10-2022

On the optimal forecast with the fractional Brownian motion

Xiaohu WANG

Chen ZHANG
Singapore Management University, chenzhang@smu.edu.sg

Jun YU
Singapore Management University, yujun@smu.edu.sg

SMU ECONOMICS &
STATISTICS

On the Optimal Forecast with the Fractional Brownian Motion

Xiaohu Wang, Chen Zhang, Jun Yu

Oct 2022

Paper No. 12-2022

On the Optimal Forecast with the Fractional Brownian Motion*

Xiaohu Wang
Fudan University
Jun Yu, Chen Zhang
Singapore Management University

October 28, 2022

Abstract

This paper examines the performance of alternative forecasting formulae with the fractional Brownian motion based on a discrete and finite sample. One formula gives the optimal forecast when a continuous record over the infinite past is available. Another formula gives the optimal forecast when a continuous record over the finite past is available. Alternative discretization schemes are proposed to approximate these formulae. These alternative discretization schemes are then compared with the conditional expectation of the target variable on the vector of the discrete and finite sample. It is shown that the conditional expectation delivers more accurate forecasts than the discretization-based formulae using both simulated data and daily realized volatility (RV) data. Empirical results based on daily RV indicate that the conditional expectation enhances the already-widely known great performance of fBm in forecasting future RV.

JEL classification: C12, C22, G01
Keywords: Fractional Gaussian noise; Conditional expectation; Anti-persistence; Continuous record; Discrete record; Optimal forecast

*We wish to thank Shuping Shi, Weilin Xiao and Masaaki Fukasawa for helpful comments. Xiaohu Wang, School of Economics, Fudan University, Shanghai, China, and Shanghai Institute of International Finance and Economics, Shanghai, China, Email: wang_xh@fudan.edu.cn. Jun Yu, School of Economics and Lee Kong Chian School of Business, Singapore Management University, 90 Stamford Road, Singapore 178903. Email: yujun@smu.edu.sg. Chen Zhang, School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903. Email: chenzhang@smu.edu.sg. Wang acknowledges the support by Shanghai Pujiang Program under No.22PJC022. Yu and Zhang would like to acknowledge that this research/project is supported by the Ministry of Education, Singapore, under its Academic Research Fund (AcRF) Tier 2 (Award Number MOE-T2EP402A20-0002).

1 Introduction

Starting with the seminar paper by Gatheral et al. (2018), studies in the volatility literature have shown that continuous-time models based on the rough fractional Brownian motion (fBm) provide accurate forecasts relative to many popular discrete-time models. For example, Gatheral et al. (2018) find that the rough fBm model with the Hurst parameter (denoted by H) being 0.14 for log realized volatility (RV) provides more accurate forecasts of log RV and RV than the HAR model of Corsi (2009). Wang et al. (2021) find that rough fBm and rough fractional Ornstein-Uhlenbeck (fOU) processes for log RV yield more accurate forecasts of log RV and RV than the HAR model and the ARFIMA model. In Wang et al. (2021), the Hurst parameter H is estimated recursively by the method-of-moments in the fBm and fOU processes from daily log RV data and found to be always significantly less than 0.5. Fukasawa et al. (2021) and Bolko et al. (2022) continue to find evidence of $H < 0.5$ when the measurement error in RV is taken into account. Bennedsen et al. (2022) find evidence of $H < 0.5$ when the composite likelihood method is used to estimate the fOU process.

Since fBm with $H < 0.5$ produces a sample path rougher than the standard Brownian motion, when it is used to model volatility, it is called the rough fractional stochastic volatility (RFSV) model in the literature. The RFSV literature has received a great deal of attention in mathematical finance, financial engineering, and financial econometrics. The 2021 risk award was presented to Jim Gatheral and Mathieu Rosenbaum for introducing such a model. There is a website dedicated to the literature that collects more than 200 papers on this subject from different perspectives.[1] In addition to the applications in forecasting volatility as mentioned earlier, RFSV models have been used to price options (Livieri et al., 2018; Bayer et al., 2016; Garnier and Sølna, 2017) and variance swaps (Bayer et al., 2016), in portfolios choice (Fouque and Hu, 2018), and in dynamic hedging (Euch and Rosenbaum, 2018).

This paper examines the performance of alternative forecasting formulae with fBm based on a discrete and finite sample. We show that the existing forecasting formula, albeit being optimal when a continuous record is available, when it is discretized, does not generate the optimal forecast anymore. The optimal forecast is the conditional expectation of the target variable on the vector of the discrete

[1] https://sites.google.com/site/roughvol/home/risks-1.

and finite sample. We quantity the loss in efficiency relative to the optimal forecast, in terms of the root mean square error (RMSE), when a discrete and finite sample is available. Findings in our paper suggest that rough fBm can produce even more accurate forecasts of log RV than what has been reported in the literature, and hence, enhance the already-widely known great performance of fBm in forecasting volatility.

The rest of the paper is organized as follows. Section 2 reviews the model and discusses the alternative forecasting formulae. Section 3 reports the forecast results based on simulated data. Section 4 discusses how to use the maximum likelihood (ML) method to estimate parameters in fBm and reports the forecast results based on real RV data. Section 5 concludes. The appendix gives proof of the Proposition in the paper.

2 The Model and Forecasting Formulae

The fBm, denoted by $\sigma B^H(t)$, is a continuous-time Gaussian process with zero mean and the autocovariance function of

$$Cov\left(B^H(t), B^H(s)\right) = \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right), \quad \forall t, s \in (-\infty, +\infty), \qquad (1)$$

where $\sigma > 0$ is a constant scale parameter and $H \in (0,1)$ is called the Hurst parameter of the fBm. Another definition of $B^H(t)$, given by Mandelbrot and Van Ness (1968), is

$$B^H(t) = \frac{1}{\Gamma(H+0.5)}\left\{\int_{-\infty}^{0}\left[(t-s)^{H-0.5} - (-s)^{H-0.5}\right]dW(s) + \int_0^t (t-s)^{H-0.5}\,dW(s)\right\},$$

where $\Gamma(\cdot)$ denotes the Gamma function and $W(t)$ is a standard Brownian motion. It is clear to see that, when the Hurst parameter $H = 0.5$, $B^H(r) = W(r)$ becomes a standard Brownian motion.

Let y_t denote the increment of the fBm, that is, $y_t = \sigma\left(B^H(t) - B^H(t-1)\right)$. The sequence $\{y_t\}$ is known in the literature as the fractional Gaussian noise (fGn), which is a normally distributed stationary process with the following autocovariances:

$$\begin{aligned} Cov\left(y_t, y_{t+k}\right) &= \frac{\sigma^2}{2}\left((k+1)^{2H} + (k-1)^{2H} - 2k^{2H}\right), \text{ for any } k \geq 0 \qquad (2)\\ &\sim \sigma^2 H(2H-1)k^{2H-2} \text{ for large } k. \end{aligned}$$

where $\sim$ denotes asymptotic equivalence. Equation (2) shows that when $H \in (0.5, 1)$, $\{y_t\}$ has positive serial dependence, and the autocovariances of $\{y_t\}$ are not absolutely summable. As a result, $\{y_t\}$ is called a long-memory process when $H > 0.5$. In contrast, if $H \in (0, 0.5)$, it can be proved that, $\{y_t\}$ has negative autocovariances and

$$\sum_{k=-\infty}^{+\infty} Cov\left(y_t, y_{t+k}\right) = 0.$$

In this case, $\{y_t\}$ is called an antipersistent process. The corresponding fBm has sample paths that are rougher than those of the standard Brownian motion. Considering that the RFSV literature has found ample evidence supporting roughness (i.e. $H < 0.5$), our paper focuses on the study for the case of $H < 0.5$.

2.1 Forecasts based on the infinite-past-formula

Gatheral et al. (2018, Equation 5.1) propose to use $X(t) = \sigma B^H(t)$ with $H = 0.14$ to model log RV. If the continuous record of $X(t)$ over the period of $(-\infty, T]$ is available, Nuzman and Poor (2000, Equation (34)) develop the following formula to generate the optimal k-period-ahead forecast:

$$E\left\{X\left(T+k\right)|X(t), t \in (-\infty, T]\right\} = \int_{-\infty}^{T} \frac{\cos(H\pi)k^{H+0.5}}{\pi(T-s+k)(T-s)^{H+0.5}} X(s)ds. \quad (3)$$

Being the conditional expectation, this formula generates the optimal forecast and is a weighted average of the entire history of $X(t)$ over $(-\infty, T]$. To simplify notations, let

$$w_1(s) = \frac{\cos(H\pi)k^{H+0.5}}{\pi(T-s+k)(T-s)^{H+0.5}} = \frac{\cos(H\pi)k^{H+0.5}}{\pi} \widetilde{w}_1(s),$$

where

$$\widetilde{w}_1(s) = \frac{1}{(T-s+k)(T-s)^{H+0.5}}.$$

We rewrite the forecasting formula in (3) as

$$\int_{-\infty}^{T} w_1(s)X(s)ds = \frac{\cos(H\pi)k^{H+0.5}}{\pi} \int_{-\infty}^{T} \widetilde{w}_1(s)X(s)ds. \quad (4)$$

The following proposition shows that for any $T, k > 0$, the weights in the forecasting formula integrated to unity.

