Inaccurate Value at Risk Estimations: Bad Modeling or Inappropriate Data?

Evangelos Vasileiou[1]

Abstract

Forecasting accurate Value-at-Risk (VaR) estimations is a crucial task in applied financial risk management. Even though there have been significant advances in the field of financial econometrics, many crises have been documented throughout the world in the last decades. An explanation for this discrepancy is that many contemporary models are too complex and cannot be easily understood and implemented in the financial industry (Fama in Financ Anal J 51:75–80, 1995; Ross in AIMR conference proceedings, vol. 1993, no. 6, pp. 11–15, Association for Investment Management and Research, 1993). In order to bridge this theory–practice gap, we present a computational method based on the leverage effect. This method allows us to focus on financial theory and remove complexity. Examining the US stock market (2000–2020), we provide empirical evidence that our newly suggested approach, which uses only the most appropriate observation period, significantly increases the accuracy of the Conventional Delta Normal VaR model and generates VaR estimations which are as accurate as those of advanced econometric models, such as GARCH(1,1).

1 Introduction

The role of the financial system is to allocate resources from lenders/investors to borrowers/enterprises in an efficient way in order to maximize welfare for all (lenders/investors, borrowers/enterprises, and society). A significant aspect of this flow of funds is how the financial system allows risk to be shared and who bears it (Allen et al. 2004). Therefore, the protection of lenders/investors is vital to the stability of

---

✉ Evangelos Vasileiou
e.vasileiou@ba.aegean.gr

[1] Department of Financial and Management Engineering, School of Engineering, University of the Aegean, 45 Kountouriotou Str, 82100 Chios, Greece

the financial system and largely relies on the ability of lenders/investors to accurately assess the financial risk in their transactions/decisions. For these reasons, the Committee of European Securities Regulators (CESR) introduced a set of regulations (CESR 2010) which focus on financial markets, and the Basel Committee on Banking Supervision (BCBS) issued recommendations (Basel Accords I, II, and III) which focus on the banking industry.

The dominant statistical measure for estimating financial risk is the Value at Risk (VaR). The VaR of an investment is an estimation of the loss that will not be exceeded, with a given significance level, at a specific timeframe. It is a percentile, usually 1% or 5%,[1] of a profit/loss distribution which presents some differences/variances depending on the model the researcher/risk analyst uses, e.g. when the Historical VaR model is applied, VaR is the percentile of the actual x-day past returns; under the Delta Normal VaR model, it is the percentile of the normalized distribution of the last x-day returns (Jorion, 2007).

The first widely known and complete VaR system was Risk Metrics (Longerstaey & Spencer, 1996). The need for financial risk information and an increased interest by policy makers on financial risk issues and VaR, in combination with advances in financial econometrics and computer science over the last decades, have led to significant advances in VaR modeling: Monte Carlo (Berkowitz et al. 2011), the popular GARCH family models (Angelidis et al. 2004; Degiannakis et al. 2012; Engle, 2004), and the Markov Switching Regime models (Billio & Pelizzon, 2000). However, some contemporary suggestions, such as Fuzzy VaR, Expected Shortfall models with elliptical distributions (Moussa et al. 2014, and Extreme Learning Machine (Zhang et al. 2017), are too complex to be implemented (explained) in the financial industry by (to) non-mathematicians and computer scientists.

The oxymoron is that despite strict legislation and significant advances in financial econometrics, several crises have emerged throughout the world in the last decades: USA savings and loan crisis (1989–1991), Asian financial crisis (1997–1998), Russian financial crisis (1998), Argentine crisis (1999–2002), Global financial crisis (2007–2008), European Sovereign Debt Crisis (2010-today) etc. Advanced VaR models show increased accuracy in forecasting risk. So, why are there so many crises?

A reason could be that, in practice, the most popular models applied in the financial industry are the conventional and simplest VaR models: the Historical, the Variance–covariance (or Delta Normal), and the Monte Carlo Simulation.[2] Conventional models generate low VaR estimations during the days before a crisis and high VaR estimations when a crisis has already emerged (Vasileiou & Pantos, 2020; Vasileiou

---

[1] Depending on the confidence level of our risk estimation.

[2] These models are considered as the most widely studied in finance (Jorion, 2007), while even the regulators mention that these models are the most popular (BCBS, 2006; CESR, 2010).

& Samitas, 2020).[3] However, the purpose of VaR should be to timely and accurately inform market participants that a stress period is approaching.

To offer an explanation as to why the advantages of financial econometrics have not been adequately utilized in the industry, some scholars suggest that wider use of advanced models in the financial industry is hindered by the complexity and the increased cost that such advanced systems involve (Vasileiou, 2016). Eminent scholars highlight the mathematical complexity issue in their work: Fama (1995) notes that in some cases contemporary models are extremely complex for non-mathematicians, and Ross (1993, p. 11) characteristically describes such a case using the phrase: "If you torture the data long enough, it will confess to any crime".

Therefore, this study's main objective is to suggest a methodology that could bridge this theory–practice gap and improve the accuracy of VaR estimations without the need for mathematical complexity. Our goal is to provide investors with accurate, representative, and easy to understand and analyze VaR estimations. This way, investors will be adequately informed of the risk they bear when they invest in a financial asset. This will help address a long-standing problem in financial markets and promote financial stability. In order to achieve our goal, we do not use complex mathematics in the output stage, but we filter the data in the input stage and select the most appropriate data set. In some way, we try to listen to what the data want to say to us, as opposed to torturing the data.

