



暨南大学
JINAN UNIVERSITY

项目学术报告

基于我国粮食经济发展战略背景下的小麦价格 ARIMA-LSTM-
XGBoost 组合预测模型研究

学生姓名：苏晓钰

指导老师：林少萍

职称：讲师

暨南大学

2023 年 3 月

摘要

国家粮食价格不仅是调节粮食产品供需关系的信号和手段，而且也与民众生活成本和农民收入水平密切相关。作为三大主粮之一的中国小麦，年产量超过 1 亿吨，占据世界市场近 25%，处于世界粮仓的重要战略地位。然而，自 2018 年起，中美贸易摩擦不断，给中国小麦进口带来更多不确定性；2020 年新冠疫情袭来，国内农产品生产受阻，小麦价格一度上涨；而国际疫情反复不断引起的市场波动，加剧了国际市场价格波动的风险。目前国内研究范围和方法存在局限性，大多仅使用传统的时间序列分析进行预测，并且很少将国际动荡事件纳入影响因素，结论过时。在当前国家粮食经济发展战略背景下，构建更精准的小麦期货价格预测模型有利于更有效地反应小麦市场，保障小麦最低收购价政策顺利实施。

本项目基于《中国农产品价格调查年鉴》、《中国统计年鉴》与国家粮油信息中心的公开数据，从小麦国际和国内市场角度出发，探求影响小麦价格的重要因素，旨在对小麦发展趋势做出准确的预测。首先根据训练集数据，基于随机森林特征重要性排序构建了全面的小麦期货价格的影响因素指标体系，并构建出传统的时间序列分析模型与单变量的 LSTM 模型与单变量 XGBoost 模型。为了进一步构建更准确的小麦期货价格预测模型，我们基于所建立的影响因素指标体系，构建 LSTM 模型与 XGBoost 模型。最后考虑将构建的 ARIMA、LSTM 和 XGBoost 三种模型集成一种多变量预测组合模型：将组合模型的预测值作为支持向量机回归 SVR 模型的输入，并采用贝叶斯优化算法对 SVR 模型进行调参，解决影响因素未来值获取问题并显著提高预测精度。综合组合模型与单一模型的预测结果，从绝对误差、相对误差、平均相对误差等指标比较模型，并选择出预测效果最好的模型。基于结果提出相应政策建议，有助于及时调节市场供求关系，稳定粮食经济；并进一步推广到农产品价格预测领域，对我国农业经济发展具有重大意义。

关键词： 时间序列预测；小麦价格；ARIMA；LSTM；XGBoost；农业经济

1 选题背景

1.1 国家战略

中国是一个粮食生产和消费大国，粮食安全问题对国家战略至关重要。粮食价格是一个重要的关注点。党的十九大报告强调了确保国家粮食安全的重要性。中央一号文件连续十二年关注“三农问题”，并提出发展农产品期货交易。小麦作为三大主粮之一，在中国的年产量和消费量都居世界前列。自 2006 年起，我国实行小麦最低收购价政策，以稳定主粮价格、改善供求关系，保障农民利益。在近十年的国家政策扶持下，国产小麦产量提升、粮食库存充足，有效保障市场需求，小麦现货价格稳步上涨。

1.2 国内外市场

中国小麦年总产量超过 1 亿吨，自给率达到 95%以上，在国民经济中占有重要地位。2021 年，小麦总产量达到 13694.6 万吨，增长率为 2%。随着人口增长和经济发展，小麦需求量逐年上涨。得益于临储和“托市”政策，我国小麦价格从 2009 年 1 月的 1790 元/吨上涨至 2013 年 9 月的 2519.29 元/吨。2020 年以来，粮食稳定成为国家安全战略部署的重要手段，国内小麦减产、储备性需求激增导致价格明显上涨。国内中筋小麦品种占比超过 55%，是国产小麦的

重要来源。进口成为补充我国优质小麦缺口的重要途径。

小麦是国际农产品贸易中最活跃、用途最广的农产品之一，其种植分布、消费及总贸易额占世界粮食作物第一。自 2001 年起，我国逐步从小麦净出口国转变为净进口国，主要进口优质小麦和饲料用小麦以调剂粮食结构。2021 年，小麦进口量达到 977 万吨，创新高。长期以来，美国、加拿大和澳大利亚进口小麦占我国进口量的八成以上，依存度高。随着一带一路倡议的提出，中国与俄罗斯、哈萨克斯坦的小麦贸易关系逐步建立，寻求多元化的进口途径，降低对少数国家的粮食依存度。

1.3 其他影响

2020 新冠肺炎疫情袭来，粮食安全成为关注焦点。疫情爆发初期，国内受疫情影响，交通封锁、停工停产等使粮食运输受限导致原料短缺、面粉生产加工停摆。随着疫情转移到国外，2020 年 7 月以后，国际谷物价格进入了上升通道。受疫情影响，各国封锁下进出口贸易受阻，外部供给环境恶化；谷物出口大国担忧粮食供应，相继限制粮食出口。由于国内小麦现价与国际价格紧密关联，国际疫情引起的市场波动往往对我国小麦生产销售市场产生影响。

2 方案论证

2.1 项目简介

本项目基于《中国农产品价格调查年鉴》中 2009 年 1 月 4 日至 2021 年 12 月 17 日共 3255 个样本的我国小麦期货价格序列日数据，以及《中国统计年鉴》和国家粮油信息中心关于农业经济投入、人力投入、我国财政收支、小麦对外贸易、小麦国际进出口量等多个影响因素，结合中国人民银行、国际清算银行、国际货币基金组织关于国家经济政策、国内外经济环境、现货基本条件等信息，建立数学模型预测小麦期货价格。

2.2 理论依据

2.2.1 ARIMA 模型

整合自回归滑动平均模型 ARIMA(Autoregressive Integrated Moving Average)是自回归和滑动平均模型的综合，其中心思想是运用数学模型近似描述一个随机的时间序列，可以处理非平稳性的、具有线性趋势和季节、循环成分的时间序列。ARIMA 模型的建模预测主要包括以下步骤：

1. 通过时序图、ACF 和 PACF 图结合来判断序列是否平稳，同时通过 ADF 单位根检验法构造 ADF 统计量，进行平稳性检验。
2. 当序列不平稳时，为消除显著的线性趋势和季节趋势的影响，我们对时间序列进行平稳化处理，如差分法、数据的分解与平滑、对数化等。
3. 通过自相关系数图(Autocorrelation Function, ACF)和偏自相关系数图(Partial Autocorrelation Function, PACF)定阶法，观察两图的截尾和拖尾情况，初步确定 p(P) 和 q(Q) 的阶数，并结合 AIC 信息准则找到适合模型的最佳阶数 P, Q 的值。AIC 准则是由 K-L 信息量推导而来，其一般形式为：

$$AIC = -2\ln(L) + 2m$$

L 是极大似然函数，m 是待估参数个数，AIC 值越小代表模型效果越好

表 1 ARIMA 模型定阶原则

模型定阶	AR(p)	MR(q)	ARMA(p,q)
自相关系数($\hat{\rho}(k)$)	拖尾	P 阶截尾	拖尾
偏自相关系数(ϕ_{kk})	Q 阶截尾	拖尾	拖尾

- 考虑到小麦期货价格序列的季节效应、长期趋势效应和随机波动之间存在复杂的交互影响关系，简单的 ARIMA 模型不足以提取其中的相关关系，因此我们采用乘积季节模型对模型进行优化，根据最小 AIC 与似然比统计量确定模型。
- 通过对残差序列的随机性检验与正态性检验评价模型是否拟合恰当，本项目将采取观察残差标准差图、残差序列 Q-Q 图、Ljung-Box 法来检验模型的显著性。Ljung-Box 法白噪声检验的假设条件为：

原假设 $H_0: \rho(1) = \rho(2) = \dots = \rho(m) = 0, \forall m \geq 1$

备择假设 H_1 : 至少存在某个 $\rho(k) \neq 0, \forall m \geq 1, k \leq m$

检验统计量 Q_{LB} 取为: $Q_{LB} = n(n+2) \sum_{k=1}^m \left(\frac{\hat{\rho}^2(k)}{n-k} \right) \sim \chi^2(m), \forall m > 0$

