Smarter than the Options-Market?
A Real-Measure GARCH Option Pricing Model
with Volatility Regime Simulation

Chrilly Donninger
Chief Scientist, Sibyl-Project
Sibyl-Working-Paper, March 2014

http://www.godotfinance.com/

In economics, the majority is always wrong.
(John Kenneth Galbraith)

Abstract:

This working paper uses as a starting point the filtered historical simulation (FHS) approach developed by Barone-Adesi et al. ([1],[2]). One builds a GRJ-GARCH model and generates Monte-Carlo return/price paths with normalized returns. This introduces a severe drift-bias. The Volatility Regime Simulation (VRS) avoids the bias by sampling from the same volatility regime.
Barone-Adesi et al. transform the real-world into the risk-neutral measure. They calibrate the GARCH model to the market prices of plain-vanilla options.
The current model stays in the real-measure. One simulates a realistic trading behavior by hedging the options along the Monte-Carlo paths. The model generates the stylized facts of S&P-500 index options. The overall agreement with market-prices is quite good. According the model Calls are somewhat under-, Puts are somewhat overpriced.
The second part of the paper demonstrates the promising application of the model for index options trading.

The GARCH-Model:

The purpose of GARCH is to model volatility. Volatility is mean-reverting and persistent. Everybody in finance knows what volatility is. But nobody knows how to measure it. The traditional proxy is the squared daily return. But the squared-returns are not persistent. Even in a high-volatility regime there can be days with a low squared return. On the other hand there can be spikes in low volatility phases. One can interpret volatility-models as smoothers or filters of the proxy. As the true value of volatility is not known, there is no clear-cut criterion for selecting a model. This is a very fortunate state for the academic world. Thousands of papers have been written about the virtues of dozens of different GARCH -variants. Following [1] I implemented the GRJ-GARCH model of [4]. GRJ is a shortcut of the authors (**G**losten**, J**agannathhan, **R**unkle). This is one of the more popular GARCH models.

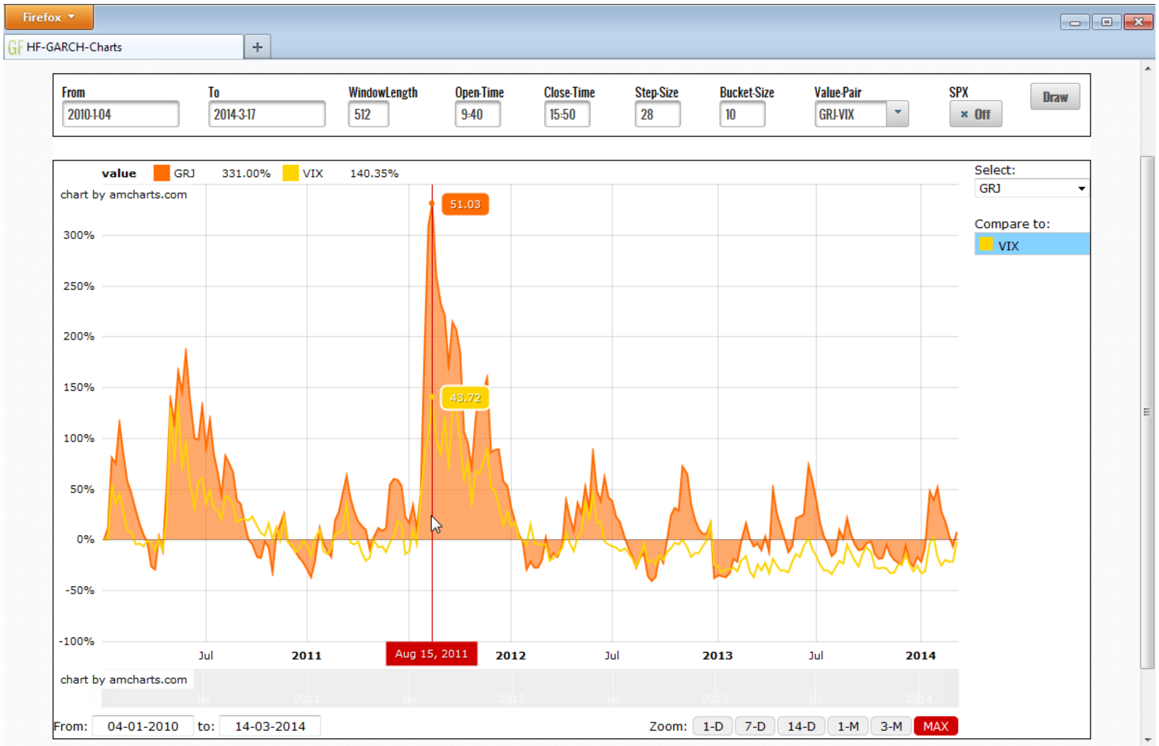
```
σ^2(t) = a0 + a1*σ^2(t-1) + a2*R^2(t) + a3*I(t)*R^2(t)
I(t) = (R(t)<0)
```

a0 is a sort of ground-level. In combination with the other parameters it is also responsible for the mean-reverting behavior. a1 is the persistence or smoothing factor. The value is typically between 0.8 and 0.95. a2 models the impact of the squared daily

return. This factor is typically negative. a3 is the impact of negative returns. The indicator I is zero for positive returns and 1 for negative. a3 is positive and larger than the absolute value of a2. a2 and a3 model the different impact of negative and positive returns. The plain GARCH model has only the parameters a0, a1, a2. Negative and positive returns have the same effect on volatility. This is reasonable for FX, but not for stocks and stock-indexes like the S&P-500. There are several other approaches to model the asymmetric behavior of stock-indexes. GRJ-GARCH is the simplest one.
The main drawback of GRJ is: If a2 is negative (and a2 is in most cases negative), the expression on the left side can become negative. This is clearly impossible. The parameter-estimation crashed into this problem in the recovery phase after the flash-crash. The flash-crash was too short to drive volatility to a very high level. The large recovery following the crash made the overall result therefore negative. As a simple fix the implementation clamps the value to the minimum of 0.8*VIX in the historic time window. This time window was for all reported results 512 plus a burn-in phase of 30 trading-days. The start value is set to 0.8*VIX on the corresponding day. But the first 30 days are not used for the GARCH parameter estimation and for VRS. The factor 0.8 was chosen, because GARCH-volatility is usually somewhat smaller than the VIX. The VIX contains a risk-premium.
Note: To convert from VIX to σ one has to scale by 1.0/(sqrt(252.0)*100.0)
If there is no implied volatility-measure like the VIX available, one could instead use plain realized volatility. The purpose is to have a reasonable starting value and a lower-bound.