We tested our assumptions using the easiest to apply and communicate VaR model, the Delta Normal (or Variance–Covariance) model (Jorion, 2007). We examined the world's most significant stock market, the US stock market and particularly the S&P500 Index, for the period 2000–2020. During this period, several crises emerged: (a) the dot.com crisis in early 2000, (b) the crisis of 2007–2008, and (c) the crisis of the COVID-19 pandemic in 2020. The empirical results confirm our assumptions: the data filtering procedure significantly improves the accuracy of Conventional Delta Normal VaR (CDNVaR) estimations without the use of complex mathematics.

The rest of the paper goes as follows: Sect. 2 analyses the motivation for this study, Sect. 3 presents the theoretical framework of the newly suggested estimation procedure, and Sect. 4 provides empirical evidence of the accuracy of the CDNVaR model and the backtesting results under the current legislation. Section 5 presents the theoretical framework of the new approach/model and empirically shows that the new model is not only significantly more accurate than the CDNVaR models, but also that its estimations are comparable to advanced models such as the widely applied GARCH(1,1) model. Section 6 concludes the study.

---

[3] This issue is linked to the optimal transparency problem (Bouvard et al., 2015), according to which transparency increases during crises (and this leads to increased VaR estimations), but it is lower during normal periods (and this has a destabilizing effect).

2 The Motivation for this Study

Computer science and advanced estimation techniques enable us to predict and optimize several of our procedures (Aggarwal et al. 2020; Bhardwaj et al. 2020; Sagar et al. 2020). However, such advances in modeling procedures may sometimes have the opposite than the desired outcome in the social sciences. Computer science enables scholars and financial engineers to use mathematical fitting procedures and allows us to design a complex model which is capable of describing a dataset, but which lacks the appropriate theoretical background. This approach may be accurate when a specific sample is tested, but a model that lacks solid theory will prove inaccurate in the long run.

Moreover, the complex models: (1) are expensive to implement in the industry, (2) need expertise to take advantage of all of their capabilities, which means highly paid risk analysts, and (3) often require a large number of costly data. These costs could be (and usually are) limited or avoided, especially in medium or small-sized financial institutions.[4]

Taking into consideration all the above mentioned, we understand that in order to narrow the theory–practice gap we should present an easy to apply and communicate method which succeeds in producing more accurate VaR estimations. The reason is simple: no matter how accurate and representative the academic models are, only when these models can be easily applied in the financial industry, will they contribute to financial stability.

3 Data Filtering Approach Based on Financial Theory for More Accurate VaR Models

There are studies in the financial literature that either filter the models (Sarma et al. 2003) or the data (Vasileiou, 2017, 2019). This paper follows the data filtering approach: it assumes that even the simplest and conventional models, such as the CDNVaR model, could be accurate when the appropriate data inputs are chosen. Computer scientists usually describe this issue using the phrase "Garbage in, Garbage out".

When we use the CDNVaR model, we assume that the upcoming performance is similar to the last x days (data inputs). The x-day observation period is usually subject to some limitations because: (a) legislation suggests that at least a 250-day observation period should be used in order to estimate VaR (BCBS, 1995, p. 11; CESR, 2010, p. 26),[5] and (b) for consistency issues, scholars and practitioners

---

[4] The risk department in a financial institution and/or an asset management company is usually a cost center department, thus such companies try to reduce its costs. The cost benefit analysis is beyond the scope of this paper.

[5] Basel accepts a 125-day average time lag of the individual observations when a weighting scheme is used (BCBS, 2006, p. 195–196).

usually use a standard x-observation data inputs period.[6] In our perspective at least two counterarguments emerge:

- why is an observation period equal to or longer than 250 days considered more appropriate than a shorter one?, and
- why should a x-day period be presumed to generate the most accurate VaR estimations? In one sub-period, longer (or smaller) data inputs may generate more accurate VaR estimations than the VaR that is estimated by the last x days.

Hendricks (1996) provides empirical evidence that simple and conventional VaR models capture market trend changes faster when they are based on the most recent observations than the respective models that are based on longer-term historical data. However, as he shows, the conventional VaR models that are based on short observation periods present an increased number of low deviated violations/overshootings due to measurement error. Therefore, according to his findings, there is a trade-off between representativeness and accuracy. This trade-off is significantly influenced by the length of the observation period which is used in the estimation process. Vasileiou and Samitas (2020) apply the legal requirements and confirm the findings by Hendricks (1996) on short vs. long observation periods.[7]

Our goal is to examine whether a model, the CDNVaR model in our study, could become more representative (by capturing market trend changes) and more accurate (by presenting fewer overshootings) when the appropriate data length is used as inputs in the estimation process.

4 Theoretical Background, Statistics, and Empirical Evidence of How the Conventional Delta Normal Model works in the Legislative Framework

The CDNVaR method assumes that all asset returns are normally distributed, and the VaR is estimated as below

$$VaR_{t+1} = -E(r) - 2.326 \times s.d. \ of \ the \ daily \ returns \ of \ the \ Index_{t,t-x}, \quad (1)$$

where E(r) is the Expected Change in Index Value which is usually measured by the average daily returns of the last x days, 2.326 is the coefficient at the 99% confidence level (c.l.),[8] and s.d. is the standard deviation of the daily returns of the Index on day t, when x number of lags are used in the estimation process.

