

From:	朱若兰	To:	投资部
Date:	2022 年/12 月/9 日	Class:	Week1 测试
Title:	AZRL001001	Version:	1.0.0

1 摘要

本文测试了期货跨品种套利，利用两个走势相近品种价差的均值回归特性，在价差偏离均值水平较大时，采取做空或做多价差操作，当价差回归正常水平时，对之前操作进行平仓获取价差回归的收益。用期货连续合约 2020 年 1 月 5 日至 2022 年 12 月 7 日的日频数据，对所有分类后的期货品种进行协整性和相关性检验，共筛选出黑色系 5 对、能源化工类 5 对和农产品类 9 对。对通过检验的品种对进行套利并对策略进行改进，最后通过敏感性分析选择最有参数并对策略进行评估。

2 策略主要思想

2.1 策略主要思想

跨品种的统计套利主要依据两个走势相近品种间价差存在均值回复的特征，因此当价差超出价差上轨（价差均值加一个标准差）时，我们可以做空价差，即买入低估品种卖空高估品种。反之，当价差低于价差下轨（价差均值减去一个标准差）时，我们可以做多价差。当价差恢复到正常水平时，对之前操作进行平仓，获取价差回归收益。

具体来说，如果 A 和 B 两个资产的价格时间序列呈现协整关系，那么便存在实数 a 和 b，使得线性组合 $spread = aA - bB$ 是一个平稳序列（spread 满足均值回归）。若 spread 值向上偏离很高，根据弱平稳性质，spread 将均值回归，这时应该卖出 a 份 A 并买入 b 份 B 来做空 spread，并在 spread 回归均值时赚取差价。反之，如果 spread 值偏低，那么应该买入 a 份 A 同时卖出 b 份 B 来做多 spread，并在 spread 回归时平仓赚取收益。

本文将先采取传统跨品种套利，通过对期货历史价格进行回归，得到固定的价格比，并利用该价格比确定固定的价差均值、价差上下限和价差正常范围。当价差超出上限时，做空价差，当价差跌破下限时，做多价差，当价差回归正常范围时平仓。

其次本文对传统策略进行改进，传统策略有以下几个弊端，第一个方面品种间价格比例系数并不是一成不变，第二个方面价差均值也不一定维持固定值不变，第三方面固定价差上下限并不可靠，当价差由于基本面变化超出正常范围时，固定上下限可能导致策略不能及时抓住价差变化，从而导致回撤较大获利较少。

保持均值回归思想不变，对一下三个方面进行改变

- (1) 首先品种间的比例系数通过滚动回归获得，滚动窗口的长度根据敏感性分析确定；
- (2) 价差的均值水平采用 n 日移动均值，参数 n 也由敏感性分析确定；
- (3) 价差的上下轨和价差正常上下限都不是固定不变，价差上下轨采用移动价差均值的加减一倍标准差，价差正常上下限为移动价差均值加减 0.5 倍标准差。

2.2 交易配对选择逻辑

根据本地数据的期货日频数据获取全品种信息，将 41 种可交易期货品种根据类型和交易所进行分类为黑色系、有色金属、贵金属、能源化工、农产品和金融，其中黑色系有 8 种，有色金属 6 种，贵金属 2 种，能源化工 10 种，农产品 12 种，金融 3 种，具体如下表：

表 1 全品种分类表

类别	交易所	商品代码	名称
黑色系	郑商所 (CZC)	ZC	动力煤
		SM	锰硅
		SF	硅铁
	大商所 (DCE)	JM	焦煤
		J	焦炭
		I	铁矿石
	上期所 (SHF)	RB	螺纹钢
		HC	热卷
有色金属	上期所 (SHF)	CU	铜
		AL	铝

		ZN	锌
		PB	铅
		NI	镍
		SN	锡
贵金属	上期所 (SHF)	AU	黄金
		AG	白银
能源化工	郑商所 (CZC)	TA	PTA
		FG	玻璃
		MA	甲醇
		SA	纯碱
	大商所 (DCE)	EG	乙二醇
		V	聚氯乙烯
		L	线性低密度聚乙烯
		PP	聚丙烯
	上期所 (SHF)	BU	沥青
		RU	橡胶
农产品	郑商所 (CZC)	WH	强麦
		CF	棉花
		SR	白糖
		OI	菜籽油
		RM	菜籽粕
	大商所 (DCE)	C	玉米
		CS	玉米淀粉
		A	大豆 1 号
		M	豆粕
		Y	豆油
		P	棕榈油
		JD	鸡蛋
金融	中金所 (CFE)	IC	中证 500 股指期货
		IH	上证 50 股指期货
		IF	5 年国债期货

2.2.1 协整性检验

用 2020 年 1 月 2 日到 2022 年 12 月 6 日开盘价，对四个大类（黑色系、有色金属、能源化工、农产品）进行协整检验，在 5% 的置信水平下，黑色系品种里有五对通过了协整检验，分别是：动力煤和硅铁（ZC-SF）、硅铁和焦煤（SF-JM）、硅铁和焦炭（SF-J）、焦

煤和焦炭（JM-J）、螺纹钢和热卷（RB-HC）。有色金属类有两对通过了协整检验，分别是铅和镍（PB-NI）、铅和锡（PB-SN）。能源化工类也有五对通过协整检验PTA和纯碱（TA-SA）、PTA和沥青（TA-BU）、甲醇和纯碱（MA-SA）、乙二醇和聚氯乙烯（EG-V）、聚丙烯和橡胶（PP-RU）。农产品有九对通过协整检验，菜籽油和大豆一号（OI-A）、菜籽油和豆油（OI-Y）、菜籽油和棕榈油（OI-P）、玉米淀粉和大豆一号（CS-A）、玉米淀粉和豆油（CS-Y）、玉米淀粉和鸡蛋（CS-JD）、大豆一号和豆油（A-Y）、大豆一号和棕榈油（A-P）、大豆一号和鸡蛋（A-JD）。

