

# 基于组合核函数高斯过程回归的股指波动率预测研究

兰秋军, 周慧灵

(湖南大学工商管理学院, 湖南省、长沙市, 410000)

**摘要:** 股指期货是金融市场的重要组成部分, 对于风险对冲和复合投资策略的设计不可或缺, 而波动率是股指期货最重要的参数之一, 因此, 预测股指期货波动率具有重要的理论和实际意义。但是, 由于已实现波动率序列包含周期性, 不规则波动和时变性噪声等特征, 使得对其进行准确预测成为一个具有挑战性的难题。

本文在较广泛的国外和国内文献综述的基础上, 提出了一种新的基于组合核函数的高斯过程回归模型 (GPR-CPK)。该模型可以有效地提高波动率的预测精度并提供区间预测。该模型分为两个阶段: 组合核函数 CPK 的构建阶段和模型的超参数优化阶段。在组合核函数的构建阶段, 对高斯过程回归中不同的核函数进行分析、比较和组合, 基于其基核函数的基本性质和组合性并根据已实现波动率序列的性质选择最适合于波动率预测建模的核函数组合方式。具体来说, 采用高斯核函数、周期核函数、有理二次核和噪声核函数以和与乘积的组合方式构建了组合核函数, 以期提高波动率的预测精度和模型的泛化性能。在超参数优化阶段, 采用共轭梯度法求取训练样本对数似然函数的极大值以自适应地获得最优超参数。从而, 利用训练好的模型对已实现波动率进行预测获得预测点的均值和方差。

为了验证本文提出模型的有效性, 以十二组世界各国具有代表性的股票指数波动率作为实验数据展开了对比实验。实验结果表明, 无论是在点预测还是在区间预测, 本文提出的 GPR-CPK 模型均优于人工智能模型、计量经济模型和基于单核的高斯过程回归模型。

**关键词:** 股票指数; 已实现波动率; 高斯过程回归; 组合核函数

**中图分类号:** C93      **文献标识码:** A

## 1. 引言

股票指数在各国股票市场中扮演着重要的角色, 反映着其股票市场的总体价格水平及其变动趋势, 主要用于规避股票市场系统性波动风险。和其他金融产品及商品期货类似, 股指期货也有着内生的波动性, 且波动性是最重要的特征之一。因此, 运用合适的模型更加准确地估计和预测股指波动率, 对于金融风险管理、衍生品定价以及投资组合的构建等来说, 都具有十分重要的现实意义和应用价值<sup>[1-2]</sup>。因此, Andersen 和 Bollerslev (1998) 利用高频数据提出了以日内高频收益的平方和计算的已实现波动率作为真实波动率的估计量<sup>[3]</sup>。相较于低频数据, 已实现波动率是可以直接观测的显性变量, 包含了更多的日内交易信息。

众所周知, 股票价格的变化往往受到各种信息的传递以及资本流动的影响, 例如宏观经济<sup>[4]</sup>, 经济政策<sup>[5]</sup>和投资者情绪<sup>[6]</sup>等。这些因素导致了股指已实现波动率序列呈现出非线性变化的特征<sup>[7]</sup>。并且, 大量的金融实证研究表明股指波动率时常呈现出周期性变化特征。其次, 金融资产的波动序列不是白噪声序列, 即波动具有聚集性, 金融市场的较大的波动往往伴随着另一次的较大的波动; 而且对数收益序列存在尖峰厚尾现象, 并不服从于正态分布。

这一发现表明波动率建模过程中的残差呈现出异方差性<sup>[7]</sup>，因此在建模过程中也不能忽略残差对于模型预测效果的影响。

此外，现有文献中的大多数预测方法多为确定性预测，即只提供某一时刻的已实现波动率的点预测值。尽管确定性预测方法精确地估计了未来已实现波动率在某一时间点的取值，但它并未包含关于未来的不确定性信息，导致相关部门或机构不能对市场风险进行有效评估，降低了波动率预测的应用价值<sup>[8]</sup>。倘若可以提前获悉波动率的波动范围，准确地预测未来波动率会在何种程度的区间内变动，无疑对股票市场的风险管理、市场交易等起到至关重要的服务和指导作用<sup>[9]</sup>。

综上所述，波动率序列复杂多变且拥有众多特性，本文的研究目的是通过对股指波动率进行拟合及预测实证分析，找到对股指波动率具有较好解释能力和预测效果的模型，同时可以提供波动率的预测区间。在此基础上，依据模型所得预测结果分析股指已实现波动率变化及其对金融市场的影响，为投资者和市场风险管理者提供有效的参考。

## 2. 文献综述

从已实现波动率提出以来，以高频交易数据为研究样本进行波动率预测的研究得到了快速发展。现有的已实现波动率预测模型主要可以分为两类：计量经济学模型和机器学习模型。

### 2.1 计量经济学模型

计量经济模型是已实现波动率预测领域中广泛使用的模型。鉴于高频数据比低频数据更容易刻画波动率，有研究尝试将已实现波动率作为外生变量纳入到条件收益波动率模型中，如 SV-RV<sup>[10]</sup>和 GARCH-X<sup>[11]</sup>模型，但这类模型不能直接对已实现波动率进行预测。Hansen 等（2012）提出的 Realized GARCH 模型改进了不能直接预测已实现波动率这一缺点，该模型将已实现波动率和条件方差项通过一个测量方程连接起来，实现了模型的封闭。同时，其研究结果也表明加入已实现波动率的新模型的预测能力明显强于传统波动率模型<sup>[12]</sup>。国内学者王天一、黄卓研究了改进的 Realized GARCH 模型在中国股市高频数据波动率建模方面的应用<sup>[13]</sup>；刘威仪等（2016）在 GARCH-X 框架下，将波动率静态估计拓展成动态预测模型，发现将价格极差和开盘价信息纳入静态估计能提高 GARCH 模型的预测能力<sup>[14]</sup>。然而，无论是在成熟的资本市场或者是新兴的资本市场，已实现波动率序列表现出很强的长记忆性。虽然上述研究提出的模型改善了传统波动率模型的预测性能，但仍然不能捕捉到已实现波动率的长记忆性。

因此，为研究已实现波动率序列的长记忆性特征以提高预测的精度，Andersen 等（1998）提出了自回归分整移动平均过程模型（ARFIMA），指出 ARFIMA 模型可以进一步提高波动率的样本外预测精度相比起基于低频数据的长记忆参数模型<sup>[15]</sup>。然而，Corsi（1998）指出 ARFIMA 模型的估算过程非常复杂，缺乏清晰的经济解释<sup>[16]</sup>。因此，他提出了基于异质性市场假说<sup>[17]</sup>和 HARCH 模型<sup>[6]</sup>的 HAR 模型，实证结果表明 HAR 模型可以很好地描述已

实现波动率序列的长记忆性，并且线性模型的形式避免了 ARFIMA 模型的复杂参数估计。同时，国内学者魏宇（2010）综合比较了 ARFIMA 模型、GARCH 模型和 SV 模型对沪深 300 股指期货波动率的预测效果，并用 SPA 检验评估了各种波动率模型的预测精度，发现 ARFIMA 模型和 SV 模型的预测精度最高<sup>[18]</sup>。文风华等（2012）在 HAR 模型的基础上提出 LAHR 模型对沪深 300 指数的波动率跳跃现象进行研究，结果表明该模型能较好地捕捉中国股票市场波动率的长记忆性和杠杆效应<sup>[19]</sup>。陈浪南、杨科（2013）研究了 HAR 拓展模型对中国股市高频波动率预测效果，发现高频波动率的连续性成分具有很强的长记忆性，而跳跃性成分长记忆性较弱<sup>[20]</sup>。

上述模型虽然获得了较好的样本外预测结果，但都没有考虑到金融市场的复杂性和高频数据中的微观结构噪声，使得已实现波动率序列往往有着非线性和时变的特征。这样的特征是具有线性模型结构的计量模型所无法捕捉到的。

## 2.2 机器学习模型

随着信息技术的迅速发展，近年来人工智能方法受到了广泛的关注。其中，神经网络 (NN)<sup>[21]</sup>和神经网络的扩展模型包括高阶神经网络(HONN)<sup>[22]</sup>、进化混合神经模糊网络(eHFN)<sup>[23]</sup>和前馈神经网络(FNN)<sup>[24]</sup>是应用较为广泛的预测已实现波动率的方法。这些研究表明，神经网络能够有效地捕捉到已实现波动率中隐藏的非线性模式，相比起计量模型在一定程度上改善了波动率的预测效果。但神经网络以经验风险最小化作为计算准则，在理论上存在三个问题：网络结构难以确定、存在局部极小值、收敛速度慢。

