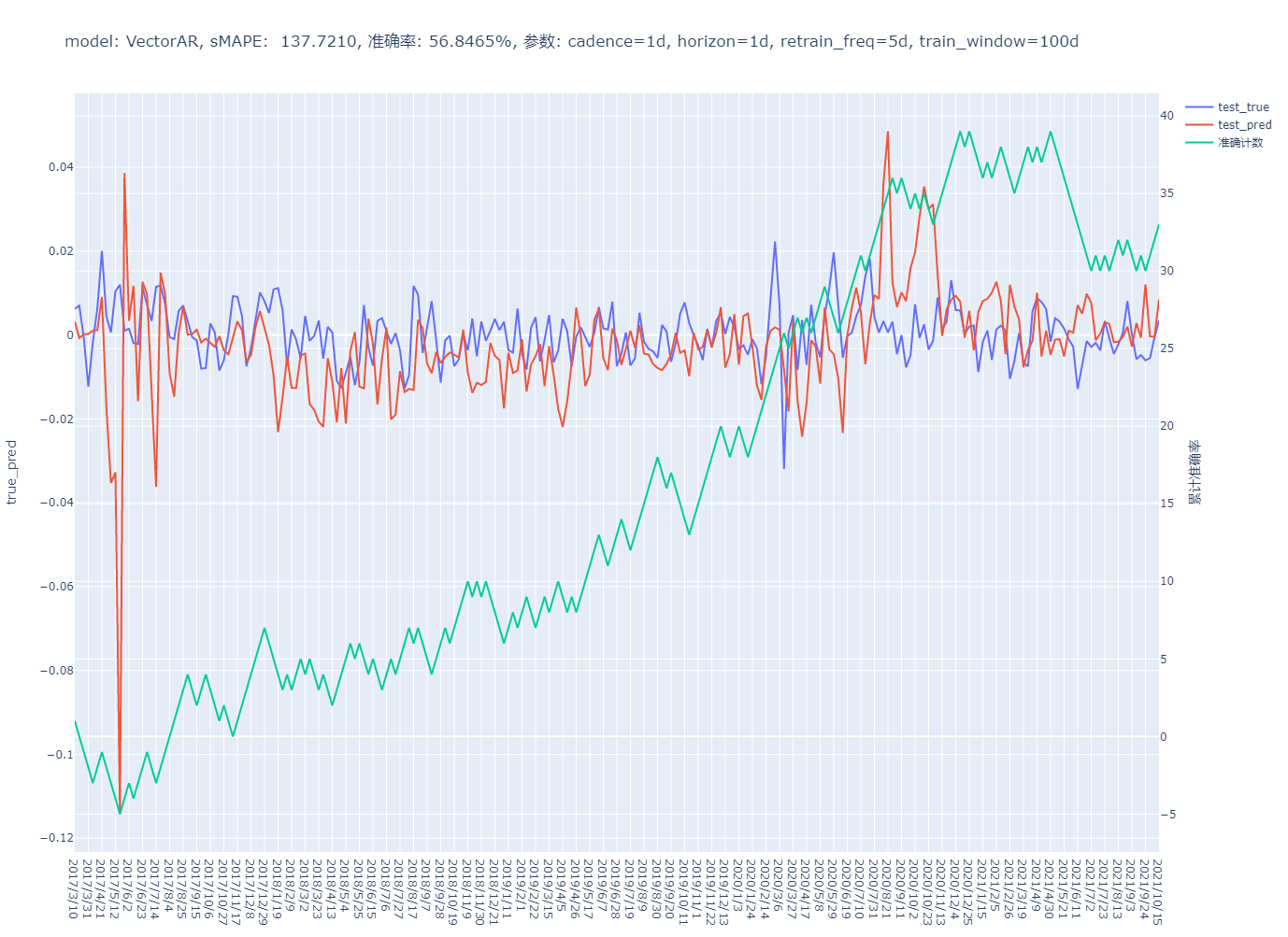
# 第1章 基于Merlion工具的多变量预测模型结果

## 1.1 Merlion工具包基本设置

* 选择因子。赵禹平提供原始数据（周频），不同模型因子不同，通过循环遍历选择。
* 训练集和测试集比例。7:3。
* 构建模型。直接调用模型相关接口，Merlion工具包对多变量预测一共提供了四种模型：Vector\_ar，ExtraTrees，RandomForest，LGBM。
* 模拟部署。Merlion工具包提供ForecastEvaluator接口，可以直接通过参数设置模型每次训练集数据量，每隔多久训练一次，预测未来几期等参数。