图 1：黑色系品种协整检验

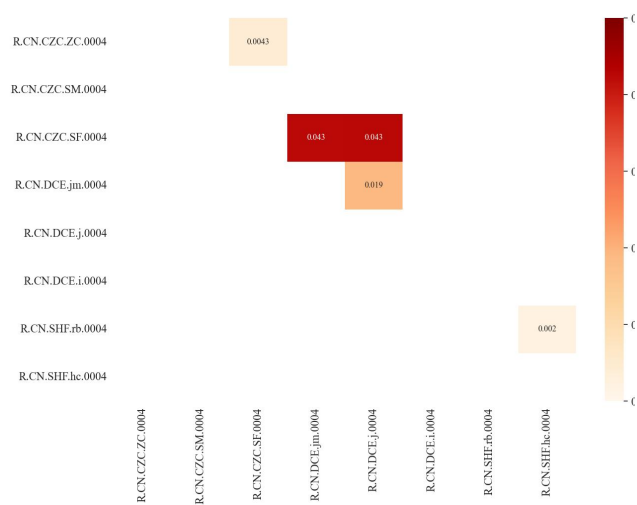


图 2：有色金属类协整检验



图 3：能源化工类协整检验

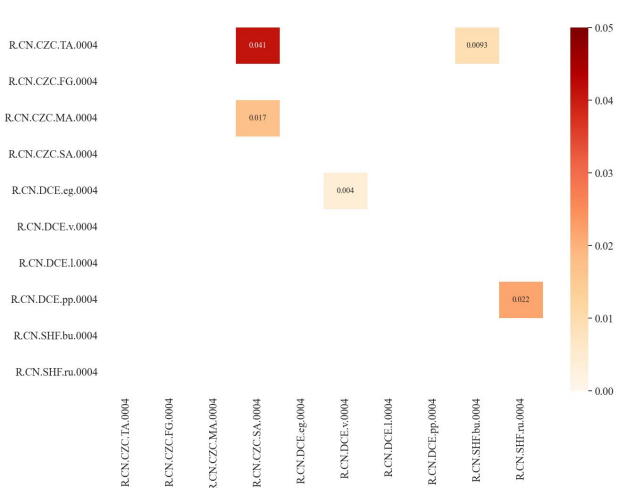
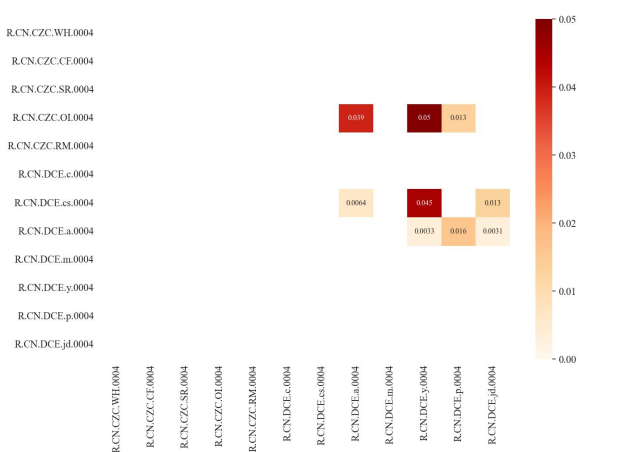


图 4：农产品类协整检验



下面将进一步检测四类期货的品种间的相关性，将 0.7 设置为价格时间序列相关性的阈值。

黑色产品配对池共有 28 待检测期货对，有 11 对通过 0.7 相关性标准；有色金属配对池共有 15 对待检测对，有 9 对通过了相关性检验；能源化工共有 45 对待检测期货对，共有 20 对通过了相关性检验；农产品共有 66 对待检测对，共有 50 对通过了 0.7 相关性阈值。从整体来看，农场品间的相关度更高，有 83%的品种相关性超过 0.7。

