Rough Volatility

Elyes Mahjoubi

HEC Lausanne

August 30, 2021

All along the thesis presentation we will study asset's and volatility's dynamic under the probability space $(\Omega,\mathcal{F}_t, \mathbb{Q})$ $\forall$t $\in \mathbb{R}_+$.The SDE to solve for the Black Scholes is :

$$dF_t = \sigma S_t dW_t^{\mathbb{Q}}$$

Solving this SDE leads us to the solution,using the Black model in order to incorporate dividends q ,such :

$$F_T = F_0 e^{\left(-\frac{\sigma^2}{2}T + \sigma W_T^{\mathbb{Q}}\right)}$$

With $\log(F_t) \sim N(-\frac{1}{2}\sigma^2 t, \sigma^2)$ for $F_0 = S_0 e^{drift(T-t)}$

The biggest advantage of the Black & Scholes model are in :

- Its economic tractability
- Its relative easiness and simplicity
- Its allowance to quote options given their implied volatility
- The Greeks that allowed to adopt trading strategies for options

However, even if it is the pioneer model of option pricing.Its drawbacks are huge such the model:

- Doesn't assume leptokurtic distribution (LTCM scandal)
- Doesn't assume a dynamic volatility
- Doesn't assume a negative asymetric skewed distribution
- Doesn't explain the difference of implied volatility between the several strikes

An alternative found in order to fix the precedent distribution flaw was to introduce jumps such the SDE becomes :

$$dF_t = -\lambda t(e^{m+\frac{v}{2}} - 1)\sigma S_{t^-} dW_t^{\mathbb{Q}} + S_{t^-} dJ_t$$

$$dJ_t = \prod_{i=1}^{dN_t} H_i - 1$$

For dN_t,a Poisson process with a probability of λdt and the log-intensity of the jump $\log(H_i) \sim N(m, \nu)$.Solving this SDE , we obtain :

$$F_T = F_0 e^{\left(-\frac{\sigma^2}{2}T + \sigma W_T^{\mathbb{Q},1}\right)} \prod_{i=1}^{N_T} H_i$$

This result into our log-forward contract return distributed as :

$$\log(F_t) \sim N(-\frac{1}{2}\sigma^2 t - \lambda t(e^{m+\frac{\nu}{2}} - 1) + \frac{n}{t}(\frac{\nu}{2} + m), \sigma^2 + n\frac{\nu}{t})$$

The advantage given by the Merton jump model are that the model :

- Doesn't assume asset prices have log-normal distribution
- The volatility is not kept constant because it is affected with jumps
- Doesn't assume a symmetric distribution
- Allows to explain the difference of implied volatility between the several strikes

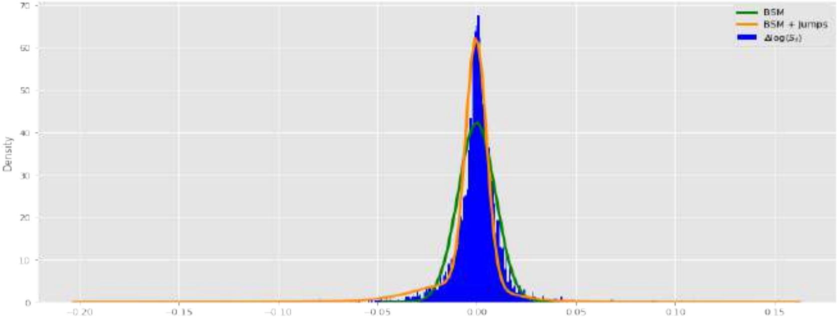

Figure 1: The Merton jump model,an alternative for the Black Sholes model

In the same probability space as before,the risk neutral SDEs are given such :

$$
\begin{aligned}
dF_t &= \sqrt{V_t} F_t dW_t^{\mathbb{Q},1} \\
dV_t &= \kappa\left(\theta - V_t\right) dt + \xi \sqrt{V_t} dW_t^{\mathbb{Q},2} \\
\mathbb{E}(dW_t^{\mathbb{Q},1} dW_t^{\mathbb{Q},2}) &= \rho dt
\end{aligned}
$$

With

- V_t,the spot variance
- ρ,the correlation coefficient between $dW^{\mathbb{Q},2}$ and $dW^{\mathbb{Q},1}$
- θ,the long term variance
- κ,the mean reverting speed parameter
- ξ,the volatility of the variance

The rise and fall of traditional stochastic volatility models

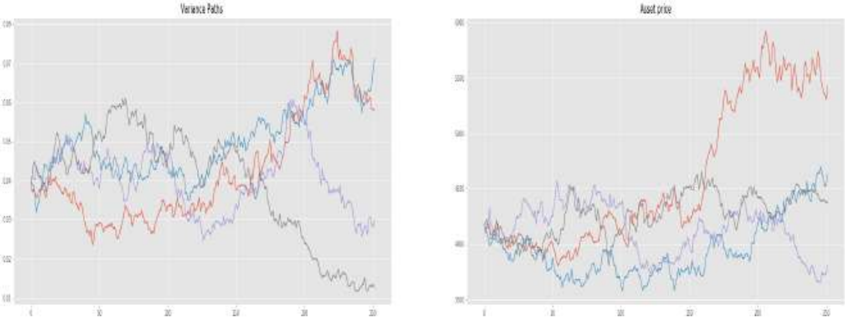

Figure 2: Euleur discretization of the asset price and volatility given the standard Heston model