---

[6] The terms x-day period, data inputs, and observations considered the same in our study.

[7] When shorter observation periods are used in a CDNVaR model, the VaR estimations are representative of the examined financial conditions, but present an increased number of low deviated violations (measurement error).

[8] Legislation requires VaR estimations at the 99% c.l., and in this study we follow this requirement.

Table 1 Descriptive statistics of the S&P500 daily returns during 2000–2020

	Daily returns
Mean	0.03%
Median	0.06%
Maximum	11.58%
Minimum	− 11.98%
SD	1.25%
Skewness	− 0.153
Kurtosis	13.720
Jarque–Bera	25,322.91 (0.00)*
Observations	5284

*Statistical significance at 1% significance level

According to the CDNVaR model, it is assumed that the previous x-day volatility[9] is representative of the future financial risk. Under the Efficient Market Hypothesis (Fama, 1970), stock prices incorporate all the available information, and the best prediction for the next day's price is the current price. Therefore, the Expected change in portfolio/asset value is usually considered to be zero, which is a reasonable assumption for a short holding period (Linsmeier & Pearson, 2000, p. 53). Thus, the following equation is usually applied in the financial industry:

$$VaR_{t+1} = -2.326 \times s.d._{t,t-x} \tag{2}$$

As Eq. 2 shows, the VaR estimation is exclusively depended on the *sd* value, which in turn depends on the observation period from which the *sd* is calculated. The legal framework requires that the history of risk factors should be at least 1 Year (BCBS, 1995, p. 9; CESR, 2010, p. 26),[10] hence there are regulatory limitations on the only parameter that can significantly influence the estimations of the VaR model.

Could the CDNVaR model become more accurate and representative if we remove these limitations? Vasileiou and Pantos (2020) and Vasileiou (2019) show that the mandatory minimum of 250 data inputs may lead to inaccurate VaR estimations,[11] and following their suggestions we remove the 250-observation threshold from our estimation process.

Table 1 presents the descriptive statistics of the S&P500 returns during the examined period, 2000–2020, and shows that the returns are not normally distributed: skewness lower than zero, Kurtosis higher than 3, and the Jarque–Bera test quantitatively confirms that the times series does not follow the normal distribution. Figure 1 shows the performance and the daily returns of the S&P500 index during the

---

[9] Statistically, standard deviation is the square root of volatility. Hereafter these terms (not the quantitative measures) are considered the same.

[10] Approximately 250 daily observations.

[11] The law requires an evaluation of the VaR estimations (backtesting procedure) on at least a monthly basis. If a model is inaccurate, it should be instantly modified, regardless of the number of observations used.
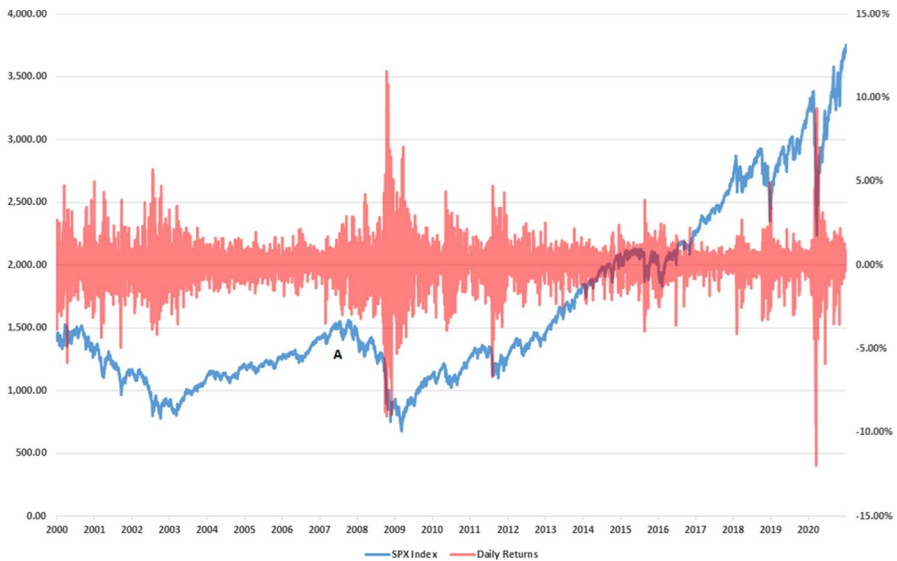
Fig. 1 The performance and the daily returns of the S&P500 Index during the 2000–2020 period

examined period (2000–2020). We can observe some features that are common in financial time series: (1) the leverage effect because when the prices of the index fall, volatility increases (spikes of the daily returns are bigger), and (2) volatility clustering because volatility appears in bunches (Brooks, 2019). These findings raise concerns about the accuracy of the CDNVaR model because it generates linear VaR estimations (Eq. 2).

