A News Event-Driven Approach for the Historical Value at Risk Method

Frederik Hogenboom[a,*], Michael de Winter[a], Flavius Frasincar[a], Uzay Kaymak[b]

[a] *Erasmus University Rotterdam, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands*
[b] *Eindhoven University of Technology, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands*

Abstract

Value at Risk (VaR) is a tool widely used in financial applications to assess portfolio risk. The historical stock return data used in calculating VaR may be sensitive to rare news events that occur during the sampled period and cause trend disruptions. Therefore, in this paper, we measure the effects of various news events on stock prices. Subsequently, we identify irregular events using a Poisson distribution, and we examine whether VaR accuracy can be improved by considering news events as an additional input in the calculation. Our experiments demonstrate that VaR predictions for rare event occurrences can be improved by removing the event-generated disturbance from the stock prices for a small, optimized time window.

Keywords: Value at Risk, News events, Historical method

1. Introduction

Over the years, Value at Risk (VaR) has become a widely adopted risk measure in the financial world and is now also a requirement for regulatory purposes despite its acknowledged limitations in terms of interpretability and mathematical properties (Artzner et al., 1999; Rockafellar & Uryasev, 2002). Such alternatives as the expected shortfall or conditional VaR (CVaR) (Rockafellar & Uryasev, 2000; Street, 2010), which measure the market risk of a portfolio and are more sensitive to the tail of the loss distribution than conventional VaR, have better properties but have not yet become standards. VaR is typically used in the field of finance to quantify the risk of loss on a portfolio of financial equities, and it is defined as a threshold value such that the probability that the loss on the portfolio over a given time horizon does not exceed a certain value at a given confidence level (Olson & Wu, 2010).

Lately, VaR has been criticized for its vulnerability in times of financial crisis (Asche et al., 2013). Research has shown that VaR estimation for emerging markets is difficult during periods of financial turmoil,

*Corresponding author; tel: +31 (0)10 408 1262; fax: +31 (0)10 408 9162

Email addresses: fhogenboom@ese.eur.nl (Frederik Hogenboom), m.r.dewinter88@live.nl (Michael de Winter), frasincar@ese.eur.nl (Flavius Frasincar), u.kaymak@tue.nl (Uzay Kaymak)

as the forecasts of most models tend to be overly conservative (Dimitrakopoulos et al., 2010). The estimates quickly become inaccurate, particularly if asset prices are highly correlated, which is the case in the oil and gas industry, among others (Asche et al., 2013). For developed markets, these effects appear to be less of an issue. From these recent discussions, we deduce that practitioners who apply VaR for risk calculations generally assume that there are no unexpected trend breaks in portfolio prices. In reality, we are regularly faced with deviations from trends that are mainly caused by emerging events, which are reported in news messages. Emerging events can be related not only to crisis situations, such as announcements of losses or even bankruptcies, but also to times of prosperity, such as profit announcements and acquisitions. All of these events have the potential to greatly impact today's financial markets, as they disrupt trendspositively or negativelyand can thus cause traders to react, seeking opportunities or minimizing daily losses.

According to the weak form of the efficient market hypothesis, news that contains information on an equity is not perfectly incorporated in the price when it is published. Studies have reported on the existence of such a delay (Fama, 1965), caused by initial over- or underreactions to the news. Additionally, news events affect the volatility of equities (Mitchell & Mulherin, 1994). The usage of information extracted from text in a financial context has proven to be a vital strategy in many financial applications (Chan, 2003; Ikenberry & Ramnath, 2002). Thus, considering news events in VaR calculations (which are based on returns distributions) could be beneficial, as volatility is the standard deviation of the distribution of returns (Byström, 2009; Engelberg & Parsons, 2009; Goonatilake & Herath, 2007; Kalev et al., 2004). It would be useful for traders to react to these news events in a timely and accurate fashion before the competition and incorporate an additional news input into the mostly numerical high-frequency trading algorithms used today. One could account for these destabilizing effects in equity stock prices in other calculations that use prices as input, such as VaR predictions. However, before we can implement such strategies, it is important to first analyze the effects of specific news events on stock prices.

In this paper, we measure the effects of various news events on stock prices. Additionally, we hypothesize that we can improve VaR computations by introducing financial news events (Borsje et al., 2010; Hogenboom et al., 2011, 2013) as an additional input. In our experiments, we use the proprietary ViewerPro (Semlab, 2013) software for the extraction of 2010 and 2011 ticker data and news events for different equities. Using a Poisson distribution, we identify the irregular events. Subsequently, we clean the ticker data of rare event-generated noise and obtain a dataset that is a more accurate representation of the expected returns distribution. We also seek to optimize the time window for which the cleaning is conducted by evaluating the accuracy of the calculated VaR for different configurations. Although the time span covered by the dataset is associated with a financial crisis, this fact should have no effect on our experiments. As stated earlier, for highly correlated stock prices, estimates may become inaccurate during times of economic crisis; however, in our dataset, we sought to have a set of companies at different stages in their economic life cycles, circumventing such negative effects. Moreover, although many events can potentially be discovered that could disrupt VaR predictions, this is also the case in times of economic expansion. The proposed methods are defined irrespective of the state of the economy and merely depend on emerging events, independent of the economic situation for which the data are collected.

This paper is a continuation of previous and ongoing efforts to improve VaR calculations (Hogenboom et al., 2012a,b). Hogenboom et al. (2012b) used a fixed time window to consider the effects of an event on stock prices. Additionally, all events in the dataset were considered to have potentially destabilizing effects on stock prices, whereas in (Hogenboom et al., 2012a), a Poisson distribution was introduced to distinguish rare events from frequently occurring events that are not likely to influence stock prices due to their regularity over time. Our current endeavors expand on this previous work by also evaluating the impact of various types of news events on stock prices on a microeconomic level. Such an analysis is necessary because news events are always related to specific equities and cannot be categorized as general market events, such as crises and consumer confidence index adjustments. After establishing the relation between news events and stock prices, we present an extension of the historical VaR method that accounts for relevant news events.

The remainder of the paper is organized as follows. First, we describe approaches related to this research in Section 2. Then, we investigate the effects of various news events on stock prices in Section 3, and we describe our proposed method for considering news events in historical VaR calculations in Section 4. In Section 5, we present an evaluation of the proposed method. Finally, in Section 6, we present our conclusions and identify directions for future work.

2. Related Work

VaR has been widely studied as a measure of the risk of loss on a specific portfolio of financial assets, represented as a single number (Holton, 2003). VaR has become widely used in practice by corporate treasurers and fund managers. It is also used by regulators to determine the capital that financial institutions are required to have to cover their risks (Hull, 2011). According to its classical definition, for a given portfolio, probability, and time horizon, VaR can be formally described as a threshold value such that the probability of a VaR break, i.e., the loss on the portfolio over a given time horizon, exceeds this threshold values at a given probability level (Jorion, 2006). For this computation, it is assumed that there is a normal market and that no trading takes place in the portfolio.

