

2023 年第四届“大湾区杯”粤港澳金融数学建模竞赛

实 测 报 告

1 数据与实测方法

1.1 数据集介绍

在本研究中，我们采用了任务二中提出的以 LSTM 网络（长短期记忆网络）为核心、全连接层为辅助的时间序列预测模型，以及任务三中提出的改进型跨市场交易策略，分析恒生科技 ETF（513130）以及任务三中提出的不同跨境 ETF 与股指期货的组合，以验证上述模型的适用性与鲁棒性。为此，我们聚焦于 2023 年 11 月 9 日至 11 月 17 日这一时间段内的交易数据，涉及 7 个交易日。此期间的数据被用于实测训练，旨在模拟高频交易环境。数据集的详细描述如下：

恒生科技 ETF（513130）及相关市场数据：

- **恒生科技 ETF 二级市场交易数据：**包括收盘价、开盘价、最高价、最低价、成交额和成交量。补充的数据来源于同花顺 iFinD 平台。
- **复权单位净值及其衍生指标：**包括恒生科技 ETF 的复权单位净值、贴水、贴水率及增长率。数据来源于同花顺 iFinD 平台。
- **申购申赎价格数据：**包括恒生科技 ETF 的第一市场申购和申赎价格数据。

跨市场套利相关数据：

- **中国大陆市场单只股指期货数据：**包括收盘价、开盘价、最高价、最低价以及成交量。补充的数据来源于同花顺 iFinD 平台。
- **香港、美国市场单只 ETF 交易数据：**包括收盘价、开盘价、最高价、最低价和成交量。数据来源于同花顺 iFinD 平台。
- **跨境 ETF 和股指期货组合：**基于前期任务三的研究，我们选择了 10 只最适合交易的跨境 ETF 和股指期货的两两组合，并按比例搭配得到了一个新的基金组合。本实测环节使用了这个新基金组合中，跨境 ETF 和股指期货两两组合的交易比例数据。

1.2 实测过程设计

在本研究中，实测过程的设计旨在评估和验证我们提出的交易策略模型在实际市场环境中的应用效果。由于涉及的数据时间粒度和均匀程度的差异，我们首先对数据进行预处

理以确保其一致性和可比性。随后，我们根据设定的模拟交易参数进行实测，以评估模型的性能。

1.2.1 数据预处理

为了应对不同数据源时间序列的不一致性，我们采取了数据拉伸操作。这一操作旨在确保所有数据在单位时间内的交易数据量保持一致，并且涵盖的总时长相同。对于拉伸后产生的数据空白，我们采用插值法进行填充。

具体而言，假设我们有一个时间序列 $y(t)$ ，其中 t 表示原始时间点。通过插值方法，我们可以得到新的时间点 t' 上的数据 $y(t')$ ，其中 t' 是在拉伸过程中新增的时间点。

插值公式可以表示为：

$$y(t') = y(t) + (t' - t) \cdot \frac{y(t+1) - y(t)}{t+1 - t} \quad (1)$$

上述公式中， $y(t+1)$ 和 $y(t)$ 表示原始时间序列中相邻两点的值，而 $t+1$ 和 t 是这两点对应的的时间。通过这种方法，我们能够拉伸后的时间序列生成连续、平滑的数据，从而提供均匀且一致的数据基础。

与此同时，对于数据中出现的毛刺等异常现象，我们进行了数据清洗操作，以保证数据能准确体现客观事实。

1.2.2 模拟交易实测方法

在模型的基础上，我们根据题目要求，考虑以下模拟交易参数：初始投资额为 150 万元，其中期货投资 100 万元，现货投资 50 万元。期货的保证金比率定为 10%，交易费率分别为期货 0.005% 和股票 0.01%，同时考虑价格滑点为 0.01%。

为了确保不同数据来源的精度一致性，我们将 9 天时间（共 7 个交易日）内的 206966 次数据交易时刻划分为不同的“时间步长”。具体而言，我们将这 206966 次数据交易时刻按照不同数据来源分别按 1 次、100 次或 2300 次作为一个时间步长，其中 100 次交易相当于每 0.81 分钟一个步长，2300 次交易则相当于每 18.7 分钟一个步长。这样的划分确保了在模拟交易中，无论数据来源的原始精度如何，每个“时间步长”内的数据都具有可比性。

在每个时间步长内，我们应用我们的模型来预测 ETF 和股指期货的价格动态，并据此进行交易决策。通过比较预测价格与实际市场价格的差异，我们可以计算每个时间步长的策略收益，进而评估模型在不同市场条件下的表现。

通过这种精细化的实测设计，我们旨在全面评估所建模型在实际交易环境中的应用效果，包括其预测准确性、收益潜力及对跨市场波动的适应性。此外，这种方法也允许我们更深入地理解我们的模型在处理高频交易数据时的行为和潜在限制。