图 1：黑色系品种相关性检验

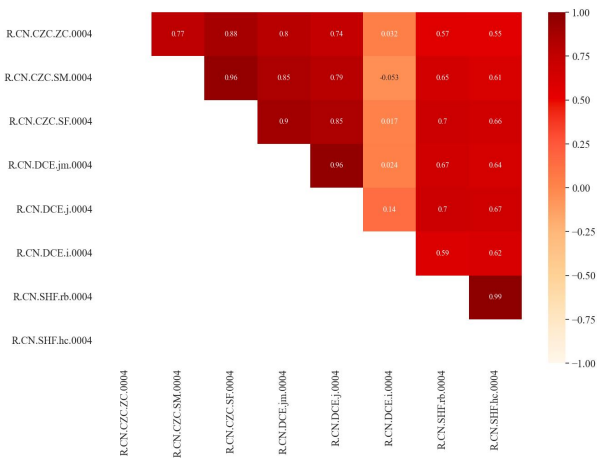


图 2：有色金属类相关性检验

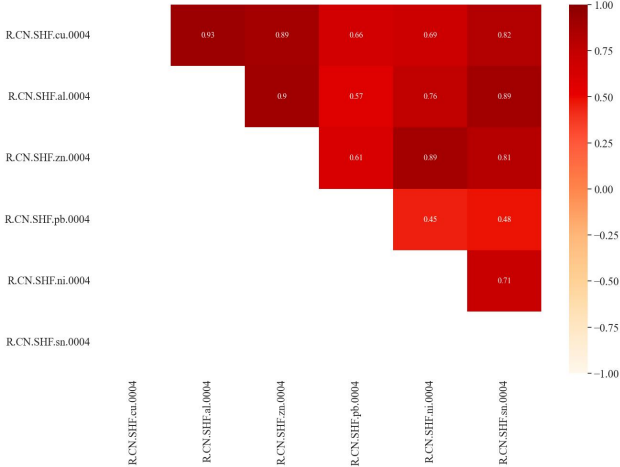


图 3：能源化工类相关性检验

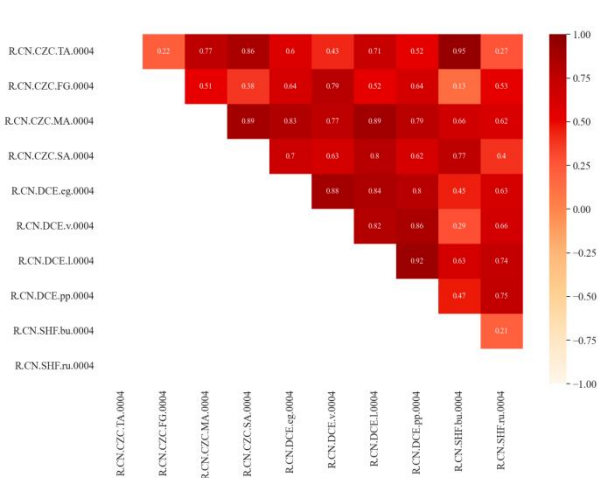
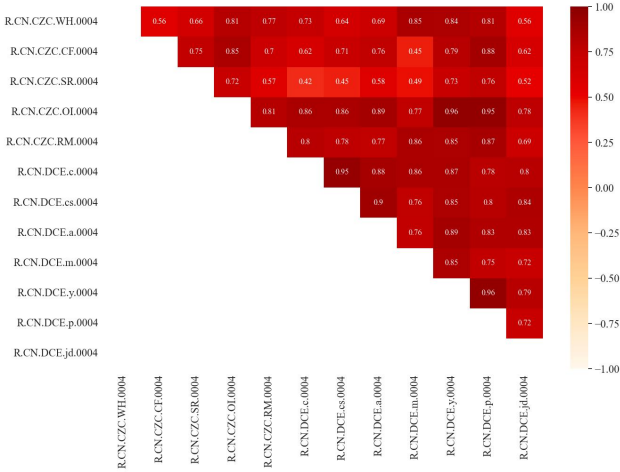


图 4：农产品类相关性检验



以协整性检验和相关性检验为筛选标准，黑色系、有色金属和农产品通过协整检验的期货对都通过了相关性检验，而两对通过协整检验的能源类产品（PB-NI 铅镍、PB-SN 铅锡）都未通过相关性检验，我们筛选出来可以进行跨品种配对套利的品种如下表：

表 2 通过协整检验和相关性检验品种

期货类别	商品代码	名称
黑色系	ZC-SF	动力煤(100)和硅铁(5)
	SF-JM	硅铁(5)和焦煤(60)

	SF-J	硅铁(5)和焦炭(100)
	JM-J	焦煤(100)和焦炭(60)
	RB-HC	螺纹钢(10)和热卷(10)
能源化工类	TA-SA	PTA(5)和纯碱(20)
	TA-BU	PTA(5)和沥青(10)
	MA-SA	甲醇(10)和纯碱(20)
	EG-V	乙二醇(10)和聚氯乙烯(5)
	PP-RU	聚丙烯(5)和橡胶(10)
农产品	OI-A	菜籽油(10)和大豆一号(10)
	OI-Y	菜籽油(10)和豆油(10)
	OI-P	菜籽油(10)和棕榈油(10)
	CS-A	玉米淀粉(10)和大豆一号(10)
	CS-Y	玉米淀粉(10)和豆油(10)
	CS-JD	玉米淀粉(10)和鸡蛋(10)
	A-Y	大豆一号(10)和豆油(10)
	A-P	大豆一号(10)和棕榈油(10)
	A-JD	大豆一号(10)和鸡蛋(10)

3 一次回归套利

3.1 价差构造方式

以焦煤和焦炭为例，通过对焦煤和焦炭的价格时间序列进行 OLS 回归，得到回归系数 b，再根据以下公示计算价差：

$$Y = bX + c$$

$$\text{spread} = Y - bX$$

图 5：焦煤焦炭价格走势

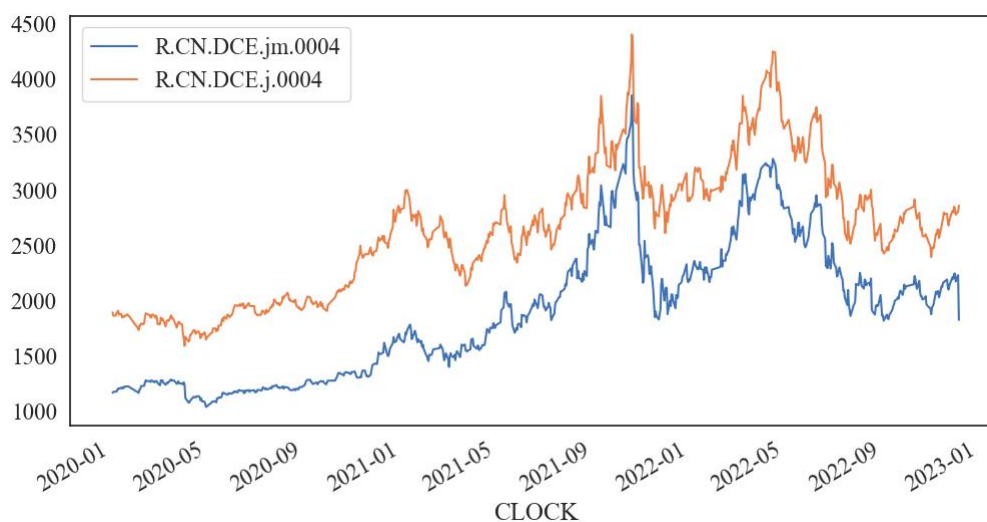
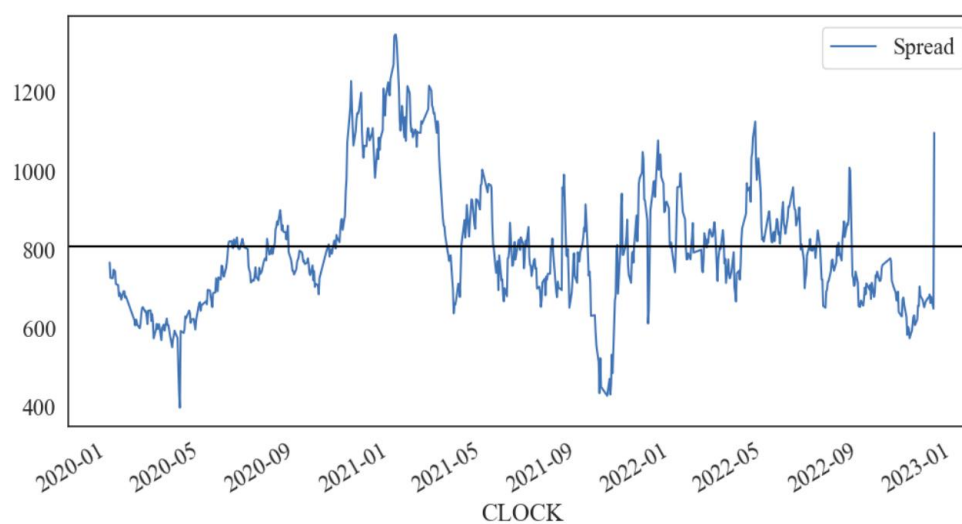


