
基于多种模型的汇率分析及预测

褚贺 PB21151829
丁诺 PB21061219
李光宇 PB21081615

University of Science and Technology of China
Department of Statistic and Finance
2024 年 1 月 23 日

摘要

本研究旨在探究美元对英镑的即期汇率，时间跨度自1970年1月至2023年12月。我们采用了五种模型：TAR、TMA、STAR、ARIMA和LSTM，通过对其建模、评估和预测，比较它们在美元对英镑汇率预测中的性能。我们选择美元对英镑的即期汇率作为研究对象，通过历史数据的分析，期望为金融从业者和决策者提供有关不同模型应用的实用建议。研究贡献在于对五种模型的综合比较，为实际市场应用提供指导。

关键词：ARIMA、TAR、TMA、STAR、LSTM、时间序列

目录

1	引言	2
2	数据预处理	2
3	模型分析	5
3.1	ARIMA	5
3.2	TAR	6
3.2.1	介绍	6
3.2.2	建模	7
3.2.3	预测	8
3.3	TMA	10
3.3.1	介绍	10
3.3.2	建模	10
3.3.3	预测	10
3.4	STAR	12
3.4.1	介绍	12
3.4.2	建模	13
3.5	STMA	13
3.5.1	介绍	13
3.5.2	模拟和预测	14
3.6	LSTM	15
3.6.1	介绍	15
3.6.2	数据处理	16
3.6.3	模型训练	16
3.6.4	建模	16
4	总结	17
4.1	研究结果	17
4.2	不足与展望	18

1 引言

随着全球化的不断推进，外汇市场在国际经济体系中扮演着至关重要的角色。其对国际贸易、金融稳定以及宏观经济政策的影响日益显著。其中，美元对英镑的即期汇率作为一个关键的货币交叉对，受到全球金融从业者、企业家和政府决策者的广泛关注。对这一货币对的准确预测不仅对金融市场参与者具有战略性的重要性，同时也对国际贸易、资本流动以及宏观经济政策的制定提出了严峻挑战。

在金融领域，对即期汇率的准确预测一直是学术界和实践者共同关注的焦点。有效的汇率预测有助于投资者更好地制定资产配置策略，企业可以更精准地规划国际贸易活动，政府也能更有效地制定货币政策以维护国家经济的稳定。然而，由于外汇市场的高度不确定性、复杂性以及受到多种因素的影响，实现准确的汇率预测一直是一个具有挑战性的问题。

为了解决这一问题，学者们提出了多种不同的汇率预测模型。在本论文中，我们将关注于五种主要的建模方法：门限自回归（TAR）、门限移动平均（TMA）、平滑转换自回归（STAR）、自回归积分滑动平均（ARIMA）和长短时记忆网络（LSTM）。这五种模型代表了传统统计模型和深度学习模型的两个主要方向，我们将对它们在美元对英镑即期汇率预测中的性能进行深入研究和比较。

在模型选择的过程中，我们充分考虑到这些模型各自的优势和局限性。TAR模型能够捕捉到汇率在不同市场条件下的非线性特征，而TMA模型则注重于滤除噪声。STAR模型使得转换函数更加平滑，对市场波动和趋势的敏感性更高。ARIMA模型是一种经典的时间序列分析方法，通过差分 and 滑动平均来捕捉数据的季节性和趋势。而LSTM作为一种深度学习模型，能够处理复杂的非线性关系，具有强大的表达能力。

我们选择美元对英镑的即期汇率作为研究对象，时间跨度涵盖了1970年1月至2023年12月。通过历史数据的回顾和模型的建立，我们旨在全面评估这五种模型在预测未来汇率变动中的表现。此外，我们还将比较不同模型的优劣，以便为决策者、投资者和研究人员提供更全面的信息，帮助他们更好地理解 and 应对外汇市场的变化。

本研究的贡献不仅在于对于美元对英镑即期汇率的深入分析，更在于对不同模型性能的全面比较。这将为学术界提供有关不同建模方法在实际金融市场中的适用性和效果的宝贵信息。通过对比传统统计模型和深度学习模型的表现，我们有望为金融从业者提供更为实用和可靠的决策支持。同时，我们的研究也为未来相关研究提供了有益的经验和启示，推动汇率预测领域的进一步发展。

2 数据预处理

本文研究的是美元对英镑从1970年1月到2023年12月的即期汇率数据，也即多少美元等价于一英镑。其中日度数据来自<https://fred.stlouisfed.org/series/DEXUSUK>。月度数据来自<https://fred.stlouisfed.org/series/EXUSUK>，为每月一号的数据。

首先对得到的时间序列数据作图：

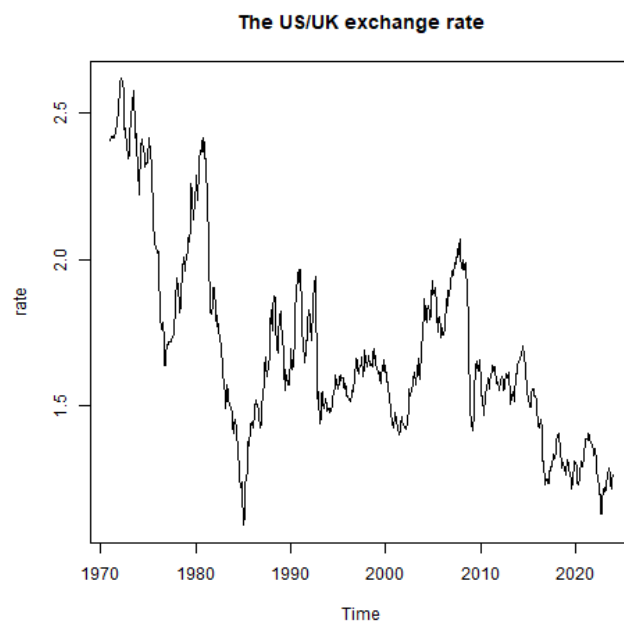


图 1: 美元对英镑即期汇率的时序图

利用decompose函数对时间序列数据进行初步分解，观察数据概况：

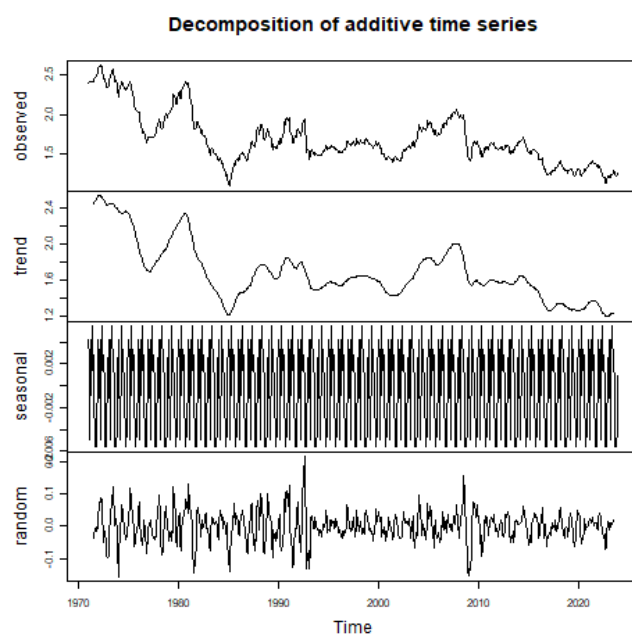


图 2: 美元对英镑即期汇率的时序分解图

可以看到整体上美元对英镑即期汇率呈下降趋势但是波动性较大，同时有一定的季节性。再考虑对数差分 $y_t = \log(x_t) - \log(x_{t-1})$ 得到序列 $\{y_t\}$ 并作时序图：