Under the stochastic volatility models,option prices have cumulative distribution equal to :

$$D_j(x_t, V_t, \tau) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} du \, \mathrm{Re}\left\{ \frac{\exp\left\{C_{SV}(u,\tau)\theta + L_j(u,\tau)V_t + iux_t\right\}}{iu} \right\}$$

Which gives the price of an option when :

$$C_{\Theta} = F_t D_1^{\Theta}(x_t, V_t, \tau) - Ke^{-r(T-t)} D_0^{\Theta}(x_t, V_t, \tau)$$

For Θ a set of parameters $(V_t, S_t, \kappa, \rho, \xi, \theta)$ giving :

- x_t, the log-moneyness such in our case $x_t = \log(\frac{F_t}{K})$
- τ, the time to maturity

An advantage of the stochastic volatility model is that we have now a potentially realistic dynamic that explains the volatility smile and the term structure of options in the market such we can get interesting market insights.

The rise and fall of traditional stochastic volatility models

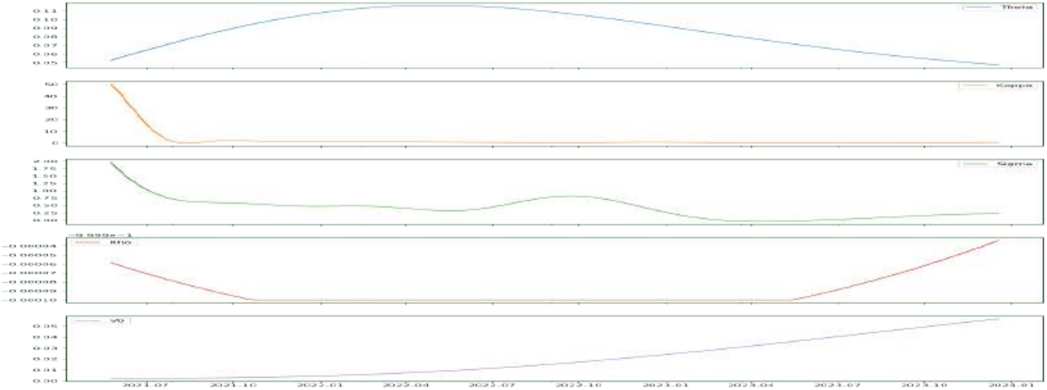

Figure 3: Some market given by the Heston model's calibration

However,this model still has a lot of drawbacks such as :

- Volatility is taken as Normal.Empirically it has shown that it is not.
- Extremely heavy to calibrate because the model contains a lot of parameters,which forces us to use a global optimizer such as the differential algorithm.Moreover,a global optimizer does not guarantee us a certain convergence.
- Very bad fitter,due to the constrains (Feller condition),bounds and the dimension of the minimization problem.Fits are generally not good specially for the very short maturity options.
- The term structure is horrible for very short maturity options.

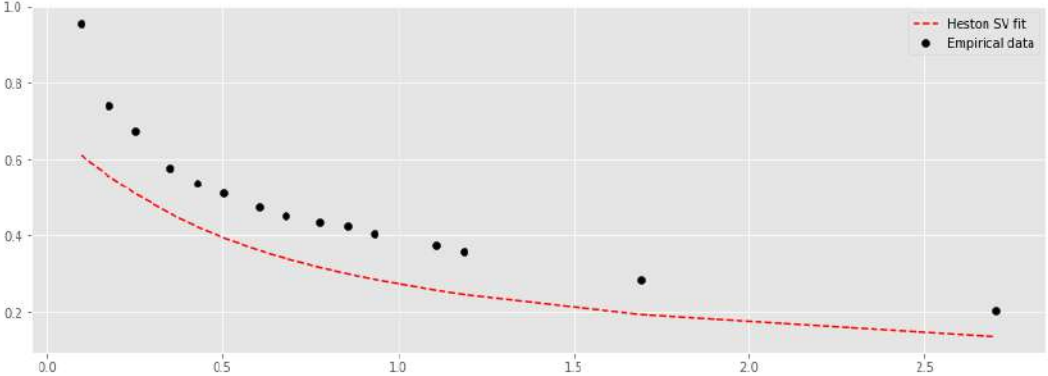

Figure 4: At-the-money empirical term structure and its Heston fit

The optimization has been made using the *IVMSE* function such for the surface :

$$argmin_{\Theta} IVMSE = \sum_{\tau} \sum_{k} \left(\sigma_{IV}^2(\tau, s, k, \Theta) - \sigma_{IV}^{mkt2} \right)$$

And individual maturities :

$$argmin_{\Theta} IVMSE = \sum_{k} \left(\sigma_{IV}^2(\tau, s, k, \Theta) - \sigma_{IV}^{mkt2} \right)$$

However because the convergeance of the algorithm is slow.We added some weights in order to facilitate the optimization such our optimization becomes for the surface :

$$argmin_{\Theta} IVMSE = \frac{1}{\sum_{\tau} \sum_{k} w_{k,\tau}} \sum_{\tau} \sum_{k} w_{k,\tau} \left(\sigma_{IV}^2(\tau, s, k, \Theta) - \sigma_{IV}^{mkt2}(\tau, s, k) \right)$$

For $s = \log S_t$ and $k = \log K$.

