Evaluation of Selected Models for Value at Risk Calculation

Vanessa Bormann, Matthias Gehrke and Karsten Luebke

Abstract We compared different newer models (e.g. CAViaR and one of the most recent approaches HAR-QREG) to the more traditional approaches (e.g. RiskMetrics and GARCH(1,1)) for value at risk calculation. As samples for different asset classes we chose MDAX and CDAX as representatives for the German capital market, gold, Brent crude oil, wheat, and corn for alternative investments, and the EUR/USD exchange rate representing the currency market. The prediction quality of each model was tested using back testing methods like the conditional coverage and dynamic quantile test. It turned out that the newer models are able to outperform the traditional approaches, but all fail to model corn return due to an extreme price drop.

Vanessa Bormann
FOM University of Applied Science, Leimkugelstraße 6, 45141 Essen, Germany
✉ bormann.vanessa@gmail.com

Matthias Gehrke
FOM University of Applied Science, Leimkugelstraße 6, 45141 Essen, Germany
✉ matthias.gehrke@fom.de

Karsten Luebke
FOM University of Applied Science, Leimkugelstraße 6, 45141 Essen, Germany
✉ karsten.luebke@fom.de

Archives of Data Science, Series A
(Online First)
KIT Scientific Publishing
Vol. 4, No. 1, 2018

DOI 10.5445/KSP/1000085951/08
ISSN 2363-9881

1 Introduction

Value at Risk (VaR) is the amount of loss for a given day t based on the asset price p_{t-1} of the day before, that will not be exceeded with a probability $1 - \alpha$. Banks in particular are obliged by regulation to estimate the risk of their investments. VaR is (still) the method of choice to fulfil these regulatory requirements.

The estimation of future risk is based on the historical development of asset prices. Thus, the challenge is to forecast future asset prices and the related maximum loss. Especially for market risks of trading assets the continuous monitoring of market development and risk potential is indispensable. Owing to regulatory requirements, banks have to compare daily their assets price development and their estimated risk by performing back tests (Hannemann et al, 2013).

1.1 Literature Review

Some research had been done in regard to the forecasting precision of various VaR estimation methods. For example[1] Hansen and Lunde (2005) compared several ARCH methods against GARCH(1,1) for DM/USD exchange rates and IBM returns and found no evidence that GARCH(1,1) can be outperformed. Bao et al (2006) investigated several models, among them RiskMetrics and CAViaR, for emerging markets before, during, and after the financial crisis of 1997–1998. Results show that the RiskMetrics approach performed quite well in tranquil periods while CAViaR seemed to be more stable over the various time periods. The focus of the research by Allen et al (2012) was the CAViaR approach applied to some Australian stocks and Australian stock exchange indices. This was compared to GARCH(1,1), RiskMetrics, and an APARCH model. Overall the CAViaR approach seemed to be superior. Bilandi and Kudła (2016) compared GARCH(1,1) and several other GARCH approaches, (filtered) Historical Simulation, and Extreme Value Theory using major international stock exchange indices. Extreme Value Theory and Historical Simulation performed best in their studies. Finally, Haugom et al (2016) proposed a new approach HAR-QREG adopting quantile regression and compared it to Historical

[1] The selection is not meant to be complete but more in regard to our study approach. Also, the sorting is not a qualitative judgement, but just by publication year.

Simulation, APARCH, RiskMetrics, and CAViAR models applied to USD/GBP exchange rates, S&P 500, and IBM returns. HAR-QREQ outperformed Historical Simulation and RiskMetrics and showed simular performance to the more sophisticated models like APARCH and CAViaR.

Our contribution to this field of research is picking the most promising modern approaches like CAViaR and HAR-QREG and comparing it to classical methods like Historical Simulation, RiskMetrics, and GARCH(1,1). Furthermore, we substantially widened the range of assets. This gives some new insight of applicability of the various VaR estimation methods for practitioners.

1.2 VaR Estimation

Even if VaR is conceptually easy to understand, the estimation is statistically quite challenging. As empirical results show, financial time series possess some notable characteristics that should be considered by VaR models. These stylized facts are weak or non-stationary processes, clustered volatilities, and left-skewed return distributions that determine fat tails. VaR models in general differ in distribution assumptions, estimation parameters, and in the overall treatment of available historical information. Therefore, several methods for the estimation of VaR exist. We compare eight different models for VaR estimation.

