



华泰期货量化专题报告：

利用限制性多项式改进混频时间序列模型

混频时间序列模型简介

金融时间序列通常包含多个维度，不同维度数据的采样频率也不一致。例如螺纹钢研究员通常关心螺纹钢的因素有日频更新的现货螺纹钢价格，周频更新的螺纹钢库存，高炉开工率和线螺采购量。如果其中某些可观测因子发生了变化，投资者对未来螺纹钢期货涨跌的预期也应发生变化，但是如何处理这些不同频率的数据是量化模型的一大难题。本系列的量化专题报告里面曾介绍一种可利用低频数据预测高频数据的时间序列模型：RU-MIDAS (Reverse Untricted Mixed Data Sampling)，中文可以译作反无限限制混合数据采样模型，这种模型是 Eric Ghysels 等人研究的 MIDAS 模型的反向应用，MIDAS 模型则是利用高频数据预测低频数据，而 RU-MIDAS 则是利用低频数据预测高频数据，因此被称为反面。

RU-MIDAS 模型并没有对高频延时项系数数量和分布形态进行限制，如果包含的高频延时项过多，则会大大增加 RU-MIDAS 模型的待调参数，如果样本数量不够多则和很容易造成过度拟合的现象。因此 Eric Ghysels 等人通过引入 MIDAS 多项式来代替高频项系数，这种方法通常能把高频项系数的可调参数降低到 2-3 个。然而 Claudia Foroni 等人在论文 Using Low Frequency Information for Predicting High Frequency Variables 中提出如果低频和高频数据的采样频率差别不大，例如季频和月频的混合，一个季度只有 3 个月，这种情况下使用 MIDAS 多项式进行参数限制意义不大，因为高频项的回溯期通常小于等于 3 个月，即其系数会小于等于 3，这时高频项参数数量与多项式参数数量已经相等了，引入多项式进行限制只会降低模型的灵活性，也起不了减少参数数量的作用。Claudia Foroni 等人还发现如果是季频和周频数据混合则 MIDAS 多项式对改进模型效果仍是有意义的。本系列专题主要研究期货交易中日频和周频基本面数据的混合，一周中通常包含 5 个交易日，而对于日频和周频数据是否应该使用 MIDAS 多项式对高频项系数进行限制这个问题，Claudia Foroni 等人的论文并没有进行研究，所以这将是本报告研究的重点。由于 Claudia Foroni 等人提出的 RU-MIDAS 模型并没有使用限制多项式，所以名称中包含了字母“U”，即 unrestricted，在本报告中使用了 MIDAS 多项式的模型将被称为 R-MIDAS。

在本报告里首先介绍了 R-MIDAS 模型的原理和结构，然后尝试利用周频和日频的螺纹钢因子数据对螺纹钢期货主力数据和其他品种数据进行预测，发现 R-MIDAS 模型在螺纹钢等品种上能取得更好的效果。

华泰期货研究所 量化策略组

陈维嘉

量化研究员

☎ 0755-23991517

✉ chenweijia@htfc.com

从业资格号：T236848

投资咨询号：TZ012046

相关研究

利用混频 RU-MIDAS 多因子模型预测期货日频收益率

2018-03-15

R-MIDAS 研究背景

金融时间序列通常包含多个维度，不同维度数据的采样频率也不一致。例如螺纹钢研究员通常关心螺纹钢的因素有日频更新的现货螺纹钢价格，周频更新的螺纹钢库存，高炉开工率和线螺采购量。如果其中某些可观测因子发生了变化，投资者对未来螺纹钢期货涨跌的预期也应发生变化，但是如何处理这些不同频率的数据是量化模型的一大难题。目前以基本面数据为基础的量化策略的预测频率大多数是与所采用的基本面数据的最低更新频率一致，如果基本面数据是每周更新的则策略的预测频率为周，并持仓一周，如果是按月更新则预测频率为月，并持仓一个月。由于受制于预测频率和持仓时间，基于基本面数据的量化策略往往无法适应市场状况而及时做出调整。因此本系列研究如何把相对较低更新频率的基本面数据应用到较高的预测频率上，这就不可避免地涉及到混频数据的处理。

大量学术研究表明，利用高频变量能够有效提高低频变量的预测准确性。Eric Ghysels 等人研究的 MIDAS(Mi(xcd) Da(ta) S(amplng))模型就是混频数据预测领域的代表之一，这种方法原理比较简单，如下式所示

$$Y_{t+1}^Q = \mu + \sum_{j=0}^{p_Y^Q-1} \alpha_{j+1} Y_{t-j}^Q + \beta \sum_{j=0}^{q_X^Q-1} \sum_{i=0}^{N_D-1} w_{i+j*N_D}(\theta^D) X_{N_D-i,t-j}^D + \varepsilon_{t+1} \quad (1)$$

其中 Y_t^Q 是低频变量， X_t^Q 是高频变量， p_Y^Q 为低频变量的延时， p_X^Q 为高频变量的延时， μ 为截距， ε_{t+1} 是误差项。高频变量 X_t^Q 前的系数使用了多项式 $w_{i+j*N_D}(\theta^D)$ 表示，原本这些系数都是待校正参数，但是由于使用了多项式 $w_{i+j*N_D}(\theta^D)$ ，待校正参数的数量就大大降低了，从而有效防止参数过多造成的模型过度拟合。MIDAS 模型通常使用在利用高频数据预测季频 GDP 上。在相同领域预测比较成功的还有状态空间(state-space)模型，但这种模型实现比较复杂，有关研究表明 MIDAS 在 GDP 上的预测效果通常不会比状态空间模型差很多。

