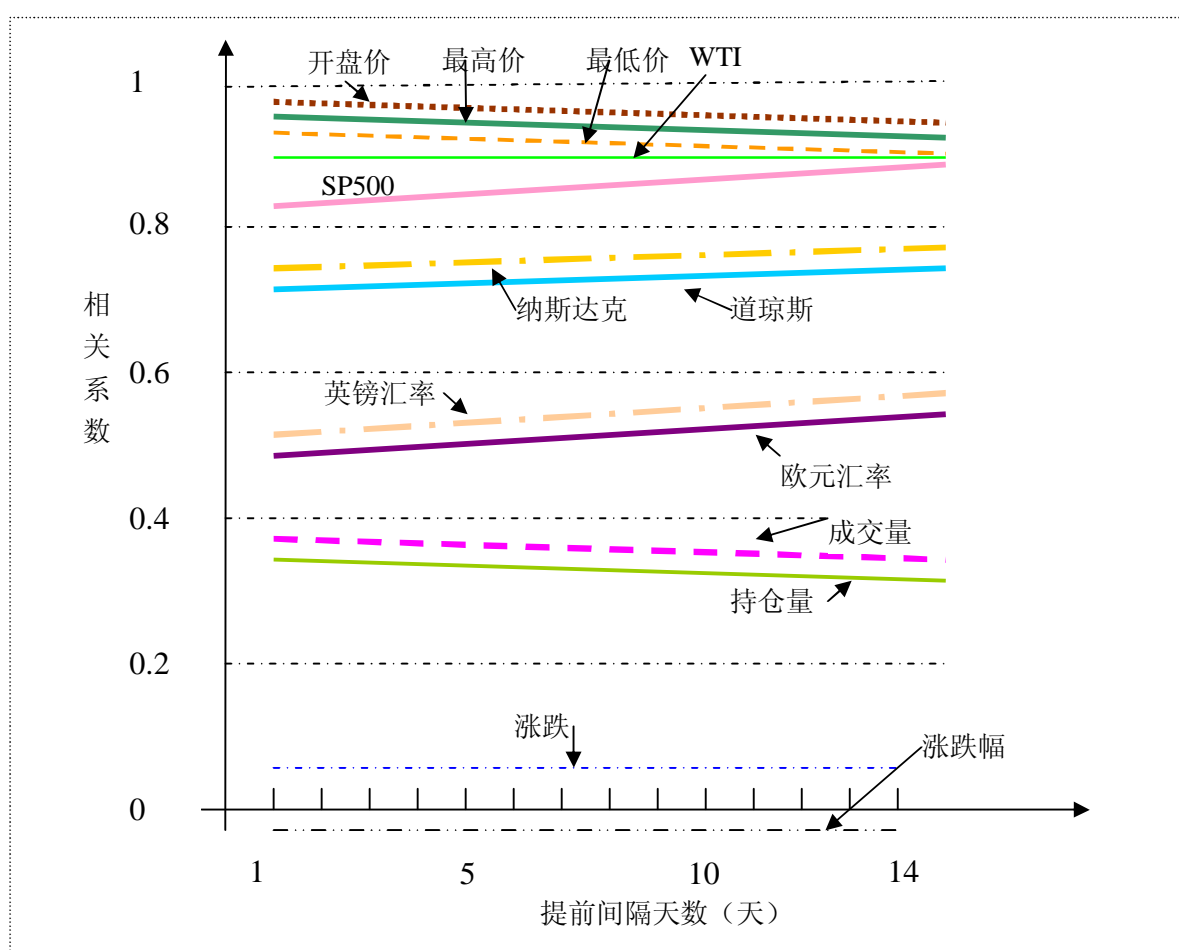


图 2.3 各数据指标相关系数图

结果显示，除Brent原油期货价格收盘价的涨跌幅与Brent原油期货价格收盘价成负相关关系外，其余的数据指标都与之成正相关关系。按照相关性大小，这些指标可以分为以下几个层次：

第一层次，WTI原油期货收盘价、Brent原油期货最高价、最低价、开盘价。这些指标与Brent原油期货收盘价相关系数最高，均大于0.9，且随着时间间隔日期（提前日期）越远，相关系数逐渐下降。

第二层次，SP500指数、纳斯达克指数和道琼斯指数。这些指标与Brent原油期货收盘价相关系数也比较高，相关系数介于0.7~0.9之间。相关系数都随时间间隔日期越远，相关系数反而增大，说明这些指标对Brent原油期货价格的影响时效较长。其中SP500指数的相关系数最大。

第三层次，美元英镑、美元欧元汇率。美元英镑、美元欧元汇率与Brent原油期货收盘价的相关系数介于0.5~0.6之间。相关系数都随时间间隔日期越远，相关系数也有增大趋势。

第四层次，成交量、持仓量。成交量和持仓量与Brent原油期货价格收盘价的相关系数也是随时间间隔日期越远相关系数逐渐下降，但是相关系数比较小，只有0.35左右。

此外，Brent原油期货价格涨跌值和涨跌幅与Brent原油期货价格收盘价相关性非常小，相关系数介于0.03~0.05之间。

通过相关分析，可以明显看出纳斯达克指数、道琼斯指数、SP500指数等价格指数与Brent原油期货价格有非常紧密的联系。而Brent原油期货价格的涨跌量及涨跌幅与Brent原油期货收盘价变化关系不大。因而在建立预测模型时，选取适当的价格指数作为预测模型输入参数可以提高预测模型的预测能力。

2.2 多元线性回归模型

多元线性回归预测模型是用的较多的统计预测方法。可用于多元线性回归预测的情况有两种：一种是变量间存在非常明显的函数关系；另一种则是变量之间存在着非确定性关系，但存在相关或回归关系。如肿瘤、心血管疾病等多种因素所致的慢性病的发生及死亡与时间之间有联系，但又非确定关系，它们之间有随机性，但又有约束性及规律性。这种统计的规律性符合数学上的最小二乘法准则，因此可用回归模型进行预测。当预测对象依时间变化呈现某种上升或下降的趋势，又无明显的季节波动，且能找到一条合适的函数曲线反映这种变化时，可用时间作自变量，时序数据作因变量，建立回归模型，当有理由认为这种趋势能够延伸到未来时，可赋给自变量所需要的值，得到相应时刻的时间序列未来值。多元线性回归预测法是分析时间序列最常用的方法之一，它适用于无周期变动的的时间序列，一般用于作短期预测。

2.2.1 多元线性回归算法

多元线性回归模型是随机变量 y 对 m 个自变量的 n 组观测值进行回归分析，基本算法如下：

(1) 确定随机变量 y 及 m 个自变量 $(x_1, x_2, x_3, \dots, x_m)$ 。建立回归表达式：

$$y = I_0 + I_1 * x_1 + I_2 * x_2 + I_3 * x_3 + \dots + I_m * x_m + w \quad (2-1)$$

式中， I_0, I_1, \dots, I_m 是模型参数， w 是误差项。 w 反映了除 m 个自变量之外的随机因素对 y 的影响。

(2) 根据样本数据，用最小二乘法，解出式中的参数 I_0, I_1, \dots, I_m 。得到回归方程：

$$y = b_0 + b_1 * x_1 + b_2 * x_2 + b_3 * x_3 + \dots + b_m * x_m \quad (2-2)$$

式中， $b_0, b_1, b_2, \dots, b_m$ 是模型参数 I_0, I_1, \dots, I_m 的估计值。

(3) 把 m 个自变量 $(x_1, x_2, x_3, \dots, x_m)$ 的观测值代入回归方程，计算出 y ，即预测值。

2.2.2 应用多元线性回归模型预测 Brent 原油期货价格

通过相关性分析，选取BRENT原油期货价格的开盘价、最高价、最低价、收盘价、英镑对美元汇率、欧元对美元汇率作为自变量 $(x_1, x_2, x_3, x_4, x_5, x_6)$ ，预测下一天的收盘价 y 。用2005年7月4日至2005年12月1日的数据作为建模数据样本，把2005年12月2日至2006年2月17日的数据作检验样本，建立了多元回归模型，解得多元回归方程为：

$$y = 0.4927 + 0.2279 * x_1 - 0.3958x_2 + 0.3671x_3 + 0.8129x_4 + 0.5358x_5 - 0.5642x_6 \quad (2-3)$$

预测结果如图所示：

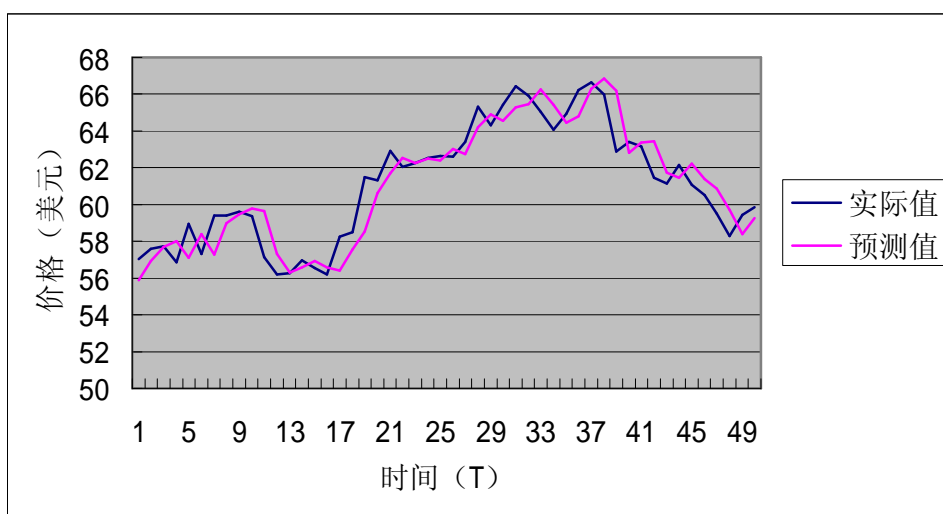


图 2.4 多元线性回归模型预测效果图

应用多元线性回归模型对Brent原油期货价格预测，平均绝对误差为0.94美元，平均相对误差为1.5447%。从预测结果图可以看出，多元线性回归模型基本反映了Brent原油期货价格走势，预测精度与随机游走相似，涨跌判断明显滞后。但多元线性回归模型算法简单，在时间序列预测中应用广泛。

2.3 灰色模型

灰色系统理论和方法是我国学者邓聚龙教授在20世纪80年代提出的用于控制和预测的新理论和方法。灰色系统理论是把一般系统论、信息论、控制论的观点和方法延伸到社会、经济、生态等抽象系统，结合运用数学方法，发展成为一套解决信息不完备系统即灰色系统的理论和方法，它对未来学的研究具有重要意义，自80年代提出以来，已在许多领域的预测工作中得到应用^[1]。

灰色系统理论是一种研究某些既含有已知信息又含有未知或未确知信息的系统理论和方法。它从杂乱无章的、有限的、离散的数据中找出数据的规律，然后建立相应的灰色模型进行预测。灰色理论的实质是对原始随机数列采用生成信息的处理方法来弱化其随机性，使原始数据序列转化为易于建模的新序列。灰色预测的基本原理就是确定一条通过系统的原始序列累加生成的点群的最佳拟合曲线^[2]。

灰色预测的特点是所需的样本数较少，计算简单，因此，比传统的预测方法具有优越性。但是，基本的灰色预测算法也存在很多缺陷，如对于光滑离散函数建模，在数据序列随机性较大时预测结果误差较大。

2.3.1 灰色预测算法

目前应用最广的灰色预测模型是关于数列预测的一个变量、一阶微分的GM(1,1)模型。其预测算法如下^[17]：

(1) 数据处理：将原始数据列 $x^{(0)}(i)$ 做一次累加，得到生成数据列 $x^{(1)}(i)$ ，即：

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i) \quad (2-4)$$

(2) 构成数据矩阵 B 与数据列 Y_N ：

$$B = \begin{pmatrix} -\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \mathbf{L} & \mathbf{L} \\ -\frac{1}{2}[x^{(1)}(n-1) + x^{(1)}(n)] & 1 \end{pmatrix} \quad (2-5)$$

$$Y_N = [x^{(0)}(2), x^{(0)}(3), \mathbf{L}, x^{(0)}(n)]^T \quad (2-6)$$

(3) 计算模型系数 a 和待辨识参数 u ：

$$\hat{\mathbf{a}} = [a, u]^T = (B^T B)^{-1} B^T Y_N \quad (2-7)$$

(4) 建立时间响应模型 $\hat{x}^{(1)}(t)$ ：

$$\hat{x}^{(1)}(k+1) = (x^{(0)}(1) - u/a)e^{-ak} + u/a \quad (2-8)$$

(5) 将时间响应离散化：

$$\hat{x}^{(1)}(k+1) = (x^{(0)}(1) - u/a)e^{-ak} + u/a \quad (2-9)$$

(6) 将 k 值代入离散模型式计算预测累加值 $\hat{x}^{(1)}(t)$ 。

(7) 将预测累加值还原为预测值:

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) \quad (2-10)$$

2.3.2 应用灰色模型预测 Brent 原油期货价格

用2005年7月4日至2005年12月1日的数据作为建模数据样本,把2005年12月2日至2006年2月17日的数据作检验样本,预测结果如图所示:

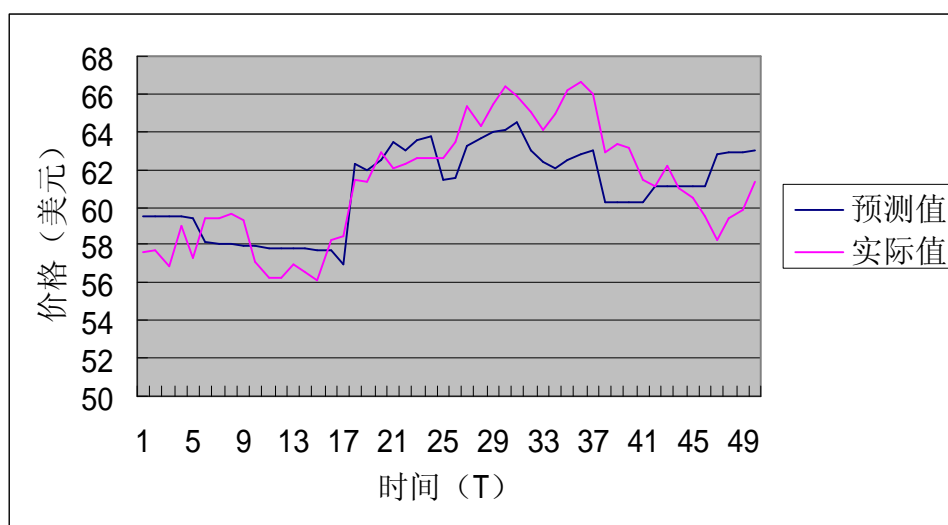


图 2.5 灰色模型预测效果图

应用GM(1,1)模型对Brent原油期货价格预测,平均绝对误差为3.8775美元,平均相对误差为6.166%。从预测结果图看出:GM(1,1)模型预测结果是一条平稳的曲线,大致反映了时间序列的走势,很难反映时间序列的波动变化,因而预测误差比较大。对于随机波动频繁,走势不平稳的Brent原油期货价格时间序列,不适合应用GM(1,1)模型进行预测。

2.4 ARIMA 模型

ARIMA模型,曾一度成为时间序列分析的主要方法。ARIMA方法在统计学上是完善的,有牢固的理论基础,有一套完整的程式化的建模方法。但同时这种方法是复杂的,对数据的性质也有一定的要求^[6]。

