Article from
Risk Management
September 2019
Issue 45

Wavelet-Based Equity VaR Estimation

By Kailan Shang

Editor's note: This article is adapted from the research paper Wavelet-Based Equity VaR Estimation, *part of the 2019 Enterprise Risk Management Symposium call for papers. It was awarded the Actuarial Foundation's ERM Research Excellence Award in Memory of Hubert Mueller for Best Overall Paper. The full report is available at* www.ermsymposium.org/wp-content/uploads/2019/05/Shang_Actuarial-Foundation-Prizewinner.pdf.

Economic risk is an important risk for insurers offering long-term products with guaranteed benefits. When estimating the magnitude of economic risk, historical data are usually used. However, an implicit assumption of this method is that the risk is time invariant. In reality, equity market volatility varies by time. It is caused by either economic cycles or economic structural changes. Figure 1 shows the annualized volatility using daily S&P 500 index return from 1990 to 2017. Assuming a time-invariant (constant) volatility, the annualized volatility is 17.7 percent. If calculating the annualized volatility on a yearly basis, the volatility could go above 40 percent, as evidenced during the 2008 financial crisis.

Another complication is the frequency of historical data to use. The annualized volatility calculated based on different frequencies varies a great deal. Table 1 shows the annualized volatility and empirical value at risk (VaR) of S&P 500 equity index return using daily, monthly and yearly data from 1990 to 2017. For simplicity, the calculation assumes that the volatility and VaR are time invariant and that the equity index follows a geometric Brownian motion. Here VaR measures the negative return value in the left tail. For example, a 99.5 percent VaR of 15 percent means that there is a 0.5 percent chance that the return will be less than –15 percent. It is the opposite of the negative return value in the left tail.

Historical equity index returns exhibit different risk levels by frequency. Annualized empirical VaR based on high-frequency data (daily and monthly) is higher than the VaR based on low-frequency data (quarterly and yearly). This phenomenon indicates the need to analyze the economic risk at different frequencies to get a holistic view.

Figure 1
S&P 500 Index Return Annualized Volatility (1990–2017)

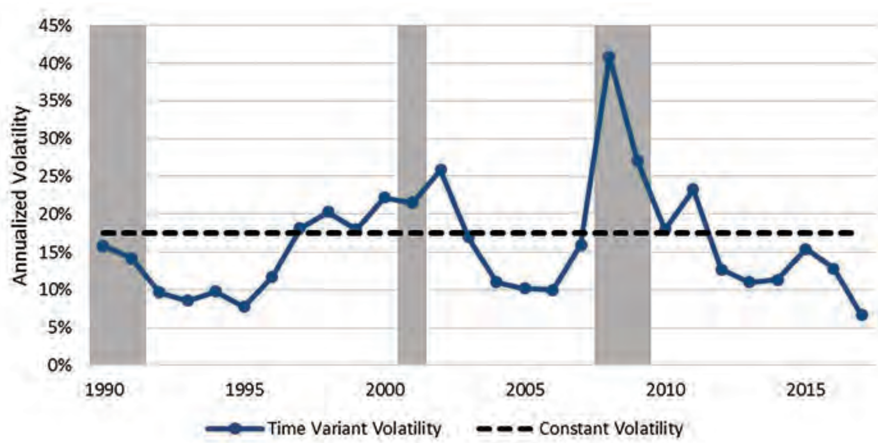


Table 1
S&P 500 Index Return Volatility and VaR by Frequency

Frequency	Time-Invariant Volatility	Annualized Volatility[1]	99.5% Empirical VaR	Annualized Empirical VaR[2]
Daily	1.1%	17.5%	3.9%	69.3%
Monthly	4.2%	14.5%	19.3%	75.3%
Quarterly[3]	7.9%	15.5%	26.9%	64.2%
Yearly	17.7%	17.5%	43.5%	43.5%

[1] Annualized volatility = time-invariant volatility $\sqrt{n}$, where n equals 250/12/4/1 for daily/monthly/quarterly/yearly frequency.
[2] Annualized empirical VaR = (99.5% Empirical VaR – Mean return) $\sqrt{n}$ – Mean return n.
[3] Minimum value of quarterly and yearly return is used for 0.5% empirical VaR because the number of data points is less than 200.

TIME SERIES MODEL

Time series models, such as generalized autoregressive conditional heteroskedasticity (GARCH) and autoregressive moving average (ARMA), can be used to capture the time-variant feature of equity volatility. An ARMA-GARCH model is used to analyze historical S&P 500 index daily returns.

$$ARMA(p,q) \sim r_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i r_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j},$$

$$\varepsilon_t = z_t \sigma_t,$$

$$GARCH(p,q) \sim \sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

where

r_t = S&P 500 index daily return. It is calculated as $log\left(S_t / S_{t-1}\right)$.

z_t = i.i.d. with zero mean and unit variance.