图 6：焦煤焦炭价差走势图



从图 6 可以看出，价差均值为 800，并且价差围绕价差均值上下波动，总体呈现均值回归水平。

3.2 指标构建及表现

3.2.1 指标的计算方法，公式等

将 2020-01-04 到 2022-01-04 的开盘价作为训练集确定焦煤焦炭的价格比、价差均值、均值上下限以及均值正常水平。首先将焦煤和焦炭从 2020-01-02 到 2022-01-04 的价格进行线性回归，得到以下方程：

$$j = 0.952 * jm + 838.76 \quad (1)$$

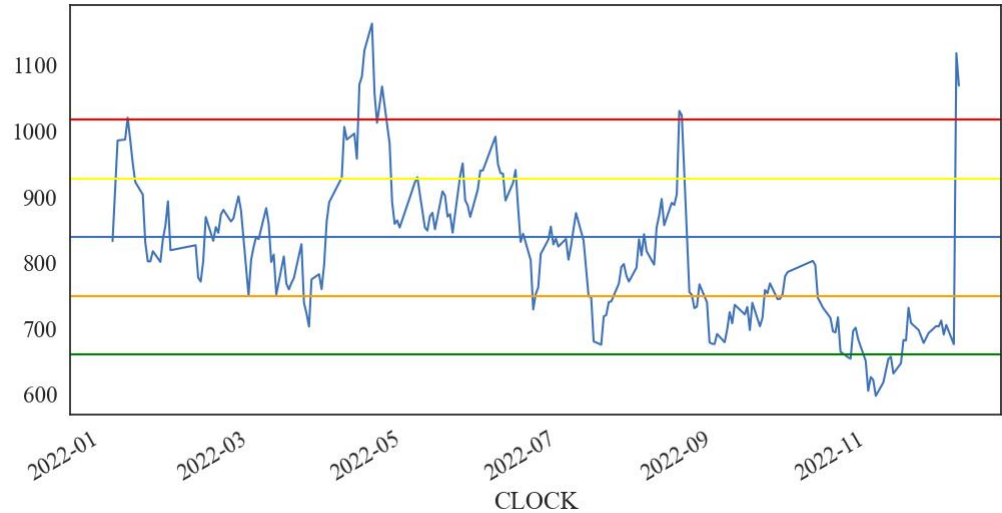
2020-01-02 到 2022-01-04 间，通过线性回顾焦炭和焦煤的配比为 0.952，价差的均值为 838.76，价差的标准差为 178.29。假设保持配比和价差均值不变（838.76），价差上限设置为价差加一倍标准差（1017.05），价差下限为价差减一倍的标准差（660.47），均值正常水平为上下 0.5 倍标准差（927.91 和 749.62）。

2022-01-04 年之后的价差用以下公式进行计算：

$$\text{spread} = j - 0.952 * jm \quad (2)$$

当价差超过 1017.05 时，认为此时价差过大，此时焦炭被高估（j），焦煤（jm）被低估，应该做空价差，即卖空一份焦炭（j）的同时，买入 0.952 份焦煤（jm）。当价差低于 660.47 时，认为价差过低，焦炭（j）被低估，焦煤（jm）被高估，应该做多价差，做多焦煤的同时，做空 0.952 份焦煤。当价差回归正常水平时，将之前的仓位平仓。

图 7 焦煤焦炭价差图



具体交易方式如下方清单：

表 3 通过协整检验和相关性检验品种

交易要素	内容
交易标的	焦煤（jm）焦炭（j）
交易合约	主力合约
交易频率	日频，信号触发时交易
交易成交时间	当天开盘成交
模拟成交价格	当天开盘价
收益计算使用价格类型	复权价格
是否考虑成本和换月	无