Historical Simulation

Historical Simulation (e.g. Allen et al, 2004; Dimitrakopoulos et al, 2010) simply takes the α-quantile Q_α of the returns, sorted in increasing order:

$$VaR_t = Q_\alpha \left(\{r_i\}_{t-1}^{w} \right), \tag{1}$$

where $\{r_i\}_{t-1}^{w}$ are the sorted returns from $t - w - 1$ to $t - 1$, w is the estimation window size. This is a non-parametric method, no assumption is made on the return distribution. The results depend strongly on window size, especially with small α values.

RiskMetrics

RiskMetrics (e.g. Bao et al, 2006; J.P.Morgan and Reuters, 1996) assumes normally distributed returns. By using the EWMA (exponentially Weighted Moving Average) approach volatility clusters will be considered in the model.

With the α-quantile Q_α^N of the normal distribution, VaR can be calculated as follows:

$$\begin{aligned} \sigma_t^2 &= \lambda \sigma_{t-1}^2 + (1-\lambda)\, r_{t-1}^2, \\ VaR_t &= Q_\alpha^N \cdot \sigma_t . \end{aligned} \tag{2}$$

In the RiskMetrics approach λ is set to 0.94, which implies a very strong dependence on the variance of the day before. σ_t^2 is calculated recursively within the estimation window, starting with the squared return of the day before as initial value for the first variance.

GARCH(1,1)

GARCH models consider two stylized facts of financial time series: Volatility clusters and fat tails. In a GARCH(1,1) model the variance of day t depends on the squared return and the variance of the day before, as well as on the long term variance σ_L^2 which ist part of the parameter ω:

$$\begin{aligned} \sigma_t^2 &= \omega + \alpha \cdot r_{t-1}^2 + \beta \cdot \sigma_{t-1}^2, \\ \omega &= \sigma_L^2 \cdot (1-\alpha-\beta)\,, && (3) \\ VaR_t &= Q_\alpha^N \delta_A . && (4) \end{aligned}$$

The parameters ω, α and β are recursively estimated from given returns in the estimation window by maximum likelihood. The first value of the variance must be set to an initial value, e.g. the squared return of the day before. VaR is calculated from the standard deviation using the α-quantile of the normal distribution. An overview of the usage of GARCH models in VaR estimation is given by e.g. Angelidis et al (2004); Chambers et al (2014); Hartz et al (2006).

Conditional autoregressive value at risk

The conditional autoregressive value at risk (CAViaR) approach by Engle and Manganelli (2004) is a set of models which calculate the quantile directly using an autoregressive process on the lagged quantile.
In the symmetric absolute value (SAV) model the following equation applies:

$$VaR_t = \beta_0 + \beta_1 \cdot VaR_{t-1} + \beta_2 \cdot |r_{t-1}|\,. \tag{5}$$

Quantile Regression is used to estimate β. To start the optimization routine, a random initialization is necessary.

In the Asymmetric Slope model positive and negative returns of the previous day are considered with different slopes:

$$VaR_t = \beta_0 + \beta_1 \cdot VaR_{t-1} + \beta_2 \cdot (r_{t-1})^+ + \beta_3 \cdot (r_{t-1})^- . \tag{6}$$

The parameter $(r_{t-1})^+$ represents a positive return of the past day, whereas $(r_{t-1})^-$ represents a respective negative return.

The indirect GARCH(1,1) (IG) approach models the VaR in dependence of the lagged squared VaR instead of the lagged variance as in standard GARCH models:

$$VaR_t = \sqrt{\beta_0 + \beta_1 \cdot VaR_{t-1}^2 + \beta_2 \cdot r_{t-1}^2}. \tag{7}$$

Contrary to the other CAViaR models, the CAViaR Adaptive includes a unit coefficient on the lagged VaR:

$$VaR_t = VaR_{t-1} + \beta_0 \cdot \left(\frac{1}{1 + e^{G \cdot \left(r_{t-1} - VaR_{t-1} \right)}} - \alpha \right). \tag{8}$$

The parameter G is some positive finite number. As in the original paper by Engle and Manganelli (2004), we set G to 10. The specific characteristic of this model is a smooth step function.

