# Volatility is rough: high-frequency modelling and forecasting in FX markets

Author: Alberto Saavedra Espinosa

Supervised by Prof. Mikko S. Pakkanen

# Executive Summary

## Objectives

This research work has two main objectives: i) investigate whether volatility in Forex Exchange (FX) markets is *rough* both at daily and intraday scales; and ii) examine whether volatility models that incorporate *both* the rough nature and the memory component of volatility can perform well on forecasting volatility in FX markets at daily and intraday scales.

## Background

Volatility modelling and forecasting are useful for many financial practices. A couple of examples are market risk measurements and making speculative bets in financial markets (e.g. through the use of derivatives such as volatility swaps). Traditional applications typically require estimates and forecasts of volatility only at daily or lower frequencies, but the advent of electronic trading in general and high-frequency trading in particular have called for the study of volatility at intraday frequencies.

Thus, during approximately the last 15 years both theoretical and empirical research has been carried out to design statistical approaches to model and forecast volatility at daily and intraday scales that exploit the availability of high-frequency data (see, for instance, Barndorff-Nielsen and Shephard (2004); Barndorff-Nielsen and Schmiegel (2009); Corsi (2009)). For this, the main empirical properties of volatility were studied and given mathematical meaning in this high-frequency framework.

Among the properties of volatility, perhaps the most popular is its high persistence, which essentially means strong dependence over long time scales. Indeed, this property has been largely discussed in academia, having both supporters (Andersen and Bollerslev, 1997) and doubters (Mikosch and Starica, 2000). Yet, the memory of volatility has always been present in the discussion of this financial phenomenon.

Notwithstanding, a new stream of research started by Gatheral et al. (2018) has been attempting to establish whether volatility exhibits, in general, another property: roughness of its sample paths. In plain terms, this essentially means that volatility exhibits an irregular behaviour at short time scales, but this notion can be given a precise mathematical meaning.

Establishing roughness as a general characteristic of volatility is important, as it has the potential to aid the design of better models for this phenomenon. In turn, this could help to improve the final outcome of applications such as volatility forecasting, pricing of financial instruments, and market risk measurements. In this regard, studies such as Gatheral et al. (2018) and Simity and Milán (2018) have empirically established that volatility does exhibit roughness of its sample paths, the first work in equities and futures markets and the second one on some FX pairs. However, these studies focused their attention only on daily and lower frequencies, leaving intraday volatility unexplored. Bennedsen et al. (2017), on the other hand, showed that volatility is rough also at intraday scales but their data consisted only of information from equities markets.

Therefore, the first aim of our work is to extend the empirical evidence on the roughness of volatility at intraday scales to a new financial market: The FX market. As far as we know, there is not yet evidence of volatility exhibiting roughness at both daily and intraday scales in this market, so our study provides the first piece of empirical evidence in this regard. This is particularly relevant for the FX market, as recent studies have shown its fast electronification and the transition of its activity towards more prevalent algorithmic and high-frequency trading (BIS Markets Committee, 2018).

On the other hand, studies such as Gatheral et al. (2018) have shown that incorporating the roughness property into the modelling of daily volatility can produce better volatility forecasts than when this feature is ignored. Furthermore, Bennedsen et al. (2017) have shown that taking into consideration both the roughness and memory components of volatility can produce even better forecasts, both at daily and intraday scales. These studies, however, focused their attention solely on equities and futures markets. Therefore, our second aim is to investigate whether volatility models incorporating roughness and memory can successfully forecast volatility in FX markets, both at daily and intraday scales. Establishing superior forecasting performance at intraday and daily scales in the FX market is key to achieve more robust financial practices in this market, given its fast paced and electronic nature.

## Table of Contents

The structure of our research work is as follows:



## Approach

Our research work consisted of estimating volatility using high-frequency data from FX markets. Specifically, we estimated volatility over time intervals ranging from a few minutes up to a full trading day, using the statistical methodology described below. For this, we used tick-by-tick observations, from April 2009 to May 2019, of the midprice quotes for four different exchange rates.[1] The exchange rates we used were: EUR/USD, GBP/USD, USD/CAD, and USD/JPY. These represent four of the most important currency pairs worldwide, which are of major relevance for any open economy. In the case of Mexico, for instance, these four currencies are among the ones that are eligible to constitute the international reserves that the Bank of Mexico manages. Thus, researching their empirical features is highly relevant in order to improve the management of this portfolio, which in turn could aid to more efficiently achieve the objectives of the Mexican exchange policy.[2]

[1] A tick (or pip for FX markets) is the smallest possible movement of a price. The term tick-by-tick indicates that a price is recorded every time it changes, which is almost equivalent to record it every time there is a trade on a given financial instrument.

[2] The results of this research could in fact benefit any Mexican enterprise or bank with investments on the main currencies of the world. We highlighted the case of the international reserves in main text due to its relevance for the economy of Mexico as a whole.

