

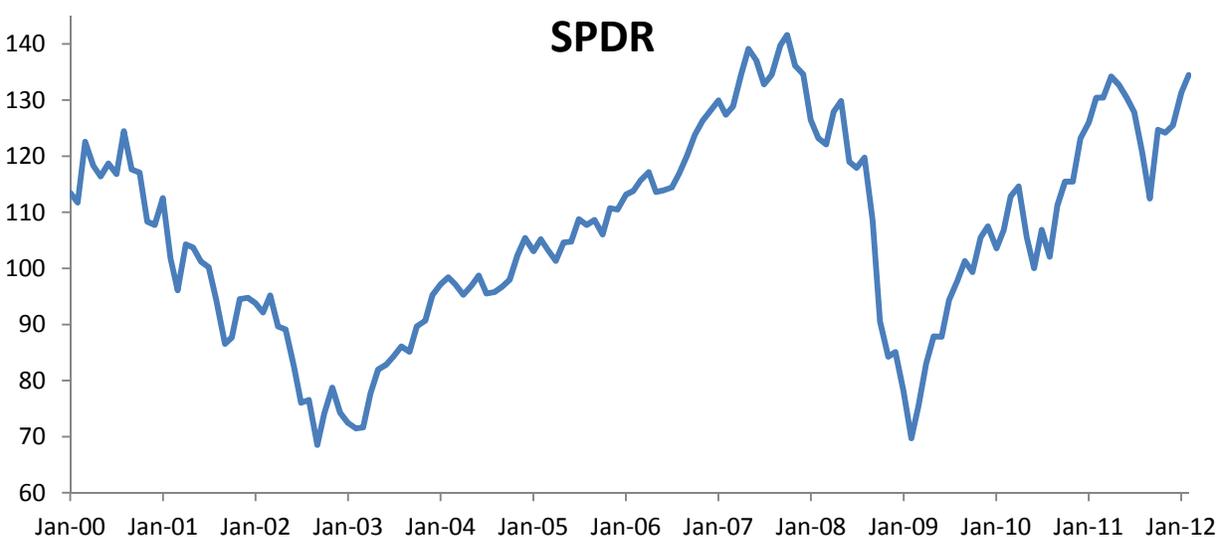
Volatility Forecasting with GARCH

What is the significance of volatility? First, the volatility, or standard deviation is an important measure of market risk. Second, it is often used to price derivatives (e.g. options) instruments.

In this paper, we will demonstrate the few steps required to convert the market index S&P 500 data into a robust volatility forecast using the NumXL Add-in within Excel.

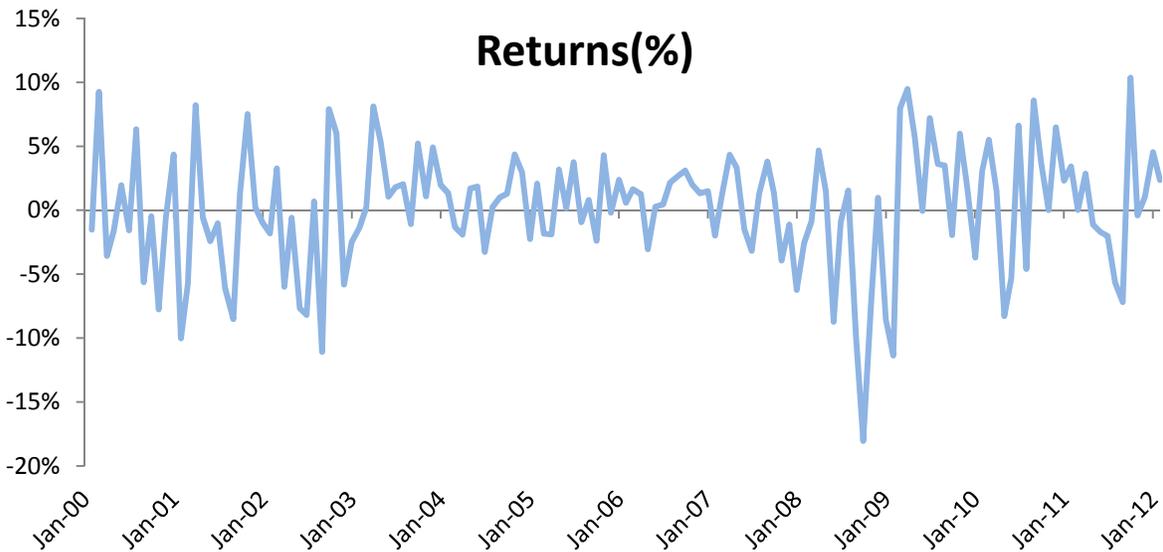
For our purpose, we are using the S&P 500 ETF (aka SPDR) prices as a proxy for the US large-cap equities market. Furthermore, we are using the monthly prices (tabulated at the beginning of the month) ranging from Jan 2000 to Feb 2012.