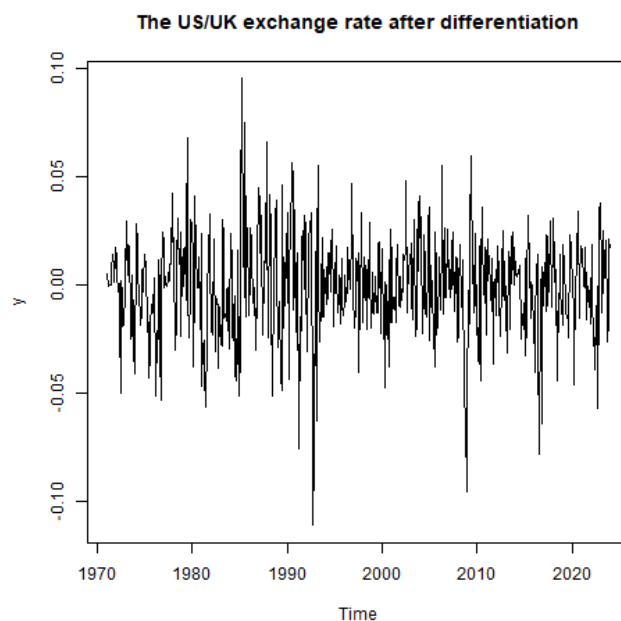


图 3: 差分序列的时序图

观察差分后数据的acf和pacf图:

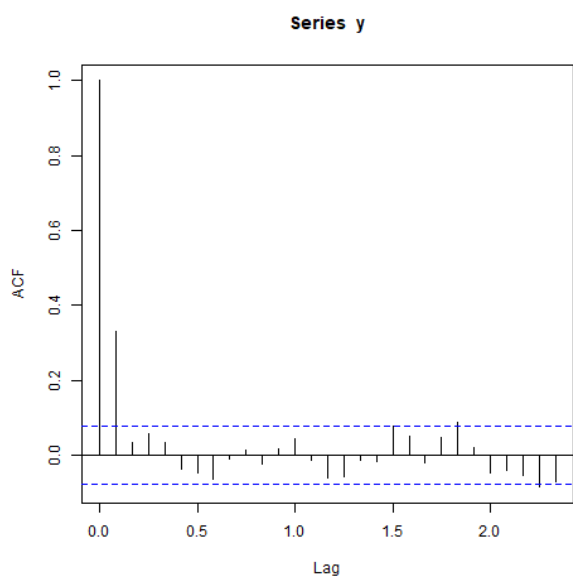


图 4: 差分后数据的acf

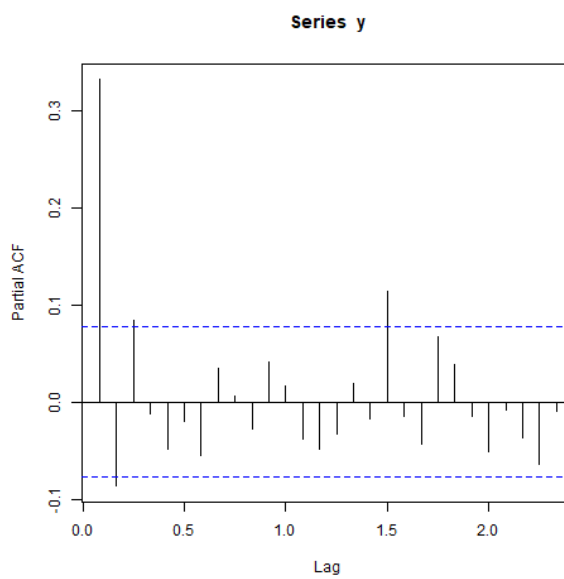


图 5: 差分后数据的pacf

可以明显看到acf截尾, pacf拖尾。

我们对数据进行基本的白噪声、单位根检验等, 数据处理即从此出发, 代码结果如下所示:

```

One Sample t-test

data: ldata
t = -0.98978, df = 13298, p-value = 0.3223
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.025081832  0.008250491
sample estimates:
 mean of x
-0.00841567

Box-Ljung test

data: ldata
X-squared = 50.199, df = 10, p-value = 2.454e-07

Warning: p-value smaller than printed p-value
Augmented Dickey-Fuller Test

data: ldata
Dickey-Fuller = -22.45, Lag order = 23, p-value = 0.01
alternative hypothesis: stationary

```

图 6: 检验

3 模型分析

3.1 ARIMA

利用auto.arima对差分数据 $\{y_t\}$ 进行ARIMA建模并预测。同时，为检验预测结果，将1971年1月至2019年12月的数据作为训练集，2020年1月至2023年12月作为测试集。运行相关代码，得到建模结果为ARIMA(0,1,1)(0,0,1)[12]，与前文的acf和pacf图像特征吻合。模型对应的AIC为-2812.38，可以看到偏差较大。利用得到的模型拟合实际数据，得到：