综上所述，本研究建立的 ARIMA 模型的应用步骤主要包括平稳性判别、模型定阶与识别、参数估计、模型的优化选择、模型的显著性检验、模型预测六个步骤。

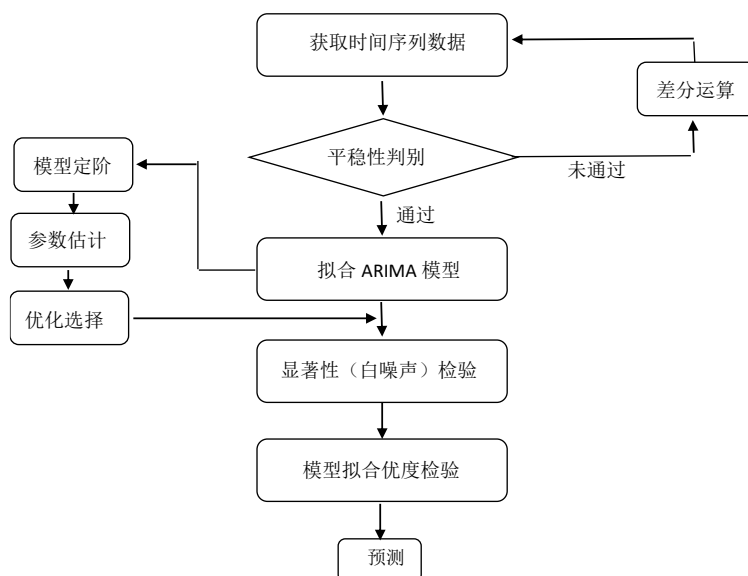


图 1 ARIMA 模型构建流程图

ARIMA 模型能较准确地预测小麦现货价格的时间序列整体趋势，但对局部拟合效果不同，且部分拟合精度较低。该模型在外推时不方便加入外生变量，未考虑结构变动带来的影响，因此仅用 ARIMA 模型可能导致未来的预测效果欠佳。

2.2.2 LSTM 神经网络

引入 LSTM 模型可以重新设置 RNN 的隐藏层模块结构，实现了对信息的长期记忆。LSTM 网络主要是通过引入门结构：遗忘门（forget gates）、输入门（input gates）和输出门（output gates），对信息进行保护和控制。

遗忘门是对于是否保留或丢弃上一时刻的 s_{t-1} 、 c_{t-1} 到当前时刻的状态 c_t 做出抉择。通过接收上一个状态的输出值 h_{t-1} 和当前时刻的输入 x_t ，使用 sigmoid 激活函数得到 0-1 之间的值，其中 1 表示全保留，0 表示全忘记，具体计算公式为：

$$F_t = \sigma(X_t W_{ih}^F + H_{t-1} W_{hh}^F + b_{ih}^F)$$

输入门的作用是判断 c_t 要留下多少当前的输入信息 x_t ，共同更新当前时刻的单元状态。首先第一层由 sigmoid 层判断需要更新的输入，即

$$I_t = \sigma(X_t W_{ih}^I + H_{t-1} W_{hh}^I + b_{ih}^I)$$

第二层由 tanh 层创建可以更新到当前状态的候选信息，即

$$\tilde{C}_t = \tanh(X_t W_{ih}^{\tilde{C}} + H_{t-1} W_{hh}^{\tilde{C}} + b_{ih}^{\tilde{C}})$$

根据遗忘门和输出门两个门结构的计算可以得到记忆细胞 c_t 的更新值：

$$C_t = F \odot C_{t-1} + I_t \odot \tilde{C}_t$$

输出门的功能是决定当前时刻 c_t 中可以输出到 s_t 的部分，由通过 sigmoid 层得到的中间变量 o_t 与经过 tanh 层处理后得到 C_t 决定：

$$O_t = \sigma(X_t W_{ih}^O + H_{t-1} W_{hh}^O + b_{ih}^O)$$

$$s_t = o_t \odot \tanh(C_t)$$

对得到的训练集预测值进行反归一化可以得到 2009.07-2021.07 年真正的预测数据，并绘制相应的拟合曲线。

2.2.3 XGBoost 模型

极端梯度提升 XGBoost(Extreme Gradient Boosting)模型是一种分布式高效梯度提升算法，可解决分类、回归等机器学习问题。它是一种集成模型，由多棵 CART 决策树作为子模型构成，采用 Gradient Tree Boosting 算法进行集成学习，最终结果为全部决策树的预测值累加。

2.3 研究意义

2.3.1 调节市场供求关系

小麦价格受供求关系影响，又反作用于供求关系。我国粮食期货市场日臻成熟，小麦期货价格成为市场价格“风向标”。通过有效价格预测，减少卖方与市场的信息不对称，提高整条小麦供应链效率、保障生产者利益，满足市场需求，构建健康的主粮市场。

2.3.2 规范国内小麦定价

我国小麦生产成本低，国家通过小麦最低收购价补贴、临储麦政策“托市”，维持了小麦供应量，从而保障消费需求。对小麦价格的有效预测，是国家加快推进小麦和稻谷定价机制、进行补贴政策和收储体制改革的重要依据；同时，帮助人们有效应对国际粮价波动带来的风险，促进国产小麦价格与国际化、标准化接轨，提升我国小麦产业国际影响力。

2.3.3 保障国家粮食安全

准确预测小麦现货价格有利于降低外部环境对于粮食供应、粮食价格的冲击，优化粮食储备。在特殊时期全球资本涌入市场避险、投机。掌握未来价格趋势的对于小麦发挥粮食作用、国家粮食供应稳定具有重要意义。

3 研究方法

3.1 构成小麦期货价格的影响因素指标体系

计算 Pearson 相关系数度量变量之间先行相关关系，变量之间的非线性相关程度通过随机森林特征重要性排序，最终筛选出少数具有影响力的因素，构成预测小麦期货价格的指标体系。

3.2 基于所构建的小麦期货价格的预测指标体系，建立模型进行预测

首先通过传统的 ARIMA 模型建立经典的单变量时间序列模型；为了提升模型预测效果，根据由随机森林提取的特征构建影响因素指标体系，并使用 LSTM 和 XGBoost 神经网络模型进行预测；然后将小麦价格序列分解为线性趋势和非线性趋势，构建 ARIMA-LSTM-XGBoost 组合预测模型对结果进行预测。

3.3 评价单一预测模型与组合模型的预测效果，并进行最优选择

通过计算得到测试集的误差、相对误差以及平均相对误差，并比较单一模型（ARIMA、LSTM、XGBoost）及组合模型的指标，最终选出最适合的预测模型，并应用于我国小麦市场发展趋势预测。

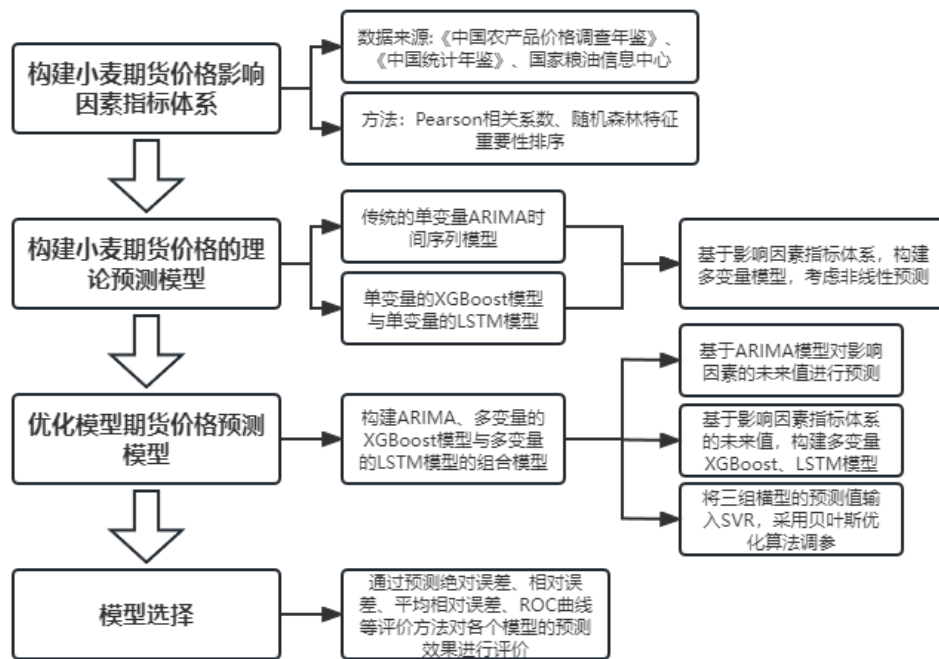


图2 小麦期货价格预测模型技术路线图

4 研究结果

4.1 ARIMA 时间序列模型小麦价格预测

本次研究将 2009 年 1 月至 2021 年 4 月作为训练样本（共 148 个月样本），2021 年 5 月至 2021 年 9 月作为测试样本（共 5 个月样本）。首先将通过绘制时间序列图观察序列的平稳性。