ARIMA法建模的基本思想是:将预测对象随时间推移而形成的数据序列视为一个随机序列,即除去个别的因偶然原因引起的观测值外,时间序列是一组依赖于时间 t 的随机变量,构成该时序的单个序列值虽然具有不确定性,但整个序列的变化却有一定的规律性,可以用相应的数学模型近似描述。这组随机变量所具有的依存关系或自相关性表征了预测对象发展的延续性,而这种自相关性一旦被相应的数学模型描述出来,就可以从时间序列的过去值及现在值预测其未来的值^[7]。

2.4.1 ARIMA 模型算法

ARIMA模型(p,d,q),汉语为差分自回归滑动平均模型。其中AR指自回归; p 为模型的自回归项数; MA为移动平均; q 为模型的移动平均项数; d 为时间序列成为平稳之前必须取其差分的次数。其一般的表达式为:

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \dots + a_p y_{t-p} + b_0 u_t + b_1 u_{t-1} + b_2 u_{t-2} + \dots + b_q u_{t-q} \quad (2-11)$$

建模步骤如下:

(1) 时间序列平稳性识别

根据时间序列的散点图、自相关函数(ACF)图和偏自相关函数(PACF)图以及ADF单位根检验观察其方差、趋势及其季节性变化规律,可识别该序列的平稳性。

(2) 数据进行平稳化处理

如果数据序列是非平稳的,并存在一定的增长或下降趋势,则需对数据进行差分处理;如果数据序列存在异方差性,则需对数据进行对数转换或者开方处理,直到处理后数据的自相关函数值和偏相关函数值无显著地异于零。

(3) 根据时间序列模型的识别规则,建立相应的模型

a.若平稳时间序列的偏相关函数是截尾的,而自相关函数是拖尾的,则可断定此序列适合AR模型。

b.若平稳时间序列的偏相关函数是拖尾的,而自相关函数是截尾的,则可断定此序列适合MA模型。

c. 若平稳时间序列的偏相关函数和自相关函数均是拖尾的，则此序列适合ARMA模型。

(4) 进行参数估计，估计暂定的模型参数，检验是否具有统计意义

(5) 进行假设检验，诊断白噪声

检验假设模型残差的ACF值和PACF值在早期或季节性延迟点处不得大于置信区间，同时残差应理想化为0均值。

(6) 利用已通过检验的模型进行预测分析

2.4.2 应用 ARIMA 模型预测 Brent 原油期货价格

用2004年2月17日至2005年12月1日的数据作为建模数据样本，把2005年12月2日至2006年2月17日的数据作检验样本。如图2.8所示：Brent原油期货价格时间序列有明显的上升趋势，从右侧的直方图看出，该数据序列不服从正态分布。

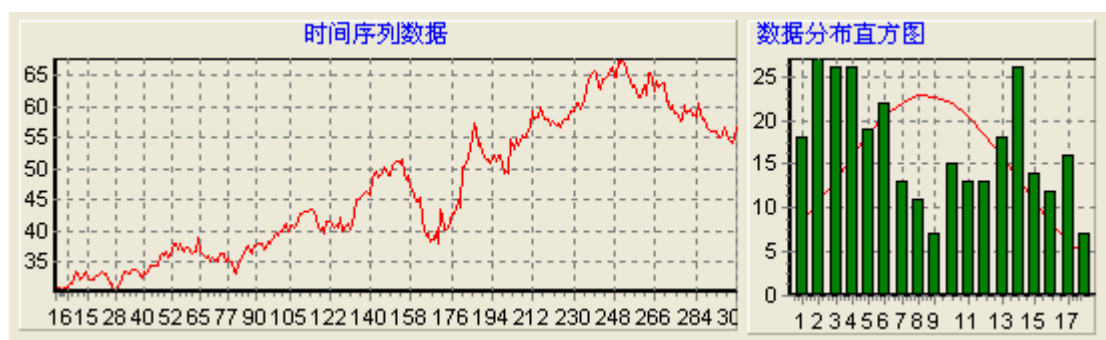


图 2.6 数据样本时间序列图

对原序列进行一次差分后，时间序列图如下：

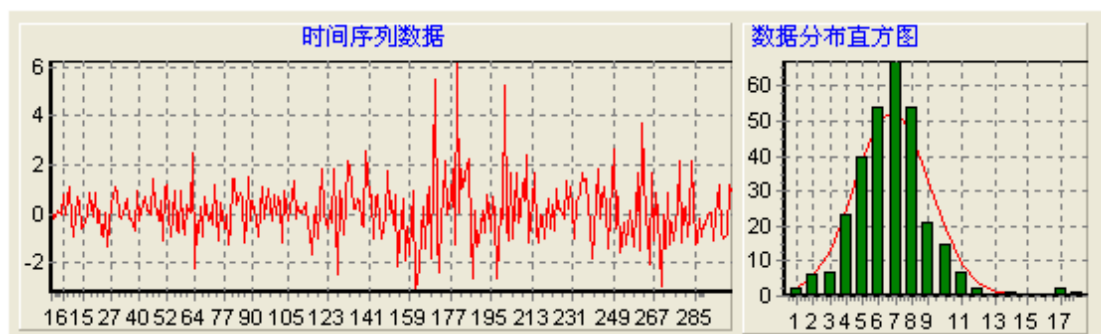


图 2.7 一次差分后时间序列图

从上图可以看出，经过一节差分，剔出了原序列的上升趋势，时间序列比较平稳，且服从正态分布。所以模型中 d 的取值为1，通过对自相关系数和偏相关系数分析， p 、 q 分别取值为4和3。预测模型结构为ARIMA(4,1,3)。

预测平均绝对误差为2.3137，平均相对误差为3.8217%。预测结果曲线图如图所示：

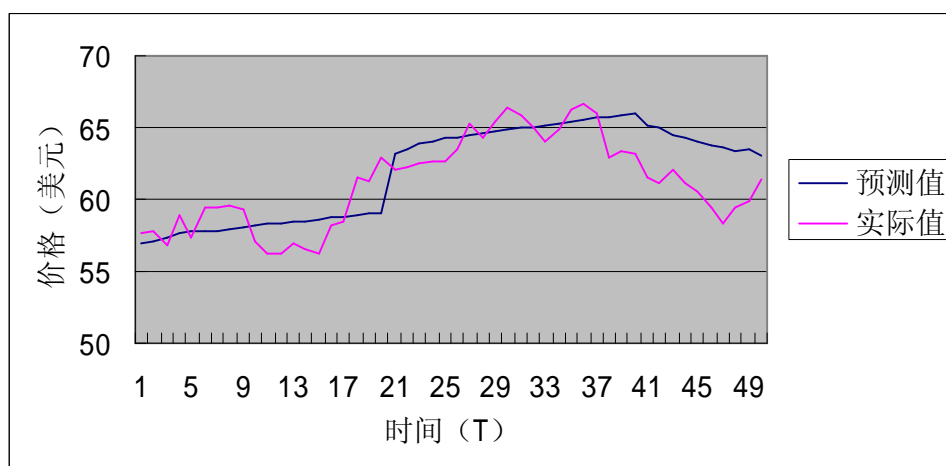


图 2.8 ARIMA 模型预测效果图

ARIMA模型是一个经典的时间序列预测模型，从预测结果图看出，预测结果非常平稳，反映了原时间序列的大致发展趋势，对随机波动不敏感，预测误差比较大，主要原因是ARIMA模型只考虑了原油期货价格时间序列本身的特性，没有充分利用隐含在其它数据指标中的市场信息。

3 人工神经网络预测理论

人工神经网络(ANN)是由大量简单处理单元广泛连接而构成的一个复杂的、非线性大系统,是对大脑神经网络结构和功能的初步模拟。它属于隐式数学处理方法,无须建立数学模型,而它将由网络训练从数据中概括出来的知识,以多组权值和阈值的形式,分别存储于各神经元中,构成网络知识,然后利用该知识来预测相似因素的结果。在科学计算、智能控制、工程预测方面人工神经网络模型得到广泛的应用。一维时间序列数据的人工神经网络预测模型也有学者作了一定研究。相对于传统的数据处理方法,它更适合处理模糊的、非线性、含有噪音及模式特征不明确的问题^[41]。

3.1 人工神经网络模型概述

3.1.1 人工神经网络的发展

人工神经网络依据数据本身的内在联系建模,具有良好的自组织性、自适应性,有很强的学习能力、抗干扰能力。它能自动从历史数据中提取有关知识,可以克服传统预测方法的许多局限以及面临的困难,同时也能避免许多人为因素的影响,因而神经网络模型是在非线性领域应用最广泛的模型。

神经网络诞生半个多世纪以来,经历了5个阶段:

(1) 奠基阶段。

早在40年代初,神经解剖学、神经生理学、心理学以及人脑神经元的电生理的研究等都富有成果。其中,神经生物学家McCulloch提倡数字化具有特别意义。他与青年数学家Pitts合作,从人脑信息处理观点出发,采用数理模型的方法研究了脑细胞的动作和结构及其生物神经元的一些基本生理特性,他们提出了第一个神经计算模型,即神经元的阈值元件模型,简称MP模型,他们认识到了模拟大脑可用于逻辑运行的网络,有一些结点,及结点与结点之间相互联系,构成一个简单神经网络模型。其主要贡献在于,结点的并行计算能力很强,为计算神经行为的某此方面提供了可能性,从而开创了神经网络的研究。这一革命性的思想,产生了很大影响。

(2)第一次高潮阶段。

1958年计算机科学家Rosenblatt基于MP模型，增加了学习机制，推广了MP模型。他证明了两层感知器能够将输入分为两类，假如这两种类型是线性并可分，也就是一个超平面能将输入空间分割，其感知器收敛定理：输入和输出层之间的权重的调节正比于计算输出值与期望输出之差。他提出的感知器模型，首次把神经网络理论付诸工程实现。

(3)坚持阶段。

神经网络理论那遥远但并非遥不可及的目标着实吸引了很多人的目光，美国军方认为神经网络工程应当比“原子弹工程”更重要，并对它的投资兴趣非常大，而对其实实践的效果也比较满意。这时，Minsky警觉的是，人工智能的发展与前途问题。以顺序离符号推导为其基本特征与神经网络大相径庭。他引发学术界的争议，导致对人工智能投资的增加。他从感知器的功能及局限性入手，在数学上进行了分析，证明了感知器不能实现XOR逻辑函数问题，也不能实现其它的谓词函数。他认识到感知器式的简单神经网络对认知群不变性无能为力。1969年Minsky和Papert在MIT出版了一本论著Perceptrons，对当时与感知器有关的研究及其发展产生了恶劣的影响，有些学者把研究兴趣转移到人工智能或数字计算机有关的理论和应用方面。这样，推动了人工智能的发展，使它占了主导地位。美国在此后15年里从未资助神经网络研究课题，前苏联有关研究机构也受到感染，终止了已经资助的神经网络研究的课题。

(4)第二次高潮阶段。

Kohonen提出了自组织映射网络模型，映射具有拓扑性质，并在计算机上进行了模拟，通过实例所展示的自适应学习效果显著。他认为有可能推广到更高维的情况。1982年Hopfield向美国科学院提交了关于神经网络的报告，其主要内容是，建议收集和重视以前对神经网络所做的许多研究工作，他指出了各种模型的实用性。从此，第二次高潮的序幕拉开了。

我国学术界大约在80年代中期关注神经网络领域，有一些科学家起到先导的作用，如中科院生物物理所科学家汪云九，姚国正和齐翔林等。从这时起，我国有些数学家和计算机科学家开始对这一领域产生兴趣，开展了一定的研究工作。

(5)新发展阶段。

1986年Rumelhart和McClelland对神经网络的进展起了极大的推动作用。他们建立了并行分布处理理论，主要致力于认知的微观研究。尤其是，Rumelhart提出了多层网络Back-Propagation法或称ErrorPropagation法，这就是后来著名的BP算法，受到许多学者的重视。

90年代，国内许多学者对Hopfield神经网络的进一步研究很感兴趣，使它得到了一定的完善和发展。

吴佑寿等人提出了一种激励函数可调的神经网络模型，对神经网络理论的发展有重要意义。

1990年Narendra和Parthasarathy提出了一种推广的动态神经网络系统及其连接权的学习算法，它可表示非线性特性，增强了鲁棒性。他们给出了一种新的辨识与控制方案，以multilayer网络与recurrent网络统一的模型描述非线性动态系统，并提出了动态BP参数在线调节方法。

从上述各个阶段发展轨迹来看，神经网络理论有更强的数学性质和生物学特征，尤其是神经科学、心理学和认识科学等方面提出一些重大问题，是向神经网络理论研究的新挑战，因而也是它发展的最大机会。上世纪90年代神经网络理论日益变得更加外向，注视着自身与科学技术之间的相互作用，不断产生具有重要意义的概念和方法，并形成良好的工具。

人工神经网络模型的发展与完善，为非线性复杂预测建模问题铺平了道路。目前，BP神经网络模型理论最为成熟，因而是应用最广泛的预测模型。但是它容易陷入局部最小，并且收敛速度慢，人们对该算法作了很多改进，产生了很多改进的BP算法。例如，把遗传算法对BP网络权值优化改进；把小波理论与神经网络理论结合，用高斯函数代替BP网络中的传递函数S型函数，可以避免陷入局部最小；采取加入动量因子和自适应学习速率的方法来克服BP神经网络易陷入局部极小、收敛速度慢的缺陷等等。此外，径向基函数网络(RBFNN)，广义回归神经网络(GRNN)和概率神经网络(PNN)等神经网络模型也是非常流行的预测模型。

总之，人工神经网络(Artificial Neural Network, ANN)是近年来迅速发展起来的一门集神经科学、计算机科学、信息科学、工程科学为一体的边缘交叉学科，是一种大规模并行的复杂的非线性动力系统，可表示极其复杂的非线性模型系统，具有高度并

行的处理机制、高度灵活可变的拓朴结构以及强大的自组织、自学习、自适应能力和处理非线性问题的能力，得到了预测科学研究的重视。自从1987年Lapedes和Farber首先将神经网络技术应用到预测研究领域以来，国内外许多研究人员开始从事神经网络预测方法的研究，涌现了许多神经网络在预测中成功运用的实例。

3.1.2 人工神经网络的结构

人工神经网络的基本处理单元称为神经元(又称处理单元、节点或短期记忆)，它是对生物神经元的简化与模拟。图3.1给出了一个基本的人工神经元结构。

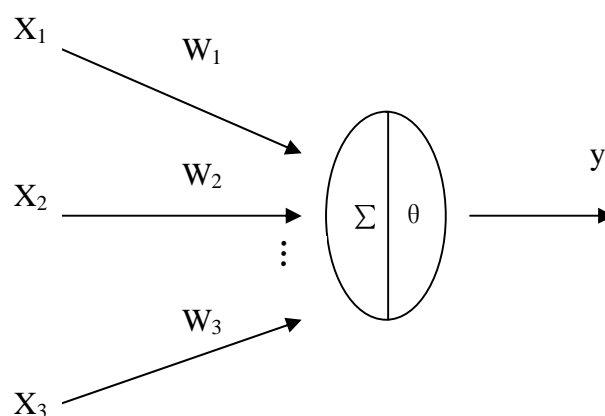


图 3.1 人工神经元结构图

每一个神经元的基本信息加工处理方式可简要说明如下：设 x_1, x_2, \dots, x_n 为神经元 n 个输入， y 为其输出，在每个输入连接上均附加权值 $w_i, i=1, 2, 3, \dots, n$ 。则神经元输出为：

$$y = f\left(\sum_{j=1}^n x_j w_j - \theta\right) \quad (3-1)$$

其中， θ 为该神经元的阈值，函数 f 称谓作用函数(又称激活函数)。为方便起见，将阈值并入权值中， $w_{n+1} = -\theta$ ，相应地增加一个输入 $x_{n+1} = 1$ ，就样就有：

$$y = f\left(\sum_{j=1}^n x_j w_j\right) \quad (3-2)$$

神经元的特性在某种程度上决定了神经网络的总体特性。每个神经元与其他神经元间通过权值连接，接受其他神经元的输出，并经过自身转换函数的转化和闭值的作用输出，从而对其他神经元产生作用。大量简单的神经元的相互连接构成了神经网络，以此来模拟大脑神经系统的结构和功能。网络的建立是依据实际数据，通过对现有数据(知识)学习(训练)。调节节点之间的权重参数，使得网络对新的同类数据具有认知能力。