We define the weight we use as :

$$w_{k,\tau} = \frac{1}{|spread_{k,\tau}|}$$

We were interested of using the ϑ weighting such :

$$w_{k,\tau} = \vartheta^2_{k,\tau}$$

For ϑ :

$$\vartheta_{k,\tau} = \frac{\partial C(k, s, \tau, \sigma_{k,\tau})}{\partial \sigma_{k,\tau}}$$

However,results were not satisfactory,under-performing our selected weight in the optimization.

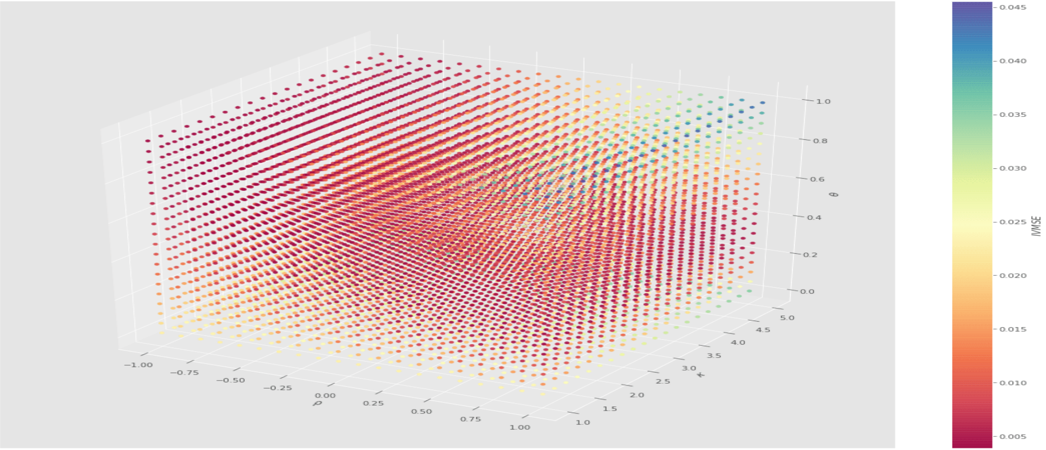

Figure 5: The Heston model converges very hardly

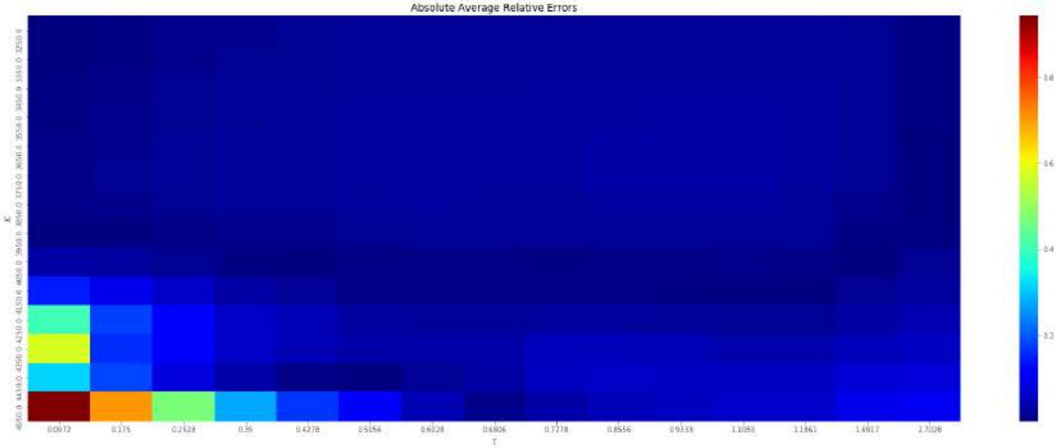

Figure 6: Average absolute relative error among different combinations of strikes and time to maturity of the Heston Fit

J.Bates tried to correct the term structure fit of the simple Heston model by adding a jump.It implies to use new SDEs , in the neutral world for the same probability space as before, such :

$$dF_t = -\lambda t(e^{m+\frac{\nu}{2}} - 1) + \sqrt{V_t} F_{t-} dW_t^{\mathbb{Q},1} + F_{t-} dJ_t$$
$$dV_t = \kappa(\theta - V_t)\, dt + \xi \sqrt{V_t} dW_t^{\mathbb{Q},2}$$
$$dJ_t = \prod_{i=1}^{dN_t} H_i - 1$$

For :

$$E(dW_t^{\mathbb{Q},1} dW_t^{\mathbb{Q},2}) = \rho dt\,,\ \log(H_i) \sim N(m,\nu)\,,\ N_t \sim Poisson(\lambda t)$$

The main advantage resides in :

- The at-the-money term structure fit is better for short maturities
- The distribution of the log-returns is closer to the reality
- The volatility skew fit is better,especially for short maturities

However, it also has its lot of drawbacks such :

- We add more complexity to an already complex optimization.
- The fits are in average worst even if there are some improvement for the close to maturity options.
- It takes more time to calibrate and converges more slowly than previously.As a result,it can't be used in real-time by financial institutions.