Once we estimated volatility in the FX market from our high-frequency data, we used the methodologies developed by Gatheral et al. (2018) and Bennedsen et al. (2017) on these estimates to empirically investigate whether its sample paths are consistent with the roughness property. The methodology from Gatheral et al. (2018) had been used before to establish roughness of volatility at daily and lower frequencies in equity, futures, and FX markets. In contrast, the methodology derived in Bennedsen et al. (2017) allowed to provide empirical evidence of volatility being rough also at intraday frequencies in equity markets. To the best of our knowledge, our study was the first to analyse volatility roughness at both daily and intraday frequencies on FX market data.

In this regard, our study not only empirically investigated whether volatility in the FX market is rough at all relevant scales for practical applications, but was also the first one to explore whether the degree of roughness of volatility varies over time in this market. This kind of behaviour had been observed in equity markets (Bennedsen et al, 2017), and stablish its generality aids to understand the microstructural foundations behind the roughness of volatility (El Euch et al, 2016), which is crucial to design better models for this phenomenon.

Once we investigated the roughness property of volatility in the FX market, we conducted a comprehensive forecasting study, where we compared volatility models that incorporate roughness in their structure with models that do not. Moreover, we also made comparisons between models only considering the roughness property with a fixed memory structure against models that *de-couple* these features and let data define the specific nature of both characteristics. The comparison of the volatility forecasts given by our models was made through robust techniques oriented towards comparisons that attempt to answer whether one forecast series is statistically better than other ones (Diebold and Mariano, 2002, Hansen et al, 2011). Our research work made these volatility forecasts comparisons across all the frequencies relevant for practical applications.

The results of our forecasting study can allow Mexican financial institutions to compute better volatility forecasts for the major currencies in the world, which in turn can lead them to better manage their risks when trading with foreign parties. Specifically, by knowing what kind of environment they might find in the FX markets in the future, these agents can make better decisions regarding their investments and exposure to risk on a foreign currency. Furthermore, as our results span daily and intraday frequencies, the spectrum of Mexican institutions that can benefit from them ranges from sophisticated trading firms to enterprises with just a moderate exposure to the FX market.

## Hypothesis

Our main hypothesis was that roughness is a genuine characteristic of volatility in FX markets both at daily and intraday scales. Our work then aimed to show this empirically, which we considered highly relevant given the size and liquidity of the FX market. Ultimately, our work would contribute to stablish roughness as a characteristic feature of volatility in any financial market.

Furthermore, we considered that, volatility being inherently rough, incorporating this feature into its modelling could yield better volatility forecasts than models that ignore this feature. This, for volatility forecast at frequencies ranging from a few minutes to a full trading day.

Finally, we believed that the memory component of volatility was also highly relevant for its modelling and forecasting, so that models considering both roughness and memory could yield better overall volatility forecasts in the FX market. This again, for volatility forecast at frequencies ranging from a few minutes to a full trading day.

## Theory and methods

### *Volatility estimation from high-frequency financial data*

The content of this subsection is based on Pakkanen (2019). In order to measure volatility, we model the log midprice of financial asset (e.g. a stock, an interest rate, or an exchange rate) $\log S = (\log S_t)_{t\geq 0}$ over a given time horizon $[0,T]$ by an Itô process

$$\log S_t = \log S_0 + \int_0^t \mu_u du + \int_0^t \sigma_u dW_u \, , t \in [0,T].$$

We measure the realisation of the volatility process $\sigma = (\sigma_t)_{t\in[0,T]}$ through an averaged quantity: the integrated variance. The integrated variance of $S$ over $[0,T]$ is given by

$$IV = \int_0^T \sigma_u^2 du.$$

The integrated variance is not directly observable since we do not actually observe $\sigma$ or even $S$ continuously. This is why the integrated variance needs to be estimated. We use the bi-power variation of Barndorff-Nielsen and Shephard (2004) to estimate the integrated variance. The realised bi-power variation of the observations $S_0, S_\Delta, \dots, S_{(n-1)\Delta}, S_{n\Delta}$, where $n\Delta = T$, is

$$BV_n = \sqrt{\frac{2}{\pi}} \frac{n}{n-1} \sum_{i=2}^{n} \left| r_{(i-1)\Delta, i\Delta} \right| \left| r_{(i-2)\Delta,(i-1)\Delta} \right|,$$

where $r_{(i-1)\Delta, i\Delta} = \log S_{i\Delta} - \log S_{(i-1)\Delta}$ is the intraday log-return over the interval $[(i-1)\Delta, i\Delta]$. Barndorff-Nielsen and Shephard (2004) prove that, as $n \to \infty$, the realised bi-power variation is a consistent estimator of integrated variance.

While this consistency would in principle suggest using $\Delta$ as small as possible, in practice this is not done since at ultra-high frequencies an Itô process is no longer a good model for prices. Essentially, this happens because, at ultra-high frequencies, prices do not evolve continuously.[3] The discrepancy between an Itô process and the observed prices at ultra-high frequency is called market microstructure noise (MMS noise). Thus, in practice we choose the parameter $\Delta$ by balancing the trade-off between mitigating the effect of MMS noise in our estimations of volatility but using the maximum possible amount of data.