The objective here is to construct a model-based forecast for the volatility over the next 12 months (i.e., to the end of Feb. 2013).

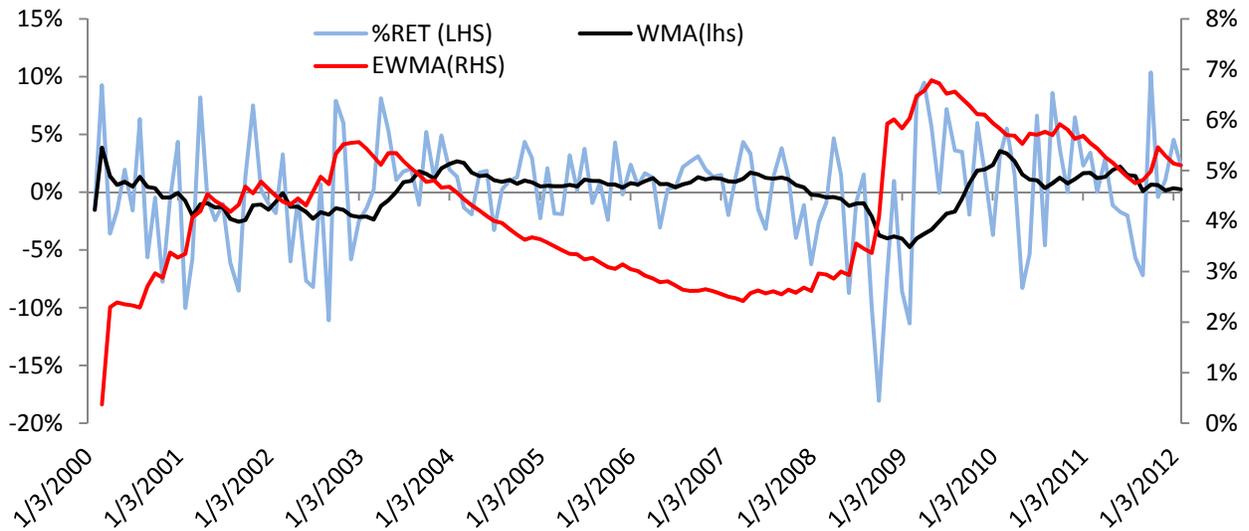


Step 1: Monthly Returns

The time series of the SPDR prices is non-stationary, and, thus is not suitable for many econometric analyses. Therefore, we first converted it to monthly returns. Furthermore, we chose the logarithmic returns over the simple returns to spread out the values of the time series as the simple returns by definition can't be lower than minus 1 (-100%).



In the graph below, we plotted the 12-month weighted moving average (WMA) and the exponential weighted volatility (EWMA) to demonstrate the variation of the mean and the volatility over time.



Please note that the volatility (proxy by EWMA) moves smoothly (unlike returns), but it is more sensitive to negative returns than it is to a positive returns market.

