

沪深300个股系统性风险预测与控制

李漪灏 黄雅仕 詹鸿羽 叶凯琪 王枝宁*

韩山师范学院数学与统计学院 广东 潮州 521041

【摘要】：基于2014~2024年沪深300个股数据，构建系统性风险预测模型与事前风控体系。首先构建8个风险计量指标：平均收益率、市场流动性指标、市场情绪指标、波动率、回撤、GARCH波动率、风险价值（VaR）和条件风险价值（CVaR），通过加权平均法构建综合系统性风险指标 SR_m ，并对序列进行平稳性检验与参数优化，确定参数ARIMA(1,0,1)，进而构建ARIMA-GARCH系统性风险预测模型。量化测度沪深300个股系统性风险，并基于精准风险预测，为投资者提供科学的风险管理策略，促进金融市场的稳定发展。

【关键词】：风险计量指标；ARIMA-GARCH模型；加权平均；综合系统性风险指标 SR_m

DOI:10.12417/2982-3382.25.04.006

中国金融市场经历了从国有金融机构占主导地位到市场化、多元化发展的深刻变革，这一转变极大地促进了资本的有效配置和市场活力的提升。然而，随着市场的迅速扩张，系统性风险也日益凸显，如2008年的全球金融危机、2015年的股市剧烈波动和2024年初的市场波动等事件，均对金融稳定性构成严峻挑战。为有效应对风险，科学的风险管理成为不可或缺的一环，其中量化测度与实时监测系统性风险是实现这一目标的核心手段。通过构建和应用多种风险指标，可以更精确地评估和预测市场风险，从而为风险管理决策提供支持。本文基于多种风险计量指标，构建ARIMA-GARCH系统性风险预测模型^[1]，运用量化方法对沪深300个股的系统性风险进行测度与预测，以其为风险预警和市场监管提供科学依据，促进金融市场的稳健发展。

1 数据来源及数据预处理

本文的数据来源于2014年至2024年沪深300个成分股的日频行情数据。部分原始数据示例见表1。

表1 2018年沪深300个股的日频行情数据

time	code	open	high	low	close	volume	amount
2018/1/2 15:00	szse.000001	11.3500	11.8431	11.3245	11.6476	2081592	2856543822.18
2018/1/3 15:00	szse.000001	11.6731	11.7836	11.2225	11.3330	2962498	4006220766.94

2018/1/4 15:00	szse.000001	11.3245	11.3670	11.1630	11.2650	1854509	2454543516.41
2018/1/5 15:00	szse.000001	11.2310	11.3500	11.1800	11.3075	1210312	1603289517.93
2018/1/8 15:00	szse.000001	11.2650	11.2990	10.9334	11.0185	2158620	2806099169.66
2018/1/9 15:00	szse.000001	11.0185	11.2225	10.9844	11.1205	1344345	1754315792.47

运用Python对2014年至2024年沪深300个股的日频行情数据进行处理。首先，合并各年度数据以构建完整的数据集；随后，验证数据完整性并剔除重复数据，同时对成交量和成交额进行非负性校验，修正无效数据；最后，按日期和股票代码排序，为后续分析和建模提供可靠的数据基础。本研究的具体算法逻辑如下：

Step 1: 使用列表推导式和条件判断，检查数据集中是否包含必需的列（如code、time、close、high、low和volume）；

Step 2: 使用drop_duplicates函数，根据date(日期)和code(股票代码)列删除数据集中的重复记录，确保每条数据唯一；

Step 3: 通过布尔索引，筛选出close(收盘价)和volume(成交量)列中值大于0的记录，剔除无效数据；

Step 4: 使用sort_values函数，按照date和code列对数据集进行排序；

作者简介：李漪灏，黄雅仕，詹鸿羽，叶凯琪均为韩山师范学院数学与统计学院在读本科生。

*通讯作者：王枝宁，男，汉族，讲师，主要从事模糊数学优化及统计学研究。

本文得到广东省2025年国家级、省级大学生创新创业训练计划项目（粤教高函〔2026〕1号、项目编号：202510578015）、2024年度广东省本科高校教学质量与教学改革工程项目—校企联合实验室：数据科学创新创业实验室（粤教高函〔2024〕30号）、2025年度广东省本科高校教学质量与教学改革工程建设项目—面向新工科的AIGC赋能《深度学习》智慧金课建设实践研究（粤教高函〔2026〕4号）的资助。

Step 5: 使用 return 语句返回清洗和整理后的数据集;

Step 6: 计算 8 个风险计量指标, 计算公式见 3.1 节;

Step 7: 构建系统性分析预测模型和事前风控体系.