目前学术界在混频预测上研究得比较多的仍然是如何利用高频变量预测低频变量，但是在期货交易上不少投资者更关心的可能是投资策略的短期回撤，而非长期回报。如果期货市场波动剧烈，偏高频的投资策略往往回撤更小，这就要求量化模型对高频变量具备预测能力。目前学术界的量化模型在利用低频变量预测高频变量上的研究并不多，之前的报告有参考 Claudia Foroni 等人在 Using Low Frequency Information for Predicting High Frequency Variables 上提出的 RU-MIDAS (Reversed Unrestricted-MIDAS)模型，进行改进，尝试利用螺纹钢的周频基本面数据预测螺纹钢主力期货的日频收益率。Claudia Foroni 等人研究的是单因子 RU-MIDAS 模型，其形式与公式(1)类似，只是反过来用，在之前的报告中我们把 Claudia Foroni 等人的单因子 RU-MIDAS 扩展到多因子上，可用如下公式表示

$$\tilde{A}_i(L)X_t = b_i(L^{k-i})Y_t + \xi_{it} \quad (2)$$

其中 X_t 为高频变量向量, Y_t 为低频变量向量, $\tilde{A}_i(L)$ 为周期矩阵, $b_i(L^{k-i})$ 为周期向量, 由于考虑到高频变量之间的协相关性, 公式(2)可用广义最小二乘法 GLS 进行参数估计。

其中 x_t 为高频变量, y_t 为低频变量。 $t = 0 + \frac{i}{k}, 1 + \frac{i}{k}, 2 + \frac{i}{k}, \dots; i = 0, \dots, k-1$ 。 k 为低频变量包含的高频变量周期数, 例如当周频和日频混合时 $k = 5$ 。 $\tilde{a}_i(L)$ 和 $b_i(L^{k-i})$ 为延时线性多项式, 例如 $\tilde{a}_i(L) = \tilde{a}_{i1}L + \tilde{a}_{i2}L^2 + \dots + \tilde{a}_{ip}L^p$ 。 因此 $\tilde{a}_i(L)$ 和 $b_i(L^{k-i})$ 都包含周期结构, 即周一至周五都有不同的多项式系数。 L 为高频延时算子, $L^j x_t = x_{t-j/k}$ 。 值得注意的是在应用中 $b_i(L^{k-i})$ 的选择可以使得 $b_i(L^{k-i})y_t$ 只包含 y_t 的低频项, 即无需通过插值等方法从低频的 y_t 获得高频的 y_t^* 。 这个模型的本质是认为当延时足够大时, 高频变量 x_t 可以看作是一个自回归 AR(p) 过程, 而低频变量 y_t 则被看作是外生变量。

公式(2)其实已经相当于向量自回归了模型(Vector Autoregression, VAR)。 在混频数据处理领域, 一种常见的方法是由 Eric Ghysels 等人首先提出的 stacked VAR 法, 这种方法是把同一低频时间节点上的高频变量时间序列和低频变量堆积(stacked)在同一个向量里, 延时项则是上一个低频时间节点上的堆积向量。 Heiner Mikosch 等人在 Real-Time Forecasting with a MIDAS VAR 里对 Eric Ghysels 等人的研究进行了一定拓展, 他们使用 Almon 多项式代替高频延时项的系数, 这个多项式的求解无需进行非线性迭代, 求解非常方便。 所以本报告也采用了 Heiner Mikosch 等人的做法, 但是作出适当简化。 在 Eric Ghysels 和 Heiner Mikosch 等人的 stacked VAR 方法里面, 低频变量是和高频变量一样作为内生变量处理的, 这造成了模型结构复杂和参数过多等问题, 而且对低频变量的预测也并非这里的重点。 所以这篇报告里仍然是把低频变量当作外生变量处理。 因此 R-MIDAS 模型仍然是保持公式(2)所使用的架构, 而只是对周期性矩阵 $\tilde{A}_i(L)$ 的矩阵系数添加 Almon 多项式进行约束。 Almon 多项式是一个线性系统, 多项式在各延时项 i 上的权重 w_i 可用如下式子表示

$$w_i(\gamma_0, \dots, \gamma_P) = \sum_{p=0}^P \gamma_p i^p \quad (3)$$

其中 $\gamma_0, \dots, \gamma_P$ 是 Almon 多项式的待定参数, 通常为了使 w_i 在各延时项上呈曲线分布 $P \geq 3$ 。 公式(3)可以用矩阵的形式更直观地展示

图 1: Almon 多项式

$$\begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2^2 & \dots & 2^P \\ 1 & 3 & 3^2 & \dots & 3^P \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & N & N^2 & \dots & N^P \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_P \end{bmatrix}$$