2 实测结果与分析

2.1 股指期货与跨境 ETF 跨市场交易实测数据

通过任务三所选取出的 10 只最适合交易的跨境 ETF 交易组合结果，进而得到每个组合的单位时间步长策略收益、单位时间相对利润指数、跨境总利润、最大回撤、收益率曲线以及跨境交易数据预测结果等评价指标，如下所示。

交易编号	名称	单位时间步长策略收益均值
IM2311.CFE, 09151.HK	中证 1000, PP 科创 50-U	84.56702073
IM2311.CFE, 09173.HK	中证 1000, PP 中新经济-U	81.23501326
IC2311.CFE, 09151.HK	中证 500, PP 科创 50-U	75.34784388
IM2311.CFE, 09031.HK	中证 1000, 海通 AESG-U	79.1468256
IM2311.CFE, 09812.HK	中证 1000, 三星中国龙网-U	76.50042332
IC2311.CFE, 09173.HK	中证 500, PP 中新经济-U	72.32218134
IC2311.CFE, 09031.HK	中证 500, 海通 AESG-U	70.42639456
IC2311.CFE, 09812.HK	中证 500, 三星中国龙网-U	68.02417915
IM2311.CFE, 09801.HK	中证 1000, 安硕中国-U	65.36713021
IM2311.CFE, 09839.HK	中证 1000, 华夏 A50-U	63.63108462

表 1 跨境 ETF 与股指期货组合单位时间步长策略收益

交易编号	名称	单位时间步长相对利润指数
IM2311.CFE, 09151.HK	中证 1000, PP 科创 50-U	61.5020968607
IM2311.CFE, 09173.HK	中证 1000, PP 中新经济-U	61.5014847812
IC2311.CFE, 09151.HK	中证 500, PP 科创 50-U	55.8953614775
IM2311.CFE, 09031.HK	中证 1000, 海通 AESG-U	61.5012273485
IM2311.CFE, 09812.HK	中证 1000, 三星中国龙网-U	61.4957045967
IC2311.CFE, 09173.HK	中证 500, PP 中新经济-U	55.8948051881
IC2311.CFE, 09031.HK	中证 500, 海通 AESG-U	55.89457122
IC2311.CFE, 09812.HK	中证 500, 三星中国龙网-U	55.8895518588
IM2311.CFE, 09801.HK	中证 1000, 安硕中国-U	61.498875752
IM2311.CFE, 09839.HK	中证 1000, 华夏 A50-U	61.5014501249

表 2 跨境 ETF 和股指期货组合的单位时间相对利润指数

根据上述 10 个组合按任务三所提出的比例在上述时间段内的模拟交易，可以得到跨境交易总组合的总利润，如下表所示。

跨境交易总组合总利润值	¥ 13519150.29
-------------	---------------

表 3 跨境交易总利润值

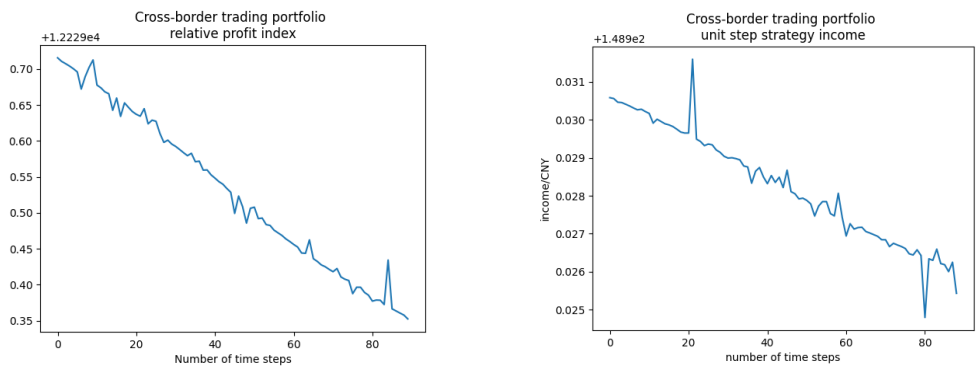


图 1 跨境交易相对利润指数(左图)，以及跨境交易单位时间策略收益(右图)

单位时间相对利润指数是用于衡量单位时间内跨境交易中的利润水平的指标，即跨境交易利润与单位时间销售额之比。而在跨境交易中，相对利润指数是用于衡量在不同时间步长内，跨境交易组合的盈利能力变化情况的一个指标。

上图显示了跨境交易组合的相对利润指数和单位时间步长策略收益的变化趋势。从图表中可以观察到，两者随时间步长的增加均呈现出一定的下降趋势。

具体来看，相对利润指数的逐步下降可能反映了交易组合随时间的盈利能力有所减少。这种下降趋势表明，在评估期内，该组合的盈利边际可能正在减弱。而单位时间步长策略收益的下降则进一步印证了这一点，表现为随时间推移策略所能实现的收益在减少。

