A RECURSIVE BAYESIAN MODEL FOR EXTREME VALUES

Douglas E. Johnston[*]

Farmingdale State College
Department of Applied Mathematics
Farmingdale, N.Y., U.S.A.
douglas.johnston@farmingdale.edu

Petar M. Djurić[†]

Stony Brook University
Department of Electrical and Computer Engineering
Stony Brook, N.Y., U.S.A.
petar.djuric@stonybrook.edu

ABSTRACT

In this paper, we propose a new approach for analyzing extreme values such as large losses in financial markets. Our goal is to compute the predictive distribution of extreme events that are clustered in time. We apply a stochastic parametrization of the generalized extreme value distribution to model the asymptotic behavior of the block-maximum and derive a Rao-Blackwellized particle filter. This reduces the parameter space, and we derive a concise, recursive solution. Using the filter, the predictive distribution, conditioned on the past data, is computed at each sample-time. We introduce a new risk-measure, p_{VaR_α}, that is a more robust estimate of the true nature of value-at-risk, and illustrate our results using both simulated data and actual stock market returns from 1928-2017.

***Index Terms*—** extreme value theory, particle filter, risk-management, VaR

1. INTRODUCTION

The modeling and statistical analysis of extreme events is a critical mission in many signal processing applications such as communication systems [1], image analysis [2], and noise cancellation [3]. Historically, extreme value analysis was developed around hydrological studies [4] but, today, it is of keen interest in the field of finance where large financial losses can lead to ruin [5]. Specifically, estimating the tail of the predictive distribution for the maximum loss over a period of time is of paramount importance.

In this paper, we introduce a new approach for modeling dynamic extreme events that exhibit long-term dependency. Utilizing the Fisher-Tippet-Gnedenko Theorem [5], we obtain a parametric form for the asymptotic general extreme value (GEV) distribution of the maximum of a block of data – the block-maximum. To allow for dependencies, such as volatility clustering, we assume the parameters of the distribution are a hidden stochastic process which results in a non-linear, non-Gaussian state-space model with unknown static parameters. To estimate the parameters of the GEV, we utilize a particle filter (PF) [6]. In particular, given the additional static parameters, we derive a Rao-Blackwellized particle filter (RBPF) [7] to marginalize these unknown, static parameters. In doing so, we derive a recursive solution for the predictive density of the GEV parameters and block-maxima. The latter is particularly important when using risk-metrics, which are typically non-linear functions of the cumulative predictive distribution.

The work presented here is an extension of some recent studies. In [8, 9], the innovations are modeled as an ARMA process. In [10], a deterministic trend is applied to a subset of the GEV parameters, while in [11, 12, 13], a GEV parameter subset is dynamic with known state equation. Most recently, [14] developed a dynamic model for the shape parameter and a Gaussian-mixture approximation to linearize the estimation problem. While the present study is related to this body of work, we have extended it in a number of ways. First, under a reasonable assumption, we reduce the number of GEV parameters and model the remaining set as a vector Markov process, with unknown system and covariance matrix. Second, we recursively compute the marginalized state density, eliminating the need to estimate unwanted nuisance parameters and, in the process, we derive recursive expressions for the necessary sufficient statistics. This allows for a fast, real-time implementation without the need to batch-process observations. Lastly, we derive the predictive distribution, for the block-maxima, that we can use to estimate p-values for risk-metrics.

The paper is organized as follows. In the next section, we formulate the problem for the time-varying block-maxima. In Section 3, we propose a solution to the problem, deriving the log-likelihood function of the block-maxima, given the density parameters, and the recursive solution for marginalizing the nuisance parameters - Rao-Blackwellization. We illustrate the efficacy of our method in Section 4 using simulated data and S&P 500 stock market returns from 1928 through the end of 2017. In addition, we introduce a new risk-metric, p_{VaR_α}, and highlight its usefulness. Lastly, we conclude the paper with ideas for further research.

2. PROBLEM STATEMENT

Let $y_k \in \mathbb{R}, k = 1, \cdots, N$ be the kth block-maximum, with block size B_k, for an underlying strictly stationary process, s_t. That is

$$y_k = \max_{N_{k-1} < t \leq N_k} s_t, \tag{1}$$

where $(N_{k-1}, N_k]$ are the time indices for the kth block ($k \geq 1$) and $N_0 = 0$. In our case, $s_t \in \mathbb{R}$ and $s_t = -r_t$, where r_t is the daily return for the S&P 500 index observed for $t \in \mathbb{N}$. The return is defined as the percentage change in the price, p_t, of an asset or $r_t = (p_t - p_{t-1})/p_{t-1}$. Thus, a large value of s_t is a large daily loss in percentage terms. In our study, block sizes are the number of trading days in a year (≈ 252) and we model the largest daily loss (negative return) for the S&P 500 over the course of a year (see Figure 1.)