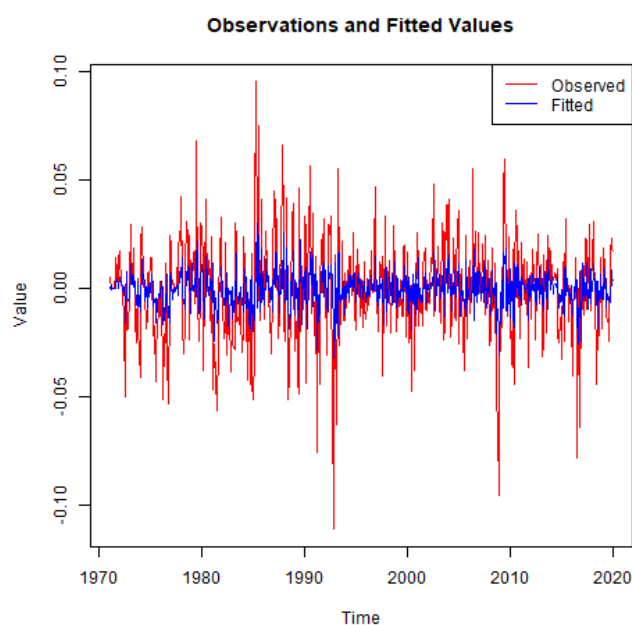


图 7: ARIMA模型下的拟合

再利用该模型在测试集上进行预测，得到：

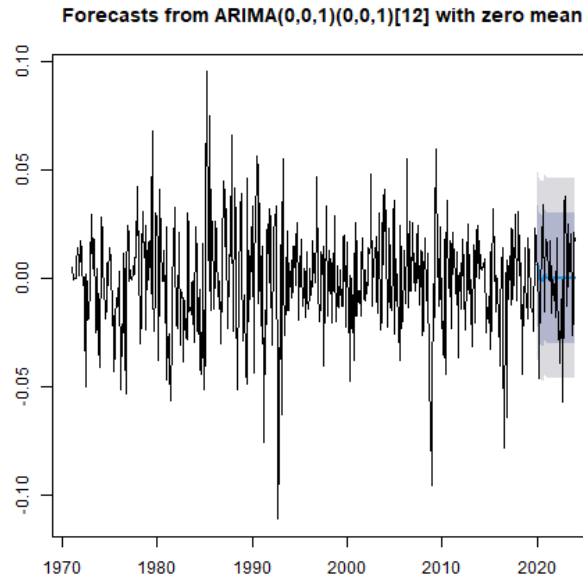


图 8: ARIMA模型下的预测

可以看到在拟合时，ARIMA模型的波动性明显小于原模型，这一点在预测时也有所体现，预测线接近水平且95%的置信区间不能将波动完全覆盖，因此拟合和预测效果不尽人意。考虑可能是因为存在结构性断裂导致使用线性模型建模和预测不准确，下面考虑使用非线性模型以及LSTM模型对数据进行建模和预测。

3.2 TAR

3.2.1 介绍

门限自回归（Threshold Autoregressive, TAR）模型作为一种被广泛运用于非线性时间序列建模的方法，通过引入门限值在不同市场条件下建立不同的回归方程，以更精准地刻画时间序列数据的非线性特性。

在TAR模型中，通常使用五个参数来描述模型的结构，分别是 $p1$ 、 $p2$ 、 d 、 a 和 b 。以下是对这些参数的简单介绍：

- $p1$ (Threshold): 门限值，表示时间序列的阈值。该值将时间序列划分为两个不同的状态，分别对应于门限以上和门限以下的市场条件。 $p1$ 的选择对于模型的性能至关重要，需要通过数据分析或其他方法来确定。

- $p2$ (Lag Order Above Threshold): 门限以上状态下的滞后阶数，表示在门限以上状态下，回归方程中考虑的去观测值的数量。 $p2$ 的选择涉及对时间序列的滞后关系的分析，以确保模型能够捕捉到适当的非线性动态。

- d (Order of Differencing): 差分阶数，表示为使时间序列平稳所需的差分操作的次数。通过差分操作，可以消除季节性和趋势成分，使数据更适合建模。

- a (Coefficient Below Threshold): 门限以下状态下的回归系数。表示在门限以下状态下，回归方程中与过去观测值相关的权重。

- b (Coefficient Above Threshold): 门限以上状态下的回归系数。表示在门限以上状态下，回归方程中与过去观测值相关的权重。

这些参数共同定义了TAR模型的结构，通过对它们的合理选择和调整，可以更好地适应时间序列数据的非线性特性。由于在实践中TAR阶数是未知的，因此需要一种允许对它们进行估计的方法。称为最小AIC（MAIC）方法：对于TAR模型，对于固定的 r 和 d ，AIC变为

$$AIC(p_1, p_2, r, d) = -2l(r, d) + 2(p_1 + p_2 + 2)$$

通过最小化AIC对象来估计参数，以便在某个时间间隔内搜索阈值参数，以使任何方案都有足够的数据进行估计。

3.2.2 建模

TAR模型的建模过程始于门限值的确定。通过对数据分布和变化趋势的深入探索，利用统计方法或图形分析选择适当的门限值。确定门限值后，将数据分割为不同的区域，并在每个区域内建立相应的回归方程。通过最大似然估计或其他参数估计方法，获取模型的参数。该过程需要仔细调整门限值和回归方程，以最大程度地贴合数据的非线性结构。

经过不断调试模型超参数，我们选取 $p_1 = 1, p_2 = 1, d = 3$ 。在最小化AIC准则下测试不同 a ， b 取值对TAR建模结果的影响。结果如下：

(a,b)	Nominal AIC
(0.1,0.9)	-85490
(0.15,0.85)	-85490
(0.2,0.8)	-85490
(0.25,0.75)	-85490
(0.3,0.7)	-85490
(0.35,0.65)	-85490
(0.4,0.6)	-85460
(0.45,0.55)	-85430
(0.5,0.5)	-85410

为了对比不同 (a,b) 对试验结果的影响，我们选取 $(0.15, 0.85)$ 与 $(0.5, 0.5)$ 两种情形建模。

首先使用残差分析完成模型诊断：

TAR模型的残差定义为：

$$\hat{e}_t = Y_t - \{\hat{\phi}_{1,0} + \hat{\phi}_{1,1}Y_{t-1} + \cdots + \hat{\phi}_{1,p}\}I(Y_{t-\hat{d}} \leq \hat{r}) - \{\hat{\phi}_{2,0} + \hat{\phi}_{2,1}Y_{t-1} + \cdots + \hat{\phi}_{2,p}\}I(Y_{t-\hat{d}} > \hat{r})$$

标准化残差定义为：

$$\hat{e}_t = \frac{\hat{e}_t}{\hat{\sigma}_1 I(Y_{t-\hat{d}} \leq \hat{r}) + \hat{\sigma}_2 I(Y_{t-\hat{d}} > \hat{r})}$$