In our application, we made the selection of $\Delta$ based on signature plots, as suggested in Pakkanen (2019). It is worth mentioning that we measured volatility over 5 different time horizons: 15 minutes, 30 minutes, 60 minutes, 126 minutes, and a full trading day of 21 hours.[4] For every period length and FX pair, an adequate $\Delta$ was chosen via signature plots.

When measuring and forecasting volatility at intraday scales (i.e. over horizons shorter than a full trading day), accounting for the intraday periodicity of the volatility process becomes relevant.[5] Before performing any analyses at this scale, it is important to control for this periodicity, since volatility estimates and forecasts could be affected if one does not take this into account (Rossi and Fantazzini, 2015). In our application, we estimated intraday periodicity of volatility via the Truncated Maximum Likelihood estimator of Boudt et al. (2011). We chose this estimator since it is robust to jumps in the price process and, being a parametric estimate, is less variable than non-parametric alternatives.

[3] Indeed, at ultra-high frequencies, prices move in discrete amounts as trades are made in the market.

[4] While the FX market never formally closes, we use working days of 21 hours, since this is the largest interval for which we have data. The seemingly odd period of 126 minutes was chosen to be able to evenly split these 21-hour trading days in a whole number.

[5] Indeed, works as Andersen and Bollerslev (1997) have shown that volatility displays a significant intraday periodicity not seen at daily or lower frequencies. These periodicity patterns are caused by daily events that affect markets' activity such as opening, lunch, and close times of financial markets.

*Roughness of volatility*

Volatility being *rough* essentially means that the sample paths of this process show an irregular behaviour at short scales. To mathematically understand the meaning of this, it is useful to remind the Hölder continuity condition and to introduce the notion of a fractional Brownian motion (fBM). Additionally, it is worth mentioning that many of the results and models developed in the Rough Volatility literature first target log-volatility processes, since they enjoy better statistical and mathematical properties. Notwithstanding, results are always extended to volatility processes themselves, since this is the phenomenon of financial interest.

Let $f: \mathbb{R} \to \mathbb{R}$ be a function. We say that $f$ is Hölder continuous of exponent $r$ if $\sup_{x, y \in \mathbb{R}} \frac{|f(x)-f(y)|}{|x-y|^r} < \infty$. Hölder continuity is about the roughness of the path of the function $f$. Specifically, the smaller the index $r$, the rougher the path of the function can be, since it allows for more variability in the image of function for a given small distance of points in its domain.

On the other hand, let $0 < H < 1$, and let $w_0$ be an arbitrary real number. We call process $W^H = (W_t^H)_{t\in\mathbb{R}}$ a fractional Brownian motion with Hurst index $H$ and starting value $w_0$ at time $t = 0$. For $t > 0$, $W_t^H$ is defined by

$$W_t^H = w_0 + \frac{1}{\Gamma(H+1/2)}\left\{\int_{-\infty}^{0}\left[(t-s)^{H-1/2} - (-s)^{H-1/2}\right]dW_s + \int_0^t\left[(t-s)^{H-1/2}\right]dW_s\right\},$$

where $W = (W_t)_{t\in\mathbb{R}}$ is a standard Brownian motion (BM).[6] The Hurst index describes the raggedness of the resultant motion since the sample paths of $W^H$ are Hölder continuous with exponent $r$ for any $r < H$. Thus, higher values of $H$ lead to smoother motions.

The roughness property can be mathematically captured through scaling properties of the increments of the log-volatility process. Specifically, if the paths of the log-volatility process are rough (Rosenbaum, 2011), then the increments

$$m(q,h) = E[|\log\sigma_h - \log\sigma_0|^q], \tag{1}$$

satisfy the key relationship $m(q,h) \propto h^{\xi_q}$. This relationship suggests a method for investigating whether a volatility process has rough sample paths: compute $m(q,h)$ for different values of $h$ and regress $\log m(q,h)$ against $\log h$. If this relation is linear, we would have evidence of roughness of the volatility process. To estimate the roughness parameter $H$, Gatheral et al. (2018) explore the relationship between $\xi_q$ and $q$, and empirically find that: i) their relationship is linear; and ii) the value of $H$ does not depend on $q$. This allowed them to identify that $\xi_q = qH$, and thus to estimate $H$ from the slopes of the regressions of $\xi_q$ against $q$.[7]

Gatheral et al. (2018) argue that, in terms of the smoothness of the volatility process, classic stochastic volatility models such as the proposed in Heston (1993) offer only volatility trajectories with regularity close to that of BM. Furthermore, they argue that previous works that had attempted to offer other kind of regularity for volatility processes had started from the stylised fact that volatility is a long memory process. However, they also argue that since the frameworks on which the memory of volatility had been measured were rather restrictive, results obtained previously could be spurious. Thus, they do not include this feature in their modelling and use a fBM with a Hurst index less than ½, which is formally a short memory process, to capture the observed roughness of the sample paths of volatility.[8] The model they developed is reviewed in the next subsection. In this regard, in the following two subsections, models for log-volatility will be denoted as a stochastic process $X = (X_t)_{t\in\mathbb{R}}$, and models for volatility itself will then

[6] Note that the fBM is a generalisation of the BM since by choosing $w_0 = 0$ and $H = 1/2$ we recover the standard BM.