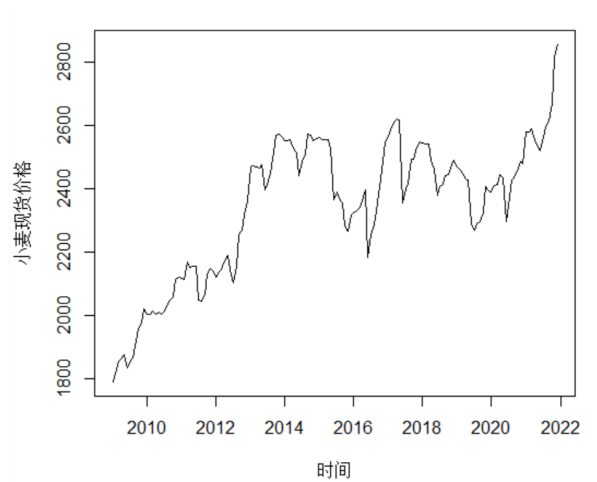


图3 小麦现货价格原始数据时序图

分析表明，小麦现货价格呈现逐年递增的趋势，并具有明显的季节性趋势，表现为12个月为一个周期。这表明该时间序列是非平稳的。由于存在季节效应，选择模型时应考虑其影响，排除非季节性的 $ARMA(p, q)$ 模型。通过运用单位根检验法对小麦现货价格时序的平稳性进行验证，结果表明，在显著性水平为5%下，ADF 检验的卡方统计量为-2.3799，且 P-value 值为0.4182（大于0.05），因此判定原序列存在单位根，为非平稳序列，需要进行差分处理。

表2 小麦现货价格时间序列的单位根检验

	Dickey-Fuller	P-value
ADF test statistic	-2.3799	0.4182

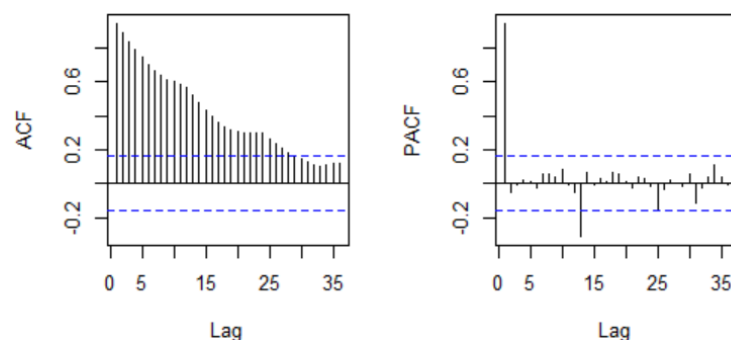


图4 小麦现货价格原始数据自相关图与偏自相关图

通过观察小麦现货价格原始数据的自相关图和偏自相关图，我们发现 ACF 图显示指数衰减趋势，即拖尾现象，且衰减速度缓慢，仍有大量值大于0.5。这表明序列 y_t 与之前的时刻具有强烈的相关性，且相关项范围较大，可判断为非平稳时间序列。而 PACF 图呈现一阶截断，显示为截尾，因此运用 $ARIMA(p, d, q)$ 模型建模。

为了消除线性趋势效应，我们对原时序数据进行差分处理，取每一项与前一项的差。一阶差分后的时序图在0附近波动，但仍存在较大的波动起伏。ACF 图的滞后12期、滞后24期、滞后36期处显示序列存在强自相关，推测存在一个12阶的周期性影响因素。为了消除趋势效应，我们再次进行差分，并通过时序图、ACF 与 PACF 图检验和单位根检验来判断差分后的序列是否平稳。

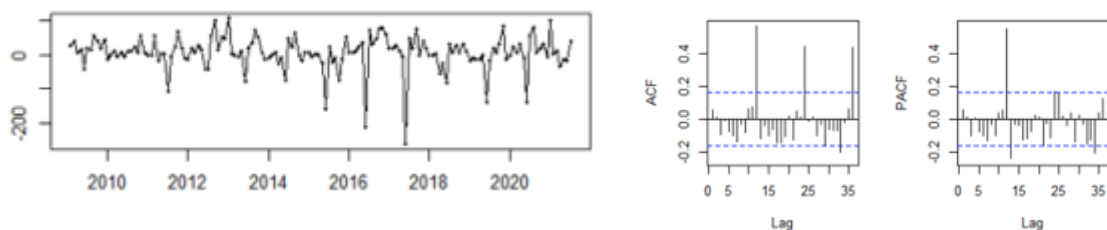


图5 一阶差分和季节差分后的小麦现货价格时序图、ACF 图和 PACF 图

表3 季节差分后的小麦现货价格单位根检验

	Dickey-Fuller	P-value
ADF test statistic	-5.6747	0.01

从上图可见一阶差分后的时序图在 0 附近波动，2015 年后的差分序列存在较为显著的波动起伏大的情况，由 ADF 检验可以看出，在 5%的水平下，卡方统计量值为-5.6747，且 P 值大于 0.05，因此可以拒绝原假设，认为经过一阶差分的序列是平稳序列，符合 ARMA 时间序列模型的建立。由 ACF 图可以看出，在已经消除了单位根的情况下，滞后 12 期、滞后 24 期、滞后 36 期处仍然存在序列强自相关，因此考虑建立 ARIMA(12,1,0)。在模型的优化选择与参数估计方面，根据最小信息量准则可以得到最优的模型。

表4 ARIMA 模型检验

	似然比统计量	AIC
ARIMA((1,2,12),1,0)	-754.35	1516.7
ARIMA((1,2,3,12),1,0)	-754.18	1518.35
ARIMA((1,2,6,12),1,0)	-754.28	1518.55
ARIMA((1,2,3,12),1,0)	-750.95	1511.91

根据简练原则，选取最小 AIC 与似然比统计量对应的模型，可得 ARIMA((1,2,12),1,0)是最优的模型，以下是该模型的参数估计,可见系数在 1%的显著性水平下显著，因此模型参数是有意义的。

表5 ARIMA((1,2,12),1,0)模型系数表

	ar1	ar2	ar12
Coeff	0.0224	-0.0111	0.5916
标准差	0.0650	0.0647	0.0685

由系数表可得最后估计出来的模型为：

$$y_t = 0.0224y_{t-1} - 0.0111y_{t-2} + 0.5914y_{t-12} + \varepsilon_t$$

ARIMA 模型的显著性（白噪声）检验

为检验模型的有效性，采取观察残差标准差图、Ljung-Box 法来检验模型的显著性。若上述各种检验方法不能拒绝原假设，则可以认为 ARIMA((1,2,12),1,0)模型的残差序列是白噪声序列，该模型有意义。

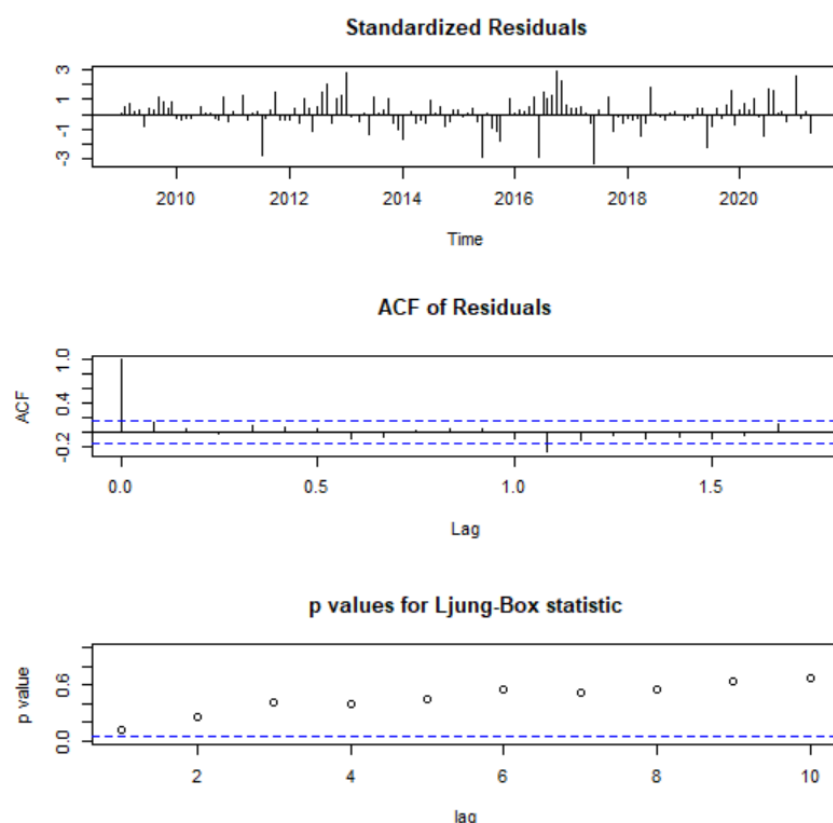


图6 残差标准差、ACF图、Ljung-Box 检验图

表6 Ljung-Box 检验

	X-squared	p-value
model residuals	8.9772	0.7049

从标准化后所得到的残差图可见残余误差都在零均值附近范围内波动，具有均匀的方差，没有明显的异常值。残差标准差的自回归函数都为近似0（两虚线内），Ljung-Box 检验图中随着滞后阶数的增大，p 值增加并总体在 0.5 上方，因此可以判断残差序列大致是白噪声序列。LB 检验的 p 值为 0.7049（大于 0.05），不能拒绝原假设 H_0 ，可以认为该序列为白噪声序列，拟合模型 $ARIMA((1,2,12),1,0)$ 为显著有效的模型。