如果TAR模型是真正的数据机制，则标准化残差图应看起来是随机的。可以通过检查标准化残差的样本ACF来检查标准化误差的独立性假设。

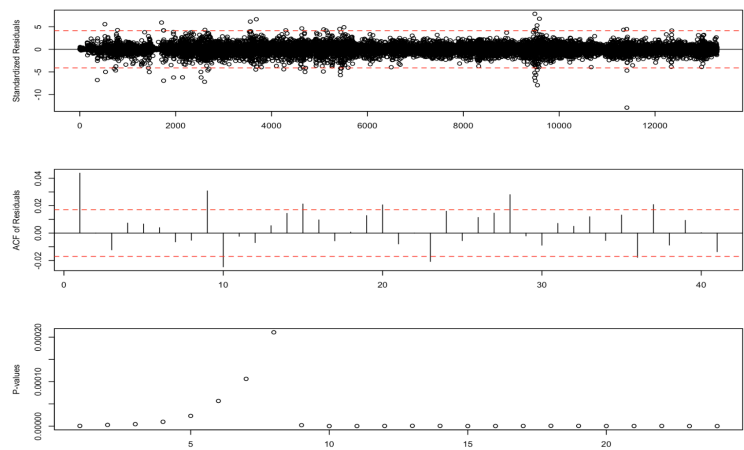


图 9: TAR(1,1,3,0.1,0.9)残差分析图

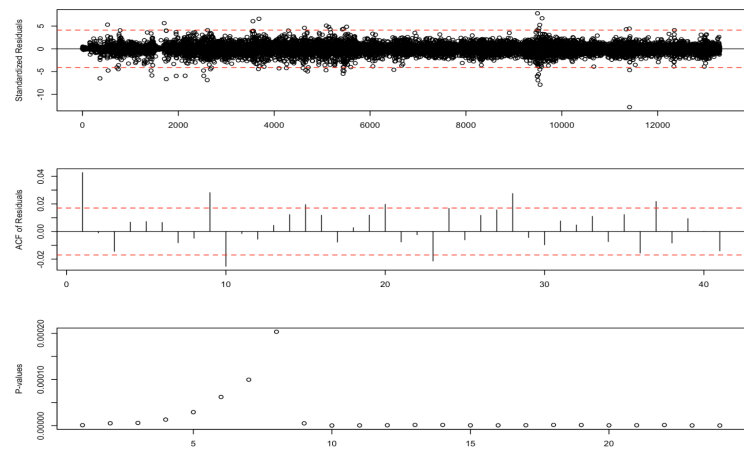


图 10: TAR(1,1,3,0.5,0.5)残差分析图

二者残差分析图表明，建模过程正确，可以继续使用TAR模型进行后续预测。

3.2.3 预测

为全面评估TAR模型的预测性能，我们采用了多个标准的评估指标。这包括均方误差（MSE）、均方根误差（RMSE）、平均绝对误差（MAE）等，以量化模型在样本内和样本外的预测准确性。同时进行残差分析，验证模型的残差序列是否符合独立同分布的假设，以及是否存在系统性模式。通过这一步骤，我们能够识别潜在的模型改进点，并进一步优化TAR模型的性能。此外，为了对比TAR与AR模型的性能，我们对比了两种模型对于数据的拟合性能。结果如下：

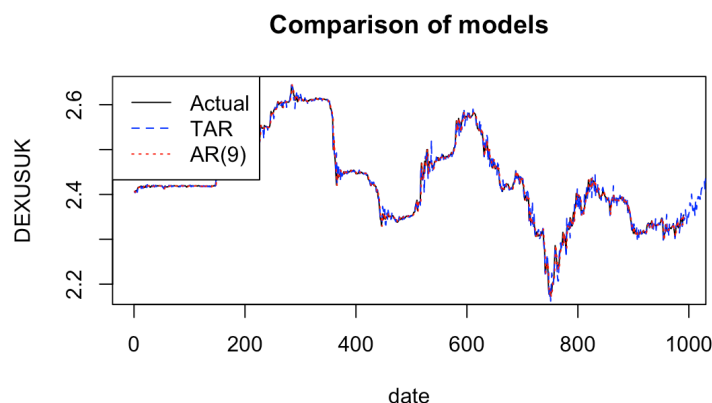


图 11: $TAR(1,1,3,0.1,0.9)$ vs. $AR(9)$

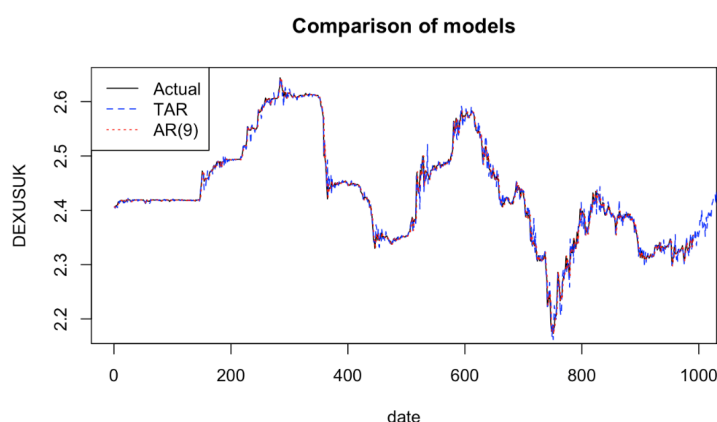


图 12: $TAR(1,1,3,0.5,0.5)$ vs. $AR(9)$

我们发现，TAR模型与AR模型相比具有更好的数据拟合性能。并且TAR模型预测数据能力更强，更符合真实的数据波动情况。

我们通过深入分析两种模型的特性，总结了TAR模型相对于AR模型的几个显著优点。

首先，TAR模型具备更强大的非线性建模能力。相较于AR模型的线性假设，TAR模型能够更灵活地捕捉时间序列数据中的非线性关系。其在不同阈值范围内采用不同的回归方程，允许模型更好地适应数据的动态变化。

其次，TAR模型适应性更强。通过调整阈值，TAR模型能够适应不同市场状态和结构性变化，使其在处理不同时间段的数据时更具鲁棒性。这一特性使得TAR模型在面对复杂和动态的时间序列数据时表现更为出色。

另外，TAR模型在捕捉时间序列中的拐点和结构性突变方面具有优势。当数据存在非线性结构的拐点时，TAR模型能够更准确地拟合这种变化，相较于AR模型更具优势。

最后，TAR模型在处理异方差性的数据时更为适用。通过在不同市场状态下使用不同的回归方程，TAR模型能够更好地应对异方差性的情况，提高了模型的稳健性。

通过以上详细步骤，我们深入了解了TAR模型在美元对英镑即期汇率预测中的表现。这一细致的分析方法不仅有助于理解模型的优势和局限性，也使我们能够为模型选择提供更为科学和实用的建议。

3.3 TMA

3.3.1 介绍

门限移动平均（Threshold Moving Average, TMA）模型[4]是一种时间序列分析方法，结合了移动平均的平滑性和门限模型的非线性特性。通过引入阈值，在不同市场状态下应用不同的移动平均值，以更准确地捕捉时间序列数据的非线性动态特性。TMA模型的优势在于能够更灵活地适应不同市场状态下的非线性动态，通过引入阈值，使模型更贴近实际市场情况。这对于研究美元对英镑即期汇率的波动至关重要，因为这样的波动常常受到多种复杂因素的影响。表达式如下：

$$y_t = \begin{cases} \beta_0 + \beta_1 x_{t-1} + \epsilon_t, & \text{if } x_{t-1} \leq \text{阈值} \\ \gamma_0 + \gamma_1 x_{t-1} + \epsilon_t, & \text{if } x_{t-1} > \text{阈值} \end{cases} \quad (1)$$

其中， y_t 是时间序列的观测值， x_{t-1} 是滞后一期的时间序列值， ϵ_t 是误差项。当 x_{t-1} 小于等于阈值时，采用第一种线性关系，否则采用第二种线性关系。

3.3.2 建模

TMA模型的建模过程涉及阈值的选择、移动平均窗口的设定和模型参数的估计。我们选择了美元对英镑即期汇率的数据，通过复杂的拐点分析，识别了数据中存在的几处明显拐点。这些拐点可能反映了宏观经济因素、政治事件或其他外部冲击，对市场产生了结构性变化。对这些拐点进行深入分析有助于更全面地理解模型的应用背景。

拐点分析可以通过寻找时间序列的一阶导数或曲率的变化来识别。数学上，一阶导数表示变化率，曲率表示变化率的变化率。拐点即对应这些变化的位置。模型参数的估计涉及最大似然估计等方法，通过最大化似然函数来找到最符合观测数据的参数值。针对这些拐点，我们分别采用阈值为2、2.2和2.5的TMA模型。通过高度技术性的模型训练和参数估计，得到了在不同市场状态下的移动平均值。这一策略旨在更好地捕捉数据的非线性动态，提高模型的预测性能。