2 符号说明

表 2 公式符号说明

符号	说明
P_i	第 i 天的收盘价
v_i	第 i 天的成交量
r_i	日收益率
c_i	前 i 个交易日累计收益率

*其余符号在文中说明

3 研究方法

3.1 风险计量指标

风险计量指标是指用于量化金融风险的指标, 帮助投资者、风险管理者和监管机构评估和管理潜在的风险. 这些指标通常用于衡量市场风险、信用风险、流动性风险等不同类型的风险. 以下是本研究中用到的 8 个风险计量指标及其计算方法.

风险计量指标 1: 平均收益率, 其计算公式如下:

$$R_i = \frac{P_i - P_{i-1}}{P_{i-1}} \quad (1)$$

风险计量指标 2: 市场流动性指标, 以收盘价、成交量和日收益率计算流动性比率, 作为市场流动性指标^[2], 其计算公式如下:

$$l_i^1 = \frac{(p_i \times v_i)}{|r_i|} \quad (2)$$

风险计量指标 3: 波动率 (日对数收益率标准差), 其计算公式如下:

$$V_i^m = \sqrt{\frac{1}{m-1} \sum_{i=1}^m (r_i - \frac{1}{m} r_m)} \quad (3)$$

风险计量指标 4: VaR (参数法-正态分布计算)^[3], 其计算公式如下:

$$VaR_\alpha = -(\mu + z_\alpha \sigma) \quad (4)$$

风险计量指标 5: CVaR (参数法-正态分布计算)^[4], 其计算公式如下:

$$CVaR_\alpha = -(\mu + z_\alpha \frac{\sigma}{1-\alpha}) \quad (5)$$

风险计量指标 6: 市场情绪指标^[5], 其计算公式如下:

$$S_i = P_R V_R \quad (6)$$

其中, P_R 表示当日价格范围除以收盘价以标准化, 计算公式为:

$$P_R = \frac{H-L}{P_i} \quad (7)$$

V_R 表示交易量变化率, 为方便计算通常变化率加上 1, 计算公式为:

$$V_R = \frac{v_i - v_{i-1}}{v_{i-1}} + 1 \quad (8)$$

风险计量指标 7: 回撤 Draw^[6], 其计算公式如下:

$$Draw_i = \frac{R_c - R_{max}}{R_{max}} \quad (9)$$

其中 R_c 表示累计收益率, 计算公式为:

$$R_c = (1 + E_i) R_{i-1} \quad (10)$$

R_{max} 表示一段时间内累计收益率最大值, 计算公式为:

$$R_{max} = \max(R_{i-n+1}, R_{i-n+2}, \dots, R_i) \quad (11)$$

风险计量指标 8: GARCH 波动率: 本文使用 GARCH(1,1) 模型预测波动率, 其计算公式如下:

$$G\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (12)$$

未来波动率预测为:

$$G\hat{\sigma}_{t+h} = \sqrt{\hat{\sigma}_{t+h}^2} \quad (13)$$

其中

$$G\hat{\sigma}_{t+h}^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \quad (14)$$

3.2 ARIMA-GARCH 模型

3.2.1 模型准备

ARIMA-GARCH 模型是一种结合了 ARIMA 模型（自回归积分滑动平均模型）和 GARCH 模型（广义自回归条件异方差模型）的混合模型。它可同时对时间序列数据的均值和波动率进行建模。ARIMA 模型用于捕捉时间序列的均值动态，而 GARCH 模型则在波动率建模方面具有显著优势，可对金融市场的异方差性进行精准刻画。

(1) ARIMA 模型^[7]

ARIMA 模型是自回归 (AR)、差分 (I) 和移动平均 (MA) 三部分的组合，它的标准形式为 ARIMA(p,d,q)，其中 p 是自回归的部分阶数，p 是差分次数，q 是移动平均的阶数。

①自回归部分 (AR)：描述当前值与过去值之间的关系，其形式为：

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p \quad (15)$$

其中 B 是滞后算子，满足 $B^k SR_t = SR_{t-k}$ 。

②差分部分 (I)：用于使非平稳时间序列变得平稳，其形式为：

$$\nabla^d = (1 - B)^d \quad (16)$$

其中 ∇^d 为差分算子，满足 $\nabla SR_t = SR_t - SR_{t-1}$ 。

③移动平均部分 (MA)：描述当前值与过去误差项之间的关系，其形式为：

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q \quad (17)$$

将上述三个部分结合起来即是 ARIMA 模型，其综合公式为

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p)((1 - B)^d) = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \varepsilon_t \quad (18)$$