The Fisher-Tippet-Gnedenko (FTG) theorem states that the only non-degenerate limiting distribution for the normalized block-maximum of i.i.d. random variables (RVs) are in the generalized

[*]D.E.J. thanks FSC and Dr. Carlos Marques for their support.
[†]P.M.D. thanks the support of NSF under Award CCF-1618999.

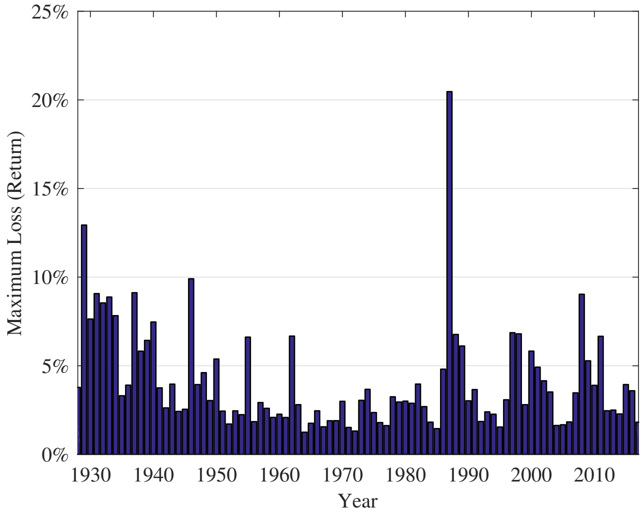

Fig. 1. Annual block maximum of S&P 500 return loss.

extreme value (GEV) family [15] with parameter ξ. One form of the GEV cumulative distribution function (cdf) is

$$H_\xi(y) = \exp(-(1+\xi y)^{-1/\xi}); \;\; \xi > 0, \; y \geq -1/\xi, \qquad (2)$$

where ξ is the shape parameter of the distribution.

With ξ being strictly positive, as in (2), the distribution is the standard Fréchet distribution which is the limiting class for underlying heavy-tailed distributions that exhibit regular variation [16] and are typically employed in finance. A more general form of the distribution has three parameters, a location parameter $\mu \in \mathbb{R}$, a scale parameter $\beta > 0$, and the shape parameter $\xi > 0$. Thus, we asymptotically model each of the block-maxima $y \sim H_\xi(\frac{y-\mu}{\beta})$, where $y \in [\mu - \beta/\xi, \infty)$. While the FTG theorem was originally stated for i.i.d. rvs, it holds for most strictly stationary processes [5] with the so-called extremal index absorbed by the parameter β.

We propose to model the three parameters of the GEV distribution, ξ, β, μ, as a stochastic process. In order to insure positivity, we define a state vector $\mathbf{x}_k = [\log(\xi_k) \;\; \log(\beta_k) \;\; \mu_k]^\top)$ and the state equation

$$\mathbf{x}_k = \boldsymbol{\eta} + \boldsymbol{\Theta}\mathbf{x}_{k-1} + \mathbf{C}^{1/2}\mathbf{u}_k, \qquad (3)$$

where $\mathbf{x}_k \in \mathbb{R}^3$, $k \in \mathbb{N}_0$, $\mathbf{u}_k \sim \mathcal{N}(0, \mathbf{I}_3)$, $\boldsymbol{\eta} \in \mathbb{R}^3$, and $\boldsymbol{\Theta}, \mathbf{C} \in \mathbb{R}^{3\times 3}$. The unknowns are $\mathbf{x}_k, \boldsymbol{\eta}, \boldsymbol{\Theta}$, and $\mathbf{C}$.

The objective is to determine the predictive distribution of y_{k+1}, i.e., $p(y_{k+1}|y_{1:k})$, where $y_{1:k} \equiv \{y_1, y_2, \cdots, y_k\}$ represents the set of past block-maximum observations. We formally obtain the predictive distribution from

$$p(y_{k+1}|y_{1:k}) = \int p(y_{k+1}|\mathbf{x}_{k+1}) p(\mathbf{x}_{k+1}|y_{1:k}) d\mathbf{x}_{k+1}, \qquad (4)$$

where the likelihood function is $p(y_k|\mathbf{x}_k) = \frac{d}{dy_k} H_{\xi_k}(\frac{y_k - \mu_k}{\beta_k})$ and $p(\mathbf{x}_{k+1}|y_{1:k})$ is the predictive distribution of $\mathbf{x}_{k+1}$ based on past observations.