[7] Estimating volatility over short time horizons (e.g. 15 minutes) allows us to estimate the roughness parameter more frequently (e.g. daily), since we can run the corresponding regression using these shorter-horizon estimations. This is highly relevant in order to assess whether volatility roughness varies over time.

[8] The empirical finding of Gatheral et al. (2018) was that the estimates of the Hurst index for the data they had were not only less than ½, but actually quite smaller than this threshold. Specifically, they found the value of $H$ to range between 0.08 and 0.2 for their equities and futures markets data.

be given by the exponential of $X$. Forecasting formulas will be derived for $X$, and predictions for volatility itself will be obtained from the log-volatility forecasts.

*Rough Fractional Stochastic Volatility (RFSV) model*

Having shown that the increments of log-volatility enjoy a scaling property with constant smoothness parameter, Gatheral et al. (2018) proposed the following model for these increments:

$$\log \sigma_t - \log \sigma_s = \upsilon(W_t^H - W_s^H), \tag{2}$$

where $t, s \in [0, T]$, $t \geq s$, $W^H$is a fBM with Hurst parameter equal to the measured smoothness of the volatility, and $\upsilon$ is a positive constant. However, model (2) is not stationary, stationarity being desirable both for mathematical tractability and to ensure reasonableness of the model at very large times.

Gatheral et al. (2018) then imposed stationarity by modelling the log-volatility as a fractional Ornstein–Uhlenbeck process with a very long reversion timescale. A stationary fractional Ornstein–Uhlenbeck process $X = (X_t)_{t\in\mathbb{R}}$ is defined as the stationary solution of the stochastic differential equation $dX_t = \upsilon dW_t^H - \alpha(X_t - m)dt$, where $W^H$is a fBM with Hurst parameter equal to the measured smoothness of the volatility, $\upsilon$ and $\alpha$ are a positive constants, and $m \in \mathbb{R}$. As for usual Ornstein–Uhlenbeck processes, there is an explicit form for the solution which is given by

$$X_t = \upsilon \int_{-\infty}^{t} e^{-\alpha(t-s)} dW_t^H + m. \tag{3}$$

Gatheral et al. (2018) then arrived at the final specification of their Rough Fractional Stochastic Volatility (RFSV) model, which models volatility on a time interval of interest $[0, T]$ as:

$$\sigma_t = \exp(X_t), t \in [0, T], \tag{4}$$

where $X_t$ is given by (3). Model (4) is indeed stationary. However, if $\alpha \ll 1/T$, the log-volatility behaves locally (at time scales smaller than $T$) as a fBM. Gatheral et al. (2018) then argued that in the RFSV model, if $\alpha \ll 1/T$, and if we confine ourselves to the interval $[0, T]$, we can proceed as if the log-volatility process were a fBM. Indeed, simply setting $\alpha = 0$ in (3) gives (at least formally) $X_t - X_s = \upsilon(W_t^H - W_s^H)$ and we immediately recover the simple non-stationary fBM model (2). Gatheral et al. (2018) also proved that the fractional Ornstein–Uhlenbeck process approximately reproduces the scaling property of the fBM, which allows to estimate the smoothness parameter for the RFSV model as explained before.

The key formula to forecast volatility with the RFSV model is given by

$$\widehat{\sigma_{t+h}^2} = E\big(\exp(\widehat{X_{t+h}}) | \mathcal{F}_t^{W^H}\big) = \exp\Big(E\big(\log \sigma_{t+h}^2 \,|\widehat{\mathcal{F}_t^{W^H}}\big) + 2c\upsilon^2(h)^{2H}\Big), \tag{5}$$

where $\upsilon^2$ is estimated as the exponential of the intercept in the linear regression of $\log m(2, h)$ vs $\log h$, $c = \frac{\Gamma(3/2-H)}{\Gamma(H+1/2)\Gamma(2-2H)}$, and

$$E\big(\log \sigma_{t+h}^2 \,|\mathcal{F}_t^{W^H}\big) = \frac{\cos(H\pi)}{\pi}(h)^{H+1/2} \int_{-\infty}^{t} \frac{\log \sigma_s^2}{(t-s+h)(t-s)^{H+1/2}} ds,$$

where $h > 0$, and $\mathcal{F}_t^{W^H}$ is the filtration generated by $W^H$. For the derivation details of these formulas, see Gatheral et al. (2018).

*De-coupling models*

We now review the models proposed in Bennedsen et al. (2017), where in contrast with the RFSV model of Gatheral et al. (2018), the objective is to *de-couple* the rough nature of volatility from its persistence component (long-term dependence structure). It is worth mentioning that these models were proposed to model and forecast both daily and *intraday* volatility, whereas Gatheral et al. (2018) focused their attention solely on daily or lower frequencies.

