BACKTESTING EXPECTED SHORTFALL WITH A SKEWED EXPONENTIAL POWER DISTRIBUTION IN ELECTRICITY MARKETS

Aymeric Thibault and Pascal Bondon

Laboratoire des Signaux et Systèmes, CNRS - CentraleSupélec - Université Paris-Sud, France.

ABSTRACT

Interest in risk measurement for spot price has increased since the worldwide deregulation and liberalization of electricity started in the early 90's. This paper is focused on quantifying risk for the Swedish spot price. Our analysis is based on a generalized autoregressive conditional heteroskedastic (GARCH) process with skewed exponential power innovations to model the stochastic component of the price. A Expected Shortfall (ES) backtesting procedure is presented and our model performance is compared to commonly used distributions in risk measurement. We show that the skewed exponential power distribution outperforms the competitors for the upside risk, which is of high interest as electricity spot prices are positively skewed.

***Index Terms*—** Electricity markets, Tail dependence, Asymmetric distributions, Expected shortfall.

1. INTRODUCTION

The worldwide deregulation and liberalization of electricity markets that started in the early 90's led market participants to move from monopolies to competitive markets. As a consequence, the process did lower the wholesale price of electricity but at the cost of higher volatility [1]. Modelling and forecasting spot price is a subject of great importance among the electrical engineering community as high price movements create a market risk for both producers and consumers. Risk measurement is therefore a critical issue to focus on.

Electricity has specific characteristics not shared by other commodities (oil, gas) [1]. Its limited storability and transportability greatly influence spot price behaviour: it is strongly dependent on spatial and temporal distribution and demand conditions (climate and weather conditions, business activity). For these reasons, day-ahead prices exhibit special features from the signal processing point of view. Mean-reverting processes have been observed in electricity spot prices. However, the rate of reversion is very high compared to other financial markets. Another aspect is the presence of price spikes. Due to the non-storability of electricity, times of high demand greatly affect spot prices. On the opposite side, negative prices have been introduced to enable the market to respond with appropriate price signals in the event of oversupply. Besides of price spikes, volatility clustering has been noticed. The climate and weather conditions impact electricity supply and spot prices display a strong seasonal behaviour on a daily, weekly and yearly basis demanding special attention.

Electricity spot price modelling and forecasting have been covered by various approaches [2], including statistical models such as autoregressive moving average (ARMA) processes [3], GARCH processes [4, 5], reduced-form models such as Markov regime-switching models [6], and computational intelligence models such as neural networks [7, 8].

This paper focuses on risk measurement in Nord Pool, which is the world's first international commodity exchange for electrical power. The Nord Pool markets are divided into several bidding areas. The different bidding areas help point out constraints in the transmission systems, and ensure that local market conditions are reflected in the price. Sweden is divided into four bidding areas and we are interested in quantifying risk for the Swedish bidding area SE3. For this purpose, we estimate the ES with the help of an ARMA-GARCH model whose innovations follow a skewed exponential power (SEP) distribution.

The paper is organized as follows. Section 2 identifies the trend and seasonal components of the SE3 price using a wavelet decomposition and a MA technique. Section 3 is focused on modelling the price stochastic component through an ARMA-GARCH model with SEP innovations. The model is fitted by maximum likelihood estimation (MLE) and an in-sample analysis is presented. Section 4 gives out-of-sample results for the ES estimation and backtesting. It is shown that the SEP distribution exhibits the best performance for ES estimation among other distributions used in risk forecasting. Concluding remarks can be found in Section 5.

2. IDENTIFYING TREND AND SEASONAL COMPONENTS

We consider the price of the Swedish bidding area SE3 (P_t) from January 1, 2013 to February 1, 2016, totalling 1127 daily observations. Figure 1 displays the series and its main statistical characteristics are presented in Table 1. Figure 1 shows that (P_t) is nonstationary and has many spikes, especially at the end of the period. Following the standard classical (or

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

frequentist) methodology in time series analysis [9], we decompose (P_t) as $P_t = m_t + s_t + Y_t$ where (m_t) and (s_t) are the deterministic trend and seasonal components, respectively, and (Y_t) is the stochastic component, see e.g. [9].