	Summary
MSE	669.304
$AARE(in\%)$	6.862
$RMSE$	25.871
$IVMSE$	4.430×10^{-4}
N	210
ρ	-0.878
σ	0.635
θ	0.042
κ	4.745
v_0	0.014
λ	0.024
m	-0.106
ν	0.088

Table 1: SVJ Heston Calibration data

	Summary
MSE	153.992
$AARE(in\%)$	4.648
$RMSE$	12.409
$IVMSE$	3.460×10^{-4}
N	210
ρ	-0.914
σ	0.479
θ	0.061
κ	1.888
v_0	0.023

Table 2: SV Heston Calibration data

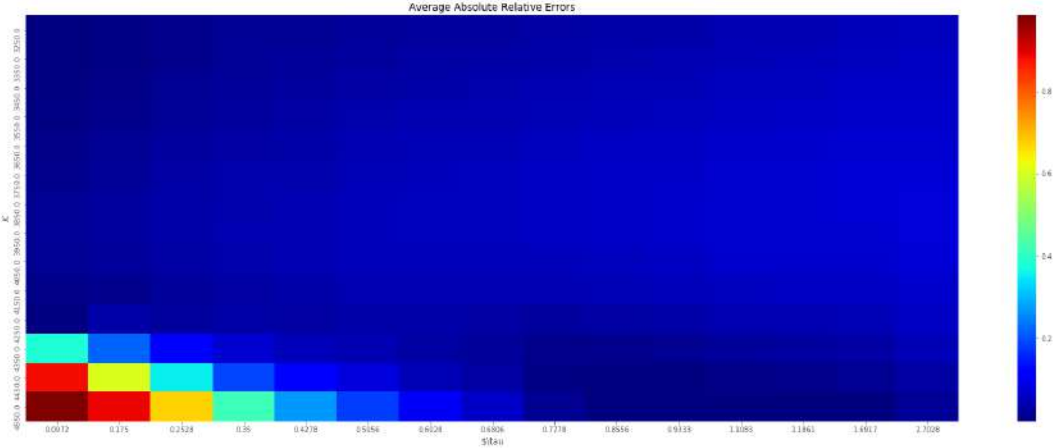

Figure 7: Average absolute relative error among different combinations of strikes and time to maturity of the Bates Fit

The rise and fall of traditional stochastic volatility models

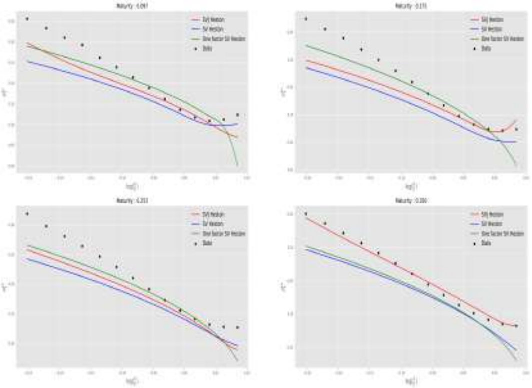

Figure 8: Volatility Skew comparison between the SV models

A parametrization derived by J.Gatheral,using a limit case of the simple SV Heston model in the early twenties using the total implied volatility such :

$$w_{\mathrm{IV}}^{\mathrm{SVI}}(x) = a + b\left(\rho(x-m) + \sqrt{(x-m)^2 + \sigma^2}\right)$$
$$x = \log(\frac{K}{F_t})$$

The originality of this model allows us to remove the multiple kinds of arbitrage other models can be subject to such :

- Calendar spread arbitrage

$$\frac{\partial w}{\partial \tau} \geq 0 \ \forall x \in \mathbb{R}, \forall \tau \in [0, \infty[$$

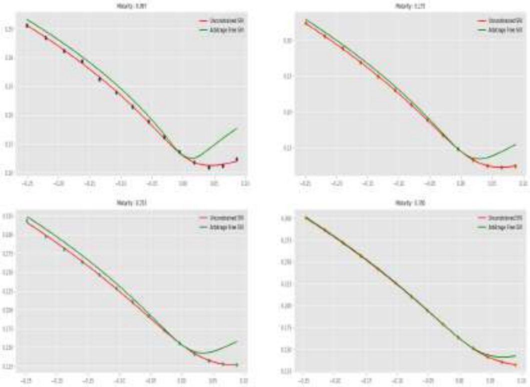

Figure 9: SVI's smile fit

- Butterfly arbitrage

$$g(x) = 1 - \frac{x}{w}\frac{\partial w}{\partial x} + \frac{1}{2}\frac{\partial^2 w}{\partial x^2} + \frac{1}{4}\left(-\frac{1}{4} - \frac{1}{w} + \frac{x^2}{w}\right)\left(\frac{\partial w}{\partial x}\right)^2 \geq 0$$
$$\forall x \in \mathbb{R}, \forall \tau \in [0, \infty[, g : R \to R$$

The results aren't bad in our case,nor extremely good.However,Largely better than what we could obtain by using standard Heston models.Moreover,the calibration,which is Zeliad's white paper, is very efficient and fast enough to implement it in real time trading(Which Goldman Sachs did).

Nonetheless,the dynamic of the process doesn't change which is problematic as,empirically,variance is not an Ornstein-Uhlberg semi-Martingale process.

The rise and fall of traditional stochastic volatility models

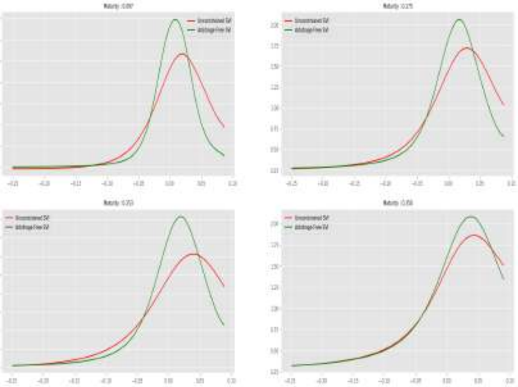

Figure 10: Durleman's condition

The topic we made our thesis on is based on the use of Fractional Brownian motions to model the variance.A fractional Brownian motion is in fact a non-Markovian Brownian motion with a covariance matrix.Defined firstly by Levis as :

$$
\begin{aligned}
W_H(t) = I^{-(H-\frac{1}{2})} f(t) &= \sqrt{2H} \int_0^t (t-s)^{H-\frac{1}{2}} f(s) ds \\
&= \sqrt{2H} \int_0^t (t-s)^{H-\frac{1}{2}} dW_s
\end{aligned}
$$