3. PROPOSED METHOD

A reasonable, and simplifying, assumption is that the block-maximum, y_k, has support $[0, \infty)$ which allows the state vector to be reduced as $\mu_k = \beta_k/\xi_k$. The implication, which is true in our case, is the S&P 500 has a non-negative loss at least one day each year. We can then specify the observation as a scaled and translated GEV RV

$$y_k = e^{x_{2_k} - x_{1_k}} + e^{x_{2_k}} w_k, \qquad (5)$$

where $e^{x_{1_k}} = \xi_k$, $e^{x_{2_k}} = \beta_k$, and w_k is distributed according to (2). Combined, equations (3) and (5) form our non-linear state-space model reminiscent of stochastic-volatility models that are also used in finance [17].

With particle filtering [18], an approximation for the predictive density of the state, $p(\mathbf{x}_{k+1}|y_{1:k})$, is

$$p(\mathbf{x}_{k+1}|y_{1:k}) \approx \sum_{m=1}^{M} w_k^{(m)} p(\mathbf{x}_{k+1}|\mathbf{x}_{0:k}^{(m)}), \qquad (6)$$

where $\mathbf{x}_{0:k}^{(m)}$ is the mth particle (vector) stream and $w_k^{(m)}$ is the associated weight. Assuming equal weights, the particles are samples from $p(\mathbf{x}_{0:k}|y_{1:k})$ that recursively evolves according to

$$p(\mathbf{x}_{0:k}|y_{1:k}) \propto \{p(y_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{x}_{0:k-1})\} \; p(\mathbf{x}_{0:k-1}|y_{1:k-1}). \quad (7)$$

At this point, the stumbling block is the set of unknown, static parameters in the state equation (3). Our approach is to marginalize $p(\mathbf{x}_{k+1}|\eta, \boldsymbol{\Theta}, \mathbf{C}, \mathbf{x}_{0:k})$ and implement a Rao-Blackwellized particle filter. For an efficient solution, we need a recursive expression for

$$\int_{\boldsymbol{\eta}, \boldsymbol{\Theta}, \mathbf{C}} p(\mathbf{x}_{k+1}|\boldsymbol{\eta}, \boldsymbol{\Theta}, \mathbf{C}, \mathbf{x}_{0:k}) dP(\boldsymbol{\eta}, \boldsymbol{\Theta}, \mathbf{C}|\mathbf{x}_{0:k}). \qquad (8)$$

The resulting predictive density, $p(\mathbf{x}_{k+1}|\mathbf{x}_{0:k})$, which is used to extrapolate particles, can be solved analytically for the case of the general linear model, with a non-informative prior, and results in a multivariate Student-t distribution [19]. Importantly, we show that the necessary sufficient statistics, the mean vector and scale matrix, can be recursively solved so that a computationally efficient implementation is derived.

Our problem can be stated as a general linear model

$$\mathbf{X}_k = \boldsymbol{\Phi}\mathbf{H}_k + \mathbf{U}_k, \qquad (9)$$

where $\mathbf{X}_k = [\mathbf{x}_k, \mathbf{x}_{k-1}, \cdots, \mathbf{x}_1] \in \mathbb{R}^{p\times k}$, $\boldsymbol{\Phi} \in \mathbb{R}^{p \times q}$ is an unknown matrix, $\mathbf{H}_k = [\mathbf{h}_k, \mathbf{h}_{k-1}, \cdots, \mathbf{h}_1] \in \mathbb{R}^{q\times k}$ has columns given by $\mathbf{h}_k^\top = [1 \;\; \mathbf{x}_{k-1}^T]$, and $\mathbf{U}_k \in \mathbb{R}^{p\times k}$ is a random Normal matrix whose columns are i.i.d. $\mathcal{N}(\mathbf{0}, \mathbf{C})$ with $\mathbf{C}$ unknown. In our case, the columns of $\mathbf{X}_k$ are formed by the particles from 1 to k, $\boldsymbol{\Phi} = [\;\boldsymbol{\eta} \quad \boldsymbol{\Theta}\;]$, and the columns of $\mathbf{H}_k$ include the past particles from 0 to $k-1$. Therefore, (9) is the state transition equation that describes the evolution of the particle vector stream, where each particle stream operates under its own version of this multivariate regression model.