与基于神经网络的方法相比，深度学习方法在神经网络内部具有反馈连接以学习时间模式，是刻画数据间复杂非线性关系的有效工具。特别是长短期记忆（Long Short-Term Memory, LSTM）模型不仅可以拟合数据的非线性关系，还能够很好地刻画时间序列的长记忆性<sup>[25]</sup>。因此，Kim 和 Won 使用 LSTM 来预测 KOSPI (Korea Composite Stock Price Index) 200 指数的已实现波动率<sup>[26]</sup>。Bucci 证明了 LSTM 相对于现有的计量模型可以获得更高的预测精度<sup>[7]</sup>。在国内，陈卫华首次运用 LSTM 对波动率进行样本外预测，以提高波动率的预测精度，并将预测结果与 19 种经典模型作对比以评价预测效果<sup>[27]</sup>。在此基础上，陈卫华和徐国祥将 LSTM 和股票论坛数据引入高频数据波动率预测中，提供了新的预测方法和预测思路<sup>[28]</sup>。上述研究为波动率预测提供了一种更精确和更稳健的实现方法，并拓展了深度学习在金融领域的运用。此外，一些研究提出了将神经网络与计量模型相结合的组合模型<sup>[29-30]</sup>，以便能够充分利用不同模型的预测信息捕捉已实现波动率序列既显示出线性特征和非线性特征。

深度学习方法凭借其复杂的多层网络结构结构可以很好地抓取数据见的特征，所以具有更强的预测能力相比较起传统模型。但是正因为如此，该类方法也就需要更多的数据来避免过拟合的发生。而金融市场受到交易日的限制，特别是对于中国等新兴金融市场，样本容量的有限性是一个普遍现象。所以将深度学习方法用到金融市场的波动率建模中也不可避免地会面临过拟合的问题。

### 3. 基于 GPR-CPK 的股指波动率预测模型构建

#### 3.1 高斯过程回归及核函数

##### 3.1.1 高斯过程回归模型定义

高斯过程回归具有结构简单、参数少、预测输出具有概率特性等特点，且在处理非线性、高复杂度的数据上有较好应用。它是通过处理训练数据的变化找到规律，并且估计生成先验分布，根据数据先验分布结合贝叶斯原理完成对后验分布的计算。因此高斯过程回归也可视作一种非线性概率模型，通过选用不同的核函数来无限逼近真实数据，在解决实际问题时，输出均值的同时给出置信区间，使结果的有效性不断增强。正是因为这些优势，高斯过程回归算法近几年来受到越来越多的关注，本文也在前人研究的基础上针对高斯过程回归展开研究。

高斯过程的定义为，在一个随机过程中，任意两个或多个随机变量都服从多维联合高斯分布，则这个随机过程被称为高斯随机过程，简称高斯过程[68]。

给定数据集， $D_i = \{(x_i, y_i) | i = 1, 2, \dots, n\}$ ，其中  $x_i$  代表训练集中的第  $i$  个输入向量， $y_i$  代表训练集的第  $i$  个输出， $n$  为训练集的样本数量。高斯过程回归建模就是对训练集中的输入向量与目标输出的函数关系  $f$  进行推断的过程，最后可以得到目标输出的条件分布。

高斯过程回归建立在一条重要假设基础上，即假设  $f(x_1), f(x_2), \dots, f(x_n)$  符合联合正态分布，且设该联合正态分布的均值向量为  $m$ ，协方差函数为  $k$ ，因此有

$$f \sim GP(m, k) \quad (3.1)$$

高斯过程在数据建模上的应用主要可以分为回归问题和分类问题：输出变量为有限个离散变量的预测称为高斯过程分类；输入数据和输出数据均为连续变量的预测问题称为高斯过程回归问题。下面从函数空间的角度对高斯过程回归进行阐述：

在 (2.3) 式中，设均值函数为  $m(x)$ ，协方差函数为  $k(x, x')$ ，则均值函数为  $m(x) = E(f(x))$ ，协方差函数  $k(x, x') = E\{[f(x) - m(x)][f(x') - m(x')]\}$ ，于是高斯过程可以描述为：

$$f \sim GP(m(x), k(x, x')) \quad (3.2)$$

根据贝叶斯原理，这里可以设一个贝叶斯回归函数： $f(x) = \phi(x)^T \omega$ ，其中  $\omega$  为权重向量，且  $\omega$  的先验分布服从  $\omega \sim N(0, \Sigma_p)$ ，根据先验分布可知  $E(\omega) = 0$ ， $E(\omega \omega^T) = \Sigma_p$ ，由此可得  $f(x)$  的数学期望和方差为：

$$\begin{aligned} E[f(x)] &= \phi(x)^T E(\omega) = 0 \\ E[f(x)]E[f(x')] &= \phi(x)^T E(\omega \omega^T) \phi(x') = \phi(x)^T \Sigma_p \phi(x') \end{aligned} \quad (3.3)$$

从式 (2.6) 可以看出，高斯过程  $f(x)$  与  $f(x')$  都服从联合高斯分布  $N(0, \phi(x)^T \Sigma_p \phi(x'))$ ，于是对任意  $k$  个向量输入  $x_1, \dots, x_k$ ，他们对应的函数值  $f(x_1, \dots, x_k)$  都服从联合高斯分布

$N(0, \phi(x)^T \Sigma_p \phi(x'))$ ，可以以平方指数函数作为其协方差函数，协方差的形式：

$$\text{cov}(f(x_p), f(x_q)) = k(x_p, x_q) = \exp\left(-\frac{1}{2}|x_p - x_q|^2\right) \quad (3.5)$$

不难看出式 (2.7) 是通过输入数据的协方差来表示输出数据的协方差的，也就是说通过输入随机变量的相关性来描述输出变量的相关性。首先考虑  $n$  对已知的输入输出数据（训练集） $X$  可知其对应函数值  $f$  服从高斯分布  $f \sim N(0, k(x, x'))$ ，然后在估计得函数分布进行采样得到个已知的输入  $X$  作为测试集，同样其输出  $f_*$  服从高斯分布  $f_* \sim N(0, k(x_*, x_*))$ 。在没有噪声的情况下，根据联合高斯分布的性质以及概率的知识可得，训练集的输出  $f$  和测试集的输出  $f_*$ 。服从如下联合分布：

$$\begin{bmatrix} f \\ f_* \end{bmatrix} = N\left(0, \begin{bmatrix} K(X, X) & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right) \quad (3.6)$$

上述联合分布中，协方差  $K(X, X_*)$  是  $n \times n_*$  的矩阵，协方差  $K(X, X)$  为  $n \times n$  的矩阵，协方差  $K(X_*, X)$  为  $n_* \times n$  的矩阵。根据贝叶斯原理，由联合高斯分布的先验表达式，可以推出  $f_*$  的条件后验分布为：

$$f_* | X_*, X, f \sim N(K(X_*, X)K(X, X)^{-1}f, K(X_*, X_*) - K(X_*, X)K(X, X)^{-1}K(X, X_*)) \quad (3.7)$$

在已知训练集的输入输出以及预测样本的输入时，便可通过上式来估计预测样本输出  $f_*$  的均值与协方差。因为噪声的混入，很有可能使观测值  $f_*$  受到影响，从而观测值与实际输出值  $f_*$  有一定的偏差。所以为了与实际情况相似，应该将噪声加入到  $f_*$  的概率分布中。

因为噪声的混入，观测值为  $y = f(x) + \varepsilon$ ，其中  $\varepsilon$  为高斯噪声且独立同分布，也符合高斯分布，即： $\varepsilon \sim N(0, \sigma^2)$ ，

$$\text{cov}(y) = K(X, X) + \sigma_n^2 I \quad (3.8)$$

上式中  $I$  为单位矩阵。所以含有噪声的观测值与函数预测值服从如下联合分布：

$$\begin{bmatrix} y \\ f_* \end{bmatrix} = N\left(0, \begin{bmatrix} K(X, X) + \sigma_n^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right) \quad (3.9)$$

将式 (2.9) 与式 (2.11) 进行对比，可推得输出  $f_*$  的后验分布即为高斯过程回归的预测输出，即：

$$f_* | X, y, X_* \sim N(\bar{f}_*, \text{cov}(f_*)) \quad (3.10)$$

式 (2.12) 中均值为：

$$\bar{f}_* = E[f_* | X, y, X_*] = K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1}y \quad (3.11)$$

方差为：

$$\text{cov}(f_*) = K(X_*, X_*) - K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1}K(X, X_*) \quad (3.12)$$

### 3.1.2 高斯过程回归的核函数

在高斯过程建模中可以选择不同种类的核函数，每一种类型的核函数都有着各自不同的结构，对数据处理的结果也不同，在 GPR 模型中，其性质完全由核函数决定[35]。另外，在第一章中介绍过高斯过程回归是非参数模型，但是 GPR 的核函数内部含有参数，称为超参数。本节主要介绍常用的核函数类型和核函数的组合方法，以及超参数的获取方式。

#### 3.1.2.1 单一核函数

高斯过程回归的均值核函数与协方差核函数统称为核函数。均值函数 $m$ 表示无任何观测值的情况下函数 $f$ 的期望。一般情况下均值函数 $m = 0$ ，表示对于任何输入，初始的函数输出总是理想化为0，但是这并不是必须的，均值函数可以设定为任何常数或其他形式的函数，但不管 $m$ 的大小或者函数形式如何， $m$ 始终表示函数值 $f$ 的期望。

协方差函数尤其是空间中两个随机变量输入点对应的随机变量输出的中心矩，用于衡量训练样本与测试样本之间的接近或者相似程度，是影响高斯过程回归预测性能的关键因素。因为均值函数一般都取为零均值函数，所以一般说核函数就是指协方差函数。