ARIMA 模型的预测效果

利用训练好的模型对未来我国小麦的月度价格进行预测，得到 2021 年 5 月到 2021 年 9 月的月度价格预测值。

表7 2021 年 5 月到 2021 年 9 月价格预测值

月份	真实值	预测值	误差	相对误差	平均相对误差
2021-05	2540.00	2549.449	9.45	0.37%	
2021-06	2522.22	2466.864	-55.36	-2.19%	
2021-07	2561.67	2496.633	-65.04	-2.54%	1.89%
2021-08	2595.56	2545.214	-50.35	-1.94%	
2021-09	2612.78	2549.592	-63.19	-2.42%	

运用均方根误差（RMSE）与平均绝对百分比误差（MAPE）两个模型预测的评价指标对小麦现货价格的预测结果进行评估。该 $AIRIMA$ 模型的 RMSE 较小，平均绝对百分比误差 MAPE 仅为 1.0347%，结果显示该模型拟合效果较好。

表 8 ARIMA 模型检验

评价标准	结果
MAE	28.3136
RMSE	40.0983
MAPE	1.2230

根据以上 ARIMA 价格预测图以及预测结果可以看出，ARIMA 模型能较准确地预测小麦现货价格的时间序列整体趋势。



图 7 ARIMA 模型预测图

4.2 LSTM 模型预测结果

均方根误差可以衡量模型的预测性能。对于 LSTM 网络模型，训练集的均方根误差为 45.5，绘制训练集的拟合情况如下图所示：

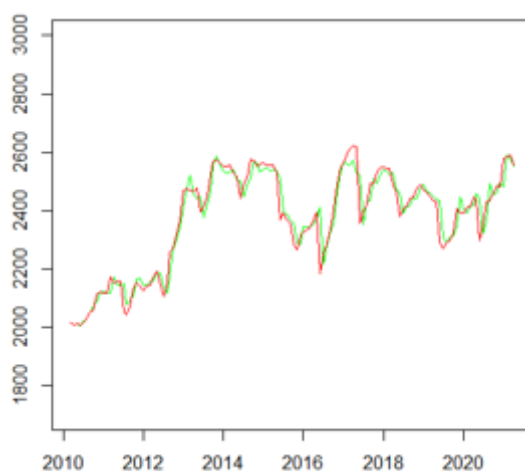


图 8 LSTM 模型训练集预测拟合情况

根据训练集拟合情况图，可以看出数据拟合程度较好。ARIMA 模型的训练集均方根误差为 28.36，与 ARIMA 模型相比，仅用 LSTM 模型在该训练集的拟合效果与 ARIMA 模型同样欠佳。再用测试集对训练集拟合的模型进行验证，得到测试集的均方根误差为 40.9，比测试集略大。其详细预测结果见下表：

表 9 LSTM 模型测试集预测效果

	真实值	预测值	误差	相对误差
2021-05	2540.00	2494	-46.00	-1.81%
2021-06	2522.22	2502	-20.22	-0.80%
2021-07	2561.67	2507	-54.67	-2.13%
2021-08	2595.56	2551	-44.56	-1.72%
2021-09	2612.78	2584	-28.78	-1.10%

根据表格结果，该模型平均相对误差 MAE 结果为 1.51%，LSTM 模型的预测值与实际值相比偏低，且误差逐渐变大。而 ARIMA 模型的未来 5 期预测值平均相对误差为 1.89%，从预测精度来看，LSTM 模型预测结果比 ARIMA 模型预测结果更好。

4.3 XGBoost 模型预测

将数据分成 2 部分，2009 年 1 月至 2021 年 4 月的数据作为训练数据来训练神经网络模型，2021 年 5 月至 2021 年 9 月的数据作为测试集用来检验神经网络模型。从 ACF 检验中推测小麦价格很可能存在一个滞后 12 期的影响因素(即小麦价格受到一年前同样月份的价格影响)，本次研究采取尝试分别将小麦价格滞后 1、2、3、12 期作为自变量加入进行 XGBoost 网络模型的建模并取其中均方根误差最小的一个模型进行小麦价格的预测。

表 10 不同滞后期数的组合模型预测效果对比

	均方根误差（训练集）	均方根误差（测试集）
滞后 1、12 期	37.8	34.2
滞后 1、2、12 期	34.2	16.7
滞后 1、2、3、12 期	39.4	98.8

整体而言将小麦价格滞后 1、2、12 期加入自变量时，其训练集和测试集的模型均方根误差均为最小，预测效果最优。而将因变量滞后 1、2、3、12 期加入自变量时，其训练集均方根误差虽然也较小，但测试集预测效果呈现出不稳定的现象，向外预测的效果不佳。绘制训练集的拟合情况如下图所示：



图 9 XGBoost 模型在训练集上预测拟合的情况

用测试集对训练集拟合的模型进行验证，预测结果见下表：

表 11 XGBoost 模型测试集预测效果

	真实值	预测值	误差	相对误差
2021-05	2540	2531	-9	-0.35%
2021-06	2522	2498	-24	-0.95%
2021-07	2562	2565	3	0.12%

2021-08	2596	2621	25	0.96%
2021-09	2613	2602	-11	-0.42%

相比单一 ARIMA 和 LSTM 模型，XGBoost 模型的预测效果更好，平均相对误差为 0.56%，且修正了传统 ARIMA 模型预测值较真实值偏低的特点。

4.4 ARIMA-LSTM-XGBoost 组合预测模型

在实际应用中，通常可以结合各个模型的结构优势来构建组合模型，以达到更稳定高效的模型结构。此外，可以使用交叉验证等方法来评估组合模型的预测能力，并根据实际情况对模型进行调整。我们考虑将三种模型的优点结合，构建组合模型的思路如下：首先使用 ARIMA 模型来捕捉时间序列数据中的线性趋势和季节性模式。然后，使用 LSTM 模型来捕捉非线性关系和长期依赖性。最后，使用 XGBoost 模型来整合 ARIMA 和 LSTM 模型的预测结果，以提高预测精度。

由构建的 ARIMA 模型进行线性趋势预测，得到 2009 年 1 月至 2021 年 4 月的原始小麦价格序列与 ARIMA 预测值的残差值如下图：

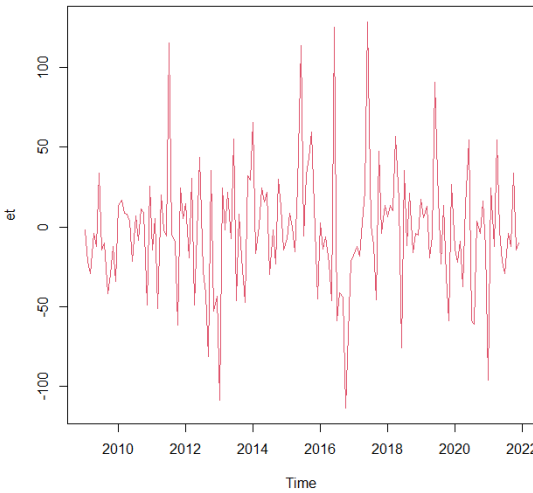


图 1 残差序列图

由残差序列图可以看出，残差值大的波动往往跟随着大的波动，小的波动往往跟随着小的波动，因此构造残差序列的自回归方程有利于帮助我们修正原 ARIMA 模型与真实值之间的差距。

非线性趋势预测通过结合 LSTM 和 XGBoost 两种模型实现。对残差值预测时得到的预测值进行反归一化处理，得到残差序列的真正预测值。将得到的预测值加和，即为 2009 年 1 月至 2021 年 4 月训练集最终的预测值，使用 2021 年 5 月至 2021 年 9 月作为测试集数据进行验证，结果见下表：