神经网络按照其结构可以分为前向型、反馈性、随机性和自组织竞争型，其中前向神经网络包括两种基本形式：感知机和多层前向神经网络。感知机由一个输入层和一个输出层组成，中间没有隐层，只能解决线性可分的分类问题。多层前向神经网络由一个输入层、一个或多个隐层、一个输出层组成，可以解决非线性可分的分类问题。目前基于神经网络的时间序列预测研究中，大多数都是使用多层前向神经网络和误差反向传播算法(Back Propagation Algorithm，简称BP算法)。

3.2 BP 神经网络预测模型

3.2.1 BP 神经网络概述

反向传播算法也称BP算法。由于这种算法在本质上是一种神经网络学习的数学模型，有时也称为BP模型。BP算法是为了解决多层前向神经网络的权系数优化而提出来的。BP算法也通常暗示着神经网络的拓扑结构是一种无反馈的多层前向网络，有时也称无反馈多层前向网络为BP模型。1986年，Rumelhart提出了反向传播学习算法，即BP(backpropagation)算法。这种算法可以对网络中各层的权系数进行修正，故适用于多层网络的学习。BP算法是目前最广泛用的神经网络学习算法之一，在自动控制中是最有用的学习算法。研究人员指出在现有的人工神经网络模型中，BP网络是大多数预测应用中最好的工具之一，其预测应用范围很广，包括混沌时间序列序列、天气预报等方面，在一定的领域和某些问题，可以得到比其它方法更好的结果^[8]。

3.2.2 BP 神经网络算法

BP算法是用于前馈多层网络的学习算法。它含有输入层、输出层以及处于输入输出层之间的中间层。中间层有单层或多层，由于它们和外界没有直接的联系，故也称为隐层。在隐层中的神经元也称隐单元。隐层虽然和外界不连接。但是，它们的状态则影响输入输出之间的关系。也就是说，改变隐层的权系数，可以改变整个多层神经网络的性能。对于输入信息，要先前向传播到隐含层节点，经过各单元的特性为Sigmoid型的激活函数(又称作用函数\转换函数或影射函数等)运算后，把隐含层节点的输出信息传播到输出节点，最后给出结果。网络的学习过程由正向和反向传播两部分组成。在正向传播过程中，每一层神经元的状态只影响下一层神经元。若输出层不能得到期望输出，则转入反向传播过程，同时将误差沿原来的连接通路返回，就是实际输出值与期望输出值之间有误差，那么转入反向传播过程，将误差信号沿原来的连接通路返回，通过修改各层神经元连接权值，逐次的向输入层传播去进行计算，再经过正向传播过程，这两个过程的反复运用，使得误差信号最小。实际上误差达到人们所希望的要求时，网络的学习过程就结束。通过训练，神经网络对原始数据产生了“识别”和“记忆”能力，当网络受到曾经学习过或相似的刺激后，网络的输出端便产生相应的输出模型。神经网络这种独具的学习、记忆和识别能力，可用来建立预测模型。

BP算法实质是求取误差函数的最小值问题。这种算法采用非线性规划中的最速下降方法，按误差函数的负梯度方向修改权系数。为了说明BP算法，首先定义误差函数E。取期望输出和实际输出之差的平方和为误差函数。则有：

$$E = \frac{1}{2} \sum_{k=1}^n (y_k - \hat{y}_k)^2 \quad (3-3)$$

其中 \hat{y}_k 为网络实际输出， y_k 为期望输出。设含有 n 个节点的任意网络，各节点之特性为Sigmoid型。为简便起见，指定网络只有一个输出 y ，任一节点 i 的输出为 O_i 。并设有 N 个样本 (x_k, y_k) ($k=1,2,3, \dots, n$)，对某一输入 x_k ，网络输出为 y_k ，节点 i 的输出为 O_{ik} ，节点 j 的输入为：

$$net_{jk} = \sum_i^n w_{ij} O_{ik} \quad (3-4)$$

$$o_{jk} = f(net_{jk}) \quad (3-5)$$

其中 k 表示当前的输入样本为第 k 个样本, w_{ij} 为从神经元 i 到神经元 j 的连接权值, O_{ji} 为神经元 j 输出。 f 为非线性可微非递减函数, 一般为 Sigmoid 型:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3-6)$$

BP 模型把一组样本的 I/O 问题变为一个非线性优化问题, 它使用的是优化中最普通的梯度下降法。这个积和误差对权重的(负)微商成正比(又称梯度下降法), 把它称作权重误差微商。权重的实际改变可由权重误差微商一个模式一个模式地计算出来, 即它们可以在这组模式集上进行累加。如果把神经网络看成输入到输出的映射, 则这个映射是一个高度非线性映射。实际应用时, 学习时要输入训练样本, 每输入一次全部样本称为一个训练周期, 学习要逐个周期地进行, 直到目标函数达到最小值或小于某一给定误差。

用 BP 算法训练网络, 寻求目标函数的极小值有两种方法, 即逐个处理和成批处理。所谓逐个处理, 即随机依次输入样本, 每输入一个样本都进行连接权的调整。所谓成批处理, 是在所有样本输入后计算其误差进行的。通常, 逐个处理要比成批处理更快和更有效, 特别是对于训练样本数很大的情况。这可解释为: 成批处理需要存储器; 能引进随机扰动, 有助于逃脱局部极小等。对于许多应用, 特别是对于需要高精度映射的情况, 成批处理仍然是种选择。总之, 究竟选择哪种方法要视具体问题而定。

3.2.3 BP 神经网络结构

BP 神经网络, 通常由输入层、输出层和若干隐含层构成, 每层由若干个结点组成, 每一个结点表示一个神经元, 上层结点与下层结点之间通过权联接, 同一层结点之间没有联系。实践证明三层神经网络已经足以模拟输入与输出之间的非线性映射关系。更多的网络层次固然能够提高神经网络学习复杂映射关系的能力, 但占用计算机资源过多, 网络学习收敛慢, 因为随着网络层次的增加, 大大增加了神经网络连接权矩阵的规模和网络学习的时间。对于三层神经网络而言, 输入层、输出层神经元的个数已由所求问题本身确定。但中间层神经元个数并不能确定, 需经对特定问题进行试探确定。三层 BP 神经网络的基本结构是由输入层、输出层和隐含层组成的。如图 3.2 所示, 同层之间的神经元没有连接, 相邻层之间的神经元两两连接; 每个神经元都有阈值, 相邻层间的神经元存在连接权 W_{ij} 和 V_{jt} , 分别表示输入层对隐含层、隐含层对输出层神经元影响的大小。正权值表示影响的增加, 负权值表示影响的减弱。每一个神经元

的输出是下一个神经元的输入，神经网络通过对权值和阈值的不断调整输出正确结果。

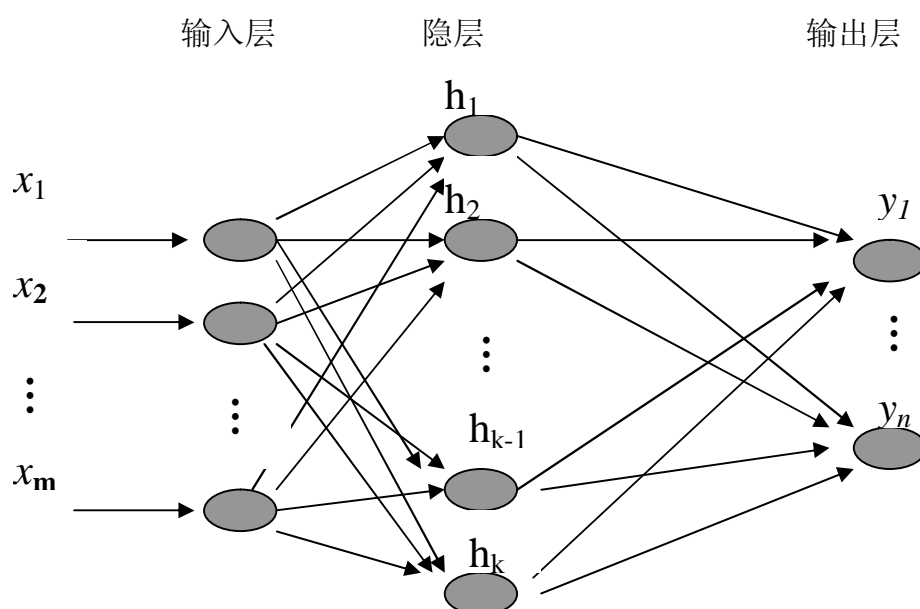


图 3.2 三层 BP 神经网络结构图

3.2.4 BP 神经网络预测模型

BP神经网络预测模型的性能，主要由输入层节点数、隐含层节点数、样本数、学习率、误差指标、训练步长、最大训练循环次数等参数决定。

(1) 网络层数

模型层数的确定尚无统一方法，实践表明，四层网络的结果比三层网络更易进入局部最小。网络层数越多，计算越复杂，越难收敛，因此在实际应用中选用最为常见的三层BP神经网络模型。

(2) 网络节点数

输入输出层节点数由具体问题决定。隐含层节点数的选取尚无理论上的指导，当隐含层节点数太少，则建模不充分，预测的精度无法得到保证；节点数太多，网络在

训练学习中又容易陷入局部极小值。因此，在不减小网络性能的前提下，选择一个最佳的网络结构成为网络设计的关键，这是一个值得进一步研究的问题。隐含层节点数的确定难度较大。目前，基本的作法是实验试凑法，即用同一样本对具有不同隐结点数的网络进行训练，直到权值不再变化，网络稳定为止，然后，依据测试误差最小，确定网络的隐结点数。显然，这种方法比较费时。以下是从事神经网络研究的资深科学家提出来的计算隐层节点数的几个经验公式：

$$m = \sqrt{n + o + a} \quad (3-7)$$

$$m = \log_2^n \quad (3-8)$$

$$m = \sqrt{n \times o} \quad (3-9)$$

m 是隐层节点数， n 是输入节点数， o 是输出节点数， a 是1~10之间的常数，通常需要经过多次试算，根据具体问题选取隐层节点数。

(3) 初始权值选取

由于系统是非线性的，初始值对于学习是否达到局部最小，是否能够收敛以及训练时间的长短的关系很大。如果初始权太大，使得加权后的输入落在了 s 型激活函数的饱和区，从而导致其导数 $f'(x)$ 非常小，由计算权值修正公式可知，权值修正调节过程几乎停顿下来，一般总是希望经过初始加权后的每个神经元的输出值都接近于零，这样可以保证每个神经元的权值都能够在它们的 s 型激活函数变化量最大之处进行调节，所以一般取初始值在 $(-1, 1)$ 之间的随机数。

(4) 功能函数

又称神经元的传递函数或激励函数，用于对求和单元的计算结果进行函数运算，得到神经元的输出。典型的神经元传递函数：线性函数 $f(x) = kx$ ，对数Sigmoid函数，双曲正切Sigmoid函数。

其中对数Sigmoid函数(S型函数)具有非线性放大系数功能，它可以输入从 $-\infty$ 到 $+\infty$ 的信号，变成 -1 到 1 之间输出，对较大的输入信号，放大系数较小，而对较小换输入信号，放大系数较大，所以采用 s 型激活可以去处理和逼近非线性的输入/输出关系。

(5) 学习速率的选择

学习速率决定每一次循环训练中所产生的权值变化量。大的学习速率使得权重变化大，收敛速度就快，可能导致系统的不稳定性；学习速率小，虽然最终可以趋于最小误差值，但是将导致网络训练时间变长，收敛速度减慢。由于学习速率小能保证网络的误差值不跳出误差表面的低谷而最终处于最小的误差值。所以在一般情况下，倾向于选取较小的学习速率以保证系统的稳定性。学习速率的选取范围在0.001-0.8之间。

与隐层单元数的选取过程一样，在一个神经网络的设计过程中。网络要经过几个不同的学习速率的训练，通过观察每次的训练后的目标函数的下降速率来判断所选定的学习速率是否合适。如果目标函数下降很快，则说明学习速率合适，若目标函数出现振荡现象，则说明学习速率过大。

(6) 期望误差

期望误差是在网络的训练过程中，期望输出与实际输出之间允许存在的最大差值，理论上的取值是相对于隐含层节点数目的多寡，因为较小的期望误差值是要靠增加隐含层的节点数，以及训练时间来获得的。实际中一般取0.001-0.0001，若期望误差值过小，则会导致网络的“过拟合”，降低网络的泛化能力(即外推预测性能)。

BP神经网络算法中采用误差平方和最小作为目标函数，在研究中发现，它会过分突出极大值的影响和作用，特别是样本输出最大值和最小值相差较大时，大大降低了实际输出较小的训练样本学习性能或网络预测能力。针对这个问题，许多研究人员给出了不同处理方法，例如陆金桂等人提出了采用绝对误差量和相对误差量的总和作为目标函数的方法，但在一定程度上增加了误差曲面的非线性程度，同时误差在很多场合并不服从正态分布。在本文研究中，采用了一种更简便易行的方法，采用均方差最小作为目标函数，在实验中达到了提高网络性能和预测能力的目的。

(7) 迭代次数

由于神经网络计算并不能保证在各种参数配置下迭代结果收敛，因此应选取一个适合的最大学习次数，若在规定的学习次数内，目标函数仍不能达到预先设定的极小值，则学习失败。

3.2.5 BP 神经网络的不足和改进

BP网络存在的不足，主要收敛速度较慢、易陷入局部极小。

BP算法是利用误差函数对权值的一阶导数信息来指导权值调整，以求最终误差达到最小。在执行过程中，网络参数每次调整的幅度，均以一个与网络误差函数或其对权值的导数大小成正比的项乘以固定的学习因子。这样，在误差曲面曲率较高处，这一偏导数值较大，网络参数调整的幅度也大，以至于在误差函数最小点附近会发生过调整现象，使权值调节路径变为严重的锯齿形，难以收敛到最小点。为保证算法的收敛性，学习率必须很小。这样在误差曲面较平坦处，由于偏导数值本身已很小，网络参数调整的幅度就更小，以至于需要经过多次调整才能将误差函数曲面降低。因此，收敛速度缓慢的直接原因是：固定的学习率、学习过程中样本与神经元的收敛速度不平衡以及学习过程中出现的“平台”或“瘫痪”现象。BP算法是以梯度下降法为基础的非线性优化方法，不可避免地存在局部极小问题。从结构上讲，网络输入输出间的非线性关系致使网络的误差或能量函数是一个具有许多局部极小点的非线性空间，常常导致网络落入局部最小点不能自拔，而达不到全局最小点^[21]。