The backtesting procedure could confirm (or not) our concerns that the CDNVaR model may not fit our data sample. According to the legal framework, a VaR model is considered accurate when up to four overshootings per 250 VaR estimations (approximately a year) are documented (BCBS, 1996; CESR, 2010). More than 4 overshootings per year ("4 overshootings rule") mean that the model is not accurate at the 99% c.l. and a revision is required, and/or penalties should be imposed.[12]

Table 2 shows several versions of the CDNVaR model, when 20-, 125-, 250-, and 500-day observation periods are used. Part (a) of Table 2 presents some descriptive statistics on the VaR backtesting procedure that are not examined by the legislation, but are useful to our analysis. The results indicate:

---

[12] The four overshootings rule per year (250 observations) backtesting procedure that the legal framework suggests is similar to the Kupiec test (1995) which is based on the binary distribution and examines only the number of oveshootings during an examined period. Some scholars suggest that if the "accuracy issue" is limited only to the "less than four violations rule", several moral issues emerge. Popularizing some more advanced backtesting suggestions (Mc Neil and Frey 2000; Mc Neil et al. 2005), in this study we examine not only the overshooting number but also the difference/deviation between (a) the real losses and the estimated VaR, and (b) the real returns and the estimated VaR. The smaller these deviations are, the more representative of the examined financial conditions the VaR estimations are.

Table 2 Back-testing results (99% c.l.) of the conventional delta-normal models

Total period	20-day CDNVaR	125-day CDNVaR	250-day CDNVaR	500-day CDNVaR
(a) Descriptive statistics of the CDNVaR versions				
Average VaR	− 2.43%	− 2.57%	− 2.62%	− 2.65%
SD	1.65%	1.38%	1.21%	1.01%
Min	− 14.23%	− 8.52%	− 6.69%	− 5.20%
Max	− 0.48%	− 0.93%	− 0.97%	− 1.35%
Violations	134	120	129	130
Average value of deviations when an overshooting is reported	− 0.69%	− 0.97%	− 0.99%	− 1.16%

Year	20-day CDNVaR	125-day CDNVaR	250-day CDNVaR	500-day CDNVaR
(b) Backtesting results per trading year				
2000	*4*	*6*	*5*	4
2001	3	2	3	3
2002	4	*5*	*5*	*5*
2003	2	0	0	1
2004	*6*	0	0	0
2005	4	2	1	0
2006	4	*5*	4	4
2007	*10*	*12*	*15*	*15*
2008	*8*	*12*	*21*	*28*
2009	*6*	0	0	2
2010	*8*	*7*	*6*	0
2011	*7*	*10*	*10*	*9*
2012	*5*	3	1	0
2013	*9*	3	2	0
2014	*11*	*10*	*10*	*7*
2015	*6*	*7*	*7*	*9*
2016	2	2	4	*5*
2017	*5*	4	3	0
2018	*10*	*16*	*15*	*20*
2019	*8*	3	4	*5*
2020	*12*	*11*	*13*	*13*
Sum of overshootings	123	107	116	118
More than 4 overshootings per year	14	11	10	10

Part (b) reports the backtesting results. Italics indicate the years in which the "4 overshootings rule" is violated

- there is an increased number of violations especially after the 2007 crisis, which means that stock markets have presented more non-linearities and fat tails since then,

- the shorter versions of the CDNVaR models (20- and 125-day CDNVaR) present increased adjustment to the financial environment (the highest standard deviations of the VaR estimations), and the lowest average value of deviations[13] when an overshooting is reported (hereafter "average deviation"). However, due to the measurement error violate the "4 overshootings rule" more times than the other CDNVaR models,
- the longer versions (250- and 500-day CDNVaR) seem to be the most accurate versions amongst the examined, but when overshootings are observed the models present the highest average deviation, and

According to the legislation, the 250-day and the 500-day CDNVaR models are considered the most accurate amongst the examined models. Both of them fail to meet the 4 overshootings rule in 10 out of the 21 years examined,[14] but they still outperform the 20- and the 125-day CDNVaR which fail to meet the 4 overshootings rule 14 and 11 times, respectively. However, in our opinion, the 250-day CDNVaR model is better than the 500-day CDNVaR model for at least two reasons. Firstly, the average deviation in the 250-day CDNVaR is narrower than the respective measure in the 500-day CDNVaR (on average 0.99% vs 1.16%).[15] Secondly, the narrower and average deviations cannot be attributed to more conservative estimations (the average value of the VaR estimations, hereafter "average VaR", are 2.62% and 2.65%, for the 250-day and 500-day CDNVaR, respectively).[16]

We consider the 250-day CDNVaR the best amongst the examined models when the total period is examined. However, it is not always the most accurate amongst the examined versions in some sub-periods, e.g. in 2010 and 2016 the 500-day version does not violate the 4 overshootings threshold. Therefore, taking into consideration all the above mentioned findings and comments, the basic hypothesis of our study emerged: could adjustments to the number of observations lead to more accurate VaR estimations?

5 The Newly Suggested Data Filtering Approach: Theory, Practice, and Improved Accuracy

(a) Could the newly suggested approach improve the accuracy of the CDNVaR model?

---

[13] A deviation is equal to next day's real loss minus VaR estimation.

[14] The high number of violations is an indication that the linear CDNVaR model is not appropriate for our dataset.

[15] When overshootings are documented, Vasileiou (2019) argues that a model should not present increased deviations, provided that the model is not too conservative relative to other examined models.

[16] If a model is much more conservative than the CDNVaR models we present in Table 2(a), e.g. a model with an average VaR value 4%, it will possibly present fewer overshootings than the examined model. It will be more accurate according to the law, but in practice it is just more conservative.