选择不同的核函数，将会直接影响 GP 模型所学习的函数类型，并影响其泛化能力 [35]。所以核函数的选择是依赖于具体问题的。核函数决定了度量两个样本相似性的方法。当专家针对待近似函数进行一定的假设，或希望在 GP 模型之中加入一定的先验知识的时候，可通过选择不同的核函数取得相应的效果。以下将介绍几种在回归问题中使用的核函数。

#### (1) 平方指数核函数

在文献中广泛使用的平方指数 (squared exponential, SE) 核函数具有以下形式：

$$k_{SE}(x, x') = \sigma_f^2 \exp\left(-\frac{2}{l^2} \|x - x'\|^2\right) \quad (3.13)$$

其中 $\sigma_f^2$ 与 $l$ 分别表示平方指数核函数的信号方差与放缩长度两个超参数。

上述核函数是无限可微的，则使用该核函数的 GP 将会非常光滑。应用该核函数，代表专家在模型中加入了待学习函数的一种光滑性先验假设。

#### (2) 有理平方核函数

有理平方 (Rational Quadratic, RQ) 核函数的形式为：

$$k_{RQ}(x, x') = \left(1 + \frac{\|x - x'\|^2}{2\alpha l^2}\right)^{-\alpha} \quad (3.14)$$

其中， $\alpha > 0$ 。有理平方核函数可以视为是具有不同放缩长度的平方指数核函数之和。当 $\alpha \rightarrow \infty$ 时，有理平方核函数的极限形式是一个平方指数核函数，其放缩长度为1。显然，有理平方核函数也是无限可微的，因此也是光滑的。

#### (3) 周期核函数

由于待近似函数可能存在一定的周期性，若需要在模型中加入周期性先验假设，可应用下列周期核 (Periodic, PER) 函数：

$$k_{PER}(x, x') = \sigma_f^2 \exp\left(-\frac{2}{l^2} \sin^2\left(\pi \frac{\|x - x'\|}{p}\right)\right) \quad (3.15)$$

#### (4) 线性核函数

前述核函数主要针对的是非线性情况，当待近似函数具有近似于线性的快速变化时，同样可以应用如下线性核函数：

$$k_{LIN}(x, x') = \sigma_f^2 (x - c)(x' - c) \quad (3.16)$$

可见，线性变化趋势同样可以应用 GP 模型进行良好地捕捉。

### 3.1.2.2 核函数的组合

协方差函数描述了各个数据之间的基本关系。在实际应用中，一个基本的核函数不能满足 GPR 的建模，也就是说需要两个或多个核函数的组合来完成 GPR 建模。所以，GPR 的核函数需要对相关的核函数进行处理和组合，通过组合构造出新的核函数来满足实际中的需求。因为数学上的基本运算方法可以运用到核函数的运算上，所以本课题所用的核函数基本上是由高斯过程的协方差函数组合而成。接下来介绍协方差函数的组合方法。

#### (1) 任意两个核函数之和仍然是一个新的核函数

证明：假设随机过程  $f(x) = f_1(x) + f_2(x)$ ， $f_1(x)$  和  $f_2(x)$  是独立的，根据两个独立随机过程的协方差函数运算法则，可得新的随机过程的协方差函数为  $k(x, x') = k_1(x, x') + k_2(x, x')$ 。所以任意不同特征的核函数都可以通过和的形式组成新的核函数。

#### (2) 任意两个核函数之积仍为一个新的核函数

证明：假设随机过程  $f(x) = f_1(x) \times f_2(x)$ ， $f_1(x)$  和  $f_2(x)$  是相互独立的，根据两个独立随机变量的性质可以计算新函数的协方差为  $k(x, x') = k_1(x, x') \times k_2(x, x')$ ，即两个核函数之积仍为一个新的核函数。对上式扩展，可以得到，对于任意  $p$  个高斯过程核函数  $k^p(x, x')$  之积，也能组合成为一个新的核函数。

上文介绍了核函数的运算法则，了解了这些运算法则之后，我们就可以根据这些运算法则构造出更符合实际要求的核函数，并且，新的核函数有拥有许多基本核函数的一般特性。所以，在使用高斯过程建模的时候，一定要根据原始数据的基本特征，对比基本核函数的性质，使用核函数的组合方法，来构造出符合实际要求的核函数。

## 4. 基于 GPR-CPK 的股指波动率预测实验结果分析

前文对组合核函数高斯过程回归模型的理论和方法进行了阐述分析。本章将通过实证研究分析该模型的效果。首先，选取了具有代表性的世界各国的股票指数作为本次实验的研究对象；其次，从点预测和区间预测两个方面，选取了当下热门的人工智能模型和计量经济模型在多组数据中进行了对比实验，最后验证了模型结果的有效性。

### 4.1 样本的获取与描述性统计

本文选取了美国、英国、日本、韩国、德国、法国、澳大利亚、欧洲、中国香港和中国等国家和地区的十二组主要股票市场指数：美国的道琼斯工业平均指数（DJIA）和标准普尔 500 指数（S&P500）、英国的金融时报指数（FTSE）、日本的日经 225 指数（N225）、韩国综合指数（KOSPI）、德国 DAX 指数（DAX）、法国 CAC40 指数（FCHI）、澳大利亚普通股指数（AORD）、欧洲斯托克 50 指数（STOXX50）、香港的恒生成分股指数（HSI）以及中国的上证综指（SCI）和沪深 300 指数（CSI 300）。其中国际股票指数来自于 OMI 量化金融研究院（Oxford-Man Institute of Quantitative Finance）提供的 5 分钟股指收益率数据，而中国的股票指数则来自于万得数据库提供的 5 分钟高频收益率数据。众所周知，万得是一家金融数据与分析工具服务提供商，为国内金融证券市场各交易品种提供实时报价、信息与数

据分析与提取工具。样本数据为 2017 年到 2019 年的 5 分钟收益率数据，其中 2017 - 2018 年的数据作为样本内数据，2019 年的数据作为样本外数据。由于时差的原因，各国或地区的股票市场存在交易不同步的现象，因此各组样本数据用于样本外预测的数量有所不同，但观察值时间间隔都是 3 年的交易时间。我们使用 5 分钟股指收益率数据计算得到已实现波动率，基于这一时间段的数据对比不同模型在对股票市场波动率上的预测精度，能够有效判别模型的有效性，并对今后股票市场波动率预测实践给予较强的指导。

表 1 是各股指波动率的统计描述表。如表 1 所示，股指已实现波动率不符合正态分布，其序列出现严重偏态和过剩峰度。相比之下，对数股指已实现波动率 RV 峰度更低没有出现偏锋情况，其更接近正态分布。为此，根据高斯过程回归模型的特点，我们在后续建模时将考虑使用对数已实现波动率作为数据样本。

表 4.1 股指已实现波动率和对数已实现波动率序列的描述性统计结果

| RV      | 最小值     | 最大值     | 平均值    | 标准差    | 偏度      | 峰度       |
|---------|---------|---------|--------|--------|---------|----------|
| SCI     | 0.0589  | 12.3487 | 0.7816 | 1.3324 | 5.5288  | 37.6391  |
| CSI 300 | 0.0619  | 18.0139 | 1.1297 | 1.8003 | 4.6813  | 27.7810  |
| AORD    | 0.0000  | 0.0005  | 0.0000 | 0.0000 | 7.3861  | 78.6145  |
| DAX     | 0.0000  | 0.0004  | 0.0000 | 0.0000 | 3.2871  | 17.9430  |
| DJIA    | 0.0000  | 0.0010  | 0.0000 | 0.0001 | 5.9427  | 51.5672  |
| FCHI    | 0.0000  | 0.0004  | 0.0003 | 0.0081 | 3.1641  | 15.5523  |
| FTSE    | 0.0000  | 0.0011  | 0.0000 | 0.0001 | 10.2583 | 171.4883 |
| HSI     | 0.0000  | 0.0003  | 0.0000 | 0.0000 | 2.7548  | 10.9596  |
| KOSPI   | 0.0000  | 0.0002  | 0.0000 | 0.0000 | 3.3403  | 16.6041  |
| S&P500  | 0.0000  | 0.0007  | 0.0000 | 0.0001 | 4.7349  | 29.7568  |
| STOXX50 | -0.0369 | 0.0378  | 0.0002 | 0.0079 | 4.2883  | 30.0127  |
| N225    | 0.0000  | 0.0007  | 0.0000 | 0.0001 | 5.8290  | 53.8207  |
| Log RV  | 最小值     | 最大值     | 平均值    | 标准差    | 偏度      | 峰度       |
| SCI     | 0.0000  | 0.8266  | 0.3189 | 0.1386 | 0.8407  | 1.0650   |
| CSI 300 | 0.0000  | 0.8002  | 0.3305 | 0.1363 | 0.6179  | 0.4052   |
| AORD    | 0.0000  | 0.9098  | 0.3257 | 0.1201 | 0.9745  | 2.1813   |
| DAX     | 0.0522  | 0.7649  | 0.3724 | 0.1166 | 0.0763  | 0.0560   |
| DJIA    | 0.0000  | 0.7801  | 0.3094 | 0.1329 | 0.5250  | 0.0281   |
| FCHI    | 0.0037  | 0.7247  | 0.3355 | 0.1102 | 0.3306  | 0.0851   |
| FTSE    | 0.0000  | 0.8037  | 0.3850 | 0.0860 | 0.3693  | 1.3382   |
| HSI     | 0.0308  | 0.7532  | 0.3852 | 0.1111 | 0.3229  | 0.1386   |
| KOSPI   | 0.0500  | 0.7439  | 0.3315 | 0.1162 | 0.3871  | 0.5975   |
| S&P500  | 0.0000  | 0.7852  | 0.3534 | 0.1348 | 0.5506  | 0.0829   |
| STOXX50 | 0.0000  | 0.8255  | 0.6230 | 0.0595 | -1.3783 | 15.5995  |
| N225    | 0.0000  | 0.8313  | 0.3667 | 0.1120 | 0.5408  | 0.9112   |