The distribution of z_t that can more flexibly capture skewness and heavy tails should be chosen. In this example, z_t is assumed to follow to the skewed generalized error distribution (SGED). It has the following probability density function:

$$f_{SGED}(x;\mu,\sigma,\lambda,p) = \frac{p e^{-\left\{\frac{|x-\mu+m|}{v\sigma[1+\lambda sign(x-\mu+m)]}\right\}^p}}{2v\sigma\Gamma(1/p)},$$

where

μ = location parameter. It is zero for z_t,

σ = scale parameter. It is one for z_t,

λ = skewness parameter,

p = shape parameter,

$$m = \frac{2^{\frac{2}{p}} v\sigma\lambda\Gamma\left(0.5 + \frac{1}{p}\right)}{\sqrt{\pi}} \text{ if the mean of variable equals } \mu,$$

$$v = \sqrt{\frac{\pi\Gamma\left(\frac{1}{p}\right)}{\pi\left(1+3\lambda^2\right)\Gamma\left(\frac{3}{p}\right) - 16^{\frac{1}{p}}\lambda^2\Gamma\left(0.5+\frac{1}{p}\right)^2\Gamma\left(\frac{1}{p}\right)}} \text{ if the}$$

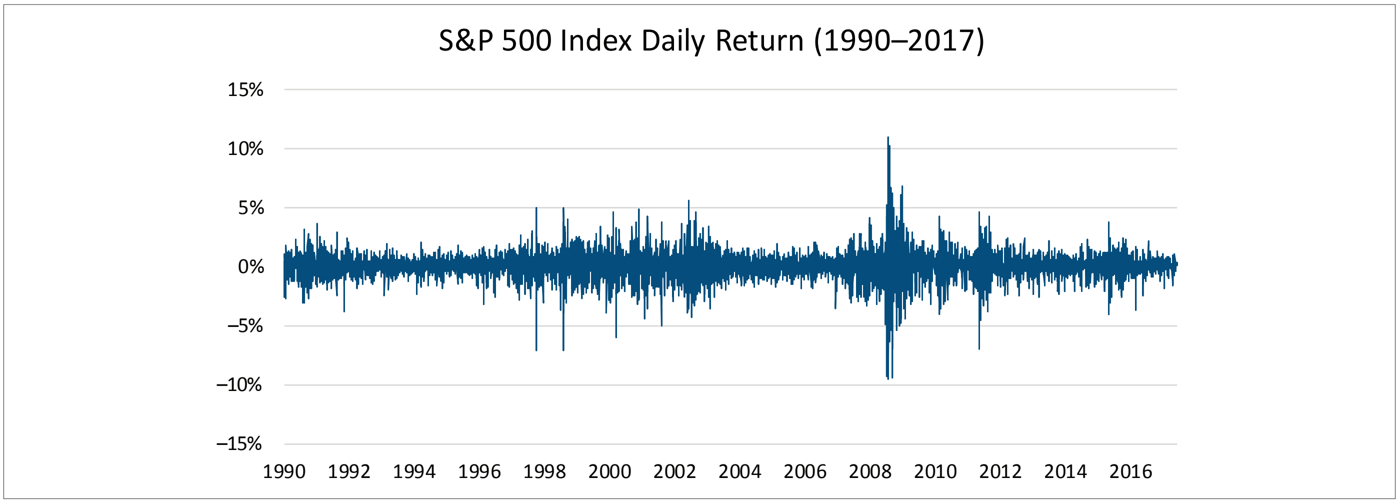
volatility of variable x equals σ.

ARMA(3,3) and GARCH(2,2) with the SGED are used to analyze historical S&P 500 daily index returns from 1990 to 2017. The orders (p and q) are chosen based on Akaike information criterion (AIC).

Figure 2 shows the daily return and the conditional volatility σ_t based on the ARMA-GARCH model. The conditional volatility varies greatly, with the highest value observed during the 2008 financial crisis.

With the fitted model, future daily VaR can be predicted. Figure 3 shows the results based on 1,000 simulations for the 251 trading days from October 2017 to September 2018. Actual daily returns are compared with the projected ranges. While 10.4 percent of actual returns fall out of the middle 90 percent range (5th percentile to 95th percentile), 1.6 percent of actual returns fall out of the middle 99 percent range (0.5th percentile to 99.5th percentile). Although the SGED generates a better range prediction than the normal distribution, it still underestimates the probability of extreme returns for the projection period.