Proposition 2.1 *For any $T, k > 0$, it has*

$$\int_{-\infty}^{T} w_1(s)ds = \int_{-\infty}^{T} \frac{\cos(H\pi)k^{H+0.5}}{\pi(T-s+k)(T-s)^{H+0.5}} ds = 1.$$

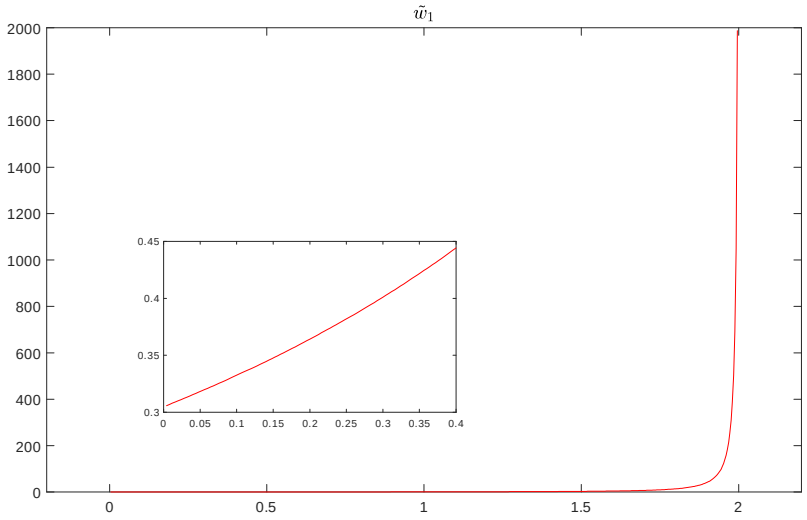

Figure 1: Plot of $\widetilde{w}_1(s)$ as a function of s when $H = 0.2$, $k = 1/50$, $T = 2$. The picture-in-picture plots $\widetilde{w}_1(s)$ when s takes values close to zero.

Figure 1 plots $\widetilde{w}_1(s)$ as a function of $s \in [0, T]$ when $H = 0.2$, $k = 1/50$, and $T = 2$. To help understand the near-zero behavior of $\widetilde{w}_1(s)$, we plot $\widetilde{w}_1(s)$ when s takes values close to zero in the same Figure. The figure shows that $\widetilde{w}_1(s)$ is a monotonically increasing function that goes to ∞ as $s \to T$. This property suggests that nearer-distant history is more important than further-distant history in forecasting X_{T+k}.

However, in practice, it is infeasible to directly implement the forecasting formula in (3) for two reasons. First, we live in a digital world. Instead of having a continuous record, $X(t)$, the log RV process, can only be observed at discrete time points. Second, we cannot access to information in the infinite past. Therefore, in reality, people can only generate forecasts based on a discrete record of $X(t)$ collected over the finite time-interval $[0, T]$. Without loss of generality, we use $\{X_0, X_1, ..., X_T\}$ to denote the data available.

Gatheral et al. (2018) suggest making adjustments to the formula in (4) in order to generate forecasts based on the discrete record $\{X_0, X_1, ..., X_T\}$. The adjusted

formula is

$$IPLA = \frac{\sum_{s=0}^{T-1} w_1(s)X_s}{\sum_{s=0}^{T-1} w_1(s)} = \frac{\sum_{s=0}^{T-1} \widetilde{w}_1(s)X_s}{\sum_{s=0}^{T-1} \widetilde{w}_1(s)}. \tag{5}$$

From the formula in (4) to the formula in (5), three adjustments are made. First, the integral (4) is truncated from below at zero. Second, the integral is replaced with the left Riemann sum. Third, the weight function is normalized by $\sum_{s=0}^{T-1} \widetilde{w}_1(s)$ so that the new weights sum to unity, mimicking the feature in Proposition 2.1. The last adjustment is to take into account of the truncation error introduced by the first adjustment. In this paper, the formula given by (5) is called the infinite-past-left-adjustment (IPLA) method.

There are other types of adjustments that can be made to the forecasting formula in (4). Keeping other adjustments in IPLA intact, if we only change the integral with the right Riemann sum, we can have[2]

$$IPRA = \frac{\sum_{s=1}^{T-1} \widetilde{w}_1(s)X_s + \widetilde{w}_1(T-1)X_T}{\widetilde{w}_1(T-1) + \sum_{s=1}^{T-1} \widetilde{w}_1(s)}. \tag{6}$$

We name the forecasting formula in (6) the infinite-past-right-adjustment (IPRA) method.

The third adjusted method is to replace the integral with the trapezoidal sum, which gives the forecasting formula

$$\begin{aligned} IPTA &= \frac{\frac{1}{2}\sum_{s=0}^{T-2}\left(w_1(s)X_s + w_1(s+1)X_{s+1}\right) + w_1(T-1)(\frac{1}{2}X_{T-1} + \frac{1}{2}X_T)}{\frac{1}{2}w_1(0) + \sum_{s=1}^{T-2} w_1(s) + \frac{3}{2}w_1(T-1)} \\ &= \frac{\frac{1}{2}\widetilde{w}_1(0)X_0 + \sum_{s=1}^{T-1} \widetilde{w}_1(s)X_s + \frac{1}{2}\widetilde{w}_1(T-1)X_T}{\frac{1}{2}\widetilde{w}_1(0) + \sum_{s=1}^{T-2} \widetilde{w}_1(s) + \frac{3}{2}\widetilde{w}_1(T-1)}. \end{aligned} \tag{7}$$

This method is called the infinite-past-trapezoidal-adjustment (IPTA) method.

Another adjusted formula can be obtained by replacing the integral with the middle Riemann sum applied to the weight only, that is,

$$IPMA = \frac{\sum_{s=1}^{T} \widetilde{w}_1\left(\frac{1}{2}+s-1\right)X_s}{\sum_{s=1}^{T} \widetilde{w}_1\left(\frac{1}{2}+s-1\right)}. \tag{8}$$

We name this adjustment method the infinite-past-middle-adjustment (IPMA) method.

It is important to remark that, although Formula (4) yields the optimal forecast given a continuous record over $(-\infty, T]$ being available, the alternative discretizations given in (5)-(8) may not lead to the optimal forecast when only finite number

[2]Note that the weight function is not well-defined at the point $s = T$. Therefore, to approximate the integral $\int_{T-1}^{T} \widetilde{w}_1(s)X(s)\,ds$, we use $X(T)$ with weight $\widetilde{w}_1(T-1)$.

discrete-time observations $\{X_0, X_1, ..., X_T\}$ are available. In fact, they are not designed to minimize any objective function of the forecast errors.

2.2 Forecasts based on the finite-past-formula

If $X(t)$ has a continuous record over the finite time-span of $[0, T]$ available, Theorem 4.4 in Nuzman and Poor (2000) gives another formula for the k-period-ahead optimal forecast at period T:

$$E\left\{X\left(T+k\right)|X(t), t \in [0,T]\right\} = \int_0^T m_T\left(\frac{k}{T}, T-s\right) X(s)ds, \tag{9}$$

where

$$m_T(c,t) = T^{-1}\left(\frac{t}{T}\right)^{-H-0.5}\left(1-\frac{t}{T}\right)^{-H-0.5} \times \left[(0.5-H)B_{c/(c+1)}\left(H+0.5, 1-2H\right) + \frac{c^{H+0.5}(1+c)^{H-0.5}(1-t/T)}{c+t/T}\right],$$

and $B_{c/(c+1)}(\cdot,\cdot)$ denotes the incomplete beta function that takes the form of

$$B_{c/(c+1)}\left(H+0.5, 1-2H\right) = \int_0^{c/(c+1)} z^{H-0.5}\left(1-z\right)^{-2H} dz.$$

It can be verified that, as noted on page 443 in Nuzman and Poor (2000), the weights in the forecasting formula (9) integrated to unity:

$$\int_0^T m_T(c,t)dt = 1 \quad \text{for any} \quad T, c > 0.$$

This feature is the same as that in Proposition 2.1. To simply notations, we let

$$w_2(s) = m_T\left(\frac{k}{T}, T-s\right),$$

and rewrite the forecasting formula in (9) as

$$E\left\{X\left(T+k\right)|X(t), t \in [0,T]\right\} = \int_0^T w_2(s)X(s)ds. \tag{10}$$

Being the conditional expectation, this formula also generates the optimal forecast and is a weighted average of the entire history of $X(t)$ over $[0, T]$.

From the definition of $w_2(s)$, it is easy to get that $\lim_{s\to 0} w_2(s) = \infty$ and $\lim_{s\to T} w_2(s) = \infty$ when $H \in (0, 0.5)$. As a result, $w_2(s)$ is not a monotonic function.

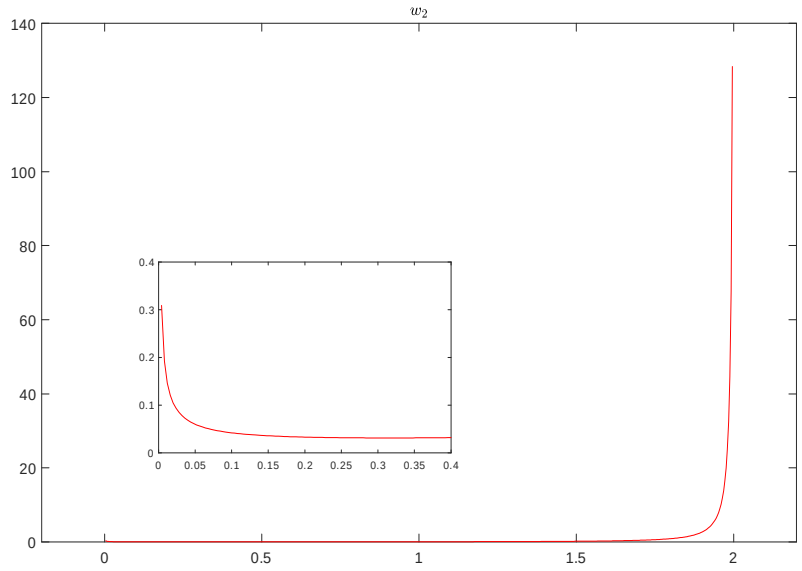

Figure 2: Plots of $w_2(s)$ as a function of s when $H = 0.2, k = 1/50, T = 2$. The picture-in-picture plots $w_2(s)$ when s takes values close to zero. The picture-in-picture plots $w_2(s)$ when s takes values close to zero.