Step 2: Summary Statistics

Let's now calculate the descriptive statistics of the monthly returns sample: average, standard deviation, etc., to help us better understand the data. Built-in functions from NumXL can be utilized as shown to generate a set of statistics to summarize past market trends.

Using the summary statistics wizard, enter the input data set (e.g., returns cells range in column H) into the "Time Series" tab, the starting cell into the output range, and then click OK.

The screenshot shows the NumXL software interface. The 'Descriptive Statistics' wizard dialog box is open, showing the 'Input Data set' as 'monthly!\$H\$3:\$H\$147' and the 'Output Range' as '\$M\$19'. The 'OK' button is highlighted with a yellow callout '4'. The resulting output table is shown below the wizard, with a yellow callout '1' pointing to the 'Descriptive Statistics' header cell.

	G	H	I	J	K
1	Adj Close	%RET	WMA	EWMA	STDEV
2	113.45	#N/A	#N/A	#N/A	
3	111.73	-1.53%	-1.53%	#N/A	
4	122.56	9.25%	3.86%	0.37%	
5	118.25	-3.58%	1.38%	2.30%	
6	116.39	-1.59%	0.64%	2.39%	
7	118.68	1.95%	0.90%	2.35%	
8	116.82	-1.58%	0.49%	2.33%	
9	124.45	6.33%		2.29%	
10	117.63	-5.64%	0.35%	2.71%	
11	117.07	-0.48%	-0.46%	2.97%	
12	108.33	-7.76%	-0.47%	3.38%	
13	107.73	-0.56%	-0.47%	3.38%	
14	112.52	4.35%	-0.07%	3.28%	4.871%
15	101.79	-10.02%	-0.78%	3.35%	5.656%
16	96.08	-5.77%	-2.03%	4.07%	4.838%
17	104.29	8.20%	-1.05%	4.19%	5.626%
18	103.71	-0.56%	-0.96%	4.53%	5.624%
19	101.23	-2.42%	-1.33%	4.40%	5.560%
20	100.2	-1.02%	-1.28%	4.31%	5.560%
21	94.26	-6.11%	-2.32%	4.18%	5.158%
22	86.56	-8.52%	-2.56%	4.32%	5.389%
23	87.69	1.30%	-2.41%	4.68%	5.475%
24	94.53	7.51%	-1.14%	4.55%	5.878%

Descriptive Statistics			Significance
AVERAGE:	0.12%		Target
STD DEV:	4.73%		0.00
SKEW:	-0.61		0.00
EXCESS-KURTOSIS:	1.02		0.00

The generated output table is shown below. Please note that cells in the output table are connected to the input data sources; the Summary statistics wizard writes the formulas of each output using the labels specified in the first row of each data column.

Descriptive Statistics		Significance Test			5.00%	Test	p-value	SIG?
		Target	P-Value	SIG?		White-noise	9.06%	TRUE
AVERAGE:	0.12%	0.000	38.30%	FALSE		Normal Distributed?	0.09%	FALSE
STD DEV:	4.73%					ARCH Effect?	0.01%	TRUE
SKEW:	-0.61	0.000	0.15%	TRUE				
EXCESS-KURTOSIS:	1.02	0.000	1.05%	TRUE				
MEDIAN:	0.68%							
MIN:	-18.06%							
MAX:	10.36%							
Q 1:	-1.94%							
Q 3:	3.07%							

Significant -ve skew
and excess kurtosis

Examining the output table demonstrates that the distribution of the log returns exhibits negative skew (skewed to the left) and fat-tails. Furthermore, the white-noise test result indicates the absence of any significant serial correlation between returns. In sum, these results indicate that these data can be well presented by a GARCH-type model.

Step 3: GARCH Modeling

Early on, we noted that volatility (proxy EWMA) reacted differently to negative returns (downturn) than to positive ones. Fortunately, the exponential GARCH (E-GARCH) captures this phenomenon.