Table 3 Correlations between next day returns and the standard deviations

Period	Next day return and standard deviation			
	20-day standard deviation (%)	125-day standard deviation (%)	250-day standard deviation (%)	500-day standard deviation (%)
(a) Correlation between next day returns and the current day standard deviation				
1979–1999	2.51	2.20	1.05	0.86
2000–2020	0.05	− 0.12	0.60	0.95
(b) Correlation between next day losses and the current day standard deviation				
1979–1999	− 26.75	− 20.75	− 18.13	− 12.64
2000–2020	− 47.35	− 37.57	− 30.04	− 22.12

The previous results point to certain drawbacks in the CDNVaR model and underline the need to revise current legal guidelines. Regarding the model, it is known that the CDNVaR may be accurate in normal periods, but fails when fat tails and non-linearities are observed (Jorion, 2007). The empirical evidence confirmed our assumptions that a linear model, such as the CDNVaR, is not appropriate for our data sample (the best versions of the model fail to meet the 4 overshootings rule in 10 out of the 21 years). How can we make the CDNVaR model non-linear and more accurate?

One avenue to explore is to waive the legal requirement that requires at least 250 observations for estimating VaR and examine whether this can potentially generate more accurate VaR estimations. The findings of Table 2 show that shorter observation periods may produce to more accurate VaR estimations when a financial trend changes because fewer observations lead to significant changes in the s.d. and VaR estimations from day to day (Eq. 2). This allows the model to capture market changes faster than the models that use longer observation periods (Vasileiou & Samitas, 2020). However, the CDNVaR models that are heavily based on recent observations fail when normal growth periods are examined due to the measurement error (Hendricks, 1996). Can we reduce the measurement error and improve the CDNVaR estimations? A data filtering approach could be the solution. When the financial conditions are normal the model should adjust data length in order to avoid the measurement error issue. Similarly, when the financial trend changes the model should use a short observation period in order to accurately capture these changes in the financial conditions. So, we should not use a constant data length in order to estimate VaR.

Therefore, our goal is to find a procedure/algorithm which selects the most appropriate dataset, so we need to examine in detail how the CDNVaR model works. According to Eq. 2, s.d. is crucial for accurate VaR estimations, thus data length is a significant parameter for accurate VaR estimations. For this reason, we calculate the standard deviations using 20-, 125-, 250-, and 500-day observations, and we empirically test their correlations to the next day's returns. (Following Hull and White's (1998) suggestion, we first present the specific relationship for the previous 21-year period (1979–1999), and then for the 21-year period (2000–2020) under

examination. The results are presented in Table 3(a) and they show that the correlation between volatility and next day returns does not provide any significant information for the upcoming performance of the stock market, even though Fig. 1 shows the existence of the leverage effect. This means that we should make some adjustments in order to appoint the leverage effect.

The CDNVaR model relies on past volatility in order to estimate a future worst-case scenario. The specific relationship should be negative because according to the leverage effect standard deviations should increase when crises are coming, and vice versa. VaR tries to capture the worst-case scenario, so perhaps the negative relationship between next day returns and the volatility of the current day can become clearer when we only include negative return days in the correlation estimations.

Applying this adjustment to the examined dataset, the leverage effect is revealed in Table 3(b) where the correlation between next day losses and volatility is negative. Moreover, the results show that the standard deviations of the shorter-term periods (20-day and 125-day) are more representative of the real financial risk than the volatility that is estimated using longer-term observation periods (250-day and 500-day).

As far as the optimal data length period is concerned, we can observe that a constant period of data length is not always appropriate. We assume that we are examining the VaR in a period during which the market trend changes, such as the one noted by point A in Fig. 1, and we apply the 250-day version of the CDNVaR model. The 250-day CDNVaR model in this case uses the last 250 observations of a growth period with low standard deviation in order to estimate the VaR of an upcoming crisis period. In this case, the longer the observation period is, the more inaccurate the VaR estimations (Garbage In, Garbage Out). As the results of Table 2 show, during the year 2007–2008 (to which point A belongs) the longer the observation periods are, the higher the number of overshootings. The model is the same; the data observation period makes the different versions either more accurate or less accurate. Thus, we should not always use a minimum of 250 observations and we should not always use the same number of observations.

The idea of the non-standard number of data observations has been documented in previous studies. Hull and White (1998) suggest that volatility should not only be incorporated in the VaR estimation procedure, but it should also be adjusted in order to reflect the changes in volatility over the historical time period. Duffie and Pan (1997) suggest that when the financial trend changes, each day in the data series should not carry the same weight. Vasileiou (2017) shows that the more representative the data inputs, the better (more accurate and representative) the VaR estimations of the Historical VaR model. Bekiros et al. (2019) show that parametric models are more appropriate for long-term investment horizons, but are not accurate as non-parametric methods for short term investment strategies.

Therefore, in the case of the CDNVaR model, a longer period may be appropriate when there is a long-term trend because it reduces the measurement error. However, when the trend changes, from growth to recession (and vice versa), a version that is based on fewer observations may be more appropriate/representative of the current financial conditions because it can capture changes in financial trends faster than a version which is based on longer observation periods.

The aforementioned studies suggest that there should be a differentiation in the observation weightings and/or the number of observations. Which criterion and which model is the most appropriate? Halbleib and Pohlmeier (2012) provide empirical evidence that data driven combination weights deliver accurate VaR forecasts and Ziggel et al. (2014) show that structural breaks instantly capture market trend changes and improve the accuracy of VaR estimations. However, the latter are not applied in the financial industry due to mathematical complexity.