3.3.3 预测

为评估TMA模型的预测性能，采用了多个标准的评估指标，如均方误差（MSE）、均方根误差（RMSE）、平均绝对误差（MAE）等。通过复杂的比较模型在训练集和测试集上的表现，全面评估其对未来值的准确性。三处拐点拟合图如下：

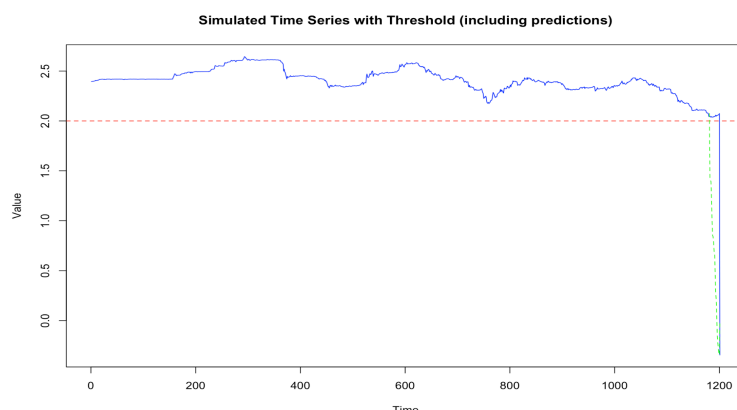


图 13: TMA(2)

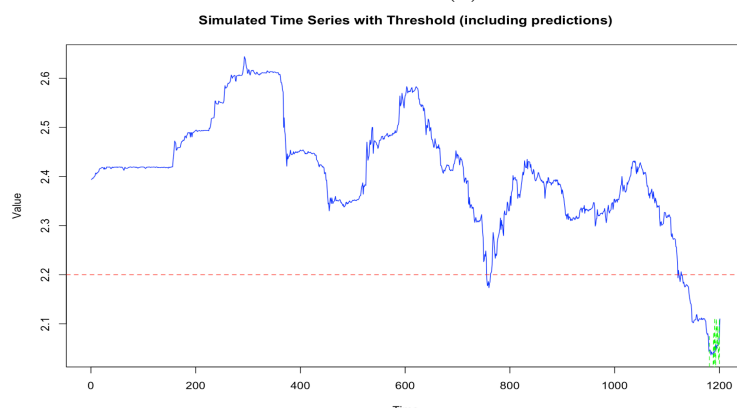


图 14: TMA(2.2)

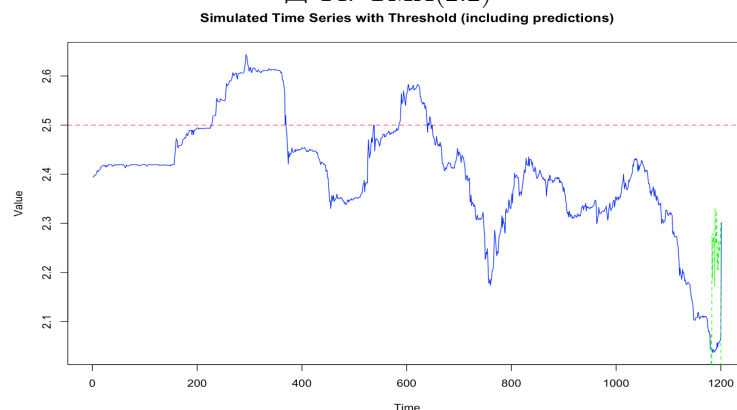


图 15: TMA(2.5)

在对预测性能的解释中，我们关注了美元对英镑即期汇率的短期波动。曲线的上升可能代表着美元对英镑汇率的短期上涨趋势，反映出市场对美元的偏好或其他相关因素。相反，曲线的下降可能表示短期下跌趋势，反映出市场对英镑的偏好或其他相关因素。拐点的出现可能意味着市场结构性变化，例如金融危机、政策调整等。通过这些解释，我们更深入地理解了模型对汇率波动的捕捉和预测性能的评估。这为我们在实际应用中更好地利用模型提供了有益的实证支持。

结果表明选取的三处拐点通过使用不同阈值的TMA模型拟合均取得了较好的预测效果。

通过将TMA模型应用于美元对英镑即期汇率的建模，我们深入分析了数据中存在拐点的原因，为模型的应用提供了更深层次的解释。这既加强了模型的解释，也为更全面地理解和利用汇率波动

提供了有益的实证支持。这进一步凸显了TMA模型在处理具有非线性特性的金融时间序列中的潜力。

3.4 STAR

3.4.1 介绍

平滑过渡自回归(Smooth Transition Autoregression, STAR)模型最初由Chan和Tong(1986)[1]提出,并由Teräsvirta(1994)[8]及其合作者进一步发展。其中,STAR模型的一些变形包括多状态STAR模型(van Dijk和Franses, 1999)[3]、时变STAR模型(Lundbergh等人, 2003)[7]、矢量STAR模型(Rothman等人, 2001)[5]等等。在过去的研究中,STAR模型已经成功地解释了各种宏观经济时间序列的行为,如商业周期不同阶段的产出、汇率和失业率, van Dijk等人(2002)[2]和Teräsvirta等人(2010)[9]也对STAR模型进行了广泛的调查。因此,在本文中,也考虑使用STAR模型对美元对英镑的汇率数据进行建模。

在STAR模型中, p 阶单变量时间序列的平滑过渡自回归模型STAR(p)一般定义如下[11]:

$$y_t = \Phi_{10} + \Phi_1^T w_t + (\Phi_{20} + \Phi_2^T w_t)G(y_{t-d}) + \epsilon_t$$

$$\epsilon_t \sim i.i.d(0, \sigma^2)$$

$$\Phi_j = (\Phi_{j1}, \dots, \Phi_{jp})^T, j = 1, 2, \Omega_t = (y_{t-1}, y_{t-2}, \dots, y_{t-p})$$

其中函数 $G(x)$ 是一个在0到1之间有界的连续函数,不同函数的选择会导致不同的结果。在过去的研究中,人们往往采用以下两种过渡函数,它们分别为:

(1) 一阶逻辑函数:得到的模型称为logistic STAR (LSTAR)模型,其中

$$G(x) = (1 + \exp(-\gamma(x - c)))^{-1}, \gamma > 0$$

在函数中,随着 x 的增加,逻辑函数从0到1单调变化, $G(c) = 0.5$, 参数 c 可以被解释为两种状态之间的阈值,参数 γ 决定了逻辑函数值变化的平滑性,从而决定了从一个状态过渡到另一个状态的平滑性。当 γ 非常大时,逻辑函数 $G(x)$ 接近示性函数 $I[x > c]$,因此,门限自回归(TAR)模型可以视为LSTAR模型的一个特例。

(2) 指数函数:得到的模型称为Exponential STAR (ESTAR)模型,其中

$$G(x) = 1 - \exp(-\gamma(x - c)^2), \gamma > 0$$

下面通过对几个示例函数进行作图来体会平滑性的作用:

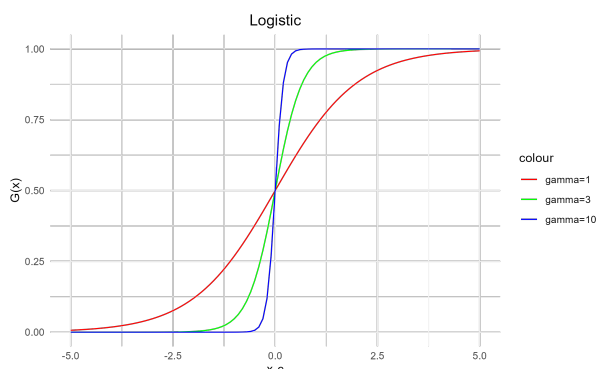


图 16: 不同 γ 下的一阶逻辑函数

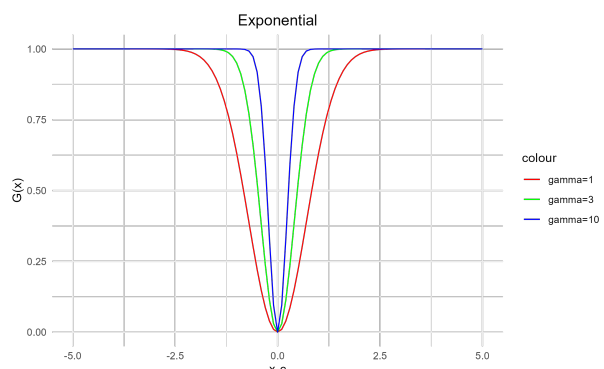


图 17: 不同 γ 下的指数函数

可以看到 γ 的增大均增加了函数的平滑性。

除此之外，也可考虑利用二阶逻辑函数 $G(x) = (1 + \exp(-\gamma(x - c_1)(x - c_2)))^{-1}, \gamma > 0$ 进行建模。但在这里我们主要使用一阶逻辑函数。

3.4.2 建模

在分析过程中发现日度数据存在少量的缺失会导致在数据处理过程中产生NA，因此在这里采用月度数据进行分析，由于数据量充足，采用月度数据和采用日度数据得到的结果基本上可以视作是一致的。利用LSTAR模型对进行汇率数据建模拟合，得到：

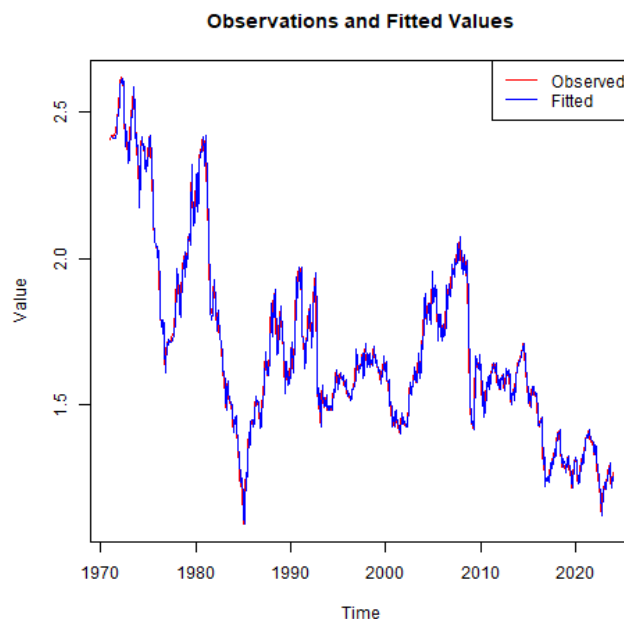


图 18: 全部数据在STAR模型下的拟合结果

可以看到汇率数据在STAR模型下的拟合效果很好。同时得到该模型下，Mean Absolute Percentage Error（平均绝对百分比误差/MAPE）为1.662%，AIC为-4199。

除此之外，通过检验一个全阶LSTAR模型和一个全阶AR模型之间的非线性关系，得到检验结果中F值为1.4959，p值为0.22485。这里的F值用于衡量LSTAR模型相对于AR模型的拟合优度，如果F值较大，则说明LSTAR模型相对于AR模型的拟合效果较好。而p值则用来评估F值的显著性。在这个例子中，p值为0.22485，大于通常的显著性水平0.05，因此我们没有足够的证据拒绝原假设，即不能确定LSTAR模型相对于AR模型是更好的拟合。

3.5 STMA

3.5.1 介绍

AR(Autoregression)模型通常反应过往数据的延续性，譬如股票市场的动量影响；而MA(Moving average)模型主要揭示了一系列具有延迟效应的不可预测事件的冲击效应，如意外消息引起的市场冲击。而由随机波动引起的不对称效应可能导致NMA模型，本文[10]即提出一个新的NMA(Non-linear Moving Average)模型子类——STMA(Smooth Transition Moving Average)。与TMA(Threshold moving average)相比，我们状态切换的过渡是平滑的。

讨论STMA(1):

$$y_t = \epsilon_t + \left\{ \phi + \psi F\left(\frac{y_{t-1} - c}{\gamma}\right) \right\} \epsilon_{t-1} \quad (2)$$

其中 ϵ_t 是独立白噪声， $F()$ 是平滑函数（常使用Sigmoid函数）， γ 为平滑子，可以看到当 $\gamma \rightarrow \infty$ 时，参数为 ϕ ；当 $\gamma \rightarrow 0$ 时，参数为 $\phi + \psi$ 。即 $\phi \rightarrow \phi + \psi \rightarrow \phi + \psi * function$ ，即由 $MA \rightarrow TMA$ 的过程。

$$\theta_t(\lambda) \equiv \phi + \psi F\left(\frac{y_{t-1} - c}{\gamma}\right) \quad (3)$$

优化的目标函数为:

$$L_n(\lambda) = \sum_{t=1}^n \tilde{\epsilon}_t^2(\lambda) \quad (4)$$

这是基于所有历史新息的替代，其中 $\tilde{\epsilon}_t(\lambda) = y_t - \left\{ \phi + \psi F\left(\frac{y_{t-1} - c}{\gamma}\right) \right\} \tilde{\epsilon}_{t-1}(\lambda)$ 。我们采用的估计即 $\hat{\lambda} = \arg \min_{\lambda} L_n(\lambda)$ 。

3.5.2 模拟和预测

我们使用的数据集为每日美国/英国的汇率（以美元到1英镑为单位）是从1971年到2020年的最后一日，共计 $n = 13045$ 次观察。此外，我们使用从2021年至今的数据进行预测，共计253次观测。假设原始汇率为 x_t 。我们将增长率系列定义为 $y_t = 100(\log x_t - \log x_{t-1})$ 。汇率和增长率如图所示。