Its well known, that estimating the GARCH-parameters is a nasty numerical problem [5]. There are many combinations of the parameters which result in a similar filter behavior. Hence the ML-function has many local minima and is in the solution area very flat. Although the Amoeba optimization method of Press et al. [6] is rather robust (and slow), I had in a previous study [3] convergence problems (Amoeba is generally known as the downhill simplex method due to Nelder and Mead).
Winker&Maringer solve in [5] the problem with their favorite method threshold-accepting. My favorite is Differential-Evolution (DE). The number of individuals was set to 16. The number of generations to 200. DE generated always slightly better results than Amoeba and has no convergence problems. DE is – like every meta-heuristic – slower than Amoeba. But optimizing 4 parameters with a meta-heuristic is on modern processors no issue.

Graphic-1 shows the GRJ-GARCH volatility and the VIX from 2010-01-04 till 2014-03-14. To get the same scale σ is multiplied by sqrt(252)*100.0. The general level of σ is lower than the VIX. But the GARCH volatility reacts faster and stronger. This can be best seen in the August 2011 crash (red line in Graphic-1). On 2011-08-15 σ was 51.0, the VIX at 43.7. The peak on the left is the flash crash.
The GARCH parameter are re-estimated every 20 trading-days. The parameters change over time. But a different re-estimation frequency has only a minor influence. The ML-function is - as already noted - very flat. Different parameters generate usually a similar overall behavior.



Graphic-1: GRJ-GARCH (orange) and VIX (yellow) 2010-01-04 till 2014-03-14

Filtered Historical Simulation:

Historical simulation is the simple idea: One samples from the past N returns randomly and generates in this way Monte-Carlo paths. The method is – besides the window-length – parameter free. One avoids the nasty question about the statistical distribution of the returns. But this naive form destroys the volatility-clustering. One stacks returns from different volatility-regimes together. Filtered historical simulation addresses this point. One first normalizes the returns by dividing with the GARCH-volatility. Normalized returns are also much closer to a normal-distribution than the returns itself. One could interpret the market behavior as a brownian motion with a time-varying temperature of the liquid. The invisible hand turns the Bunsen-burner on and of**f.**

```
nR(t) = R(t)/σ(t)
```

For path generation one picks randomly a normalized return and multiplies with the current volatility. The generated return is plugged in the GRJ-GARCH formula and a new σ is calculated.

```
R(T) = nR(t)*σ(T)
```

There is one problem with this approach. The mean of the returns aka the drift is weakly dependent on the volatility regime. If one picks in a high-volatility phase a normalized return from a low-volatility-regime, one has a systematic wrong drift. FHS generates in high-volatility-regimes in the mean a strong bull market. One can argue that the drift is anyway hedged away. This is – in a realistic setting – only partly true. But another shortcoming of FHS creates even in a complete market a significant bias. The frequency of the historic positive returns is almost always higher than the negative ones. The market moves up in many small steps and goes done in a few larger ones. Hence FHS samples in high-volatility-regimes the wrong ratio of up- and down-moves. GRJ-GARCH reacts by construction different to negative and positive returns. The combination of FHS and GRJ-GARCH is in contrast to the claims in [1] and [2] not well suited.

Volatility Regime Simulation

Volatility Regime Simulation (VRS) tackles the problems of FHS by sampling from the same historic regime. One possible solution is to define volatility buckets. Each historic return is classified according its GRJ-Volatility. One could define 3 buckets with σ below 13%, between 13 and 20 and above 20%. If the current σ is 15, one samples only from the mid-bucket, if it is above 20, one takes a historic return from the high-volatility regime. The bucket approach is simple to implement, but there is the problem of defining appropriate bucket-boundaries. There are also significant boundary effects. If the current σ is 12.8%, one starts the MC-path from the low-volatility regime. The chance of staying in this regime is considerable higher than for a starting value of 13.1%. One has to add noise to blur the boundary effect. The logical consequence of this is to avoid the boundaries at all. For the current implementation of VRS the returns are sorted according their corresponding σ. One selects from the historic sorted returns the entry with the closest σ. Then one adds some noise to the index and selects the corresponding return. Adding noise is necessary. Otherwise the paths would be identical. There is also in real-live a market-external "noise". The exact calculation is:

```
[1] nInd(t) = findNearestIndex(σ(t-1))
[2] bsz = max(min(N-nInd(t),nInd(t))/2,10);
[3] ind(t) = nInd(t)+ NV(0,1)*bsz;
[4] R(t) = sortedReturns[ind(t)].
N is number of historic returns.
```

One calculates in step [2] the amplitude of the added noise. The value depends on the current index. If the current σ is around the median of the historic volatility, the added noise is larger than on the boundaries. The calculated index in step [3] can be below 0 or above N-1. There are several strategies to deal with this over- or underflow. One can clamp a negative value to 0, a large value to N-1. One can reflect the index. If ind(t) is -10, one would set it to +10. The same reflection happens at the upside. Or one can repeat step [3] till ind(t) is in the valid range [0,N-1].

The first method – clamping – has the effect that one selects with a relative high-frequency the extreme values. This is an unrealistic behavior. The effect is most noticeable on the high-side. This entry is a crash-day. Hence R is highly negative and σ explodes. In the next step one selects with probability 0.5 again the crash entry. The largest entry becomes an absorbing state.

The reflection and repeat method are clearly better. There is no clear cut advantage for one over the other. The current implementation uses the repeat approach.

The VRS has one drawback with the plain historical simulation in common. The range of returns is restricted to the returns in the historic window (the last 2 years). This is a critical point for VaR calculations. One can address this problem by adding at position N an especially large crash-day. The current option model uses only the historic sample. In contrast to VaR one wants to model average behavior and not a worst-case scenario.

Options Calculation:

VRS generates Monte-Carlo Paths for the returns and the corresponding σ. One trades now for each path the option. The option is delta-hedged by the underlying. The delta is calculated with Black-Scholes. In contrast to real-life trading one can't use the market-implied volatility for the delta-calculation. Instead the GARCH σ is used. This gives on average a somewhat smaller delta (see also Graphic-1). One could think about more sophisticated calculations. Calculate in the first iteration with the GARCH σ and use in subsequent iterations the previous value. Determine from the value the implied-volatility and the corresponding delta. This approach was not examined in detail. The calculation is already for the non-iterative approach relative time-consuming. Graphic-2 shows a scatter plot of market IV (x-axis) versus the model-IV for OTM Puts with a delta of -0.15. The maturity is from 22 to 31 trading days. The Puts were traded from 2010-01-03 till 2012-12-12.