The multivariate Student-t distribution of a p-dimensional random vector is denoted as $t_p(v, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto [1 + (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})/v]^{-(v+p)/2}$ with v degrees of freedom, mean $\boldsymbol{\mu}$, and scaling matrix $\boldsymbol{\Sigma}$. In our case, $p = 2$ and $q = p + 1$. When we use the non-informative prior, $p(\boldsymbol{\eta}, \boldsymbol{\Theta}, \mathbf{C}^{-1}) \propto |\mathbf{C}|^{(p+1)/2}$, we can write the marginal predictive density for $\mathbf{x}_{k+1}$ as a multivariate Student-t distribution with $v_k = n - 2p$ degrees of freedom, i.e.,

$$p(\mathbf{x}_{k+1}|\mathbf{X}_k, \mathbf{H}_k, \mathbf{h}_{k+1}) \sim t_p(v_k, \hat{\mathbf{x}}_{k+1}, \hat{\boldsymbol{\Sigma}}_{k+1}). \qquad (10)$$

The mean of this distribution is derived as,

$$\hat{\mathbf{x}}_{k+1} = \mathbf{S}_{XH}^k (\mathbf{S}_{HH}^k)^{-1} \mathbf{h}_{k+1} = \hat{\boldsymbol{\Phi}}_k \mathbf{h}_{k+1}, \qquad (11)$$

and scale matrix as

$$\hat{\mathbf{\Sigma}}_{k+1} = \frac{1 + \mathbf{h}_{k+1}^\top (\mathbf{S}_{HH}^k)^{-1} \mathbf{h}_{k+1}}{v_k} (\mathbf{S}_{XX}^k - \hat{\mathbf{\Phi}}_k {\mathbf{S}_{XH}^k}^\top), \quad (12)$$

with the matrices defined as

$$\mathbf{S}_{XX}^k = \mathbf{X}_k \mathbf{X}_k^\top, \quad \mathbf{S}_{XH}^k = \mathbf{X}_k \mathbf{H}_k^\top, \quad \mathbf{S}_{HH}^k = \mathbf{H}_k \mathbf{H}_k^\top. \quad (13)$$

Letting $\mathbf{P}_k = (\mathbf{S}_{HH}^k)^{-1}$, we derive recursive expressions for the mean and the scale matrix as follows:

$$\mathbf{K}_k = \mathbf{h}_k^\top \mathbf{P}_{k-1} [1 + \mathbf{h}_k^\top \mathbf{P}_{k-1} \mathbf{h}_k]^{-1}, \quad (14)$$

$$\mathbf{P}_k = \mathbf{P}_{k-1} [\mathbf{I} - \mathbf{h}_k \mathbf{K}_k], \quad (15)$$

$$\hat{\mathbf{\Phi}}_k = \hat{\mathbf{\Phi}}_{k-1} + (\mathbf{x}_k - \hat{\mathbf{x}}_k) \mathbf{K}_k, \quad (16)$$

$$\hat{\mathbf{\Sigma}}_{k+1} = \frac{v_k - 1}{v_k} \left\{ \frac{1 + \mathbf{h}_{k+1}^\top \mathbf{P}_k \mathbf{h}_{k+1}}{1 + \mathbf{h}_k^\top \mathbf{P}_{k-1} \mathbf{h}_k} \right\} \times \left\{ \hat{\mathbf{\Sigma}}_k + (\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^\top / (v_k - 1) \right\}, \quad (17)$$

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{\Phi}}_k \mathbf{h}_{k+1}. \quad (18)$$

In our case, $\mathbf{x}_k (= \mathbf{x}_k^{(m)})$ is the mth particle vector that is available after updating the particle filter with the measurement at time k and $\mathbf{h}_{k+1}^\top = [1 \ \mathbf{x}_k^\top]$, which is needed to update both the mean and scale matrix, is also available. With this result, we can readily implement the RB particle filter (RBPF) [20]. For each particle stream, we retain the sufficient statistics, as defined above, recursively estimate the mean, $\hat{\mathbf{x}}_{k+1}$, and scale matrix, $\hat{\mathbf{\Sigma}}_{k+1}$, and then sample $\mathbf{x}_{k+1}^{(m)} \sim t_p(v_k, \hat{\mathbf{x}}_{k+1}, \hat{\mathbf{\Sigma}}_{k+1})$ to generate new particles over time, between observations. The particles are then updated (re-sampled) according to the likelihood function. The log-likelihood function is given by (here we drop the time subscript for simplicity)

$$\ln(p(y|\mathbf{x})) = -x_2 - (ye^{x_1 - x_2})^{-e^{-x_1}} - (1 + e^{-x_1}) \ln(ye^{x_1 - x_2}), \quad (19)$$

and to initialize the filter, we use random samples generated from (9) and the batch equations (11) - (13) to initialize the requisite filter parameters.