(1) 对于收敛速度慢的改进

a. 动量法

此方法将上一次权值调整量的一部分迭加到按本次误差计算所得的权值调整量上，作为本次的实际权值调整量，即：

$$\Delta W(n) = -h \Delta E(n) + a \Delta W(n-1) \quad (3-10)$$

式3-10中， a 为动量系数，通常 $0 < a < 0.9$ ； $a \Delta W(n-1)$ 是增添的动量项。

b. 学习率 h

BP算法的有效性和收敛性，在很大程度上取决于 h 值。 h 的最优值与具体问题有关，没有对任何问题都适合的 h 值。即使对某一特定问题，也很难找到一个自始至终都合适的 h 值。训练开始时较合适的 h 值，后来不一定合适。因此，为解决 h 的取值问题，可采用自适应学习率。对于成批处理可以根据总误差变化的信息进行启发式调整，其规则是：若总误差 E 减小(即新误差比老误差小)，则学习速率增加(例如将实际值乘以因子 $a=1.05$)；若总误差 E 增加(即新误差比老误差大)，则学习速率减小，当新

误差与老误差之比超过一定值(例如1.40), 则学习速率快速下降(例如将实际值乘以因子 $b=0.7$)。

(2) 克服局部极小的方法

BP网以梯度信息指导权值调整机制使其易陷入局部最小的缺点无法自行克服, 必须有新的思想的引入。采用模拟退火算法的随机神经网络无疑是经典的, 它的基本思想是赋予网络一定的“爬坡”能力, 使其有可能跳出“局部低谷”而最终落入“全局低谷”, Boltzmann机是此算法的代表。

近年来, 随着遗传算法的兴起, 对BP算法的改进方面的研究已取得了长足的进展。遗传算法是模拟达尔文的遗传选择和自然淘汰的生物进化过程的计算模型。它的思想源于生物遗传学和适者生存的自然规律, 是具有“生存+检测”的迭代过程的搜索算法。遗传算法以一种群体中的所有个体为对象, 并利用随机化技术指导对一个被编码的参数空间进行高效搜索。其中, 选择、交叉和变异构成了遗传算法的遗传操作, 参数编码、初始群体的设定、适应度函数的设计、遗传操作设计、控制参数设定五个要素组成了遗传算法的核心内容。遗传BP网络的主要机理是利用遗传算法全局搜索能力强的特点, 先用遗传算法对BP网络的权值和闭值进行全局粗搜索, 定位最优解区域, 使得权值和阈值种群聚集在参数解空间的某几处, 再用BP算法分别对其进行梯度细搜索, 最终求得最优解。

3.3 径向基神经网络预测模型

神经网络具有极强的非线性映射能力, 能够逼近任意非线性函数, 为模拟非线性时间序列提供了一种有力的工具, 在预测领域中显示了很大的优势。

3.3.1 径向基神经网络的基本原理

1985年, Powell提出了多变量插值的径向基函数(Radial-Basis Function, RBF)方法。1988年, Broom head和Lowe首先将RBF应用于神经网络设计, 构成了径向基函数神经网络, 即RBF神经网络。用径向基函数(RBF)作为隐单元的“基”构成隐含层空间, 对输入矢量进行一次变换, 将低维的模式输入数据变换到高维空间内, 通过对隐单元输出的加权求和得到输出。这就是RBF网络的基本思想。

3.3.2 径向基神经网络结构

径向基函数(RBF)法是将输入矢量扩展到或预处理到高维空间中的神经网络学习方法。RBF神经网络结构是一种三层前向网络。输入层由信号源结点组成；第二层为隐含层，单元数视所描述的问题而定；第三层为输出层，它对输入模式的作用做出响应。

RBF神经网络利用了这样的数学基础：分类问题在高维空间（在某种特殊意义下的）比在低维空间中更可能是线性可分的。构成RBF网络的基本思想是：用RBF作为隐层神经元的“基”构成隐含层空间，这样就可将输入矢量直接（即不通过权连接）映射到隐空间。当RBF的中心确定后，这种映射关系也就确定了。而隐含层空间到输出层空间的映射是线性的，即网络的输出是隐层神经元输出的线性加权和，此处的权为网络的可调参数。这样网络的权就可由线性方程组解出或用RLS（递推最小二乘）方法递推计算，从而大大加快学习速度并避免局部极小问题。

3.3.3 径向基神经网络预测算法

RBF神经网络要学习的参数有三个：基函数的中心和方差以及权值。根据径向基函数中心选取方法的不同，RBF网络有多种学习方法，其中最常用的四种学习方法有：随机选取中心法、自组织选取中心法、有监督选取中心法和正交最小二乘法。

(1) 随机地从输入样本中选取RBF中心，中心固定后，可以计算高斯函数的宽度：

$$s = \frac{d_m}{\sqrt{2m}} \quad (3-11)$$

这里 d_m 为所选中心之间的最大距离， m 为隐层神经元个数，这时隐层神经元的函数输出也就已知了，于是权值就可通过解线性方程组得到。

(2) 自组织学习选取RBF中心。在这种算法中，首先RBF的中心通过自组织学习确定其位置，一般常采用模式识别中广泛应用的k-均值聚类算法，是无监督的学习算法，学习的目的是使RBF的中心位于输入空间重要的区域，并根据算法(1)确定基函数宽度；而输出层的线性权通过有监督的误差校正学习算法，如最小二乘法等。所以这是一类混合学习算法。

(3) 有监督的学习选取RBF中心。在这种方法中RBF的中心以及网络的其他参数都是通过有监督的学习来确定的。一般可用简单有效的梯度下降法。该算法虽简单直接,但思想类似于BP网络的BP 算法,训练结果存在收敛性问题,跟初始化参数有关。

(4) 正交最小二乘法 (OLS) 选取中心。OLS 方法来源于线性回归模型,在该算法中,首先把网络的求解问题转化为一个线性回归模型,对回归进行正交化,然后确定权值和中心。

RBF网络时间序列预测建模的基本步骤如下:

- A. 样本选取: 将已知的时间序列样本数据分成训练段和检验段。
- B. 网络结构设计: 确定网络的输入向量、输出向量的维数,隐含层节点数。
- C. 网络拓扑设计: 选择网络输出层的激活函数(隐含层选用高斯基函数)。
- D. 网络初始化: 设置网络的初始权值、目标误差值以及高斯基函数宽度参数。
- E. 网络训练: 利用聚类算法确定RBF隐层中心、最小二乘算法确定连接权值,训练网络,拟合学习段时间序列,直到网络收敛于一定的标准。否则,可重新改变网络的初始权值甚至网络的拓扑结构,直至训练结果满意。
- F. 网络预测: 利用检验段数据检验训练好的网络模型对未来时间序列进行预测。

4 组合预测理论

4.1 组合预测简介

在预测实践中，对同一问题常采用不同的预测方法。不同的预测方法提供不同的有用信息，其预测精度往往也不同。如果简单地将预测误差较大的一些方法舍弃掉，将会丢失一些有用的信息，这种做法对信息是一种浪费，应予以避免。一种更为科学的做法是，将不同的预测方法进行适当的组合，从而形成所谓的组合预测方法。组合的主要目的是综合利用各种方法所提供的信息，尽可能地提高预测精度。

早在1954年，美国人Schmitt 就曾经用组合预测方法对美国37个大城市的人口进行过预测，使预测精度有所提高。1959年J. M. Bates 和C. W. J. Granger 对组合预测方法进行了系统的研究，其研究成果引起了预测学者的重视。进入70年代，组合预测的研究更被预测工作者所重视，发表了一系列关于组合预测的论文。组合预测方法广泛应用于汇率预测、股价预测、城市用水量预测、电力负荷预测等领域。

组合预测研究的实质就是求出单项预测模型的加权平均系数，使得组合预测模型更加有效地提高预测精度。经过几十年的研究和实践，人们提出了各种各样的组合方法，组合预测的方法主要有简单平均组合预测方法、最优组合预测方法和变权重预测方法^[23]。

4.2 简单平均组合预测模型

简单平均组合预测方法是不变权重组合预测方法之一。设对同一预测问题有 m 种预测方法，实际观测值有 n 期， y_t 为第 t 期的统计观测值， f_{it} 为第 i 种方法第 t 期的拟合预测值，其预测误差 $e_{it} = y_t - f_{it}$ ，第 i 种方法的预测误差向量

$$E_i = (e_{i1}, e_{i2}, e_{i3}, \dots, e_{in})^T, i = 1, 2, \dots, m。$$

$$f_t = \sum_{i=1}^m w_i^{(t)} f_{it} \quad (4-1)$$

式中 $w_i^{(t)} = \frac{1}{m}, 1 \leq i \leq m。$

简单的组合预测模型算法易于理解、计算过程简单，在预测实践中得到广泛应用。

4.3 最优组合预测方法

最优组合法的思想是根据“过去一段时间内组合预测误差最小”这一原则来求取各个单项预测方法的权系数向量的。

设对同一预测问题有 m 种预测方法， y_t 为第 t 期的实际值 ($t=1,2,3,\dots,n$)， g_{it} 为采用第 i 种方法对第 t 期拟合预测值，其预测误差 $e_{it} = y_t - g_{it}$ ，向量 $W = (w_1, w_2, w_3, \dots, w_m)^T$ 分别为 m 种预测方法在组合预测中的权重，则有组合预测模型为：

$$\hat{y}_t = \sum_{i=1}^m w_i * g_{it} \quad (4-2)$$

$$e_{it} = y_t - g_{it} \quad (4-3)$$

$$e_t = \sum_{i=1}^m w_i (y_t - g_{it}) = \sum_{i=1}^m w_i e_{it} \quad (4-4)$$

$$E = \sum_{t=1}^n e_t^2 = W^T E(n) W \quad (4-5)$$

式中， $E(n) = (E_{ij})_{mm}$ 是预测误差协方差阵，依据误差平方和最小准则，则可构造以下数学规划模型：

$$\begin{aligned} \min E &= W^T E(n) W \\ \text{s.t.} \sum_{i=1}^m R_n^T w_i &= 1 \end{aligned} \quad (4-6)$$

$R(n) = (1, 1, \dots, 1)^T$ 为元素全部为1的 n 维列向量。采用拉格朗日乘法可以得到模型的唯一最优解：

$$W = \frac{E^{-1} R_m}{R_m^T E^{-1} R_m} \quad (4-7)$$

但是可能会出现负权重，负权重是否具有现实意义的问题还存在争议。如果给权重加上非负限制，则原问题变成求解下列非线性规划问题：

$$\begin{aligned} \min E &= W^T E(n) W \\ \text{s.t. } \sum_{i=1}^m R_n^T w_i &= 1 \\ w_i &\geq 0 \end{aligned} \quad (4-8)$$

文献[42]证明，上述非线性规划有唯一解。

4.4 变权重预测方法

由于单项预测模型经常表现出“时好时坏”性，因此合理变权重比不变权重更为科学。变权重的组合预测方法已经成为组合预测领域的热点之一。

最早的变权重组合预测方法有Diebold(1987)年提出的组合权重的最小方差估计方法和Gunter(1989)年提出的多步确定方法。经过10多年的研究，到目前为止，随时间变化的组合权重的确定方法有以下几种：

不变权重组合预测方法的权重系数反映的只是各单项预测方法在整个预测中的平均重要程度，变权重组合预测方法的权重系数则能够随着各单项预测方法在组合预测中重要程度的不同而不断调整，因而能更充分的利用各单项预测方法所提供的信息。

变权重组合预测方法的权重系数是不断变化的，这使求解变得更复杂和困难，在预测实践中的应用也不如不变权重组合和预测方法广泛。对变权重组合和预测方法的研究主要是变权系数的确定，对该方面的研究可分为两大类：一类是利用模糊逻辑理论求解变权重系数，另一类是通过建立数学规划模型求解权重系数。

5 基于概率统计的组合预测模型

上文中对Brent原油期货收盘价时间序列进行了深入分析，并分析了它与纳斯达克综合指数、SP500指数、欧元对美元汇率、英镑对美元汇率等数据序列的相关性。本文根据Brent原油期货收盘价时间序列的特点，引入了神经网络理论和组合预测理论，提出了基于概率统计的组合预测模型。主要内容包括组合算法、组合权系数的确定和单项预测模型的选择。

5.1 模型算法

从图2.1可以看出，Brent原油期货价格有明显的上升趋势，对该序列一次差分后，时间序列变为平稳的随机波动序列，没有明显的周期性和季节性，进一步说明Brent原油期货价格是趋势序列与随机波动序列叠加而成的。

因此，本文用一个线性的时间序列模型来描述其趋势成分，然后用神经网络模型来描述其非线性的成分。本文采用的指数平滑方法是最典型的时间序列趋势描述方法之一，指数平滑参数取0.75，对原序列进行拟合，把原序列分离为两部分：

$$y = y_L + y_N \quad (5-1)$$

y_L 为指数平滑拟合序列，作为原序列的趋势部分， y_N 为原序列 y 与 y_L 的差值序列，是随机波动的部分。趋势部分与波动部分曲线图如图5.1所示：

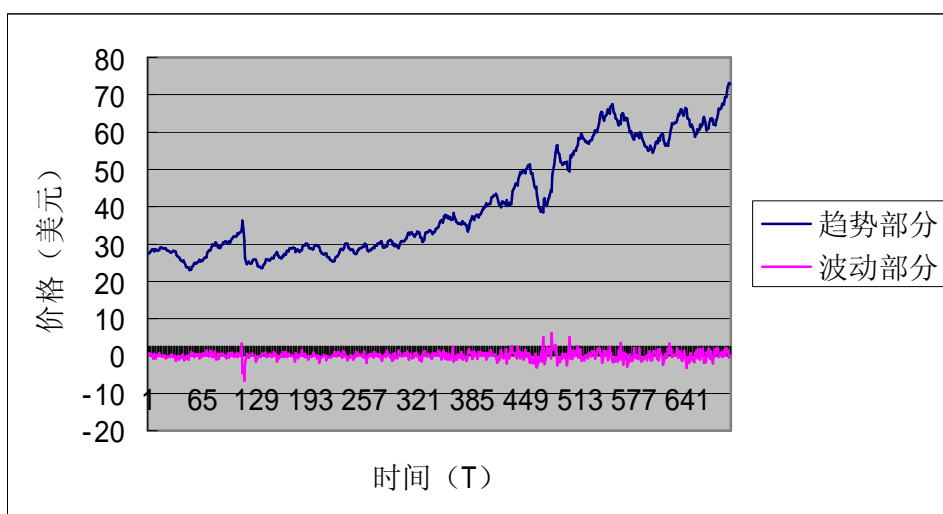


图 5.1 序列趋势与波动分离图

图5.2是波动部分 y_N 序列分析图，随机波动序列在零值附近随机波动，并服从正态分布。

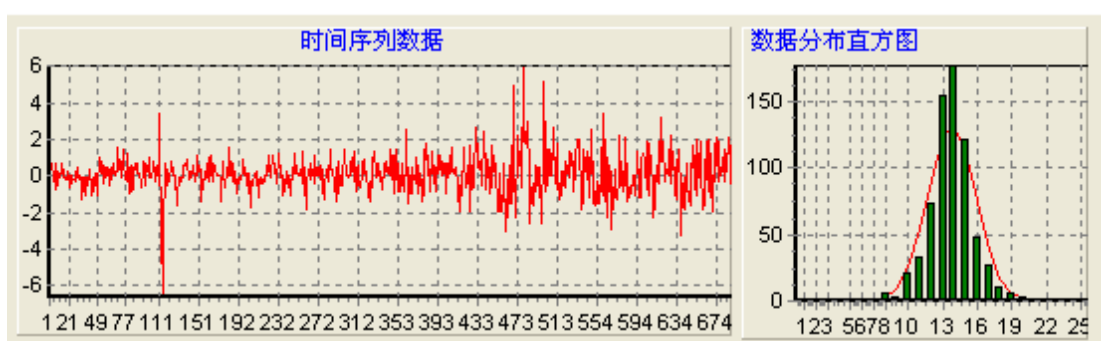


图 5.2 BRENT 原油期货价格随机波动序列图

5.1.1 组合算法

为了提高预测的精确度以及预测稳定性，本文采用多个神经网络模型对 y_N 进行预测，总的预测模型为：

$$y_t = y_{L_t} + \sum_{i=1}^m w_{it} y_{it} \quad (5-2)$$

y_t 表示第 t 期组合预测值, y_{Lt} 表示第 t 期的指数平滑值, y_{it} 为第 i 个神经网络模型对随机波动第 t 期的预测值, w_{it} 为第 i 个神经网络模型第 t 期的权系数。本文基于概率统计的组合预测模型的重点就是组合权系数的确定算法和单项预测模型的选择。