In order to estimate the optimal and the most representative observation period, but at the same avoid using complex mathematical equations, we have created an optimization algorithm based on financial theory, which works as follows.

i.   it calculates the *s.d.* using i-lag-periods, where *i* is more than 20 days, and up to 1,000-day observations,

ii. it calculates the correlations of the last 20 and up to 250 pairs of negative returns[17] and their respective standard deviations of the previous day,

iii. from the abovementioned correlations, the algorithm selects the optimal i*-period which corresponds to the correlation with the minimum value in order to capture the negative fat tails and the leverage effect

$$\min(correlation_n(negative\ return\ days;\ i-day\ standard\ deviation), \qquad (3)$$

where n = 20, 21,…,250 last pairs of negative returns and standard deviations values, and

iv. the optimal VaR following Eq. (2) is estimated by the optimal lag

$$VaR_{t+1} = -2.326 \times s.d._{\cdot t, t-i*}, \qquad (4)$$

where t = the VaR estimation for the day t, and i* = the optimal sd period of each day.

The algorithm examines the data and suggests an optimal data length observations period; this leads to an optimal *sd* for more accurate VaR estimations. That is why we use the name Optimal Standard Deviation Delta Normal VaR (OSDDN-VaR) to refer to this new version of the Delta Normal model. This way depending on the financial trend, the data input period varies, and the CDNVaR model turns from parametric to non-parametric.

The newly suggested model is not mathematically complex. It just employs an additional data filtering method which is based on the leverage effect and slightly modifies the CDNVaR model and turns it into the OSDDNVaR. With this filtering procedure we do not torture the data, but we listen to…what they want to tell us. Figure 2 graphically presents the steps of the estimation procedure of the OSDDN-VaR model, and Table 3 shows the results of the OSDDNVaR model.

If we compare the results of the CDNVaR models in Table 2 and the results of the OSDDNVaR model in Table 4, we observe that the newly suggested approach

---

[17] We use the correlations between s.d. and the negative return days in our optimization process because this combination better captures the leverage effect (Table 3).
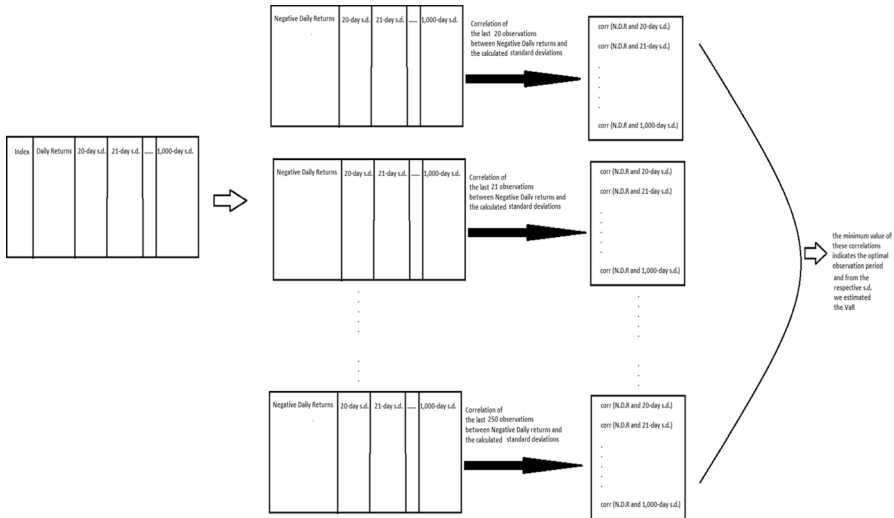
Fig. 2 Graphical presentation of the estimation procedure of the OSDDN VaR model

significantly improves the accuracy of CDNVaR models. Particularly, the OSDDN-VaR model presents the lowest number of overshootings during the examined period and violates the "4 overshootings per year" threshold in 6 years out of the 21 years examined, while the best CDNVaR model, the 250-day CDNVaR model, violates the respective rule in 10 years in the same examined period. Moreover, it presents on average lower deviations when an overshooting is recorded (the average value the OSDNVaR Violations Differences is -0.79%, and the respective value of the 250-observation CDNVaR model is -0.99%).

(b) Could the estimations of the OSDDNVaR model be comparable to the estimations of advanced models?

The main goal of our study has been accomplished: the filtering procedure significantly improved the accuracy of the CDNVaR model. The next step in our study is to examine whether the accuracy of OSDDNVaR model could be compared to advanced VaR model such as the GARCH (1,1). We use the following set of equations

$$r_t = \mu + \varepsilon_t \tag{5}$$

$$\varepsilon_t = z_t \times \sigma_t \tag{6}$$

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + a_2 \sigma_{t-1}^2, \tag{7}$$

where ε_t is the error term, z_t is an i.i.d. random variable with zero mean and variance one. σ_t^2 is the conditional variance. The results of the GARCH (1,1) model which uses 500 observations are presented in Table 3. In contrast to Table 2, we do not

Table 4 Back-testing results (99% c.l.) of the OSDDNVaR and GARCH(1,1) models

Total period	OSDDNVaR model	GARCH(1,1) 500-day
(a) Descriptive statistics of the OSDDNVaR and GARCH(1,1) versions		
Average VaR	−2.67%	−2.72%
SD	1.46%	1.68%
Min	−14.19%	−19.57%
Max	−0.71%	−0.65%
Violations	85	75
Average value of deviations when an overshooting is reported	−0.79%	−0.65%