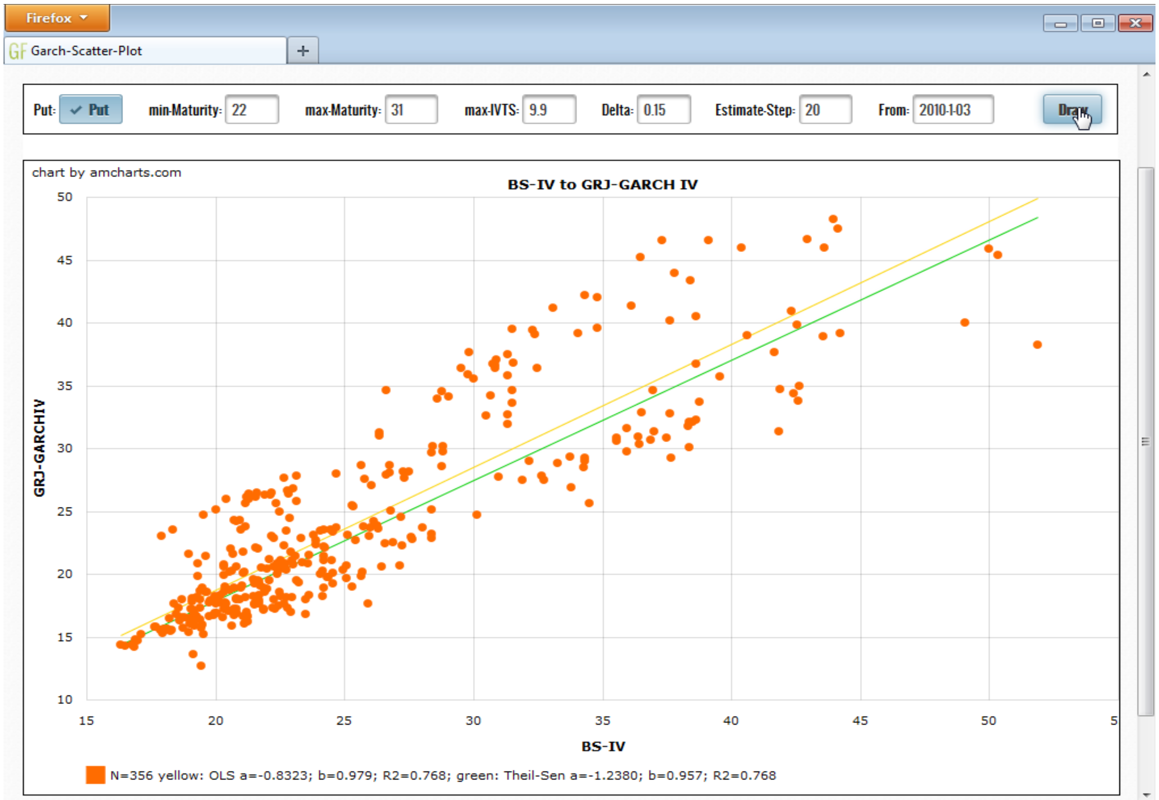
Note: I have bought for a previous study a database of complete options values. This database (unfortunately) ends at 2012-12-12 (see also [7],[8]).

Graphic-3 shows the same calculation for OTM Calls with a delta of 0.15. Generally the fit is for Puts better than for Calls. The scatter-plots look for other delta values similar. The yellow line is an OLS-, the green line a robust Theil-Sen-Regression (see [9]). The deviation of OLS to Theil-Sen indicates the existence of outliers.

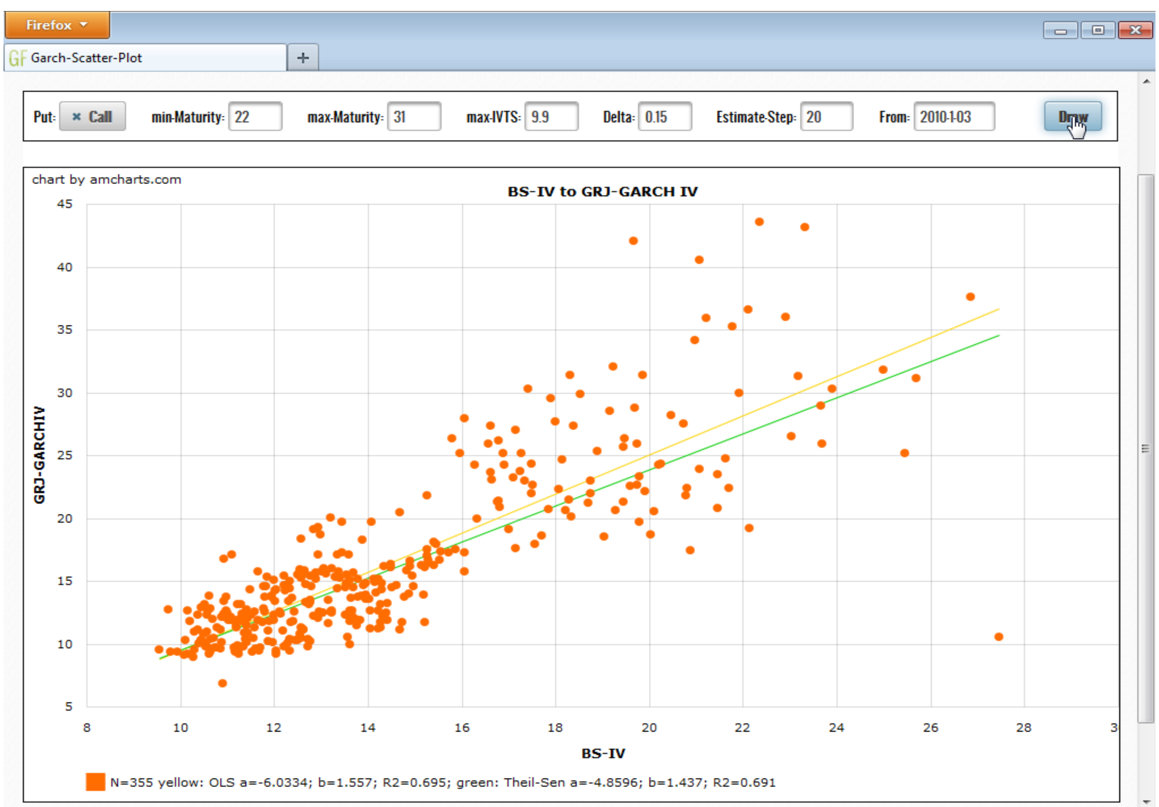
Option	Mean Market	Mean Model	Median Market	Median Model
Put Delta -0.15	25.70	24.33	22.90	21.68
Put Delta -0.25	23.17	22.48	20.82	19.96
Put Delta -0.35	21.51	21.28	19.36	18.99
Put Delta -0.45	19.78	19.88	17.95	17.89
Call Delta 0.45	16.90	18.91	15.25	16.96
Call Delta 0.35	15.73	17.85	14.38	15.76
Call Delta 0.25	15.01	17.15	13.70	15.08
Call Delta 0.15	14.40	16.33	13.34	14.32

Table-1: Mean and Median Market- and Model-IV.