Figure 2
S&P 500 Index Daily Return and Conditional Volatility



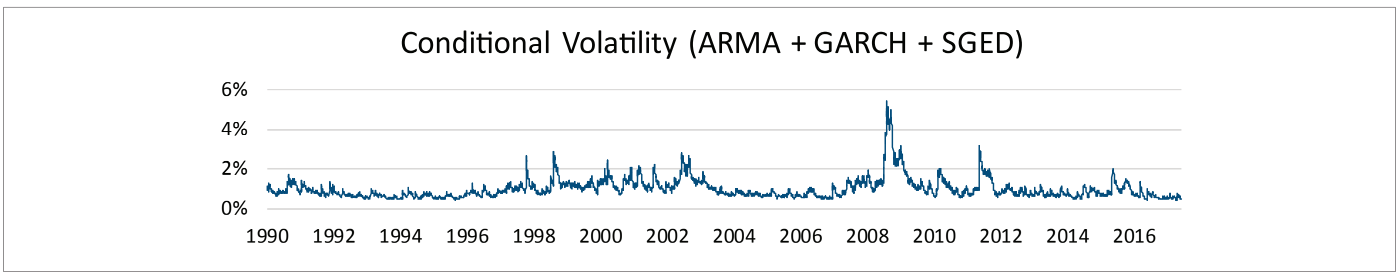
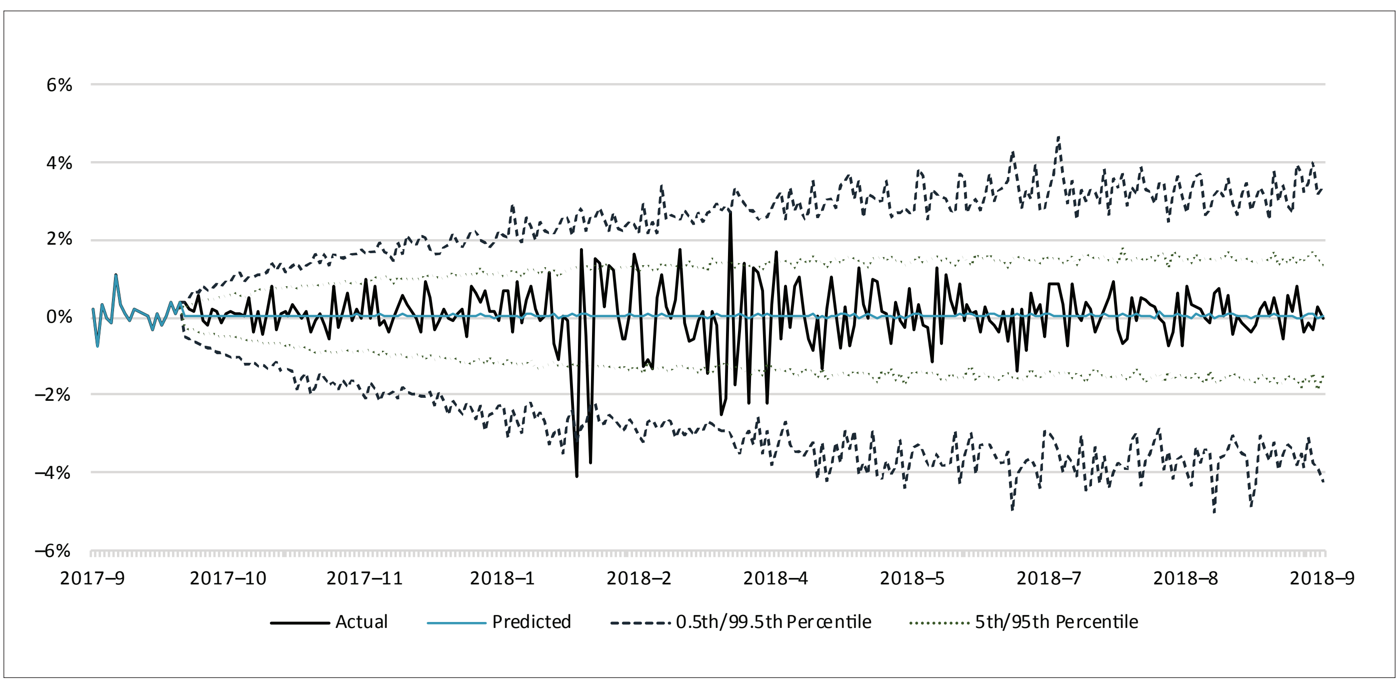


Figure 3
S&P 500 Index Daily Return Range Estimation



Instead of using closed-form formulas, annual VaR can be estimated based on simulated daily returns, as shown in Table 2. In this example, the SGED has a heavier left tail than the normal distribution.

Table 2
S&P 500 Index Return Annual VaR Estimation

	95% VaR	99.5% VaR
SGED	4.6%	24.2%
Normal distribution	5.0%	14.1%

WAVELET ANALYSIS

If the evolving of risk is driven by a few forces with different frequencies, a pure time series model may not be able to capture all the different patterns. When predicting the return and conditional volatility, the ARMA-GARCH model reflects only the direct impact of returns and volatilities in the past three days. The model cannot effectively capture the impacts for medium- and long-term patterns. People may argue that less frequent (such as annual) data can be used to estimate annual VaR. However, historical data may not be sufficient for a credible estimate, and valuable information in high-frequency data is lost.

Wavelet analysis can be used to analyze the historical data from two dimensions (time and frequency) at the same time. Wavelet analysis can be considered a combination of time series analysis and Fourier transform. Fourier transform analyzes the data purely from the frequency domain, assuming that patterns are time invariant. As shown in Figure 4, wavelet analysis keeps more time information for high-frequency data and less time information for low-frequency data.