3.2.2 模拟表现

通过以上指标构造公式以及开平仓逻辑，我们可以得到 2022-01-05 至 2022-12-07 焦煤、焦炭的交易信号图和仓位模拟图如图 8-图 11。

图 8：价差信号图

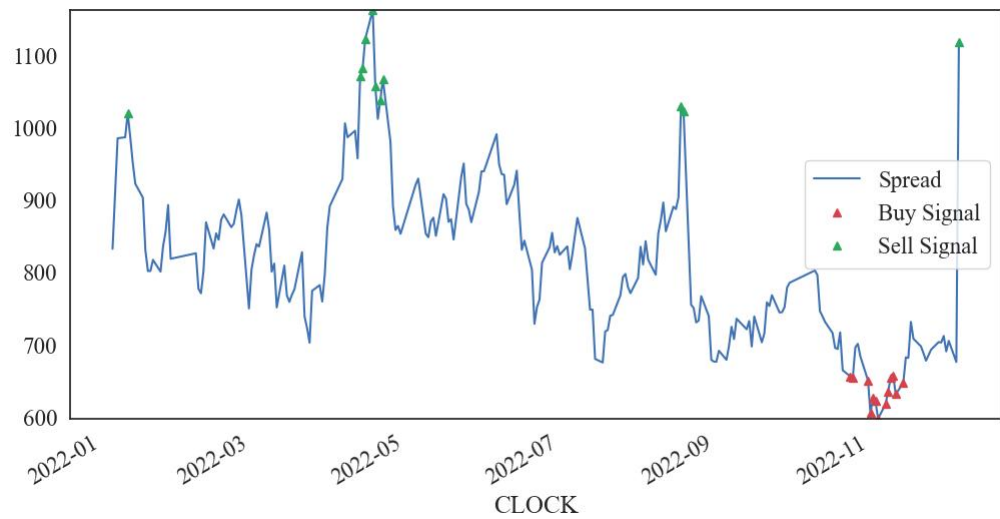


图 9：期货价格信号图

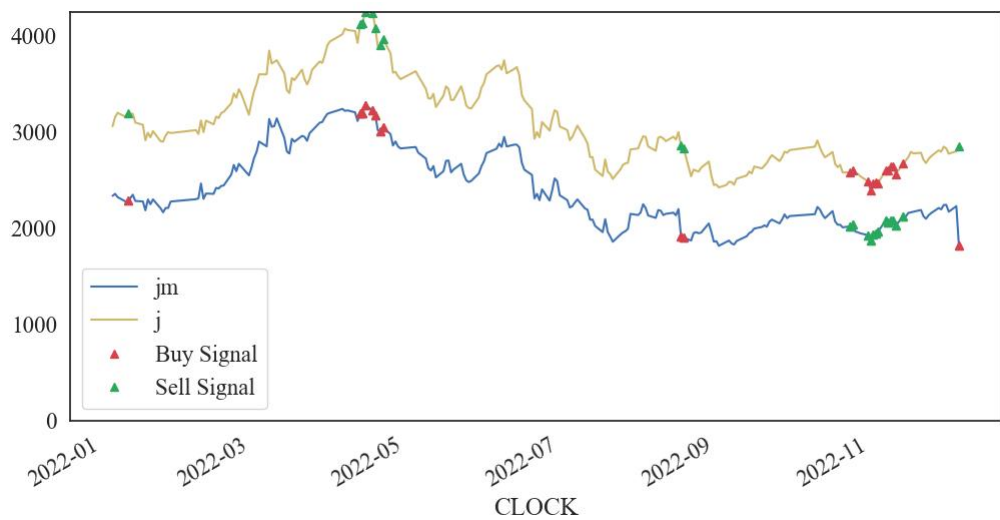
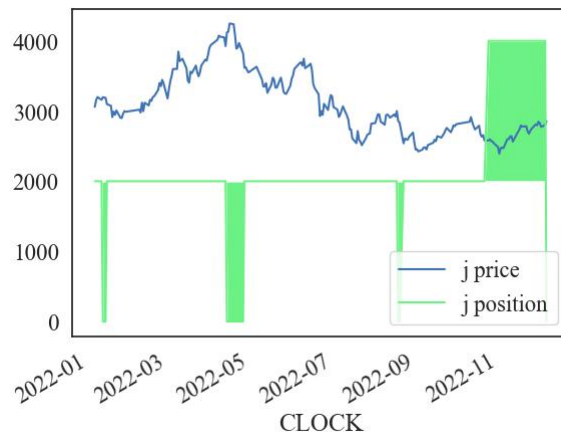
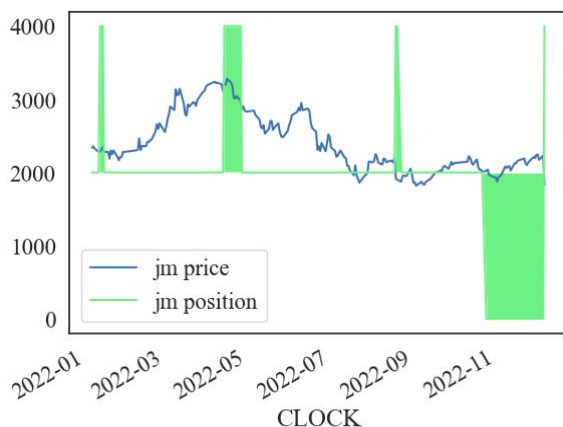


图 10：焦煤模拟仓位图

图 11：焦炭模拟仓位图



从信号模拟图可以看出价差指标触发的交易次数较少,从仓位模拟图中可以看出 2022 年 11 月的交易持仓周期较长,从图 7 可以看出,从 9 月开始价差均值水平下移到 700 左右,原来的价差下限和价差的下正常轨道处于高估水平,因此在触发做空价差时,长时间没有办法触发平仓信号。如果此时基本面进一步恶化,价差持续低位,可能会引发更长周期的持仓和回撤。

3.3 仓位管理方法

3.3.1 信号描述:

在开盘时计算焦煤和焦炭的价差:

如果 $\text{Spread} < \text{价差下轨}$, 做多

如果 $\text{Spread} > \text{价差上轨}$, 做空

如果 Spread 回归价差均值正常水平, 平仓

3.3.2 仓位方法描述:

账户初始资产为 1500W

表 4 每日交易资金和仓位

信号	交易品种	每次交易金额	仓位 (手)
做多价差	做多焦炭期货 (j)	2500W	$2500 W / (P2 * \text{合约乘数})$
	做空焦煤期货 (jm)	$2500W / Pj * \text{配比} * Pjm$	$-2500 W * \text{配比} / (P2 * \text{合约乘数})$
做空价差	做空焦炭期货 (j)	2500W	$-2500 W / (P2 * \text{合约乘数})$

	做多焦煤期货 (jm)	$2500W / P_j^* \text{ 配比} * P_{jm}$	$2500 W^* \text{ 配比} / (P_2^* \text{ 合约乘数})$
平仓	清空焦煤焦炭	根据当前开盘价和仓位	0

3.3.3 开平仓交易日志

时间为: 2022-01-11 00:00:00,差价太大, S1 低估, S2 高估, 买入S1: 124 手, 卖出S2: 78. 手, 现有S1 头寸为: 124, 现有S2 头寸为: -78, 买卖差为: 7975108.143

时间为: 2022-01-14 00:00:00,价差回归, 平仓, 现有S1 头寸为: 0, 现有S2 头寸为: 0,买卖差为: -7212360.575

时间为: 2022-04-13 00:00:00,差价太大, S1 低估, S2 高估, 买入S1: 96 手, 卖出S2: 61 手, 现有S1 头寸为: 96, 现有S2 头寸为: -61,买卖差为: 6497331.923

时间为: 2022-04-26 00:00:00,价差回归, 平仓, 现有S1 头寸为: 0, 现有S2 头寸为: 0,买卖差为: -5411055.940

时间为: 2022-08-18 00:00:00,差价太大, S1 低估, S2 高估, 买入S1: 139 手, 卖出S2: 87 手, 现有S1 头寸为: 139, 现有S2 头寸为: -87,买卖差为: 9012963.220

时间为: 2022-08-22 00:00:00,价差回归, 平仓, 现有S1 头寸为: 0, 现有S2 头寸为: 0,买卖差为: -6606384.610

时间为: 2022-10-24 00:00:00,差价太小, S1 高估, S2 低估, 卖出S1: 154 手, 买入S2: 97 手, 现有S1 头寸为: -154, 现有S2 头寸为: 97,买卖差为: -6355748.159