Table-1 shows the mean and median for the market- and the model-IV for different strikes (delta). The model IV (and hence the price) is for Puts somewhat lower, for Calls higher than the market values. The trades are done on the same trading-days. Table-1 shows therefore also the mean/median-smile. The market smile is from 25.70/22.90 for OTM Puts down to 14.40/13.34 for OTM Calls. The model-smile is from 24.33/22.90 down to 16.33/14.32. The model-smile is in relative good agreement with the market.



Graphic-2: Market-IV versus Model-IV for OTM Puts

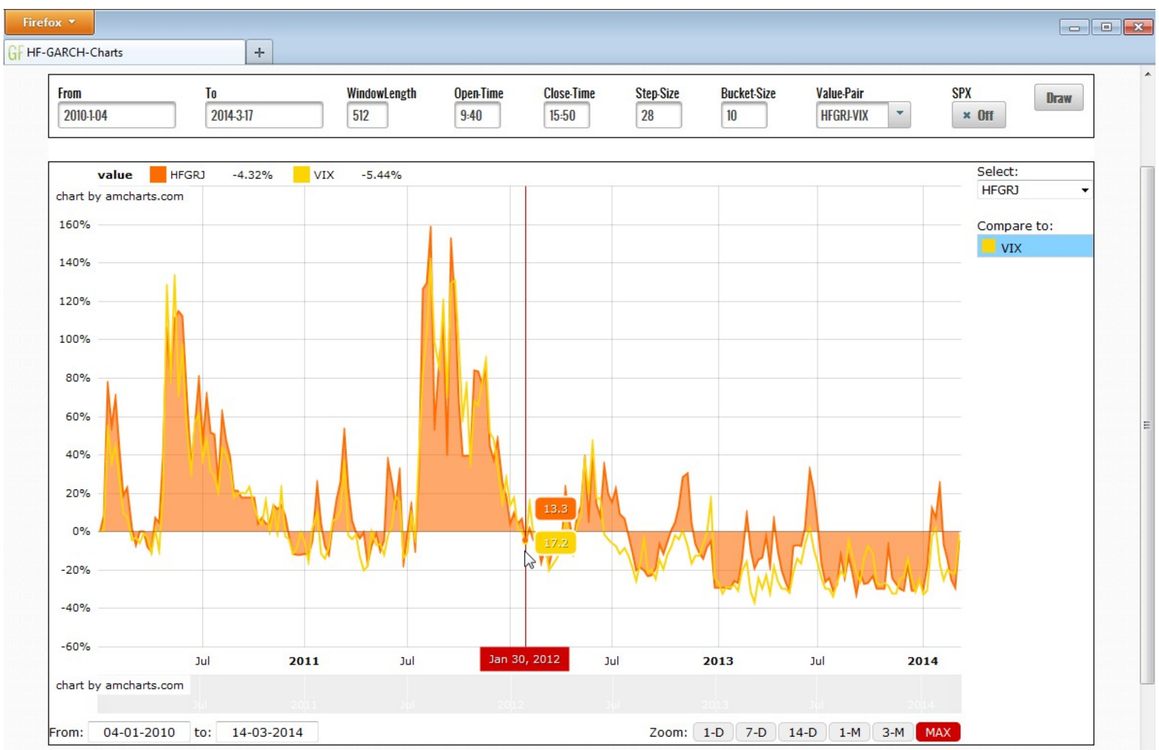


Graphic-3: Market-IV versus Model-IV for OTM Calls

High-Frequency GRJ-GARCH:

The GRJ-GARCH model generates only end-of-day values. In [7] I analyzed the effect of trading and hedging after the open or before the close. For the first case, called the Lark-strategy, one can/does not use the latest overnight return information. The High-Frequency GRJ-GARCH takes this into account. One splits the daily return up into an overnight- and an intraday-part. The GRJ-GARCH formula and the parameter estimation is the same. One has just the double number of observations. For comparisons with the VIX one has now to scale by sqrt(2*252.0)*100.0. The two returns have not the same distribution. The intraday-return is usually somewhat larger. The overnight-return has a larger skew. I tried to extend the GRJ-GARCH model by another factor with an indicator function of 1 for intraday- and 0 for overnight-returns. But this does not improve the ML-function significantly. The weight is also quite unstable. It changes almost randomly from one estimation to the next. The sign is sometimes negative, sometimes positive. So I dropped this term again.

The HF-GRJ-GARCH is smoother than the model with daily-returns. The overall level is also somewhat higher. One could extend the idea by considering even smaller time-steps. For VRS one samples the overnight-step from the overnight and the intraday-step from the intraday-returns. Although the GARCH behavior differs, the option evaluation is similar.



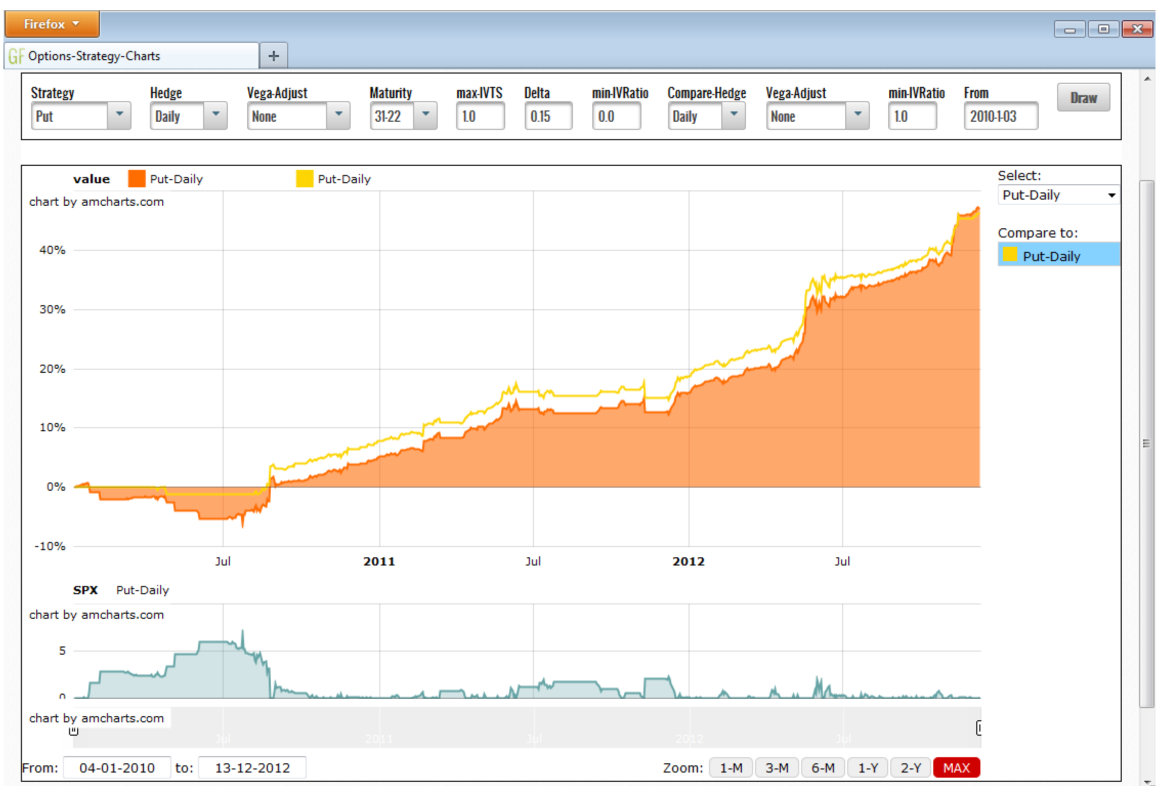
Graphic-4: HF-GRJ-GARCH (orange) and VIX (yellow) 2010-01-04 till 2014-03-14