数据来源: Eric Ghysels, Macroeconomics and the Reality of Mixed Frequency Data

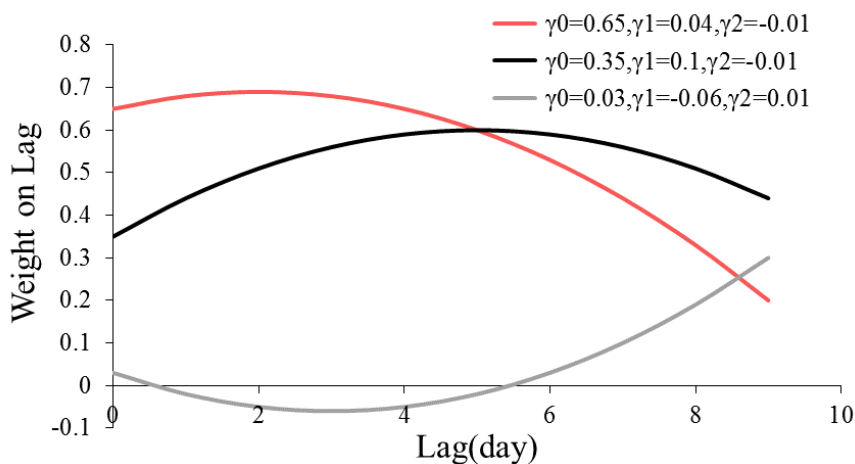
上图中 Almon 多项式的矩阵系数已经固定了, 只与延时数 $N + 1$ 和待定参数数量 P 有关。通过 Almon 多项式约束, R-MIDAS 模型中高频项的待定参数数量就由 $N + 1$ 下降到 P 了。R-MIDAS 模型可以写作如下形式:

$$\tilde{W}(L)\Gamma_i X_t = b_i(L^{k-i})Y_t + \xi_{it} \quad (4)$$

其中 $\tilde{W}(L)$ 就是固定的 Almon 多项式, Γ_i 为高频项的待定参数。

下图是 Almon 多项式取不同参数值 γ 时在不同延时项上的权重分布。这里参数值 γ 的数量只取到 3, 延时只取到 10。由图可见, Almon 多项式的权重取值非常灵活, 红色线所显示的一组参数值, 权重是随着延时的增加而减少, 黑色线那组则是先增后减, 灰色线那组是先减少再增加。这三组曲线都有类似于抛物线的特点, 并且只有一个拐点。

图 2: Almon 多项式



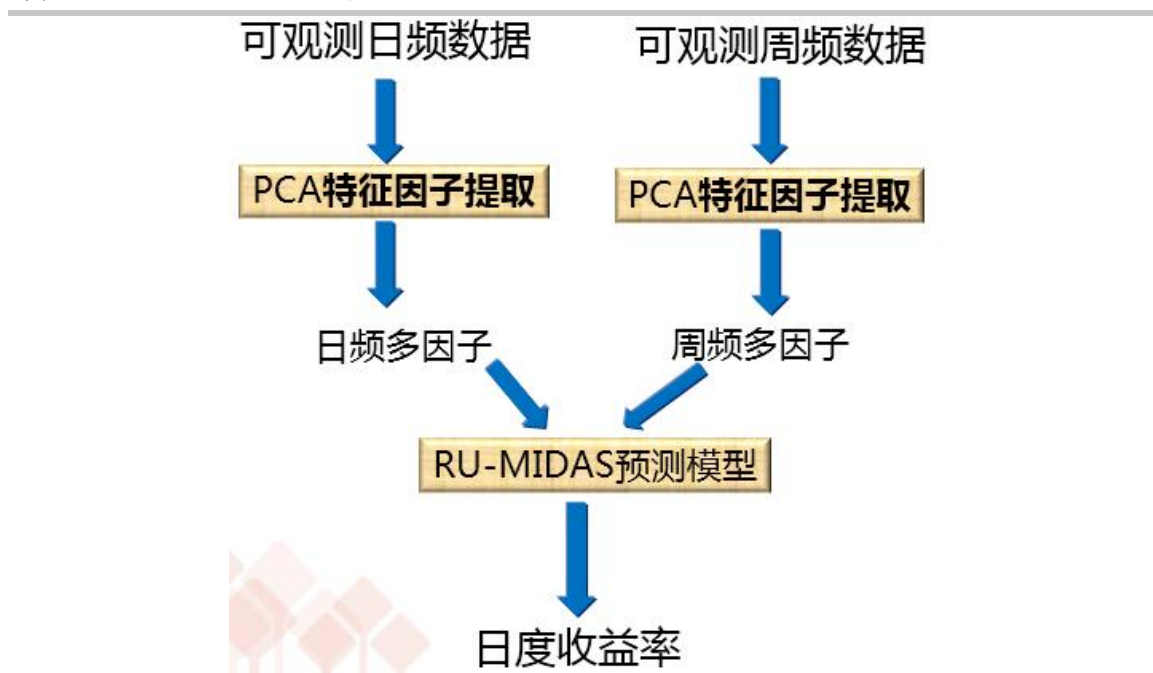
数据来源: 华泰期货研究院

下面就以螺纹钢、焦炭、PVC、甲醇和豆油这 5 个品种为代表, 对比 RU-MIDAS 和 R-MIDAS 模型在添加限制性 Almon 多项式后的预测效果。由于螺纹钢交易最为活跃, 因此本报告以螺纹钢会作为重点研究对象。

R-MIDAS 模型应用实例

在这一节里首先尝试使用 R-MIDAS 预测螺纹钢期货主力合约的日频收益率。预测是基于每周发布的螺纹钢库存、高炉开工率和钢厂产能等因子以及日度发布的螺纹钢、铁矿石、焦煤日收益率等因子。由于可观测因子的个数较多，这里先使用主成分分析(Principal Components Analysis, PCA)进行降维处理，得到的模型结构可以用下图表示