For VaR calculations, one can distinguish among three main methods. First, the parametric method assumes a specific distribution of equity returns. Commonly employed distributions for the parametric method are the normal distribution and the log-normal distribution, which offer simplicity while maintaining robustness. However, in practice, equity returns are almost never normally distributed (Andersen et al., 2001). Second, the Monte Carlo simulation-based method predicts future returns by fitting a distribution based on historical data. In contrast to the parametric method, Monte Carlo simulation does not assume a normal distribution because it randomly samples the historical data multiple times to approximate its distribution. However, this random sampling renders the method computationally intensive, and thus, real-time application is difficult to achieve. The last method, i.e., the historical method, is the most popular method for calculating VaR because of its real-time applicability even though the computed values for some applications could contain little information on future volatilities (Pérignon & Smith, 2010). In this paper, we employ the historical method for VaR prediction, which we extend to consider information on news events.

A recent example of a parametric method for VaR calculation is the work of Huang & Lee (2013). The authors seek to predict future daily returns by considering high-frequency 5 minute data and demonstrate the merits of using such high-frequency intra-day data over using low-frequency data alone. Their proposed method involves a single parameterized model, which is constructed based on averaged or merged high-frequency data. Thus, the model is constructed over multiple forecasts that are generated using multiple lower-frequency (daily) datasets constructed from higher-frequency (intra-day) data. In their work, three data merging methods are evaluated, i.e., combining forecast, subsample averaging, and bootstrap averaging. Although this work is similar to ours in that high-frequency data are used as input, the work differs from our current endeavors in that we seek to predict returns within minutes or hours, whereas daily returns are predicted in (Huang & Lee, 2013). Moreover, the authors do not consider news events to smooth out irregularities in the high-frequency data. Finally, our approach does not rely on an underlying parametric model but rather on historical data alone.

Another example of a parametric approach is given in (Bormetti et al., 2012). The authors propose a Bayesian methodology for VaR computation that employs parametric production partition models. Such models rely on clustering structures and allow one to identify anomalies in the data under the assumption that the data are normally distributed. The approach is evaluated by comparing it to maximum likelihood-based approaches. The overall performance is the same, but the authors claim that the proposed methodology provides richer information about the clustering structure and outliers. The drawbacks of the proposed approach are related to scalability and dimensionality issues resulting from an increased number of assets and parameters.

The semi-parametric approach proposed by Mancini & Trojani (2011) does not assume a normal distribution but estimates predictive distributions of (parametric) GARCH models. Evaluation of the method through a Monte Carlo simulation and empirical application illustrates that the method provides more accurate VaR forecasts than classical methods, such as the historical method, particularly for longer horizons of several days and in the case of outliers. An important difference between that approach and our proposed approach is that Mancini & Trojani employ a semi-parametric approach with a Monte Carlo simulation, whereas we employ a non-parametric method, the historical method.

GARCH models have also been popular in attempts to improve historical VaR calculations. For instance, Hull & White (1998) improve the VaR calculation by updating the volatility in the historical method using GARCH/EWMA models to reflect the difference between the volatility at the time of the observation and the current volatility. An important difference between the work of Hull & White (1998) and ours is that we only observe portfolios with a fixed composition (i.e., fixed weight factors) rather than regular multi-equity portfolios. This decision is motivated by the fact that we seek to avoid having interdependencies between variances (heteroscedasticity), which is often the case when observing portfolios containing various financial equities. Hull & White (1998) propose a method to update the volatility during an appropriate time interval so that the volatility becomes a more dynamic factor in the VaR calculation. Based on the mean absolute percentage error (MAPE), their work is compared to another method that involves the assignment of weights to more recent observations (Boudoukh et al., 1998). The method proposed by Hull & White (1998) appears to outperform the traditional historical method and the method proposed by Boudoukh et al. (1998) for exchange rates, but the results are mixed for stock indices.

The existence of a strong relationship between the stock market and news events has been acknowledged in many previous studies (Byström, 2009; Engelberg & Parsons, 2009; Goonatilake & Herath, 2007).

Additionally, there is a proven correlation between number of news events and trading activity (Mitchell & Mulherin, 1994). However, many of the existing VaR calculation methods still do not consider these events. In practice, news information is not always fully and immediately processed in the value of shares (Rosenberg et al., 1985), and thus, for traders, reacting to news and estimating the portfolios' VaR in a timely and accurate manner is of utmost importance.

Antweiler & Frank (2006) note that those studies that do consider events often do not have extensive robustness checks. Typically, these studies employ inter-day data and are limited to a time horizon of several days before and after an event. Moreover, very little evidence is provided about what occurred in the evaluated time period. In contrast, long-term event studies are often subject to publication bias and typically focus on a single news event type. Antweiler & Frank (2006) employ a naive Bayes algorithm to classify news while employing various time windows and demonstrate that there is an initial under- or overreaction to news events, followed by directly opposite stock price movements. The authors claim that news events with a mainly negative connotation, e.g., merge announcements, equity repurchases, or declining earnings, appear to yield large abnormal returns. As with most comparable studies, a serious drawback is the use of inter-day data. In contrast, in our research, we employ data that is more fine-grained, i.e., we analyze intra-day data at hourly and 5 minute intervals.

Another method developed to improve technical trading algorithms by using information extracted from news is discussed by Zhai et al. (2007). The authors use a simple text classification algorithm with a supervised learning method. However, they also integrate general market news in combination with technical indicators instead of only microeconomic news events, which is the focus of our work. Based on a real-life market simulation, the authors conclude that technical indicators and news events alone are inaccurate as estimators but that the combination of the two could possibly yield better results.

3. Impact of Events on Stock Prices

To check the feasibility of our proposed approach, i.e., accounting for news events in VaR calculations, we assess the impact of news on stock prices. We first introduce our dataset and our metrics and then continue by exploring the impact of news events on stock prices.

3.1. Data

For our experiments, we consider a dataset collected using the ViewerPro tool (Semlab, 2013) for 363 equities from January 2010 to June 2011. ViewerPro is a proprietary application of SemLab, used to ex-

Table 1: Examples of news event types identified by the ViewerPro software

Type	Sub-types
CEO	hiring, resignation
Acquisition	consideration, start, completion, stop
Bid	receival, consideration, acceptance, drop, raise
Profit	down, up
Legal conflict	loss, resolution, win
Bankruptcy	–

tract various types of events from text-based data, such as news messages. These events, some examples of which are given in Table 1, can be used to determine the impact of a news item on an equity. ViewerPro converts large quantities of data about unstructured news items into structured trading information. After feeding unstructured news items into the ViewerPro system, several proprietary processing steps are executed to filter out unwanted information. The main procedures are metadata filtering, parsing, gazetteering, stemming, and automatic pattern matching.