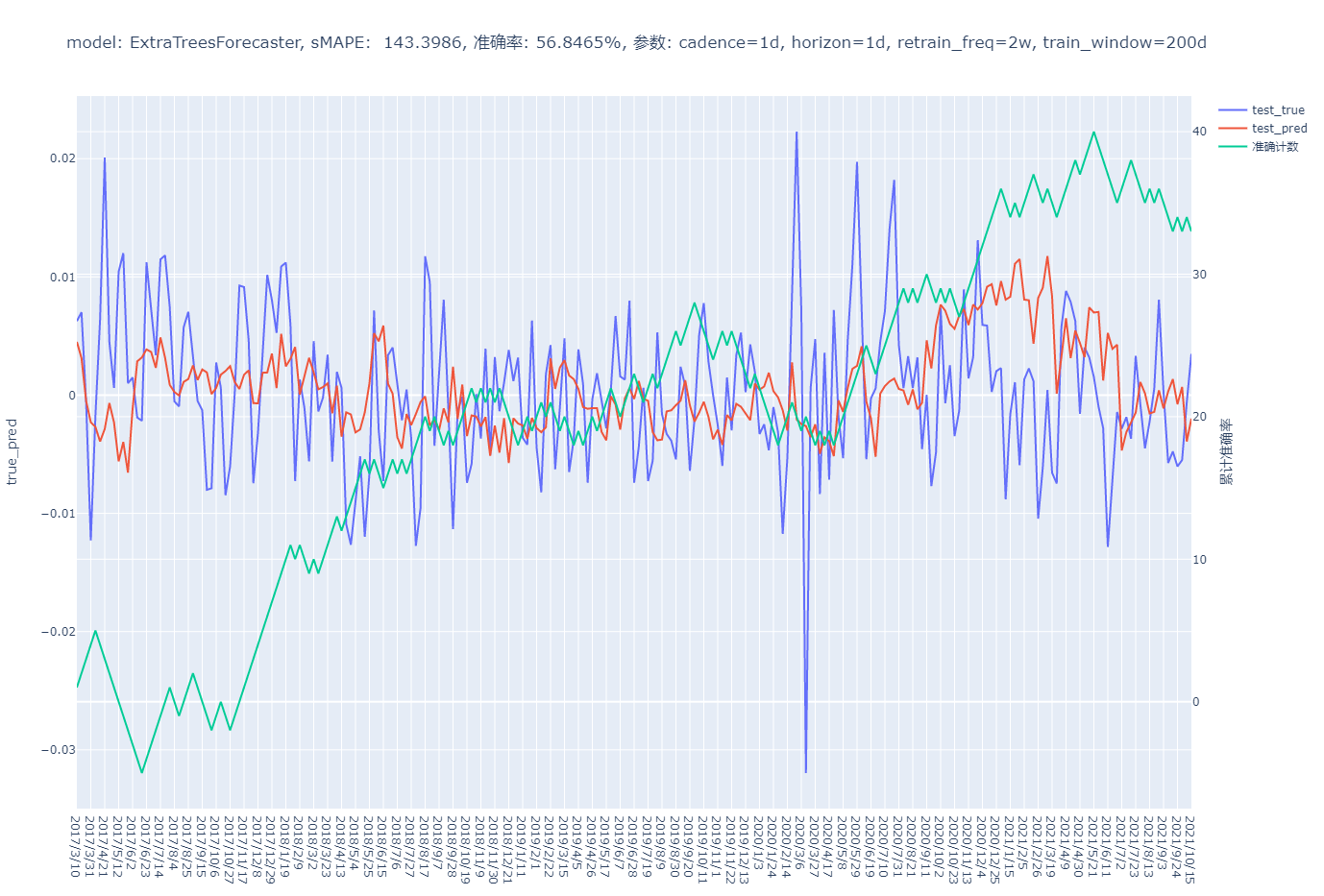
## 1.2 Vector\_ar模型

* 选择因子['NonRept\_SI\_traders', 'NonComm\_SI\_report', 'NonComm\_Net\_Pct\_report','NonComm\_Net\_Position\_report']
* 每5期重新训练一次模型，模型训练使用过去100期数据，每次预测未来1期数据
* 结果误差sMAPE: 137.721
* 方向准确率: 0.5684