For W_s,a standard Brownian motion of Hurst parameter $H = \frac{1}{2}$.
Then later defined by Mandelbrot and Van Ness as :

$$
W_t^H = C_H \Big(\int_\infty^0 [(t-s)^{H-1/2} - (-s)^{H-1/2}] \mathrm{d} W_s + \int_0^t (t-s)^{H-1/2} dW_s \Big) \quad \forall t \in \mathbb{R}_+
$$

With :

$$C_H = \left(\int_{-\infty}^{0} \left\{ (t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \right\} dW_s + \int_{0}^{t} \left\{ (t-s)^{H-\frac{1}{2}} \right\} dW_s \right)^{-\frac{1}{2}}$$

By using Ito's isometry,the variance of our fractional Brownian motion becomes equal to :

$$\mathbb{E}((W_t^H)^2) = t^{2H}$$

The fantastic thing about that is at the end of the day we obtain Gaussian increments such $W_t^H \sim N(0, t^{2H})$ with a non-Markovian process.Therefore,the covariance of a fractional Brownian motion between 2 different dates are given such :

$$\text{Cov}\left(W_t^H, W_s^H\right) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), t, s \in \mathbb{R}$$

What does the Hurst parameter change ? :

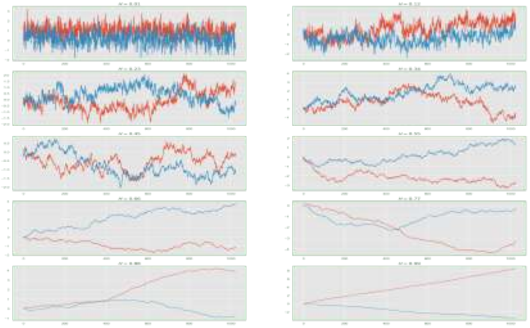

Figure 11: Fractional Brownian motions given different H

- The Hurst parameter affects the 'roughness' of the Fractional Brownian motion

Fractional Brownian motions can be efficiently derived using the fast Fourier algorithm.

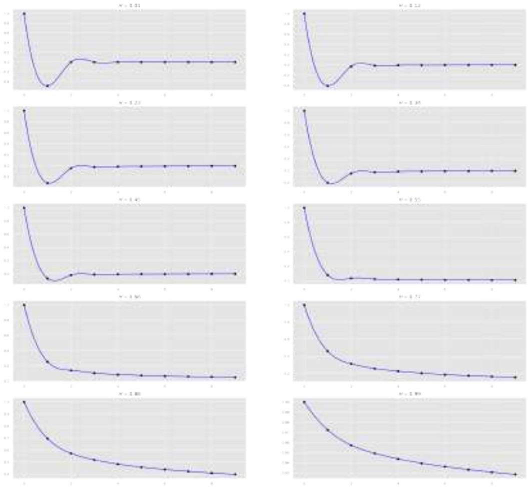

Figure 12: Covariance matrix of a fractional Brownian motion with daily lags

The efficient implementation is done using Z, an $2(n-1) \times L$ standard normal matrix :

$$\mathrm{Re}\left(\sqrt{2(2n-2)} \times \mathrm{ifft}\left(\sqrt{\mathrm{fft}\left(M_{1,:}\right)} \times Z\right)\right)$$

For M, a circulent $(2n-2) \times (2n-2)$ matrix :

$$M = \left(\begin{array}{cccc|cc}
\rho(0) & \rho(1) & \dots & \rho(n-1) & \dots & \rho(1) \\
\rho(1) & \rho(0) & \ddots & \rho(n-2) & \ddots & \rho(2) \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\rho(n-1) & \rho(n-2) & \ddots & \rho(0) & \dots & \rho(n-1) \\
\hline
\rho(n-2) & \rho(n-3) & \dots & \dots & \ddots & \rho(n-2) \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\rho(1) & \rho(2) & \dots & \rho(n-2) & \dots & \rho(0)
\end{array}\right)$$

Fractional Brownian motions

This technique allows us to extract directly Gaussian increments,our most needed thing:

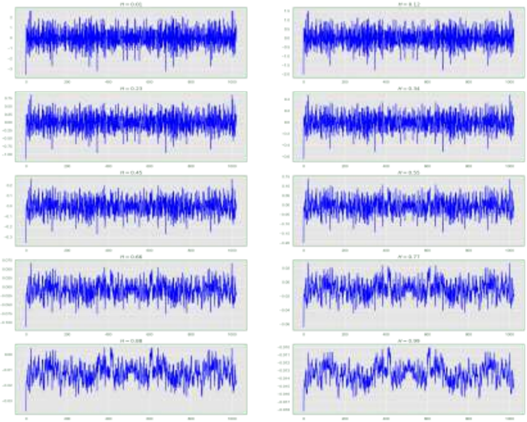

Figure 13: Gaussian noise

The use of fractional Brownian motion to model volatility dynamic isn't from yesterday,but from 1998.its first SDEs are presented such :

$$\frac{dF_t}{F_t} = \sigma_t dW_t^{\mathbb{Q},1}$$
$$d\log\sigma_t = \kappa(\log\sigma_t - \theta)dt + \nu dW_t^{H,\mathbb{Q}}$$

With the solutions :

$$F_T = F_0 e^{-\frac{1}{2}\sigma_T^2 T + \sigma_T W_T^{\mathbb{Q}},1}$$
$$\sigma_T = e^{\theta + e^{-kT}(\log\sigma_0 - \theta) + \nu\int_0^T e^{-\kappa(T-s)} dW_s^{H,\mathbb{Q}}}$$

In the past,volatility was mistakenly thought to be positively correlated with past values "period of volatility follow up".For this reason,the Hurst parameter was chosen to be above $\frac{1}{2}$ on the FSV models.