The ViewerPro system relies on a domain-specific knowledge repository, i.e., a domain ontology with properties and lexical representations of financial entities (companies). First, concepts from the domain ontology are matched to news items. Second, the list of concepts is segmented into groups of related concepts using heuristics based on semantic, morphological, syntactical, and typographical data. Finally, the application identifies events using pattern matching with the previously extracted information.

In our dataset, for each news event of an associated stock, we record 3 hours of price data before and after the event's announcement with a 5 minute polling interval (a common interval in high-frequency time series, e.g., (Huang & Lee, 2013)), yielding $(3 \times 60) \div 5 = 36$ data points before and after an event announcement (in both sets, the price at the event occurrence is not included). Due to occasionally missing recorded stock rates, our dataset contains incomplete time series that must be discarded. A preliminary analysis indicates that for larger time windows, there is an increasingly higher probability that time series are incomplete, and thus, we do not consider time frames of more than 3 hours, as the number of usable (complete) time series decreases drastically. Events that have missing values within the time window of 3 hours before and after the announcement are discarded, which leaves us with a total of $5,435$ events and their associated prices for all equities monitored from January 2010 to June 2011.

3.2. Metrics

Based on the prices in our dataset, we calculate the average price for news event instance i as follows:

$$\overline{price}(i) = \frac{\sum_{t=1}^{l} prices_t}{l} \,, \tag{1}$$

where $prices$ is a vector of prices for time $t = 1, \ldots, l$ before or after news event instance i (note: the price at the event occurrence, i.e., when $t = 0$, is discarded from $prices$), and l is the length of the vector, i.e., length($prices$). Next, we calculate the percentile difference between the average price before a news event announcement and the average price after an announcement, i.e., $\overline{price}(i)_{before}$ and $\overline{price}(i)_{after}$, respectively. The percentile difference for prices for a specific event i, $return(i)$, is defined as:

$$return(i) = \left(\frac{\overline{price}(i)_{after} - \overline{price}(i)_{before}}{\overline{price}(i)_{before}} \right) \times 100 \,. \tag{2}$$

For positive values of $return(i)$, the prices increased after a particular news event i, whereas negative values indicate that prices went down after the announcement of event instance i. Next, we calculate the mean difference in percentile differences over all event instances i per event type e to prevent trends occurring in our data:

$$\overline{return}(e) = \frac{\sum_{j=1}^{m} return(i)_j}{m} \,, \tag{3}$$

where m is the number of instances i of a specific news type e, and $return(i)_j$ is the return of the j-th event instance i. To perform a better analysis of our data, we filter out the event types e with $m < 50$, i.e., the event types that occur disproportionally infrequently compared to more frequent events, which generally appear several hundreds of times. Moreover, we remove all instances with $\overline{price}(i)_{before} = 0$ or $\overline{price}(i)_{after} = 0$, i.e., those that do not invoke price changes.

3.3. Impacts

We evaluate the impacts of the events identified by the ViewerPro software by examining the magnitude and significance of the mean differences in returns for various time windows, ranging from 10 minutes to 3 hours before and after event announcements, and we provide a more detailed analysis for time windows of 30 minutes, 1 hour, and 3 hours.

When analyzing the impact of events for time windows of 30 minutes before and after their announcements, depicted in Table 2, we observe that several events have a statistically significant high positive or negative difference in mean returns, indicating large impacts. Share value and profit (expectation) increases

Table 2: Impacts in terms of $\overline{return}$ for a time window of 30 minutes before and after a news event type

Event type	$\overline{return}$	σ	# Events	p-value
Shares up	0.5416	1.1931	1618	0.0000
Profit exceeds expectations	0.2840	1.1143	177	0.0009
Profit up	0.1175	0.8692	604	0.0010
Profit down	−0.1552	0.9267	299	0.0041
Revenue up	0.1322	0.7490	239	0.0069
Warning	−0.3144	1.5813	186	0.0075
Division transaction	0.1023	0.5994	244	0.0084
Profit announced	0.0952	0.6880	347	0.0105
Earnings down	−0.2657	1.0363	81	0.0245
Contract win	0.0419	0.3809	414	0.0258
Earnings up	0.1177	0.7998	220	0.0305
Rating positive	0.0962	0.5654	139	0.0477
Company exceeds expectations	0.2492	1.7875	174	0.0684
Deal start	0.0556	0.4345	193	0.0775
Profit fails to meet expectations	−0.2824	1.5675	78	0.1180
Acquisition completed	0.0876	0.4935	73	0.1366
Company meets expectations	0.1205	0.7899	85	0.1656
Expectations confirmed	−0.0931	0.6871	105	0.1699
Securities issuance	−0.0382	0.5594	402	0.1726
Earnings fail to meet expectations	−0.2643	1.4082	55	0.1735
Shares transaction	0.0992	1.9765	685	0.1899
Bid raised	0.8134	4.8321	62	0.1935
Collaboration start	0.0337	0.4571	280	0.2190

yield large improvements in returns, whereas decreases in profit and earnings, as well as warnings about companies, tend to yield losses in mean returns. The standard deviations of events with high impacts are quite high; however, the p-values from our significance tests (based on a one-sample two-sided Student's t-test with a significance level of 10%) indicate that the returns after the associated events are significantly

Table 3: Impacts in terms of $\overline{return}$ for a time window of 1 hour before and after a news event type

Event type	$\overline{return}$	σ	# Events	p-value
Shares up	0.9150	1.7283	1496	0.0000
Profit exceeds expectations	0.4438	1.5456	163	0.0003
Warning	−0.4911	1.8887	181	0.0006
Rating down	−0.1518	1.6733	1319	0.0010
Division transaction	0.1705	0.8339	233	0.0021
Earnings down	−0.5832	1.7192	75	0.0047
Exceeds expectations	0.4094	2.1175	166	0.0140
Profit up	0.1202	1.1703	551	0.0163
Deal start	0.1124	0.6168	178	0.0164
Initial rating positive	0.1408	0.7092	134	0.0236
Collaboration start	0.0859	0.6480	276	0.0287
Contract win	0.0574	0.5508	389	0.0407
Earnings up	0.1446	1.1071	206	0.0630
Profit fails to meet expectations	−0.4161	1.8311	68	0.0673
Earnings exceed expectations	0.4192	1.7973	61	0.0758
Expectations negative	−0.2587	1.6546	120	0.0906
Profit announcement	0.0937	1.0281	335	0.0966
Earnings fail to meet expectations	−0.3959	1.7751	55	0.1070
Profit meets expectations	−0.4073	1.9370	56	0.1246
CEO resigns	0.1748	0.9984	75	0.1363
Bid raised	1.1299	6.0080	62	0.1470
Shares buy back	0.0538	0.5652	229	0.1523
Revenue up	0.1410	1.4824	225	0.1559
Securities issuance	−0.0485	0.7219	395	0.1828
Shares transaction	0.0442	0.8332	566	0.2082
Jobs down	0.6665	5.2082	93	0.2227