It is a little surprising to see that $w_2(s)$ approaches infinity as $s \to 0$. This feature is sharply distinct from the weights function $\widetilde{w}_1(s)$ used in the forecast formula (4) that converges to a small constant number when $s \to 0$, as discussed in Subsection 2.1. This difference is due to the fact that the new forecasting formula in (9) is based on a continuous record over $[0, T]$ rather than $(-\infty, T]$. When the record over $[0, T]$ being available, the forecasting formula (9) places the optimal weight $X(s)$ for $s \in (0, T)$. Hence, it can generate more accurate forecasts than using the forecasting formula (4) with a truncation:

$$\int_0^T w_1(s)X(s)ds, \quad \text{or} \quad \frac{\int_0^T w_1(s)X(s)ds}{\int_0^T w_1(s)ds} = \frac{\int_0^T \widetilde{w}_1(s)X(s)ds}{\int_0^T \widetilde{w}_1(s)ds}. \tag{11}$$

Figure 2 plots $w_2(s)$ as a function of s when $H = 0.2, k = 1/50$, and $T = 2$. It shows that when T is finite, $w_2(s)$ is U-shaped, placing higher weights when s approaches T or 0.

We plot normalized $w_2(s)$ and $\widetilde{w}_1(s)$, i.e. $w_2(s)/\int_0^T w_2(\tau)d\tau$ and $\widetilde{w}_1(s)/\int_0^T \widetilde{w}_1(\tau)d\tau$, as a function of s in Figure 3, for $k = 1/250, 1/50$ and $T = 2, 4$.[3] The normalized

[3]Although $\int_0^T w_2(\tau)d\tau = 1$, normalization of $w_2(\tau)$ is needed as the numerical approximation

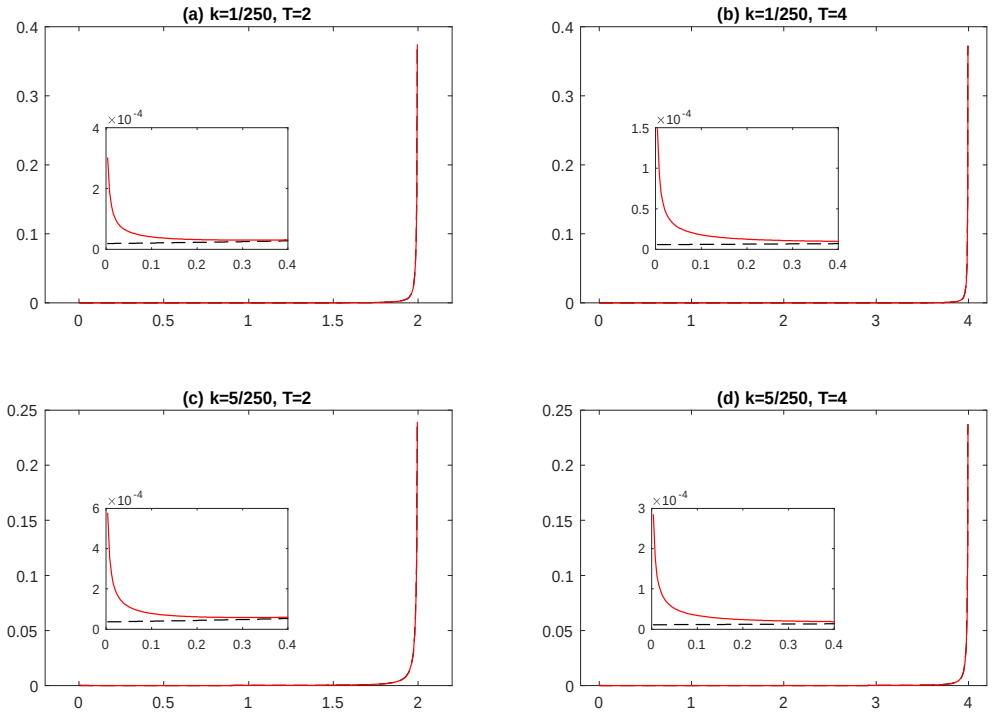

Figure 3: Plots of normalized $\widetilde{w}_1(s)$ and $w_2(s)$ as a function of s when $H = 0.2$, $k = 1/250, 1/50$ and $T = 2, 4$. The solid line is for normalized $w_2(s)$ while the dash line is for normalized $\widetilde{w}_1(s)$. The picture-in-picture plots each weight function when s takes values close to zero.

weight function ensures that the weights calculated at discrete points sum to unity. These plots show that, compared to the normalized $\widetilde{w}_1(s)$, the normalized $w_2(s)$ places higher weights on observations near period 0.

In practice, to apply the formula (10) to generate forecasts based on $\{X_0, ..., X_T\}$, some adjustments have to be made. Again, there are various adjustment methods. The first method is to replace the integral in (10) by the left Riemann sum and then normalize the corresponding weights of every data observation. This method is called the finite-past-left-adjustment (FPLA) method, which gives the following forecasting formula:[4]

to $\int_0^T w_2(\tau)d\tau$ can be different from one.

[4]Note that $w_2(0) = \infty$. Therefore, the left Riemann sum is not applicable to approximate the integral $\int_0^1 w_2(s)X(s)\,ds$. To approximate this integral, we use $X(0)$ with the weight $w_2(1)$.

$$FPLA = \frac{\sum_{s=2}^{T} w_2(s-1)X_{s-1} + w_2(1)X_0}{w_2(1) + \sum_{s=2}^{T} w_2(s-1)}. \tag{12}$$

An alternative method is to replace the integral with the right Riemann sum, which is named as the finite-past-right-adjustment (FPRA) method. The resulting forecasting formula is[5]

$$FPRA = \frac{\sum_{s=1}^{T-1} w_2(s)X_s + w_2(T-1)X_T}{w_2(T-1) + \sum_{s=1}^{T-1} w_2(s)}. \tag{13}$$

The third method is called the finite-past-trapezoidal-adjustment (FPTA) method, which replaces the integral with the trapezoidal sum and gives the forecasting formula of

$$FPTA = \frac{\frac{1}{2}w_2(1)(X_0 + X_1) + \frac{1}{2}\sum_{s=1}^{T-2}\left(w_2(s)X_s + w_2(s+1)X_{s+1}\right) + \frac{1}{2}w_2(T-1)(X_{T-1} + X_T)}{\sum_{s=2}^{T-2} w_2(s) + \frac{3}{2}\left(w_2(1) + w_2(T-1)\right)}. \tag{14}$$

The fourth adjustment method is to replace the integral with the middle Riemann sum applied to the weight only. This method is called the finite-past-middle-adjustment (FPMA) method and yields the following forecasting formula:

$$FPMA = \frac{\sum_{s=1}^{T} w_2\left(\frac{1}{2} + s - 1\right)X_s}{\sum_{s=1}^{T} w_2\left(\frac{1}{2} + s - 1\right)}. \tag{15}$$

As before, although Formula (9) yields the optimal forecast given a continuous record over $[0, T]$ being available, the alternative discretizations given in (12)-(15) may not lead to the optimal forecast when only finite number discrete-time observations $\{X_0, X_{1,}..., X_T\}$ are available. None of these discretization-based formulae is designed to minimize any objective function of the forecast errors.

Table 1 reports the weights applied to all observations in the discrete sample under alternative forecasting formulae. A distinctive feature is that both IPLA and FPLA give zero weight to the most recent observation, X_T. This feature has important implications for the finite sample performance of the two methods, as will be remarked on later.

[5]Note that $w_2(T) = \infty$. Therefore, we approximate the integral $\int_{T-1}^{T} w_2(s)X(s)\,ds$ by using $X(T)$ with the weight of $w_2(T-1)$.

Table 1: Weights on the sample under different forecasting formulae before the normalization

Methods	X_0	X_1	X_2	$\cdots$	X_{T-2}	X_{T-1}	X_T
IPLA	$\widetilde{w}_1(0)$	$\widetilde{w}_1(1)$	$\widetilde{w}_1(2)$	$\cdots$	$\widetilde{w}_1(T-2)$	$\widetilde{w}_1(T-1)$	0
IPRA	0	$\widetilde{w}_1(1)$	$\widetilde{w}_1(2)$	$\cdots$	$\widetilde{w}_1(T-2)$	$\widetilde{w}_1(T-1)$	$\widetilde{w}_1(T-1)$
IPTA	$\frac{1}{2}\widetilde{w}_1(0)$	$\widetilde{w}_1(1)$	$\widetilde{w}_1(2)$	$\cdots$	$\widetilde{w}_1(T-2)$	$\widetilde{w}_1(T-1)$	$\frac{1}{2}\widetilde{w}_1(T-1)$
IPMA	0	$\widetilde{w}_1(0.5)$	$\widetilde{w}_1(1.5)$	$\cdots$	$\widetilde{w}_1(T-2.5)$	$\widetilde{w}_1(T-1.5)$	$\widetilde{w}_1(T-0.5)$
FPLA	$w_2(1)$	$w_2(1)$	$w_2(2)$	$\cdots$	$w_2(T-2)$	$w_2(T-1)$	0
FPRA	0	$w_2(1)$	$w_2(2)$	$\cdots$	$w_2(T-2)$	$w_2(T-1)$	$w_2(T-1)$
FPTA	$\frac{1}{2}w_2(1)$	$w_2(1)$	$w_2(2)$	$\cdots$	$w_2(T-2)$	$w_2(T-1)$	$\frac{1}{2}w_2(T-1)$
FPMA	0	$w_2(0.5)$	$w_2(1.5)$	$\cdots$	$w_2(T-2.5)$	$w_2(T-1.5)$	$w_2(T-0.5)$

2.3 Optimal forecast based on a discrete and finite sample

This subsection introduces the optimal forecasting formula based on finite discrete-time observations $\{X_0, X_\Delta ..., X_{T\Delta}\}$, where Δ is the sampling interval and $T+1$ is the number of historical observations. We summarize the data into a column vector $X = (X_0, X_\Delta ..., X_{T\Delta})'$. Our goal is to generate the optimal forecast for $X_{(T+k)\Delta}$, that is, k-period-ahead forecast. Let the covariance matrix of X be $\Sigma_{0:T\Delta}$. Elements in $\Sigma_{0:T\Delta}$ can be easily obtained from the formula in (1). Let $\gamma_{k\Delta}^{(0:T\Delta)} = (Cov(X_{(T+k)\Delta}, X_0), ..., Cov(X_{(T+k)\Delta}, X_{T\Delta}))'$ denote the vector of covariances between $X_{(T+k)\Delta}$ and the elements in X. Again, elements in $\gamma_{k\Delta}^{(0:T\Delta)}$ can readily be obtained from (1).