NumXL supports three (3) types of distributions for the residuals: (1) Gaussian, (2) Generalized Error Distribution (GED), and (3) Student's t-distribution. The sample data exhibits relatively low excess kurtosis, so the GARCH model will capture the entire excess kurtosis, yet, permitting the residuals to be normally distributed (i.e. Gaussian).

The screenshot shows the NumXL software interface with a spreadsheet of financial data and an open dialog box for GARCH modeling. The spreadsheet has columns for 'Adj Close', '%RET', 'WMA', and 'EWMA'. The dialog box, titled 'GARCH/EGARCH/GARCH-M', has the following settings:

- Input Data set: monthly!\$H\$3:\$H\$147
- Time series: monthly!\$H\$3:\$H\$147
- Ascending:
- Model Specification: ARCH Order (p) = 1, GARCH Order (q) = 1, Model = EGARCH
- Innovations: Normal Distribution, Student t-Distribution, Generalized Error Function (GED)
- Options: Guess parameters, Goodness of Fit, Residuals Diagnosis
- Significance Level (e.g. 5%): 0.05
- Output Range: monthly!\$M\$64

After entering the input data set into Time Series and the output range cell, the model type can be selected and must be primed by entering some model-specific parameters. Please note that while these values are not yet known, a crude and intelligent guess should be entered.

EGARCH(1,1)		Goodness-of-fit			Residuals (standardized) Analysis						
Param	Value	LLF	AIC	CHECK	AVG	STDEV	SKEW	KURTOSIS	Noise?	Normal?	ARCH?
μ	0.00	-3591037.29	7182080.59	1	8.25	222.40	-0.74	5.53	TRUE	FALSE	TRUE
α_0	-10.97				Target	0.00	1.00	0.00	0.00		
α_1	1.07				SIG?	FALSE	TRUE	TRUE	TRUE		
γ_1	0.00	2.385%									
β_1	-0.35										

As in the summary statistics, the cells in the E-GARCH output table are connected to the source input data via the formulas.

Step 4: E-GARCH Calibration

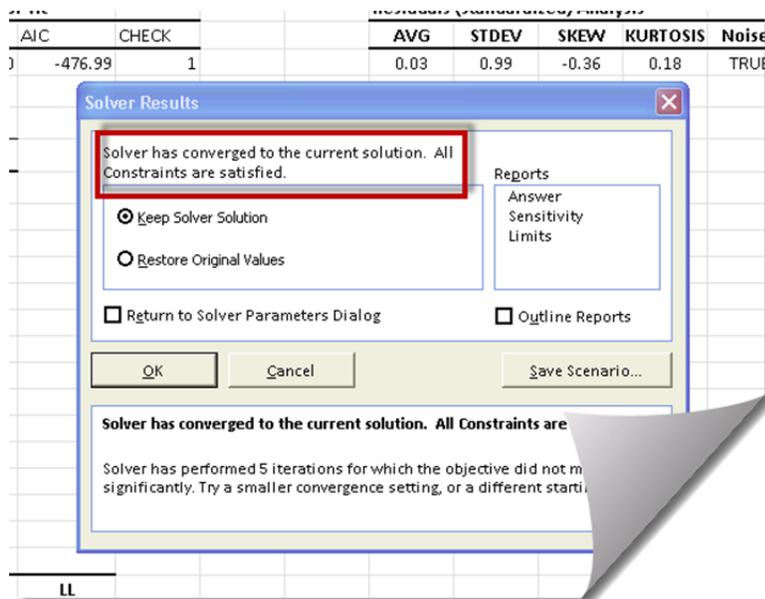
To fit (i.e. calibrate) the model with our sample data:(1) select the cell labeled “EGARCH(1,1)”, (2) click on the Calibrate icon or menu item, and finally, (3) click on Solve button.