Equipped with the RBPF approximation for the state predictive density, $p(\mathbf{x}_{k+1}|y_{1:k})$, as in (6), we may compute the predictive density of y_{k+1} conditioned purely on past observations $y_{1:k}$ via (4). That said, we require the cdf of this predictive density

$$F(y_{k+1}|y_{1:k}) = \int_{-\infty}^{y_{k+1}} p(\zeta|y_{1:k}) d\zeta, \quad (20)$$

which we can readily obtain as

$$F(y_{k+1}|y_{1:k}) = E_{\mathbf{x}_{k+1}|y_{1:k}} [H_{\mathbf{x}_{k+1}}(y_{k+1})]. \quad (21)$$

The cumulative predictive distribution is the expected GEV distribution conditioned on the state predictive density $p(\mathbf{x}_{k+1}|y_{1:k})$.

4. RESULTS

To begin, we simulated the stochastic GEV parameters from an uncoupled, mean-reverting, AR(1) process as follows

$$\log \xi_k = (1 - .95) \times \log(.45) + .95 \log \xi_{k-1} + .1 u_{1,k}, \quad (22)$$

$$\log \beta_k = (1 - .9) \times \log(.012) + .9 \log \beta_{k-1} + .1 u_{2,k}, \quad (23)$$

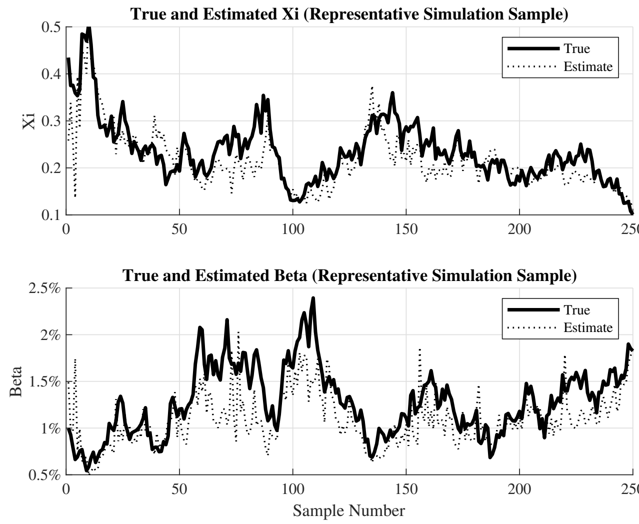

Fig. 2. True and estimated GEV parameters for one run.

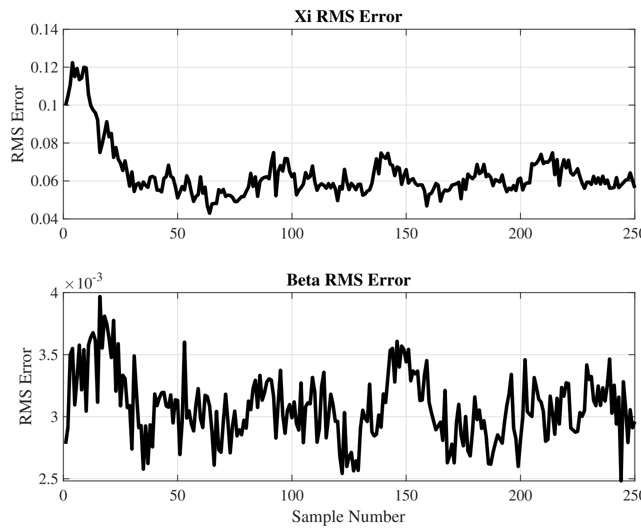

Fig. 3. RMS estimation error for GEV parameters over 100 runs.

where $u_{1,k}$ and $u_{2,k}$ are independent $\mathcal{N}(0, 1)$. We produced 250 samples for each simulation run and we used $M = 5000$ particles. At each point in time we computed estimates for the GEV parameters as $\hat{\mathbf{x}}_k = \sum_{m=1}^M w_k^{(m)} \mathbf{x}_k^{(m)}$. A representative simulation run is shown in Figure 2 and, for this example, the filter tracks the true parameters quite well.