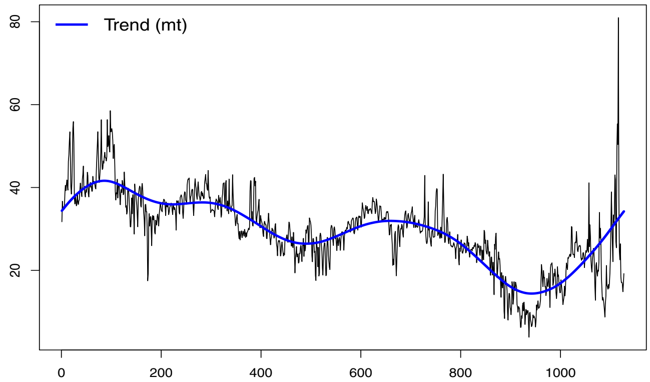

Fig. 1: Price of the Swedish bidding area SE3 price (P_t) (EUR/MWh) and its trend (m_t) from January 1, 2013 to February 1, 2016.

First, we capture the trend (m_t) with a wavelet decomposition as discussed in [10]. A so-called wavelet low pass filter is known to be more robust to outliers and to provide a less periodic alternative to Fourier analysis. A wavelet family includes a pair of a father and a mother wavelet, like the Daubechies wavelets of order 4 used here. A signal f can be decomposed as a sequence of projections onto one father and a sequence of mother wavelets $f = u_J + v_J + v_{J-1} + \cdots + v_1$ where 2^J is the maximum scale sustainable by the number of points of the signal. Here, we approximate the trend (m_t) by u_J. Higher levels of refinement can be obtained by adding successively to u_j the mother wavelets $v_J, v_{J-1}, \ldots$. In our analysis, we use a u_7 approximation, which roughly corresponds to a four-month period smoothing ($2^7 = 128$ days). The trend (m_t) is displayed on Figure 1 and captures fairly well the long-term behavior of the series.

Then, we estimate the seasonal component. Since the data is daily collected a weekly periodicity is expected. For each $k = 1, \ldots, 7$, we compute the average w_k of the deviations $(P_{k+7j} - m_{k+7j})$ for $1 \leq k + 7j \leq 1127$. Since these average deviations do not necessarily sum to zero, we estimate the seasonal component as

$$s_k = w_k - \tfrac{1}{7}\textstyle\sum_{i=1}^{7} w_i, \quad k = 1, \ldots, 7,$$

and $s_t = s_{t-7}$ for $t > 7$.

Finally, the stochastic component (Y_t) is obtained by $Y_t = P_t - m_t - s_t$. The main statistical characteristics of (Y_t) are presented in Table 1. We observe that (Y_t) is skewed and has a fat-tailed distribution.

3. MODELLING THE STOCHASTIC COMPONENT

Our purpose is to model the heavy tails and the excess kurtosis of (Y_t) in order to provide accurate estimation of a risk measure. More precisely, we fit to (Y_t) an ARMA(p, q)-GARCH(r, s) model with SEP innovations,

$$\begin{aligned} Y_t &= \sum_{i=1}^{p} \phi_i Y_{t-i} + X_t + \sum_{i=1}^{q} \theta_i X_{t-i}, \\ X_t &= \sigma_t \varepsilon_t, \\ \sigma_t^2 &= a_0 + \sum_{i=1}^{r} a_i X_{t-i}^2 + \sum_{i=1}^{s} b_i \sigma_{t-i}^2, \end{aligned} \tag{1}$$

where the polynomials $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ and $\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$ have no common zeros and neither $\phi(z)$ nor $\theta(z)$ has zeros in the closed unit disk $\{z \in \mathbb{C} : |z| \leq 1\}$, σ_t is the positive square root of σ_t^2, $a_0 > 0$, all coefficients (a_i, b_i)'s are nonnegative, $\sum_{i=1}^{\max(r,s)}(a_i + b_i) < 1$, and (ε_t) is a sequence of independent and identically distributed (iid) random variables satisfying $\mathrm{E}(\varepsilon_t) = 0$ and $\mathrm{E}(\varepsilon_t^2) = 1$. The innovation ε_t follows a SEP distribution whose density is defined by