Since X_t is a Gaussian process, $(X_0, X_\Delta, ..., X_{T\Delta}, X_{(T+k)\Delta})'$ follows a multivariate normal distribution. The conditional expectation, which is known as the optimal forecast, has a closed-form expression. In particular, the conditional mean is given by

$$E(X_{(T+k)\Delta}|X) = \left(\gamma_{k\Delta}^{(0:T\Delta)}\right)' \Sigma_{0:T\Delta}^{-1} X. \tag{16}$$

This is the optimal forecast when a discrete record over the finite past $[0, T\Delta]$, $(X_0, X_\Delta ..., X_{T\Delta})'$, is available because it minimizes the RMSE of the forecast.

Figure 4 plots the weight function $\left(\gamma_{k\Delta}^{(0:T)}\right)' \Sigma_{0:T\Delta}^{-1} := w^*(s)$ as a function of $s \in \{0, \Delta, 2\Delta, ..., T\Delta\}$ when $H = 0.2$, $k\Delta = 1/50$, $T\Delta = 2$. It shows that similar to $w_2(s)$, the weight function in the optimal forecast $\left(\gamma_{k\Delta}^{(0:T\Delta)}\right)' \Sigma_{0:T\Delta}^{-1} X$ is also U-shaped, indicating that the initial observation X_0 is relatively important in forecasting $X_{(T+k)\Delta}$.

Figure 5 plots the log of the ratio of $w^*(s)$ to the normalized $\widetilde{w}_1(s)$ (the weights used in IPLA) and the log of the ratio of $w^*(s)$ to the normalized $w_2(s)$ (the weights

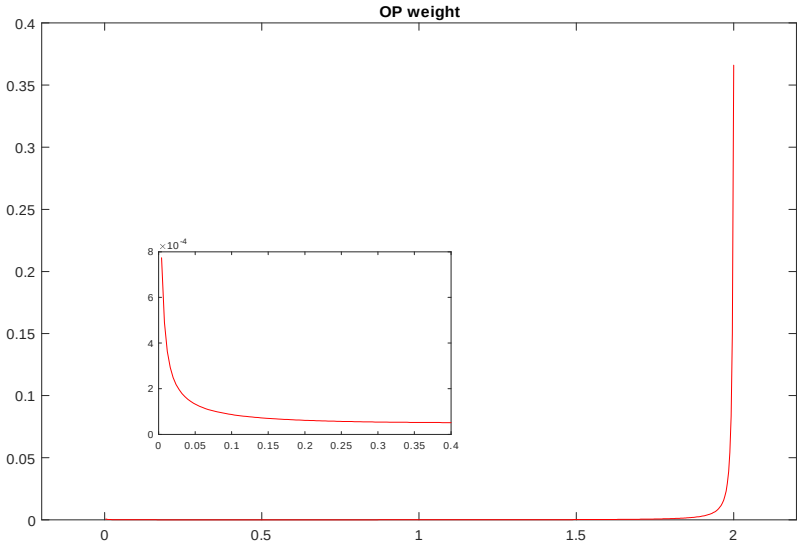

Figure 4: Optimal weight as a function of $s \in \{0, \Delta, 2\Delta, ..., T\Delta\}$ when $H = 0.2$, $k\Delta = 1/50$, $T\Delta = 2$.

used in FPLA), both as a function of $s \in \{0, \Delta, 2\Delta, ..., T\Delta\}$ when $H = 0.2$, $k\Delta = 1/50$, $T\Delta = 2$. The figure shows that, compared to the optimal weight $w^*(s)$, the IPLA method and the FPLA method give too less weight when s is near 0 and too greater weight when s is near $T\Delta$. Comparing the IPLA method and the FPLA method, the weight implied by the IPLA method is further away from the optimal weight $w^*(s)$ than that implied by the FPLA method, suggesting that it is difficult for the IPLA method to outperform the FPLA method. This comparison indicates that although the IPLA method has been the method used in the literature (e.g., Gatheral et al., 2018; Wang et al., 2019), it is sub-optimal when a discrete record over the finite period $[0, T\Delta]$ is available. Improvements over the IPLA method are possible.

3 Comparison based on Simulated Data

We now design Monte Carlo experiments to examine the performance of alternative forecasting formulae when a discrete and finite sample is simulated from the fBm. In all experiments, we set $\sigma = 1$, $H \in \{0.05, 0.1, 0.15, 0.2, 0.25, 0.4\}$, $\Delta = 1$, and assume

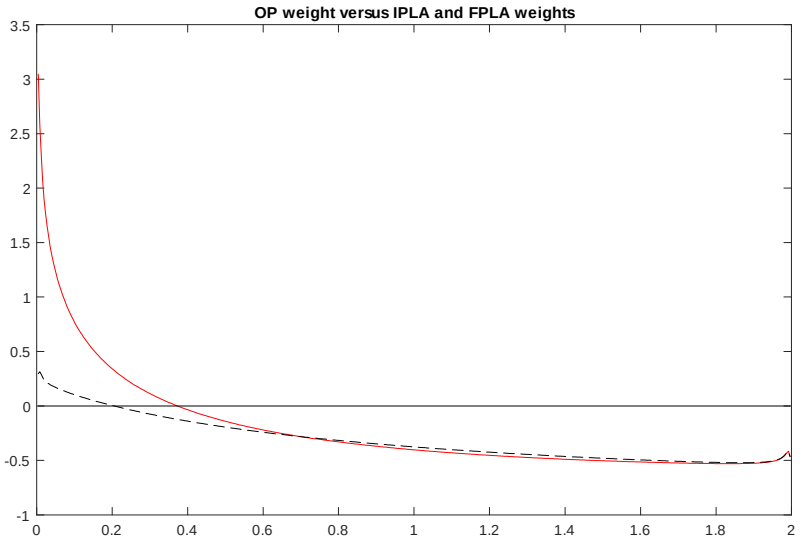

Figure 5: The log of the ratio of $w^*(s)$ to the normalized $\widetilde{w}_1(s)$ in IPLA and the log of the ratio of $w^*(s)$ to the normalized $w_2(s)$ in FPLA, both as functions of s for $H = 0.2$, $k\Delta = 1/50$, and $T\Delta = 2$.

both σ and H are known.[6] As a result, no estimation is needed. The number of replications is 100,000 and used to calculate the RMSE. For each replication, we simulate 511 observations. The first 501 observations (hence $T = 501$) are used to generate forecasts. Alternative forecasting formulae are used to generate k-step-ahead-forecast with $k = 1, ..., 10$. The RMSEs from these forecasting formulae are reported in Tables 2-7. The last row in each table shows the theoretical RMSE of the optimal forecast, which is obtained as

$$RMSE^* = \sigma^2 + w^{*\prime}\Sigma_{0:T}w^* - 2w^{*\prime}\gamma_k^{(0:T)} = \sigma^2 - w^{*\prime}\Sigma_{0:T}w^*, \tag{17}$$

where $\sigma^2 = Var(X_{T+k})$ and $w^* = \Sigma_{0:T}^{-1}\gamma_k^{(0:T)}$.

Several conclusions can be drawn from these tables. First, the optimal forecast (OP) always yields the lowest RMSE, which is always close to the theoretical RMSE.

Second, the smaller the k is, the bigger the improvement of the performance of OP over the other forecasting formulae. For example, when H is 0.25, OP improves IPLA by 16.19%, 10.94%, 8.50%, 6.72%, 5.98%, 2.27%, 4.52%, 4.18%, 4.07% and 3.69% for $k = 1, ..., 10$, respectively.

[6] Due to the self-similarity property, by setting $\Delta = 1/250$ will not change the empirical results reported below.

Third, the larger the H is, the bigger the improvement of OP's performance over the other forecasting formulae. For example, when $k = 1$, OP improves IPLA by 3.36%, 6.00%, 9.68%, 14.70%, 20.04%, 41.73% for $H = 0.05$, 0.1, 0.15, 0.2, 0.25, 0.4, respectively.

Fourth, the relative performance of the four discretization schemes depends on the value of H. When H is small, say $H = 0.05$, the right adjustment yields the smallest RMSE. As H increases, the middle adjustment performs the best among the discretization-based methods.

Fifth, the finite-past formulae always yield smaller RMSEs compared to their infinite-past counterparts. However, the magnitude of the improvement is small. For example, FPLA improves IPLA by 0.038%, 0.012%, 0.011%, 0%, 0.010% and 0.043% for $H = 0.05$, 0.1, 0.15, 0.2, 0.25, 0.4 when $k = 1$.

Last but not least, IPLA and its infinite-past counterpart, FPLA, always perform the worst, although IPLA has been used in practice.

Table 2: RMSE of the alternative forecasting formulae for k-day-ahead-forecast when $H = 0.05$.

k	1	2	3	4	5	6	7	8	9	10
IPLA	0.8132	0.8202	0.8280	0.8352	0.8406	0.8455	0.8505	0.8551	0.8583	0.8619
IPRA	0.7929	0.8088	0.8199	0.8279	0.8353	0.8410	0.8464	0.8514	0.8553	0.8592
IPTA	0.7965	0.8101	0.8206	0.8288	0.8356	0.8412	0.8466	0.8516	0.8553	0.8592
IPMA	0.8053	0.8169	0.8256	0.8317	0.8387	0.8439	0.8487	0.8533	0.8572	0.8608
FPLA	0.8127	0.8199	0.8276	0.8349	0.8402	0.8452	0.8501	0.8547	0.8578	0.8615
FPRA	0.7926	0.8085	0.8196	0.8276	0.8349	0.8406	0.8460	0.8510	0.8548	0.8587
FPTA	0.7962	0.8098	0.8203	0.8285	0.8352	0.8409	0.8463	0.8512	0.8549	0.8587
FPMA	0.8048	0.8164	0.8251	0.8312	0.8382	0.8434	0.8482	0.8528	0.8566	0.8602
OP	0.7868	0.8043	0.8160	0.8246	0.8324	0.8384	0.8439	0.8490	0.8529	0.8568
Theoretical	0.7863	0.8044	0.8162	0.8250	0.8323	0.8384	0.8437	0.8485	0.8527	0.8566

4 Comparison based on Real Data

4.1 ML for fBm

Since $X \sim N(0, \Sigma_T)$ with elements of Σ_T being found from (1), the log likelihood function of fBm is,

$$\ln L(H, \sigma^2) = -\frac{\ln \sigma^2}{2} - \frac{1}{2\,(T+1)} \ln |\Sigma_T| - \frac{1}{2\,(T+1)\,\sigma^2} X' \Sigma_T^{-1} X. \qquad (18)$$

When we numerically maximize $\ln L(H, \sigma^2)$, $|\Sigma_T|$ is obtained as the product of the eigenvalues of Σ_T.