The screenshot shows the MS Excel Solver interface. The Solver Parameters dialog box is open, with the following settings:

- Set Objective: \$Q\$60
- To: Max Min Value Of: 0
- By Changing Variable Cells: \$O\$60:\$O\$64
- Subject to the Constraints: \$Q\$60 >= 0.99999
- Make Unconstrained Variables Non-Negative:
- Select a Solving Method: GRG Nonlinear
- Solving Method: Select the GRG Nonlinear engine for Solver Problems that are smooth nonlinear. Select the LP Simplex engine for linear Solver Problems, and select the Evolutionary engine for Solver problems that are non-smooth.

The 'Solve' button is highlighted with a yellow box and the number 3.

The MS Excel Solver will maximize the log-likelihood function (LLF) by altering the coefficients values.



Step 5: Residuals Diagnosis

Once the E-GARCH model's coefficients are calibrated, we can examine the model's standardized residuals to make sure that they satisfy the underlying assumptions of the model (i.e., normally distributed).

EGARCH(1,1)		Goodness-of-fit			Residuals (standardized) Analysis						
Param	Value	LLF	AIC	CHECK	AVG	STDEV	SKEW	KURTOSIS	Noise?	Normal?	ARCH?
μ	0.12%	241.50	-476.99	1	0.03	0.99	-0.36	0.18	TRUE	TRUE	TRUE
α_0	-9.23				Target	0.00	1.00	0.00	0.00	✓	✓
α_1	0.31				SIG?	FALSE	FALSE	FALSE	FALSE		
γ_1	-0.58										
β_1	-0.46										
	VL										
	4.66%										

Using the residuals diagnosis table, we note all tests pass with the mere exception of the ARCH test which suggests the presence of a higher order (i.e. quadratic) dependency. For the purpose of this paper, we will accept the calibrated model.

The GARCH family of models captures a common and important phenomenon for volatility: mean-reversion. Using our E-GARCH model, the long-term monthly volatility is estimated at 4.66% (or 16.14% annually).

Step 6: Volatility Forecast

The GARCH-family of models describes the variation of one-step (i.e., local) volatility over time, but, in practice, we need volatility values that span multi-steps (i.e., global or term). In this paper, we will prepare both the local and the term volatilities over the next 12 months.

To accomplish this, (1) select the cell with “**EGARCH(1,1)**” Text, (2) click on “**Forecast**” icon or menu, select the latest (3) realized returns and (4) volatilities, (5) change the forecast horizon and (6) specify output location. Finally, select “OK.”

The screenshot shows the 'Forecasting' dialog box with the following settings:

- Model: \$M\$58
- Input Data set:
 - Time series: monthly!\$H\$124:\$H\$147
 - Realized Vol: monthly!\$K\$124:\$K\$147
- Output:
 - Max Steps: 18
 - Upper & Lower Limits? Significance Level (e.g. 5%)
 - Volatility Term structure? 0.05
 - Output Range: monthly!\$M\$68

The data table below shows the input data for the forecast:

	I	J	K	L	M
56	-3.27%	1.04%	4.00%	2.572%	
57	0.24%	0.89%	3.96%	2.562%	
58	1.00%	1.06%	3.84%	2.485%	
59	1.28%	0.74%	3.73%	2.122%	
60	4.35%	1.01%	3.63%	2.366%	
61	2.97%	0.85%	3.68%	2.130%	
62	-2.26%	0.49%	3.64%	2.273%	
63	2.06%	0.56%	3.57%	2.307%	
64	-1.84%	0.51%	3.50%	2.350%	
65	-1.90%	0.51%	3.42%	2.348%	
66	3.18%	0.64%	3.35%	2.452%	
67	0.15%	0.50%	3.34%	2.425%	
68	3.75%	1.08%	3.24%	2.275%	
69	-0.94%	0.98%	3.27%	2.340%	