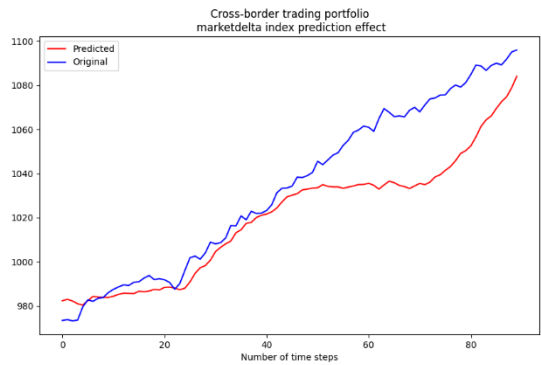


图 2 任务三跨境交易模型数据预测结果

从上图可以看出，累计跨境交易组合指数预测效果的两条曲线随时间步长的增加均呈现上升趋势（其中蓝色线代表实际值，红色线代表预测值），其预测曲线与实际曲线的趋势

大致相同，但实际值的增幅大于预测值，说明预测模型能够在一定程度上反映市场趋势，但在精确程度上还有改进空间。

2.2 恒生 ETF 交易实测数据

针对任务二的模型所进行的模拟交易，我们详细记录并分析了以下关键指标：ETF 的总交易利润、单位时间步长内的策略收益、ETF 的平均利润率、恒生科技 ETF 的最大回撤率，以及 ETF 在第二市场的交易价格预测结果，如下所示。

ETF 交易总利润是指在一定时期内，通过交易 ETF（交易所交易基金）所获得的利润总额。这个指标反映了投资组合在 ETF 交易中的表现，是评估投资效益的重要参数。针对恒生科技 ETF 在对应时间的实测数据，我们可以得到套利总利润：

恒生科技 ETF 套利总利润	¥ 2471863.03
----------------	--------------

表 4 恒生科技 ETF 套利总利润

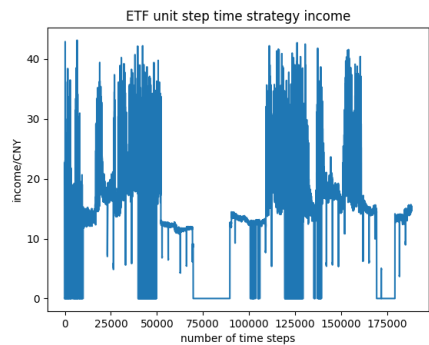


图 3 ETF 单位步长时间策略收益

上图展示了 ETF 单位时间步长策略收益随时间的变化情况。通过分析可以看出，策略收益在整个时间序列中呈现出较大的波动性，这突显了策略在响应市场变化时的反应度。特别是在序列的早期，收益率高点出现的频率较高，而在后续时间步长中，这些高点的出现频率及其幅度均呈现出下降趋势。这种趋势的变化可能指示了市场条件的演变或策略在不同市场环境中的表现具有变异性。

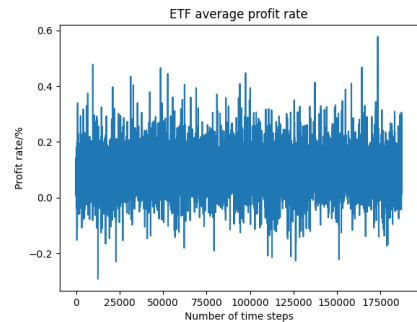


图 4 ETF 平均利润率

该图表展示了 ETF 平均利润率随时间步长的变化，利润率定义为在每个特定时间步长内实现的利润与投资本金之间的比率。此比率是衡量投资回报率的关键指标。

图中数据揭示了 ETF 的平均利润率具有显著的波动性，表现出在不同时间步长的利润率波动幅度较大，且包含正负两方面的变化。这表明在某些时间步长期间投资策略实现了正的回报，而在另一些时间步长内则出现了损失。这种波动性的存在可能反映了市场的不确定性，也可能是交易策略对市场变动反应的直接体现。

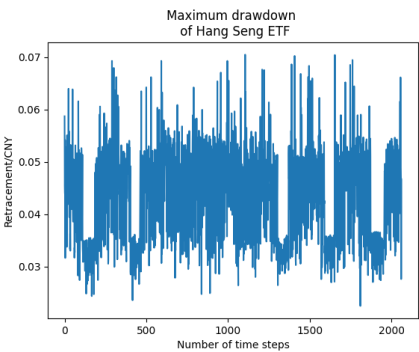


图 5 恒生 ETF(513130) 各时间步长最大回撤

上图展示了恒生 ETF 在各个时间步长内最大回撤率的波动情况。最大回撤率，作为衡量资产在一定时间内所经历的最大潜在损失的指标，为我们提供了关于资产风险的重要信息。