Year	OSDDNVaR model	GARCH(1,1) 500-day
(b) Backtesting results per trading year		
2000	4	2
2001	2	2
2002	3	2
2003	0	0
2004	4	1
2005	1	2
2006	2	3
2007	*10*	*8*
2008	*8*	*5*
2009	0	1
2010	2	4
2011	*7*	*6*
2012	0	2
2013	2	2
2014	4	*5*
2015	*6*	*7*
2016	2	2
2017	2	4
2018	*11*	*7*
2019	4	4
2020	*11*	*6*
Sum of overshootings	85	75
More than 4 overshootings per year	6	7

Part (b) reports the backtesting results. Italics indicate the years in which the "4 overshootings rule" is violated

include 20-day, 125-day, and 250-day observation versions for the GARCH (1,1) model due to issues with the Hessian matrix. Most studies that examine the issue of the appropriate length of inputs in GARCH models use more than 250 observations

because shorter observation periods lead to problems of convergence of Maximum Likelihood Estimator (MLE) for obtaining GARCH parameters (Angelidis et al. 2004; Lundbergh & Teräsvirta, 2002).[18]

The results show that the GARCH(1,1) model is more conservative than the OSDDNVaR model (average value of VaR estimations with GARCH (1,1) model is -2.72% and -2.67% with the OSDDNVaR model). This could be an explanation for the lower number of violations that the GARCH (1,1) model presents during the examined period relative to the OSDDNVaR model (number of total overshootings GARCH(1,1) 75 vs OSDDNVaR 85), as well as an explanation for the lower deviations when violations are observed (GARCH average deviation 0.79% vs 0.65%).

On the one hand, the OSDDNVaR outperforms the GARCH(1,1) according to the legislation because the former model fails to meet the 4 overshootings threshold per year in 6 out of the 21 examined years, while the latter fails in 7 out of the 21 years. On the other hand, we should mention that when the OSDDNVaR model fails to meet the "4 overshootings" rule, it presents in most of the cases a higher number of overshootings per year than the GARCH(1,1).[19]

Therefore, the newly suggested approach which is based on financial theory and does not use complex mathematics, not only significantly improves the VaR estimation of a conventional model, but also appears to be as accurate as the advanced and widely accepted GARCH(1,1) VaR model. The non-constant observation period offers flexibility to the model and allows it to use the most appropriate data length period in order to provide the best possible VaR estimations. This approach could be applied in advanced models, such as the GARCH family models. The filtering method will again suggest the optimal data observation period and the accuracy of the estimations will further improve.

6 Conclusions

This study shows that a model may be more accurate when the appropriate data are used in the estimation process. When the CDNVaR model is applied, the accuracy of the model's estimation depends on the appropriate data length. On the one hand, a short observation period may be more appropriate for accurate VaR estimations during periods of changes in the financial trend, but during normal periods the short observation data inputs may lead to measurement errors. On the other hand, longer observation periods may be a solution for the measurement error, but when the financial trend changes, there are a lot of inappropriate data included in the estimation process, which leads to inaccurate VaR estimations for the upcoming period.

Thus, why should we always use the same data length? We employ a new algorithm which is based on the "leverage effect" and which enables us to use the most

---

[18] Thus, in contrast to the CDNVaR model for which fewer than 250 observations may lead to more accurate VaR estimations, the GARCH(1,1) model needs larger data samples in order to be econometrically valid and to generate more accurate VaR estimations.

[19] e.g. in 2007, 2018, and 2020, the OSDDNVaR model presents more than 10 overshootings per year, while the max number of overshootings per year for the GARCH model is 8.

appropriate data length depending on the financial conditions. The results show that this procedure significantly improves the accuracy of the CDNVaR model and the new OSDDNVaR version is as accurate as a GARCH (1,1) model. This procedure, with the appropriate adjustments, could also be used with other advanced models, such as a GARCH family model, in order to improve their accuracy.

Our conclusion is that in many cases the models are not inaccurate; the data inputs make them inaccurate (Garbage In, Garbage Out). Complex mathematical models may generate to more accurate VaR estimations, but they have many drawbacks in their implementation. These drawbacks lead to theory–practice gaps. A data filtering approach could provide a computational solution based on financial theory and produce more accurate VaR estimations without resorting to complex mathematical models. This way we present an approach that significantly improves the accuracy of the existing VaR models, can be easily applied and understood, and can contribute to financial stability.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Informed consent The research is based only on data, no human/animal participation.

References

Aggarwal, A., Alshehri, M., Kumar, M., Alfarraj, O., Sharma, P., & Pardasani, K. R. (2020). Landslide data analysis using various time-series forecasting models. *Computers & Electrical Engineering, 88*, 106858

Allen, F., Chui, M. K., & Maddaloni, A. (2004). Financial systems in Europe, the USA, and ASIA. *Oxford Review of Economic Policy, 20*(4), 490–508

Angelidis, T., Benos, A., & Degiannakis, S. (2004). The use of GARCH models in VaR estimation. In *Statistical methodology* (vol. 1, pp. 105–128).

Basel Committee on Banking Supervision. (1995). An internal model-based approach to market risk capital requirements. Basel, Switzerland. Retrieved from https://www.bis.org/publ/bcbsc224.pdf.

Basel Committee on Banking Supervision. (1996). Supervisory framework for the use of backtesting in conjunction with the internal models approach to market risk capital requirements. Basel, Switzerland.