5.1.2 组合权系数的确定

通过统计分析,波动序列 y_N 全体样本服从 $N(0.094, 1.1)$ 正态分布。在测试集附近取400个样本进行统计分析,波动序列也服从正态分布,参数为 $N(0.15, 1.4)$, 95%置信区间为 $(0.0315, 0.2605)$, 直方图如图5.3、5.4所示,

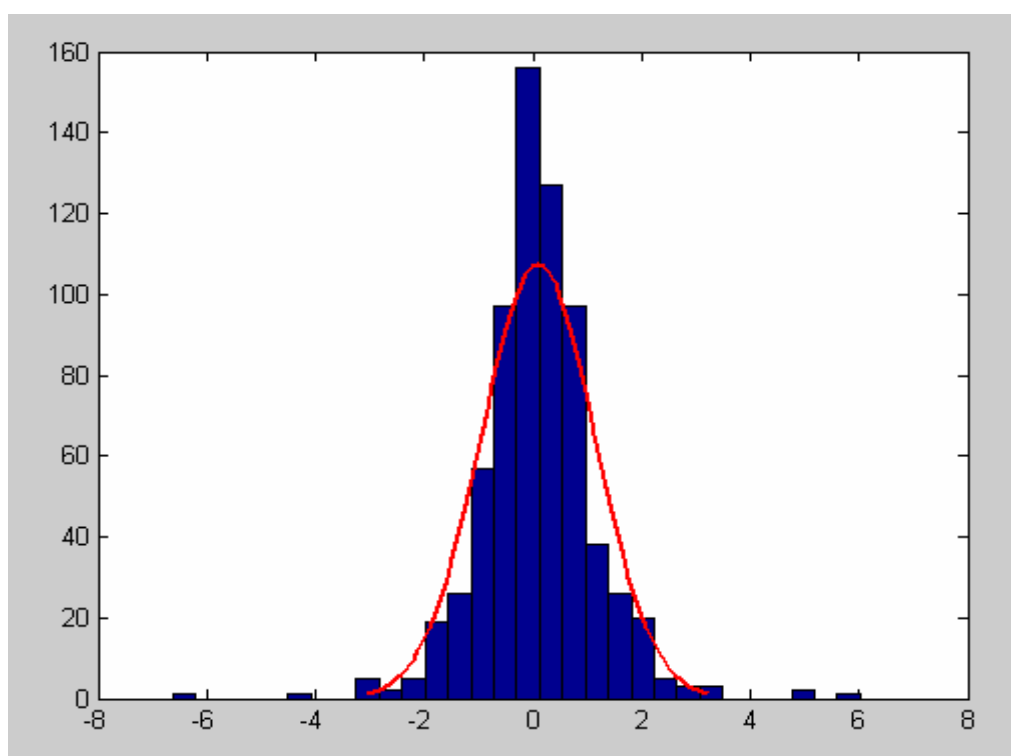


图 5.3 波动序列全体样本直方图

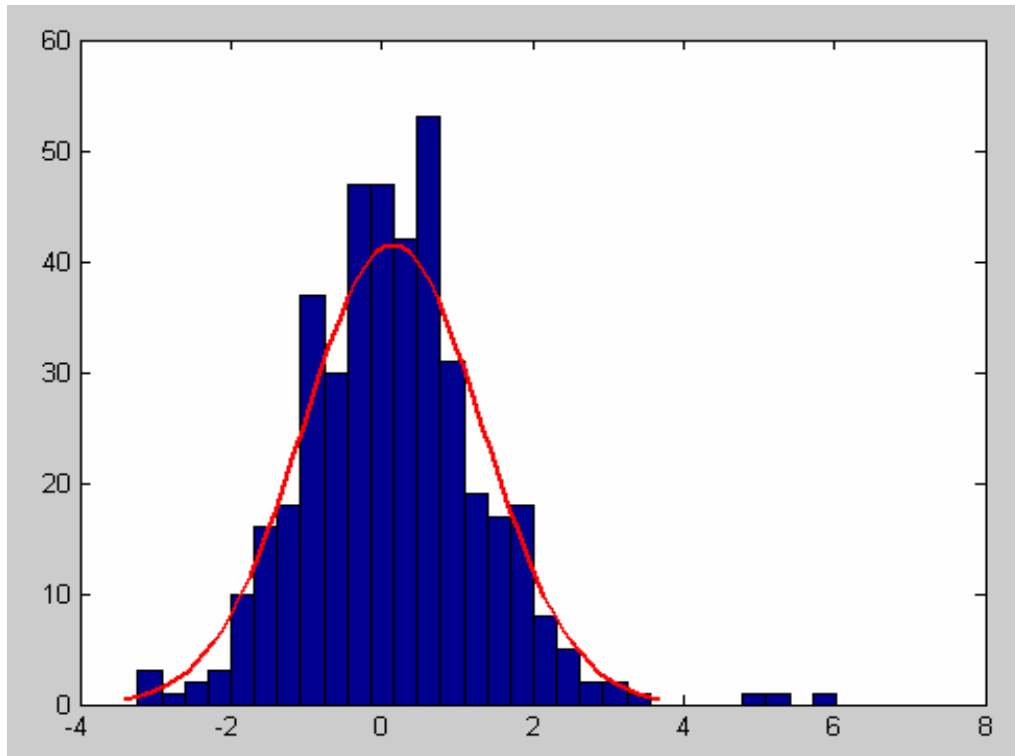


图 5.4 测试集附近波动序列直方图

从图5.4看出，波动绝对值大于4美元的概率接近于零，越接近正态分布的中心，波动值出现的概率越大。因此，本文根据这一特点将组合权系数 w_{it} 取值与该单项预测模型预测值出现概率结合起来，出现概率越大，权系数越大。理论上，基于概率统计的组合方式预测效果不比通常使用的简单平均组合方式差。

本文把波动值分布区间划分为若干个子区间，并计算出波动值落在该区间的概率，公式5-2中 w_{it} 的值大小由第 i 个神经网络模型第 t 期预测值所在区间的概率决定。表5-1是区间划分及其概率情况：

表 5-1 随机波动概率区间分布

区间	$(-\infty, -4]$	$(-4, -2]$	$(-2, -1]$	$(-1, 0]$	$(0, 1]$	$(1, 2]$	$(2, 4]$	$(4, +\infty)$
概率	0	0.0217	0.1184	0.3068	0.3599	0.1401	0.0459	0.0072

设有 m 个预测模型，在第 t 期 m 个模型的权系数算法为：

$$\hat{y} = (\hat{y}_1, \hat{y}_2, \mathbf{L}, \hat{y}_m) \quad (5-3)$$

$$p = (p_1, p_2, \mathbf{L}, p_m) \quad (5-4)$$

$$K = \sum_{i=1}^m p_i \quad (5-5)$$

$$W = (w_1, w_2, \mathbf{L}, w_m) = \left(\frac{p_1}{K}, \frac{p_2}{K}, \mathbf{L}, \frac{p_m}{K}\right) \quad (5-6)$$

其中 \hat{y} 表示 m 个模型第 t 期的预测值向量。 p 表示 m 个预测值所在区间对应表 5-1 中的概率组成的向量， W 为权系数向量。

举例说明：M1, M2, M3 分别表示建立的三个预测模型，在第 t 期 3 个模型的波动预测值分别为 -1.5、0.5、2.5，3 个预测值对应的概率分别为 0.1184、0.3599、0.0459，对 3 个预测值归一化值分别为 0.2259、0.6866、0.0875。作为 3 个模型第 t 期的权系数，即 $W_t = (w_{1t}, w_{2t}, w_{3t}) = (0.2259, 0.6866, 0.0875)$ 。

5.1.3 单项模型的选择标准

组合预测的目的就是要充分利用不同模型隐含的信息片断，有些模型具有较强的拟合能力，能够更好地推断时间序列的发展趋势，有些模型善于发现各个影响因素对预测值指导关系。所以不能一味地选择预测误差最小的模型，而是要选择最有特点的若干个模型进行组合，尽可能多地搜集到各种指导信息片断，使预测系统更准确地反映真实情况。在选择单项预测模型时，通常评价预测模型优劣的指标有如下几个：

(1) 平均误差和平均绝对误差：

$$\text{平均误差的公式： } ME = \frac{\sum_{i=1}^n e_i}{n} \quad (5-7)$$

$$\text{平均绝对误差的公式： } MAE = \frac{\sum_{i=1}^n |e_i|}{n} \quad (5-8)$$

(2) 平均相对误差和平均相对误差绝对值

$$\text{平均相对误差的公式： } MPE = \frac{1}{n} \sum_{i=1}^n \frac{y_i - \hat{y}_i}{y_i} \quad (5-9)$$

$$\text{平均相对误差绝对值的公式： } MPAE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (5-10)$$

(3) 预测误差的方差和标准差

$$\text{预测误差的方差公式: } MSE = \frac{1}{n} \sum_{i=1}^n e_i^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5-11)$$

$$\text{标准差公式: } SDE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5-12)$$

上述几个指标都是对预测精确度的不同描述方式，它更多地说明了该预测模型对原序列的拟合能力，然而预测精度只是预测模型预测效果好坏的指标之一。因此，本文还引入了涨跌判断准确率指标，用来描述预测模型对未来价格涨跌的判断能力。

5.2 单项预测模型

这里说的单项预测模型其实也是一个组合预测模型，如公式5-13所示， y_L 部分用指数平滑预测， y_N 部分用一个神经网络模型预测， y 为两个预测模型叠加值。

$$y = y_L + y_N \quad (5-13)$$

因为此公式5-13中只用了一个神经网络预测模型，相对于后面提到的组合预测模型，把这个组合预测模型当作单项预测模型。

在分析各个因素的相关性，以及对Brent原油期货价格时间序列分析的基础上，用Matlab神经网络工具箱建立多个BP神经网络和径向基预测模型。

本文采用含有一个隐层的3层BP神经网络，输入参数是上述各种指标历史数据，输出节点只有一个，那就是Brent原油期货价格下一天的收盘价。输入层与隐层之间的传递函数是“tansig”函数，隐层与输出层的之间的传递函数是“logsig”，采用“trainlm”训练算法，训练最大次数是300次，训练目标误差是0.003，训练学习率是0.1。输入节点数由不同预测模型的参数个数决定，由于预测模型的隐层节点数与输入节点数有关，因而每个预测模型的隐层节点数都不一样。训练样本集为前590个样本，测试集取第591~690个样本。

为了区别每个预测模型，给每个模型取了一个简洁明了的名称。

5.2.1 模型 L_bp_all

L_bp_all预测模型是一个指数平滑模型与一个BP神经网络预测模型叠加的结构。其中的BP神经网络输入参数为Brent原油期货开盘价、最高价、最低价、收盘价、成交量、持仓量、WTI原油期货收盘价、纳斯达克指数、SP500指数、美元英镑汇率、美元欧元汇率、道琼斯指数，共12个输入节点，隐层节点数为9。经过225次训练达到收敛。神经网络预测模型输出结果与趋势部分指数平滑值叠加后，预测平均绝对误差0.80504，平均相对误差0.013318，涨跌判断准确率56.566%，预测效果如图5.5所示：

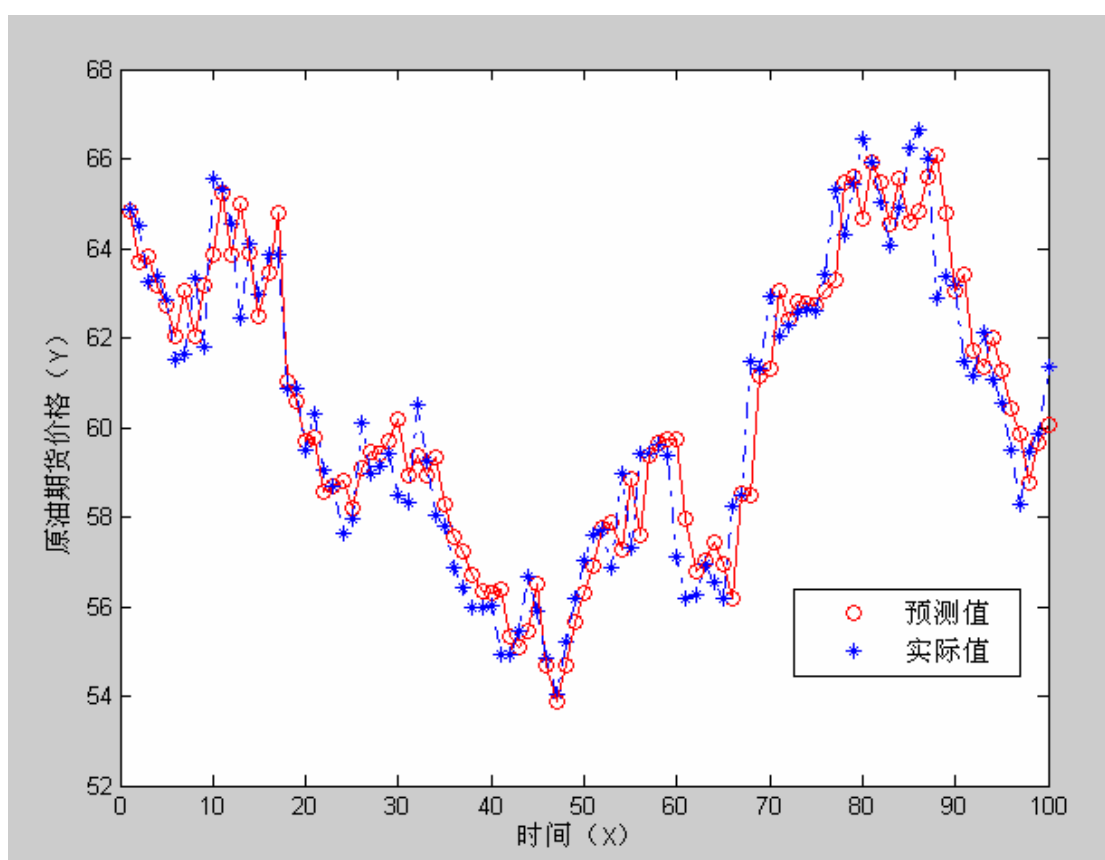


图 5.5 L_BP_ALL 模型预测效果图

5.2.2 模型 L_bp4

L_bp4预测模型是一个指数平滑模型与一个BP神经网络预测模型叠加的结构。其中的BP神经网络输入参数为Brent原油期货开盘价、最高价、最低价、收盘价，用来预测下一天的收盘价的随机波动部分，隐层节点数为8。与趋势部分指数平滑值叠加