Table 4: Impacts in terms of $\overline{return}$ for a time window of 3 hours before and after a news event type

Event type	$\overline{return}$	σ	# Events	p-value
Shares up	0.6614	1.4762	300	0.0000
Rating down	−0.0993	1.0965	639	0.0225
Earnings up	0.3155	1.3428	91	0.0283
Warning	−0.6579	2.8594	80	0.0442
Profit exceeds expectations	0.3258	1.4145	56	0.0933
CEO resigns	0.3256	1.3341	50	0.0939
Division transaction	0.2598	1.7251	124	0.0974
Collaboration start	0.1175	0.8659	150	0.0996
Profit down	−0.2264	1.6264	136	0.1082
Shares transaction	0.1762	1.6416	201	0.1305
Deal start	0.4498	3.3296	99	0.1842
Initial rating positive	0.1352	0.8510	72	0.1848
Revenue up	0.1480	1.1388	104	0.1900
Expectations negative	−0.3902	2.437	64	0.2084

different. When using a time window of 30 minutes, non-significant impacts are mainly related to expectations that are not met or that are exceeded.

When the time window is increased to 1 hour (Table 3), we obtain a similar list of event impacts. In general, positive and negative impacts are higher than those obtained for the 30 minutes time window. When we increase the time window to 3 hours (Table 4), fewer types of events are included due to the increased data scarcity inherent in larger window sizes. However, we again obtain a similar set of statistically significant impacts.

This observation is confirmed when evaluating the top five most significant impacts of news events for all time windows, ranging from 10 minutes to 3 hours (in increments of 5 minutes), as shown in Figure 1. Some events have a substantial impact in most of the time windows, whereas some only appear in a few time windows. An announcement of increased company shares has the most significant impact (35 times). Warnings about companies also represent a substantial share of the top five most significant impacts (30 times), followed by profits exceeding expectations (26 times). Whenever there are announcements of rev-

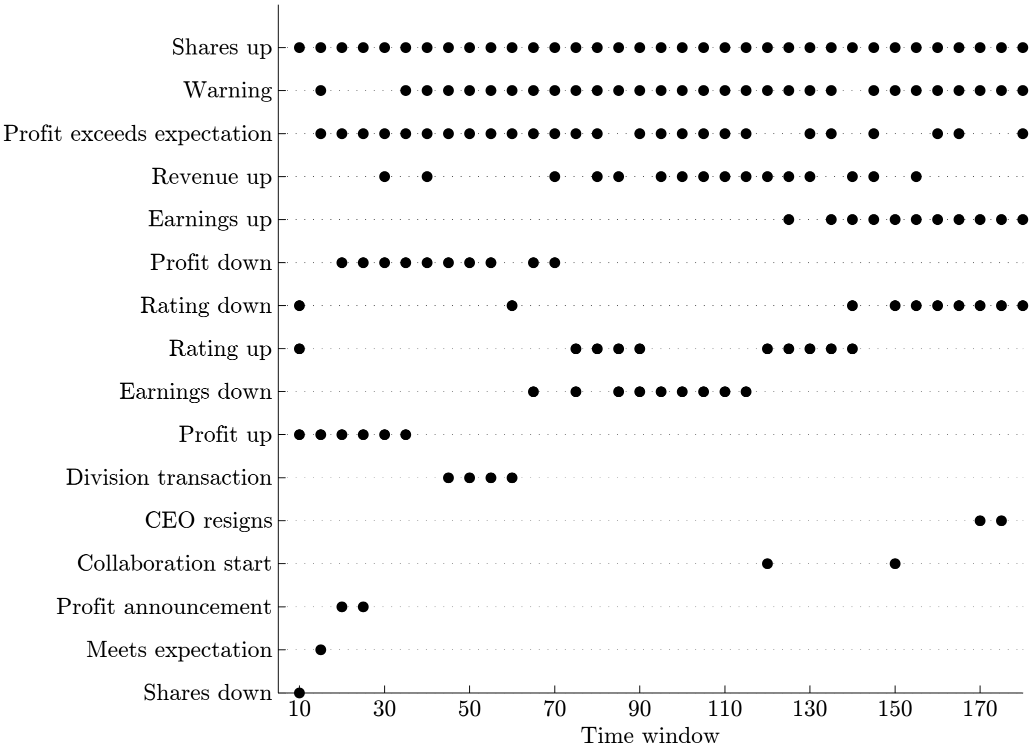


Figure 1: Top five most significant impacts per time window

enues going up (16 times), earnings going up (11 times), earnings going down (9 times), profits going down (10 times), or ratings going up or down (10 times each), impacts appear to be significant as well. The other events occur less frequently.

An important observation here is that some of these events often have an impact regardless of the time window is (e.g., company shares going up and company warnings), whereas other events mainly have an impact in the short term (profits going up or down) or longer term (ratings going down or earnings going up). In addition, there are some events that appear to have an impact on medium-term time windows only (i.e., earnings going down and ratings going up).

The results are intuitive. In the stock market, changes in prices are rarely more than $3-5\%$. In our experiments, we measured changes between -1% and 1%. In addition, we observed that negative news event types generate negative average percentile changes, whereas positive news event types generated positive average percentile changes, which is in accordance with similar observations in related work (Antweiler & Frank, 2006).

4. Event-Based Historical VaR Method

To assess whether the incorporation of news into the calculation of VaR of equities improves the overall quality of the prediction, we propose the framework depicted in Figure 2. In contrast to Section 3, in which we described the analysis of 5 minutes interval prices before and after events, the framework is based on two inputs, namely, an unmodified complete list of historical stock prices, $prices_{hist}$, and a list of financial events, *events*. These inputs are extracted from several feeds using the ViewerPro application described in Section 3. Before using the collected equity prices, a cleaning procedure is followed, by which prices that are recorded within stock markets' opening times are retained for further computations. In addition, to decrease the computational complexity, the time intervals between individual prices are defined per hour, which is in contrast to the short 5 minute intervals between the prices used in our analysis in