Choosing $q = 2$ in (1) we can further characterise the roughness of volatility through its Autocorrelation Function (ACF) as

$$1 - \rho(h) := 1 - \text{Corr}(\log \sigma_t, \log \sigma_{t+h}) \sim c|h|^{2\alpha+1}, \ |h| \to 0, \tag{6}$$

for a constant $c$ and some $\alpha \in \left(-\frac{1}{2}, 0\right)$. Here "$\sim$" indicates that the ratio between the left- and right-hand side tends to one. Bennedsen et al. (2017) call $\alpha = H - 1/2$ the roughness index of the log-volatility process, which can then be computed for any covariance stationary log-volatility process satisfying $m(2, h) \propto p|h|^{2\alpha+1}$, as $h \to 0$, for a constant $p$ and some $\alpha \in \left(-\frac{1}{2}, 0\right)$.[9]

Bennedsen et al. (2017) proposed two different models for the log-volatility process: The Cauchy and the Brownian Semistationaty (BSS) processes. The main property of these models is their capacity to have an arbitrary roughness index $\alpha \in (-\frac{1}{2}, \frac{1}{2})$ and a memory structure that is independent of the value of $\alpha$.

The Cauchy process is a Gaussian process that decouples the short and long-term behaviour by using the Cauchy class of autocorrelation functions. The resulting Cauchy process is a centred and stationary Gaussian process $X = (X_t)_{t \in \mathbb{R}}$ with ACF given by

$$\rho(h) = (1 + |h|^{2\alpha+1})^{-\frac{\beta}{2\alpha+1}}, \ h \in \mathbb{R}, \tag{7}$$

where $\alpha \in (-\frac{1}{2}, \frac{1}{2})$ and $\beta > 0$. In particular, the process $X$ satisfies (6). However, Bennedsen et al. (2017) pointed out that the limitation of this process, from a modelling point of view, is its inherent Gaussianity.[10]

The BSS process was introduced in Barndorff-Nielsen and Schmiegel (2009). This stochastic process is able to capture roughness, strong persistence, stationarity, non-Gaussianity, and is defined as

$$X_t = \int_{-\infty}^{t} g(t-s) \upsilon_s dW_s, \ t \geq 0, \tag{8}$$

where $W = (W_t)_{t \in \mathbb{R}}$ is a standard BM, $g: (0, \infty) \to \mathbb{R}$ is square-integrable kernel function, and $\upsilon = (\upsilon_t)_{t \in \mathbb{R}}$ is an adapted, covariance-stationary volatility (of volatility) process. If $\upsilon$ is deterministic, $X$ is Gaussian whereas if it is stochastic $X$ is non-Gaussian, thus overcoming the modelling weakness of the Cauchy process.

By covariance stationarity of $\upsilon$, the autocovariance function of the general BSS process (8) is

$$\text{Cov}(X_t, X_{t+h}) = E(\upsilon_0^2) \int_0^{\infty} g(x) g(x + |h|) dx, \ h \in \mathbb{R}. \tag{9}$$

Bennedsen et al. (2017) argued that under the assumptions given above, the process $X$ is already well-defined and covariance stationary. However, they introduced additional technical conditions on kernel function $g$ that allowed them to show that the BSS process $X$ defined by (8) has $\alpha$ as its roughness index in the sense of equation (6). Bennedsen et al. (2017) then used two specific kernel functions, satisfying the afore mentioned technical conditions, to complete the specification of their volatility models:

- The power law kernel is a function $g: (0, \infty) \to \mathbb{R}$ given by $g(x) = x^{\alpha}(1-x)^{-\alpha-\gamma}$, $x > 0$, $\alpha \in \left(-\frac{1}{2}, \frac{1}{2}\right)$, $\gamma \in \left(\frac{1}{2}, \infty\right)$. A BSS process with power law kernel function has roughness index $\alpha$ and memory properties controlled by $\gamma$. In what follows, we will refer to the BSS process with the power law kernel as the Power-BSS process.
- The gamma kernel is a function $g: (0, \infty) \to \mathbb{R}$ given by $g(x) = x^{\alpha} e^{-\lambda x}$, $x > 0$, $\alpha \in \left(-\frac{1}{2}, \frac{1}{2}\right)$, $\lambda \in (0, \infty)$. A BSS process with gamma kernel function has roughness index $\alpha$ and memory properties controlled by $\lambda$. Below, we will refer to the BSS process with the power law kernel as the Gamma-BSS process.

[9] In general, $\alpha$ takes values in $(-\frac{1}{2}, \infty)$ but, as suggested by the empirical findings of Gatheral et al. (2018), only negative values of $\alpha$ will be relevant for us since these are the ones associated with rough volatility processes.

[10] Indeed, empirical research such as that of Bennedsen et al. (2017) for equity markets, and our own for the FX market, has shown that volatility, especially at intraday frequencies, is not Gaussian. The tails of its distribution are heavier than those of a Normal distribution. In contrast, the Normal-Inverse Gamma (NIG) distribution seems to offer an excellent fit to the increments of log-volatility processes, both at daily and intraday frequencies.