Notes

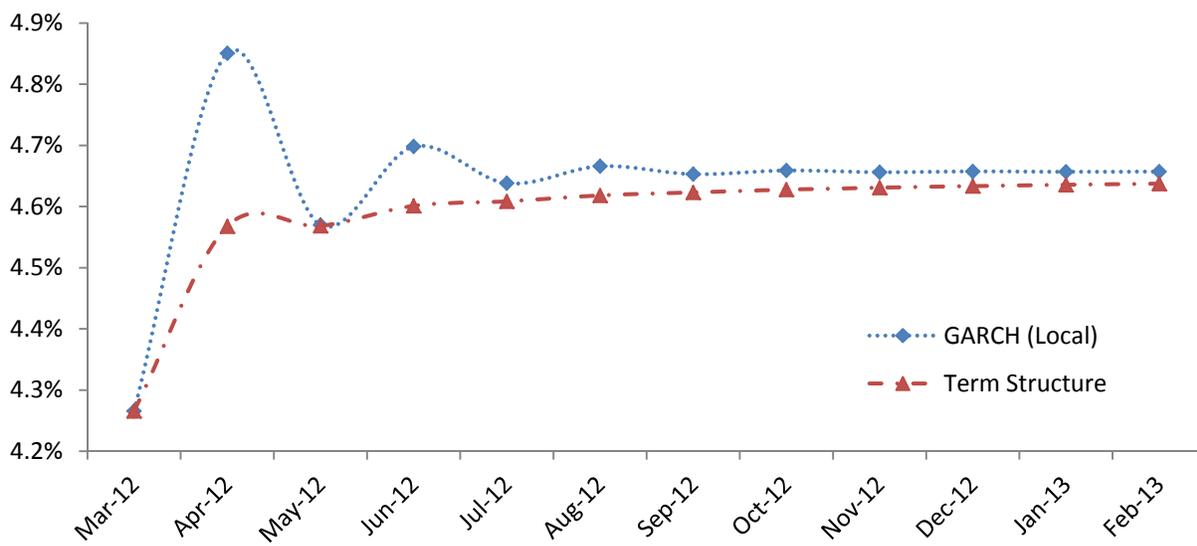
1. The input data should represent the most recent observations. For the E-GARCH (1, 1) model, at least one or two observed returns are required.
2. The Realized volatility (input data) is the most recent volatility. Since volatility is not directly observed, you would need to compute it using your preferred method. In this example, the 12 month window standard deviation was used.

The table output by the NumXL forecast is:

	M	N	O	P	Q	R
Step	Mean	STD	TS	UL	LL	
1	0.12%	4.27%	4.27%	8.53%	-8.29%	
2	0.12%	4.85%	4.57%	9.60%	-9.36%	
3	0.12%	4.57%	4.57%	9.09%	-8.85%	
4	0.12%	4.70%	4.60%	9.32%	-9.08%	
5	0.12%	4.64%	4.61%	9.21%	-8.97%	
6	0.12%	4.67%	4.62%	9.26%	-9.02%	
7	0.12%	4.65%	4.62%	9.24%	-9.00%	
8	0.12%	4.66%	4.63%	9.25%	-9.01%	
9	0.12%	4.66%	4.63%	9.25%	-9.01%	
10	0.12%	4.66%	4.63%	9.25%	-9.01%	
11	0.12%	4.66%	4.64%	9.25%	-9.01%	
12	0.12%	4.66%	4.64%	9.25%	-9.01%	
13	0.12%	4.66%	4.64%	9.25%	-9.01%	
14	0.12%	4.66%	4.64%	9.25%	-9.01%	
15	0.12%	4.66%	4.64%	9.25%	-9.01%	
16	0.12%	4.66%	4.64%	9.25%	-9.01%	
17	0.12%	4.66%	4.64%	9.25%	-9.01%	
18	0.12%	4.66%	4.64%	9.25%	-9.01%	



The E- GARCH model states that we are currently in a historically low-volatility arena, and it forecasts a rise (mean reversion) in the overall volatility to its *historical* level (4.66% /mo. Or 16.14%/yr.).



More specifically, these results indicate that for the month of February -2012 (i.e., ending March 1st, 2012), we forecast a lower volatility than Jan 2012 because the value is less than the 4.66% baseline. However, this volatility is expected to increase in March as it reverts to its long-term mean of 4.66%.