Maximal overlap discrete wavelet transform (MODWT) is used to illustrate enhanced risk analysis based on wavelets. The MODWT is chosen over many other wavelets because its decomposition at different scales can easily be compared with original time series. The MODWT is also less sensitive than other wavelet transforms to the starting point of a time series. This is helpful to understand the patterns at different frequencies: short term, medium term or long term. Following the definition of Percival and Walden (2000), the MODWT of a time series X_t, $t = 1,2,\ldots,N$ to the jth level works as the following:

Wavelet coefficient $$\tilde{W}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \text{ } MOD \text{ } N},$$

Scale coefficient $$\tilde{V}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{g}_{j,l} X_{t-l \text{ } MOD \text{ } N},$$

where $\tilde{h}_{j,l}$ = wavelet filter constructed by convolving j filters composed of $\tilde{g}_l$ and $\tilde{h}_l$. It suffices the following conditions:

$$\sum_{l=0}^{L-1} \tilde{h}_l = 0 \quad \sum_{l=0}^{L-1} \tilde{h}_l^2 = \frac{1}{2} \quad \sum_{l=-\infty}^{\infty} \tilde{h}_l \tilde{h}_{l+2n} = 0 \text{ for all integers } n > 0,$$

$\tilde{g}_{j,l}$ = scale filter constructed by convolving j filters composed of $\tilde{g}_l$. It suffices the following conditions:

$$\sum_{l=0}^{L-1} \tilde{g}_l = 1 \quad \sum_{l=0}^{L-1} \tilde{g}_l^2 = \frac{1}{2} \quad \sum_{l=-\infty}^{\infty} \tilde{g}_l \tilde{g}_{l+2n} = 0 \text{ for all integers } n > 0,$$

$$\sum_{l=-\infty}^{\infty} \tilde{g}_l \tilde{h}_{l+2n} = 0 \text{ for all integers } n,$$

$L_j = (2^j - 1)(L-1) + 1$. L is the width of the base level filter.

The maximum number of levels depends on the available data points. Table 3 lists the frequency of the first eight levels.

Figure 4
Wavelet Analysis Concept

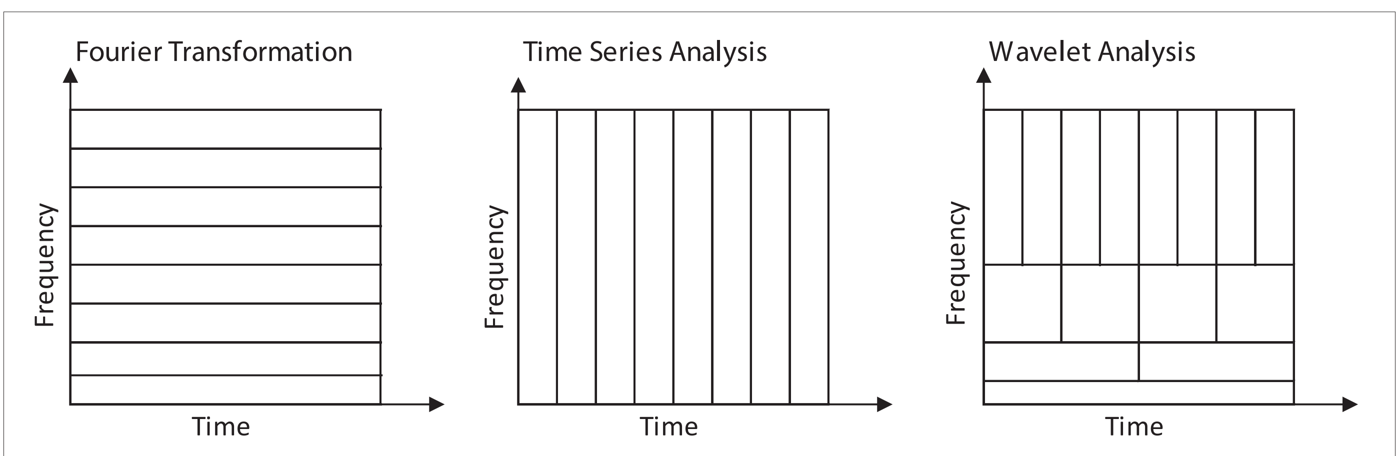


Table 3
Frequency of Decomposition Levels

Level (j)	Frequency	Scale (1/Frequency)*
1	[1/4,1/2]	2–4 days
2	[1/8,1/4]	4–8 days
3	[1/16,1/8]	8–16 days
4	[1/32,1/16]	16–32 days
5	[1/64,1/32]	32–64 days
6	[1/128,1/64]	64–128 days
7	[1/256,1/128]	128–256 days
8	[1/512,1/256]	256–512 days

* The scale is measured in business days.

To analyze the equity risk, LA(8) (Daubechies least asymmetric filter with L = 8) is used to define $\tilde{h}_{j,l}$ and $\tilde{g}_{j,l}$. Figure 5 shows the wavelet filters $\tilde{h}_{j,l}$ and scale filters $\tilde{g}_{j,l}$ for the first three levels. The wavelet dampens out with larger width as the level goes up. The same pattern applies when the level goes higher than level three.

The original time series (S&P 500 index daily return) is decomposed into eight levels. Figure 6 shows the wavelet coefficients ($\tilde{W}_{j,t}$) for all eight levels and the scale coefficients ($\tilde{V}_{j,t}$) for the eighth level. The wavelet coefficients are smoother at a higher level, representing longer-term volatility. The scale coefficients at the highest level represent the volatility that is not explained by wavelet coefficients.