### 4.3 模型评估指标和对比模型

对 GPR-CPK 模型的区间预测结果的评价应从两个方面进行考虑。首先, 预测结果的期望值应尽可能接近实际值, 这可以参考传统的点预测算法评价指标。其次, 预测区间的宽度要尽可能小, 覆盖率要尽可能大。基于以上考虑, 本文首先采用损失函数和预测能力检验方法对预测值的精度进行评估, 其次选取了常用区间预测指标进一步地对预测区间的质量进行评价。

### 4.3.1 损失函数

从现有的文献来看, 大多数研究在比较波动率模型预测能力都使用平均绝对值误差 MAE, 均方根误差 RMSE 和平均绝对百分比误差 MAPE。三种损失函数的计算方法如下:

$$MAE = \frac{1}{T} \sum |\widehat{RV}_t - RV_t| \quad (4.1)$$

$$RMSE = \frac{1}{T} \sum (\widehat{RV}_t - RV_t)^{1/2} \quad (4.2)$$

$$MAPE = \frac{100\%}{T} \sum |1 - \widehat{RV}_t / RV_t| \quad (4.3)$$

式中:  $\widehat{RV}_t$  为  $t$  时刻预测的波动率,  $RV_t$  为  $t$  时刻实现的波动率,  $T$  为预测的完整时间长度。

### 4.3.3 区间预测指标

为了定性地评价预测模型的性能, 需要考虑预测区间的可靠性和平均宽度以及负荷平均值的精度, 因此, 选用预测区间覆盖率、预测区间归一化宽度和综合指标这三个指标作为模型的评价指标<sup>[70, 71]</sup>, 其定义如下。

#### (1) 预测区间覆盖率 (Prediction Interval Coverage Probability, PICP)

PICP 表示的是实际值落在预测区间的统计概率。根据 PICP 的定义, 真实预测点将按照规定的概率 (置信度  $((1 - \alpha)\%)$ ) 落在预测区间内。理想情况下区间对真实值的覆盖率将大于等于规定的置信度。因此, 预测区间覆盖率 (PICP) 对于评估模型所构建的区间的可信度具有重要意义, 其计算方式如下:

$$PICP = \frac{1}{n} \sum_{i=1}^n c_i \quad (4.4)$$

式中:  $n$  为预测样本中的数据个数;  $c_i$  为预测结果的判别指标, 当真实值  $RV_i$  位于给定的上下边界之间时, 其取值为 1, 否则取值为 0, 如公式所示。

$$c_i = \begin{cases} 1, & RV_i \in [L(RV_i), U(RV_i)] \\ 0, & RV_i \notin [L(RV_i), U(RV_i)] \end{cases} \quad (4.5)$$

式中:  $U(RV_i)$  和  $L(RV_i)$  分别为  $RV_i$  对应的预测区间的上、下限。

不难看出, PICP 指标能够直观反映区间预测结果的准确性, 其数值越高, 表明负荷的真实值落在预测区间中的概率越大, 预测结果越好。因此, 如果 PICP 的值远小于其置信度则证明该 PICP 对应的模型所产生的预测区间准确度不高。

#### (2) 预测区间归一化平均宽度 (Normalized Mean Prediction Interval Width, NMPIW)

如果模型产生的预测区间过于宽, 甚至在这种情况下 PICP 达到 100%。那么该区间虽

然包含了所有的真实值但是趋于保守，不能提供有关预测点区间变动不确定性的有效信息，也就是说该预测区间没有实用价值。因此，我们需要再寻找一种指标来反映区间的宽度。

预测区间平均宽度（Mean Prediction Interval Width, MPIW）表示的是样本中所有数据对应的区间宽度均值，其计算公式如下式所示。

$$MPIW = \frac{1}{n} \sum_{i=1}^n (U(RV_i) - L(RV_i)) \quad (4.6)$$

式中： $U(RV_i)$ 和 $L(RV_i)$ 分别为 $RV_i$ 对应的预测区间的上、下限， $n$ 为预测样本中的数据个数。为了保证可以对不同模型产生的预测区间进行统一的比较，我们使用区间的平均宽度对 MPIW 进一步地进行归一化。

则预测区间归一化平均宽度（Normalized Mean Prediction Interval Width, NMPIW）的计算公式如下。

$$MPIW = \frac{1}{nR} \sum_{i=1}^n (U(RV_i) - L(RV_i)) \quad (4.7)$$

式中， $R$ 为区间的平均宽度。一般来说，NMPIW 和 PICP 有着直接的关系，当得到的预测区间越宽的时候也意味着该区间可以包含更多的真实值。因此在同等条件下，较大的 NMPIW 通常会导致较高的 PICP。

### (3) 综合指标（Coverage Width-based Criterion, CWC）

预测精度较高的预测区间是具有较高的覆盖概率（较大的 PICP）的窄预测区间（较小的 NMPIW）。然而，从理论上来看这两个目标是相互矛盾的。因为一旦预测区间的宽度减小，部分点的真实值就会处于区间之外，这样也就导致了 PICP 的减小。因此，为了找到具有合适 NMPIW 和 PICP 的预测区间，需要一种综合指标来定量表示二者的大小权衡情况。由于 PICP 是衡量预测区间可信度的基本特征，因此综合指标应更加重视 PICP 的变化。换句话说，对于 PICP 小于其相应的置信度  $((1 - \alpha)\%)$  的情况，综合指标也应有较大的变化幅度。综合指标（CWC）的提出解决了以上问题，该指标计算方式如下。

$$CWC = NMPIW \times (1 + \gamma \times e^{(-\eta \times (PICP - \mu))}) \quad (4.8)$$

$$\gamma = \begin{cases} 0, & PICP \geq \mu \\ 1, & PICP < \mu \end{cases} \quad (4.9)$$

式中：其中  $\eta$  和  $\mu$  是两个超参数。 $\eta$  决定在低覆盖概率的预测区间分配多少惩罚。 $\mu$  对应于与区间相关的置信水平，通常设置为  $(1 - \alpha)$ 。 $\eta$  的作用是放大 PICP 和  $\mu$  之间的差距，通常其取值应较大。

$\gamma$  则是判断是否需要当前评估的预测区间进行惩罚的参数。因此当 PICP 低于  $\mu$  时，此时若不对预测区间进行相应惩罚，将导致出现误导性的预测区间，所以当  $PICP < \mu$  时， $\gamma = 1$  以保证对区间处以指数项的惩罚。反之，当 PICP 大于  $\mu$  时，则不需要对区间进行惩罚，此时综合指标 CWC 的值就等于区间的平均宽度 NMPIW。总的来说，CWC 值越小，区间预测结果越好。

## 4.4 单值预测实验结果及分析

#### 4.4.1 损失函数比较

为了验证本文构建的 GPR-CPK 模型在波动率点预测方面的性能, 本文使用十二组世界各国股指波动率数据进行实验, 分别比较了计量经济模型中的典型模型: 朴素贝叶斯 (Naive), HAR 和人工智能模型中的 LSTM 模型在波动率点预测中的结果。此外, 为了突显组合核函数的优越性, 本文还选择了带有高斯核的高斯过程回归模型 (GPR-SE)、带有周期核的高斯过程回归模型 (GPR-PER) 和带有有理二次核的高斯过程回归模型 (GPR-RQ) 作为比较模型, 从而说明使用组合核函数在提升预测精度方面的优越性。

对于股指已实现波动率的点预测, 我们使用三种损失函数 RMSE, MAE 和 MAPE 作为结果比较的衡量指标, 表 4.2、4.3 和 4.4 分别记录了不同模型在各个指数中的预测结果得到的 RMSE、MAE 和 MAPE 值。根据三种损失函数, 可以看出各国指数的波动率预测结果中, RMSE 和 MAE 结果均显示本文构建的 GPR-CPK 模型的预测精度优于 HAR、Naive、LSTM、GPR-SE、GPR-PER 和 GPR-RQ 模型。虽然在 MAPE 结果中显示 HAR 模型在一些股指数据中预测损失较小, 但是 GPR-CPK 模型的 MAPE 与其相差不大。因此, 根据损失函数比较的结果, 可以发现在大多数情况下 GPR-CPK 模型的预测能力较强。

详细来说, 在除 HAR 模型之外的其他对比模型中, GPR-CPK 在十二组股指数据上的预测结果均取得了最小的损失函数值, 从而也显示了 GPR-CPK 模型较好的预测性能。特别是将 GPR-CPK 与 LSTM 在 DJIA 指数上的预测结果进行比较, 其 MAPE 大约减小了 25%。因此从结果也可以反映出当面对小样本时, GPR-CPK 可以缓解因过度拟合而导致预测不准确的问题。