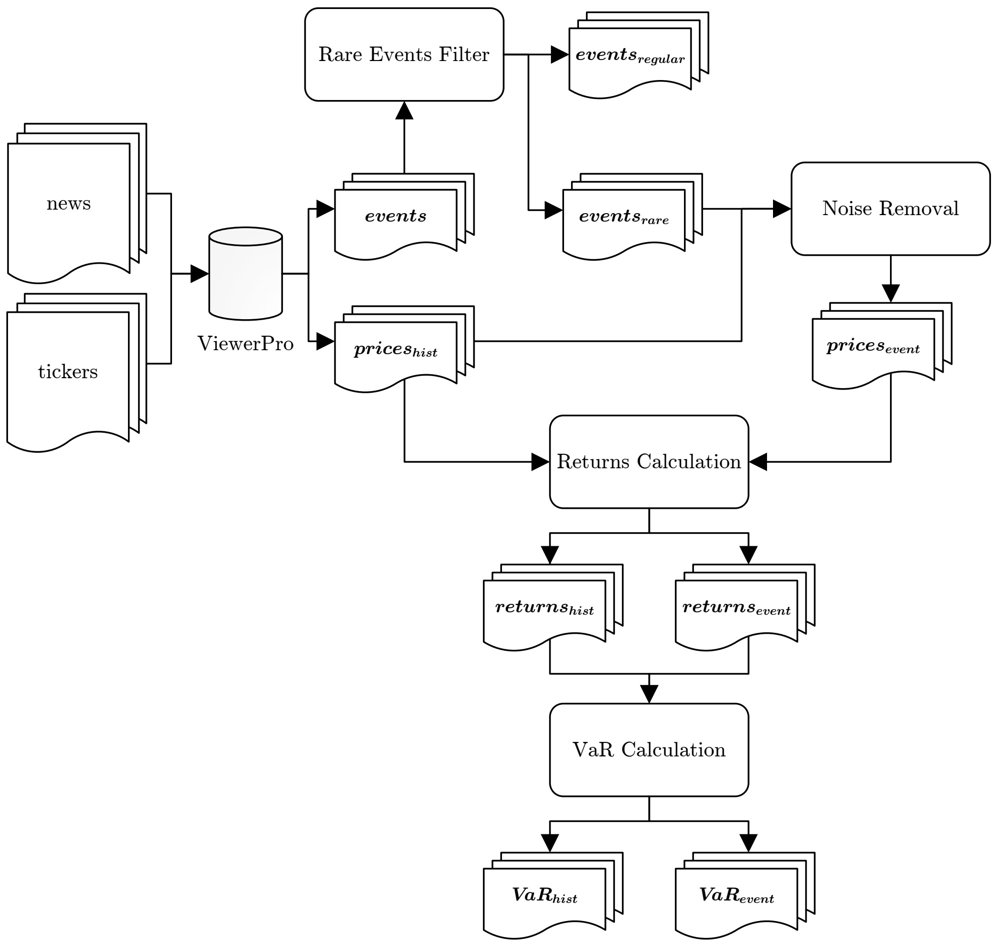


Figure 2: Overview of data flows and processing steps within the event-based historical VaR method

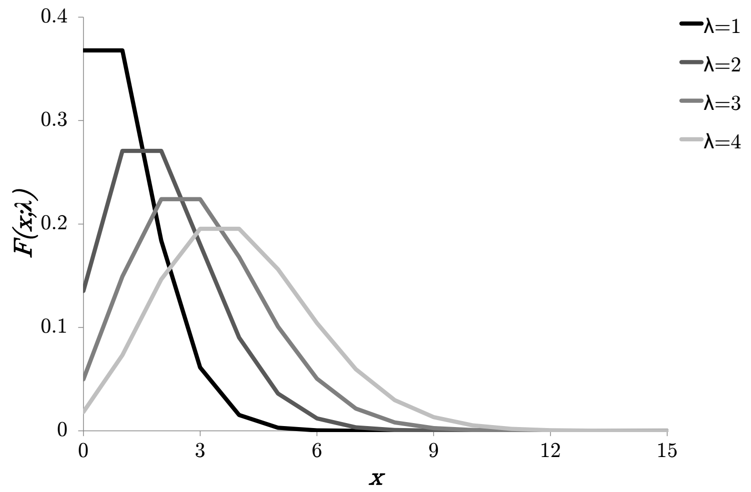


Figure 3: Poisson distributions for various measured and expected occurrences, i.e., x and λ, respectively

Section 3. News is parsed using ViewerPro's computational linguistics, semantic analysis, and formal logic procedures, which determine the positive and negative impacts of the information described in the news on the equities.

Additionally, we identify the irregularly occurring event types from our event set, as these events are not likely to occur again and thus cause significant noise in stock rates. Poisson distributions are used in many fields to model the number of events that occur in a certain time interval, and therefore, we apply a Poisson distribution F to a test set *test*:

$$F(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!} , \tag{4}$$

where x and λ represent the measured and expected number of occurrences in the test set *test*. For a threshold α of 0.05, for $x = 0$ (no occurrence), $F(x; \lambda) < \alpha$ for $\lambda \geq 3$, as depicted in Figure 3. For a training set *train*, the expected number of occurrences λ' is obtained by scaling λ by the proportion of the set cardinalities, i.e., $\lambda' = \lambda \times \vartheta$, with $\vartheta = |train|/|test|$. Therefore, in this paper we consider event types that occur $\geq 3 \times \vartheta$ in the training set as regular events, and events occurring $< 3 \times \vartheta$ as rare events. Rare events are stored in $events_{rare}$ and are used in further processing steps, whereas regular events in $events_{regular}$ are discarded.

Next, noise removal is performed using the identified rare events in set $events_{rare}$, which are associated with the times at which they occurred and their respective stock rates. We adjust the collected prices $prices_{hist}$ for a time window to account for the generated noise by changing their values to the previously measured value, resulting in a list of event-corrected prices $prices_{event}$. This process is illustrated in Algorithm 1, in which a list of chronologically ordered hourly recorded historical prices *prices* (containing

Algorithm 1 Stock price cleaning based on news events

```
Require: prices = array of historical stock prices and associated times
Require: events = array of rare events and associated times
Require: window = integer representing time window
previousprice.value = prices[1].value
for all price in prices do
    for all event in events do
        if impact > 0 then
            impact = impact − 1
            price.value = previousprice.value
        end if
        if price.time = event.time then
            impact = window
        end if
    end for
    previousprice.value = price.value
end for
```

all prices in $prices_{hist}$ for a specific stock) is processed into a set of prices $prices_{event}$ that is cleaned from the effects of an emerging event event from event list events (containing the rare events from $events_{rare}$). For each historical stock price *price* in *prices*, the algorithm checks for event occurrences by comparing the stock price time with the time of each rare event *event* stored in events. When an event occurrence is identified, we set *impact* to the window size *window*, which causes the value of the subsequent *price* items to be set to the current value (we assume that no two events occur simultaneously). The value of *impact* is decreased by one every next *price* in price list *prices*, so subsequent price values are updated until the window size has been reached. In the case of overlapping events, the *impact* counter is reset to the window size *window*.