Table 3: RMSE of the alternative forecasting formulae for k-day-ahead-forecast when $H = 0.1$.

k	1	2	3	4	5	6	7	8	9	10
IPLA	0.8845	0.9123	0.9349	0.9550	0.9698	0.9846	0.9971	1.0098	1.0176	1.0278
IPRA	0.8400	0.8837	0.9149	0.9374	0.9558	0.9719	0.9855	0.9984	1.0080	1.0186
IPTA	0.8533	0.8919	0.9203	0.9424	0.9596	0.9754	0.9888	1.0018	1.0107	1.0213
IPMA	0.8392	0.8846	0.9164	0.9379	0.9568	0.9723	0.9856	0.9982	1.0081	1.0185
FPLA	0.8844	0.9122	0.9347	0.9548	0.9696	0.9844	0.9969	1.0095	1.0173	1.0274
FPRA	0.8399	0.8836	0.9147	0.9372	0.9556	0.9716	0.9852	0.9981	1.0077	1.0182
FPTA	0.8533	0.8919	0.9202	0.9422	0.9594	0.9752	0.9886	1.0015	1.0104	1.0209
FPMA	0.8390	0.8844	0.9161	0.9376	0.9565	0.9720	0.9852	0.9978	1.0077	1.0180
OP	0.8344	0.8810	0.9131	0.9358	0.9546	0.9706	0.9842	0.9970	1.0067	1.0172
Theoretical	0.8341	0.8819	0.9126	0.9357	0.9544	0.9702	0.9840	0.9962	1.0071	1.0171

Table 4: RMSE of the alternative forecasting formulae for k-day-ahead-forecast when $H = 0.15$.

k	1	2	3	4	5	6	7	8	9	10
IPLA	0.9562	1.0110	1.0527	1.0895	1.1166	1.1424	1.1637	1.1824	1.2011	1.2184
IPRA	0.8872	0.9648	1.0164	1.0576	1.0898	1.1181	1.1421	1.1631	1.1833	1.2016
IPTA	0.9102	0.9800	1.0284	1.0684	1.0989	1.1264	1.1494	1.1697	1.1894	1.2074
IPMA	0.8728	0.9576	1.0113	1.0523	1.0856	1.1142	1.1390	1.1605	1.1811	1.1995
FPLA	0.9563	1.0110	1.0527	1.0895	1.1166	1.1424	1.1636	1.1824	1.2010	1.2183
FPRA	0.8873	0.9648	1.0163	1.0576	1.0897	1.1180	1.1419	1.1630	1.1831	1.2014
FPTA	0.9103	0.9800	1.0284	1.0684	1.0988	1.1264	1.1493	1.1696	1.1892	1.2072
FPMA	0.8728	0.9575	1.0112	1.0522	1.0855	1.1141	1.1388	1.1603	1.1808	1.1991
OP	0.8718	0.9567	1.0105	1.0518	1.0852	1.1138	1.1384	1.1600	1.1802	1.1985
Theoretical	0.8725	0.9562	1.0105	1.0516	1.0850	1.1134	1.1382	1.1602	1.1801	1.1982

Table 5: RMSE of the alternative forecasting formulae for k-day-ahead-forecast when $H = 0.2$.

k	1	2	3	4	5	6	7	8	9	10
IPLA	1.0379	1.1202	1.1833	1.2338	1.2778	1.3171	1.3520	1.3826	1.4131	1.4371
IPRA	0.9402	1.0513	1.1282	1.1876	1.2372	1.2801	1.3182	1.3516	1.3835	1.4103
IPTA	0.9746	1.0756	1.1478	1.2041	1.2518	1.2935	1.3305	1.3630	1.3944	1.4201
IPMA	0.9073	1.0315	1.1136	1.1763	1.2277	1.2714	1.3102	1.3445	1.3765	1.4045
FPLA	1.0380	1.1202	1.1834	1.2339	1.2779	1.3171	1.3520	1.3826	1.4131	1.4371
FPRA	0.9402	1.0513	1.1282	1.1876	1.2372	1.2800	1.3181	1.3515	1.3833	1.4102
FPTA	0.9746	1.0757	1.1478	1.2041	1.2519	1.2935	1.3305	1.3629	1.3943	1.4201
FPMA	0.9073	1.0315	1.1136	1.1763	1.2277	1.2713	1.3101	1.3443	1.3763	1.4043
OP	0.9049	1.0294	1.1115	1.1744	1.2257	1.2691	1.3082	1.3425	1.3740	1.4024
Theoretical	0.9049	1.0290	1.1109	1.1737	1.2252	1.2692	1.3079	1.3424	1.3737	1.4024

Table 6: RMSE of the alternative forecasting formulae for k-day-ahead-forecast when $H = 0.25$.

k	1	2	3	4	5	6	7	8	9	10
IPLA	1.1196	1.2358	1.3272	1.3961	1.4629	1.5176	1.5642	1.6105	1.6567	1.6929
IPRA	0.9932	1.1425	1.2496	1.3307	1.4029	1.4631	1.5151	1.5642	1.6114	1.6508
IPTA	1.0388	1.1765	1.2782	1.3550	1.4254	1.4837	1.5337	1.5818	1.6288	1.6671
IPMA	0.9409	1.1085	1.2230	1.3099	1.3837	1.4460	1.5003	1.5505	1.5976	1.6384
FPLA	1.1195	1.2357	1.3271	1.3960	1.4628	1.5175	1.5641	1.6103	1.6566	1.6928
FPRA	0.9931	1.1424	1.2495	1.3306	1.4027	1.4629	1.5149	1.5640	1.6112	1.6506
FPTA	1.0387	1.1764	1.2781	1.3548	1.4253	1.4835	1.5335	1.5816	1.6287	1.6669
FPMA	0.9408	1.1085	1.2229	1.3099	1.3836	1.4459	1.5001	1.5504	1.5975	1.6383
OP	0.9327	1.1006	1.2144	1.3023	1.3754	1.4382	1.4935	1.5432	1.5893	1.6304
Theoretical	0.9325	1.1006	1.2140	1.3021	1.3751	1.4381	1.4937	1.5436	1.5892	1.6312

Table 7: RMSE of the alternative forecasting formulae for k-day-ahead-forecast when $H = 0.4$.

k	1	2	3	4	5	6	7	8	9	10
IPLA	1.4011	1.6381	1.8308	2.0009	2.1445	2.2779	2.3928	2.4944	2.6039	2.7015
IPRA	1.1730	1.4537	1.6677	1.8508	2.0059	2.1456	2.2680	2.3768	2.4892	2.5898
IPTA	1.2581	1.5235	1.7302	1.9091	2.0602	2.1979	2.3176	2.4236	2.5352	2.6348
IPMA	1.0403	1.3573	1.5876	1.7790	1.9410	2.0850	2.2121	2.3258	2.4395	2.5423
FPLA	1.4005	1.6373	1.8299	1.9999	2.1434	2.2767	2.3915	2.4931	2.6025	2.7000
FPRA	1.1725	1.4531	1.6670	1.8500	2.0051	2.1447	2.2671	2.3758	2.4882	2.5887
FPTA	1.2576	1.5228	1.7294	1.9082	2.0593	2.1968	2.3165	2.4225	2.5340	2.6335
FPMA	1.0400	1.3569	1.5871	1.7784	1.9404	2.0843	2.2114	2.3251	2.4387	2.5415
OP	0.9886	1.3024	1.5302	1.7162	1.8771	2.0184	2.1467	2.2637	2.3730	2.4756
Theoretical	0.9881	1.3020	1.5304	1.7165	1.8764	2.0181	2.1462	2.2638	2.3729	2.4749

Let $\hat{H}_{ML}$ and $\widehat{\sigma}^2_{ML}$ denote the maximum likelihood (ML) estimator of H and σ^2, respectively. The ML method is expected to deliver more efficient estimates than method-of-moment (MM), the latter of which was proposed in Lang and Roueff (2001), Barndorff-Nielsen et al. (2013), Brouste et al. (2020), and Wang et al. (2021), although MM is computationally cheaper to implement. It is also expected to be more efficient than the composite likelihood method of Bennedsen et al. (2022) although the latter method is applicable to more complicated models.

4.2 Results

We now examine the performance of alternative forecasting formulae based on real data. We download daily RV time series of the S&P 500 market ETF and nine industry ETFs from the Risk Lab constructed by Dacheng Xiu.[7] The sample period is from October 1, 2017 to September 30, 2022. We model the log of each RV time series by fBm. The ML method is then applied to estimate H and σ^2. A two-year rolling window is used to fit fBm and based on the ML estimates of H and σ^2 in fBm, we obtain k-day-ahead-forecast of log RV with $k = 1, 2, ..., 10$ for the remaining three years. We set $\Delta = 1/252$ in this exercise to reflect 252 trading days in a year.

The top panel of Figure 6 plots the full sample of the log RV of S&P500 EFT. The bottom panel plots the rolling window estimates of H . These estimates fluctuate within the interval $[0.2, 0.3]$ and are much lower than 0.5, indicating roughness.