图 3: R-MIDAS 模型结构

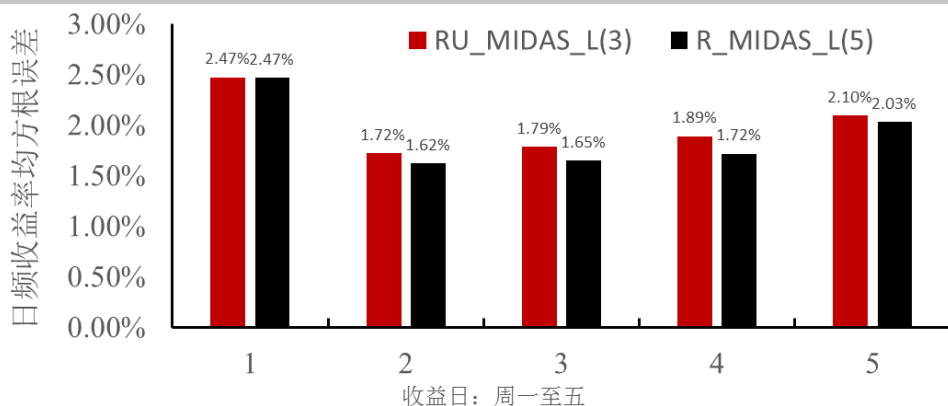


数据来源：华泰期货研究院

首先从 Wind 上获取螺纹钢相关的可观测数据, 然后对日频和周频数据分别进行 PCA 转换, 提取出主成份, 然后把日频主成份 X_t 和周频主成份 Y_t 分别放入 RU-MIDAS 模型的高频变量和低频变量当中进行模型训练。由于模型参数较多所以训练方式使用递进法, 即每过一周便把新一周归入样本中, 所以随着时间增加, 训练的样本数量也将不断增加。每训练完一周即保持模型参数, 在下一周的每个交易日进行日频收益率的预测。低频变量在每周五预测下周一收益率时进行更新, 在周二至周四预测时不更新周频变量, 而高频变量则在每日预测时都会更新。模型包含的超参数包括低频、高频变量的因子个数和延时等, 这些超参数跟据样本外的预测效果进行选择。为了方便对比, 这里选取的 RU-MIDAS 模型只包含了 3 个延时项, 而 R-MIDAS 模型则包含 5 个延时项, 因为 Claudia Foroni 等人的研究表明高频项不超过 3 项的时候 RU-MIDAS 模型的效果较好, R-MIDAS 模型则在高频项较多的时候才能体现效果, 但值得注意的是, 由于使用了 Almon 多项式, RU-MIDAS 和 R-MIDAS 模型所使用的参数数量是一样的, 因为公式(3)中的 P 值已被固定为 5。

由于 RU-MIDAS 和 R-MIDAS 模型都包含周期结构，这里首先考察该模型在周一至周五螺纹钢开盘价收益率预测的均方根误差。由图 4：可以看出这两个模型预测周一的日度收益率误差最大为 2.47%，这可能是因为隔了周六日，市场未来走势的分歧较大，所以预测收益率准确度不高。但是这两个模型周二收益率的误差随即下降，然后逐渐升高，这可能是因为使用了周频因子作为固定的外生变量，随着时间的流逝该部分信息逐步被市场消化，从而导致预测效果逐渐变差。从误差的变化趋势上看，这两个模型的表现较为一致。而且从整体上看这两个模型预测的均方根误差在 2% 左右，处在较高水平，但是使用了 Almon 限制性多项式后，R-MIDAS 模型的均方根误差有了小幅下降。

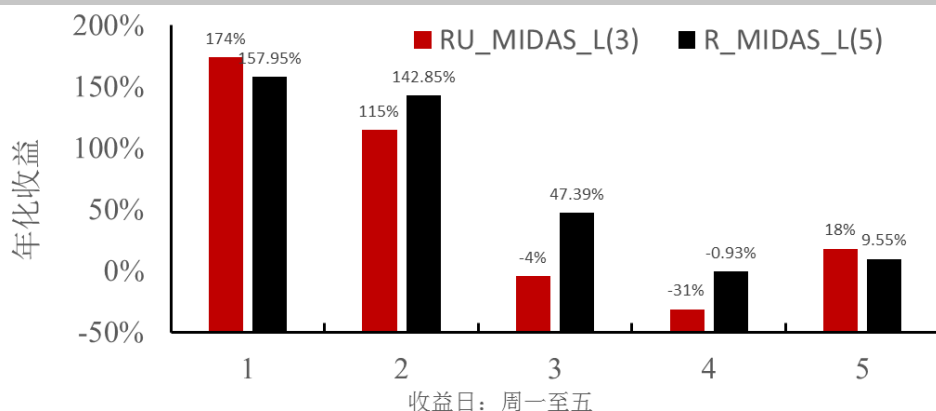
图 4： RU-MIDAS 和 R-MIDAS 日频收益率均方根误差



数据来源：华泰期货研究院

图 5：列出了该部分收益在周一至周五的分布，虽然预测周一的均方根误差较大，但是这两个模型的主要收益都来自于周一，年化收益分别高达 174% 和 157.95%，其次是周二的收益分别高达 115% 和 142.85%，周三和周四的平均年化收益 RU-MIDAS 模型均表现为负值，但是 R-MIDAS 模型则有了明显的改善，尤其是周三的收益，由 -4% 提高到了 47.39%。在周五的收益两者均为正值。总体上来说，降低收益率预测的均方根误差对提高收益仍然是有一定帮助的，但是并非收益率预测误差小，收益就一定会高。另外 R-MIDAS 模型的收益在周一至周五内分布比 RU-MIDAS 相对更为均衡，并没有出现较大的负值。