$$f(x|\gamma, \epsilon) = \frac{1}{\rho} f_\gamma \left[\frac{x - \mu}{(1+\epsilon)\rho}\right] \mathbb{1}_{\{x<\mu\}} + \frac{1}{\rho} f_\gamma \left[\frac{x - \mu}{(1-\epsilon)\rho}\right] \mathbb{1}_{\{x\geq\mu\}},$$

where $f_\gamma(x) = c\exp(-|x|^\gamma)$, $\gamma > 0$ is the shape parameter, $c^{-1} = 2\Gamma(1+1/\gamma)$ and Γ is the Gamma function, $\epsilon \in (-1, 1)$ is the skew parameter, $\mu \in \mathbb{R}$ is the location parameter and $\rho > 0$ is the scale parameter. Since

$$\begin{aligned} \mathrm{E}(\varepsilon_t) &= \mu - 2\epsilon\rho d_1 = 0, \\ \mathrm{Var}(\varepsilon_t) &= \rho^2[(1 + 3\epsilon^2)d_2 - 4(\epsilon d_1)^2] = 1, \end{aligned}$$

where

$$d_r = \int_{-\infty}^{+\infty} |x|^r f_\gamma(x) dx = \Gamma\left(\frac{r+1}{\gamma}\right) \Big/ \Gamma\left(\frac{1}{\gamma}\right),$$

the parameters (μ, ρ) are determined from (γ, ϵ), see [11].

Time correlation motivates the use of an ARMA model. The sample partial autocorrelation function of (Y_t) is displayed in Figure 2(a). All the partial autocorrelation coefficients at lag $h \geq 7$ are not statistically significantly different from 0. We compute the Bayesian information criterion (BIC) for each ARMA(p, q) model with $0 \leq p, q \leq 6$, and the smallest value is 1919.43 and is obtained with $(p, q) = (6, 0)$. Therefore, we choose the ARMA$(6, 0)$ model for (Y_t).

The study of the residuals (X_t) exhibits positive skewness, heavy tails and conditional heteroskedasticity as shown in Table 1 and Figures 2(b) and (c). These features encourage the use of a GARCH model. In order to avoid overparameterisation, we fit a GARCH(1,1) model to (X_t) which is usual in the literature.

Signal	Mean	Standard Deviation	Min	Max	Skewness	Excess Kurtosis
(P_t)	31.01	10.05	3.51	88.1	0.07	0.96
(Y_t)	0	6.08	-20.6	53.72	0.99	7.38
(X_t)	-0.02	3.19	-10.97	11.33	0.11	1.59

Table 1: Statistical characteristics of (P_t), (Y_t) and (X_t).

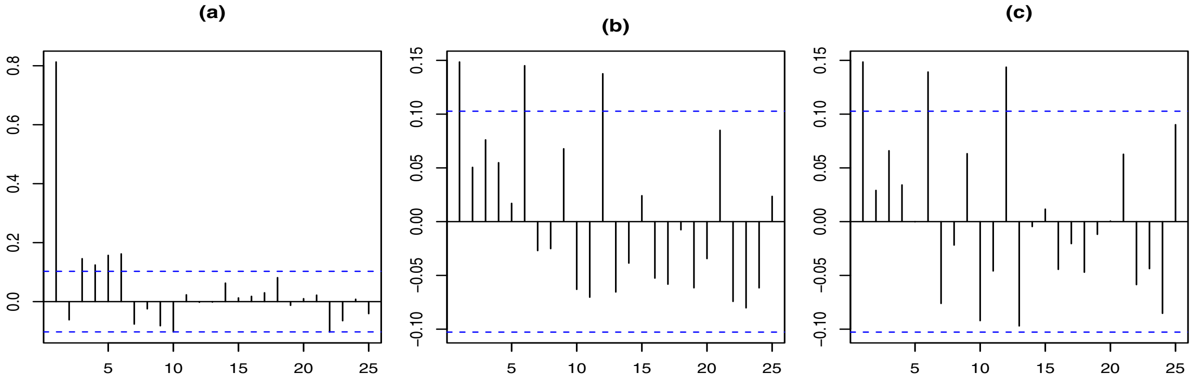

Fig. 2: (a) Sample partial autocorrelation function of (Y_t); (b) Sample autocorrelation function of (X_t^2); (c) Sample partial autocorrelation function of (X_t^2).