Improving the Trade-Entry:

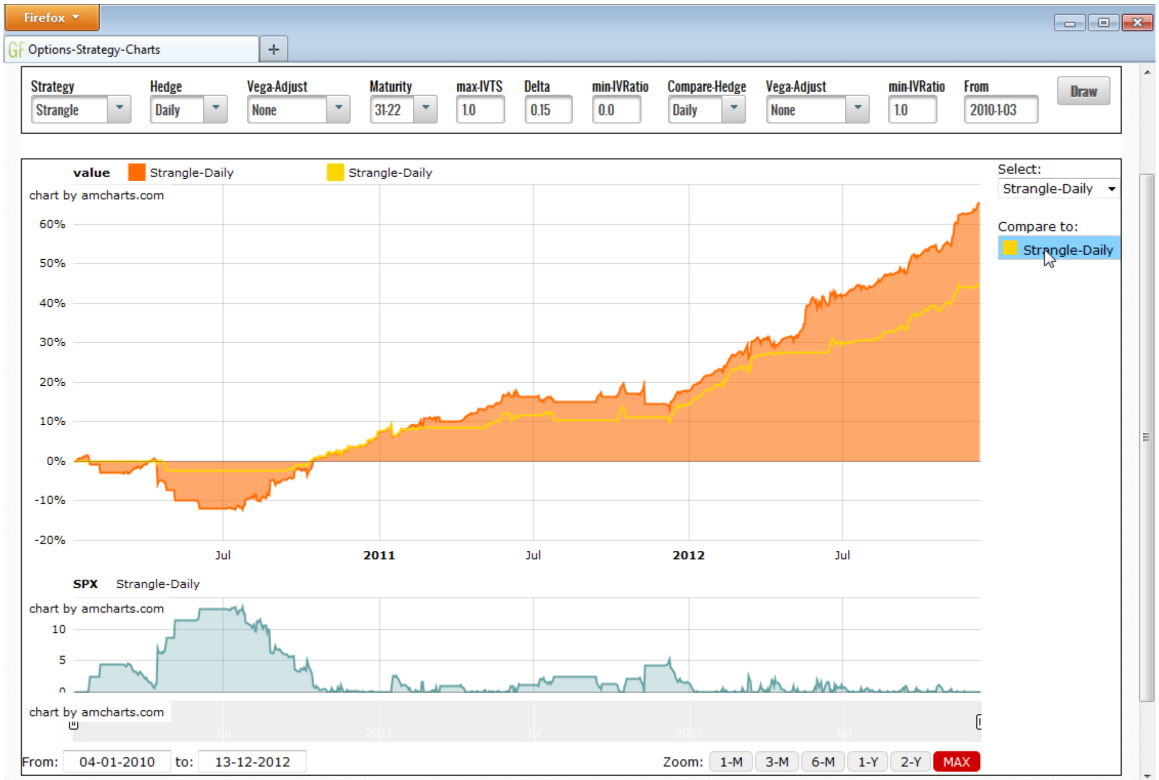
A fundamental question in options trading is to find a good entry point for the trade. In [7],[8] the Kir-criterion was developed. One goes on the sideline if the Implied-Volatility-Term-Structure (IVTS) is above 1.0. The IVTS is defined as:

```
IVTS(t) = VIX(t)/VXV(t)
```

It is the ratio of the 1-month and the 3-month implied volatility. As an immediate consequence one does not enter a position if the IVTS is above 1.0. Options selling is too dangerous in times of troubles. But otherwise the rules in [7],[8] provide no further hint for the entry problem.
One application of the GRJ-GARCH model is to answer the question, if an option is cheap or expensive. One sells only expensive options. The market-IV (price) must be greater than the model-value.
One of the interesting strategies in [7],[8] was an OTM-Put with a delta of -0.15 and a maturity between 31 to 22 trading days. Starting at 31 trading days before maturity, one enters the position the first time the Kir (and GARCH) criterion is met. Trading is done from 2010-01-04 till 2012-12-12.
Graphic-5 shows the result. The GARCH-criterion stays from May to August 2010 on the sideline (left part of chart). This is a correct decision. But it lags behind in the last phase on the right. The position is not avoided at all, but entered somewhat later. It misses a part of the risk-premium.



Graphc-5: OTM-Put with Kir (orange) and Kir+GARCH (yellow).



Graphc-6: OTM-Strangle with Kir (orange) and Kir+GARCH (yellow).