表 1 ARIMA 与组合模型在测试集预测效果

	真实值	ARIMA	预测残差	组合模型	误差	相对误差	平均相对误差
2021-05	2540.00	2538.918	-1.18	2540.10	0.10	0.00%	1.25%
2021-06	2522.22	2457.344	-31.90	2489.24	-32.98	-1.31%	
2021-07	2561.67	2503.054	-84.00	2587.05	25.38	0.99%	
2021-08	2595.56	2541.991	-11.70	2553.69	-41.87	-1.61%	
2021-09	2612.78	2539.56	-11.70	2551.26	-61.52	-2.35%	

相较于使用单一模型 ARIMA 模型进行预测，各年份误差、相对误差均有所下降，整体的预测效果更优，这主要应该这是由于 ARIMA 模型预测结果往往较真实值偏小，而 ARIMA-LSTM-XGBoost 模型对此进行了一定程度的修正。

4.6 模型比较

将本次研究建立的模型进行比较，结果如下表所示：

表 2 预测效果的模型结果比较

	真实值	ARIMA	LSTM	XGBoost	组合模型
2021-05	2540	2549	2494	2531	2540
2021-06	2522	2467	2502	2498	2489
2021-07	2562	2497	2507	2565	2587
2021-08	2596	2545	2551	2621	2554
2021-09	2613	2549	2584	2602	2551
平均相对误差		1.89%	1.51%	0.56%	1.25%

根据结果得出结论：在本项目中，LSTM 模型在训练集上的拟合性能不如 ARIMA 模型，但在测试集上的预测精度稍高。加入影响因素后，XGBoost 模型能够有效提升预测精度。ARIMA-LSTM-XGBoost 组合预测模型在小麦价格序列上的预测精度较高，稳定性更好，但不及 XGBoost 网络模型。尽管如此，组合模型仍具有很高的应用前景，能够更好地捕捉时间序列数据中的非线性关系、噪声和异常值、长期依赖关系以及多元关系。

5 项目特色与创新点

5.1 选题具有现实意义

当前全球各地疫情连续不断，阻碍各国进行有规律的农业生产，而近日俄乌局势动荡，乌克兰作为欧洲粮仓的地位被打破，导致以小麦为代表的国际与国内粮食价格被不断推高，直接增加了老百姓的生活成本与市场风险，也使国家难以对小麦价格进行调控。而对小麦价格的准确预测有利于国家积极应对市场波动并采取有效的政策策略，从而降低市场风险。

5.2 研究方法具有创新性

本课题基于集成学习算法中分类器的特点，创新性地将传统的 ARIMA 模型以及 LSTM、XGBoost 模型组合成多变量、非线性的预测模型。目前 XGBoost 算法和 LSTM 算法在多个领域的预测问题上取得了较好的成果，但二者在经济预测方面的研究较少，结合机器学习方法在经济序列预测的较好成绩，考虑引用影响因素、使用组合模型进行预测能有效地提高模型预测效果。

本项目建立了有关小麦价格预测的指标体系，并利用相关系数与随机森林进行变量筛选，综合考虑了变量间的线性和非线性相关程度，筛选出对小麦现货价格变化贡献较大的特征，有利于提高模型的预测能力与拟合速度。

本项目考虑到小麦现货价格同时具有线性与非线性特征，选择了适用于线性预测、短期预测精度较高的 ARIMA 模型拟合趋势，以及选择在处理非线性时间序列数据方面性能较好的其他模型，建立的组合预测模型相比于单一预测模型具有提高精度的优势。综合考虑了小麦价格时间序列本身的自相关图确定滞后阶数，发现小麦价格的影响因素（例如存在一个滞后 12 期，即小麦价格受到一年前同样月份的价格影响），从而帮助确定最优模型参数。

本项目分别建立 ARIMA 模型、LSTM 模型、XGBoost 模型以及它们的组合模型进行预测精度的比较，并得出局部最优模型的结构，挖掘集成模型在预测小麦价格用途的可能性，丰富了有关市场价格预测的理论研究。

6 参考文献

- [1]朱晶,张庆萍.中国利用俄罗斯、乌克兰和哈萨克斯坦小麦市场分析[J].农业经济问题,2014,35(04):42-50+111.DOI:10.13246/j.cnki.iae.2014.04.007.
- [2]杨惠珍,韦敬楠,张立中.我国粮食期货市场价格发现功能的实证分析——以玉米和小麦市场为例[J].价格月刊,2017(05):19-23.DOI:10.14076/j.issn.1006-2025.2017.05.04.
- [3]刘平,方旖旎.中国小麦进口结构及竞争力比较研究[J].市场周刊,2021,34(08):144-146.
- [4]高云,孙一铮,郭新宇,矫健.国内外小麦价格相关性及其影响因素研究[J].价格理论与实践,2018(05):71-74.DOI:10.19851/j.cnki.cn11-1010/f.2018.05.018.
- [5]林春霞.国际粮价将呈高位震荡趋势[N].中国经济时报,2021-09-29(001).DOI:10.28427/n.cnki.njjsb.2021.002172.
- [6]赵荣,曹洁,朱婷婷.浅议中国小麦的国际贸易及竞争力提升[J].滁州学院学报,2021,23(01):14-19+41.
- [7]朱聪,曲春红,王永春,赵伟.新一轮国际粮食价格上涨:原因及对中国市场的影响[J/OL].中国农业资源与区划:1-13[2022-03-08].<http://kns.cnki.net/kcms/detail/11.3513.S.20211130.1225.010.html>.
- [8]钟钰,陈希,牛坤玉.粮食出口限制政策的实施效果与我国应对——来自部分小麦出口国的证据[J].经济纵横,2021(08):29-39+137.DOI:10.16528/j.cnki.22-1054/f.202108029.
- [9]潘寅茹.乌克兰为何被称为“欧洲粮仓”?玉米小麦出口均占全球10%以上[N].第一财经日报,2022-03-04(010).
- [10]曹慧.新冠肺炎疫情对国内外小麦产业的冲击:趋势、问题与建议[J].世界农业,2021(01):4-10+27.DOI:10.13856/j.cn11-1097/s.2021.01.001.
- [11]申洪源.2021年中国小麦市场分析[N].粮油市场报,2022-02-15(B02).DOI:10.28553/n.cnki.nlysc.2022.000245.
- [12]邵海鹏.小麦也“疯狂”:进口首次突破配额,价格站上历史高位[N].第一财经日报,2022-03-02(A06).
- [13]邵海鹏.CBOT小麦期货连续涨停,国内外小麦价格倒挂影响几何[N].第一财经日报,2022-03-04(A10).
- [14]孙致陆.贸易开放背景下国际小麦贸易市场势力实证分析[J].华中农业大学学报(社会科学版),2019(04):1-14+169.DOI:10.13300/j.cnki.hnwkxb.2019.04.001.
- [15]陈植.芝商所小麦玉米期货迭创新高 对冲基金押宝“新避险投资”[N].21世纪经济报道,2022-03-04(005).
- [16]Haile M G, Kalkuhl M, Braun J V Worldwide Acreage and Yield Response to International Price Change and Volatility: A Dynamic Panel Data Analysis for Wheat, Rice, Corn, and Soybeans M//Food Price Volatility and Its Implications for Food Security and Policy. Springer International Publishing, 2016
- [17]Pal D, Mitra S K Interdependence between crude oil and world food prices: A detrended cross correlation analysis. Physica A Statistical Mechanics & Its Applications, 2018
- [18]Ahumada H, Cornejo M Forecasting food prices: The case of corn, soybeans and wheatly International Journal of Forecasting. 2016.
- [19]柳燕子.基于多模型分析的小麦价格预测研究[J].山西农经,2018(04):148.DOI:10.16675/j.cnki.cn14-1065/f.2018.04.103.
- [20]李雪,韩一军,付文阁.最低收购价政策对小麦市场价格波动影响的实证分析[J].华中农业大学学报(社会科学版),2018(02):1-7+154.DOI:10.13300/j.cnki.hnwkxb.2018.02.001.