TIME-INVARIANT RISK ANALYSIS

Wavelet analysis can be used to attribute the total volatility to different levels. The total variance can be calculated as the sum of the variances at each level:

$$\sigma_X^2 = \sum_{J=1}^{J_M} \sigma_X^2(j),$$

where

σ_X^2 = total variance of the original time series,

$\sigma_X^2(j)$ = variance of the decomposition at level j,

J_M = number of levels used in wavelet analysis.

Figure 5
LA(8) Wavelet and Scale Filters for MODWT

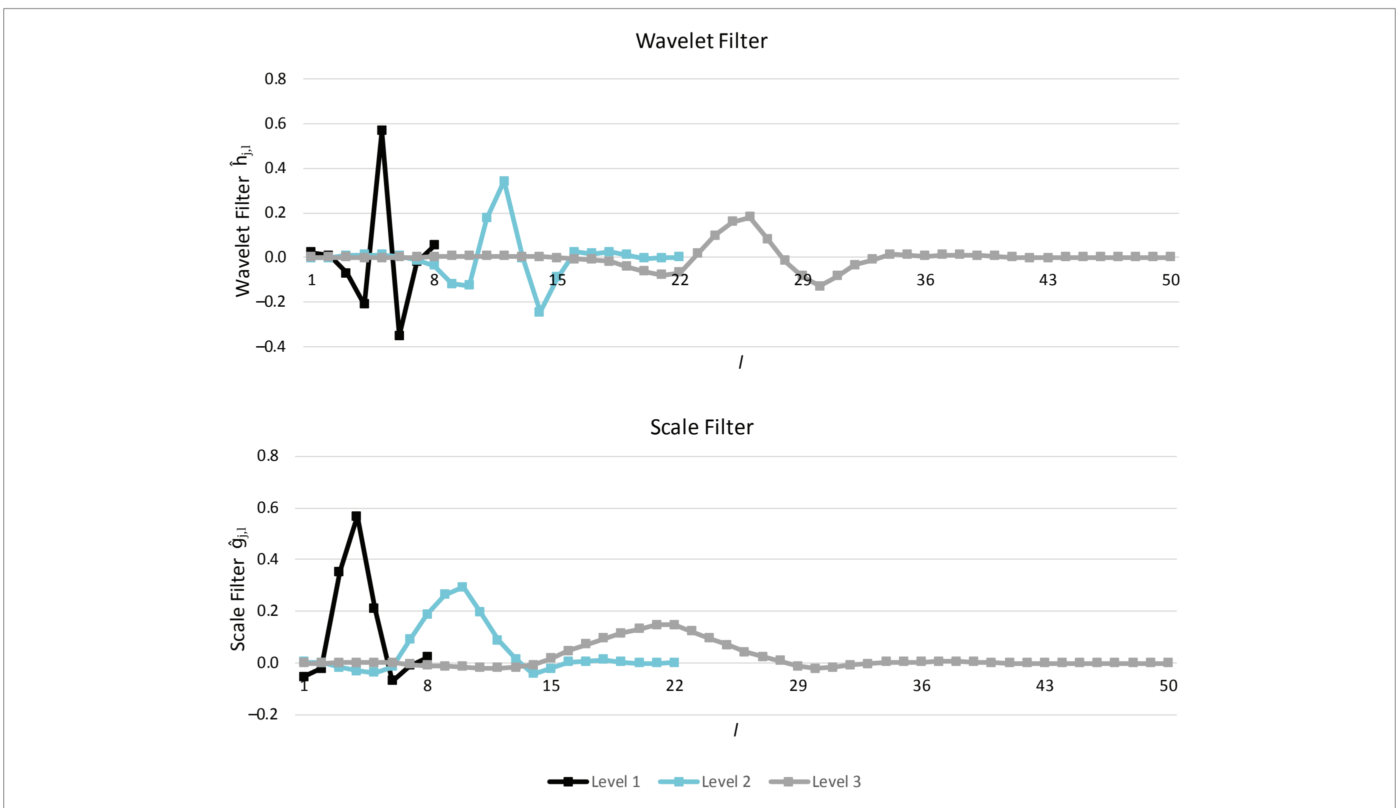
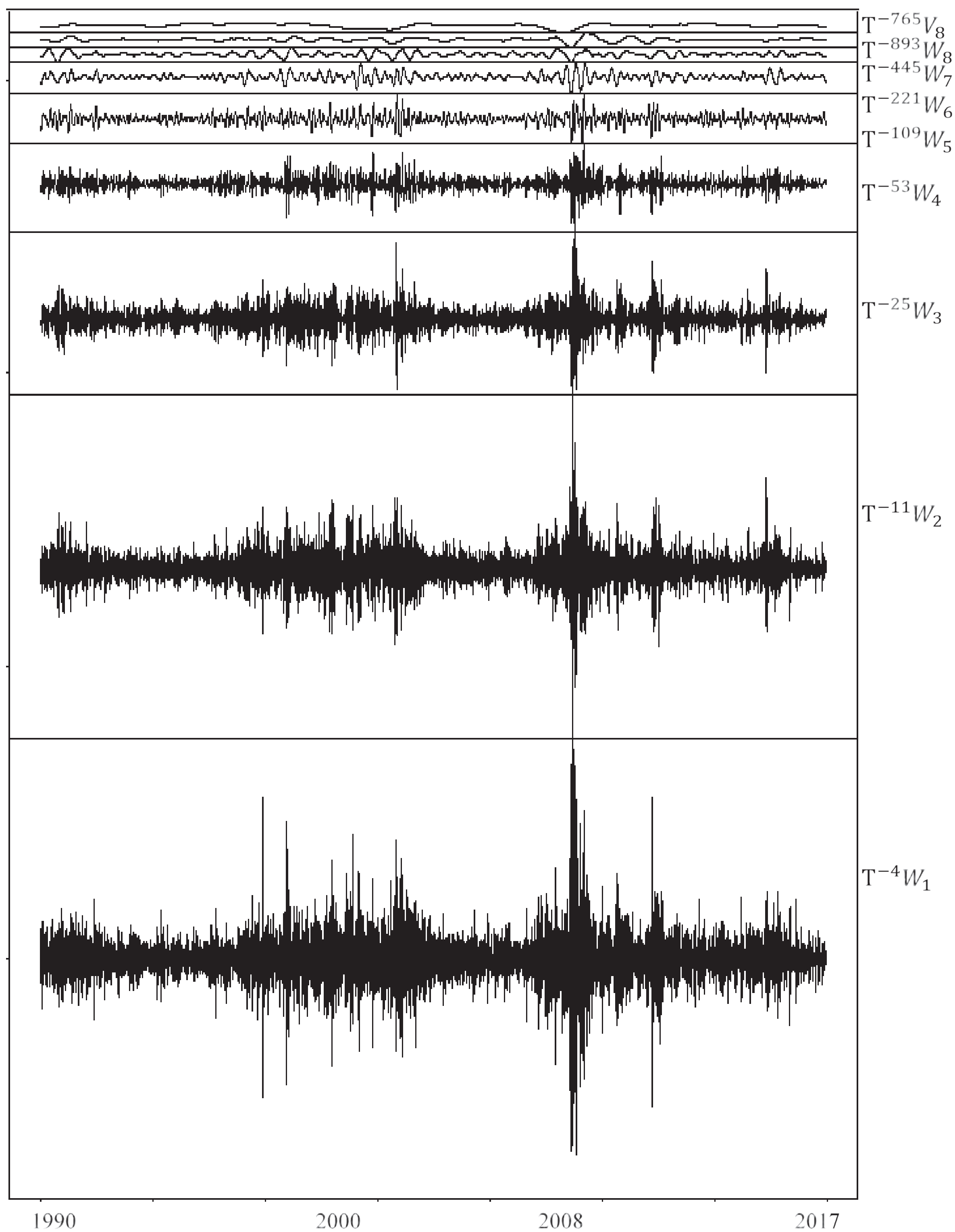