Finally, the parameters of our AR(6)-GARCH(1,1) model with SEP innovations are $\eta = (a_0, a_1, b_1, \gamma, \epsilon, \phi)$ where $\phi = (\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6)$. We fit this model to the in-sample data set which ranges from January 1, 2013 to December 31, 2013, totalling $T = 365$ observations. The parameters are estimated by maximizing the conditional log likelihood

$$L_T(\eta) = \sum_{t=7}^{T} \ln f\left(\frac{Y_t - \phi_1 Y_{t-1} - \cdots - \phi_6 Y_{t-6}}{\sigma_t}\right).$$

The estimates are given in Table 2 where we see that they are all statistically significant. Standard deviations are computed using the function optim in R. Figure 3 shows that the residuals (ε_t) are independent. Observe that $a_1 + b_1 < 1$ and therefore, (X_t) is strictly stationary and ergodic.

4. BACKTESTING ES

Backtesting ES consists in designing statistical tests to compare actual losses and ES calculations. While in financial applications, ES usually refers to the lower tail of the return distribution, the residuals hereby demonstrate positive skewness, see Table 1. Thus, the estimation of upside risk is of great interest to participants in electricity markets. ES examines the upper tail end of a distribution and determines the mean of the upper tail of the distribution that exceeds the value at risk. For each trading day t with $1 \leq t \leq n$, the value at risk VaR_α^t at the probability level α, $0 < \alpha < 1$, is the $(1-\alpha)$ quantile of (X_t), with a negative value corresponding to a loss. Equivalently,

$$\Pr(X_t \leq \mathrm{VaR}_\alpha^t) = 1 - \alpha. \tag{2}$$

The upside risk at the probability level α is

$$\mathrm{ES}_\alpha^t = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_p^t \, dp \tag{3}$$

and coincides with the conditional value at risk CVaR_α^t defined by

$$\mathrm{CVaR}_\alpha^t = \mathrm{E}[X_t \mid X_t > \mathrm{VaR}_\alpha^t], \tag{4}$$

since the SEP distribution is continuous, see [12].

Here, we evaluate the out-of-sample performance of our model. Upside risk is predicted one-day-ahead. The out-of-sample data set covers the time period from January 1, 2014 to February 1, 2016 and contains $n = 1127 - 365 = 762$ daily observations. The performance of the AR(6)-GARCH(1,1) model with SEP innovations is compared to the performance of AR(6)-GARCH(1,1) models whose innovations follow other distributions, such as the Johnson SU, the generalized hyperbolic and the skewed Student-t.

The ES backtesting procedure used here is the one presented in [13]. This procedure presents a coverage test for any spectral risk measure such as ES. For each trading day t with $1 \leq t \leq n$, a spectral risk measure M_ϕ^t with admissible risk measure ϕ is defined by

$$M_\phi^t = \int_0^1 \mathrm{VaR}_p^t \, \phi(dp), \tag{5}$$

Parameter	a_0	a_1	b_1	γ	ϵ	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6
Estimate	0.825	0.223	0.636	1.138	-0.052	0.842	-0.190	0.088	-0.007	0.045	0.135
Standard deviation	0.193	0.007	0.008	0.014	0.002	0.005	0.028	0.009	0.001	0.003	0.001