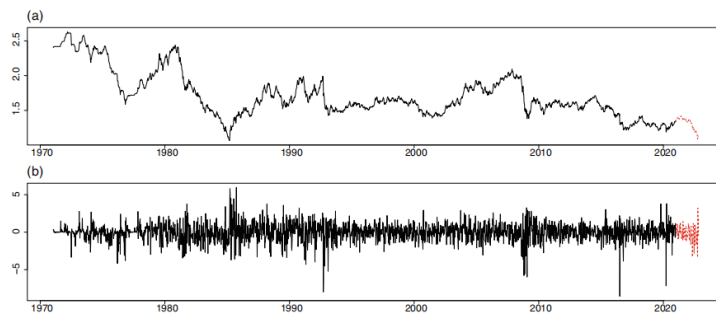


图 19: 1970-至今汇率和增长率

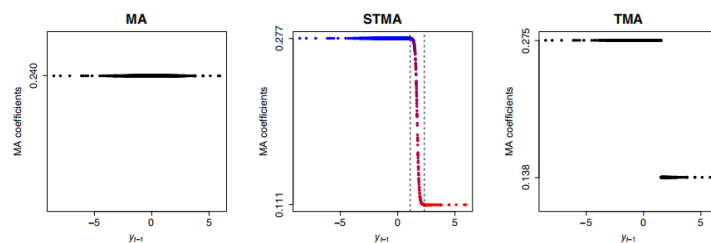


图 20: 拟合结果

我们可以看到，大多数观察结果都落在较低的时期。在较高的状态下，当汇率迅速上升时，对增长率的冲击效应较小，主要包括英镑迅速升值的三个时期：

(1)1985–1990.1985年，美英双方达成了“Plaza Accord”，允许美元对英镑和其他货币大幅贬值。与此同时，1980年至1985年英镑的贬值有利于英国的国际贸易和外汇储备。

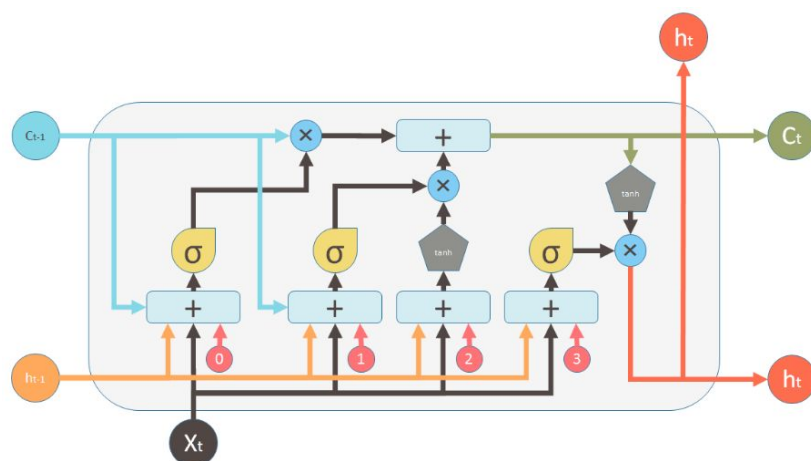


图 21: LSTM图示

(2)2009-2010.在2008年的全球金融危机期间，英镑大幅贬值。在2009-2010年后经济正在缓慢复苏。

(3)2020.可能的原因是英国脱欧造成的不稳定。

上述背景信息表明，在整个波动模型下存在一种开关机制。切换是完全突然的还是完全平稳的，这还需要进一步的研究。

3.6 LSTM

3.6.1 介绍

LSTM，全称Long Short Term Memory（长短期记忆）是一种循环神经网络（RNN）的变体，用于处理和预测时序数据。相对于传统的RNN，LSTM引入了一种特殊的记忆单元，可以更好地捕捉和长期记忆数据中的依赖关系。

LSTM的关键思想是通过使用记忆单元和门控机制来控制信息的流动。记忆单元允许网络选择性地记住或忽略输入数据中的某些信息。门控机制由三个门控单元组成：输入门（input gate）、遗忘门（forget gate）和输出门（output gate）。

输入门决定将多少新信息纳入记忆单元，遗忘门决定将多少旧信息从记忆单元中删除，输出门决定将多少记忆输出到下一个时间步。这些门控单元通过使用可学习的权重和激活函数（如sigmoid函数）来控制信息的流动。

具体来说，LSTM的计算过程如下：

- 1.输入数据和前一个时间步的隐藏状态被馈送到输入门和遗忘门。输入门决定了新信息的重要程度，遗忘门决定了保留多少旧信息。
- 2.输入门和遗忘门的输出被用于更新记忆单元的内容。旧的记忆被遗忘，新的记忆被加入。
- 3.输入数据和前一个时间步的隐藏状态被馈送到输出门。输出门决定了将多少记忆输出到下一个时间步。
- 4.记忆单元的内容被通过输出门进行筛选，得到当前时间步的隐藏状态。

重复上述步骤，处理下一个时间步的输入数据。

LSTM的门控机制使得它能够有效地处理长期依赖关系，避免了传统RNN中的梯度消失和梯度爆炸问题。因此，LSTM在处理自然语言处理、语音识别、时间序列预测等任务时表现出色，并成为了深度学习中重要的模型之一[6]。

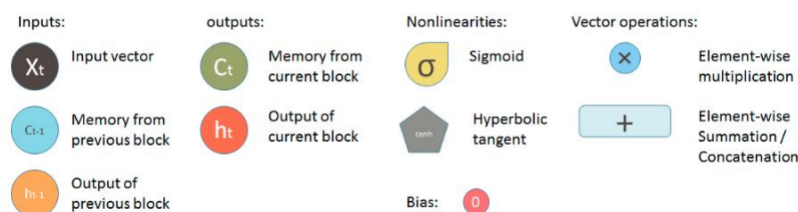


图 22: 其中相应的元

3.6.2 数据处理

读入数据后，转换数据格式，将其进行基本的平稳性、独立性等检验，知需要一阶差分。处理后的数据如前所示。

将时间序列问题转换为机器学习(监督学习)的问题主要在于数据的处理，前述已知，我们需要用历史信息，推断未来信息，此处我们选择一次的推移：把前一个数作为输入，后一个数作为输出。

拆分数据集为训练集与测试集。

将训练集合测试集中的数据缩放到 $[-1, 1]$ 之间，加快收敛速率及出于LSTM对Scale的敏感性，在这里我们借助MinMaxScaler缩放器。

3.6.3 模型训练

初始化LSTM模型并开始训练，设置神经元核心的个数，设置训练时输入数据的格式等等。详情参见APPENDIX。

3.6.4 建模

将数据的输入和输出列分开，并且将输入进行变换，传入到预测函数中进行单步预测，得到预测值后对其进行逆缩放和逆差分，将其还原到原来的取值范围内，遍历全部测试集数据，对每行数据执行以上操作，并将最终的预测值保存下来。