从图中可以看出，恒生 ETF 的最大回撤率在整个观测期内表现出相对高的波动性。尽管在某些时间步长，如 0-100 和 150-250 区间，最大回撤率的波动较为显著，显示出市场的快速变动或较大的价格压力，但在其他时间段，例如 100-150 步长区间，最大回撤率相对平稳，表明市场在这一阶段相对稳定。这种波动的存在可能指出 ETF 价值随市场波动而变化的特性，以及在不同时间点上投资者可能面临的不同风险水平。

最大回撤率的波动性强调了该跨境 ETF 相较于其他境内 ETF 可能具有更高的市场波动性。这一现象与我们的模型优化策略相一致，暗示了跨境 ETF 在市场波动性上的潜在优势，以及由此可能产生的较大套利机会。因此，这些实证结果有效地支持了任务二模型在捕捉高波动性市场中套利机会方面的优异性能。

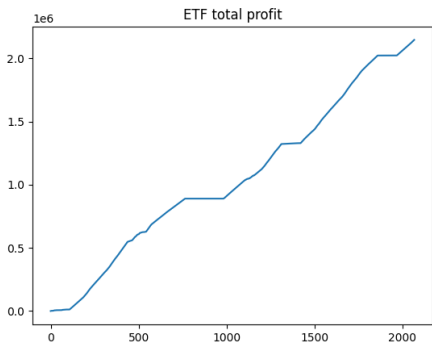


图 6 恒生 ETF 总时间策略收益

上图描绘了恒生 ETF 单位步长时间策略在连续时间步长中累计利润的变化轨迹，体现了交易策略随时间推进的总收益情况。在所观察的时间序列中，总利润呈现出了一致的增长趋势，透露出策略在整个时期内稳定的盈利能力。

值得注意的是，在整个时间序列中，总利润的增长并未出现显著的波动，这可能表明策略能够有效应对市场的波动，或者策略设计考虑到了市场波动的平稳对冲。尽管没有尖峰或深谷的剧烈波动，这一稳定上升的利润曲线仍然显示出在特定时期策略捕捉到连续的交易机会，或对市场趋势有一致的正确判断。

图表所示的总利润增长趋势，尤其是在时间序列的后段，维持了平稳的上升态势，这强化了策略在长期内保持盈利能力的观点。这种持续的增长趋势表明，尽管市场可能存在不确定性，策略仍然能够稳健地积累收益，展现出在不同市场条件下的稳定表现。此外，总利润的持续增长也暗示了策略在实现长期价值增长方面的有效性。

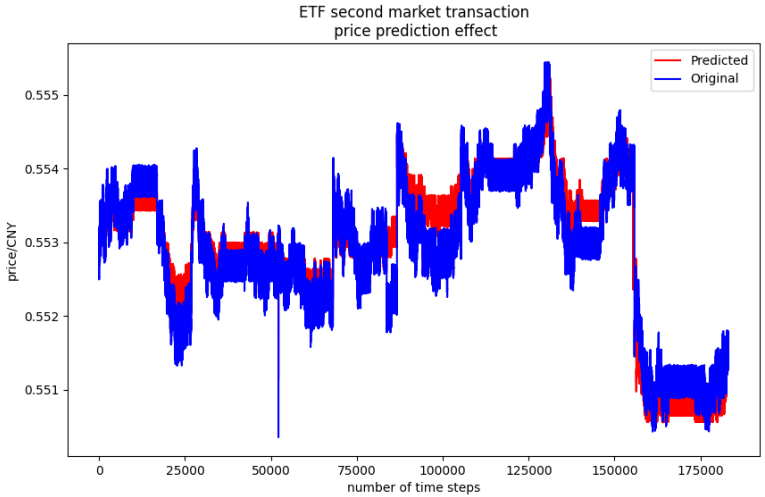


图 7 ETF 第二市场交易价格预测结果

图表展示了 ETF 在第二市场交易价格的预测效果，其中红色线代表预测价格，蓝色线代表实际观测价格。整体而言，预测轨迹与实际价格轨迹在大多数时间步长上保持较为紧密的一致性，显示了预测模型具有一定的准确性。尽管如此，预测与实际观测值之间的差异在特定时间区间内变得更为显著。具体地，在时间步长 10000 至 30000 的区间内，以及时间步长 95000 至 110000，预测价格与实际价格之间的偏离增大，这表明在这些区间内，模型的预测准确性受到影响。需要指出的是，尽管在上述特定时间区间预测准确性降低，但整体来看，模型对于价格波动的预测捕捉了实际趋势的主要变化点。