- [21]张瑞娟,高芸.国内外小麦价格联动关系研究[J].价格理论与实践,2016(07):112-115.DOI:10.19851/j.cnki.cn11-1010/f.2016.07.029.
- [22]新冠肺炎疫情对中国粮食价格的影响效应及作用机制.崔海莹
- [23] 战玉锋,刘放. 我国小麦期货价格影响因素实证分析[J]. 沈阳工业大学学报(社会科学版), 2012, 5(3): 223-231.
- [24] 张恒,赵宇洋,安起光. 基于 BEKK-GARCH 模型的小麦期现货市场价格波动溢出性分析[J]. 粮食科技与经济, 2022, 47(1): 40-44.
- [25]张鑫, 吴海涛, 曹雪虹. Hadoop 环境下基于随机森林的特征选择算法[J]. 计算机技术与发展, 2018, 28(07):88-92.
- [26] 顾鑫. 基于 ARIMA-LSTM-XGBoost 组合模型的中国经济发展预测[D]. 青岛大学, 2021.
- [27] GAO W J, LI Z, MIN X, et al. Application system response time pre-diction based on ARIMA-LSTM composite mode [J] . Computer & Digital Engineering, 2021, 49 (5): 880-885.高文俊, 李志, 闵星, 等 . 基于 ARIMA-LSTM 组合模型的应用系统响应时间预测 [J] . 计算机与数字工程, 2021, 49 (5): 880-885.
- [28] XU D D, JIANG Z X. End-to-end speech recognition based on deep-ly optimized residuals convolutional neural network [J] . Application Research of Computers, 2020, 37 (S2): 139-141.徐冬冬, 蒋志翔 . 基于深度优化残差卷积神经网络的端到端语音识别 [J] . 计算机应用研究, 2020, 37 (S2): 139-141.
- [29]Du Y. Predicting China GDP Based on ARIMA-BP Neural Network Model[A]. Information Engineering Research Institute(USA) 、 Asia Pacific Human-Computer Interaction Research Center(Hong Kong).Proceedings of 2018 IEA 3rd International Conference on Frontier of Computer Science and Information Engineering(ICFCI 2018)[C].Information Engineering Research Institute(USA) 、 Asia Pacific Human-Computer Interaction Research Center(HongKong):智能信息技术应用学会, 2018:6.

7 附录

附录 1: R 语言代码节选

#相关系数图

```
library(PerformanceAnalytics)
library(corrplot)
```

```
x=read.table('clipboard',header=T);x
chart.Correlation(x, histogram=TRUE, pch=19)
corrplot.mixed(cor(x,method="pearson"),lower.col="black", number.cex=
.7,mar = c(2,2,3,2))
```

#随机森林特征重要性排序

```
wine = read.table("clipboard",header=T);wine
#将数据集分为训练集和测试集, 比例为7:3
train_sub = sample(nrow(wine),7/10*nrow(wine))
train_data = wine[train_sub,]
test_data = wine[-train_sub,]
```

```

library(pROC) #绘制 ROC 曲线
library(randomForest)
#随机森林
wine_randomforest <- randomForest(yt~x1+x2+x3+x4+x5+x6+x7+x8+x9,
                                   data = train_data,
                                   ntree =500,
                                   mtry=3,
                                   importance=TRUE ,
                                   proximity=TRUE)

#查看变量的重要性
wine_randomforest$importance
varImpPlot(wine_randomforest, main = "variable importance",font = 2,lw
d = 5,col = "red",col.axis = "blue",font.axis = 2,cex = 1,cex.axis = 5,
pch = 16,cex.lab = 0.7,font.lab = 2)

#arima 模型
library(tseries)
library(forecast)
library(fUnitRoots)
library(fGarch)
y1=read.table("clipboard")
head(y1)
na.omit(y1)
y11<-ts(y1,start=c(2009,1),end=c(2021,12),frequency=12)
plot(y11,type="l",xlab="时间",ylab="小麦现货价格")
ts.plot(diff(y11))
acf(y11,main="自相关图")#拖尾
pacf(y11,main="偏自相关图")#拖尾
trainy11<-ts(as.vector(y11[1:148]),frequency=12,start=c(2009,1))
#看拆掉之后的训练集图形
plot(trainy11,type="l",xlab="时间",ylab="小麦现货价格")
acf(trainy11,main="自相关图") #画出自相关图
pacf(trainy11,main="偏自相关图") #画出偏自相关图
tsdisplay(trainy11)
adf.test(trainy11)#接受原假设, 存在单位根
s1<-diff(trainy11,1)
plot(s1,type="l",xlab="时间",ylab="一阶差分结果")
acf(s1,main="自相关图") #画出自相关图
pacf(s1,main="偏自相关图") #画出偏自相关图
#进一步使用adf 检验, 看一下是否存在单位根(验证平稳性)
adf.test(s1)#拒绝原假设, 序列平稳
tsdisplay(s1)

a1=arima(trainy11,order=c(2,1,1),method="ML")

```



```

a2=arima(trainy11,order=c(12,1,1),fixed=c(NA,NA,0,0,0,0,0,0,0,0,0,NA,NA),transform.pars = FALSE)#AIC=907.17
res2<-a2$residuals^2
tsdiag(a2)
Box.test(a2$residuals, type="Ljung-Box",lag=12)#接受  $H_0$ 
Box.test(res2, type="Ljung-Box",lag=12)#拒绝  $H_0$ 
#然后使用tsdiag 看一下各自的结果, 图中表明残差标准差基本都在[-1,1]之间,
#残差的自回归都为0 (两虚线内), Ljung-Box 检验的p 值都在0.05 之上, 结果不错。
qqnorm(a2$residuals)
qqline(a2$residuals)

```

```

fore2<-forecast(a2,h=5,level=c(99.5))
plot(fore2,ylim=c(1700,3000))
lines(fore2$fitted,col="green")
lines(y11,col="red")

```

#神经网络

```

library(neuralnet)
options(digits=3)
x=read.table("clipboard",header=T);x
maxs=apply(x,2,max);maxs
maxs[1]
mins=apply(x,2,min);mins
mins[1]
m=(maxs-mins)[1];m
x1=as.data.frame(scale(x,center=mins,scale=maxs-mins));x1
attach(x1)
train.cse=x1[1:146,];train.cse
test.cse=x1[147:151,];test.cse

```

#滞后1 期

```

cse.net=neuralnet(yt~yt.1+ yt.12+x1+x2+x4+x5+x7+x8,data=train.cse,hidden=c(5,4),
threshold=0.01,act.fct="tanh",linear.output=T)

```

#滞后2 期

```

cse.net=neuralnet(yt~yt.1+yt.2+ yt.12+x1+x2+x4+x5+x7+x8,data=train.cse,hidden=c(5,4),
threshold=0.01,act.fct="tanh",linear.output=T)

```

#滞后3 期

```

cse.net=neuralnet(yt~yt.1+yt.2+yt.3+ yt.12+x1+x2+x4+x5+x7+x8,data=train.cse,hidden=c(5,4),
threshold=0.01,act.fct="tanh",linear.output=T)

```

#neuralnet 包的误差函数默认为sse, 没有mse, 本次研究只有1 个输出神经元, 可以

理解为mse

#单变量预测

```
cse.net=neuralnet(yt~yt.1+yt.2+yt.3+yt.12,data=train.cse,hidden=c(5,
4),
threshold=0.01,act.fct="tanh",linear.output=T)
ythat=cse.net$net.result[[1]]*m+mins[1];ythat
plot(cse.net)
```

```
rmse=function(x,y){sqrt(mean((x-y)^2))}
tyt=train.cse$yt*m+mins[1];tyt
rmse(ythat,tyt)
```

#测试集

```
ythat2=compute(cse.net,test.cse[,2:10])$net.result*m+mins[1];ythat2
tyt2=test.cse$yt*m+mins[1];tyt2
rmse(ythat2,tyt2)
```

#将数据结果写入

```
write.table(ythat[,1], file = "C:\\Users\\mujin1922\\Desktop\\Rresult1
", append = FALSE, quote = TRUE, sep = " ",
eol = "\n", na = "NA", dec = ".", row.names = TRUE,
col.names = TRUE, qmethod = c("escape", "double"),fileEncoding = "")
```