此外, 从和基于单一核函数的 GPR 模型的对比结果可以看出, GPR-CPK 的损失函数值明显低于基于单一核函数的 GPR 模型。因此, 本文根据波动率性质设计的组合核函数是适用于波动率预测建模的, 并且其预测性能优于单一核函数 GPR 模型。

表 4.2 RMSE

| 模型     | RMSE          |        |               |        |        |        |         |
|--------|---------------|--------|---------------|--------|--------|--------|---------|
|        | GPR-CPK       | Naive  | HAR           | LSTM   | GPR-SE | GPR-RQ | GPR-PER |
| SZ     | 0.0787        | 0.0857 | <b>0.0774</b> | 0.0949 | 0.0815 | 0.0809 | 0.0809  |
| HS     | 0.0974        | 0.1185 | <b>0.0967</b> | 0.1249 | 0.1008 | 0.0999 | 0.1042  |
| AORD   | <b>0.1187</b> | 0.1471 | 0.1190        | 0.1271 | 0.1197 | 0.1197 | 0.1196  |
| DAX    | <b>0.0919</b> | 0.1120 | 0.0926        | 0.1084 | 0.0932 | 0.0932 | 0.0952  |
| DJIA   | <b>0.0768</b> | 0.0874 | 0.0768        | 0.0974 | 0.0770 | 0.0775 | 0.0795  |
| FCHI   | <b>0.1011</b> | 0.1248 | 0.1018        | 0.1138 | 0.1014 | 0.1019 | 0.1020  |
| FTSE   | <b>0.0654</b> | 0.0815 | 0.0664        | 0.0804 | 0.0663 | 0.0672 | 0.0657  |
| HSI    | 0.0814        | 0.1023 | <b>0.0812</b> | 0.0969 | 0.0831 | 0.0831 | 0.0872  |
| KOSPI  | <b>0.0820</b> | 0.1020 | 0.0826        | 0.0972 | 0.0827 | 0.0835 | 0.0828  |
| S&P500 | 0.0836        | 0.0923 | <b>0.0830</b> | 0.0931 | 0.0848 | 0.0851 | 0.0958  |

|         |               |        |               |        |        |        |        |
|---------|---------------|--------|---------------|--------|--------|--------|--------|
| STOXX50 | 0.0507        | 0.0635 | <b>0.0506</b> | 0.0649 | 0.0529 | 0.0517 | 0.0514 |
| N225    | <b>0.0837</b> | 0.1007 | 0.0844        | 0.0992 | 0.0844 | 0.0863 | 0.0843 |

表 4.3 MAE

| 模型      | MAE           |        |               |        |        |        |         |
|---------|---------------|--------|---------------|--------|--------|--------|---------|
|         | GPR-CPK       | Naïve  | HAR           | LSTM   | GPR-SE | GPR-RQ | GPR-PER |
| SZ      | 0.0603        | 0.0683 | <b>0.0595</b> | 0.0760 | 0.0615 | 0.0626 | 0.0617  |
| HS      | <b>0.0709</b> | 0.0904 | 0.0714        | 0.0931 | 0.0735 | 0.0724 | 0.0776  |
| AORD    | 0.0909        | 0.1112 | <b>0.0900</b> | 0.0962 | 0.0915 | 0.0914 | 0.0915  |
| DAX     | <b>0.0715</b> | 0.0855 | 0.0720        | 0.0843 | 0.0721 | 0.0721 | 0.0731  |
| DJIA    | 0.0626        | 0.0708 | <b>0.0621</b> | 0.0773 | 0.0628 | 0.0624 | 0.0642  |
| FCHI    | <b>0.0798</b> | 0.0972 | 0.0807        | 0.0935 | 0.0805 | 0.0806 | 0.0871  |
| FTSE    | <b>0.0496</b> | 0.0640 | 0.0505        | 0.0631 | 0.0504 | 0.0508 | 0.0495  |
| HSI     | 0.0627        | 0.0781 | <b>0.0620</b> | 0.0759 | 0.0644 | 0.0644 | 0.0673  |
| KOSPI   | <b>0.0638</b> | 0.0826 | 0.0646        | 0.0772 | 0.0645 | 0.0647 | 0.0651  |
| S&P500  | 0.0670        | 0.0741 | <b>0.0664</b> | 0.0721 | 0.0674 | 0.0677 | 0.0735  |
| STOXX50 | <b>0.0376</b> | 0.0472 | 0.0377        | 0.0515 | 0.0404 | 0.0388 | 0.0383  |
| N225    | <b>0.0664</b> | 0.0797 | 0.0665        | 0.0807 | 0.0665 | 0.0680 | 0.0678  |

表 4.4 MAPE

| 模型      | MAPE         |              |              |       |        |        |         |
|---------|--------------|--------------|--------------|-------|--------|--------|---------|
|         | GPR-CPK      | Naïve        | HAR          | LSTM  | GPR-SE | GPR-RQ | GPR-PER |
| SZ      | 17.11        | 19.64        | <b>16.84</b> | 21.20 | 16.90  | 17.81  | 17.24   |
| HS      | <b>17.82</b> | 23.13        | 18.39        | 22.91 | 17.82  | 17.96  | 19.11   |
| AORD    | 32.92        | 39.30        | <b>32.52</b> | 35.87 | 33.12  | 33.09  | 33.08   |
| DAX     | 22.54        | 25.38        | <b>21.80</b> | 26.62 | 23.11  | 23.11  | 23.21   |
| DJIA    | 22.75        | 23.93        | <b>21.56</b> | 27.63 | 22.82  | 23.14  | 23.37   |
| FCHI    | 55.00        | <b>49.40</b> | 52.81        | 60.03 | 56.85  | 55.59  | 54.99   |
| FTSE    | <b>12.96</b> | 16.33        | 13.08        | 16.31 | 13.14  | 13.17  | 12.89   |
| HSI     | 18.18        | 21.62        | <b>16.69</b> | 22.45 | 19.14  | 18.67  | 20.31   |
| KOSPI   | 18.61        | 23.74        | <b>18.27</b> | 22.51 | 18.86  | 18.89  | 18.88   |
| S&P500  | 22.28        | 22.15        | <b>20.96</b> | 22.84 | 23.25  | 23.35  | 24.68   |
| STOXX50 | <b>6.13</b>  | 7.63         | 6.29         | 8.39  | 6.64   | 6.34   | 6.25    |
| N225    | 22.01        | 24.83        | <b>21.18</b> | 27.06 | 22.36  | 22.90  | 22.54   |

为了进一步直观地比较不同模型点预测的结果，图 4.2 和 4.3 分别记录了十二组股指波动率对应七个模型的预测结果。为了图片展示效果更加清晰明了，本文将十二组数据拆分成两张图，其中每张图中的一个小组对应了一个股指数据的预测结果。图中黑色虚线代表了股

指波动率的真实值，而其他颜色则一一对应了其他模型（见图例）。从图 4.2 和 4.3 可以看出，GPR-CPK 模型的预测结果较为接近真实值，且出现极端预测值的情况较少。因此，该图片很好地反映了 GPR-CPK 在十二组波动率数据上的预测性能是较为优越的。