简化后的公式为：

$$\varphi(B)\nabla^d SR_m = \theta(B)\varepsilon \quad (19)$$

其中 ε 是白噪声数据。该模型可以捕捉时间序列数据的复

杂动态特性，能够适应不同市场条件和周期性变化，于是本文通过此模型来预测未来系统性风险。

(2) GARCH 模型^[8]

GARCH 模型由以下 3 个部分组成，分别为均值模型、异方差的分布假定、条件异方差模型。其中， r_t 是时间序列的观测值， μ 是均值， ε_t 是条件误差项， σ_t 是条件方差， z_t 是独立同分布的随机误差项，通常假设为标准正态分布。

$$r_t = \mu + \varepsilon_t \quad (20)$$

$$\varepsilon_t = \sigma_t z_t \quad (21)$$

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (22)$$

其中， $\alpha_0 > 0$ 是常数项， $\alpha_i, \beta_j \geq 0$ 分别衡量了过去的误差项平方和过去的条件方差对当前条件方差的影响， $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$ 且假定 α_i, β_j 满足一定条件使得 ε_t 的条件方差随时间变化是有限的。

3.2.2 模型的参数研究

本文通过建立 ARIMA 模型以预测未来的系统性风险。为提升模型的预测能力，首先采用 ADF 检验方法对综合系统性风险指标序列进行平稳性检验，通过对比检验结果的统计量与临界值，判断序列是否存在单位根，并根据检验结果确定差分阶数 d；随后，结合信息准则优化自回归阶数 p 与移动平均阶数 q，通过网格搜索法遍历 (p,q) 组合，并基于赤池信息准则 (AIC) 评估模型拟合优度，最终选取使 AIC 值最小的参数组合 (p,q)，从而确定 ARIMA 模型的最终参数。参数选择过程确保了模型对时间序列数据的精准拟合。检验方法如下：

(1) ADF 检验回归方程：

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^p \delta_i \Delta y_{t-i} + \varepsilon_t \quad (23)$$

γ 是单位根检验的关键系数， $\Delta y_t = y_t - y_{t-1}$ 为时间序列一阶差分，p 为滞后阶数， ε_t 为误差项， α 是常熟项， β 趋势项系数。

(2) 统计量：

$$t_\gamma = \frac{\hat{\gamma}}{SE(\hat{\gamma})} \quad (24)$$

其中原假设为：若 $t_\gamma >$ 临界值或 $t_\gamma =$ 临界值，序列存在单位

根，序列非平稳；备择假设为：若 $t_\gamma <$ 临界值，序列不存在单位根，序列平稳。

4 结果与分析

4.1 风险计量指标计算与标准化

根据 3.1 节的公式计算出市场整体的 8 个风险计量指标值后，为消除不同指标间的量纲差异，便于比较和分析，我们对这些指标进行了标准化处理。标准化处理后的指标更适用于模型构建和风险评估。标准化后市场整体的 8 个风险计量指标值的部分数据如表 3 所示。

表 3 市场整体的 8 个风险计量指标值

时间	平均收益率	市场流动性指标	市场情绪指标	波动率
2014-01-02	-1.356213	-0.019493	-0.183297	-0.595408
2014-01-03	-0.817632	-0.019493	0.012408	-0.595408
2014-01-06	-1.966544	-0.019493	0.146510	-0.653476
...
2024-10-31	0.501845	-0.019493	-0.232647	2.651079

时间	回撤 Draw	GARCH 波动率	VaR	CVaR
2014-01-02	2.358618	-0.101503	0.101503	0.101503
2014-01-03	2.358618	-0.101503	0.101503	0.101503
2014-01-06	2.034505	-0.101503	0.101503	0.101503
...
2024-10-31	0.719879	2.988731	-2.988731	-2.988731

4.2 综合系统性风险指标 SR_m 构建

根据上文计算得到的市场整体的 8 个风险计量指标值，本文通过加权平均的方法，对这 8 个风险计量指标值进行综合计算，进而得到综合系统性风险指标值。具体的加权平均计算公式如公式 (25) 所示。

$$SR = 0.15E_i + 0.15V_i + 0.075VaR_\alpha + 0.075CVaR_\alpha + 0.10S_i + 0.15Draw_i + 0.15L_i + 0.15\hat{\sigma}_{i+hh} \tag{25}$$