将测试集的y值和预测值绘制在同一张图表中，代码如下。这个问题的数据集非常大，LSTM的训练效果非常好，标准差大概为2，预测结果符合预期。

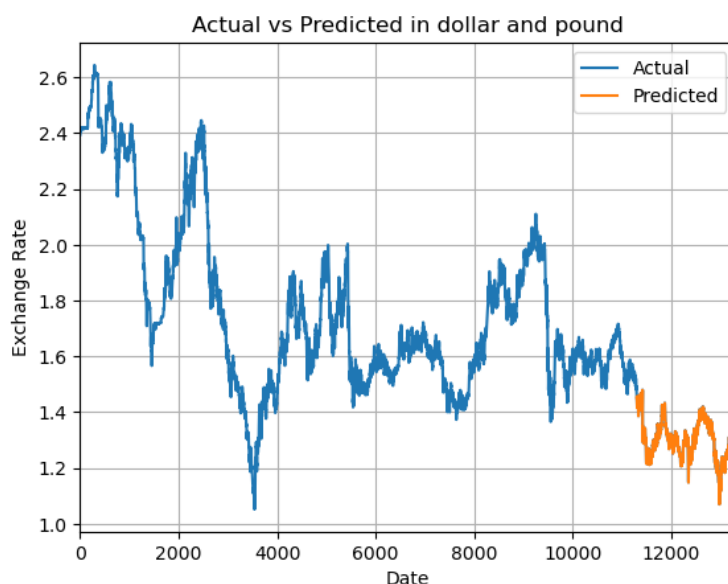


图 23: Forecast

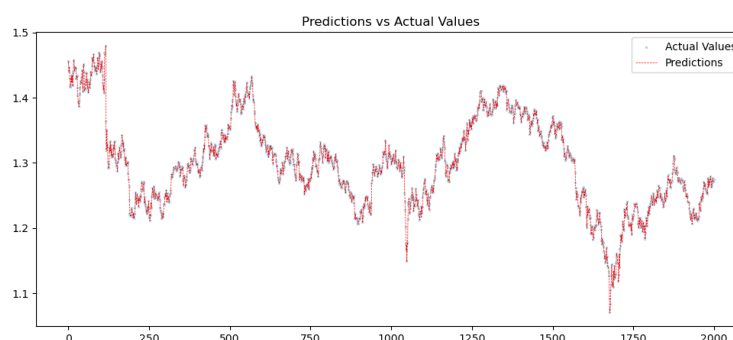


图 24: Behavior on Test Set

4 总结

4.1 研究结果

1、从整体上看，曲线呈下降趋势，但是波动性很大，季节性也比较明显。同时，由于某些年份发生了重大事件导致数据产生剧烈波动，比如说1985数据断崖式下降后又有所回升，这与英美双方达成的“Plaza Accord”有关。汇率是观察金融风向的重要指标之一，因此，通过利用多种时序模型对汇率数据进行建模与预测分析有助于人们提前调整金融行为，降低金融风险；

2、通过ARIMA模型建模，可以发现ARIMA得到的模型整体上波动性小于真实情况，且预测的准确度也有待提高，猜测可能是因为存在结构性断裂，因此线性模型在这里不再适用，考虑非线性模型以及非参数模型；

3、通过TAR模型建模，发现它与AR模型相比具有更好的数据拟合性能和更强的预测能力，从而更契合真实数据的波动情况。同时，通过调整阈值，TAR可以适应不同市场状态和结构性变化，在捕捉拐点时也更具有优势并可用于处理异方差性数据，应用范围也更加广泛；

4、通过TMA建模建模，我们深入分析了数据中存在拐点的原因，为模型的应用提供了更深层次的解释。这既加强了模型的解释，也为更全面地理解和利用汇率波动提供了有益的实证支持，从

而进一步凸显了TMA模型在处理具有非线性特性的金融时间序列中的潜力；

5、通过TSAR模型建模，采用一阶逻辑函数作为过渡函数，可以发现其对已有数据的拟合效果很好，MAPE也较低，但AIC方面高于ARIMA模型同时在检验中没有足够的证据证明其优于AR模型；

6、通过TSMA模型建模并结合背景信息，可以发现在整个波动模型下存在着一种开关机制，同时由预测结果可以发现其对美元对英镑的即期汇率预测有较好的效果；

7、通过利用LSTM对数据进行分析与处理，可以发现由于研究所用的数据集非常大，因此该模型的训练效果很好，预测结果与实际情况的误差很小，这也表明在分析时间序列数据时也可以考虑引入深度学习，这往往会在预测中发挥意想不到的作用；

8、通过使用多种模型对汇率数据进行分析，我们可以发现不同模型对汇率数据都具有一定的解释作用，同时也可能存在一定的缺陷导致在该数据集上并不适用，在不断进行检验、建模和预测的过程中，我们可以从多个视角对汇率数据有一个更深的了解，对不同模型的性能也有一个全面的比较，从而有望为金融从业者提供更为实用和可靠的决策支持，推动汇率预测领域的进一步发展。

4.2 不足与展望

1、数据取自官网，本身具有权威性，但仍然有可能存在一些系统性缺陷，如果深入对数据点进行分析判断是否有异常点可能使结果更具稳健性，但由于汇率数据量庞大，因此该过程可能过于复杂，这也是本文没有考虑的原因；

2、模型参数可能还可以通过一些技巧进行进一步优化，从而在建模和预测中获得更好的效果，这是对各种建模和预测方法更加深入的研究，期待后面的研究者取得新的进展；

3、对汇率数据的分析还可以考虑其他模型的使用，可能也可以达到很好的效果，为行文的简洁性这里也不再赘述，期待其他研究者进行探究。

参考文献

- [1] K. S. Chan and H. Tong. On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 7(3):179–190, 1986.
- [2] Timo Teräsvirta Dick van Dijk and Philip Hans Franses. Smooth transition autoregressive models — a survey of recent developments. *Econometric Reviews*, 21(1):1–47, 2002.
- [3] Dick van Dijk and Philip Hans Franses. Modeling multiple regimes in the business cycle. *Macroeconomic Dynamics*, 3(3):311–340, 1999.
- [4] Alain Guay and Olivier Scaillet. Indirect inference, nuisance parameter, and threshold moving average models. *Journal of Business & Economic Statistics*, 21:122 – 132, 2003.
- [5] Philip Rothman, Dick van Dijk, and Philip Hans. Multivariate star analysis of money–output relationship. *Macroeconomic Dynamics*, 5(4):506–532, 2001.
- [6] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun Woo. Convolutional lstm network: a machine learning approach for precipitation nowcasting. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*, NIPS’15, page 802–810, Cambridge, MA, USA, 2015. MIT Press.

-
- [7] Timo Teräsvirta Stefan Lundbergh and Dick van Dijk. Time-varying smooth transition autoregressive models. *Journal of Business & Economic Statistics*, 21(1):104–121, 2003.
 - [8] Timo Teräsvirta. Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89(425):208–218, 1994.
 - [9] Timo Teräsvirta, Dag Tjøstheim, and Clive W. J. Granger. *Modelling Nonlinear Economic Time Series*. Oxford University Press, 12 2010.
 - [10] Xinyu Zhang and Dong Li. Smooth transition moving average models: Estimation, testing, and computation. *Journal of Time Series Analysis*, (Sep), 2023.
 - [11] Jia Zhou. Smooth transition autoregressive models : A study of the industrial production index of sweden, 2010.