Figure 6
MODWT Wavelet Coefficients and Scaling Coefficients



Note: T^{-i} means that the series of the coefficients is shifted by *i* positions backward so that all series are on the same timeline.

Also, $\sigma_X^2(j)$ has an unbiased estimator:

$$\hat{\sigma}_X^2(j) = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \tilde{W}_{j,t}^2,$$

where

$$M_j = N - L_j + 1.$$

Skewness and kurtosis of each level can be estimated as well:

$$\textbf{Skewness} \quad \hat{S}_X(j) = \frac{\frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \tilde{W}_{j,t}^3}{\hat{\sigma}_X^3(j)},$$

$$\textbf{Kurtosis} \quad \hat{K}_X(j) = \frac{\frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \tilde{W}_{j,t}^4}{\hat{\sigma}_X^4(j)}.$$

Table 4 lists the mean, variance, skewness and kurtosis for each decomposition level and the original time series. Low levels (high frequency/short term) contribute most of the variance of the original return series. Skewness and kurtosis are quite different among the eight levels, which indicates that the patterns at different frequencies are different, and it may be beneficial to model them separately.

The empirical VaR of the original time series can be approximated by aggregating the VaR at each decomposition level as follows:

$$VaR_{Agg} = \sqrt{\sum_{j=1}^{J_M} VaR_j^2},$$

where

VaR_{Agg} = aggregated VaR,

VaR_j = VaR at level j.

In this example, aggregated empirical VaR is 3.94 percent, compared to 3.93 percent calculated directly from the original time series. The non-normality of the original time series is preserved well by the wavelet coefficients in this example.