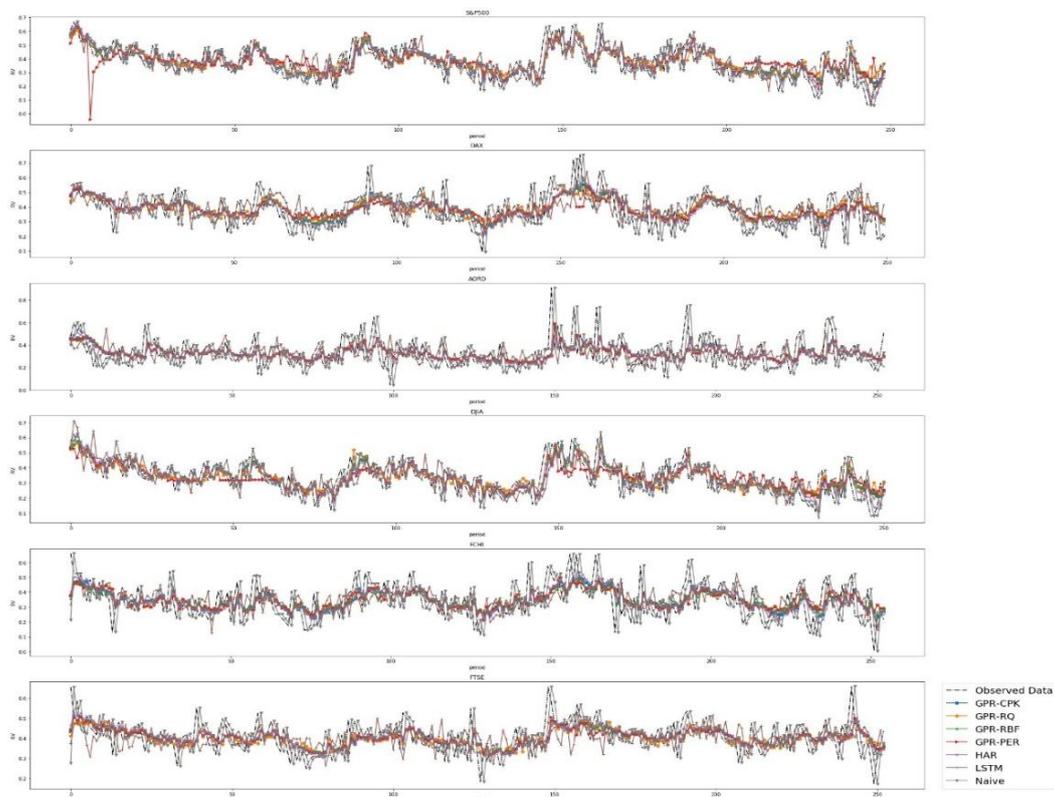


图 4.1 股指波动率点预测结果（部分一）

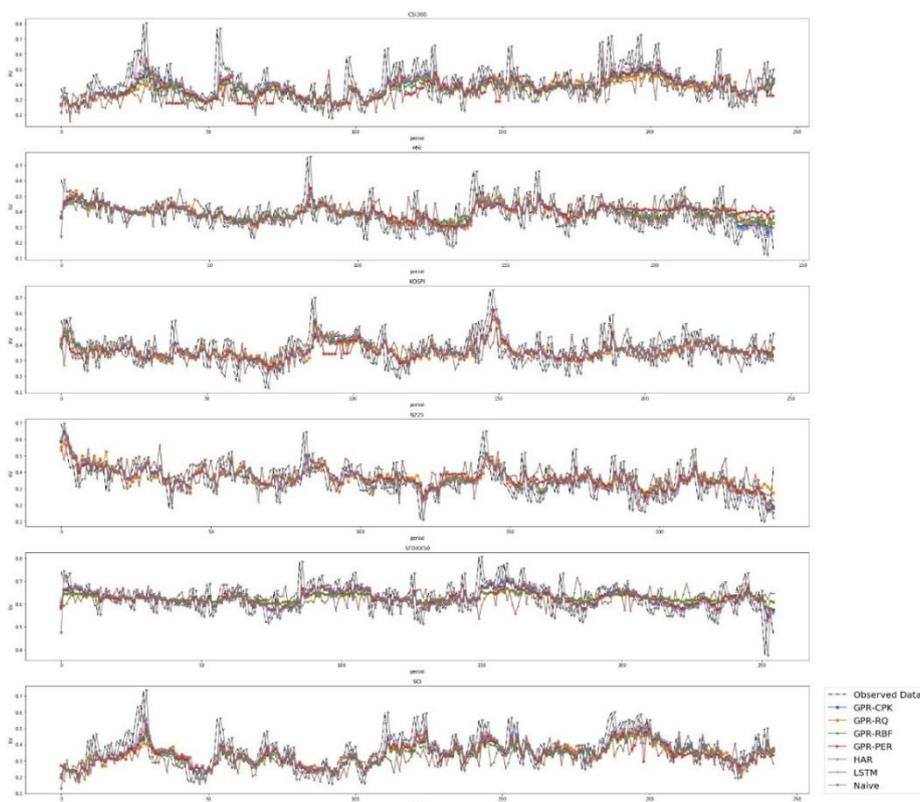


图 4.2 股指波动率点预测结果（部分二）

## 4.5 区间预测实验结果及分析

### 4.5.1 区间预测结果比较

从上述章节的点预测结果可以看出，GPR-CPK 模型比 LSTM 模型和 Naïve 方法有着更高的预测精度，并且通过 DM 检验也证明了其有着更为优越的预测能力。然而，在与 HAR 模型比较中，虽然在损失函数值的比较中 GPR-CPK 模型在数值上更胜一筹，但 DM 检验的结果显示两模型的预测能力没有差异。因此有必要进一步比较 GPR-CPK 与 HAR 模型的区间预测能力。此外，GPR 模型的优点就在于其可以直接提供预测区间，因此为了验证组合核函数在区间预测上的优越性，也需要与单一核函数的 GPR 模型进行对比。

因此，为验证本文提出模型的有效性，选取 95% 的置信水平对 2019 年的各国股指波动率进行区间预测。各种模型在不同股指下的区间预测结果如表 4.6、4.7、4.8 所示。通过对比各模型在不同股指数据下的区间预测结果 PICP、PINAW 和 CWC 可以得出以下结论。

首先，表 4.6 和 4.7 中区间覆盖率 PICP 和区间平均宽度 PICP 的结果显示，GPR-CPK 和 HAR 模型具有差不多的区间归一化宽度 NMPIW，而模型 GPR-CPK 的区间覆盖率 PICP 在大多数情况下是更高的。根据区间覆盖率 PICP 的定义，覆盖率代表了模型能否保证产生的预测区间包含的真实值的比例是合理的，这就需要比较模型区间对应的 PICP 是否大于规定的置信水平。在本文的实验中，所有模型的置信水平都为 95%，因此为了直观地显示 PICP 达标的结果，本文将小于 95% 的 PICP 采用双下划线并加粗的形式进行表示。因此，根据表 4.6 中划线的结果，HAR 模型和基于单一核函数的 GPR 模型——GPR-RBF 模型、GPR-PER 模型和 GPR-RQ 模型的区间覆盖率 PICP 有多组股指数据小于置信水平 95%，这反映了模型产生的区间的可信度是较低的。特别是对于 HAR 来说，近 60% 的数据集的 PICP 不大于 95%，这说明 HAR 模型的预测区间在大多数情况下不满足 95% 的置信区间要求。因此，与 HAR 模型相比，GPR-CPK 模型可以保证相对较高的 PICP，同时保证更窄的 NWPIW。

为了对比单一核函数和组合核函数在区间预测上的性能，分别使用基于 RBF、PER 和 RQ 单核函数和组合核函数 CPK 四种 GPR 模型进行股指区间预测。从表 4.6 和 4.7 对比结果可以得到，组合核函数 GPR-CPK 模型的 PICP 和 NWPIW 预测区间指标基本上均优于单一核函数 SE、PER 和 RQ。其中，与 GPR-RBF 模型相比，GPR-CPK 模型在大多数数据集中具有更高的区间覆盖率 PICP 和更低的 NMPIW 区间宽度，这意味着本文构建的组合核函数在波动率预测方面优于使用高斯核的高斯过程回归模型。此外，与 GPR-PER 和 GPR-RQ 模型相比，GPR-CPK 具有较高的 PICP 和较小的 NWPIW。这表明组合核函数 GPR-CPK 模型能够提高单核函数的预测效果，主要原因是已实现波动率序列有着周期性、不规则波动和时变性噪声的性质，而组合核函数可以更加充分地获取数据中的各种特性，从而证明了本文构建的组合核函数可以提高区间预测精度。

从现实角度来看,我们希望模型得到的预测区间是具有较高的覆盖率(即较大的PICP),并且该区间的平均宽度较窄(即较小的NMPIW)非常重要。但是,理论上这两个目标是相互矛盾的,因为减少预测区间的宽度通常会导致PICP减小,这是由于一些真实值由于区间宽度的减小而无法被包含在预测区间中。因此,很难根据PICP和NMPIW两个方面的大小来判断哪个模型更好。这就需要一个综合指标CWC来综合权衡PICP和NWPIW,所以从PICP和NMPIW的模型比较结果我们可以了解各模型得到的预测区间的大体情况。而衡量预测区间最重要的指标则是综合指标CWC,从其定义可知当CWC越小时说明该模型的预测区间越佳,即该模型的区间预测性能越好。从表4.8的CWC结果可得,GPR-CPK模型在接近90%的股指数据集上的CWC明显小于其他模型。该结果表明,GPR-CPK模型得到的预测区间更具有实际意义。

综上所述,本文构建的GPR-CPK模型在95%的置信水平下和各组股指数据中均能保持良好的区间预测性能,其主要原因有:高斯过程回归模型本身就是一种无参数的概率性预测方法,由该模型可以得到较为合理的预测区间。而本文采用的组合核函数可以很好地捕捉股指波动率特征,从而相比起其他模型获得了更好的预测结果,进一步提高了预测精度。

表 4.6 PICP

| 模型      | PICP                 |                      |                      |                      |                      |
|---------|----------------------|----------------------|----------------------|----------------------|----------------------|
|         | GPR-CPK              | HAR                  | GPR-SE               | GPR-RQ               | GPR-PER              |
| SZ      | 95.47%               | 96.30%               | 100.00%              | 95.88%               | 95.88%               |
| HS      | 97.53%               | <b><u>93.83%</u></b> | 100.00%              | <b><u>93.83%</u></b> | <b><u>94.24%</u></b> |
| AORD    | 97.23%               | <b><u>94.07%</u></b> | <b><u>91.30%</u></b> | <b><u>91.70%</u></b> | <b><u>93.68%</u></b> |
| DAX     | 99.20%               | <b><u>90.80%</u></b> | 100.00%              | 100.00%              | <b><u>92.00%</u></b> |
| DJIA    | 100.00%              | <b><u>94.42%</u></b> | <b><u>91.63%</u></b> | 95.22%               | 98.01%               |
| FCHI    | 98.43%               | <b><u>92.16%</u></b> | 100.00%              | <b><u>92.55%</u></b> | <b><u>92.55%</u></b> |
| FTSE    | 96.05%               | 96.44%               | 100.00%              | 96.44%               | 97.23%               |
| HSI     | 95.02%               | 95.02%               | 100.00%              | 95.02%               | 96.27%               |
| KOSPI   | 95.51%               | 95.92%               | 100.00%              | 95.51%               | 95.51%               |
| S&P500  | 100.00%              | 94.38%               | 100.00%              | 90.76%               | 95.18%               |
| STOXX50 | 96.47%               | 95.69%               | 100.00%              | 95.29%               | 96.08%               |
| N225    | <b><u>94.56%</u></b> | <b><u>94.14%</u></b> | 100.00%              | <b><u>94.14%</u></b> | <b><u>94.98%</u></b> |