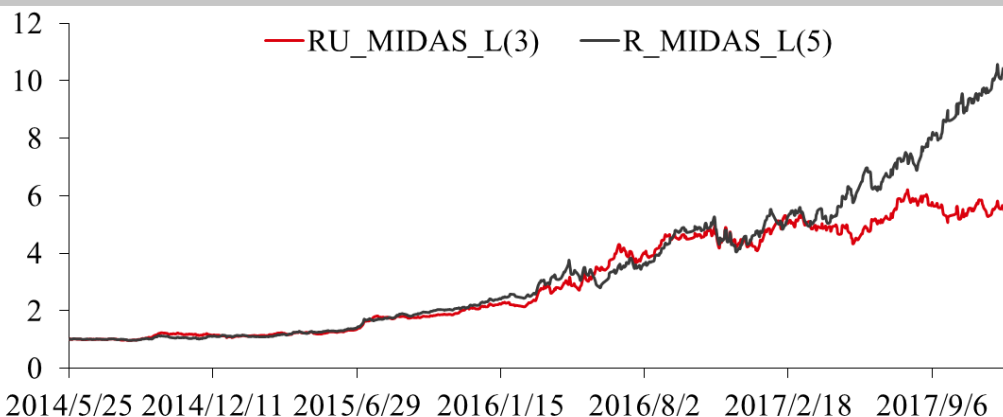
图 5: RU-MIDAS 和 R-MIDAS 在不同交易日的年化收益



数据来源: 华泰期货研究院

RU-MIDAS 和 R-MIDAS 模型的预测数据全部使用每个交易日的收盘价,但使用隔天开盘价进行交易。由于这两类模型属于日频策略交易费用的影响可能较大,这里的交易费用设定为单边万分之二。净值曲线如图 6: 所示,从中可见这两条净值曲线的走势在 2017 年 2 月之前是非常接近的,但在 2017 年 3 月之后 R-MIDAS 模型的表现则明显较好。这可能是因为 R-MIDAS 模型使用了更多的延时数从而更好地捕捉趋势性或者反转性行情。

图 6: 螺纹钢使用 RU-MIDAS 和 R-MIDAS 模型的净值曲线



数据来源: 华泰期货研究院

螺纹钢 RU-MIDAS 和 R-MIDAS 模型的表现统计如下表格 1 所示, R-MIDAS 模型的年化收益达到 67.18%, 比 RU-MIDAS 提高 18%, 在年化收益的提高上表现较为明显, 而波动率则变化不大, 两者都在 29.8% 左右。R-MIDAS 模型的夏普率提高也较为明显, 能达到 2.25。在胜率方面, R-MIDAS 模型能达到 57.2%, 比 RU-MIDAS 模型提高了 3.2%, 提高幅度较为可观, 但是 R-MIDAS 模型的回撤也较大, 达到 25.5%, 而最大回撤期则缩短为 30 个交易日。因此把 R-MIDAS 模型应用在螺纹钢上整体收益提升还是比较明显的, 虽然最大回撤有所增加, 但整体波动没有增大, 所以夏普率会有所提高。

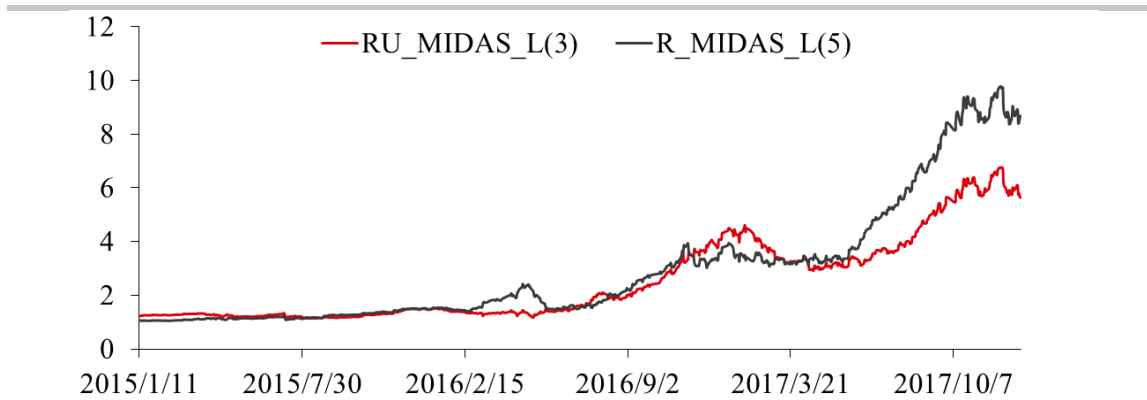
表格 1 螺纹钢 RU-MIDAS 和 R-MIDAS 模型收益与风险统计

	RU-MIDAS	R-MIDAS
年化收益	49.11%	67.18%
年化波动率	29.9%	29.8%
夏普率	1.64	2.25
胜率	53.9%	57.2%
最大回撤	19.2%	25.5%
最大回撤期	49 交易日	30 交易日