TIME-VARIANT RISK ANALYSIS

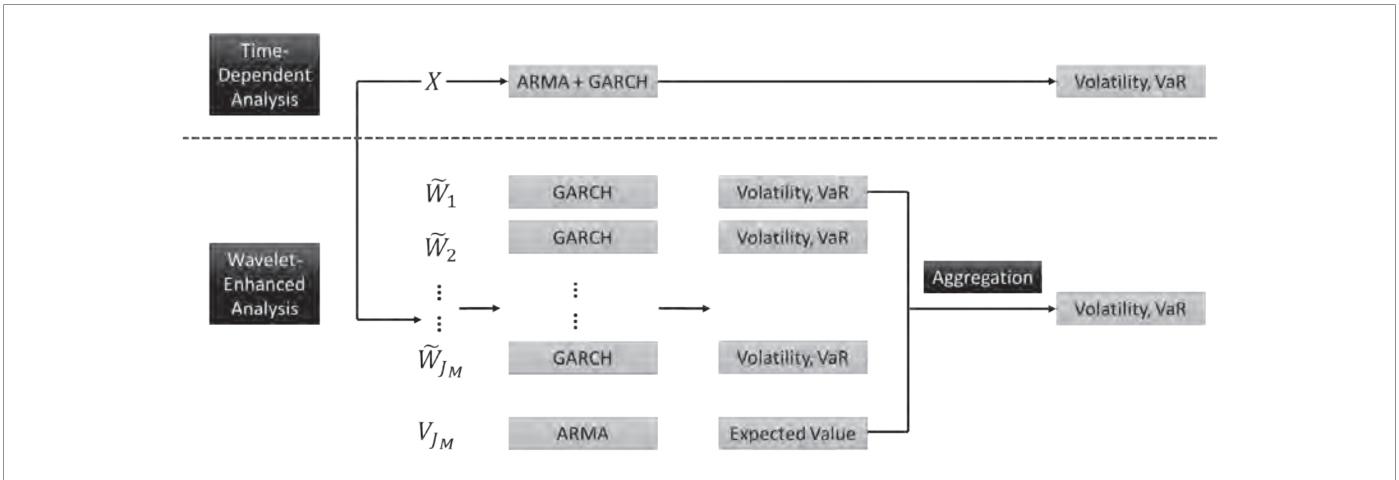
The wavelet analysis in the previous section assumes a constant volatility. Time-variant risk analysis can be enhanced with wavelet analysis as well to reflect different patterns at each wavelet decomposition level. This section builds on the ARMA-GARCH example to include analysis at each decomposition level. As shown in Figure 7, instead of modeling the original time series with one model, wavelet-enhanced time-dependent analysis studies wavelet coefficients at each level separately to understand the risk in different ranges of frequency. Wavelet coefficients are fitted into a GARCH model to get the volatility and VaR information. Scale coefficients at the highest level are fitted into ARMA and GARCH models to understand the trend of the time series. They are aggregated to get the predicted return, total volatility and VaR.

Following the simulation method used in the time series model to simulate future equity returns, wavelet coefficients can be simulated at each decomposition level. Conditional volatility and VaR can be projected for each level according to the

Table 4
Descriptive Statistics at Different Decomposition Levels

	Mean	Volatility	Variance Contribution	Skewness	Kurtosis	99.5% Empirical VaR	99.5% VaR (Normal)
Level 1	0.0000%	0.8%	53.5%	0.3	12.7	3.0%	2.1%
Level 2	0.0000%	0.6%	24.9%	0.2	11.3	2.0%	1.4%
Level 3	−0.0001%	0.4%	12.3%	0.1	7.6	1.2%	1.0%
Level 4	0.0000%	0.2%	5.0%	−0.1	6.3	0.9%	0.6%
Level 5	−0.0001%	0.2%	2.3%	0.1	5.5	0.5%	0.4%
Level 6	−0.0002%	0.1%	1.2%	0.03	5.2	0.4%	0.3%
Level 7	0.0001%	0.1%	0.4%	−0.2	3.7	0.2%	0.2%
Level 8	−0.0001%	0.1%	0.3%	−0.3	6.4	0.2%	0.2%
Original	0.0274%	1.1%	—	−0.2	11.9	3.93%	2.84%

Figure 7
Wavelet-Enhanced Time-Dependent Analysis Structure



calibrated GARCH model. They can be aggregated to predict the total VaR:

$$VaR_{Agg,T+l} = \sqrt{\sum_{j=1}^{J_M} VaR_{j,T+l}^2} - \mathbb{E}(r_{T+l}),$$

$$VaR_{j,T+l} = -\sigma_{j,T+l} SGED_j^{-1}(1-p),$$

where

$VaR_{Agg,T+l}$ = aggregated daily *VaR* at $T + l$, l periods ahead of T,

$VaR_{j,T+l}$ = daily *VaR* at $T + l$ at decomposition level j. The expected value of wavelet coefficients is zero and therefore is not included in the formula,

$\sigma_{j,T+l}$ = projected conditional volatility of level j wavelet coefficient at $T + l$,

$SGED_j^{-1}(1-p)$ = the $[100 \times (1 - p)]$th percentile of fitted SGED for level j wavelet coefficients.

Figure 8 shows the daily return range prediction based on 1,000 simulations for 250 trading days from the beginning of October 2017. Actual daily returns till September 2018 are compared with the projected ranges. While 10.2 percent of actual returns fall out of the middle 90 percent range (5th percentile to 95th percentile), 0.7 percent of actual returns fall out of the middle 99 percent range (0.5th percentile to 99.5th percentile). Compared to a pure time-dependent prediction, as in Figure 3, wavelet-enhanced prediction has a wider predicted range for extreme returns (0.5th percentile and 99.5th percentile).