Table 8 reports the RMSEs of the alternative forecasting formulae for the k-day-ahead-forecast of log RV of SPY with $k = 1, 2, ..., 10$. It can be seen that OP always yields the lowest RMSE. The smaller the k is, the biggest its improvement over other forecasting formulae. For example, it improves the RMSE over the IPLA method, which is the method used by Gatheral et al. (2018) and Wang et al. (2021), by 20.6%, 12.4%, 7.1%, 5.4%, 5.7%, 4.2%, 3.6%, 4.0%, 3.6%, 3.8% for $k = 1, ..., 10$, respectively. The magnitude of these improvements is consistent with our findings based on simulated data. Judged by the forecast results provided in the literature (Andersen et al. 2003, Gatheral et al. 2018, Wang et al. 2021), 20.6% for 1-day-ahead-forecast is large and economically significant. Both IPLA and FPLA perform relatively poorly. The reasons for the relative poor performance of IPLA and FPLA can be found in Figure 5 and Table 1, including zero weight on $X_{T\Delta}$, too less weight on X_s when s is near 0 and too greater weight when s is near $T\Delta$. Moreover, the

[7] See https://dachxiu.chicagobooth.edu/#risklab.

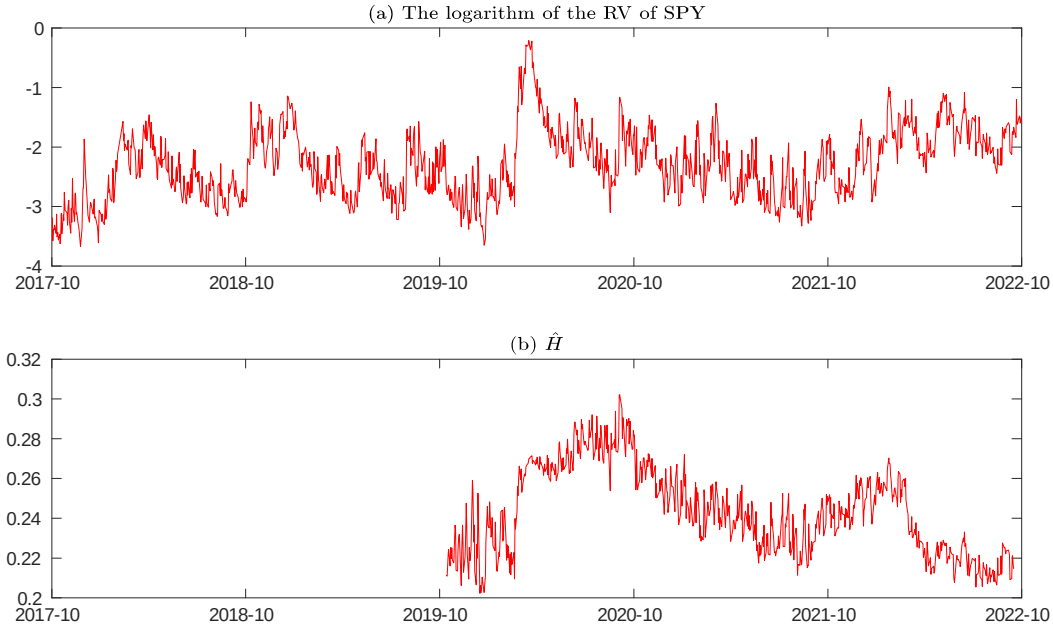

Figure 6: (a) The log RV of S&P 500 ETF between October 1, 2017 and September 30, 2022; (b) Rolling window estimates of H for log RV of S&P 500 ETF

second best is FPMA, which is followed closely by IPMA. The finite-past methods always perform slightly better than the infinite-past methods. All these empirical findings are consistent with the findings obtained from the simulated data.

Figure 7 plots the rolling window estimates of H for the log RV series of the nine industry ETF. These estimates suggest that the estimates are fluctuating within the interval $[0.1, 0.35]$. These values are lower than 0.5, indicating roughness.

Tables 9-17 report the RMSEs of the alternative forecasting formulae for the k-day-ahead-forecast of log RV of the nine industry ETFs with $k = 1, 2, ..., 10$. It can be seen that OP always yields the lowest RMSE with the only exception of XLK and $k = 3, 4$. The smaller the h is, the biggest its improvement over other forecasting formulae. For example, when $h = 1$, its improvement over the IPLA method is $21.5\%, 25.2\%, 24.0\%, 23.3\%, 19.3\%, 24.3\%, 27.6\%, 21.7\%, 19.2\%$ for the nine industry ETF, respectively. These improvements are large and economically significant. Again, both IPLA and FPLA perform relatively poorly. Moreover, the second best methods are FPMA and IPMA. Interestingly, Table 13 shows that the OP method is not the best when $k = 3, 4$ for XLK. One possible explanation is that there may exist model mis-specification for this particular RV series.

Table 8: RMSE of the alternative forecasting formulae for k-day-ahead-forecast of log RV of SPY between October 1, 2017 and September 30, 2022. Boldface corresponds to the lowest RMSE.

k	1	2	3	4	5	6	7	8	9	10
IPLA	.3332	.3684	.3917	.4095	.4279	.4420	.4540	.4664	.4777	.4893
IPRA	.2951	.3406	.3721	.3932	.4115	.4285	.4421	.4540	.4664	.4778
IPTA	.3089	.3508	.3791	.3990	.4176	.4334	.4464	.4587	.4706	.4822
IPMA	.2783	.3292	.3659	.3887	.4059	.4245	.4390	.4500	.4629	.4737
FPLA	.3331	.3683	.3916	.4093	.4277	.4417	.4537	.4662	.4774	.4890
FPRA	.2950	.3405	.3721	.3931	.4114	.4284	.4419	.4538	.4662	.4776
FPTA	.3088	.3507	.3790	.3989	.4174	.4332	.4462	.4584	.4703	.4819
FPMA	.2783	.3292	.3659	.3886	.4058	.4244	.4388	.4498	.4626	.4735
OP	**.2763**	**.3279**	**.3658**	**.3885**	**.4049**	**.4242**	**.4383**	**.4483**	**.4609**	**.4715**

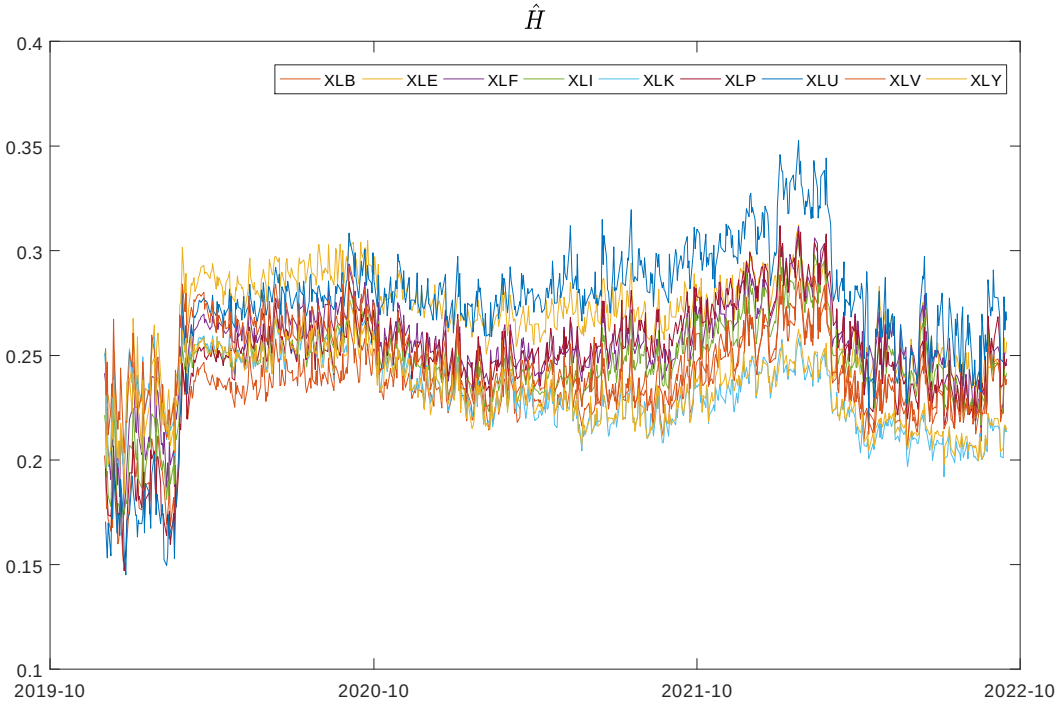

Figure 7: Rolling window estimates of H for log RV of 9 industry ETFs

Table 9: RMSE of the alternative forecasting formulae for k-day-ahead-forecast of XLB. Boldface corresponds to the lowest RMSE.

k	1	2	3	4	5	6	7	8	9	10
IPLA	.2512	.2796	.2997	.3161	.3308	.3411	.3508	.3599	.3688	.3775
IPRA	.2220	.2575	.2833	.3018	.3180	.3315	.3415	.3507	.3596	.3687
IPTA	.2327	.2657	.2893	.3071	.3229	.3350	.3449	.3542	.3631	.3721
IPMA	.2090	.2481	.2774	.2968	.3133	.3289	.3390	.3480	.3567	.3657
FPLA	.2512	.2797	.2997	.3161	.3309	.3412	.3508	.3600	.3689	.3776
FPRA	.2220	.2575	.2833	.3018	.3180	.3316	.3416	.3508	.3597	.3687
FPTA	.2327	.2657	.2893	.3071	.3229	.3351	.3450	.3542	.3632	.3722
FPMA	.2090	.2481	.2775	.2969	.3133	.3290	.3390	.3480	.3568	.3658
OP	**.2068**	**.2460**	**.2760**	**.2955**	**.3121**	**.3283**	**.3377**	**.3462**	**.3542**	**.3631**

Table 10: RMSE of the alternative forecasting formulae for k-day-ahead-forecast of XLE. Boldface corresponds to the lowest RMSE.

k	1	2	3	4	5	6	7	8	9	10
IPLA	.2118	.2371	.2562	.2700	.2826	.2929	.3025	.3121	.3219	.3306
IPRA	.1849	.2174	.2406	.2577	.2711	.2829	.2927	.3022	.3119	.3214
IPTA	.1948	.2247	.2464	.2622	.2754	.2866	.2964	.3060	.3158	.3250
IPMA	.1717	.2088	.2343	.2537	.2672	.2798	.2898	.2988	.3082	.3182
FPLA	.2118	.2370	.2562	.2700	.2826	.2929	.3025	.3121	.3219	.3306
FPRA	.1848	.2173	.2405	.2577	.2711	.2829	.2927	.3022	.3119	.3215
FPTA	.1948	.2246	.2464	.2622	.2754	.2866	.2964	.3060	.3158	.3250
FPMA	.1717	.2088	.2343	.2537	.2672	.2798	.2898	.2989	.3083	.3183
OP	**.1692**	**.2067**	**.2328**	**.2528**	**.2659**	**.2783**	**.2875**	**.2957**	**.3045**	**.3147**