数据来源：华泰期货研究院

焦炭的净值曲线如下图所示，RU-MIDAS 模型和 R-MIDA 模型的走势基本一致，但是在不同时期这两个模型也会表现出不同的差异。例如在 2016 年 2 月中旬后 R-MIDAS 模型走势明显较好，但 5 月份后表现出回撤，净值与 RU-MIDAS 一致，随后在 2017 年 1 月底 RU-MIDAS 表现创新高时，R-MIDAS 模型一直走势平稳，直至 2017 年 5 月底 R-MIDAS 模型的收益才一路领先，到 2017 年 11 月这两个模型的走势又趋于一致。由此可见，在焦炭这个品种上，这两个模型的走势存在一定差异，在某些时期，可能某个模型收益会高些，但是从更长远的时间周期来看 R-MIDAS 模型在焦炭上表现较好，遇到较大趋势性行情时捕捉得更好。但是在 2017 年 10 月后的震荡行情中这两个模型表现也都不好。

图 7： 焦炭使用 RU-MIDAS 和 R-MIDAS 模型的净值曲线



数据来源：华泰期货研究院

这两个模型在焦炭上的表现总结如下表,使用 R-MIDAS 模型年化收益达到 57.9%,比 RU-MIDAS 模型提高 13%, 年化波动率方面则没有太大变化,都是 33.7%, 由于收益的提高, R-MIDAS 模型的夏普率也有所提高达到 1.71。而胜率的提升则不明显只有 0.1%, 达到 52.4%, 最大回撤也有所增加达到 39.1%,但是最大回撤期有所下降,只有 28 个交易日。与螺纹钢相比 R-MIDAS 模型在焦炭上的收益和风险增加都不多,但是夏普率还是有提升的,而且 R-MIDAS 的表现也并非一直领先,在上图的净值曲线上从 2016 年 11 月底至 2017 年 5 月 R-MIDAS 的表现有较长的一段平台期,但从长期收益和夏普率来看 R-MIDAS 模型仍有一定优势。

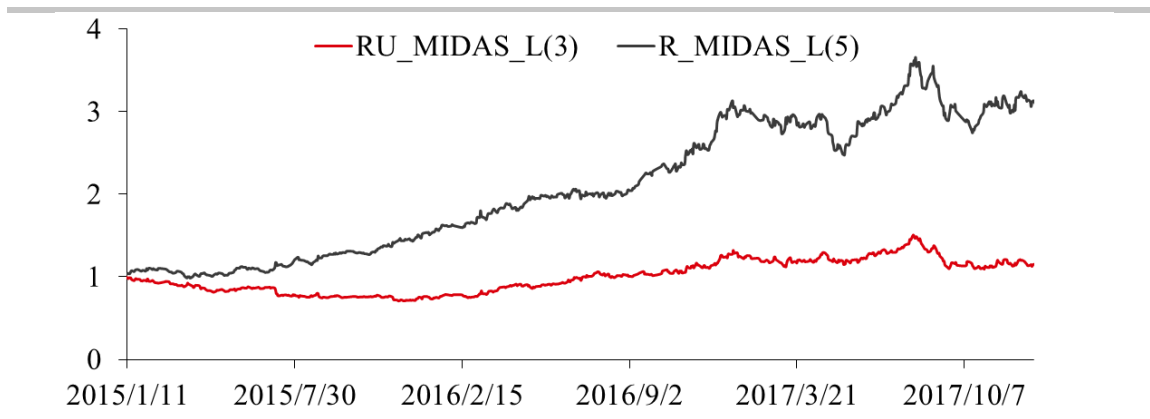
表格 2 焦炭 RU-MIDAS 和 R-MIDAS 模型收益与风险统计

	RU-MIDAS	R-MIDAS
年化收益	44.49%	57.9%
年化波动率	33.7%	33.7%
夏普率	1.31	1.71
胜率	52.3%	52.4%
最大回撤	36.7%	39.1%
最大回撤期	53 交易日	28 交易日

数据来源：华泰期货研究院

以下是这两个模型在 PVC 上的表现，从图上看 RU-MIDAS 模型表现较差并没有明显的收益增长，而 R-MIDAS 模型则表现明显较优，从 2015 年 1 月开始 R-MIDAS 模型便表现出较为持续的增长势头而且直到 2017 年 1 月一直没有出现较大的回撤，但是 2017 年 1 月后 R-MIDAS 模型便进入了一个较长的回撤期直到 2017 年 5 月，之后有较大的反弹然后保持震荡上行的趋势。因此利用 R-MIDAS 模型增加了延时和 Almon 多项式后，在 PVC 这个品种上收益的提高非常明显。

图 8: PVC 使用 RU-MIDAS 和 R-MIDAS 模型的净值曲线



数据来源：华泰期货研究院

这两个模型在 PVC 上的表现统计如下表，利用 R-MIDAS 模型年化收益能达到 30.7%，远高于 RU-MIDAS 模型的 4.4%，因此在 PVC 这个品种上 R-MIDAS 模型提高收益的作用是非常明显的，夏普率也大幅提高到了 1.48，而且胜率从 48.68%提高到了 51.5%。最大回撤也下降到了 25%，而最大回撤期则只有 43 个交易日。所以 R-MIDAS 在 PVC 上的表现可以说是非常突出的。