绝对收益率: 0.550190747291337

年化收益率: 0.5976774486944584

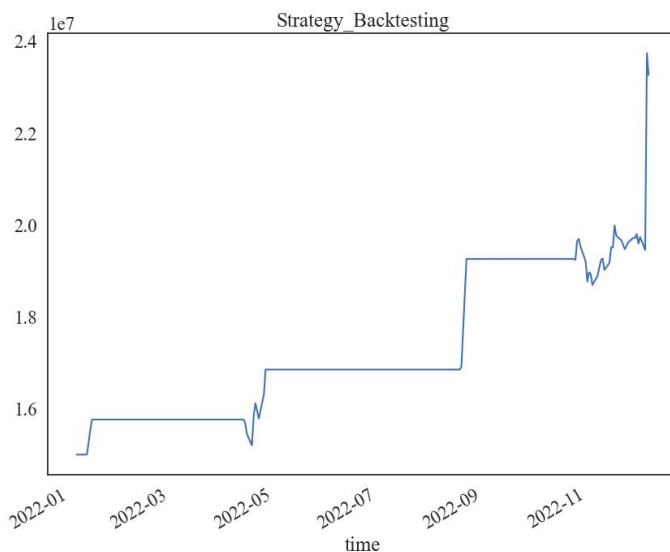
最大回撤: 0.050993022790226494

夏普比率: 2.1820685824902557

Calmar: 11.720769156069943

平均持仓周期: 6.7 天 (未包含最后未平仓头寸)

图 12：传统方法焦炭焦煤套利净值曲线



4 滚动回归套利

以上是传统的跨品种套利，根据两个品种历史价格的相关关系设定一个固定的配比以及均值回归水平进行套利。以第三节的焦煤焦炭套利为例，根据 2022 年 1 月以前的相关关系，确定一吨焦炭配 0.952 吨焦煤，差价均值为 838.76，且价差均值上下限都为固定的数字。这种传统的方法的效果并不好，焦煤和焦炭的价格比例并不会保持在过去的 1:0.952，差价的均值水平也并不会保持在过去的 838.76，另外均值上下限也不应该固定不变。基本面的变化会导致差价的上下限超出正常的水平，从而导致配对策略不能及时抓住差价的变化，从而导致策略的回撤较大。

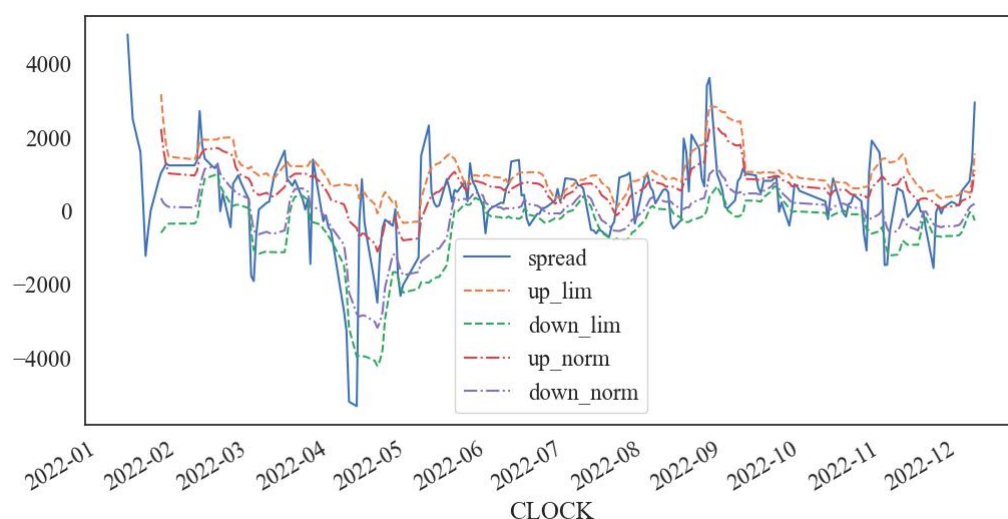
为了克服传统方法的缺陷，对以下三方面进行了改进：

- (1) 价格比例通过滚动回归获得，最优滚动窗口由敏感性分析确定；
- (2) 价差的均值水平由过去 n 日价差平均值确定， n 由敏感性分析确定；
- (3) 价差的上下限改为价差 n 日均值加减 1 倍标准差，价差上下正常水平为 0.5 个标准差。

4.1 滚动回归窗口 5 价差均值窗口 10

在保持均值回归思想不变下，仍然以焦煤焦炭进行跨品种套利，回测时间为 2022 年 1 月 5 日到 2022 年 12 月 7 日。设置滚动回归窗口为 5，价差均值窗口为 10。每日的价格配比由今日开盘价以及前四天开盘价确定，并由滚动回归得到的价格配比计算每日价差。其中，价差均值为 10 日价差的平均水平，价差均值上下限为 10 日价差均值加减 10 日价差的 1 倍标准差，价差均值上限正常范围为 10 日价差均值加减 10 日价差的 0.5 倍标准差。