Bennedsen et al. (2017) remarked that the power law and gamma kernels exemplify the theoretical distinction between long and short memory. Specifically, the Gamma-BSS process has short memory, whereas the Power-BSS process will have polynomially decaying ACF and, in particular, the long memory property when $\gamma < 1$. Moreover, although the Gamma-BSS process has short memory in theory, by selecting very small values of $\lambda$ it is possible to specify processes with a very high degree of persistence, mimicking long memory on finite time intervals (as with the RFSV model). Empirically, these two models then allow to assess if there is any gain from using a model with true long memory, as opposed to a highly persistent model with (technically) short memory, particularly in terms of forecasting accuracy.

Finally, Bennedsen et al. (2017) proved that the roughness and memory properties of a Gaussian BSS process for log-volatility carry over the volatility process itself, and conjectured that this also holds for more general BSS processes, under suitable assumptions.

Regarding the estimation of the parameters of the de-coupling models, the roughness index $\alpha$ can be estimated semiparametrically by the OLS regression suggested by (1).[11] The memory parameters $\gamma$, $\lambda$, and $\beta$ of the Power and Gamma-BSS processes and the Cauchy process, respectively, can be estimated by a method-of-moments procedure. Specifically, we can estimate them by fitting the theoretical ACF of each model to the empirical ACF of a log-volatility time series. For this, we use the parametric ACFs of the Cauchy (7) and BSS (9) processes, respectively.[12] As suggested by Bennedsen et al. (2017), in our application we used $L = \lceil n^{1/3} \rceil$ lags in the ACF for this estimation procedure, where $n$ denotes the length of the log-volatility time series whose ACF is being modelled.

In order to forecast log-volatility using the de-coupling models, we use conditional expectation as predictor. Assuming we have $m$ past observations of log-volatility, our formula to forecast log-volatility with the decoupling models is

$$E(\widehat{X_{t+h}|\mathcal{F}_t^X}) = \Upsilon_{1,2}{\Upsilon_2}^{-1} x_{\text{obs}}, \tag{10}$$

where $x_{\text{obs}} = (X_{t-1}, \dots X_{t-m})$ is our vector of observations, and $\Upsilon_{1,2}$ and $\Upsilon_2$ are obtained from the theoretical correlation structure of the process being forecasted, as implied by the estimated parameters. The shape of these matrices is

$$\Upsilon_2 = \begin{pmatrix} \widehat{\rho(0)} & \widehat{\rho(1)} & \cdots & \widehat{\rho(m-1)} \\ \widehat{\rho(1)} & \widehat{\rho(2)} & \cdots & \widehat{\rho(m-2)} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\rho(m-1)} & \widehat{\rho(m-2)} & \dots & \widehat{\rho(0)} \end{pmatrix} \text{ and } \Upsilon_{1,2} = \begin{pmatrix} \widehat{\rho(h+1)} \\ \widehat{\rho(h+2)} \\ \vdots \\ \widehat{\rho(h+m)} \end{pmatrix}.$$

The details on how this formula is derived can be consulted in Bennedsen et al. (2017). Notwithstanding, it is worth noting that they are derived under the assumption that $E(X_t) = 0$ for all $t \in \mathbb{R}$. Thus, in practice we first centre our data, then make use of equation (10), and finally add the subtracted mean to the forecasts to obtain the final log-volatility predictions.

Finally, to obtain forecasts of the squared volatility process itself, we assume Gaussianity and use the formula

$$\widehat{\sigma_{t+h}^2} = E(\exp\widehat{(X_{t+h})}|\mathcal{F}_t^X) = \exp\left(E(\widehat{X_{t+h}|\mathcal{F}_{t-1}^X}) + \frac{1}{2}\text{Var}(\widehat{X_{t+h}|\mathcal{F}_{t-1}^X})\right),$$

where $E(\widehat{X_{t+h}|\mathcal{F}_{t-1}^X})$ is given by (10) and we compute the convexity correction factor as $\text{Var}(\widehat{X_{t+h}|\mathcal{F}_{t-1}^X}) = \widehat{\text{Var}(X_t)}\left(1 - \Upsilon_{1,2}{\Upsilon_2}^{-1}\Upsilon_{2,1}\right)$, with $\Upsilon_{2,1} = \Upsilon_{1,2}^T$. The term $\text{Var}(X_t)$ is estimated through the sample variance of the observations $x_{\text{obs}}$.

*Forecasting evaluation*

[11] This is, by regressing $\log m(q,h)$ against $\log|h|$ for $q = 2$.

[12] In order to numerically estimate $\gamma$ and $\lambda$ in the case of the Power-BSS and Gamma-BSS processes, in practice we use the exact shape of the covariance function for the Gamma and Power kernel functions, which are derived in Bennedsen et al. (2017).