附录 2：实验分析数据集

	yt	x1	x2	x3	x4	x5	x6	x7	x8	x9
2009-01	1790.00	104.82	18.79	15.50	8000.86	101.00	101.30	3247.00	634.00	163.40
2009-02	1817.00	107.62	20.48	15.50	7062.93	98.40	101.00	92033.00	625.25	164.58
2009-03	1855.00	107.41	25.51	15.50	7608.92	98.80	100.30	28063.00	577.25	168.80
2009-04	1862.00	104.93	25.89	15.50	8168.12	98.50	100.50	35876.00	540.75	180.07
2009-05	1878.00	101.56	25.74	15.50	8500.33	98.60	101.20	71562.00	548.60	179.39
2009-06	1836.00	99.58	28.46	15.50	8447.00	98.30	101.00	192905.00	623.00	151.80
2009-07	1853.00	99.18	28.42	15.50	9171.61	98.20	102.10	85771.00	576.25	152.04
2009-08	1866.00	98.40	28.53	15.50	9496.28	98.80	102.70	8436.00	593.67	150.59
2009-09	1920.00	97.47	29.31	15.50	9712.28	99.20	102.80	74838.00	566.25	167.28
2009-10	1960.00	95.93	29.42	15.50	9712.73	99.50	103.20	13241.00	566.25	171.69
2009-11	1977.00	95.72	29.74	15.50	10344.84	100.60	103.30	94683.00	606.00	164.61
2009-12	2021.00	97.26	27.68	15.50	10428.05	101.90	103.90	203476.00	598.00	167.34
2010-01	2006.00	97.98	25.98	16.00	10067.33	101.50	104.70	86638.00	584.75	161.91
2010-02	2005.00	100.45	25.52	16.50	10325.26	102.70	104.20	95851.00	540.13	158.93
2010-03	2014.00	99.00	22.50	16.50	10856.63	102.40	107.90	192283.00	502.23	157.27
2010-04	2006.00	98.95	21.48	16.50	11008.61	102.80	106.60	132287.00	472.48	163.24
2010-05	2011.00	101.43	21.00	17.00	10136.63	103.10	108.00	191412.00	458.55	152.87
2010-06	2005.00	101.91	18.46	17.00	9774.02	102.90	108.50	191419.00	470.68	163.92
2010-07	2016.00	100.81	17.60	17.00	10465.94	103.30	107.80	133085.00	486.86	175.60
2010-08	2028.00	99.77	19.20	17.00	10014.72	103.50	107.30	90221.00	519.95	205.84
2010-09	2050.00	99.76	18.96	17.00	10788.05	103.60	104.10	91340.00	532.38	235.70
2010-10	2058.00	98.30	19.30	17.00	11118.49	104.40	103.80	12237.00	543.14	238.24
2010-11	2113.00	100.19	19.50	18.00	11006.02	105.10	102.90	7657.00	536.78	250.63
2010-12	2120.00	101.46	19.72	18.50	11577.51	104.60	100.40	6294.00	528.38	265.29
2011-01	2117.00	101.81	17.20	19.00	11891.93	104.90	99.90	60063.00	532.80	293.40
2011-02	2116.00	102.36	15.70	19.50	12226.34	104.94	99.60	116923.00	508.96	290.43
2011-03	2171.00	100.78	16.60	20.00	12319.73	105.38	107.60	75949.00	500.57	318.94
2011-04	2153.00	99.65	15.30	20.50	12810.54	105.34	106.60	11042.00	500.55	308.58
2011-05	2155.00	99.81	15.10	21.00	12569.79	105.52	105.80	10893.00	518.09	310.54
2011-06	2157.00	100.12	15.90	21.50	12414.34	106.36	108.10	280648.00	546.19	300.83
2011-07	2050.00	100.43	14.70	21.50	12143.24	106.45	105.60	108482.00	577.30	310.31
2011-08	2044.00	101.25	13.50	21.50	11613.53	106.15	105.00	128664.00	615.55	296.30
2011-09	2068.00	104.47	13.00	21.50	10913.38	106.07	103.40	177664.00	602.14	274.85
2011-10	2133.00	105.76	12.90	21.50	11955.01	105.50	100.50	60649.00	609.09	274.28
2011-11	2150.00	106.45	12.70	21.50	12045.68	104.23	97.00	8633.00	580.91	258.52
2011-12	2139.00	108.08	13.60	21.00	12217.56	104.07	100.50	218443.00	541.06	272.92
2012-01	2123.00	109.58	12.40	21.00	12632.91	104.50	103.90	213976.00	547.48	279.51
2012-02	2140.00	108.11	13.00	20.50	12952.07	103.20	105.00	373195.00	577.05	280.72
2012-03	2145.71	108.36	13.40	20.50	13212.04	103.60	100.00	537103.00	585.95	274.22
2012-04	2173.57	108.20	12.80	20.50	13213.63	103.40	103.00	293963.00	612.43	268.88
2012-05	2189.29	109.09	13.20	20.00	12393.45	103.00	104.20	557627.00	606.14	267.33
2012-06	2145.71	108.92	13.60	20.00	12880.09	102.20	99.30	217779.00	578.36	333.00
2012-07	2102.86	108.83	13.90	20.00	13008.68	101.80	98.20	252649.00	582.87	332.23
2012-08	2156.43	108.39	13.50	20.00	13090.84	102.00	99.40	252154.00	590.50	320.71
2012-09	2255.71	107.56	14.80	20.00	13437.13	101.90	100.80	524906.00	584.74	321.38
2012-10	2268.57	107.81	14.10	20.00	13096.46	101.70	106.10	351450.00	590.73	321.60
2012-11	2319.29	109.70	13.90	20.00	13025.58	102.00	105.10	118941.00	565.52	308.75
2012-12	2364.29	110.04	13.80	20.00	13104.14	102.50	103.70	7293.00	573.39	303.36
2013-01	2470.00	111.58	15.90	20.00	13860.58	102.03	104.50	181388.00	574.07	302.61
2013-02	2472.14	113.43	15.20	20.00	14054.49	103.22	108.20	223876.00	564.52	309.58
2013-03	2470.71	113.77	15.70	20.00	14578.54	102.07	102.60	289987.00	553.73	279.76
2013-04	2466.43	114.89	16.10	20.00	14839.80	102.39	103.70	204264.00	552.09	295.37
2013-05	2477.14	115.87	15.80	20.00	15115.57	102.10	99.00	284646.00	546.25	297.11
2013-06	2396.43	115.92	14.00	20.00	14909.60	102.67	97.00	221425.00	538.26	279.01
2013-07	2413.57	116.90	14.50	20.00	15499.54	102.67	97.20	311137.00	509.77	234.94
2013-08	2448.57	116.97	14.70	20.00	14810.31	102.57	97.80	421056.00	470.00	207.49
2013-09	2519.29	117.30	14.20	20.00	15129.67	103.05	99.80	719874.00	453.26	201.78
2013-10	2568.57	116.05	14.30	20.00	15545.75	103.21	102.90	1310449.00	448.81	199.18
2013-11	2574.29	117.47	14.20	20.00	16086.41	103.02	98.90	972141.00	447.55	197.55
2013-12	2561.43	118.54	13.60	20.00	16576.66	102.50	102.30	395215.00	441.04	198.77
2014-01	2552.14	120.80	13.20	20.00	15698.85	102.49	101.10	731935.57	447.00	209.36
2014-02	2550.71	120.71	13.30	20.00	16321.71	101.95	103.10	583985.08	434.14	222.39
2014-03	2555.00	117.15	12.10	20.00	16457.66	102.38	107.90	541479.87	409.94	222.43
2014-04	2527.14	114.52	13.20	20.00	16580.84	101.80	104.80	375355.19	403.59	216.98
2014-05	2515.00	113.89	13.40	20.00	16717.17	102.48	102.30	226461.69	414.67	202.63
2014-06	2441.43	113.89	14.70	20.00	16826.60	102.34	104.70	114271.02	428.00	182.29
2014-07	2487.14	114.53	13.50	20.00	16563.30	102.29	104.40	120423.58	440.38	176.48
2014-08	2510.71	116.46	12.80	20.00	17098.45	101.99	103.80	36223.17	436.50	163.13
2014-09	2573.57	119.24	12.90	20.00	17042.90	101.63	105.40	134612.88	432.70	163.38
2014-10	2570.00	121.13	12.60	20.00	17390.52	101.60	103.40	41426.64	419.00	178.71
2014-11	2552.86	123.80	12.30	20.00	17828.24	101.44	105.50	45572.78	410.74	178.71
2014-12	2558.57	125.29	12.20	20.00	17823.07	101.51	105.80	53208.73	409.68	174.79
2015-01	2563.57	127.37	10.80	20.00	17164.95	100.76	105.70	78881.78	409.50	173.76
2015-02	2555.71	129.30	12.50	19.50	18132.70	101.43	109.80	134158.30	400.73	174.29
2015-03	2557.14	130.34	11.60	19.50	17776.12	101.38	107.10	154667.25	392.41	172.15
2015-04	2555.71	129.57	10.10	18.50	17840.52	101.51	107.60	245241.12	382.43	166.35