However,Now it is widely known that the volatility doesn't hold long memory such suggests this graphic :

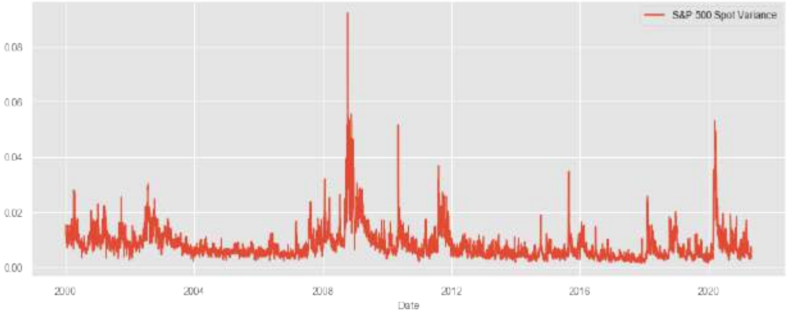

Figure 14: S&P500 spot volatility across time

As a result,an other family of volatility models with short-memory $H < \frac{1}{2}$ and with a mean reverting parameter $\kappa = 0$ was later found by Gatheral, Jaisson and Rosenbaum such we express the volatility as :

$$\sigma_T = \sigma_0 e^{\nu W_T^H} \quad \text{For } W_0^H = 0$$

- This an RFSV process

The Biggest drawback of the FSV resides in its impossibility to correctly fit the volatility emprical log-moments.Nonetheless, RFSV is fitting empirical log-moments very well :

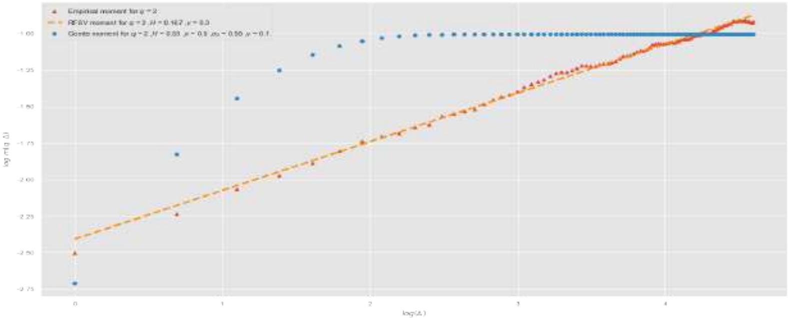

Figure 15: log-moment fit for $q = 2$

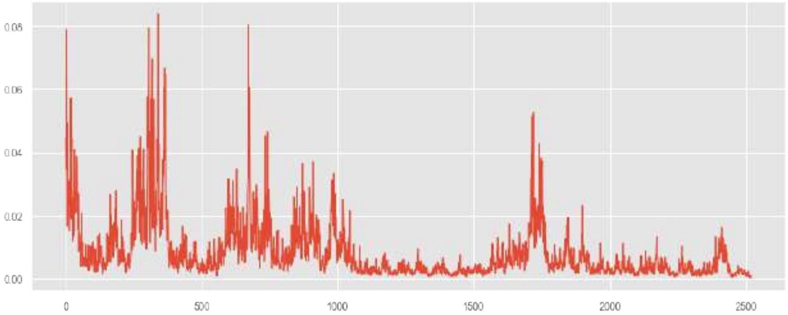

Figure 16: Generated rough Bergomi volatility

But isn't volatility log-normal in reality ?

- The answer is yes,only oil's volatility exhibits fatter tails than the log-normal fit .

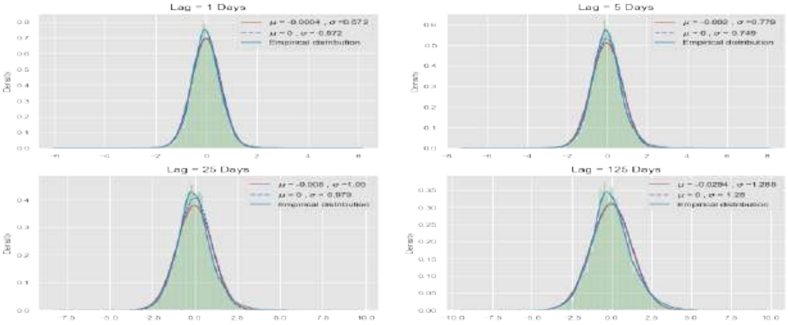

Figure 17: Log-increment volatility distribution for different lags Δ

Multiple ways were found to guess the Hurst parameter H and ν without real success such :