后，预测平均绝对误差1.0911，平均相对误差0.018045，涨跌判断准确率45.455%，预测效果如图5.6所示：

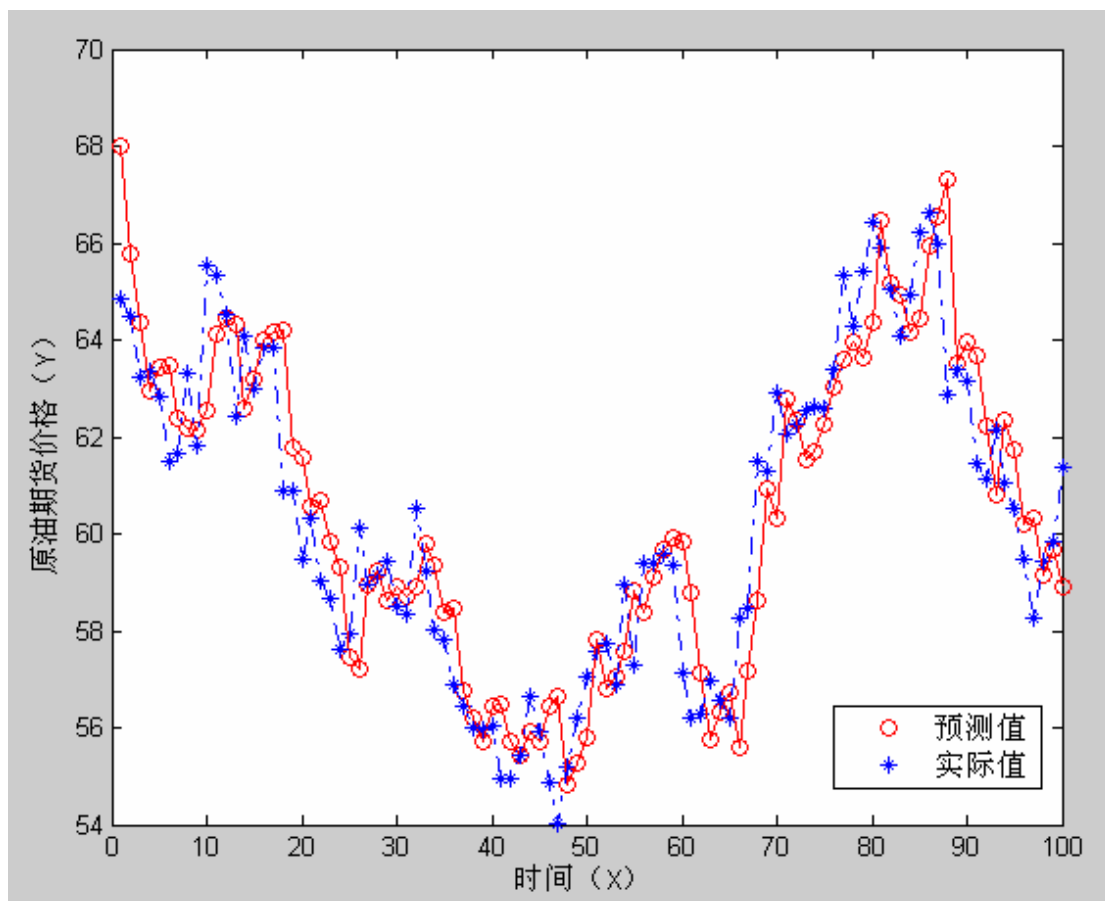


图 5.6 L_BP4 模型预测效果图

5.2.3 模型 L_bp5s

L_bp5s预测模型是一个指数平滑模型与一个BP神经网络预测模型叠加的结构。其中的BP神经网络输入参数为前五天的收盘价，共五个输入节点，隐层节点数为16。神经网络预测模型输出与趋势部分指数平滑值叠加后，预测平均绝对误差0.83713，平均相对误差0.01382，涨跌判断准确率57.576%，预测效果如图5.7所示：

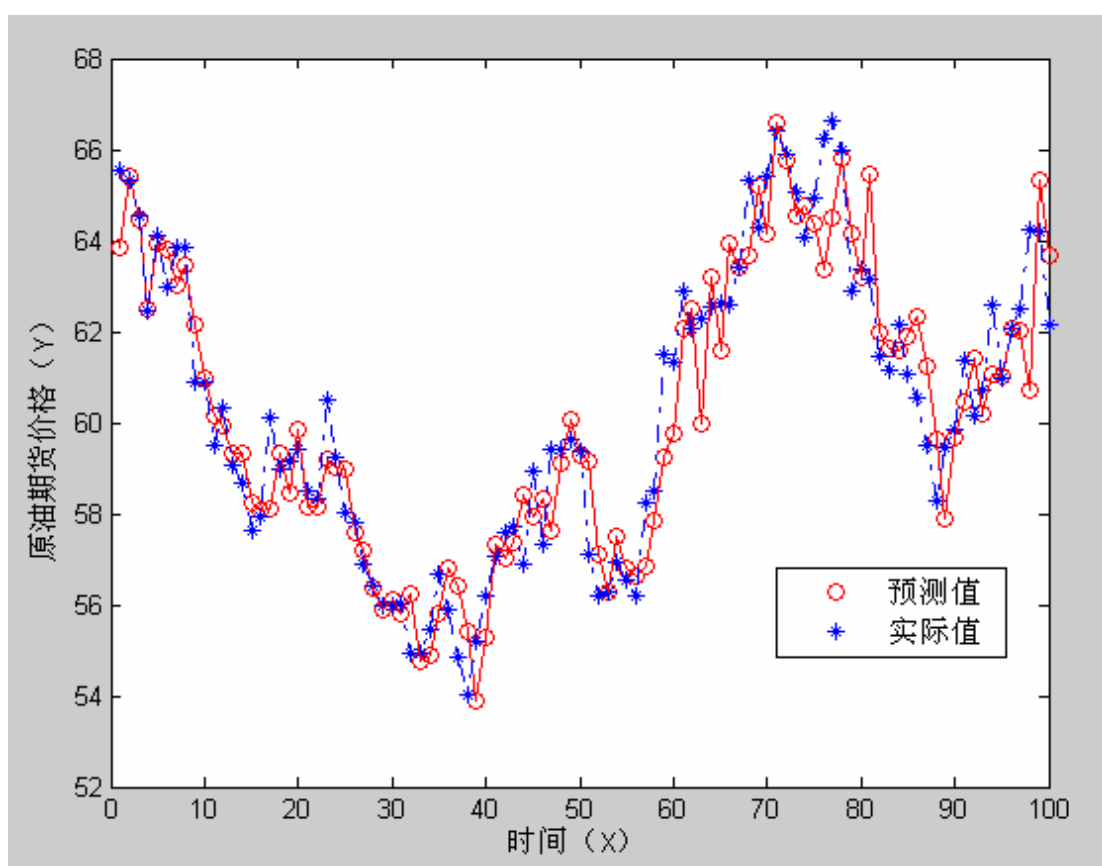


图 5.7 L_BP5s 模型预测效果图

5.2.4 模型 L_bp10s_3z

L_bp10s_3z预测模型是一个指数平滑模型与一个BP神经网络预测模型叠加的结构。其中的BP神经网络输入参数为Brent原油期货前10天的收盘价，以及WTI原油期货收盘价、纳斯达克指数、SP500指数、美元英镑汇率、美元欧元汇率、道琼斯指数前8、9、10天的值。因为这些指数与Brent原油期货收盘价的相关指数在14天内随时间间隔越远，相关系数越大，所以选择这些指数前8、9、10天的值作为BP神经网络预测模型的输入参数。输入节点数为28，隐层节点数为13，经过275次训练达到收敛，BP神经网络预测模型输出与趋势部分指数平滑值叠加后，预测平均绝对误差0.83294，平均相对误差0.013775，涨跌判断准确率59.596%。预测效果如图5.8所示：

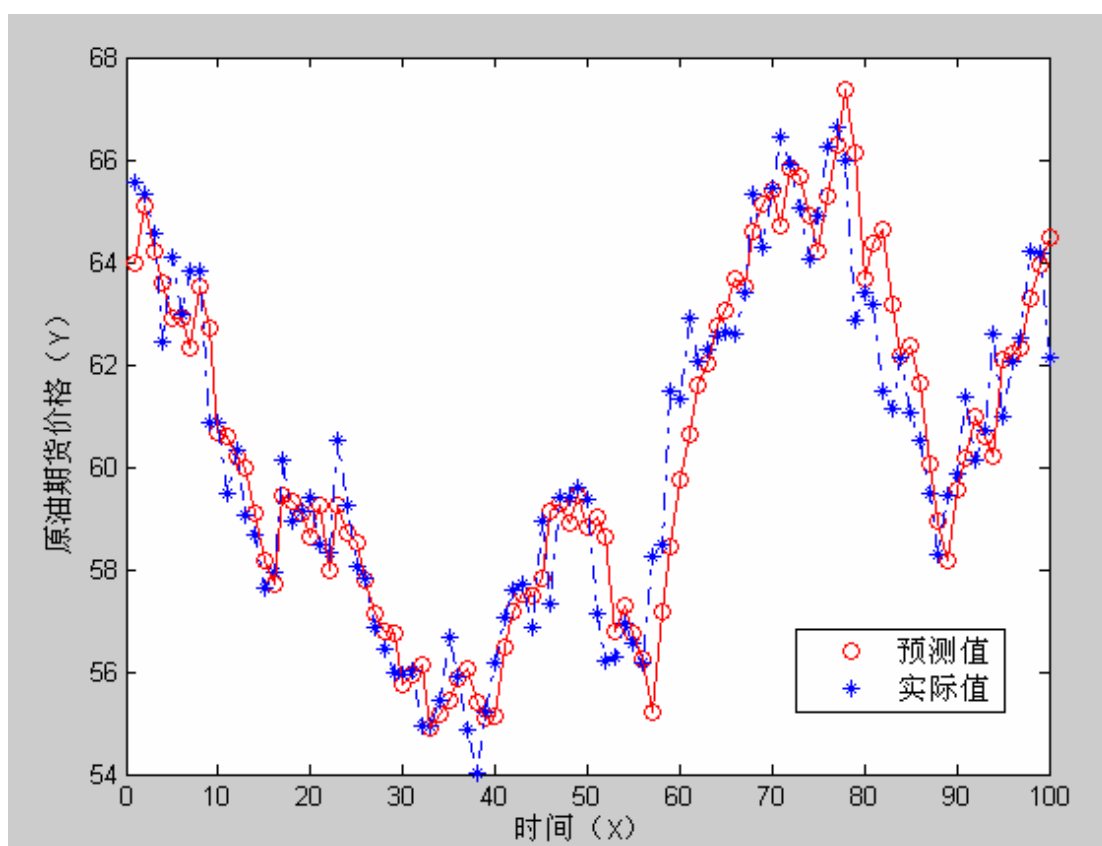


图 5.8 L_BP10s_3z 模型预测效果图

用Matlab神经网络工具箱建立径向基神经网络预测模型，训练目标误差为0.005，径向基函数分布系数为5。

5.2.5 模型 L_rbf_all

L_rbf_all预测模型是一个指数平滑模型与一个径向基神经网络预测模型叠加的结构。其中的径向基神经网络输入参数为Brent原油期货开盘价、最高价、最低价、收盘价、成交量、持仓量、WTI原油期货收盘价、纳斯达克指数、SP500指数、美元英镑汇率、美元欧元汇率、道琼斯指数，共12个输入参数。把径向基神经网络预测模型输出与趋势部分指数平滑值叠加，预测结果平均绝对误差3.0084，平均相对误差0.047996，涨跌判断准确率64.645%。预测效果如图5.9所示：

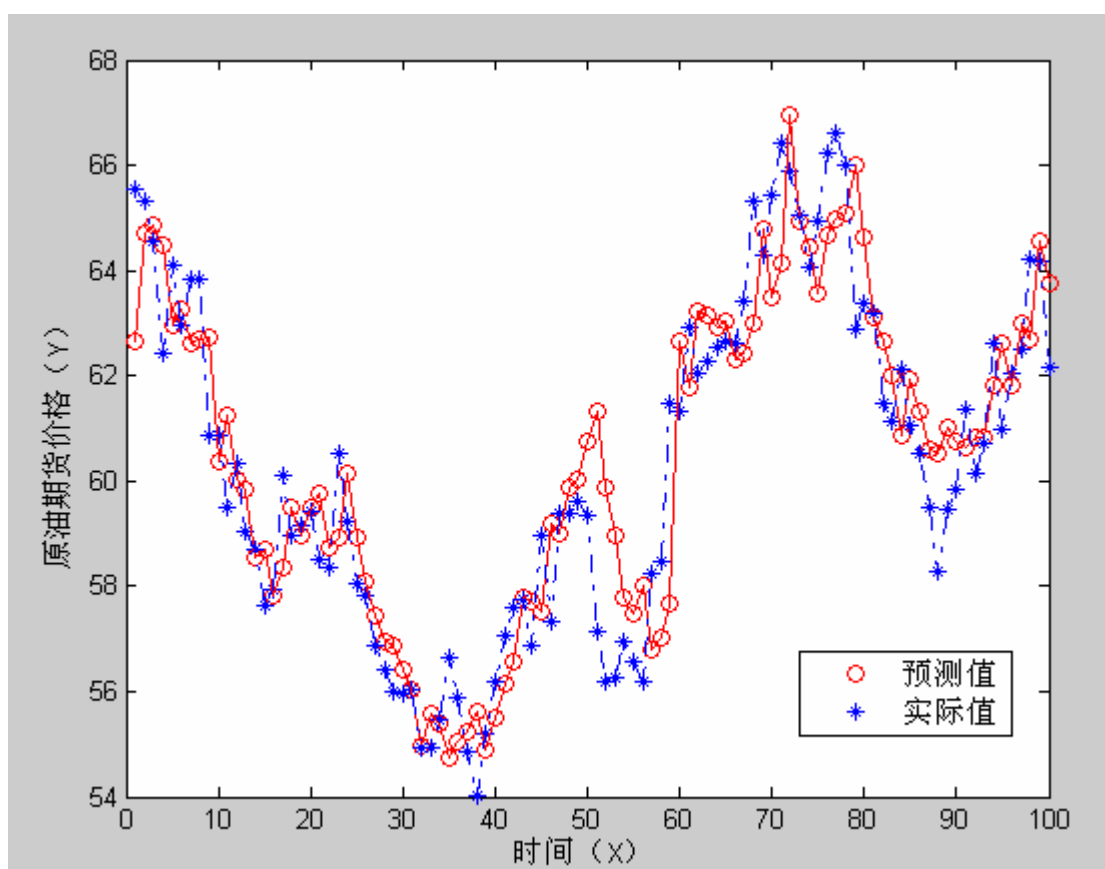


图 5.9 L_RBF_ALL 模型预测效果图

5.2.6 模型 L_rbf4

L_rbf4预测模型是一个指数平滑模型与一个径向基神经网络预测模型叠加的结构。其中的径向基神经网络输入参数为Brent原油期货开盘价、最高价、最低价、收盘价，共4个输入参数。把径向基神经网络预测模型输出与趋势部分指数平滑值叠加，预测结果平均绝对误差1.0055，平均相对误差0.016586，涨跌判断准确率53.535%。预测效果如图5.10所示：