表 4.7 NMPIW

| 模型 | NMPIW   |             |        |        |         |
|----|---------|-------------|--------|--------|---------|
|    | GPR-CPK | HAR         | GPR-SE | GPR-RQ | GPR-PER |
| SZ | 0.58    | <b>0.57</b> | 2.46   | 0.59   | 0.6     |
| HS | 0.85    | <b>0.56</b> | 2.27   | 0.58   | 0.62    |

|         |             |             |             |             |      |
|---------|-------------|-------------|-------------|-------------|------|
| AORD    | 0.63        | 0.49        | <b>0.44</b> | 0.44        | 0.5  |
| DAX     | 0.8         | <b>0.51</b> | 2.13        | 2.13        | 0.53 |
| DJIA    | 0.95        | <b>0.54</b> | 0.49        | 0.56        | 0.71 |
| FCHI    | 0.81        | 0.54        | 2.15        | <b>0.54</b> | 0.54 |
| FTSE    | <b>0.55</b> | 0.61        | 2.9         | 0.61        | 0.66 |
| HSI     | <b>0.52</b> | 0.53        | 2.23        | 0.52        | 0.56 |
| KOSPI   | <b>0.55</b> | 0.56        | 4.51        | 0.56        | 0.58 |
| S&P500  | 0.85        | 0.49        | <b>0.47</b> | 0.5         | 3.16 |
| STOXX50 | <b>0.48</b> | 0.49        | 3.28        | 0.49        | 0.49 |
| N225    | 0.53        | 0.54        | 4.07        | <b>0.46</b> | 0.63 |

表 4.8 CWC

| 模型      | CWC         |             |        |             |         |
|---------|-------------|-------------|--------|-------------|---------|
|         | GPR-CPK     | HAR         | GPR-SE | GPR-RQ      | GPR-PER |
| SZ      | 0.58        | <b>0.57</b> | 2.46   | 0.59        | 0.6     |
| HS      | <b>0.85</b> | 1.18        | 2.27   | 1.22        | 1.29    |
| AORD    | <b>0.63</b> | 1.03        | 1.07   | 1.05        | 1.07    |
| DAX     | <b>0.80</b> | 1.3         | 2.13   | 2.13        | 1.25    |
| DJIA    | 0.95        | 1.11        | 1.18   | <b>0.56</b> | 0.71    |
| FCHI    | <b>0.81</b> | 1.26        | 2.15   | 1.23        | 1.33    |
| FTSE    | <b>0.55</b> | 0.61        | 2.9    | 0.61        | 0.66    |
| HSI     | <b>0.52</b> | 0.53        | 2.23   | 0.52        | 0.56    |
| KOSPI   | <b>0.55</b> | 0.56        | 1.02   | 0.56        | 0.58    |
| S&P500  | <b>0.85</b> | 1           | 1.12   | 1.27        | 3.16    |
| STOXX50 | <b>0.48</b> | 0.49        | 3.28   | 0.49        | 0.49    |
| N225    | <b>0.96</b> | 1.13        | 0.99   | 0.96        | 1.11    |

由表 4.7 的各模型区间归一化宽度 CWC 结果可知, GPR-CPK 和 HAR 的 CWC 值相差不大。但如表 4.6 所示, 我们仍然可以得出 GPR-CPK 所构建的区间位置比 HAR 更合适, 因为两者区间在宽度差不多的情况下 GPR-CPK 模型的 PICP 更高, 这说明了该模型构建的区间可以包含更多的真实值。此外, 为了更加直观地显示本文构建的组合核函数的优势, 本文选取了亚欧地区 SCI、HSI、STOXX50 和 FTSE 指数的区间预测结果将基于单核函数的 GPR 模型和 GPR-CPK 模型进行对比显示, 详细情况见图 4.4 到 4.7。其中图中的黑色虚线代表各指数已实现波动率的观测值, 蓝色填充的区域代表 GPR-CPK 模型的预测区间, 而橙色、绿色和红色线条之间的区域则分别代表 GPR-RQ 模型、GPR-RBF 模型和 GPR-PER 模型。从图中可以看出, 与其他基于单核函数的 GPR 模型相比, GPR-CPK 的预测区间大多较窄。此外, GPR-CPK 的区间包含的真实值在基本高于 95% 的置信水平, 从图中也可以看出

大部分的真实值都可以被包含在模型预测区间内。这表明，使用 GPR-CPK 模型进行区间预测，可以确保得到的区间在相对狭窄的情况下包含更多的真实值。从 SCI、HSI、FTSE 和 STOXX50 指数的数据可以清楚地看到，单核模型的预测区间相对较宽，这说明它们的预测区间是相对保守的，该类模型为保证尽可能高的覆盖率 PICP 选择扩大区间宽度，因此牺牲了区间预测的精确度。

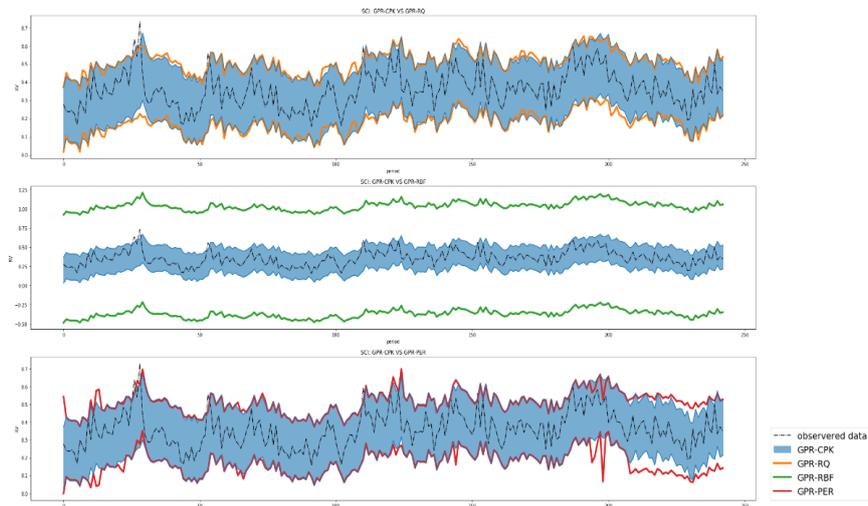


图 4.3 上证指数

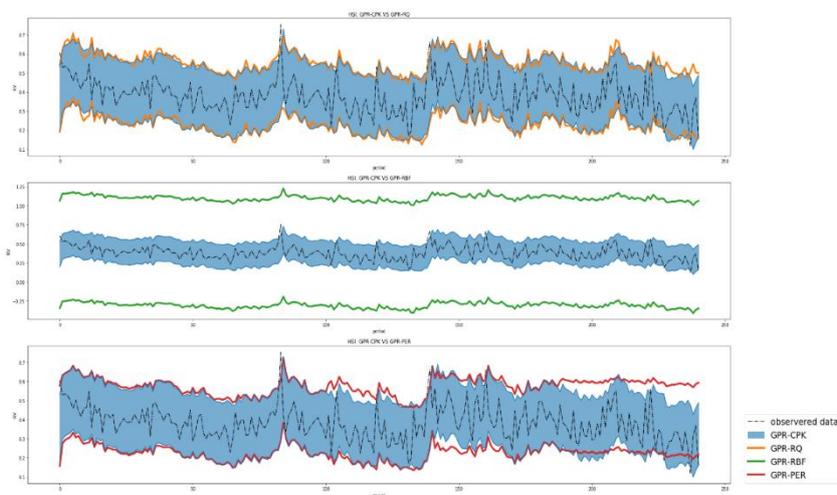


图 4.4 恒生指数

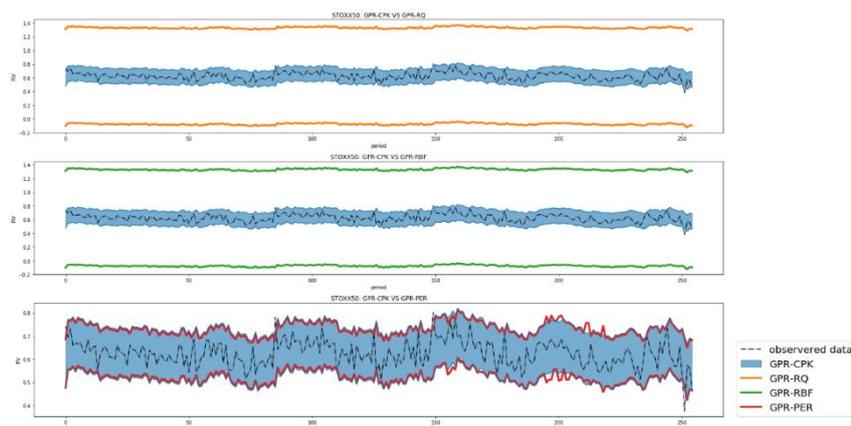


图 4.5 欧洲斯托克 50 指数 (STOXX50)