The content of this subsection is based on Pakkanen (2019). Having many possibilities to produce different forecasts of the same quantity, calls for a way to compare these different approaches, so that we can decide which one is best. In the context of volatility forecasting, what we are ultimately trying to forecast is the integrated variance $IV_t$. In practice, we do not, of course, observe $IV_t$, so we need to use a noisy proxy $\widetilde{IV}_t = IV_t + u_t$; a high-frequency estimate such as the realised bi-power variation, say.

The Diebold-Mariano (DM) test (Diebold and Mariano, 2002) is a method geared towards comparisons that attempts to answer whether one forecast series is statistically better than another one. The DM test assumes that we have two forecast series $(X_t^A)_{t=1}^T$ and $(X_t^B)_{t=1}^T$ of the time series $(X_t)_{t=1}^T$. We need to specify a loss function $L: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ measuring the quality of these forecasts. The actual test is then based on loss differentials $d_t = L(X_t, \hat{X}_t^A) - L(X_t, \hat{X}_t^B)$, $t = 1, \dots . T$. The process $d = (d_t)_{t=1}^T$ is assumed to be covariance stationary and equal forecast performance is characterised by the null hypothesis $H_0: E(d_t) = 0, t = 1, \dots, T$.

Under technical conditions restricting the memory of $d$, we have $\bar{d}_T := \frac{1}{\sqrt{T}} \sum_{t=1}^T d_t \xrightarrow{d} N(0, v_d^2)$, as $T \to \infty$, where the asymptotic variance $v_d^2$ can be estimated using a suitable consistent estimator (see Section 1.1 of Diebold and Mariano (2002)). Finally, we have under $H_0$, $DM_T := \frac{\bar{d}_T}{\hat{v}_d^2} \xrightarrow{d} N(0,1)$, as $T \to \infty$, and we can carry out one and two-sided tests of $H_0$ using the test statistic $DM_T$.

The use of a noisy proxy $\widetilde{IV}_t$ instead of $IV_t$ has implications in the context of the DM test. The main concern is whether the loss function $L$ is robust in the sense that $E\left(L(IV_t, \widehat{IV}_t^A)\right) > E\left(L(IV_t, \widehat{IV}_t^B)\right)$ holds if and only if $E\left(L(\widetilde{IV}_t, \widehat{IV}_t^A)\right) > E\left(L(\widetilde{IV}_t, \widehat{IV}_t^B)\right)$. Patton (2011) shows that, under reasonable assumptions on the proxy error $u_t$, the only robust loss functions in the above sense are the squared loss $L(x, \hat{x}) = (x - \hat{x})^2$ and the QLIKE loss $L(x, \hat{x}) = \log \hat{x} + \frac{x}{\hat{x}}$, which are the loss functions we use in our comparisons.

In practice, as discussed in Pakkanen (2019), we are often comparing several (instead of just two) competing models. Then, the DM test would have to be applied several times to do pairwise comparisons. This leads to the multiple comparisons problem, that is, accumulation of Type-I error probability. Hansen et al. (2011) solve this problem by their model confidence set (MCS) algorithm, which takes a set of competing forecasts as an input and outputs a subset of forecasts that contains the best forecast with a given confidence level. The algorithm is sequential, and at each step the current subset of forecasts is tested for $H_0$ of equal forecast performance. If $H_0$ is not rejected, the algorithm halts and outputs the current subset of forecasts, whereas if $H_0$ is rejected, the worst forecast is eliminated from the subset and the above step is repeated until $H_0$ is not rejected and the algorithm halts. Hansen et al. (2011) showed that this construction guarantees that the algorithm does not suffer from the multiple comparisons problem. In our application, we use the MCS method to choose the best volatility forecasting models.

## Conclusions

The empirical evidence presented in our research work allows us to conclude that volatility in FX markets is, indeed, rough. This is true for both intraday and daily volatility, then suggesting that this property is present at all scales relevant in applications (e.g. derivatives pricing or the implementation of electronic trading strategies). More specifically, we presented evidence of four exchange rate volatility time series where the roughness index was found to be negative, and typically around a value of $-0.4$. This is particularly relevant since, as far as we know, other studies of volatility at daily and intraday scales had been conducted exclusively in equity markets, and thus our study extends this finding to the FX market.

Additionally, we found evidence suggesting that the roughness index varies along time, which is consistent with the findings of Bennedsen et al. (2017). Specifically, the higher values of the roughness index seem to coincide with times of market turmoil, implying that volatility tends to be less rough in such periods. Moreover, in our study, periods of low roughness could be shown to coincide with specific events for a given FX pair, such as the release of macroeconomic and political announcements (e.g. the results of the Brexit referendum for the GBP/USD pair), as well as Central Banks' market interventions. Some events, like the results of the 2016 US Presidential election, had an impact on all the FX pairs we studied: all our series exhibited lower roughness and an overall higher volatility near this event. As far as we know, we provided the first evidence of this behaviour of volatility in FX markets.

On the other hand, the results of our forecasting study showed that it pays off to model *both* the rough nature and the memory component of volatility. Indeed, we found the de-coupling models of Bennedsen et al. (2017) to consistently have a better forecasting performance that the RFSV model of Gatheral et al. (2018) even at the daily scale. This contrasts with previous studies, and might be related with volatility being non-Gaussian in the FX market, at all the scales analysed in our study.