- Fitting the log-increments using the the monofractal scaling relationship such :

$$m(q,\Delta) = \frac{|\log \sigma_{t+\Delta} - \log \sigma_t|^q}{N}$$

$$\zeta_q = qH$$

$$\log m(q,\Delta) \approx \Delta^{\zeta_q}$$

- Or fitting the term structure derived by Fukasawa such :

$$\psi(\tau) \sim A\tau^{-\alpha} \quad \alpha \in (0.3, 0.5) \quad A \in (0,1)$$

Rough Volatility models : The FSV and RFSV approaches

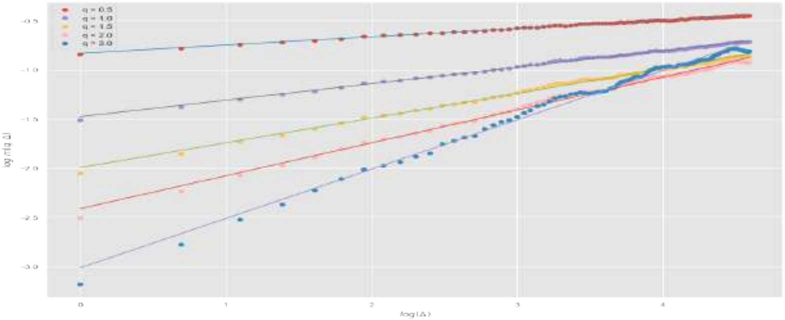

Figure 18: Log-moment of log increment volatility for different lags log(Δ)

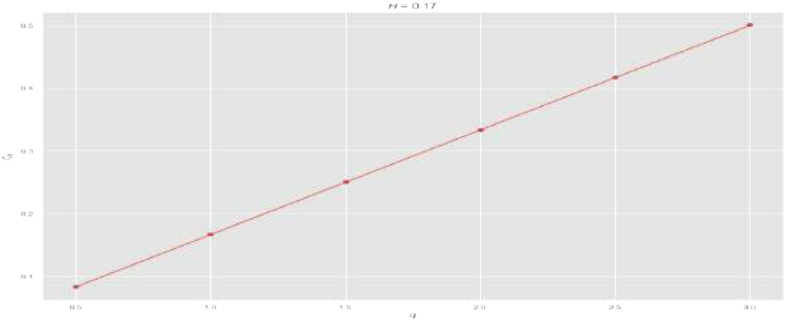

Figure 19: $\zeta_q = qH$

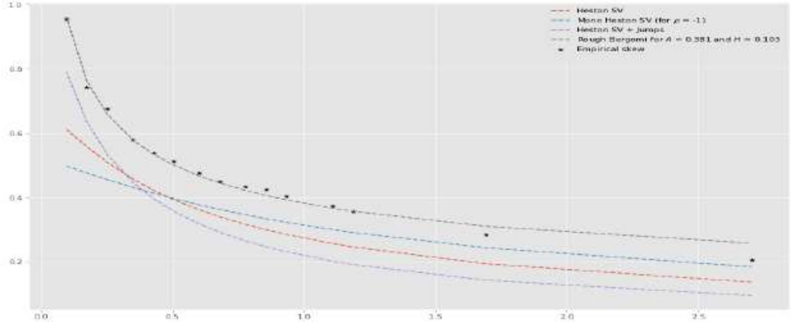

Figure 20: Comparison between the at-the-money term structures of different models

However,these fits are not satisfactory because of the reasons :

- The bias is high because we depend on the realized volatility we computed if we use log-moments.
- The pricing replication using the at-the-money term structure is bad.

Knowing the SDE of the RFSV :

$$dF_t = \sqrt{v_t} F_t dW_t^{\mathbb{Q},1}$$
$$d \log v_t = 2\nu dW_t^H$$

$$\begin{aligned}\log v_u - \log v_t &= 2\nu C_H \left(\int_t^u |u-s|^{H-\frac{1}{2}} dW_s^{\mathbb{Q},2} + \int_{-\infty}^t \left[|u-s|^{H-\frac{1}{2}} - |t-s|^{H-\frac{1}{2}} \right] dW_s^{\mathbb{Q},2} \right) \\ &= 2\nu C_H \left[M_t(u) + Z_t(u) \right]\end{aligned}$$

If we take the expectation,it leads to the expression of the forward variance at time u from time t.

$$\mathbb{E}^{\mathbb{Q}} \left[v_u \mid \mathcal{F}_t \right] = v_t \exp \left\{ \eta Z_t(u) + \frac{1}{2} \eta^2 (u-t)^{2H} \right\} \tag{1}$$

$$\xi_t(u) = \mathbb{E}\left[v_u \mid \mathcal{F}_t\right] \neq \mathbb{E}\left[v_u \mid v_t\right]$$

Because $v_u|\mathcal{F}_t$ is log-normal that :

$$v_u = v_t \exp\left\{\eta \tilde{W}_t^{\mathbb{Q}}(u) + \eta Z_t(u)\right\}$$

For $\eta = \frac{2\nu C_H}{\sqrt{2H}}$. With C_H and $\tilde{W}_t^{\mathbb{Q}}(u)$ being equal to :

$$C_H = \sqrt{\frac{2H\Gamma(3/2 - H)}{\Gamma(H + 1/2)\Gamma(2 - 2H)}}$$

$$\tilde{W}_t^{\mathbb{Q}}(u) = \sqrt{2H} \int_t^u \frac{dW_s^{\mathbb{Q},2}}{(u-s)^\gamma}$$

$$\gamma = \frac{1}{2} - H$$

From 1 and because we know :

$$v_u = v_t \exp\left\{\eta \tilde{W}_t^{\mathbb{Q}}(u) + \eta Z_t(u)\right\}$$

It is straighforward to derive :

$$v_u = \mathbb{E}^{\mathbb{Q}}\left[v_u \mid \mathcal{F}_t\right] \mathcal{E}\left(\eta \tilde{W}_t^{\mathbb{P}}(u)\right)$$