We performed 100 simulation runs and we computed an RMS error at each sample time across the 100 runs. The RMS errors are shown in Figure 3. After an initial period of roughly 30 samples, the filter performance stabilizes and appears reasonable. The average RMS error for ξ, after the initial convergence period, was about .06 which is 13% of the underlying process mean of .45. Similarly, for β, the RMS error was .3% compared to its mean value of 1.2%. In all, the simulation gives us confidence in the approach and performance of the RBPF.

We also ran the filter on the S&P 500 block-maxima data, shown in Figure 1, which are the $N = 90$ maximum annual daily losses (negative return) for the years 1928-2017. We computed the predic-

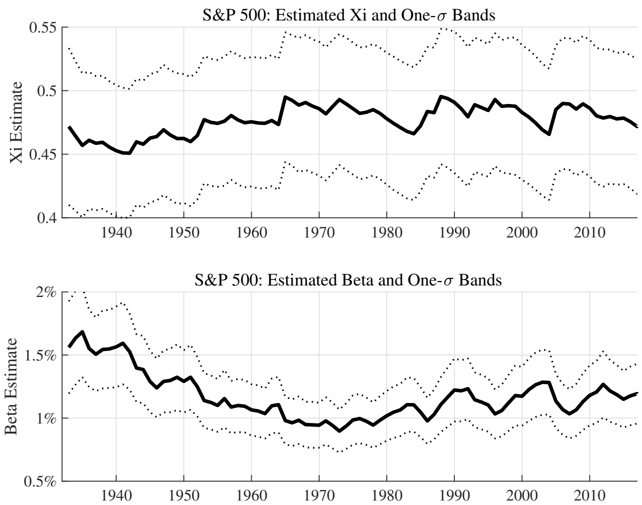

Fig. 4. GEV parameter estimates for S&P 500 data

tive estimates for our GEV parameters, $\hat{\mathbf{x}}_{k+1}$ using solely the observations up to time k. In Figure 4, we show the estimates for the GEV parameters along with computed one standard deviation bands. The shape parameter estimate, $\hat{\xi}$, appears stable over time ranging from .45 to .5. We should note that a RV whose block-maxima is in the Fréchet limit has finite Kth moment for $\xi < 1/K$. Thus, ex-post, there is a decent chance that the underlying process driving daily stock market returns has infinite variance. The scale parameter estimate, $\hat{\beta}$, appears more dynamic ranging from 1% to 1.5% with visible trends and oscillations.

To test the efficacy of our approach, we compared the actual S&P 500 block-maximum loss to quantiles predicted by the model, using (21), which is shown in Figure 5. The actual losses exceeded the predictive median 54% of the time and the 90'th percentile 8% of the time. In addition, we applied a chi-square test as described in [21]. At each time k, we partition the support of y_{k+1} into J equiprobable intervals, $\mathcal{S}_j, j = 1, 2, \cdots, J$ using the predictive cdf $F(y_{k+1}|y_{1:k})$. We then compare the $(k+1)$st observation, y_{k+1}, to each interval to find the index j such that $y_{k+1} \in \mathcal{S}_j$. In other words, we quantize the observations into model-based, equiprobable intervals to create a frequency distribution computed as

$$\widehat{p}_{k+1}(j) = \frac{1}{k+1}(k \times \widehat{p}_k(j) + I_{y_{k+1} \in \mathcal{S}_j}), \quad j = 1, 2, \cdots, J, \quad (24)$$

where I_x is the indicator function and $\widehat{p}_0(\cdot) = 0$. If the model is true, $\widehat{p}_N$ will be a sample estimate of a uniform distribution and the statistic $D_N = J(N-1)\sum_{j=1}^{J}(\widehat{p}_N(j) - 1/J)^2$ is approximately chi-square with $J-1$ degrees of freedom. For our test, we chose $J = 20$ intervals and, with $N = 89$, the p-value from the chi-square test was .96 giving confidence in the model's predictive power.