Table 11: RMSE of the alternative forecasting formulae for k-day-ahead-forecast of XLF. Boldface corresponds to the lowest RMSE.

k	1	2	3	4	5	6	7	8	9	10
IPLA	.2470	.2763	.2972	.3135	.3285	.3417	.3532	.3637	.3738	.3835
IPRA	.2165	.2532	.2801	.2989	.3148	.3293	.3422	.3533	.3633	.3736
IPTA	.2277	.2618	.2864	.3043	.3200	.3341	.3464	.3573	.3674	.3775
IPMA	.2021	.2430	.2738	.2941	.3100	.3250	.3386	.3498	.3596	.3700
FPLA	.2470	.2763	.2972	.3135	.3285	.3417	.3532	.3638	.3738	.3835
FPRA	.2165	.2532	.2801	.2989	.3148	.3294	.3422	.3533	.3634	.3737
FPTA	.2277	.2618	.2864	.3043	.3200	.3341	.3464	.3573	.3674	.3775
FPMA	.2021	.2430	.2738	.2941	.3100	.3250	.3386	.3499	.3596	.3701
OP	**.1992**	**.2404**	**.2721**	**.2923**	**.3080**	**.3232**	**.3365**	**.3474**	**.3565**	**.3670**

Table 12: RMSE of the alternative forecasting formulae for k-day-ahead-forecast of XLI. Boldface corresponds to the lowest RMSE.

k	1	2	3	4	5	6	7	8	9	10
IPLA	.2567	.2884	.3114	.3288	.3449	.3567	.3679	.3786	.3885	.3975
IPRA	.2254	.2634	.2928	.3136	.3308	.3456	.3573	.3683	.3784	.3883
IPTA	.2369	.2728	.2998	.3193	.3362	.3497	.3613	.3722	.3823	.3918
IPMA	.2111	.2521	.2856	.3082	.3255	.3423	.3540	.3649	.3748	.3852
FPLA	.2567	.2884	.3114	.3289	.3449	.3567	.3680	.3787	.3886	.3976
FPRA	.2254	.2634	.2929	.3136	.3309	.3457	.3574	.3684	.3785	.3884
FPTA	.2369	.2729	.2998	.3193	.3362	.3498	.3614	.3723	.3824	.3919
FPMA	.2111	.2521	.2856	.3082	.3256	.3423	.3541	.3650	.3749	.3853
OP	**.2082**	**.2492**	**.2839**	**.3067**	**.3242**	**.3415**	**.3527**	**.3632**	**.3726**	**.3829**

Table 13: RMSE of the alternative forecasting formulae for k-day-ahead-forecast of XLK. Boldface corresponds to the lowest RMSE.

k	1	2	3	4	5	6	7	8	9	10
IPLA	.2958	.3248	.3435	.3566	.3701	.3814	.3911	.4014	.4100	.4193
IPRA	.2632	.3016	.3282	.3444	.3578	.3709	.3815	.3914	.4004	.4097
IPTA	.2749	.3100	.3335	.3486	.3622	.3746	.3849	.3951	.4040	.4133
IPMA	.2492	.2924	**.3241**	**.3419**	.3544	.3683	.3794	.3886	.3977	.4066
FPLA	.2958	.3248	.3435	.3567	.3702	.3815	.3911	.4015	.4102	.4194
FPRA	.2632	.3016	.3283	.3445	.3579	.3709	.3816	.3915	.4005	.4098
FPTA	.2749	.3100	.3335	.3486	.3623	.3747	.3850	.3952	.4041	.4134
FPMA	.2492	.2924	**.3241**	.3420	.3544	.3683	.3795	.3887	.3978	.4067
OP	**.2479**	**.2919**	.3246	.3422	**.3541**	**.3681**	**.3786**	**.3870**	**.3954**	**.4042**

Table 14: RMSE of the alternative forecasting formulae for k-day-ahead-forecast of XLP. Boldface corresponds to the lowest RMSE.

k	1	2	3	4	5	6	7	8	9	10
IPLA	.2471	.2788	.3024	.3215	.3400	.3538	.3677	.3797	.3919	.4039
IPRA	.2163	.2542	.2828	.3042	.3234	.3405	.3544	.3677	.3800	.3928
IPTA	.2277	.2635	.2903	.3109	.3299	.3457	.3597	.3724	.3848	.3972
IPMA	.2021	.2433	.2753	.2978	.3166	.3360	.3494	.3634	.3753	.3882
FPLA	.2471	.2788	.3024	.3215	.3400	.3538	.3678	.3797	.3920	.4039
FPRA	.2163	.2542	.2828	.3042	.3234	.3405	.3544	.3677	.3801	.3928
FPTA	.2277	.2635	.2903	.3109	.3299	.3457	.3597	.3725	.3848	.3973
FPMA	.2021	.2433	.2753	.2978	.3166	.3361	.3494	.3634	.3754	.3882
OP	**.1988**	**.2397**	**.2722**	**.2947**	**.3134**	**.3336**	**.3464**	**.3605**	**.3720**	**.3854**

Table 15: RMSE of the alternative forecasting formulae for k-day-ahead-forecast of XLU. Boldface corresponds to the lowest RMSE.

k	1	2	3	4	5	6	7	8	9	10
IPLA	.2101	.2398	.2630	.2821	.2977	.3125	.3262	.3384	.3502	.3615
IPRA	.1822	.2167	.2437	.2652	.2828	.2986	.3131	.3266	.3389	.3509
IPTA	.1926	.2255	.2512	.2718	.2886	.3041	.3183	.3313	.3435	.3552
IPMA	.1689	.2059	.2355	.2582	.2769	.2931	.3079	.3220	.3343	.3465
FPLA	.2101	.2398	.2630	.2821	.2977	.3124	.3261	.3384	.3502	.3614
FPRA	.1822	.2166	.2437	.2651	.2828	.2986	.3131	.3265	.3388	.3508
FPTA	.1926	.2255	.2512	.2718	.2886	.3041	.3183	.3313	.3434	.3551
FPMA	.1689	.2059	.2355	.2582	.2769	.2931	.3079	.3220	.3342	.3465
OP	**.1647**	**.2012**	**.2313**	**.2544**	**.2734**	**.2895**	**.3042**	**.3185**	**.3304**	**.3429**

Table 16: RMSE of the alternative forecasting formulae for k-day-ahead-forecast of XLV. Boldface corresponds to the lowest RMSE.

k	1	2	3	4	5	6	7	8	9	10
IPLA	.2447	.2733	.2940	.3095	.3251	.3367	.3473	.3568	.3667	.3762
IPRA	.2160	.2511	.2772	.2955	.3115	.3262	.3373	.3474	.3573	.3673
IPTA	.2264	.2593	.2834	.3007	.3167	.3300	.3410	.3509	.3609	.3707
IPMA	.2033	.2416	.2709	.2907	.3061	.3228	.3340	.3443	.3539	.3639
FPLA	.2446	.2733	.2940	.3096	.3252	.3367	.3473	.3569	.3667	.3763
FPRA	.2159	.2511	.2772	.2955	.3115	.3262	.3373	.3474	.3573	.3673
FPTA	.2264	.2593	.2834	.3007	.3167	.3301	.3410	.3510	.3609	.3707
FPMA	.2033	.2416	.2709	.2907	.3061	.3228	.3340	.3443	.3539	.3639
OP	**.2011**	**.2392**	**.2692**	**.2890**	**.3042**	**.3217**	**.3325**	**.3424**	**.3515**	**.3615**

Table 17: RMSE of the alternative forecasting formulae for k-day-ahead-forecast of XLY. Boldface corresponds to the lowest RMSE.

k	1	2	3	4	5	6	7	8	9	10
IPLA	.2926	.3222	.3412	.3566	.3717	.3832	.3931	.4019	.4118	.4229
IPRA	.2605	.2992	.3253	.3425	.3581	.3724	.3832	.3922	.4013	.4118
IPTA	.2721	.3075	.3309	.3475	.3631	.3763	.3867	.3957	.4052	.4161
IPMA	.2469	.2903	.3210	.3386	.3537	.3696	.3809	.3900	.3983	.4079
FPLA	.2926	.3222	.3412	.3566	.3718	.3833	.3932	.4021	.4119	.4231
FPRA	.2605	.2992	.3253	.3425	.3581	.3724	.3832	.3923	.4014	.4119
FPTA	.2721	.3075	.3309	.3475	.3631	.3763	.3868	.3958	.4054	.4162
FPMA	.2469	.2903	.3210	.3387	.3537	.3697	.3810	.3901	.3984	.4081
OP	**.2454**	**.2893**	**.3205**	**.3377**	**.3522**	**.3685**	**.3791**	**.3873**	**.3945**	**.4038**

4.3 Additional empirical results

The empirical study reported earlier is based on the sample period from October 1, 2017 to September 30, 2022. It is well known that the sample period contains the financial crash in 2020 that began on February 20, 2020, and ended on April 7, 2020. This crash explains why the estimated H in the bottom panel of Figure 6 moved up from its values from the near 0.2 level in February 2020 to the near 0.3 level in April 2020.

To check the robustness of the empirical results reported earlier, we now examine the performance of alternative forecasting formulae based on the daily RV time series of the S&P 500 market ETF from January 2, 2012 to December 31, 2019. Once again, we model the log of each RV time series by fBm and use the ML method to estimate H and σ^2. A fix-year rolling window is used to fit fBm and based on the ML estimates of H and σ^2 in fBm, we obtain k-day-ahead-forecast of log RV with $k = 1, 2, ..., 10$ for the remaining two years.