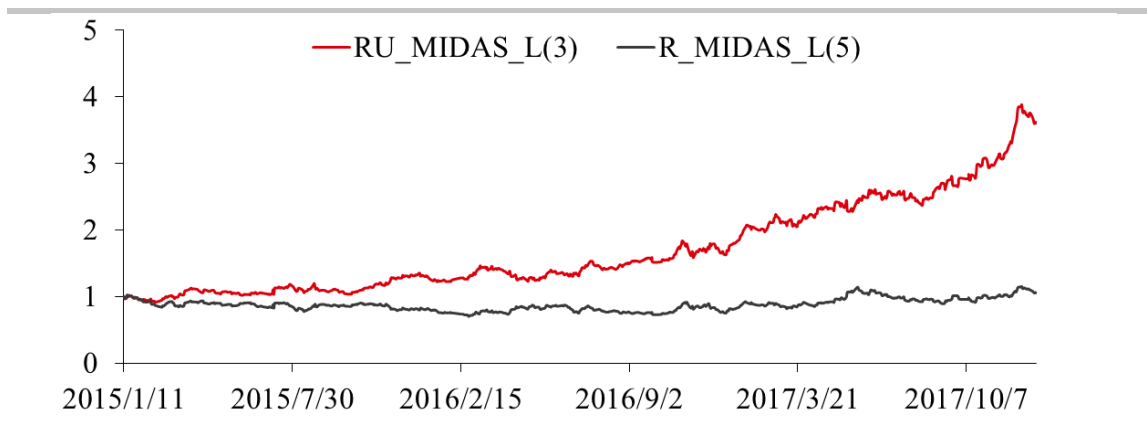
表格 3 PVC RU-MIDAS 和 R-MIDAS 模型收益与风险统计

	RU-MIDAS	R-MIDAS
年化收益	4.4%	30.7%
年化波动率	20.9%	20.7%
夏普率	0.21	1.48
胜率	48.68%	51.5%
最大回撤	35.5%	25%
最大回撤期	260 交易日	43 交易日

数据来源：华泰期货研究院

甲醇的净值曲线走势如下图所示，与之前几个品种不同，在这个品种上 RU-MIDAS 模型表现远远好于 R-MIDAS 模型。从 2015 年 1 月份起 RU-MIDAS 模型便较为稳定地走出上涨行情，直到 2017 年 12 月才出现明显回撤。而 R-MIDAS 模型一直走势相对平缓，从 2015 年 1 月便开始较长的回撤，知道 2017 年初才开始有所盈利，长期来看 R-MIDAS 模型勉强能够盈亏相抵。造成这种现象可能是因为甲醇这个品种因子之间较长时间的相关性并不强，比如说因子延时 1 天与延时 5 天的相关性很低的话，那样即使添加更多的延时也不会有太大意义。如果延时相关性不高的话，那么使用 R-MIDAS 模型增加延时，限制参数也就起不到作用了。

图 9: 甲醇使用 RU-MIDAS 和 R-MIDAS 模型的净值曲线



数据来源：华泰期货研究院

甲醇的收益与风险总结如下 RU-MIDAS 模型的年化收益达到了 44.1%，远远高于 R-MIDAS 模型，由于具有较高收益率，RU-MIDAS 模型的年化夏普率能达到 1.61，胜率也较高，达到 54.7%，最大回撤为 15.9%，最大回撤其 39 个交易日。从表上看 RU-MIDAS 模型的各项收益与风险指标都要好于 R-MIDAS 模型，所以 R-MIDAS 模型也并非对每个品种都适合。

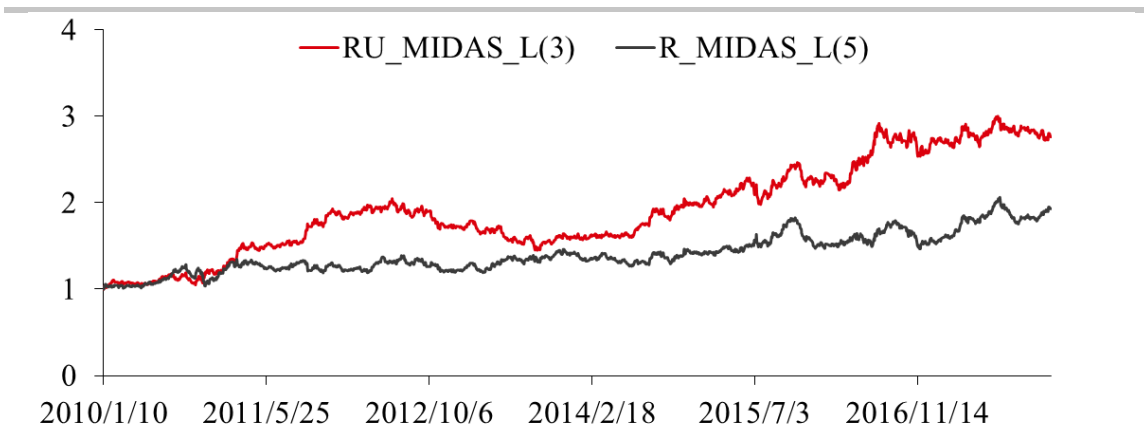
表格 4 甲醇 RU-MIDAS 和 R-MIDAS 模型收益与风险统计

	RU-MIDAS	R-MIDAS
年化收益	44.1%	1.80%
年化波动率	27.3%	27.5%
夏普率	1.61	0.065
胜率	54.7%	52.4%
最大回撤	15.9%	31.5%
最大回撤期	39 交易日	269 交易日