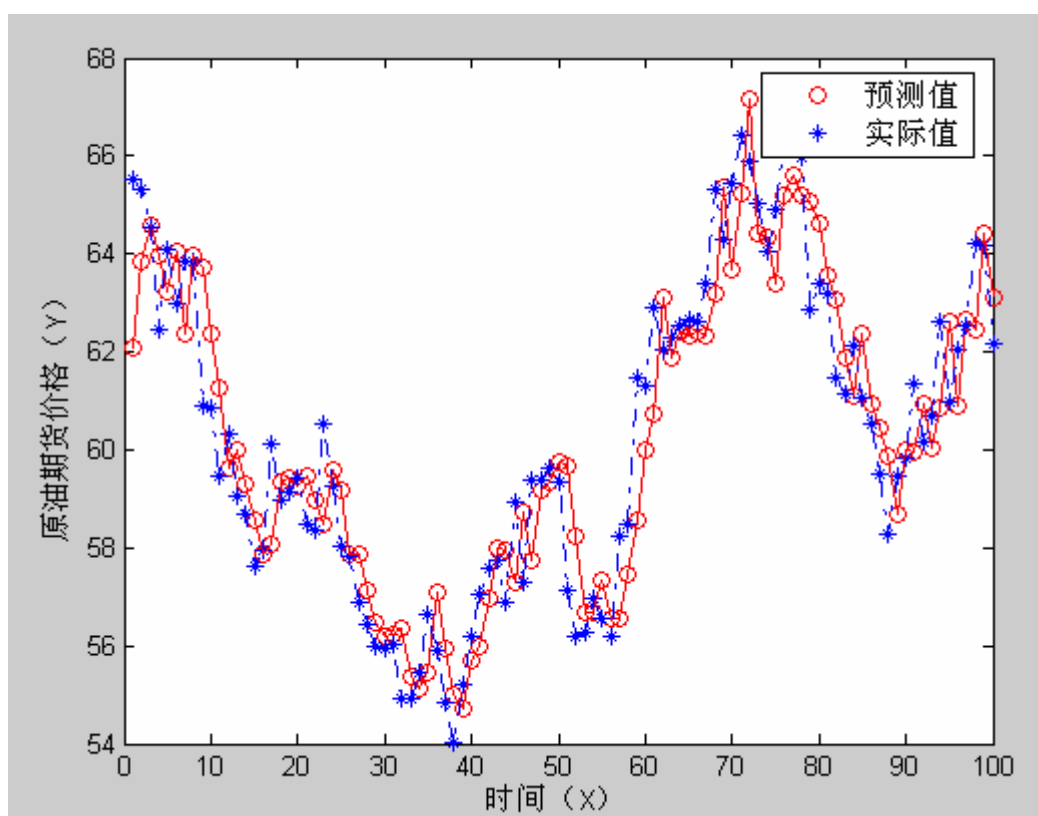


图 5.10 L_RBF4 模型预测效果图

5.2.7 模型 L_rbf5s

L_rbf5s预测模型是一个指数平滑模型与一个径向基神经网络预测模型叠加的结构。其中的径向基神经网络输入参数为前五天的收盘价，共5个输入参数。把径向基神经网络预测模型输出与趋势部分指数平滑值叠加，预测结果平均绝对误差0.94317，平均相对误差0.015625，涨跌判断准确率51.515%。预测效果如图5.11所示：

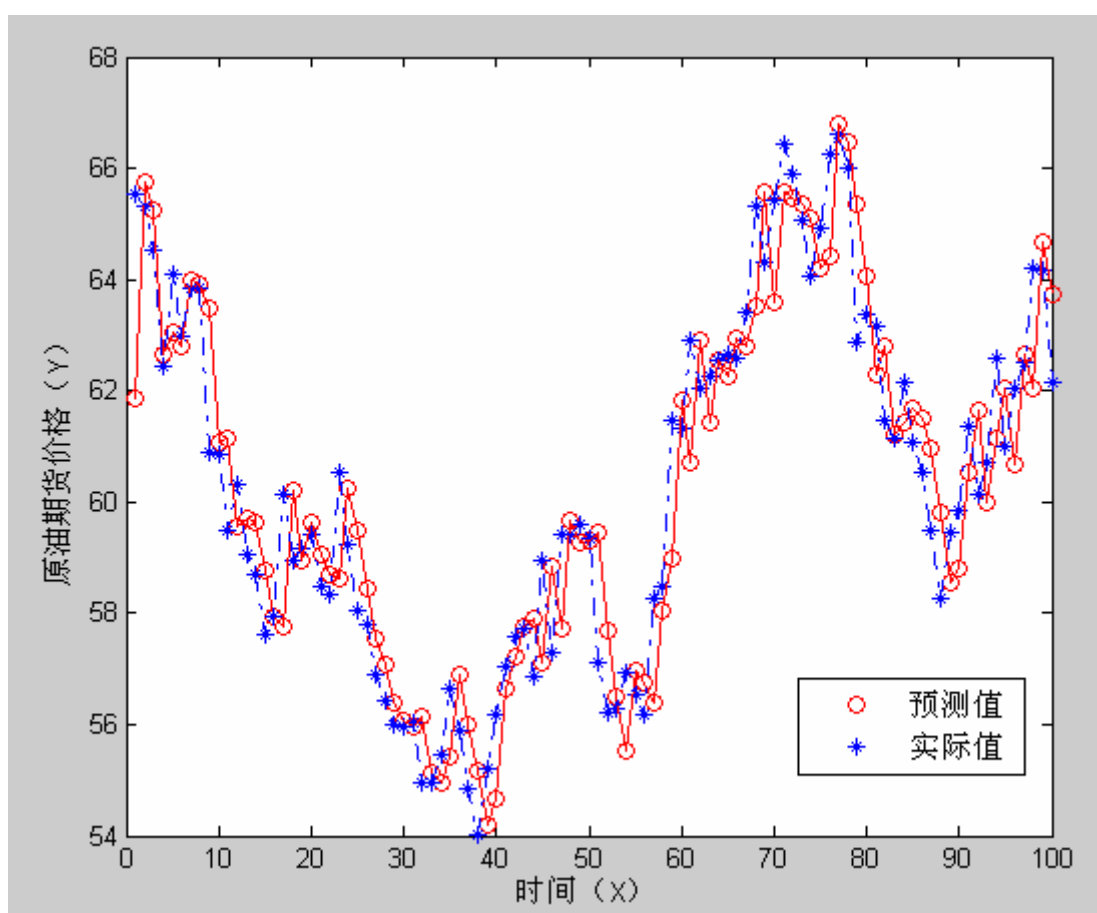


图 5.11 L_RBF5S 模型预测效果图

5.2.8 模型 L_rbf10b_3z

L_rbf10b_3z预测模型是一个指数平滑模型与一个径向基神经网络预测模型叠加的结构。其中的径向基神经网络输入参数为Brent原油期货前10天的收盘价随机波动值，以及WTI原油期货收盘价、纳斯达克指数、SP500指数、美元英镑汇率、美元欧元汇率、道琼斯指数前8、9、10天的值，共28个输入参数。把径向基神经网络预测模型输出与趋势部分指数平滑值叠加，预测结果平均绝对误差1.3345，平均相对误差0.021668，涨跌判断准确率70.707%。预测效果如图5.12所示：

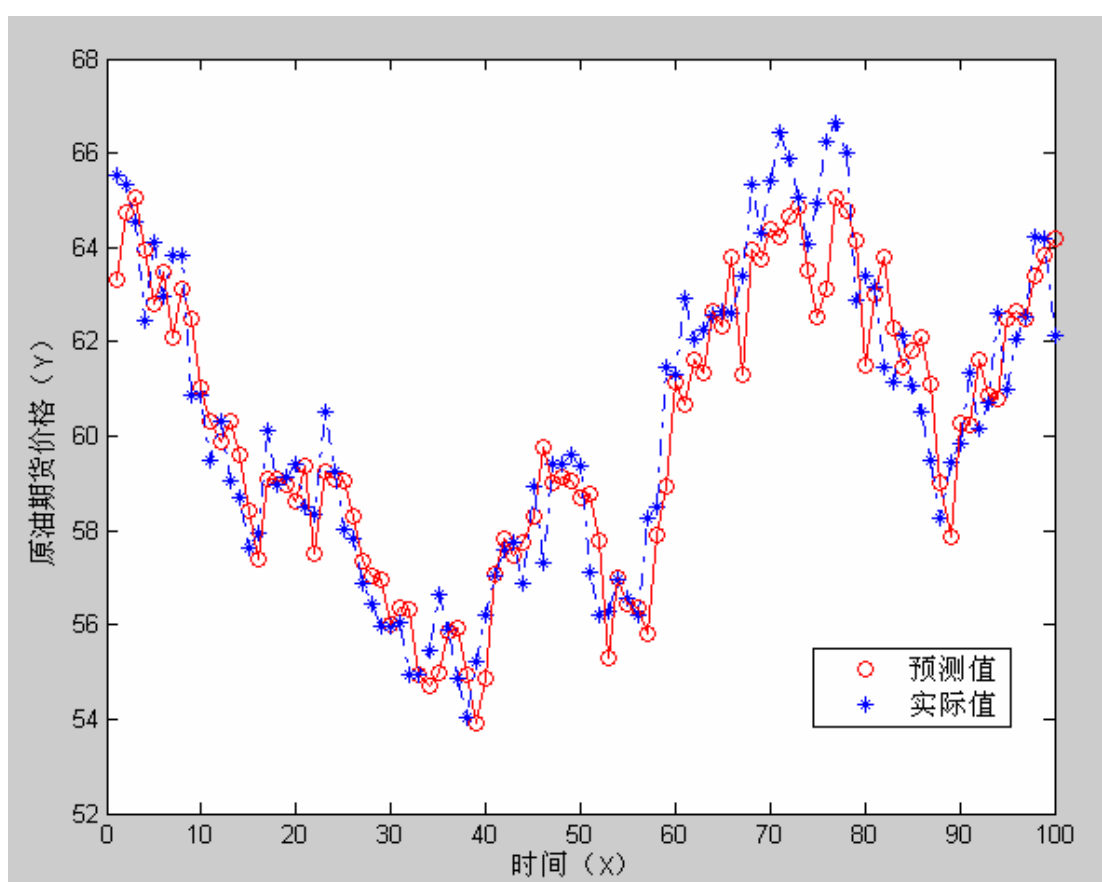


图 5.12 L_RBF10B_3z 模型预测效果图

5.2.9 单项预测模型预测性能比较

根据单项模型的选择标准，首先对各个单项预测模型的预测结果进行对比，找出其中有自身特点的、含有不同信息片断的预测模型进行组合。

表 5-2 单项预测模型性能对比

序号	模型	平均绝对误差	平均相对误差	涨跌判断正确率
1	线性回归	0.9400	1.5447%	--
2	灰色模型	3.8775	6.1656%	--
3	ARIMA	2.3137	3.8217%	--
4	L_bp_all	0.9197	1.3318%	56.57%
5	L_bp4	1.0911	1.8045%	45.46%
6	L_bp5s	0.8371	1.3820%	57.58%

7	L_bp10s_3z	0.8329	1.3775%	59.60%
8	L_rbf_all	3.0084	4.7996%	64.65%
9	L_rbf4	1.0055	1.6586%	53.54%
10	L_rbf5s	0.94317	1.5625%	51.52%
11	L_rbf10b_3z	1.3345	2.1668%	70.71%

从表5-2可以看出，用神经网络模型对Brent原油期货价格进行预测，效果明显好于线性回归、灰色预测模型和ARIMA预测模型，这是由这些模型的预测原理以及Brent原油期货价格时间序列的特点决定的。

从本文采用的神经网络模型角度看，L_bp_all与L_rbf_all、L_bp4与L_rbf4模型分别采用相同的输入参数，采用BP神经网络的L_bp_all和L_bp4模型的预测精度高于采用径向基神经网络的L_rbf_all和L_rbf4模型，但是，采用径向基神经网络的两个预测模型的涨跌判断正确率明显高于采用BP神经网络的两个预测模型。

从输入参数角度看，L_bp_all、L_rbf_all、L_rbf10b_3z等加入了纳斯达克指数、SP500指数等参数的预测模型的涨跌判断正确率明显高于输入参数只含Brent原油期货价格时间序列本身数据的预测模型，如L_bp4、L_rbf4模型，预测精度相差不大。说明了纳斯达克指数、SP500指数等等外部因素对Brent原油期货价格具有指导意义。

上述预测模型中，有些预测模型预测精度比较高，预测绝对误差小，有些预测模型对涨跌判断能力较强。例如：表5-2中L_bp_all模型预测绝对误差最小，但对涨跌判断能力不如L_rbf10b_3z模型。因此，选择多个具有不同预测效果的单项预测模型进行组合，能提高预测模型的综合预测性能。

5.3 组合预测模型

5.3.1 组合模型（一）

本文选取表5-2中预测绝对误差较小的L_bp_all、L_bp10s_3z模型与L_rbf10b_3z模型进行组合，对Brent原油期货价格随机波动部分进行组合预测，然后与指数平滑模型预测的趋势部分叠加，得到Brent原油期货价格预测值。组合算法如下：

$$y_t = y_{L_t} + (w_{1t}y_{1t} + w_{2t}y_{2t} + w_{3t}y_{3t}) \quad (5-14)$$

其中 y_t 表示Brent原油期货价格第 t 期预测值, y_{L_t} 表示第 t 期指数平滑值, w_{1t} 、 w_{2t} 和 w_{3t} 分别表示L_bp_all、L_bp10s_3z模型和L_rbf10b_3z模型第 t 期的组合权系数, y_{1t} 、 y_{2t} 和 y_{3t} 分别表示L_bp_all、L_bp10s_3z模型和L_rbf10b_3z模型第 t 期预测值。

组合模型预测结果平均绝对误差为0.73648, 平均相对误差为1.2176%, 涨跌判断准确率为62%。预测结果如图5.13所示:

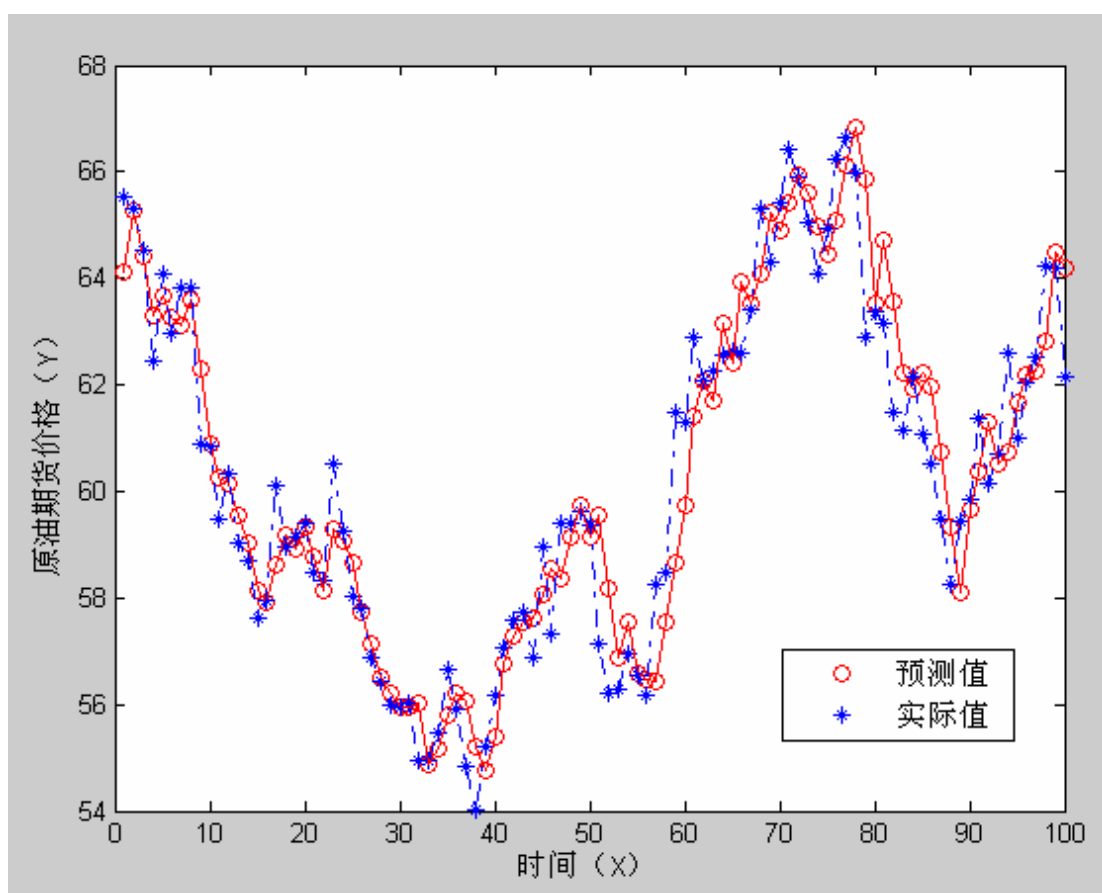


图 5.13 组合模型（一）预测效果图

从上图看出, 组合模型（一）预测精度高于其中任何一个单项预测模型, 涨跌判断准确率也比较高, 说明组合预测模型的拟合能力, 推广能力都得到提高。

5.3.2 组合模型（二）

本模型选择8个单项预测模型进行组合, 预测Brent原油期货价格随机波动部分, 然后叠加Brent原油期货价格指数平滑值。组合预测结果如图5.14所示, 平均绝对误差0.82978, 平均相对误差1.372%, 涨跌判断准确率为61%。