Basel Committee on Banking Supervision. (2006). International convergence of capital measurement and capital standards. A revised framework. Comprehensive version. Bank for International Settlements, Basel, Switzerland.

Bekiros, S., Loukeris, N., Eleftheriadis, I., & Avdoulas, C. (2019). Tail-related risk measurement and forecasting in equity markets. *Computational Economics, 53*(2), 783–816

Berkowitz, J., Christoffersen, P., & Pelletier, D. (2011). Evaluating value-at-risk models with desk-level data. *Management Science, 57*(12), 2213–2227

Bhardwaj, A., Shah, S. B. H., Shankar, A., Alazab, M., Kumar, M., & Gadekallu, T. R. (2020). Penetration testing framework for smart contract Blockchain. *Peer-to-Peer Networking and Applications, 5*, 1–16

Billio, M., & Pelizzon, L. (2000). Value-at-risk: a multivariate switching regime approach. *Journal of Empirical Finance, 7*(5), 531–554

Bouvard, M., Chaigneau, P., & Motta, A. D. (2015). Transparency in the financial system: Rollover risk and crises. *The Journal of Finance, 70*(4), 1805–1837

Brooks, C. (2019). *Introductory econometrics for finance*. Cambridge University Press.

Committee of European Securities Regulators. (2010). CESR's guidelines on risk measurement and the calculation of global exposure and counterparty risk for UCITS. CESR/10-788.

Degiannakis, S., Floros, C., & Livada, A. (2012). Evaluating value-at-risk models before and after the financial crisis of 2008: International evidence. *Managerial Finance, 38*(4), 436–452

Duffie, D., & Pan, J. (1997). An overview of value at risk. *Journal of derivatives, 4*(3), 7–49

Engle, R. (2004). Risk and volatility: Econometric models and financial practice. *The American Economic Review, 94*(3), 405–420

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383–417

Fama, E. F. (1995). Random walks in stock market prices. *Financial Analysts Journal, 51*(1), 75–80

Halbleib, R., & Pohlmeier, W. (2012). Improving the value at risk forecasts: Theory and evidence from the financial crisis. *Journal of Economic Dynamics and Control, 36*(8), 1212–1228

Hendricks, D. (1996). Evaluation of value-at-risk models using historical data (digest summary). *Economic Policy Review Federal Reserve Bank of New York, 2*(1), 39–67

Hull, J., & White, A. (1998). Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of Risk, 1*(1), 5–19

Jorion, P. (2007). *Financial risk manager handbook*. (Vol. 406)John Wiley & Sons.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives, 3*(2), 73–84

Linsmeier, T. J., & Pearson, N. D. (2000). Value at risk. *Financial Analysts Journal, 56*, 47–67

Longerstaey, J., & Spencer, M. (1996). *Riskmetricstm—technical document*. Morgan Guaranty Trust Company of New York.

Lundbergh, S., & Teräsvirta, T. (2002). Evaluating GARCH models. *Journal of Econometrics, 110*(2), 417–435

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance, 7*(3–4), 271–300

McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative risk management: Concepts, techniques and tools*. (Vol. 3)Princeton University Press.

Moussa, A. M., Kamdem, J. S., & Terraza, M. (2014). Fuzzy value-at-risk and expected shortfall for portfolios with heavy-tailed returns. *Economic Modelling, 39*, 247–256

Ross, S. A. (1993). Is beta useful? In *AIMR conference proceedings* (vol. 1993, no. 6, pp. 11–15). Association for Investment Management and Research.

Sagar, S., Stamatiadis, N., Wright, S., & Green, E. (2020). Use of codes data to improve estimates of at-fault risk for elderly drivers. *Accident Analysis & Prevention, 144*, 105637

Sarma, M., Thomas, S., & Shah, A. (2003). Selection of Value-at-Risk models. *Journal of Forecasting, 22*(4), 337–358

Vasileiou, E. (2016). Overview of the Greek value at risk (VaR) legislation framework. *Journal of Financial Regulation and Compliance, 24*(2), 213–226

Vasileiou, E. (2017). Value at Risk (VaR) historical approach: Could it be more historical and representative of the real financial risk environment? *Theoretical Economics Letters, 7*, 951–974. https://doi.org/10.4236/tel.2017.74065

Vasileiou, E. (2019). Accuracy versus complexity trade-off in VaR modeling: Could technical analysis be a solution? *Journal of Financial Management, Markets and Institutions, 7*(02), 1950003

Vasileiou, E., & Pantos, T. (2020). What do the value-at-risk measure and the respective legislative framework really offer to financial stability? Critical views and pro-cyclicality. *European Journal of Economics and Economic Policies: Intervention, 17*(1), 39–60

Vasileiou, E., & Samitas, A. (2020). Value at risk, legislative framework, crises, and procyclicality: What goes wrong? *Review of Economic Analysis, 12*(3), 345–369

Zhang, H. G., Su, C. W., Song, Y., Qiu, S., Xiao, R., & Su, F. (2017). Calculating Value-at-Risk for high-dimensional time series using a nonlinear random mapping model. *Economic Modelling, 67*, 355–367

Ziggel, D., Berens, T., Weiß, G. N., & Wied, D. (2014). A new set of improved Value-at-Risk backtests. *Journal of Banking & Finance, 48*, 29–41