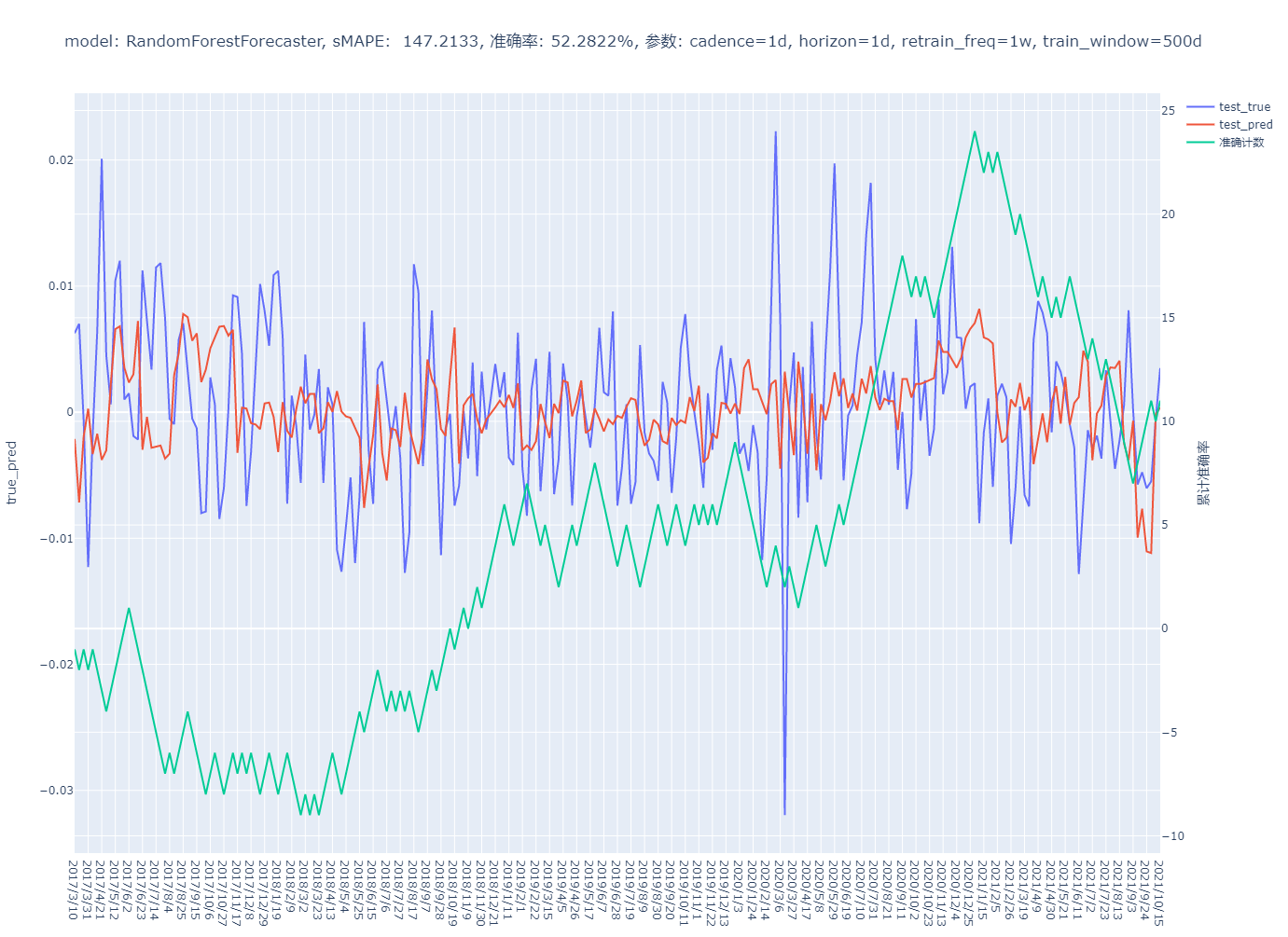
## 1.3 ExtraTrees模型

* 选择因子['NonRept\_SI\_traders', 'Asset\_Mgr\_Net\_Position\_traders', 'NonRept\_COT\_index\_report', 'Comm\_COT\_index\_report', 'Comm\_SI\_report', 'Comm\_Position\_Pct\_report', 'Comm\_Net\_Pct\_report', 'NonComm\_COT\_index\_report', 'NonComm\_Net\_Pct\_report', 'NonComm\_Net\_Position\_report', 'label']
* 每14期重新训练一次模型，模型训练使用过去200期数据，每次预测未来1期数据
* 结果误差sMAPE: 143.3986
* 方向准确率: 0.5684



## 1.4 RandomForest模型

* 选择因子['NonRept\_SI\_traders', 'NonRept\_Net\_Position\_traders', 'Dealer\_Net\_Pct\_traders', 'NonComm\_SI\_report', 'NonComm\_Position\_Pct\_report', 'NonComm\_Net\_Position\_report', 'label']
* 每7期重新训练一次模型，模型训练使用过去500期数据，每次预测未来1期数据
* 结果误差sMAPE: 147.2133
* 方向准确率: 0.5228



## 1.5 LGBM模型

* 选择因子['Lev\_Money\_Position\_Pct\_traders', 'NonRept\_Net\_Pct\_report', 'NonComm\_COT\_index\_report', 'NonComm\_SI\_report', 'NonComm\_Position\_Pct\_report', 'NonComm\_Net\_Pct\_report', 'NonComm\_Net\_Position\_report', 'label']
* 每5期重新训练一次模型，模型训练使用过去100期数据，每次预测未来1期数据
* 结果误差sMAPE: 138.9117
* 方向准确率: 0.5809