各风险计量指标的权重占比如图 1 所示。



图 1 风险计量指标权重占比

表 4 展示了综合系统性风险指标值的部分结果。

表 4 综合系统性风险指标值

时间	综合系统性风险指标
2014-01-02	0.039775
2014-01-03	0.140133
2014-01-06	-0.076121
2014-01-07	0.253227
2014-01-08	0.230144
...	...
2024-10-25	0.696148
2024-10-28	0.602203
2024-10-29	0.417630
2024-10-30	0.454764

4.3 ARIMA 模型的参数确定

基于上文得到的综合系统性风险指标值，我们建立了 ARIMA 模型以预测未来的系统性风险。ARIMA 模型能够有效捕捉时间序列的线性依赖关系，在金融时间序列预测中具有广泛的应用。为确保模型的准确性和稳健性，我们通过系统的方法对模型参数 (p, d, q) 进行了合理选择，最终确定的 ARIMA 模型参数组合为(1,0,1)，具体的参数选择方法详见下文。

(1) 差分阶数的选择

由于 ARIMA 模型要求时间序列是平稳的，本文对综合系统性风险指标 SR_m 序列进行了 ADF 检验，结果如表 5 所示。

表 5 ADF 单位根检验结果

变量	SR _m
ADF 统计量	-3.86
1%临界值	-3.96
5%临界值	-3.41
10%临界值	-3.13
滞后阶数	20
结论	拒绝非平稳假设

根据以上检验结果可知, ADF 统计量 (-3.86) 小于 1% 临界值 (-3.96), 表明在极严格条件下, 仍能以超过 95% 的置信水平拒绝非平稳假设, 这一结果证明了序列平稳性的高度可靠性, 为 ARIMA 模型的参数设定提供了直接依据. 因此, 后续建模可直接采用差分阶数 $d=0$, 避免因过度差分导致的信息损失或伪回归问题. 这表明市场整体回报率的波动围绕长期趋势呈现稳定分布特征, 为风险预测模型的参数稳定性和外推有效性提供了理论保障.

(2) 自回归阶数 p 与移动平均阶数 q 的选择

我们通过网格搜索法遍历 (p,q) 组合, 并基于赤池信息准则 (AIC) 评估模型拟合优度, 结果如图 2 所示. 根据 AIC 值最小所选取的 p 和 q 均为 1.

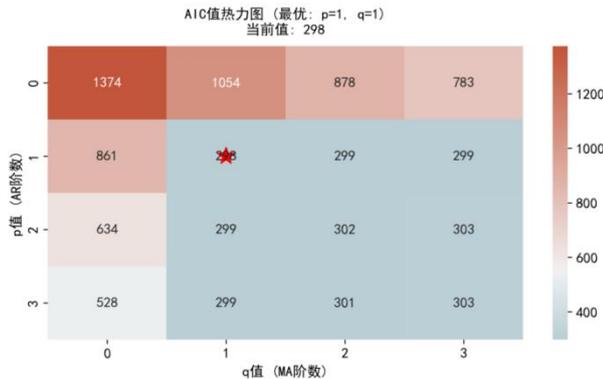


图 2 参数 p 和 q 的选择

4.4 系统性风险预测结果

(1) 未来一年系统性风险预测

基于上文建立的 ARIMA 模型, 我们对未来一年的系统性风险进行了预测. 图 3 展示了未来一年系统性风险的预测趋势.

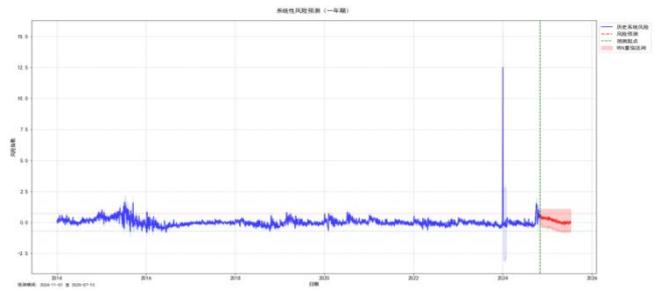


图 3 未来一年系统性风险预测趋势图

预测结果显示, 系统性风险的均值为 0.1669, 波动率为 0.1710, 且预测趋势呈下降状态 (概率为 66.63%), 表明未来一年的系统性风险较低. 基于预测结果, 我们构建了一个系统性风险预警报告 (表 6), 并提出了相应的建议措施. 该预警报告通过当前风险值判断系统所处的风险状态, 为风险管理提供参考.