Notwithstanding, we also found some simpler autoregressive models able to have a good forecasting performance in the FX market. We believe this success can be explained by a weekly, or more generally 5 to 10 observations, pattern arising in the ACF of high-frequency volatility estimates in FX markets. This then suggests that in order to achieve a better performance when forecasting volatility in FX markets, it might be necessary to further pre-model the high-frequency estimates of volatility (as opposed with equity markets, where there is no evidence of this being necessary). Then, models accounting for roughness and memory components of volatility, as the ones of Bennedsen et al. (2017), might consistently offer the best forecasting performance.

Regarding future research, further exploring the BSS model seems highly promising, given that under this model one could choose the Normal Inverse Gamma as marginal distribution for the log-volatility process, which empirically offers an excellent fit to the increments of this process both in equity and FX markets, at daily and intraday scales.

Moreover, in order to further explore the usefulness of properly modelling both the roughness and memory properties of volatility, it would be interesting to use the forecasts given by rough volatility models in further applications, such as the estimation of market risk measures, the computation of capital requirements, or the implementation of algorithmic trading strategies. In this regard, performing this kind of studies for the USD/MXN exchange rate would be particularly interesting and relevant for the Mexican financial markets and economy, as it could reveal new empirical features of Mexican financial data.[13]

Finally, theoretical research in the rough volatility realm seems highly promising, as rough volatility models have been proven to be more consistent with market data and enjoy a microstructural foundation (El Euch et al, 2016). In fact, an important amount of research is currently under development in this regard as these characteristics have very recently called the attention of both academics and practitioners around the world.[14]

## Main references


1. Barndorff-Nielsen and Shephard (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. Journal of Financial Econometrics, 2(1):1–37, 01 2004. ISSN 1479-8409. doi: 10.1093/jjfinec/nbh001.
2. Barndorff-Nielsen and Schmiegel (2009). Brownian semistationary processes and volatility/intermittency. Advanced financial modelling, 8:1–26, 2009.


[13] Performing this analysis is one of our future objectives. The limitation for this as a student was the lack of a free and reliable source of high-frequency data for this key exchange rate.

[14] For a detailed summary of the recent literature around rough volatility, see https://sites.google.com/site/roughvol/home/risks-1.

3. Corsi (2009). A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics, 7(2):174–196, 2009. ISSN 1479-8409.
4. Andersen and Bollerslev (1997). Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance, 4(2):115 – 158, 1997. ISSN 0927-5398.
5. Mikosch and Starica (2000). Is it really long memory we see in financial returns? In Extremes and Integrated Risk Management, pages 149–168. Risk Books: London, 2000.
6. Gatheral et al. (2018). Volatility is rough. Quantitative Finance, 18(6):1–17, 2018. ISSN 1469-7688.
7. Simity and Milán (2018). Rough Volatility in Foreign Exchange Markets. Corvinus Theses and Outstanding Papers. Available at: http://szd.lib.uni-corvinus.hu/11253/ [Accessed on July 18, 2020].
8. Bennedsen et al. (2017). Decoupling the short- and longterm behavior of stochastic volatility, 2017. Working paper, available at https://arxiv.org/abs/1610.00332v2.
9. BIS Markets Committee (2018). Monitoring of fast-paced electronic markets. Markets Committee Papers, Vol 10. Available at: https://www.bis.org/publ/mktc10.pdf [Accessed on July 18, 2020].
10. El Euch et al, 2016. The microstructural foundations of leverage effect and rough volatility, 2016. Working paper. available at https://arxiv.org/abs/1609.05177 [Accessed on July 18, 2020].
11. Diebold and Mariano (2002). Comparing predictive accuracy. Journal of Business Economic Statistics, 20(1):134–144, 2002. ISSN 0735-0015.
12. Hansen et al. (2011). The model confidence set. Econometrica, 79(2):453–497, 2011. ISSN 00129682.
13. Pakkanen (2019). Advanced Statistical Finance. Lecture notes. Imperial College London, London; United Kingdom, 2019.
14. Rossi and Fantazzini (2015). Long memory and periodicity in intraday volatility. Journal of Financial Econometrics, 13(4):922–961, 2015. ISSN 1479-8409.
15. Boudt et al. (2011). Robust estimation of intraweek periodicity in volatility and jump detection. Journal of Empirical Finance, 18(2):353–367, 2011. ISSN 0927-5398.
16. Rosenbaum (2011). A new microstructure noise index. Quantitative Finance, 11(6):883–899, 2011. ISSN 1469-7688.
17. Heston (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. The Review of Financial Studies, 6(2):327–343, 1993. ISSN 08939454.
18. Patton (2011). Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics, 160(1):246–256, 2011. ISSN 0304-4076.
19. Bayer et al. (2016). Pricing under rough volatility. Quantitative Finance, 16(6):1–18, 2016. ISSN 1469-7688.