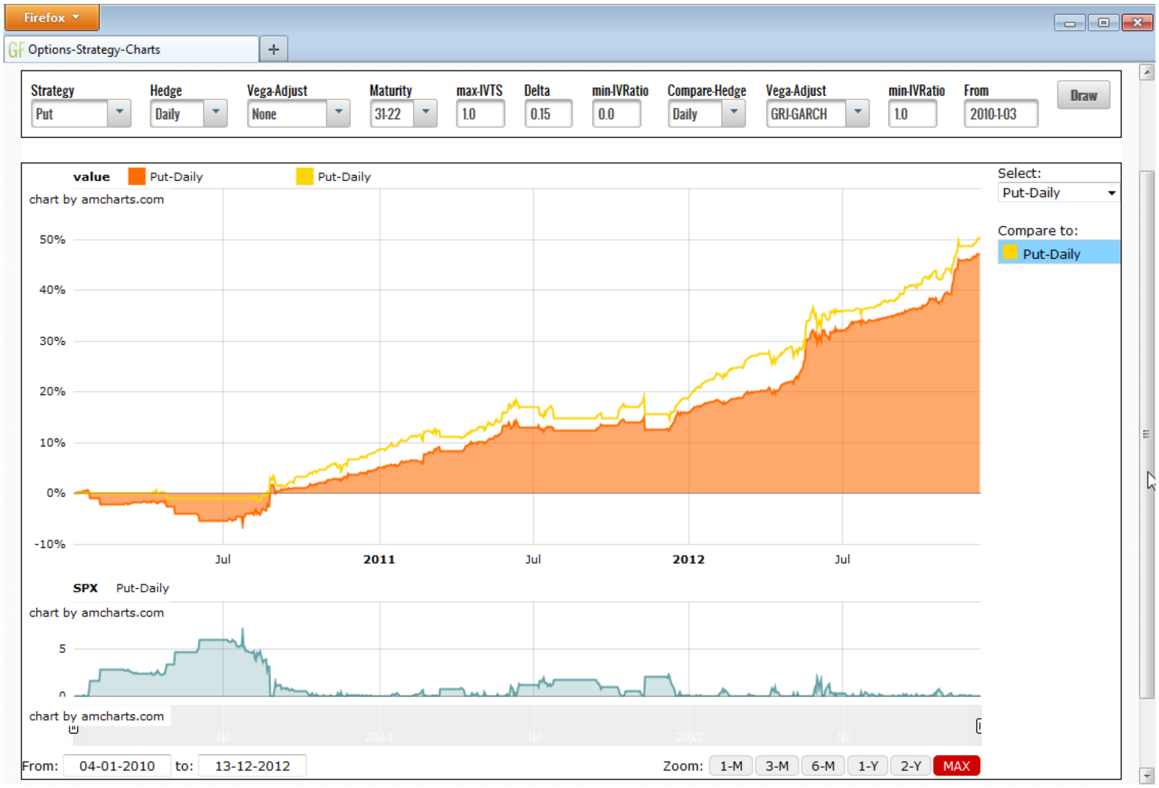
Graphic-6 shows the performance of an OTM-Strangle. The Put and the Call have a delta of +/- 0.15. Both sides must be expensive (all strategies sell the options). This criterion is quite restrictive. The model evaluates Calls generally somewhat higher than the market. The GARCH filter avoids any problems, but misses also profitable market phases.

GRJ-GARCH-Hedging:

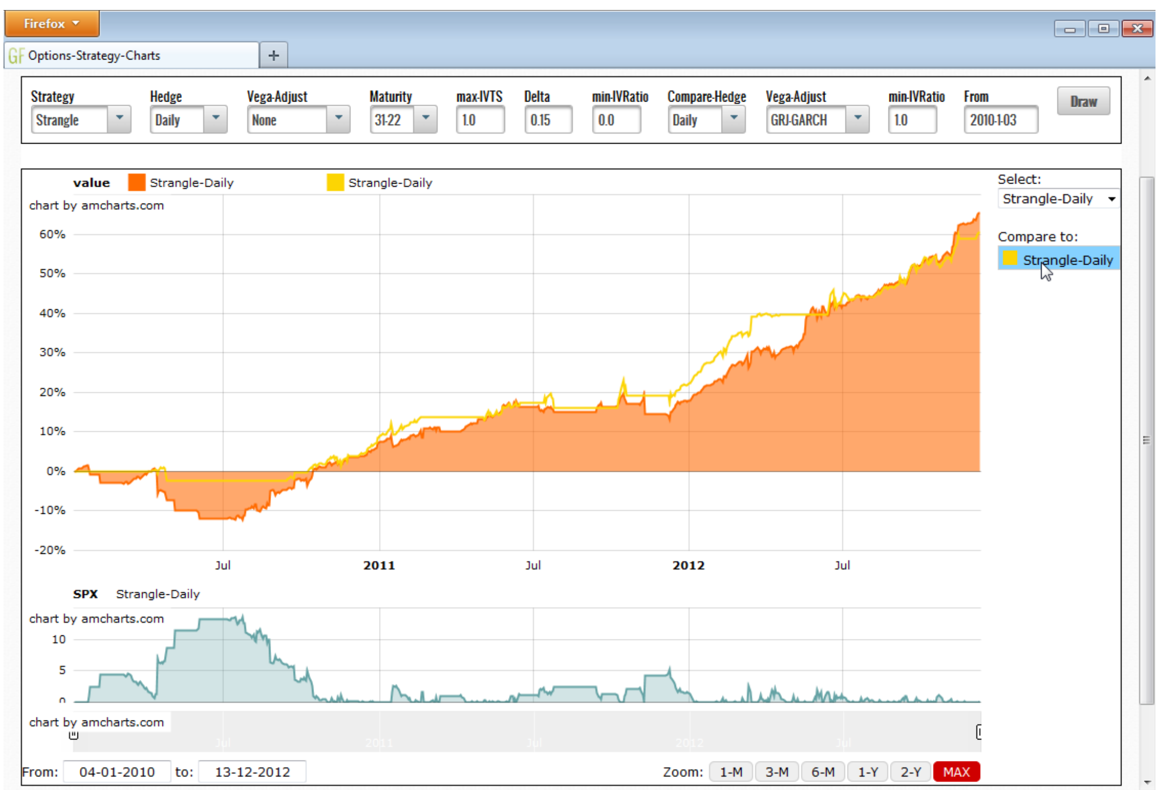
One can use the GARCH-model to calculate the options-delta. One calculates the option with the current S&P-value and starts the calculation again with a higher (Call) or lower (Put) price. The calculation is done on the same MC-paths. The S&P is changed by +/- 5.0 for the second calculation. The result is relative insensitive to different offsets.

```
Delta = (Value(S&P+Offset)- Value(S&P))/Offset.
```

Instead of the market-implied delta one hedges the position with the model-delta. Graphic-7 shows the same calculation than Graphic-5. But this time hedging is done for the Kir+GARCH strategy with the model delta. The model delta is usually somewhat smaller than the market implied value. So the Kir+GARCH stays ahead of the plain Kir. But the difference is with an overall performance of 47.1% to 50.3% not very large. Kir+GARCH misses some profits. This is compensated by the lower hedging costs.



Graphic-7: OTM-Put with Kir (orange) and Kir+GARCH+model-Delta (yellow).