With $\mathcal{E}$,the Wick exponential such :

$$\mathcal{E}(\Psi) = \exp\left(\Psi - \frac{1}{2}\mathbb{E}\left[|\Psi|^2\right]\right)$$

So in order to compute our forward rate $\mathbb{E}^{\mathbb{Q}}\left[v_u \mid \mathcal{F}_t\right]$ we should theoretically know the dynamic of the Brownian motion that is leading the stock even before its creation (As J.Gatheral likes to say we need to go back to the 'Big Bang').

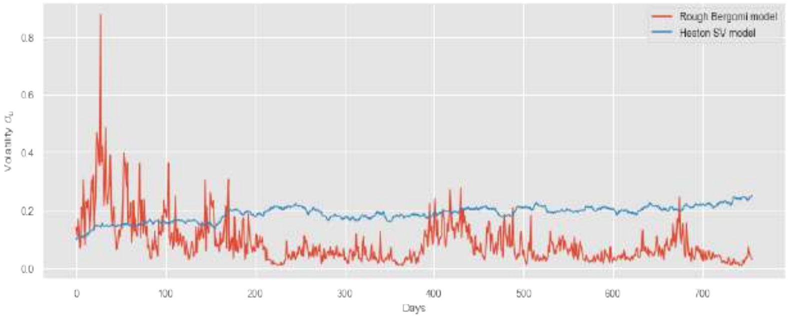

Figure 21: Comparison between two spot variance under the rough Bergomi model and the Heston SV model

Rough Volatility models : The rough Bergomi model

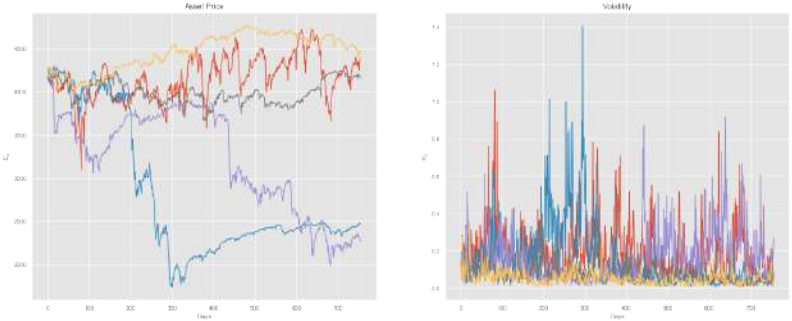

Figure 22: Evolution of the spot asset price S_u and volatility σ_u across time

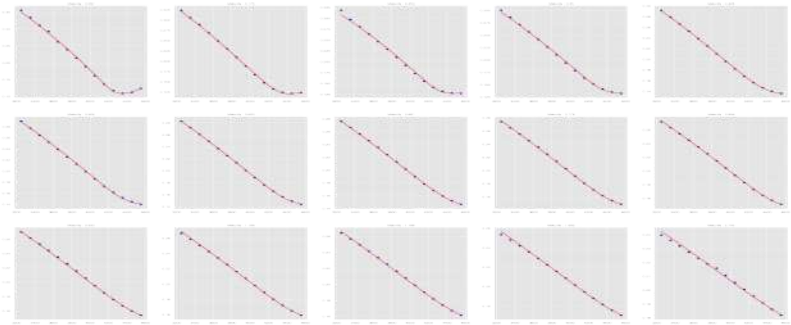

Figure 23: Rough Bergomi's volatility skew

τ	H	η	ρ	ξ_0^T	$IVMSE$
0.097	0.129	2.418	-0.825	0.031	0.249×10^{-5}
0.175	0.230	2.906	-0.722	0.037	0.139×10^{-5}
0.253	0.115	2.359	-0.871	0.047	0.349×10^{-5}
0.350	0.160	2.504	-0.810	0.055	0.150×10^{-5}
0.428	0.210	2.631	-0.754	0.059	0.040×10^{-5}
0.506	0.115	2.257	-0.870	0.062	0.068×10^{-5}
0.603	0.098	2.136	-0.919	0.063	0.025×10^{-5}
0.681	0.075	1.970	-0.996	0.061	0.016×10^{-5}
0.778	0.202	2.393	-0.795	0.076	0.007×10^{-5}
0.856	0.082	1.990	-0.999	0.068	0.021×10^{-5}
0.933	0.190	2.178	-0.826	0.077	0.025×10^{-5}

Table 3: Result of the rough Bergomi model calibration using a differential evolution algorithm

Conclusion

The main advantage of the use of rough volatility models are :

- We don't need to add jumps
- The rough Bergomi model has only 3 parameters : H , ν and ρ,the correlation between the asset price's and variance's Brownian motions.
- Tends to give a very good fit and especially for very short term maturities.
- Tends to give a very good fit.Especially true for very short term maturities.

However its main drawbacks are that :

- We need to use a global optimiser that takes a lot of time to converge (or not).
- The parameters are not stable across time .Hard to construct a volatility surface.
- A closed or semi-closed formula doesn't exist.

An open discussion consists of incorporating deep learning into the calibration process in order to calibrate the model more rapidly.