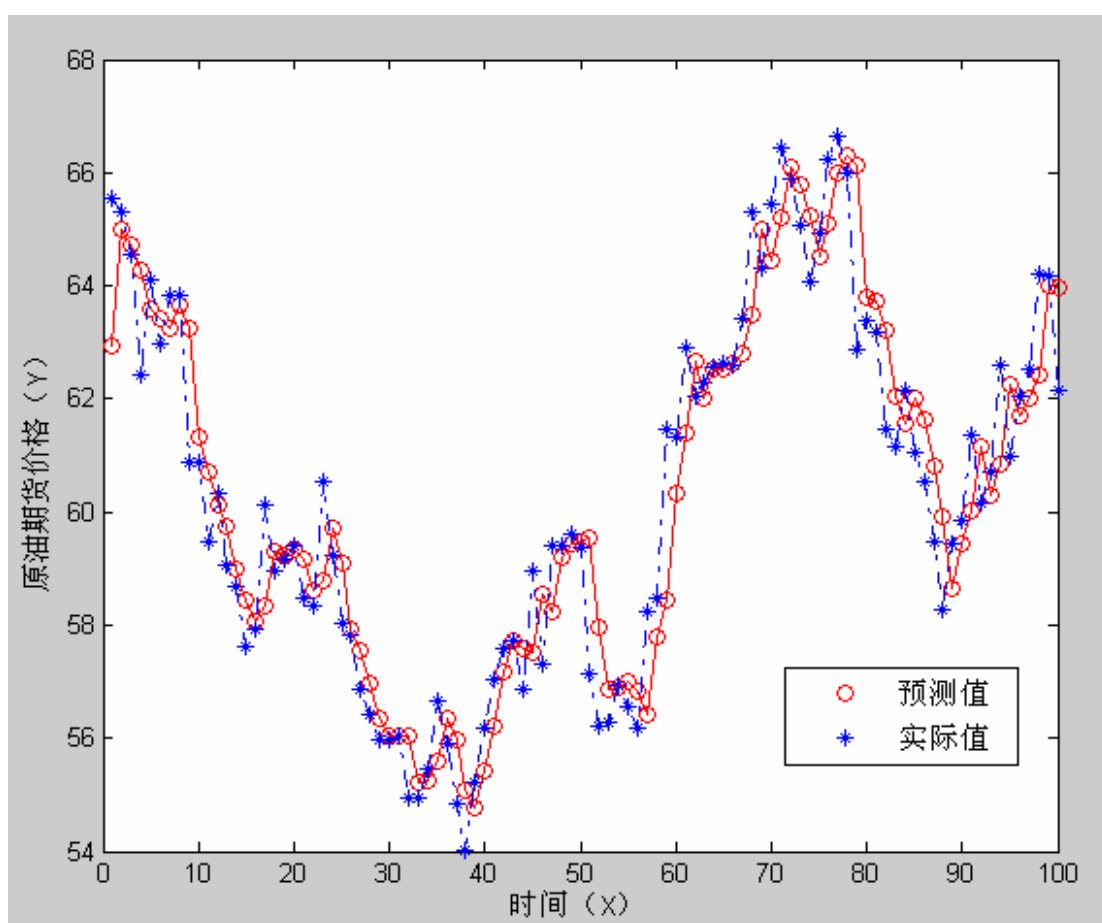


图 5.14 组合模型（二）预测效果图

从组合模型（二）预测性能看，组合预测模型精确度高于任何一个单项模型，涨跌判断准确率适中。但预测精度和涨跌判断能力都不如组合模型（一），说明组合预测中，单项预测模型并不是越多越好，在进行组合之前，选择恰当的单项预测模型是非常必要的。

5.3.3 与简单平均组合模型的比较

本文还以其他方式进行了若干次组合，组合后的模型预测精度都有提高，在此没有一一列举。把组合模型（一）、组合模型（二）中选择的单项模型，用简单平均组合方法进行组合，分别命名为M1、M2。与基于概率统计的组合预测模型的组合模型（一）、组合模型（二）进行对比，如表5-3所示：

表 5-3 组合预测模型预测结果比较

模型	平均绝对误差	平均相对误差	涨跌判断正确率
组合模型（一）	0.73648	1.2176%	62%
组合模型（二）	0.82978	1.372%	61%
M1	1.6376	2.649%	60%
M2	0.89728	1.480%	57%

从表5-3看出本文提出的基于概率统计的组合预测模型在预测精度、涨跌判断方面都优于简单平均组合预测模型。同时也发现，简单平均组合预测模型中组合的单项预测模型越多，预测精度越高，这是它与基于概率统计的组合预测模型不同的地方。

6 结论与展望

6.1 总结

在国民生产和日常生活中,经常遇到此类复杂时间序列预测问题,例如房价预测、股价预测、外汇汇率预测等。然而对复杂时间序列进行预测是一个很困难的问题,也是多年来研究的热点。时间序列预测技术很多,不能说哪一种方法好哪一种坏,而是要结合到实际问题中来,根据具体问题,选择恰当的预测方法。例如:在对企业某产品产量进行预测的时候,用传统的统计学方法也许更加简单实用。在对一些复杂非线性时间序列进行预测时,传统的统计学方法很难描述其随即波动的特征,而需要采用非线性的预测模型。

多元线性回归模型的目的是要找到自变量与因变量之间的某种函数关系,用自变量的值来估计和预测因变量的值。然而,Brent原油期货价格这类复杂的非线性时间序列,很难用一个函数来描述它与其他变量之间的关系。因此,用多元线性回归模型来预测Brent原油期货价格,效果不理想。

本文建立了灰色预测模型,预测结果并不如文献[4]中所说那样,能准确地预测时间序列的随机波动,只能反映时间序列的总体增长趋势,而不能反映其波动变化的具体特征。单一的灰色模型往往无法解决时间序列的复杂非线性关系,导致预测精度不高。

传统的ARIMA模型在对平稳时间序列进行预测,可以取得良好的效果。但是该模型只考虑时间序列内部关系,没有考虑其他含有指导信息的因素,所以它最多能外推时间序列的趋势,不能利用其他市场信息对时间序列的波动进行判断,因此本文建立的ARIMA预测模型预测效果欠佳。

神经网络预测模型可以描述复杂的非线性关系,并且可以输入外部影响因素,这是神经网络模型的优势所在。神经网络输入参数中加入WTI原油期货收盘价、纳斯达克指数、SP500指数、美元英镑汇率、美元欧元汇率、道琼斯指数等经济指标,能提高预测模型的预测能力。

但BP神经网络容易陷入局部最小，导致每次运行程序，预测结果不一定相同，预测稳定性不高。径向基网络收敛快、但可能出现过度拟合问题，导致模型泛化能力不强，因此在建模时可以通过调节训练最大次数和目标误差来解决这一问题。

每一个神经网络预测模型隐含了不同的信息片断，因此每个预测模型预测性能不尽相同。组合预测模型的提出，使这个问题迎刃而解。组合预测模型综合了单项预测模型的特点，最突出的优点是削弱了单项预测模型“时好时坏”的影响，提高了预测稳定性。

本文在对Brent原油期货收盘价时间序列深入分析的基础上，采取把趋势部分与随机波动部分相分离的办法，用指数平滑方法来拟合基本趋势，用多个神经网络模型来预测随机波动，用基于概率统计的组合算法进行组合，然后把两部分叠加作为预测值。实证分析证明，该方法是行之有效的。

6.2 进一步研究

(1) 由于数据问题，本文只考虑了WTI原油期货收盘价、纳斯达克指数、SP500指数、美元英镑汇率、美元欧元汇率、道琼斯指数等外部因素，将来还要考虑引入其它的含有更多指导信息的外部因素。

(2) 世界重大政治、经济、军事事件对原油期货价格有重大影响，但是很难界定各个事件对原油期货价格影响的程度，因此本文假设这些事件产生的影响都包含在市场价格里，这个假设是否合理，还需要进一步证实。

(3) 本文只研究了Brent原油期货下一天价格的预测方法，这些方法是否同样适用于多天、周、旬时间间隔预测，还有待研究。

致谢

在两年的学习和生活过程中，很多老师、同学和朋友给了我很多帮助和鼓励。

首先，我首先要衷心地感谢我的导师朱东华教授。从本论文的选题、研究、到论文写作都倾注了朱老师的心血。朱老师广博的学识、高瞻远瞩的学术视野、深邃的洞察力、严谨求实的治学态度、勤奋忘我的精神、诲人不倦的教育情怀和对事业的忠诚给我树立了榜样。朱老师给我提供了良好实验条件，使我能顺利进行各项研究工作。在生活上，朱老师也对我无微不至的关怀。在此向朱老师表示崇高的敬意和真诚的感谢。

感谢祝金荣师兄在技术实现上对我的指导，以及论文撰写全过程中都给予了我无私的关怀和帮助，本人深表感激。

感谢班上的许多同学，在论文写作期间给我提出的宝贵意见和建议。

本论文的研究过程参考了大量中英文文献，在此一并向所有给我以知识和启迪的作者们表示感谢。

攻读硕士期间发表的学术论文

1. 杨熙亮, 朱东华. 神经网络技术在国际原油期货价格预测中的应用. 北京理工大学学报(增刊). 2006.6。
2. 杨熙亮, 朱东华, 余丰. 我国物流规模及发展状况分析. 科技经济市场. 2006.5

参考文献

- [1]周瑞平.GM(1,1)模型灰色预测法预测城市人口规模[J].内蒙古师范大学学报.2005.3
- [2]王成山,杨,军,张崇见.灰色系统理论在城市年用电量预测中的应用[J].电网技术.1999.2
- [3]赵春雷,朱艳,李明等.灰色证券系统证券价格趋势模型的建立与分析[J].河北工业科技.2004.5
- [4]王凤兰, 闻邦椿.股价波动序列的综合预测方法研究[J].经济经纬.2005.2
- [5]耿继进.灰色预测理论若干问题研究[J].武汉测绘科技大学学报.1994.3
- [6]范正绮, 王祥云.ARIMA模型在汇率时间数列预测中的应用[J].金融市场.1997.3
- [7]朱立斌.ARIMA模型在股市预测中的应用[J].江苏统计.1999.1
- [8]刘新勇, 贺江峰, 孟祥泽等.基于神经网络的股市预测[J].南开大学学报.1999.9
- [9]王玉荣.ARIMA模型在我国出口贸易预测中的应用[J].决策参考.2004.4
- [10]金席卷,李欣然,唐福顺.基于模糊理论及神经网络的电力短期负荷预测方法[J].电力信息化.2005.2
- [11]佚名.期货市场的功能与结构[J].中国石油期货评论.2004.6
- [12]佚名.世界主要石油期货市场--NYMEX[J].中国石油期货评论.2005.3
- [13]李慕春.世界石油期货市场研究.大连期货交易所石油期货系列研究报告.2003.6
- [14]陈秀芝.当前石油市场形势分析及基本走势预测[J].中国能源.2004.11
- [15]韩冬炎,陈蕊,崔立瑶.对石油价格走势预测的数理研究[J].价格理论与实践.
- [16]董秀成, 曹文红.国际油价预测中的各种不确定因素分析[J].石油大学学报.2000.4
- [17]邓聚龙.灰色系统预测与决策[M].武汉:华中理工大学出版社,1986.
- [18]王卓,王艳辉,贾利民,李平.改进的BP神经网络在铁路客运量时间序列预测中的应用[J].中国铁道科, 2005.3
- [19]刘豹,胡代平.神经网络在预测中的一些应用研究[J].系统工程学报.1999.12
- [20]王悦.人工神经网络在经济预测中的应用[J].北京广播电视大学学报.2002.1
- [21]向小东.BP算法的几点改进[J].预测.2000,19(7)
- [22]孙见荆.组合预测参数估计方法研究[J].预测,1996,(6):54~56.
- [23]王景,刘良栋等.组合预测方法的现状与发展[J].预测,1997,(16):37~38.
- [24]冯春山,吴家春,蒋馥.石油价格的组合预测研究[J].石油大学学报.2004.2

- [25]冯旭东,陈方.基于遗传技术辅助设计的神经网络期货市场预测[J].计算机工程与应用.1998.11
- [26]钟颖,汪秉文.基于遗传算法的BP神经网络时间序列预测模型[J].系统工程与电子技术.2002
- [27]胡丹,莫智文.粗糙集理论在模糊神经网络中的应用研究[J].四川师范大学学报.2002.1
- [28]唐纪,王景.组合预测方法评述[J].预测.1997.6
- [29]BatesJM,GrangerCJ. Combination of Forecasts[J]. Operations Research Quarterly,1969,20:451~468.
- [30]Ajay Kumar Palit, D.Popovic.Nonlinear Combination of Forecasts Using Artificial Neural Network[J]. IEEE,2001.
- [31]YULean,WangShouYang.TimeSeriesForecastingWithMultipleCandidateModels:Selectingor Combining?[J].JournalofScienceandComplexity.2005.2
- [32]Azmi Omar, ShamsulMajid.Improving the Price Forecast of Crude Palm Oil Futures Using Historical Return Variances[J]. Oil Palm Industry EconomicJournal.2004.2
- [33]Saeed Moshiri,Faezeh Foroutan.Testing for deterministic chaos in futures crude oil price;Does neural network lead to better forecast?.2004.5
- [34]Cai Chang feng.Predicting Tendency of Futures by Gray Model GM(1,X1) [J].Mathematic aApplication.1999
- [35]Benjamin Tal. Background Information on our Neural Network - Based System of Leading Indicators[J].Economics&Strategy.2003.9
- [36]AlexanderV.Prasolov.On Forecast of Exchange Rate of a Foreign Currency[J]. IEEE.2000
- [37]MartinRast. Fuzzy Neural Networks for Modelling Commodity Markets[J].IEEE.2001
- [38]Greg Tkacz ,Sarah Hu. Forecasting GDP Growth Using Artificial Neural Networks. Proceedings of a conference held at the Bank of Canada.1993.3
- [39]Chakradhara Panda, V.narasimhan. Forecasting daily foreign exchange rate in India with Artificial Neural Network[J].The Singapore Economic Review.2003.2
- [40] V. I. Yukalov.Self-similar extrapolation of asymptotic series and forecasting for time series[J]. Modern Physics Letters.2000(14)
- [41]周志华, 曹存根.神经网络及其应用[M]. 清华大学出版社.2004
- [42]唐小我,曹长修,金德运.组合预测最优加权系数向量的进一步研究[J].预测.1994.2