In finance, risk-metrics are derived from the tail of the predictive cdf. A widely used risk-metric is value-at risk (VaR_α) which is the α-quantile of the loss distribution, $H_{\mathbf{x}}^{-1}(\alpha)$. Traditionally, the maximum likelihood (ML) estimates for the GEV parameters are used and $H_{\hat{\mathbf{x}}_{ML}}^{-1}(\alpha)$ is computed. However, a critical issue is that risk-metrics are unknown quantities and using ML estimates for the parameters, or even the predictive cdf $F(y_{k+1}|y_{1:k})$, can be insufficient to gauge risk. This is particularly true since the density of VaR_α is asymmetric with positive skew.

To improve risk-assessment, we introduce a new risk-metric, p_{VaR_α}, which is the threshold such that Prob($\mathrm{VaR}_\alpha > p_{\mathrm{VaR}_\alpha}$) < $1-p$. In Figure 6, we show different $\mathrm{VaR}_{99\%}$ statistics over time using the RBPF. Shown is the Expected VaR, E($\mathrm{VaR}_{99\%}$), which at time k is $E_{\mathbf{x}|y_{1:k-1}}[H_{\mathbf{x}}^{-1}(99\%)]$, as well as the $50\%_{\mathrm{VaR}_{99\%}}$ and the $90\%_{\mathrm{VaR}_{99\%}}$. As of the end of 2017, the $90\%_{\mathrm{VaR}_{99\%}} \approx 28\%$ implying model-based confidence, at the 90% level, that the largest daily loss in 2018 will not exceed that level. This is in contrast to the $50\%_{\mathrm{VaR}_{99\%}}$ = 20.4% and E($\mathrm{VaR}_{99\%}$) = 21.3%. The latter indicates a skew to the $\mathrm{VaR}_{99\%}$ distribution, which becomes even more severe at higher quantiles.

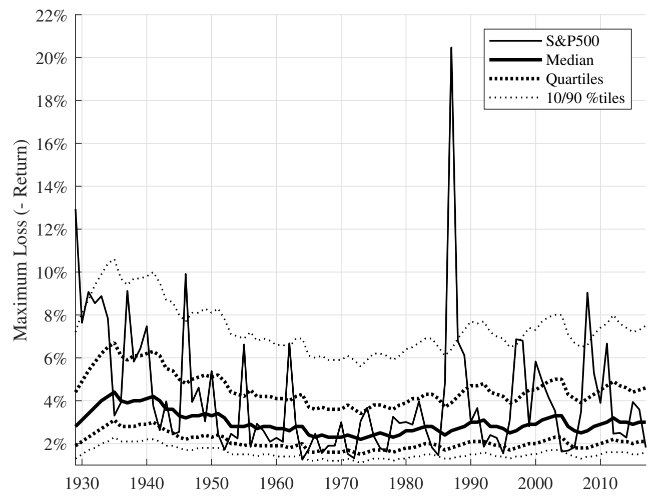

Fig. 5. S&P 500 max annual loss and predicted quantiles

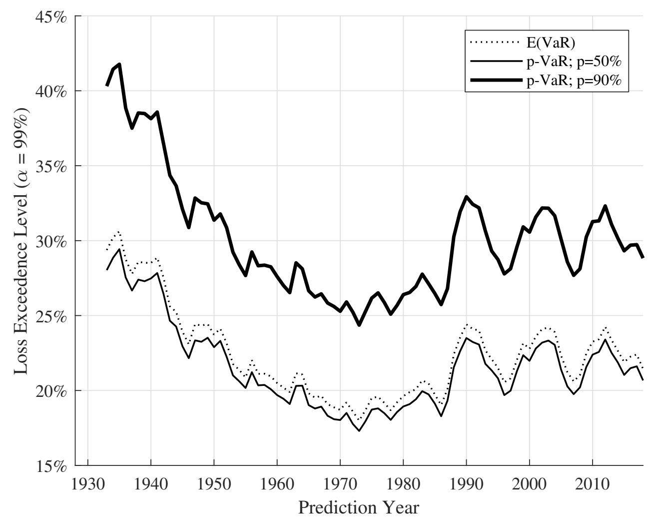

Fig. 6. $E[\mathrm{VaR}_{99\%}]$, $50\%_{\mathrm{VaR}_{99\%}}$, and $90\%_{\mathrm{VaR}_{99\%}}$ estimates

5. CONCLUSIONS

In this paper we proposed a new approach for analyzing extreme events in financial markets along with a new risk-metric, p_{VaR_α}, that more adequately describes risk. We derived the analytical solutions and a recursive implementation via RB particle filtering. We tested our model using both simulated and actual returns data and statistical tests indicate the performance of the model is quite good. In future research, we will study the effect of different block sizes, daily threshold exceedances, as well as the Cramér-Rao bound.