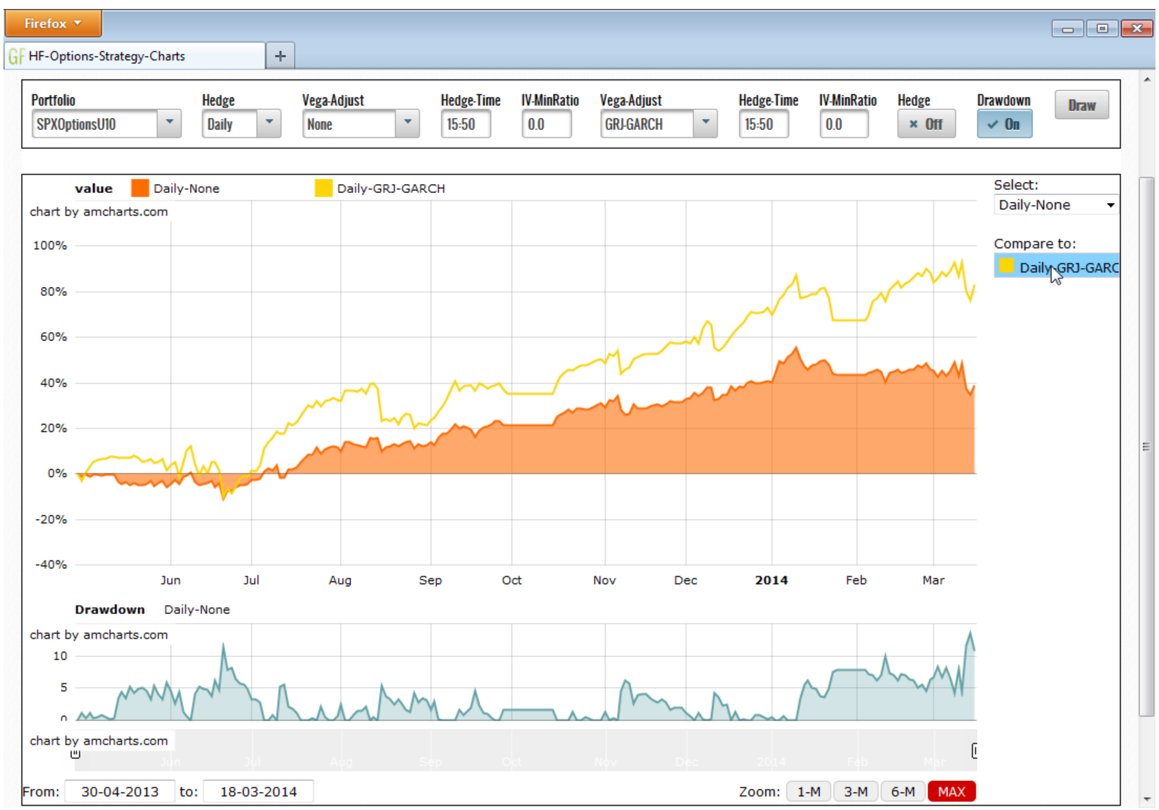


Graphic-8: OTM-Strangle with Kir (orange) and Kir+GARCH+model-Delta (yellow).

Graphic-8 shows the performance of the OTM-Strangle from Graphic-6. Kir+GARCH is with model-hedging almost on par with the plain Kir-strategy. The performance has increased from 44.9 to 60.3%.

HF-GRJ-GARCH-Hedging:

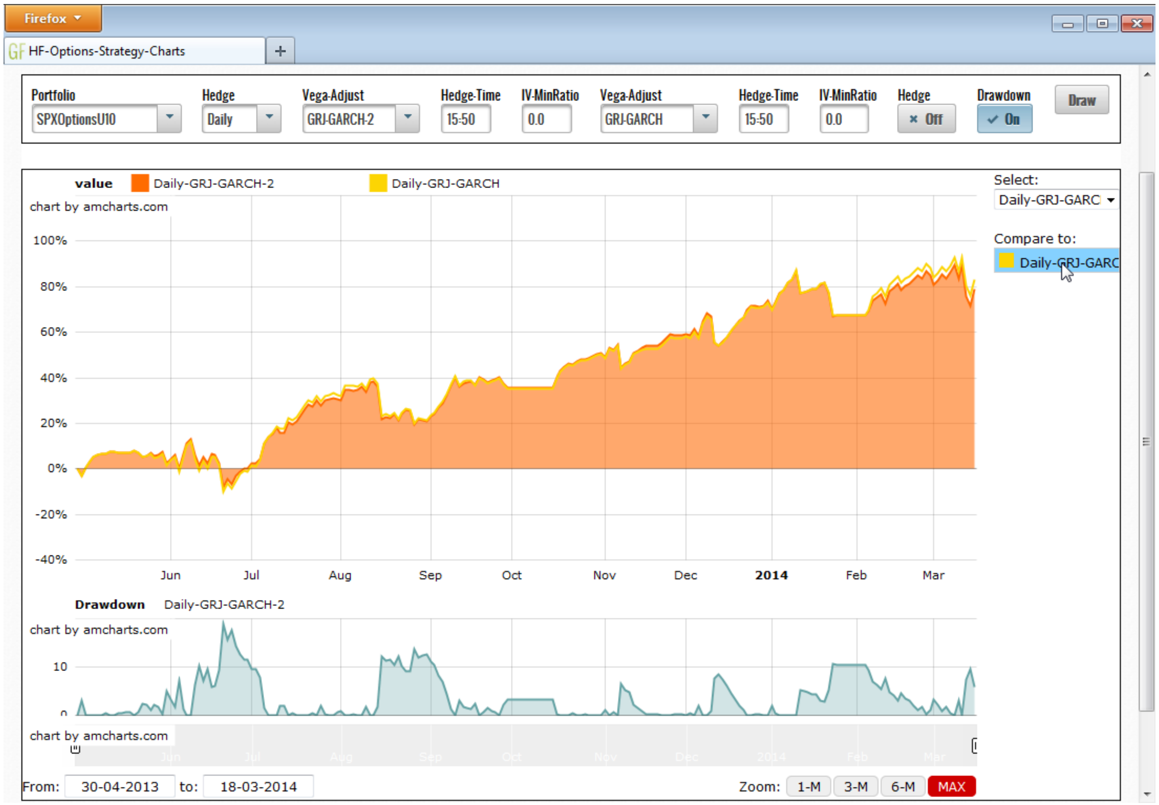
In [9] I analyzed the effect of different hedging-strategies and also the influence of hedging after the open (Lark) or near the Close (Owl). The overall message was: The hedge-refinements do not improve plain-vanilla market-delta hedging. The HF-GRJ-GARCH model was especially developed for these (real-life) scenarios.
Graphic-9 shows the performance of an OTM Strangle. The position is entered at the 1st Wednesday of the previous month. The maturity is about 6 weeks or 30 trading days. Trading is done from 2013-05-01 till 2014-03-17 (there are no longer HF-options time series available).
The model hedge (yellow) clearly outperforms the market-delta hedge (orange). But this is not entirely for free. The relative drawdown on 2013-06-20 is 11.5% for the market-hedge (left-peak in blue chart) and 18.9% for the model-delta-hedge. But the model-delta performs better on the right side. The max. relative drawdown on 2014-03-14 (Crimea crisis) is for the market delta 13.6%, the model-delta performs with a drawdown of 8.7% better.