Both sets of original ("*hist*") and cleaned ("*event*") prices are converted to sets with hourly returns. We compute the return set $returns_t$ of a price set *prices* as the relative change between the price at time $t+1$ and the previous price at time t, i.e.,

$$returns_t = \frac{prices_{t+1} - prices_t}{prices_t} \quad t = 1, \ldots, N-1 \;, \tag{5}$$

where N denotes the number of items in the list. The historical returns are used to estimate future returns. The time horizon used to compute returns is one day. After sorting the return list *returns*, we calculate the Value at Risk, *VaR*, as

$$VaR = returns' \left[\lfloor \alpha \cdot \text{length}(returns) \rfloor \right] \ , \tag{6}$$

where $returns'$ is the ordered (sorted) list of returns and α denotes the confidence level. Thus, in a dataset with 20 historical returns – with the first element located at position 1 and the last element located at position 20 – we select the return on position 19 (to be the VaR) when using a confidence level of 0.95. Using equation (6), we calculate VaR_{event} and VaR_{hist} using our adjusted method and the traditional method (i.e., the historical method without the improvements proposed in (Boudoukh et al., 1998; Hull & White, 1998)), respectively. An example is given in Table 5. Here, the results of a VaR calculation are presented based on 21 prices – with and without cleaning – with an *event* occurring at $t = 4$ (denoted by a *) while using a *window* size of five. With a confidence interval of 95% (5% probability for the VaR definition in Section 2), this would result in a VaR of -0.44 or -0.15 (printed in bold font) for returns for the historical method or the event-based historical method, respectively (the VaR position in the returns is $95\% \cdot 20 = 19$). The observed difference stems from the proposed removal of noise inherently associated with events, i.e., the noise in prices generated from time $t = 4 + 1$ to time $t = 4 + 5$. These differences can then be evaluated by assessing the quality of both predicted values.

5. Evaluation

We employ various measures to evaluate the performance of the proposed historical VaR calculation for fixed-composition portfolios using a dataset of stock rates and news events. First, we discuss our data, and then, we elaborate on the metrics used. Finally, we present our experimental results.

5.1. Data

In our experiments, we employ a dataset stemming from the ViewerPro software, which, after the processing steps described in Section 4, contains news events and stock data collected on an hourly basis for 363 equities on weekdays during the year 2010. The dataset consists of approximately $2,000$ stock data points, 119 event types, and $50 - 75$ event instances per equity. To evaluate the performance of the calculation, we predict the VaR_{event} and VaR_{hist} for 75% of our dataset. The remaining 25% is used as a test set for comparing the predicted VaR with the actual VaR.

Table 5: Example VaR calculation for returns with and without noise cleaning

	Prices				Returns				Returns (sorted)	
	t	*hist*	*event*		*t*	*hist*	*event*		*hist*	*event*
	1	0.35	0.35		1	0.14	0.14		4.77	0.46
	2	0.40	0.40		2	0.07	0.07		0.54	0.14
	3	0.43	0.43		3	−0.05	−0.05		0.36	0.14
*	4	0.41	0.41		4	0.54	0.00		0.14	0.07
	5	0.63	0.41		5	−0.79	0.00		0.14	0.05
	6	0.13	0.41		6	4.77	0.00		0.07	0.03
	7	0.75	0.41		7	−0.44	0.00		0.05	0.00
	8	0.42	0.41		8	0.36	0.00		0.05	0.00
	9	0.57	0.41		9	0.05	0.46		0.03	0.00
	10	0.60	0.60	→	10	−0.15	−0.15	→	−0.02	0.00
	11	0.51	0.51		11	−0.16	−0.16		−0.03	0.00
	12	0.43	0.43		12	−0.14	−0.14		−0.03	−0.02
	13	0.37	0.37		13	−0.03	−0.03		−0.05	−0.03
	14	0.36	0.36		14	0.14	0.14		−0.07	−0.03
	15	0.41	0.41		15	0.05	0.05		−0.08	−0.05
	16	0.43	0.43		16	−0.02	−0.02		−0.14	−0.07
	17	0.42	0.42		17	−0.07	−0.07		−0.15	−0.08
	18	0.39	0.39		18	0.03	0.03		−0.16	−0.14
	19	0.40	0.40		19	−0.08	−0.08		−0.44	−0.15
	20	0.37	0.37		20	−0.03	−0.03		−0.79	−0.16
	21	0.36	0.36							

5.2. Metrics

In contrast to common approaches to evaluating VaR calculations, our approach does not employ the Kupiec test (Kupiec, 1995), as the test is statistically weak with small sample sizes (e.g., one year). As the dataset we employed only covers 2010, we need different measures that provide insight into the effectiveness of our proposed event-based approach. Therefore, to analyze the number of equities for which our adjusted event-based historical method provides better-quality predictions than the traditional historical method, we measure each method's squared error SE for equity c. The squared error is defined as follows:

$$SE_c = \left(VaR_{c,actual} - VaR_{c,predicted}\right)^2 \ , \tag{7}$$

where $VaR_{c,actual}$ and $VaR_{c,predicted}$ represent equity c's actual VaR measured in our test set and the predicted VaR based on our training set, respectively, and $VaR_{c,predicted}$ is one of VaR_{event} or VaR_{hist}.

Subsequently, we combine the squared errors into the mean squared error (MSE) for the historical method and the event-based historical method, i.e., MSE_{hist} and MSE_{event}. The MSE is defined as the summation of the squared errors (SE) of all equities $c \in C$ divided by the number of equities, i.e.,

$$MSE = \frac{\sum_{c \in C} SE_c}{|C|} \ , \tag{8}$$

where $|C|$ denotes the total number of equities in set E (363).

Additionally, we evaluate the number of times both methods outperform the other, i.e., OPT (OutPerformed Total). To do this, we compare the computed squared errors $SE_{c,hist}$ and $SE_{c,event}$ for all equities $c \in C$:

$$OPT_{hist,event} = \sum_{c \in C} O(SE_{c,hist}, SE_{c,event}) \ , \tag{9}$$

$$OPT_{event,hist} = \sum_{c \in C} O(SE_{c,event}, SE_{c,hist}) \ , \tag{10}$$

$$O(X, Y) = \begin{cases} 1 & \text{if } X < Y \\ 0 & \text{else} \end{cases} \ . \tag{11}$$