Table 2: Parameters of the AR(6)-GARCH(1,1) model with SEP innovations

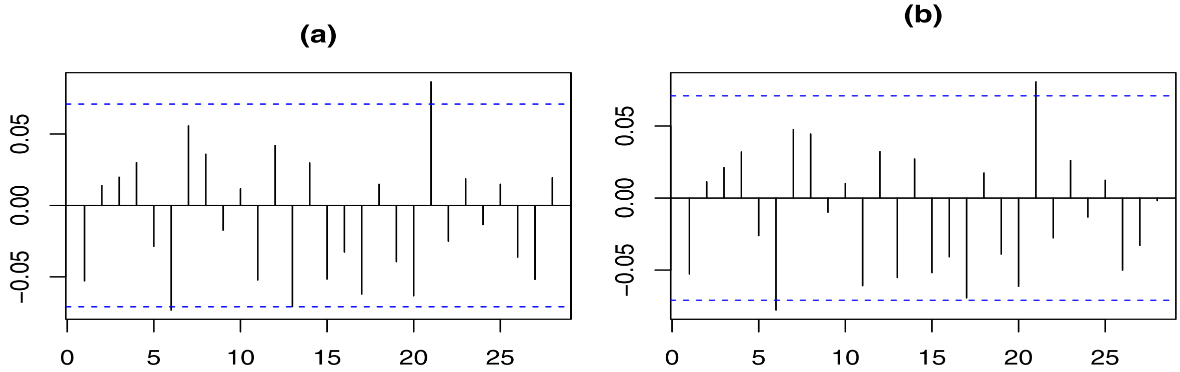

Fig. 3: (a) Sample autocorrelation function of (ε_t); (b) Sample partial autocorrelation function of (ε_t).

and the spectral risk measure failure rate is

$$\overline{X}^n_{\mathrm{SR}}(\phi) = \frac{1}{n}\sum_{t=1}^{n} X^t_{\mathrm{SR}}(\phi), \tag{6}$$

where

$$X^t_{\mathrm{SR}}(\phi) = \int_0^1 \mathbb{1}_{\{X_t \geq \mathrm{VaR}^t_p\}}\, \phi(dp).$$

The null-hypothesis for the spectral risk measure coverage test is

$$\mathrm{H}_0 : \mathsf{Pr}(X_t \leq \mathrm{VaR}^t_\alpha) = 1 - \alpha.$$

Since X_t is stricly stationary and ergodic with a finite mean, the pointwise ergodic theorem for the stationary sequence $(X^t_{\mathrm{SR}}(\phi))$ implies that

$$\overline{X}^n_{\mathrm{SR}}(\phi) \xrightarrow{a.s.} \mathrm{E}[X^t_{\mathrm{SR}}(\phi)] = \mu_\phi, \tag{7}$$

where

$$\mu_\phi = \int_0^1 p\phi(dp).$$

Under H_0 and according to the central limit theorem for stationary ergodic martingale differences [14], we have

$$Z^n_{\mathrm{SR}} = \frac{(\overline{X}^n_{\mathrm{SR}}(\phi) - \mu_\phi)}{\sigma_\phi} \xrightarrow{d} N(0,1) \tag{8}$$

where

$$\sigma^2_\phi = \frac{1}{n}\left(2\int_0^1 \phi(dp)\int_0^p q\phi(dq) - \left(\int_0^1 p\phi(dp)\right)^2\right).$$

The Z-score of Z^n_{SR} describes a Z-test for $\overline{X}^n_{\mathrm{SR}}(\phi)$ and a coverage test for M^t_ϕ.

If ϕ is the Dirac measure concentrated at α, $M^t_\phi = \mathrm{VaR}^t_\alpha$, $\mu_\phi = \alpha$ and $\sigma_\phi = \sqrt{\alpha(1-\alpha)}$. Thus, the Z-score becomes

$$Z^n_{\mathrm{VaR}_\alpha} = \sqrt{n}\frac{\overline{X}^n_{\mathrm{VaR}_\alpha} - \alpha}{\sqrt{\alpha(1-\alpha)}}. \tag{9}$$

In fact, $Z^n_{\mathrm{VaR}_\alpha}$ is the Wald variant of the likelihood ratio statistic proposed by [15, 16]. However, the Dirac measure concentrated at α is not an admissible risk measure.

Taking $\phi(p) = \frac{1}{\alpha}\mathbb{1}_{\{0 \leq p \leq \alpha\}}$, $M^t_\phi = \mathrm{ES}^t_\alpha$. The computation of μ_ϕ and σ_ϕ gives the following Z-score

$$Z^n_{\mathrm{ES}_\alpha} = \sqrt{3n}\frac{2\overline{X}^n_{\mathrm{ES}_\alpha} - \alpha}{\sqrt{\alpha(4-3\alpha)}}. \tag{10}$$

We use the model estimated in Section 3 to calculate one-day-ahead conditional means and variances for the out-of-sample data set and the ES failure rate $\overline{X}^n_{\mathrm{ES}_\alpha}$. The Z-score $Z^n_{\mathrm{ES}_\alpha}$ is computed for different values of α and for various innovation distributions. The results are presented in Table 3. We rank the performance of the models according to their Z-score and the closest Z-score to zero corresponds to the best model. We see that the SEP distribution significantly outperforms the other distributions for $\mathrm{ES}_{1\%}$, $\mathrm{ES}_{2\%}$, $\mathrm{ES}_{2.5\%}$, $\mathrm{ES}_{3\%}$ and $\mathrm{ES}_{4\%}$.