Graphic-9: 1st Wed. Trades: market-delta (orange) and model-delta (yellow)

In [9] the delta was adjusted by the first order vega effect. Something similar can be done within the model-delta-framework. The initial value of σ is adjusted/increased by a negative Return of 2 σ. The MC-Path-generation starts with a higher volatility. Graphic-

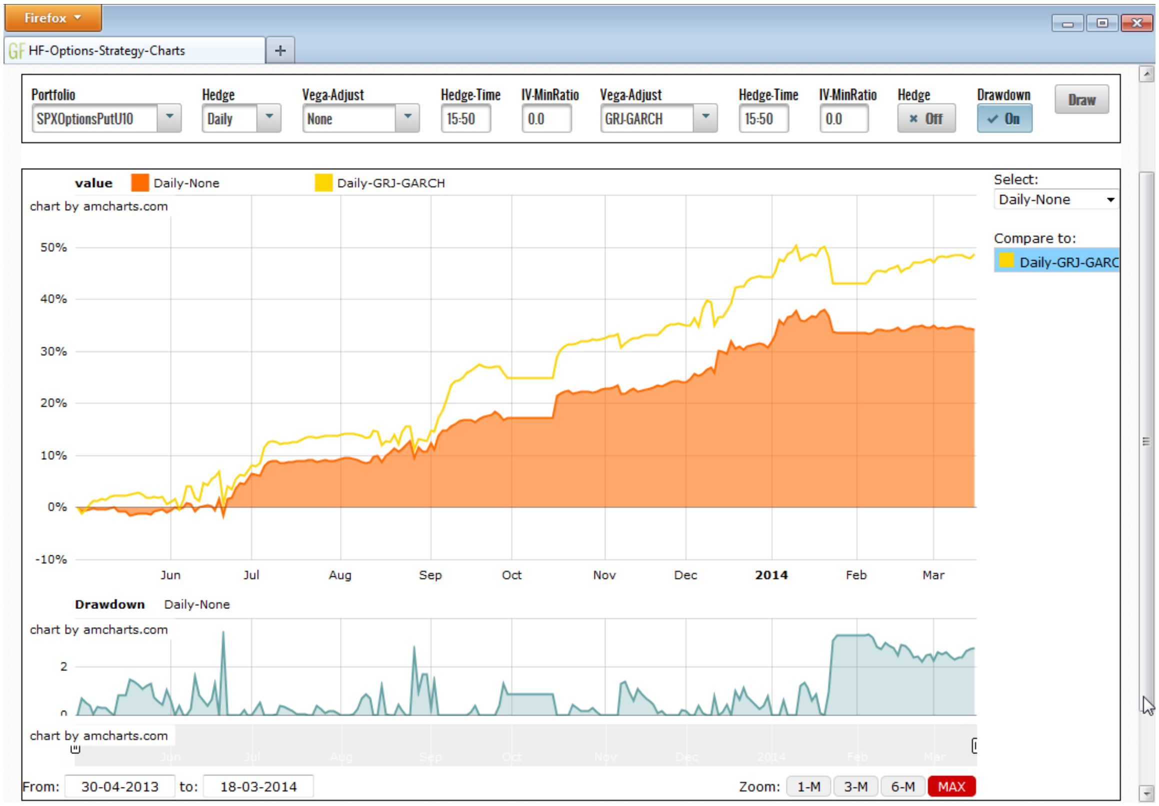
10 compares the model hedge (orange) with the 2 σ adjustment (yellow). The performance is quite similar. The adjusted hedge performs at the end slightly better.



Graphic-10: 1st Wed. Trades: model-delta (orange) and model-adjusted-delta (yellow)

Graphic-11 shows the performance of the 1st Wednesday trades. But this time only the Put-side is sold. The overall pattern is similar to the Strangle. The model-delta (yellow) performs better, but the relative drawdown is increased. The model-delta is usually somewhat smaller than the market-delta. Compensating this effect with the adjusted-model-delta does not change the behavior significantly.

In [9] other trades (3rd Wednesday and an actual traded portfolio) were analyzed. The results are (almost) the same. The model-delta outperforms the market-delta at the price of an increased risk/drawdown. The 2 σ adjustment has only a minor effect.



Graphic-11: 1st Wed. Put-Trades: market-delta (orange) and model-delta (yellow)

Conclusion:

Volatility Regime Simulation improves on Filtered Historical simulation. The behavior of the MC-Paths is more realistic. The method should also be useful for VaR calculations. If one uses the method as an entry-filter, risk is considerable reduced. But there is also less fun. For hedging it is the other way round. There is considerable more fun, but somewhat more risk.

Further Work:

An interesting extension is the direct application of the model to the VIX. This is not a trivial task, because the VIX is a mean-reverting process and has different statistical characteristics.

An even more interesting and challenging application is a joint model for the S&P and the VIX. One generates like above the S&P-paths and models along these paths the movement of the VIX. The ultimate goal would be a consistent and practical model for S&P-500 Matryoshka (S&P-Futures and -Options, VIX-Futures and -Options).

Final-Note: There are some models around which claim to solve this question. These models are in my view completely over-sophisticated and without practical value.

References:

[1] Barone-Adesi, G., Engle, R., Mancini, L.: A GARCH Option Pricing Model with Filtered Historical Simulation, Review of Financial Studies, Vol. 21, 2008
[2] Barone-Adesi, G., Giannopoulos, K., Vosper, L.: Filtering Historical Simulation. Backtest Analysis. March 2000.
[3] Donninger, Ch.: A Real-Measure Option Pricing Model with Filtered Historical Simulation. Sibyl-Working-Paper, March 2012
[4] Glosten, L., Jagannathan, R., Runkle D.: On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. Journal of Finance, Vol. 48, 1993, p. 1779-1801.
[5] Winker, P., Maringer, D.: The Convergence of Optimization Based Estimators: Theory and Application to a GARCH-Model. Univ. Erfurt Discussion Paper, 2005.
[6] Press, W., Teukolsky, S., Vetterling, W., Flannery, B.: Numerical Recipes, The Art of Scientific Computing, 3rd Edition, Chap. 10.5, p. 502 ff.
[7] Donninger, Ch.: Plain-Vanilla SPX-Options Hedging: The Effect of Smile-Adjustments and the Lark versus Owl Question. Sibyl-Working-Paper, Feb 2014
[8] Donninger, Ch. (2013). SPX Options Trading: The Kir Strategy. Sibyl-Working-Paper, Rev. 2, 2013-03-08.
[9] Wilcox, R.: Introduction to Robust Estimation and Hypothesis Testing, Chap.10.2 Theil–Sen Estimator, pp. 423–427.