表 6 简单预警报告

当前风险值(current_risk)	风险等级	预警颜色
$current_risk > 2$	极高	红色
$current_risk > 1$	高	橙色
$current_risk > 0$	中等	黄色
$current_risk \leq 0$	低	绿色

根据以上的风险等级, 本文给出以下三点建议:

- ①如果风险等级为“极高”或“高”, 建议立即减少高风险资产敞口, 增加低风险资产配置, 密切监控市场变化, 并准备应急预案;
- ②如果风险等级为“中等”, 建议保持均衡的资产配置, 适度控制风险敞口, 并定期评估市场状况;
- ③如果风险等级为“低”, 建议维持当前投资策略, 关注市场机会, 并定期检查风险指标.

(2) 未来一个月系统性风险预测

在预测未来一年系统性风险的同时, 本文也对未来一个月的系统性风险进行了预测. 图 4 展示了未来一个月系统性风险的预测趋势.

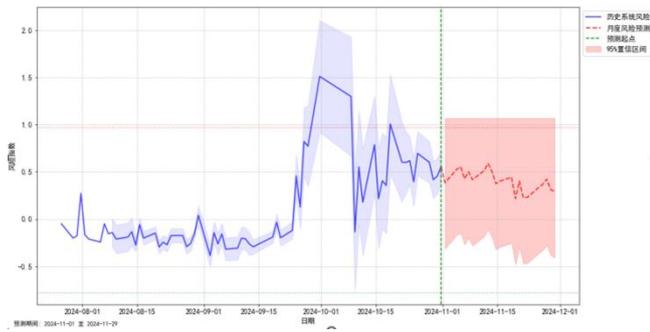


图4 未来一个月系统性风险预测趋势图

预测结果显示,未来一个月内系统性风险呈现一定的波动性,反映了市场的不稳定性.与2024年9月1日至10月1日的历史数据相比,未来一个月内的系统性风险指数出现了显著上升趋势,表明市场在此期间面临较大的压力或挑战;随后,系统性风险指数有所回落,显示出市场逐渐恢复稳定的迹象.基于此,投资者应密切关注市场动态和政策变化,及时调整投资策略以降低风险暴露;同时,政府和监管机构应加强市场监管和风险防范工作,确保市场的稳定运行.

参考文献:

- [1] 徐彩凤.基于 CoVaR 和 MES 的系统性风险预测及其在我国股市中的应用[D].北方民族大学,2024.
- [2] 杨立生,杨杰.货币政策、投资者情绪与股票市场流动性研究——基于 TVP-SV-VAR 模型的实证分析[J].价格理论与实践,2021,(09):142-145+203.
- [3] 颜翔宇.基于 ARMA-GARCH 与 VAR 模型的中信证券股票收益率分析[J].商展经济,2024,(11):98-101.
- [4] 王宗润,古慧敏.基于 CVaR 风险测度的后悔最小化投资组合模型[J].南昌大学学报(人文社会科学版),2021,52(03):41-49.
- [5] 高俊.投资者情绪对股票市场收益率的影响研究[D].合肥大学,2025.
- [6] 李思.基于回撤风险的投资组合优化和资产定价研究[D].西南财经大学,2024.
- [7] 周静雯.基于 ARIMA 模型的股票预测——以中国银行为例[J].现代计算机,2024,30(14):89-92.
- [8] 陈彦如,程敬雨.基于 K-Means 和 GARCH 模型的地方债利差分析[J].债券,2025,(02):86-92.

5 结论

系统性风险预测与控制在金融市场中具有重要意义.本文基于2014年至2024年沪深300个股数据,构建了系统性风险预测模型与事前风控体系,通过多维度分析验证了模型和体系在风险预测与控制方面的有效性.研究表明,ARIMA-GARCH模型能够较好地捕捉市场风险的动态特征.

首先,本文设计了8个风险计量指标,通过加权平均法构建了综合系统性风险指标,为风险量化提供了全面且科学的依据,基于此指标构建的ARIMA-GARCH模型在预测未来系统性风险时表现出较高的准确性,尤其是对未来一年和未来一个月的风险趋势预测,为投资者提供了及时的风险预警和决策支持.

综上所述,本文提出的系统性风险预测模型为理解和预判市场风险提供了科学工具,也为相关决策与研究提供了新的思路.未来的研究可进一步优化模型参数、引入更多风险因子,并探索结合机器学习等前沿技术以提升模型的动态预测能力.