In our experiments, we compare the MSE and OPT for the traditional and event-based VaR calculation methods, both determined for the full event dataset and for a dataset containing only the rare events, using a time window of 8 hours (determined based on initial estimates). We then determine the optimal time window size by observing event-based VaR calculation plots of (normalized) MSE and OPT values for time windows ranging from 1 to 24 (i.e., three working days of 8 hours, which is the maximum effect of a news event (Kalev et al., 2004)). In addition, we consider the number of overconfident predictions ($CONF$) of all equities $c \in C$, which is calculated as follows

$$CONF = \sum_{c \in C} Q(VaR_{c,predicted}, VaR_{c,actual}) \ , \tag{12}$$

$$Q(X,Y) = \begin{cases} 1 & \text{if } X > Y \\ 0 & \text{else} \end{cases} , \tag{13}$$

where $VaR_{e,predicted}$ represents the predicted VaR_{event} for equity c based on our adjusted dataset (containing only the rare events). To optimize the size of the time window, for each window size, we normalize its associated MSE, $OPT_{event,hist}$, and $CONF$ using min-max normalization based on the previously computed values for the full range of window sizes, resulting in MSE', OPT', and $CONF'$, respectively. Subsequently, we subtract $CONF'$ from the computed difference between OPT' and MSE', and normalize the result using min-max normalization to obtain a final score S that is to be maximized:

$$S = (OPT' - MSE') - CONF' . \tag{14}$$

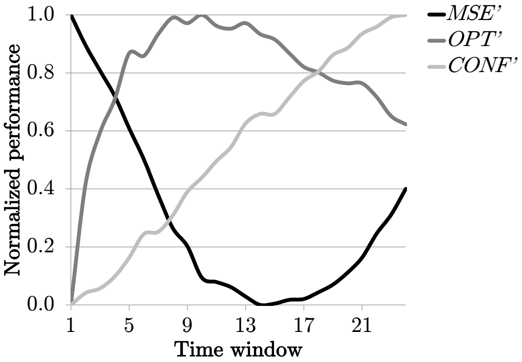
Finally, the significance of the results is assessed by performing a two-sample one-tailed Student's t-test on the sets of individual squared errors SE_{hist} and SE_{event} for our optimal configuration. To do this, we use a significance level of 0.05 to reject the null hypothesis that there is no difference between the measured MSE values.

5.3. Results

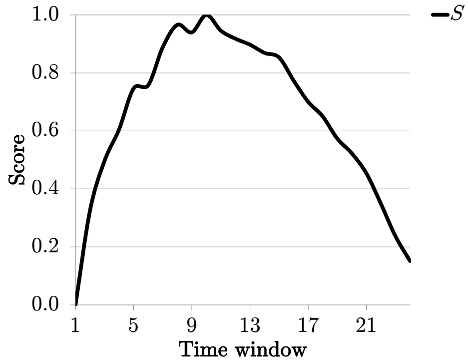
Table 6 presents the MSE and OPT values for both the traditional and event-based historical VaR calculation methods. When using all (i.e., regular and rare) events and stock rates on an hourly basis and when employing a time window of 8 hours, we observe a decrease of 21.44% in terms of MSE when accounting for event-generated noise in our stock data. Additionally, the event-based VaR calculation method outperforms the traditional historical method in 76.18% of cases. The table also indicates that removing only the noise generated by rare events yields an additional improvement over the previous results. In 79.12% of the cases, event-based historical VaR calculation outperforms the historical method. In addition, the MSE is 26.37% lower for the event-based historical VaR calculations when only rare events are considered.

Table 6: Comparison of the performance of traditional and event-based historical VaR calculations using a window of 8 hours

	All events		Rare events	
Measure	*hist*	*event*	*hist*	*event*
MSE	1.1234E−05	8.8254E−06	1.1234E−05	8.2717E−06
OPT	81	259	71	269



(a) *MSE*′, *OPT*′, and *CONF*′



(b) *S*

Figure 4: Performance of event-based VaR prediction models with various time windows

As illustrated in Figure 4, we obtain an optimized window size of 10. Although the *MSE* and *OPT* scores have a clear minimum and maximum, respectively, for our dataset, the number of overconfident predictions becomes increasingly large when the time window is enlarged. Thus, when the differences between the *OPT* and *MSE* scores for various time window sizes are small, one should focus on minimizing the number of overconfident predictions, which results in small window sizes.

As shown in Table 7, utilizing a window of 10 instead of 8 hours on a dataset with rare events indeed yields improvements when compared to the results shown in Table 6. The *MSE* of our event-based historical VaR prediction models decreases by 31.78% over the traditional historical VaR prediction method's *MSE*, and event-based VaR prediction outperforms the historical method in 77.71% of cases. Alternatively, even greater improvements are achieved when we determine the optimal cleaning window for each event type separately, with large differences in stock prices (we employ a threshold of 50.00% in our experiments) after an event occurrence indicating noise that should be cleaned and small differences indicating that the market has returned to normal. The measured *MSE* values decrease by 35.55%, and 75.59% of the event-based VaR predictions outperform the historical method.

Table 7: Comparison of the performance of traditional and event-based historical VaR calculation with rare events

	window = 10		*window* = event-based	
Measure	*hist*	*event*	*hist*	*event*
MSE	1.1234E−05	7.6633E−06	1.1234E−05	7.2402E−06
OPT	76	265	83	257

To assess the significance of the measured *MSE* improvement of 35.55%, we perform a paired two-sample one-tailed t-test based on SE_{hist} and SE_{event}, containing squared errors for all equities. We obtain a p-value of 0.0027 and reject the null hypothesis that there is no difference between the measured *MSE* values at a significance level of 0.05. Thus, the proposed event-based historical VaR calculation method (using rare events and event-based window sizes) produces more reliable VaR predictions than the traditional method.

6. Conclusions

VaR is one of the most widely used risk assessment measures in the financial world. The historical VaR is a popular method for computing VaR in real time. The historical stock return data used to calculate VaR may be sensitive to outliers caused by seldom-occurring news events that occur during the sampled period. Therefore, in this study, we investigated the impacts of various news events on stock prices. Additionally, we have proposed a way to enhance the prediction of VaR based on historical data by removing disturbances induced by such events. Removing such disturbances would enable practitioners to make better predictions of risk in terms of distributions of expected future returns.

Based on a dataset of stock rates and news events obtained using the proprietary ViewerPro software, we have identified news events that were likely to generate noise. This event-generated noise was subsequently removed from the stock rates. Based on our experiments, in which we evaluated various cleaning window sizes, we can conclude that VaR can be improved with news as an additional input. When using an arbitrary cleaning window of 8 hours, we observed that event-based historical VaR calculations produced more accurate results than traditional historical VaR calculations, resulting in lower *MSE* scores. When only rare events (identified using a Poisson distribution) were considered, the decrease in *MSE* increased from 21.44% to 26.37%, and our new method outperformed the traditional method more often (79.12% versus 76.18% of the cases). Moreover, we have optimized the cleaning window to 10 hours (when considering only rare events), resulting in an *MSE* improvement of 31.78% and event-based VaR calculation outperforming the traditional method in 77.71% of cases. Alternatively, we considered a per-event cleaning window optimization strategy that demonstrated a significant *MSE* improvement of 35.55% compared to the traditional historical method, outperforming the latter in 75.59% of cases.