数据来源：华泰期货研究院

豆油的净值曲线走势如下图所示，从图上看 RU-MIDAS 模型表现也是要优于 R-MIDAS，但是差距没有甲醇那么大。从 2010 年 1 月开始至 2011 年 2 月初，两个模型的表现基本一致。但之后 RU-MIDAS 模型开始走高到 2012 年 6 月开始回撤，之后便有较强的上涨趋势，并表现好于 R-MIDAS 模型。R-MIDAS 模型虽然没有太大回撤，也保持上涨的趋势，但是长期来看收益不如 RU-MIDAS。

图 10: 豆油使用 RU-MIDAS 和 R-MIDAS 模型的净值曲线



数据来源：华泰期货研究院

豆油的收益与风险统计如下表所示，RU-MIDAS 模型的年化收益为 12.4%比 R-MIDAS 略高，夏普率达到 0.76，最大回撤是 29.2%比 R-MIDAS 高，而且回撤期也是比较长达 296 个交易日。因此总体上讲，在豆油这个品种上虽然两者都能取得不算太低的正收益，但表现也并没有之前品种那么突出。原因可能是因为豆油属于农产品，受天气、温度、降水等因素影响较大，而这些因素目前也并没有包含在模型的因子中。另外，农产品通常也呈现出季节性的特点，而目前这两个模型都没有考虑因子的季节性。

表格 5 豆油 RU-MIDAS 和 R-MIDAS 模型收益与风险统计

	RU-MIDAS	R-MIDAS
年化收益	12.4%	8.45%
年化波动率	16.3%	16.3%
夏普率	0.76	0.52
胜率	50.8%	50.6%
最大回撤	29.2%	19.7%
最大回撤期	296 交易日	256 交易日

数据来源：华泰期货研究院

结果讨论

本报告首先对 Eric Ghysels 等人提出的 MIDAS 模型进行了介绍，接着对 Claudia Foroni 等人提出的 RU-MIDAS 模型进行了改造，包括对多因子的拓展以及限制性多项式的引入等，从而构造出 R-MIDAS 模型。由于限制性多项式的使用，R-MIDAS 模型能够实现在不增加模型参数的基础上增加延时的效果，当期货收益率的时间序列与其它因子存在较强的滞后相关性时，增加延时对策略收益率的提高能起到一定促进作用。

但是 R-MIDAS 模型并非对所有品种都能起到效果，在目前测试的几个品种中，R-MIDAS 模型对螺纹钢、焦炭和 PVC 的收益率提高比较明显，年化收益至少提高 10% 以上，而且夏普率一般也有所提升，但是在高收益的同时往往伴随的是高回撤，因此在这三个品种上是否使用 R-MIDAS 模型还需根据风险偏好来进行考量。而在甲醇这个品种上 R-MIDAS 模型是不如 RU-MIDAS 模型的，原因可能是甲醇这个品种的因子滞后相关性并不高，从而不需要太多的延时项，所以 RU-MIDAS 模型表现尚且能令人满意。但是在豆油上这两个模型表现都不十分突出，虽然他们都有正收益，但都不十分高在 10% 左右，而且回撤期特别长。

影响时间序列模型表现的因素有很多例如使用的可观测因子、降维方法甚至模型的训练方法等，在接下来的专题报告里会对这一些列因素做进一步优化。

● 免责声明

此报告并非针对或意图送发给或为任何就送发、发布、可得到或使用此报告而使华泰期货有限公司违反当地的法律或法规或可致使华泰期货有限公司受制于的法律或法规的任何地区、国家或其它管辖区域的公民或居民。除非另有显示，否则所有此报告中的材料的版权均属华泰期货有限公司。未经华泰期货有限公司事先书面授权下，不得更改或以任何方式发送、复印此报告的材料、内容或其复印本予任何其它人。所有于此报告中使用的商标、服务标记及标记均为华泰期货有限公司的商标、服务标记及标记。

此报告所载的资料、工具及材料只提供给阁下作查照之用。此报告的内容并不构成对任何人的投资建议，而华泰期货有限公司不会因接收人收到此报告而视他们为其客户。

此报告所载资料的来源及观点的出处皆被华泰期货有限公司认为可靠，但华泰期货有限公司不能担保其准确性或完整性，而华泰期货有限公司不对因使用此报告的材料而引致的损失而负任何责任。并不能依靠此报告以取代行使独立判断。华泰期货有限公司可发出其它与本报告所载资料不一致及有不同结论的报告。本报告及该等报告反映编写分析员的不同设想、见解及分析方法。为免生疑，本报告所载的观点并不代表华泰期货有限公司，或任何其附属或联营公司的立场。

此报告中所指的投资及服务可能不适合阁下，我们建议阁下如有任何疑问应咨询独立投资顾问。此报告并不构成投资、法律、会计或税务建议或担保任何投资或策略适合或切合阁下个别情况。此报告并不构成给予阁下私人咨询建议。

华泰期货有限公司 2018 版权所有并保留一切权利。

● 公司总部

地址：广东省广州市越秀区东风东路761号丽丰大厦20层、29层04单元

电话：400-6280-888

网址：www.htfc.com