> Time-variant risk analysis can be enhanced with wavelet analysis to reflect different patterns at each wavelet decomposition level.

For decision makers with a longer time horizon, annual VaR is a better measure than daily VaR for risk assessment. Multiresolution analysis (MRA) based on MODWT can be used to construct daily returns from transformed coefficients that preserve the autocorrelation of daily returns. Annual returns are then calculated based on simulated daily returns. Table 5 compares the annual VaR derived by different methods for the period from October 2017 to September 2018. Wavelet-enhanced time-dependent analysis provides a much higher annual VaR than a pure time-dependent analysis given a low volatility environment in September 2017. Wavelet analysis has a longer memory and helps preserve the long-term pattern much better than the time-dependent analysis in this example. Wavelet-enhanced time-dependent analysis also reflects current market conditions to predict the future risk in a given time horizon.

Figure 8
Wavelet-Based S&P 500 Index Daily Return Range Estimation

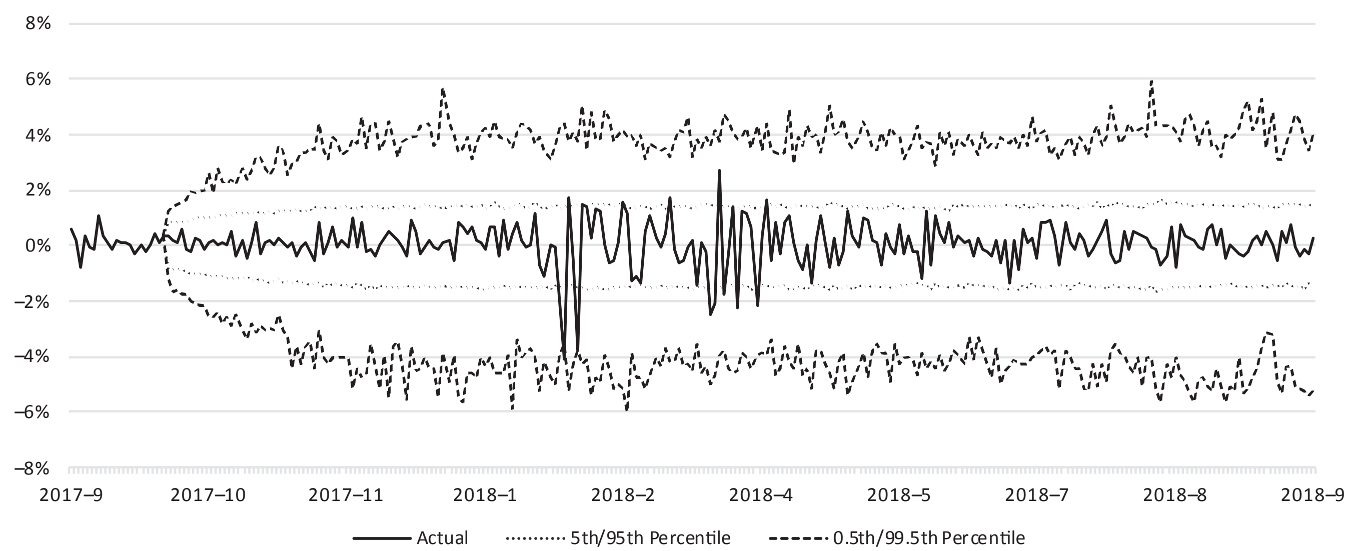


Table 5
S&P 500 Index Return Annual VaR Estimation

	Projection Type	Model	95% VaR	99.5% VaR
Time-dependent analysis	Conditional	ARMA + GARCH	4.6%	24.2%
Wavelet-enhanced time-dependent analysis	Conditional	MODWT + MRA	17.6%	39.9%
Empirical analysis (Jan. 1990–Sept. 2017)	Unconditional	Statistical analysis	26.9%	43.5%

For VaR estimation at a high confidence level, wavelet-enhanced time-dependent analysis is the best option based on the back-testing results at different volatility levels. In addition, this type of analysis can adjust itself based on new information in a timely manner.

CONCLUSION

Unlike time series analysis, wavelet analysis can be used to systematically analyze historical time series data by time and frequency concurrently. Wavelet analysis provides a decomposition of the total risk and can tell whether short-, medium- or long-term risk is dominating. It can better capture different patterns at different frequency levels to improve risk estimation. Risk measures such as volatility and VaR can be calculated directly using wavelet models.

Wavelet analysis is especially useful when time horizon has a significant impact on risk analysis. It can help refine assumptions such as volatility, tail heaviness and correlation according to the time horizon of risk analysis.

Kailan Shang, FSA, ACIA, CFA, PRM, SCJP, is an associate director at Aon PathWise Solutions Group, Canada. He can be reached at *kevin.shang@aon.com*.

REFERENCE

Percival, Donald B., and Andrew T. Walden. 2000. *Wavelet Methods for Time Series Analysis*. New York: Cambridge University Press, 159–180.