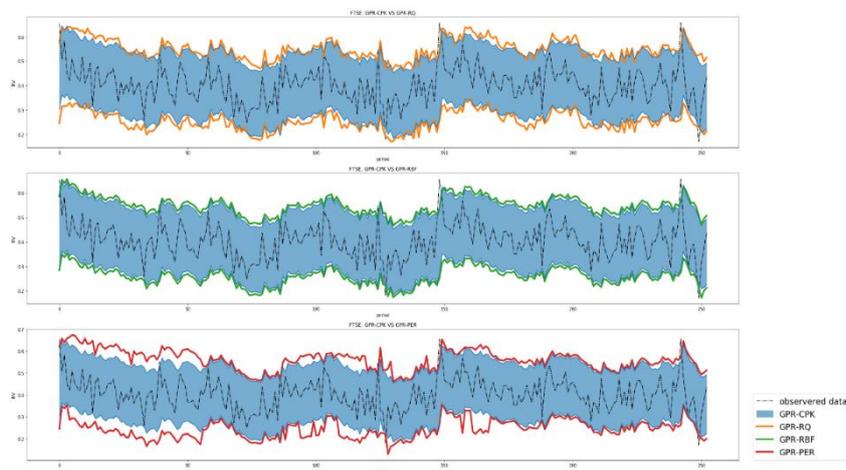


图 4.6 金融时报指数 (FTSE)

## 5. 结论

在现代金融研究中,波动率是描述和反映金融资产风险的重要指标之一,在金融衍生品定价和金融风险度量中都发挥着非常重要的作用。尤其是基于高配数据的已实现波动率更加真实和全面地反映了市场的实际情况,且无需建模和参数估计,因而被广泛的应用在了波动率的度量和预测上。而股票指数反映了各个国家金融市场的总体情况,并可以用于评判投资收益情况优良,推动了衍生品创新等新兴投资方式的诞生,为投资者提供了新的投资参考的标准。因此,准确预测股指的已实现波动率一方面对于金融风险管理、对冲策略评估以及相关资产定价都有着重要的作用,另一方面影响着投资者对于期权收益的判断,是投资过程中关注的重要指标。然而,大量研究表明股指已实现波动率序列存在着周期性的日历效应、异方差和波动率聚集效应。并且,已实现波动率序列的方差并不为常数,而是一个时变随机变量。由此可见,传统的经典金融理论模型并不能很好地刻画波动率序列的实际特征。

针对上述问题,本文通过探究股指已实现波动率序列的特征,并根据股指波动率序列的特征构建了一个新的 GPR-CPK 预测模型框架。具体而言,本文首次根据波动率序列特征设计了适用于波动率预测的组合核函数用于高斯过程回归模型,并使用该模型对十二组国家典型股指波动率数据进行预测。同时,通过比较不同预测模型的预测结果,证明本文所提出的 GPR-CPK 模型不但提高了波动率的预测精度,而且由于高斯过程回归模型的特性其产生的预测为概率性预测,即其不仅可以提供高精度的预测值,还可以对已实现波动率序列进行区

间预测以得到预测区间。从现实意义来看，有助于投资者更加全面的了解认知当前股市整体风险，同时进一步丰富了投资者技术分析的手段，帮助投资者对股市未来走势进行研判，对投资者构建更加合理和优越的投资策略有一定的参考价值。

## 参考文献

- [1] Bollerslev T, Patton A J, Quaedvlieg R. Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 2016, 192(1): 1-18
- [2] Liu L Y, Patton A J, Sheppard K. Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *Journal of Econometrics*, 2015, 187(1): 293-311
- [3] Andersen T G, Bollerslev T. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, 1998: 885-905
- [4] Baek Y, Kim H Y. ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. *Expert Systems with Applications*, 2018, 113: 457-480
- [5] Lin T, Liu D, Zhang L, et al. The information content of realized volatility of sector indices in China's stock market. *International Review of Economics & Finance*, 2019, 64: 625-640
- [6] Müller U A, Dacorogna M M, Davé R D, et al. Volatilities of different time resolutions—analyzing the dynamics of market components. *Journal of Empirical Finance*, 1997, 4(2-3): 213-239
- [7] Bucci A. Cholesky-ANN models for predicting multivariate realized volatility. *Journal of Forecasting*, 2020, 39(6): 865-876
- [8] Bansal R, Kiku D, Shaliastovich I, et al. Volatility, the macroeconomy, and asset prices. *The Journal of Finance*, 2014, 69(6): 2471-2511
- [9] Barndorff-Nielsen O E, Shephard N. Estimating quadratic variation using realized variance. *Journal of Applied econometrics*, 2002, 17(5): 457-477
- [10] Koopman S J, Jungbacker B, Hol E. Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, 2005, 12(3): 445-475
- [11] Engle R. New frontiers for ARCH models. *Journal of Applied Econometrics*, 2002, 17(5): 425-446
- [12] Hansen P R, Huang Z, Shek H H. Realized GARCH: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 2012, 27(6): 877-906
- [13] 王天一,黄卓.Realized GAS-GARCH 及其在 VaR 预测中的应用.管理科学学报,2015,18(05):79-86
- [14] 刘威仪,孙便霞,王明进.基于日度低频价格的波动率预测.管理科学学报,2016,19(01):60-71
- [15] Andersen T G, Bollerslev T, Diebold F X, et al. Modeling and forecasting realized volatility. *Econometrica*, 2003, 71(2): 579-625
- [16] Corsi F. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 2009, 7(2): 174-196
- [17] Müller U A, Dacorogna M M, Davé R D, et al. Fractals and intrinsic time—a challenge to econometricians. Unpublished manuscript, Olsen & Associates, Zürich, 1993: 130
- [18] 魏宇.沪深 300 股指期货的波动率预测模型研究.管理科学学报,2010,13(02):66-76
- [19] 文凤华,刘晓群,唐海如,杨晓光.基于 LHAR-RV-V 模型的中国股市波动性研究.管理科学学

报,2012,15(06):59-67

[20] 陈浪南,杨科. 中国股市高频波动率的特征、预测模型以及预测精度比较. 系统工程理论与实践,2013,33(02):296-307

[21] McAleer M, Medeiros M C. Forecasting realized volatility with linear and nonlinear univariate models. *Journal of Economic Surveys*, 2011, 25(1): 6-18

[22] Sermpinis G, Laws J, Dunis C L. Modelling and trading the realised volatility of the FTSE100 futures with higher order neural networks. *The European Journal of Finance*, 2013, 19(3): 165-179

[23] Rosa R, Maciel L, Gomide F, et al. Evolving hybrid neural fuzzy network for realized volatility forecasting with jumps. 2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr). IEEE, 2014: 481-488

[24] Audrino F, Knaus S D. Lassoing the HAR model: A model selection perspective on realized volatility dynamics. *Econometric Reviews*, 2016, 35(8-10): 1485-1521

[25] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural computation*, 1997, 9(8): 1735-1780

[26] Kim H Y, Won C H. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, 2018, 103: 25-37

[27] 陈卫华. 基于深度学习的上证综指波动率预测效果比较研究. *统计与信息论坛*, 2018, 33(05): 99-106

[28] 陈卫华, 徐国祥. 基于深度学习和股票论坛数据的股市波动率预测精度研究. *管理世界*, 2018, 34(01): 180-181

[29] Kim D, Baek C. Factor-augmented HAR model improves realized volatility forecasting. *Applied Economics Letters*, 2020, 27(12): 1002-1009

[30] Liu F, Pantelous A A, von Mettenheim H J. Forecasting and trading high frequency volatility on large indices. *Quantitative Finance*, 2018, 18(5): 737-748

## Forecasting of Stock Index Realized Volatility Using Gaussian Process Regression with Compositional Kernel

Qiujun Lan, Huiling Zhou

(Hunan University, Changsha/ Hunan, 410000)

**Abstract:** Advances in realized volatility (RV) forecasting immensely benefit financial market regulators and investors. Due to periodicity, small irregularities and time-varying noises, RV forecasting has always been a challenging problem. In this paper, a novel Gaussian process regression model with compositional kernel (GPR-CPK) is proposed that can efficiently improve model accuracy and provide interval estimation. In the proposed model, the kernel is defined as a compositional representation based on sums and products of four base kernels for discovering complex patterns in RV series. The compositional kernel integrates the advantages of typical single kernels to capture the periodic and irregular fluctuations in the RV series. In this paper, the hyperparameters of kernel is optimized using the gradient descent method to maximize the marginal likelihood. Thus, optimized hyperparameters are used for computing forecasting distribution of target RV. To verify the validity of the model, we test it using twelve worldwide stock indices and compare its forecasting results with those from existing popular AI-based models and traditional econometric models. The results show that the proposed model generally outperforms representative benchmark models using both point and interval evaluation measures.

**Keywords:** Compositional kernel, Gaussian process regression, Interval forecasting, Realized volatility, Stock index