	2015-05	2527.86	127.89	10.80	18.50	18010.68	101.23	109.90	368868.52	370.55	166.81
	2015-06	2366.43	128.68	11.80	18.50	17619.51	101.39	105.50	431837.16	387.74	179.61
	2015-07	2390.71	130.69	13.30	18.50	17689.86	101.65	104.48	302986.93	375.57	162.74
	2015-08	2367.86	130.10	13.30	18.50	16528.03	101.96	104.00	344444.15	358.77	166.08
	2015-09	2355.71	130.03	13.10	18.00	16284.70	101.60	105.60	209033.45	360.45	171.45
	2015-10	2282.14	128.85	13.50	17.50	17663.54	101.27	103.80	318181.06	358.38	166.09
	2015-11	2267.14	130.87	13.70	17.50	17719.92	101.49	104.10	199606.12	354.35	164.00
	2015-12	2317.86	130.21	13.30	17.50	17425.03	101.60	103.70	218955.53	359.48	161.10
	2016-01	2323.57	129.72	14.00	17.50	16466.30	101.80	104.00	240000.00	373.19	159.72
	2016-02	2331.43	130.94	13.30	17.50	16516.50	102.30	104.40	110000.00	370.48	159.23
	2016-03	2341.43	128.72	13.40	17.00	17685.09	102.30	100.00	210000.00	376.05	164.47
	2016-04	2363.57	126.33	12.80	17.00	17773.64	102.33	101.00	300000.00	408.82	169.00
	2016-05	2396.43	125.33	11.80	17.00	17787.20	102.04	99.80	440000.00	420.68	179.91
	2016-06	2183.57	123.32	11.80	17.00	17929.99	101.88	102.90	480000.00	456.24	161.82
	2016-07	2254.29	122.09	10.20	17.00	18432.24	101.77	106.80	310000.00	414.22	150.22
	2016-08	2285.71	121.19	11.40	17.00	18400.88	101.34	105.60	350000.00	384.00	148.51
	2016-09	2330.71	121.61	11.50	17.00	18308.15	101.92	104.60	440000.00	367.67	152.62
	2016-10	2406.43	121.36	11.60	17.00	18142.42	102.10	107.20	270000.00	360.82	151.33
	2016-11	2487.14	122.04	11.40	17.00	19123.58	102.25	108.60	30000.00	367.55	152.73
	2016-12	2547.14	123.12	11.30	17.00	19762.60	102.08	108.40	230000.00	372.55	160.01
	2017-01	2565.71	123.64	10.70	17.00	19864.09	102.55	109.20	340000.00	368.50	162.96
	2017-02	2585.56	122.44	10.40	17.00	20812.24	100.80	112.60	230000.00	367.78	158.98
	2017-03	2610.56	120.79	10.10	17.00	20663.22	100.90	111.00	510000.00	374.50	156.54
	2017-04	2622.22	119.76	9.80	17.00	20940.51	101.20	113.40	600000.00	402.91	158.60
	2017-05	2616.67	118.86	9.10	17.00	21008.65	101.50	112.00	510000.00	444.68	157.97
	2017-06	2354.44	119.20	9.10	17.00	21349.63	101.50	113.30	460000.00	419.95	157.61
	2017-07	2400.00	119.34	8.90	17.00	21891.12	101.40	114.60	240000.00	405.65	148.54
	2017-08	2418.89	120.07	8.60	17.00	21948.10	101.80	114.70	290000.00	413.14	147.32
	2017-09	2492.22	121.69	9.00	17.00	22405.09	101.60	118.60	440000.00	407.23	148.72
	2017-10	2492.22	122.03	8.90	17.00	23377.24	101.90	123.90	410000.00	404.05	148.79
	2017-11	2530.56	121.93	9.10	17.00	24272.35	101.70	121.30	180000.00	407.95	149.09
	2017-12	2547.78	122.14	8.10	17.00	24719.22	101.80	122.60	210000.00	422.30	155.92
	2018-01	2545.56	123.28	8.60	17.00	26149.39	101.50	122.30	250000.00	417.85	163.46
	2018-02	2542.22	126.20	8.80	17.00	25029.20	102.90	124.00	80000.00	409.77	172.08
	2018-03	2543.33	124.32	8.20	17.00	24103.11	102.10	122.30	310000.00	432.38	175.60
	2018-04	2490.00	124.61	8.30	16.00	24163.15	101.80	122.90	370000.00	433.22	179.18
	2018-05	2461.67	125.36	8.30	16.00	24415.84	101.80	122.90	630000.00	413.14	165.14
	2018-06	2378.89	124.52	8.00	16.00	24271.41	101.90	118.20	310000.00	378.68	156.55
	2018-07	2408.33	120.92	8.50	15.50	25415.19	102.10	119.70	140000.00	383.43	162.52
	2018-08	2414.44	119.84	8.20	15.50	25964.82	102.30	118.60	140000.00	383.00	154.95
	2018-09	2440.56	120.53	8.30	15.50	26458.31	102.50	118.50	280000.00	392.04	160.37
	2018-10	2445.56	120.03	8.00	14.50	25115.76	102.50	119.10	220000.00	390.82	160.88
	2018-11	2477.22	119.98	8.00	14.50	25538.46	102.20	122.10	140000.00	380.33	167.49
	2018-12	2488.89	121.09	8.10	14.50	23327.46	101.90	123.00	240000.00	383.74	166.85
	2019-01	2470.00	122.10	8.40	13.50	24999.67	101.70	123.70	610000.00	390.55	169.66
	2019-02	2460.00	124.53	8.00	13.50	25916.00	101.50	126.00	140000.00	383.14	166.32
	2019-03	2447.22	124.38	8.60	13.50	25928.68	102.30	124.10	290000.00	397.41	161.65
	2019-04	2432.22	124.01	8.50	13.50	26592.91	102.50	125.30	220000.00	390.30	171.18
	2019-05	2429.44	122.19	8.50	13.50	24815.04	102.70	123.40	210000.00	396.25	195.17
	2019-06	2291.11	120.43	8.50	13.50	26599.96	102.70	125.90	310000.00	395.96	189.49
	2019-07	2271.11	121.21	8.10	13.50	26864.27	102.80	124.40	220000.00	414.68	163.63
	2019-08	2289.44	119.84	8.20	13.50	26403.28	102.80	122.40	120000.00	415.43	157.37
	2019-09	2295.56	120.00	8.40	13.00	26916.83	103.00	124.10	160000.00	399.13	167.30
	2019-10	2324.44	120.89	8.40	13.00	27046.23	103.80	124.30	280000.00	394.57	166.36
	2019-11	2407.22	122.63	8.20	13.00	28051.41	104.50	124.60	450000.00	397.00	167.06
	2019-12	2392.22	122.25	8.70	13.00	28538.44	104.50	126.60	500000.00	428.52	171.81
	2020-01	2390.00	124.80	8.40	12.50	28256.03	105.40	126.40	340000.00	430.25	168.74
	2020-02	2410.56	126.45	8.80	12.50	25409.36	105.20	118.90	340000.00	471.68	162.18
	2020-03	2415.00	126.20	10.10	12.50	21917.16	104.30	122.20	590000.00	542.82	147.01
	2020-04	2445.56	125.74	11.10	12.50	24345.72	103.30	116.40	360000.00	498.86	143.99
	2020-05	2436.67	123.61	11.10	12.50	25383.11	102.40	115.80	810000.00	503.05	148.04
	2020-06	2296.67	121.50	11.10	12.50	25812.88	102.50	112.60	910000.00	464.48	152.59
	2020-07	2350.00	122.48	10.70	12.50	26428.32	102.70	117.20	930000.00	484.43	149.39
	2020-08	2429.44	122.79	10.40	12.50	28430.05	102.40	116.40	700000.00	491.91	166.60
	2020-09	2435.56	124.92	10.90	12.50	27781.70	101.70	120.50	1070000.00	458.86	186.77
	2020-10	2455.56	125.76	10.50	12.50	26501.60	100.50	121.70	630000.00	459.86	190.63
	2020-11	2486.67	126.25	10.70	12.50	29638.64	99.50	124.00	800000.00	499.35	199.18
	2020-12	2480.00	126.35	10.10	12.50	30606.48	100.20	122.10	880000.00	518.48	231.75
	2021-01	2580.00	127.65	9.40	12.50	29982.62	99.70	122.80	1500000.00	535.75	245.55
	2021-02	2581.11	129.57	10.10	12.50	30932.37	99.80	127.00	990000.00	502.35	245.76
	2021-03	2589.44	128.80	9.40	12.50	33874.85	100.40	122.20	440000.00	470.27	268.62
	2021-04	2555.56	127.20	8.10	12.50	34529.45	100.90	121.50	900000.00	477.19	304.30
	2021-05	2540.00	126.98	8.30	12.50	34502.51	101.30	121.80	790000.00	454.86	292.52
	2021-06	2522.22	126.86	8.60	12.50	34935.47	101.10	122.80	750000.00	414.86	278.45
	2021-07	2561.67	127.72	8.30	12.00	35360.73	101.00	117.80	880000.00	386.77	257.02
	2021-08	2595.56	127.74	8.20	12.00	33843.92	100.80	117.50	710000.00	387.00	235.82
	2021-09	2612.78	127.80	8.30	12.00	35819.56	100.70	121.20	640000.00	388.90	239.60