The top panel of Figure 8 plots the full sample of the log RV of S&P500 EFT. The bottom panel plots the rolling window estimates of H. These estimates fluctuate within the much narrower interval $[0.19, 0.22]$ and are much lower than 0.5, indicating roughness.

Table 18 reports the RMSEs of the alternative forecasting formulae for the k-day-ahead-forecast of log RV of SPY with $k = 1, 2, ..., 10$. It can be seen that OP continues to yield the lowest RMSE always. The smaller the k is, the biggest its improvement over other forecasting formulae. For example, it improves the RMSE over the IPLA method, which is the method used by Gatheral et al. (2018) and Wang et al. (2021), by 16.3%, 8.8%, 7.3%, 6.7%, 5.3%, 4.6%, 3.7%, 3.5%, 3.7%, 3.0% for $k = 1, ..., 10$, respectively. The magnitude of these improvements is consistent with our findings based on simulated data and the sample between October 1, 2017 and September 30, 2022. This finding suggests that our empirical results are not driven by the 2020 financial crash.

5 Conclusion

In this paper, we have examined the performance of alternative forecasting formulae with the fractional Brownian motion based on a discrete and finite sample. In the literature, two optimal forecasting formulae, both based on a continuous record, have

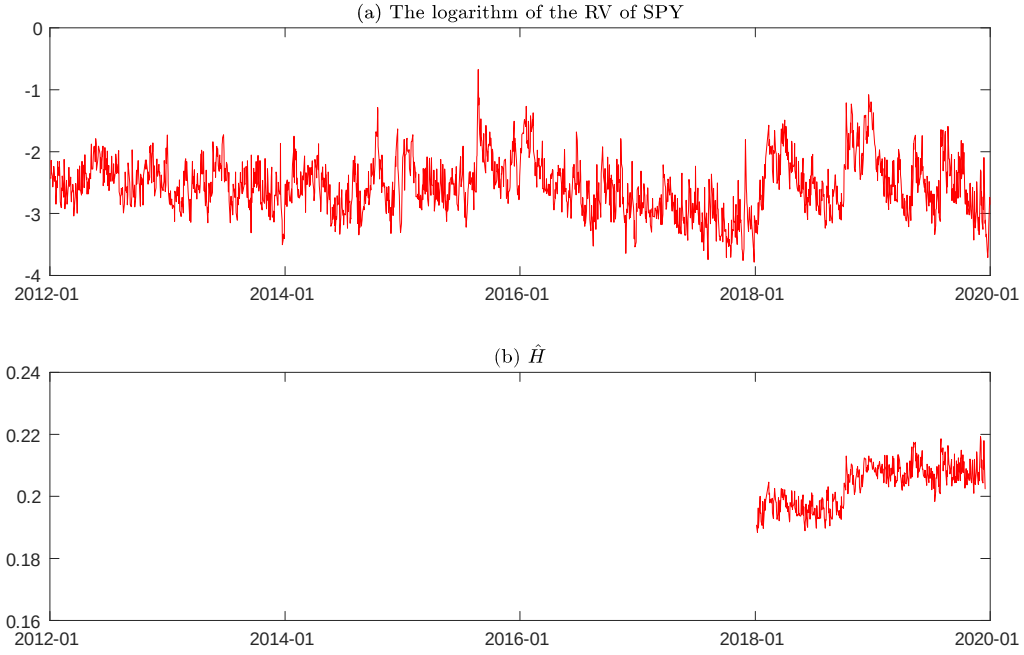

Figure 8: (a) The log RV of S&P 500 ETF between January 2, 2012 and December 31, 2019; (b) Rolling window estimates of H for log RV of S&P 500 ETF.

Table 18: RMSE of the alternative forecasting formulae for k-day-ahead-forecast of log RV of SPY between January 2, 2012 and December 31, 2019. Boldface corresponds to the lowest RMSE.

k	1	2	3	4	5	6	7	8	9	10
IPLA	.3205	.3463	.3680	.3876	.4048	.4209	.4322	.4426	.4530	.4619
IPRA	.2887	.3252	.3497	.3705	.3904	.4077	.4211	.4319	.4421	.4525
IPTA	.3001	.3327	.3563	.3769	.3958	.4127	.4253	.4359	.4463	.4561
IPMA	.2768	.3195	.3446	.3650	.3862	.4037	.4183	.4292	.4387	.4499
FPLA	.3205	.3463	.3679	.3875	.4047	.4208	.4321	.4424	.4528	.4617
FPRA	.2887	.3252	.3496	.3704	.3903	.4076	.4210	.4318	.4419	.4523
FPTA	.3001	.3326	.3563	.3768	.3957	.4126	.4251	.4358	.4462	.4559
FPMA	.2768	.3195	.3446	.3649	.3861	.4036	.4182	.4290	.4385	.4497
OP	**.2756**	**.3184**	**.3430**	**.3632**	**.3846**	**.4022**	**.4169**	**.4275**	**.4370**	**.4485**

been derived. The first one is based on an infinite past and the other is based on a finite past. When a discrete and finite past sample is available, one may discretize these two forecasting formulae by alternative schemes, leading to two classes of forecasting formulae.

Instead of discretizing these two forecasting formulae, one can use the conditional expectation to obtain the optimal forecast based on a discrete and finite past sample. The conditional expectation is optimal in the sense of minimizing the RMSE.

Via simulated data and real RV data, we show that the conditional expectation always yields more accurate forecasts than two classes of discretization-based forecasting formulae for all forecasting horizons considered. The shorter the forecast horizon is, the bigger the improvement. When the forecast horizon is one day, for example, we have found that the conditional expectation can improve the forecasting formula currently implemented in the literature by 20% or so. This improvement is large and economically significant.

6 Appendix

Proof of Proposition 2.1. Without loss of generality, let us assume $k = 1$. In this case, the weight function becomes

$$\omega_1(s) = \frac{\cos(H\pi)}{\pi} \frac{1}{(T-s+1)(T-s)^{H+0.5}}$$

We then have

$$\begin{aligned}
\int_{-\infty}^{T} \omega_1(s)\, ds &= \frac{\cos(H\pi)}{\pi} \int_{-\infty}^{T} \frac{1}{(T-s+1)(T-s)^{H+0.5}} ds \\
&= \frac{\cos(H\pi)}{\pi} \int_{0}^{\infty} \frac{1}{(1+u)u^{H+0.5}} du \\
&= \frac{\cos(H\pi)}{\pi} \pi \csc((H+0.5)\pi) \\
&= \frac{\cos(H\pi)}{\sin(H\pi + 0.5\pi)} \\
&= \frac{\cos(H\pi)}{\sin(0.5\pi)\cos(H\pi) + \cos(0.5\pi)\sin(H\pi)} \\
&= 1.
\end{aligned}$$

where $\csc((H+0.5)\pi) := 1/\sin((H+0.5)\pi)$ and the third equation comes from the Formula 361 on Page 290 of Tallarida (2015).

References

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2003. Modeling and forecasting realized volatility. *Econometrica* 71(2), 579–625.

Barndorff-Nielsen, O. E., Corcuera, J. M., Podolskij, M., 2013. Limit theorems for functionals of higher order differences of Brownian semi-stationary processes. In *Prokhorov and Contemporary Probability Theory: In Honor of Yuri V. Prokhorov*, Volume 33, 69–96. Springer Science & Business Media.

Bayer, C., Friz, P., Gatheral, J., 2016. Pricing under rough volatility. *Quantitative Finance* 16(6), 887–904.

Bennedsen, M., Christensen, K., Christensen, P., 2022. Likelihood-Based Estimation of Rough Stochastic Volatility Models, Aarhus University, manuscript.

Bolko, A. E., Christensen, K., Pakkanen, M. S. , and Veliyev, B. (2022). A GMM approach to estimate the roughness of stochastic volatility. *Journal of Econometrics*, forthcoming.

Brouste, A., Soltane, M., Votsi, I., 2020. One-step estimation for the fractional Gaussian noise at high-frequency. *ESAIM: Probability and Statistics* 24, 827–841.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7(2), 174–196.

Euch, O. E., Rosenbaum, M., 2018. Perfect hedging in rough Heston models. *Annals of Probability* 28(6), 3813–3856.

Fukasawa, M., T. Takabatake, and R. Westphal (2021). Consistent estimation for fractional stochastic volatility model under high-frequency asymptotics. *Mathematical Finance*, forthcoming.

Fouque, J. P., Hu, R., 2018. Optimal portfolio under fast mean–reverting fractional stochastic environment. *SIAM Journal of Financial Mathematics* 9(2), 564–601.

Garnier, J., Sølna, K., 2017. Correction to Black-Scholes formula due to fractional stochastic volatility. *SIAM Journal Financial Mathematics* 8(1), 560–588.

Gatheral, J., Jaisson, T., Rosenbaum, M., 2018. Volatility is rough. *Quantitative Finance* 18(6), 933–949.

Lang, G., Roueff, F., 2001. Semi-parametric estimation of the Hölder exponent of a stationary Gaussian process with minimax rates. *Statistical Inference for Stochastic Processes* 4(3), 283–306.

Livieri, G., Mouti, S., Pallavicini, A., Rosenbaum, M., 2018. Rough volatility: evidence from option prices. *IISE Transactions* 50(9), 767–776.

Mandelbrot, B.B., Van Ness, J.W., 1968. Fractional Brownian motions, fractional noises and applications. *SIAM Review*. 10, 422–437.

Nuzman, C. J., Poor, H. V., 2000. Linear Estimation of Self-Similar Processes via Lamperti's Transformation. *Journal of Applied Probability*, 37(2), 429-452.

Tallarida, R. J., 2015. Pocket Book of Integrals and Mathematical Formulas. 5th Edition, Taylor & Francis Group.

Wang, X., Xiao, W., Yu, J., 2019. Estimation and Inference of Fractional Continuous-Time Model with Discrete-Sampled Data. *Singapore Management University Economics and Statistics Working Paper Series*, 17-2019.

Wang, X., Xiao, W., Yu, J., 2021. Modeling and Forecasting Realized Volatility with the Fractional Ornstein-Uhlenbeck Process. *Journal of Econometrics* https://doi.org/10.1016/j.jeconom.2021.08.001.