6. REFERENCES

[1] S. I. Resnick and H Rootzeń, "Self-similar communication models and very heavy tails," *Ann. Appl. Probab.*, vol. 10, no. 3, pp. 753–778, August 2000.

[2] S. J. Roberts, "Extreme value statistics for novelty detection in biomedical data processing," *IEE Proc. Sci., Meas. Technol.*, vol. 147, no. 6, pp. 363–367, November 2000.

[3] G. A. Tsihrintzis and C. L. Nikias, "Fast estimation of the parameters of alphastable impulsive interference," *IEEE Trans. Signal Processing*, vol. 44, no. 6, pp. 1492–1503, June 1996.

[4] H. Hurst, "Long-term storage capacity of reservoirs," *Trans. Amer. Soc. Civil Eng.*, vol. 116, no. 1, pp. 770–799, 1951.

[5] A. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management*, Princeton Univ. Press, Princeton, NJ, 2005.

[6] M. S. Arulampalam, N. Gordon S. Maskell, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, February 2002.

[7] R. Chen, X. Wang, and J. S. Liu, "Adaptive joint detection and decoding in flat fading channels via mixture Kalman filtering," *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 2079–2094, September 2000.

[8] T. Kunihama, Y. Omori, and Z. Zhan, "Efficient estimation and particle filter for max-stable processes," *Journal of Time Series Analysis*, vol. 33, no. 1, pp. 1324–1338, 2012.

[9] J. Nakajima, T. Kunihama, and Y Omori, "Bayesian modeling of dynamic extreme values: Extension of generalized extreme value distributions with latent stochastic processes," *Journal of Applied Statistics*, vol. 44, no. 7, pp. 1248–1268, July 2017.

[10] Y. Hundecha, A. St-Hilaire, T.B.M.J. Quarda, and S. El Adlouni, "A nonstationary extreme value analysis for the assessment of changes in extreme annual wind speed over the Gulf of St. Lawrence, Canada," *Environmetrics*, vol. 24, no. 1, pp. 51–62, February 2013.

[11] G. Huerta and B. Sansó, "Time-varying models for extreme values," *Environ. and Ecol. Statistics'*, vol. 14, no. 3, pp. 285–299, September 2007.

[12] G. Toulemonde, A. Guillou, and P. Naveau, "Particle filtering for Gumbel-distributed daily maxima of methane and nitrous oxide," *Journal of Applied Meteorology and Climatology*, vol. 47, pp. 2745–2759, November 2008.

[13] Y. Wei and G. Huerta, "Dynamic generalized extreme value modeling via particle filters," *Commmications in Statistics*, vol. 46, no. 8, pp. 6324–6341, March 2017.

[14] G. Mao and Z. Zhang, "Stochastic tail index model for high frequency financial data with Bayesian analysis," *Journal of Econometrics*, vol. 205, no. 2, pp. 470–487, August 2018.

[15] P Embrechts, C. Kluppelberg, and T. Mikosch, *Modelling Extremal Events*, Springer-Verlag, New York, 2003.

[16] S. Resnick, *Heavy Tail Phenomena: Probabilistic and Statistical Modeling*, Springer-Verlag, New York, 2008.

[17] E. Jacquier, N. G. Polson, and P. E. Rossi, "Bayesian analysis of stochastic volatility models," *Journal of Business and Economic Statistics*, vol. 12, pp. 371–417, 1994.

[18] P. M. Djurić and M. F. Bugallo, "Particle filtering," in *Adaptive Signal Processing*, T. Adali and S. Haykin, Eds., pp. 271–331. Wiley & Sons, 2010.

[19] S. Geisser, "Bayesian estimation in multivariate analysis," *The Annals of American Statistics*, vol. 36, no. 1, pp. 150–159, February 1965.

[20] D. E. Johnston, I. Urteaga, and P. M. Djurić, "Replication and optimization of hedge fund risk factor exposures," in *the Proceedings of ICASSP*, Vancouver, BC, Canada, 2013.

[21] P. M. Djurić, M. Khan, and D. E. Johnston, "Particle filtering of stochastic volatility modeled with leverage," *IEEE Jounal of Selected Topics in Signal Processing*, vol. 6, no. 4, pp. 327–336, August 2012.