5. CONCLUSION

This paper aims to predict upside risk for the price of the Swedish bidding area SE3 using an ARMA-GARCH model with SEP innovations. The great flexibility of the SEP distribution allows to model a large class of data which are skewed

Distribution	$\mathrm{ES}_{1\%}$	$\mathrm{ES}_{2\%}$	$\mathrm{ES}_{2.5\%}$	$\mathrm{ES}_{3\%}$	$\mathrm{ES}_{4\%}$	$\mathrm{ES}_{5\%}$
SEP	**0.1944**	**0.5691**	**0.1579**	**0.8265**	**0.3913**	-0.7066
GHYP	1.2046	1.3172	0.3179	1.2861	0.8865	-0.3056
JSU	1.4048	1.2794	1.0379	2.2406	1.3580	0.3903
SSTD	1.4545	1.1206	0.8394	1.9504	1.0871	**0.1394**

Table 3: Backtesting ES for the SEP, Johnson SU (JSU), generalized hyperbolic (GHYP) and skewed Student-t (SSTD) distributions.

and fat-tailed. The SEP distribution captures well the spiky events in January 2016 and shows good performance in terms of ES estimation compared to commonly used distributions in risk measurement.

6. REFERENCES

[1] R. Weron, "Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach," *Hugo Steinhaus Center, Wroclaw University of Technology*, 2006.

[2] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.

[3] J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1014–1020, 2003.

[4] K. Ignatieva and S. Trück, "Modeling spot price dependence in Australian electricity markets with applications to risk management," *Computers & Operations Research*, vol. 66, pp. 415–433, 2016.

[5] R. C. Garcia, J. Contreras, M. Van Akkeren, and J. B. C. Garcia, "A GARCH forecasting model to predict day-ahead electricity prices," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 867–874, 2005.

[6] A. M. Gonzalez, A. M. S. Roque, and J. Garcia-Gonzalez, "Modeling and forecasting electricity prices with input/output hidden markov models," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 13–24, 2005.

[7] N. Amjady, "Day-ahead price forecasting of electricity markets by a new fuzzy neural network," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 887–896, 2006.

[8] P. Mandal, T. Senjyu, N. Urasaki, A. Yona, T. Funabashi, and A.K. Srivastava, "Price forecasting for day-ahead electricity market using recursive neural network," in *IEEE Power Engineering Society General Meeting*, 2007, pp. 1–8.

[9] P. J. Brockwell and R. A. Davis, *Time series: theory and methods*, Springer Series in Statistics. Springer-Verlag, New York, second edition, 1991.

[10] J. Janczura, S. Trück, R. Weron, and R. C. Wolff, "Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling," *Energy Economics*, vol. 38, pp. 96–110, 2013.

[11] P. Bondon and L. Song, "AR processes with non-gaussian asymmetric innovations," in *European Signal Processing Conference*, 2013, pp. 1–5.

[12] C. Acerbi, "Spectral measures of risk: A coherent representation of subjective risk aversion," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1505–1518, 2002.

[13] Nick Costanzino and Mike Curran, "Backtesting general spectral risk measures with application to expected shortfall," *Available at SSRN 2514403*, 2015.

[14] Patrick Billingsley, "The lindeberg-lévy theorem for martingales," *Proceedings of the American Mathematical Society*, vol. 12, no. 5, pp. 788–792, 1961.

[15] P. H. Kupiec, "Techniques for verifying the accuracy of risk measurement models," *The journal of derivatives*, vol. 3, no. 2, pp. 73–84, 1995.

[16] Sean D. Campbell, "A review of backtesting and backtesting procedures," Finance and Economics Discussion Series 2005–21, Board of Governors of the Federal Reserve System (U.S.), 2005.