In future work, we propose to investigate how to account for the type of news event in our VaR method, which could affect the influence of an event on equity prices (e.g., mergers could generate more noise than quarterly profit announcements). Another future research direction is related to accounting for general stock market events, such as financial crises, instead of company-specific news only. We would also like to build a real-life market simulation for our improved historical VaR method.

Acknowledgment

The authors are partially supported by the NWO Physical Sciences Free Competition project 612.001.009: Financial Events Recognition in News for Algorithmic Trading (FERNAT) and the Dutch national program COMMIT. We would like to thank Mark Vreijling and Wijand Nuij from Semlab for fruitful discussions and technical support during this research.

References

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The Distribution of Stock Return Volatility. *Journal of Financial Economics*, *61*, 43–76.

Antweiler, W., & Frank, M. Z. (2006). Do US Stock Markets Typically Overreact to Corporate News Stories? `http://dx.doi.org/10.2139/ssrn.878091`.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, *9*, 203–228.

Asche, F., Dahl, R. E., & Oglend, A. (2013). *Value-at-Risk: Risk Assessment for the Portfolio of Oil and Gas Producers*. UiS Working Papers in Economics and Finance 2013/3 University of Stavanger.

Bormetti, G., De Giuli, M. E., Delpini, D., & Tarantola, C. (2012). Bayesian Value-at-Risk with Product Partition Models. *Quantitative Finance*, *12*, 769–780.

Borsje, J., Hogenboom, F., & Frasincar, F. (2010). Semi-Automatic Financial Events Discovery Based on Lexico-Semantic Patterns. *International Journal of Web Engineering and Technology*, *6*, 115–140.

Boudoukh, J., Richardson, M., & Whitelaw, R. F. (1998). The Best of Both Worlds: A Hybrid Approach to Calculating Value at Risk. *Risk*, *11*, 64–67.

Byström, H. (2009). News Aggregators, Volatility and the Stock Market. *Economics Bulletin*, *29*, 2673–2682.

Chan, W. S. (2003). Stock Price Reaction to News and No-News: Drift and Reversal After Headlines. *Journal of Financial Economics*, *70*, 223–260.

Dimitrakopoulos, D. N., Kavussanos, M. G., & Spyrou, S. I. (2010). Value at Risk Models for Volatile Emerging Markets Equity Portfolios. *The Quarterly Review of Economics and Finance*, *50*, 515–526.

Engelberg, J. E., & Parsons, C. A. (2009). The Causal Impact of Media in Financial Markets. *Journal of Finance*, *66*, 67–97.

Fama, E. F. (1965). The Behavior of Stock-Market Prices. *Journal of Business*, *38*, 34–105.

Goonatilake, R., & Herath, S. (2007). The Volatility of the Stock Market and News. *International Research Journal of Finance and Economics*, *3*, 53–65.

Hogenboom, A., Hogenboom, F., Frasincar, F., Kaymak, U., van der Meer, O., & Schouten, K. (2011). Detecting Economic Events Using a Semantics-Based Pipeline. In A. Hameurlain, S. W. Liddle, K.-D. Schewe, & X. Zhou (Eds.), *Twenty-Second*

International Conference on Database and Expert Systems Applications (DEXA 2011) (pp. 440–447). Springer volume 6860 of *Lecture Notes in Computer Science*.

Hogenboom, A., Hogenboom, F., Frasincar, F., Schouten, K., & van der Meer, O. (2013). Semantics-Based Information Extraction for Detecting Economic Events. *Multimedia Tools and Applications*, *64*, 27–52.

Hogenboom, F., de Winter, M., Frasincar, F., & Hogenboom, A. (2012a). A News-Based Approach for Computing Historical Value-at-Risk. In J. Casillas, F. J. Martínez-López, & J. M. Corchado (Eds.), *First International Symposium on Management Intelligent Systems (IS-MiS 2012)* (pp. 283–292). Springer volume 171 of *Advances in Intelligent Systems and Computing*.

Hogenboom, F., de Winter, M., Jansen, M., Hogenboom, A., Frasincar, F., & Kaymak, U. (2012b). Event-Based Historical Value-at-Risk. In *IEEE Computational Intelligence for Financial Engineering & Economics 2012 (CIFEr 2012)* (pp. 66–72). IEEE Computer Society.

Holton, G. A. (2003). *Value at Risk: Theory and Practice*. (8th ed.). Academic Press.

Huang, H., & Lee, T.-H. (2013). Forecasting Value-at-Risk Using High-Frequency Information. *Econometrics*, *1*, 127–140.

Hull, J. C. (2011). *Options, Futures, and Other Derivatives*. (8th ed.). Pearson Education Limited.

Hull, J. C., & White, A. (1998). Incorporating Volatility Updating into the Historical Simulation Method for Value-at-Risk. *Journal of Risk*, *1*, 5–19.

Ikenberry, D. L., & Ramnath, S. (2002). Underreaction to Self-Selected News Events: The Case of Stock Splits. *Review of Financial Studies*, *15*, 489–526.

Jorion, P. (2006). *Value at Risk: The New Benchmark for Managing Financial Risk*. (3rd ed.). McGraw-Hill.

Kalev, P. S., Liu, W.-M., Pham, P. K., & Jarnecic, E. (2004). Public Information Arrival and Volatility of Intraday Stock Returns. *Journal of Banking & Finance*, *28*, 1441–1467.

Kupiec, P. H. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *Journal of Derivatives*, *3*, 73–84.

Mancini, L., & Trojani, F. (2011). Robust Value at Risk Prediction. *Journal of Financial Econometrics*, *9*, 281–313.

Mitchell, M. L., & Mulherin, J. H. (1994). The Impact of Public Information on the Stock Market. *Journal of Finance*, *49*, 923–950.

Olson, D. L., & Wu, D. (2010). Enterprise Risk Management Models. chapter Value at Risk. (pp. 131–141). Springer.

Pérignon, C., & Smith, D. R. (2010). The Level and Quality of Value-at-Risk Disclosure by Commercial Banks. *Journal of Banking & Finance*, *34*, 362–377.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of Conditional Value-at-Risk. *Journal of Risk*, *2*, 21–41.

Rockafellar, R. T., & Uryasev, S. (2002). Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking & Finance*, *26*, 1443–1471.

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive Evidence of Market Inefficiency. *Journal of Portfolio Management*, *11*, 9–16.

Semlab (2013). ViewerPro. `http://viewerpro.semlab.nl/`.

Street, A. (2010). On the Conditional Value-at-Risk Probability-Dependent Utility Function. *Theory and Decision*, *68*, 49–68.

Zhai, Y., Hsu, A., & Halgamuge, S. K. (2007). Combining News and Technical Indicators in Daily Stock Price Trends Prediction. In D. Liu, S. Fei, Z. Hou, H. Zhang, & C. Sun (Eds.), *4th International Symposium on Neural Networks (ISSN 2007)* (pp. 1087–1096). Springer-Verlag Berlin, Heidelberg.