图 13：滚动回归价差及信号图



从图 14 和 15 可以看出来，通过滚动回归不断更新配比、价差、价差上下限和价差的正常范围，在交易逻辑不变的情况下，可以得到更多的开平仓信号。

图 14：价差信号图

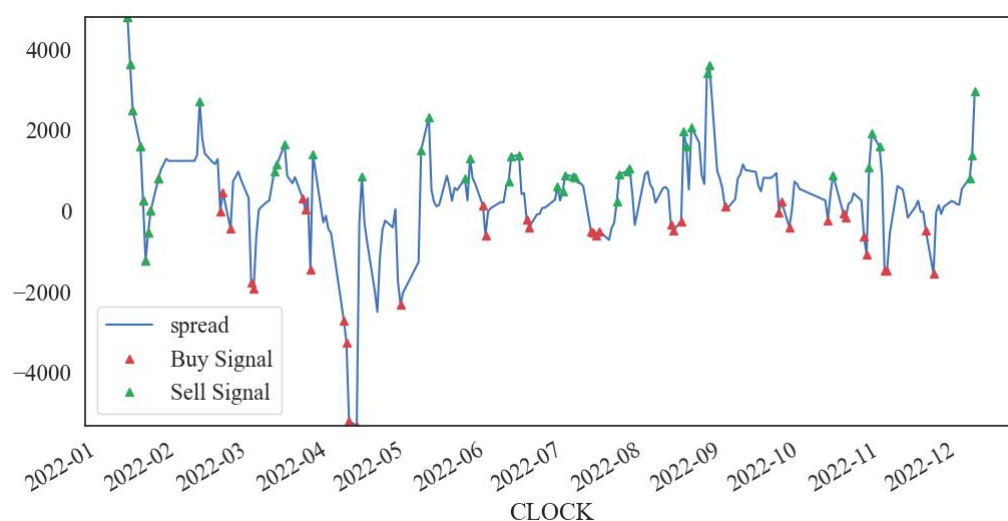
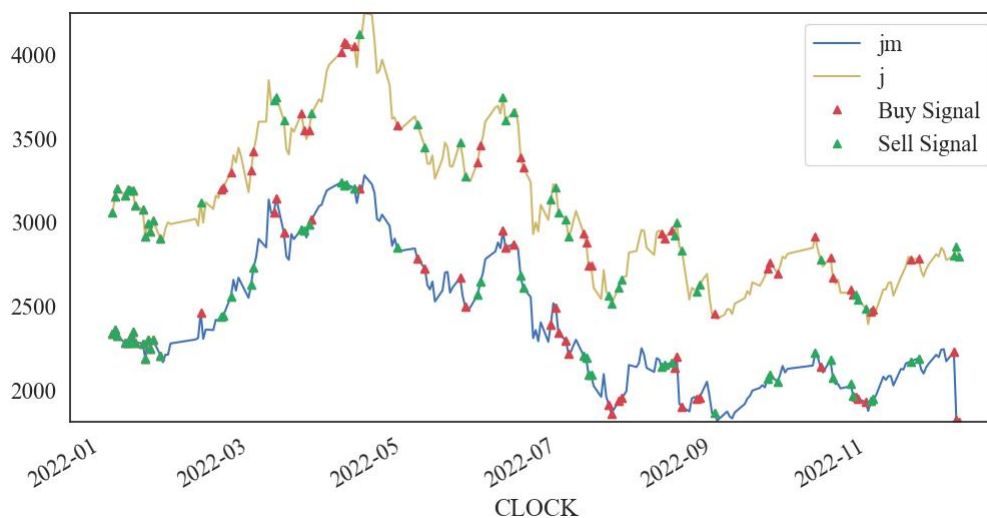


图 15：期货价格信号图



从图 16 和图 17 的仓位模拟可以看出，滚动回归的方法不仅信号更多，而且平均持仓周期更短。

图 16：焦煤模拟仓位图

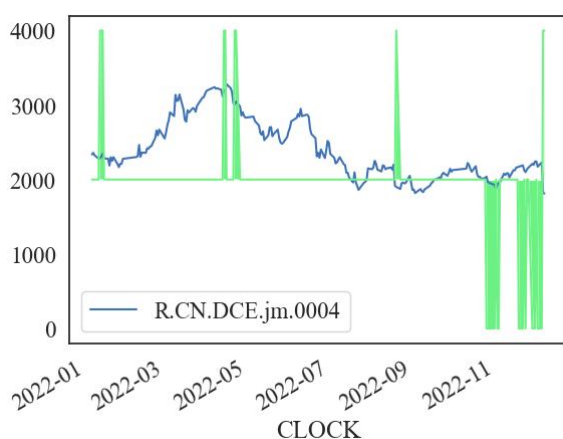
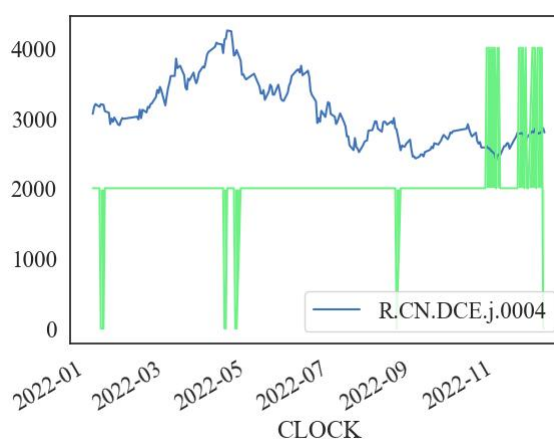


图 17：焦炭模拟仓位图



4.2 策略表现

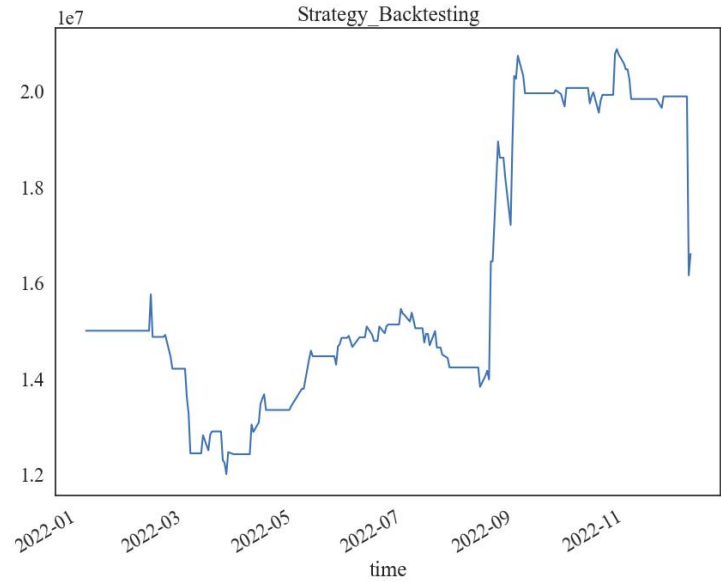
回测时间从 2022 年 1 月 5 日到 2022 年 12 月 7 日，持续时间为 336 天，总交易次数 23 次（未算最后没平仓头寸），盈利次数 12 次，最大回撤为 23.82%，是由于 2022 年 12 月 5 日认为价差太大焦煤被低估，从而做多焦煤做空焦炭，但是后两天焦煤突然进一步下跌，导致账户产生了很大的浮亏具体。回测结果如下表：

表 5 回测结果

回测指标	具体数值	回测指标	具体数值
持续时间/天	336	绝对收益	10.69%
开始时间	2022-01-05	年化收益	11.61%

结束时间	2022-12-07	最大回撤	23.82%
总交易次数	23 次	夏普比率	0.58
总盈利笔数	12 次	Calmar	0.48

图 18: 传统方法焦炭焦煤套利净值曲线



4.3 参数敏感性分析

该策略有两个参数是不确定的，一个是为了获得比例系数需要进行回归的滚动窗口长度，第二个是计算多日价差均值的窗口长度。为了得到比较好的策略结果，对策略的两个参数进行敏感性分析，并从中选出表现较好的策略参数。

图 19 不同参数下夏普比率

回归\价差均值窗口	10	20	30	40
5	0.58	0.24	1.19	1.28
15	0.61	(0.06)	1.56	1.19
25	0.76	0.27	1.56	0.95
35	1.08	0.27	1.25	0.87
45	1.21	0.07	1.02	0.01
55	0.99	0.07	0.99	0.61

图 21 不同参数下最大回撤

图 20 不同参数下年化收益率

回归\价差均值窗口	10	20	30	40
5	11.61%	0.45%	28.01%	28.25%
15	12.54%	9.00%	35.54%	25.75%
25	17.35%	1.64%	35.45%	19.88%
35	26.83%	1.54%	27.82%	17.74%
45	30.27%	-4.32%	22.10%	9.62%
55	23.58%	-4.32%	20.80%	11.15%

图 22 不同参数下胜率

回归\价差均值窗口	10	20	30	40
5	-23.82%	-28.69%	-13.15%	-14%
15	-22.47%	-30.64%	-12.52%	-14.54%
25	-21.78%	-28.46%	-12.53%	-15.13%
35	-20.54%	-28.48%	-13.17%	-15.35%
45	-20.13%	-29.64%	-13.69%	-16.27%
55	-20.95%	-29.64%	-13.82%	-16.09%

回归\价差均值窗口	10	20	30	40
5	0.57	0.57	0.50	0.53
15	0.59	0.55	0.53	0.47
25	0.62	0.55	0.50	0.50
35	0.63	0.53	0.47	0.47
45	0.67	0.50	0.44	0.38
55	0.65	0.50	0.40	0.42

从以上四个图可以看出，在价差均值窗口不变下（40 除外），随着滚动回归的窗口越长，，夏普比率、年化收益率和最大回撤表现都是先变好后变差，在固定价差均值窗口为 10 下，胜率随着回归窗口拉长胜率变高，而将价差均值固定在 20、30、40 水平下，胜率反而呈现单调递减的情况。在滚动回归窗口不变时，夏普、年化收益率和最大回撤表现也随价差均值窗口拉长先变好后变坏，而胜率呈现价差均值窗口越长胜率越低，这个结果刚好与回归窗口相反。

结合两个参数在四个维度的表现，其中滚动回归窗口 15、价差均值窗口 30 的表现最好，夏普比率为 1.56，年化收益率为 35.54%，最大回撤为-12.52%，胜率为 53%。

4.4 不同配对测试

在做动力煤和硅铁配对时，发现程序无法完成滚动线性回归，经调试发现动力煤的价格于 2022 年 10 月 31 日就维持在 898.8 水平不变，长达 22 个交易日，导致在以 15 的滚动窗口下无法完成回归，因此将该不活跃期货对剔除回测。

不同的配对上对策略进行测试，以如下形式呈现：

表 6 回测结果

配对	年化收益	最大回撤	Sharpe	Calmar	胜率
SF-JM	10.5%	27.63%	0.55	0.38	6/14
SF-J	35.73%	17.36%	1.13	2.06	8/14
RB-HC	8.6%	5.57%	0.83	1.55	8/11
TA-SA	0.78%	45.54%	0.27	0.017	4/10
TA-BU	20.39%	28.94%	0.91	0.70	8/13
MA-SA	17.70%	29.27%	-0.27	-0.60	8/14
EG-V	23.39%	16.51%	1.18	1.42	6/9
PP-RU	-4.13%	19.67%	-0.06	-0.21	5/10
OI-A	-5.05%	15.18%	-0.41	-0.33	6/12

OI-Y	-38%	47.34%	-2.07	-0.89	5/13
OI-P	-87.38%	90%	-1.43	-0.97	2/9
CS-A	-5.74%	43.75%	-0.44	-0.34	3/8
CS-Y	51.23%	15.59%	1.81	3.29	3/7
CS-JD	-17.40%	24.72%	-0.55	-0.70	6/10
A-Y	-5.49%	19%	0.04	-0.29	5/11
A-P	-25.2%	78%	41.55%	-0.32	5/10
A-JD	-17.65%	27.63%	-0.51	-0.64	4/10

5 总结

从上述测试中可以发现，黑色系和能源化工类的跨品种套利是有效的，而农产品中除了玉米淀粉和豆油有比较好的表现之外，其余的跨品种配对的年化收益都为负，而且回撤非常大，最高达到了 90%。对菜籽油(10)和棕榈油(10)进一步检验发现，当把滚动窗口拉长至 65，差价均值窗口不变 (30)，策略的年化收益变成了 30%，最大回撤为 22.39%。通过进一步检验，发现在部分时期二者呈同向变动，但在 2021 年 6 月到 2021 年 12 月，二者呈现出反向变动。基本面的剧烈变化，导致短回归窗口下策略表现很差。同样将玉米淀粉和鸡蛋配对的回归窗口改到 75 时，策略的年化收益从-17.4%增加到了 21.49%，但是交易次数明显降低，从 10 次变成了 3 次。但是胜率有 2/3。通过对农产品配对进一步测试发现，农产品里对跨品种配对更适合用长回归窗口。可能和基本面变化较大，配对关系不稳定有关。

总体来看，短回归窗口在黑色系和能源化工两大类的配对效果不错，而大部分农产品配对更适合长回归窗口进行配对。

由于本策略采用的是主力合约的复权价格，在计算盈亏时与真实情况存在一定出入。下一步将改进为使用复权价格产生交易信号，用真实价格计算盈亏。同时进一步考虑交易成本、杠杆和合约展期的问题，并在配对之前，根据流动性对合约进行初步筛选，避免产生与动力煤和硅铁同样对问题。