Heterogeneous autoregressive quantile regression

The heterogeneous autoregressive quantile regression (HAR-QREG) model by Haugom et al (2016) uses, like the CAViaR models, quantile regression as well. This model considers the expectations of various market actors by using volatilities over different time frames:

$$\sigma_{d,t} = \sqrt{r_t^2} \quad \text{(daily volatility)}, \tag{9}$$

$$\sigma_{w,t} = \sqrt{\frac{1}{5} \left(r_t^2 + r_{t-1}^2 + \cdots + r_{t-4}^2 \right)} \quad \text{(weekly volatility)}, \tag{10}$$

$$\sigma_{m,t} = \sqrt{\frac{1}{20} \left(r_t^2 + r_{t-1}^2 + \cdots + r_{t-19}^2 \right)} \quad \text{(monthly volatility)}. \tag{11}$$

VaR is then calculated as follows by estimating the quantile regression for the respective VaR level:

$$VaR_t = \beta_0 + \beta_1 \cdot \sigma_{d,t-1} + \beta_2 \cdot \sigma_{w,t-1} + \beta_3 \cdot \sigma_{m,t-1}. \tag{12}$$

1.3 VaR Back Testing

Basically back testing of VaR models counts the number of violations (i.e. exceeding the predicted VaR) v_1 and tests if this appears to be within given limits. The violation ratio VR is simply calculated as follows:

$$VR = \frac{v_1}{\alpha \cdot w}. \tag{13}$$

The parameter w defines the testing window size. The number of non-violating observations v_0 is simply $w - v_1$. VR should be close to one to prevent an under- or overforecast of the risk (Danielsson, 2001).

Unconditional Coverage Test

The unconditional coverage test (UC test) or Kupiec test (Kupiec, 1995) checks if the probability to exceed the VaR is not significantly different to the predefined significance level α by using a likelihood ratio test:

$$LR_{UC} = 2 \log \frac{(1-\hat{p})^{v_0} \cdot \hat{p}^{v_1}}{(1-p)^{v_0} \cdot p^{v_1}} \qquad \text{with} \qquad \hat{p} = \frac{v_1}{w} \quad \text{and} \quad p = \alpha. \tag{14}$$

LR_{UC} is asymptotically $\chi^2_{(1)}$-distributed under H_0.

Independence Test

It is also important to know whether the violations are serially independent. As the UC test does not check whether the violations are clustered or not, Christoffersen (1998) developed the independence test. The independence test considers how often ones (observation violates the VaR) are followed by ones (v_{11}) and so on.
The restricted likelihood function $\mathcal{L}_R$ is given as follows:

$$\mathcal{L}_R = (1-p_{01})^{v_{00}} \cdot p_{01}^{v_{01}} \cdot (1-p_{11})^{v_{10}} \cdot p_{11}^{v_{11}}. \tag{15}$$

The parameter v_{ij} represents the number of observations where j follows i, $i, j = 0, 1$, p_{ij} is the probability for that, e.g.

$$p_{00} = \frac{v_{00}}{v_{00} + v_{01}} \qquad p_{01} = \frac{v_{01}}{v_{00} + v_{01}} \qquad \cdots. \tag{16}$$

Under H_0 (no clustering) $p_{01} = p_{11} = p$ the unrestricted likelihood function $\mathcal{L}_U$ is the same as in the UC test:

$$\mathcal{L}_U = (1-\hat{p})^{v_{00}+v_{10}} \cdot \hat{p}^{v_{01}+v_{11}} = (1-\hat{p})^{v_0} \cdot \hat{p}^{v_1}. \tag{17}$$

Again, a likelihood ratio test is performed:

$$LR_{IND} = 2\log\frac{\mathcal{L}_U}{\mathcal{L}_R} \overset{asympt.}{\sim} \chi^2_{(1)}. \tag{18}$$

Conditional Coverage Test

The conditional coverage test (CC test) combines the UC and the independence test by simply adding their likelihood ratios (Christoffersen, 1998) :

$$LR_{UC} + LR_{IND} \overset{asympt.}{\sim} \chi^2_{(2)}. \tag{19}$$

Dynamic Quantile Test

The dynamic quantile test by Engle and Manganelli (2004) is also a joint test of unconditional coverage and independence. There is an in-sample version and an out-of-sample version. We used the latter one:

$$DQ = \frac{\mathbf{h}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{h}}{\alpha\cdot(1-\alpha)} = \frac{\mathbf{h}'\cdot\hat{\mathbf{h}}}{\alpha\cdot(1-\alpha)}. \tag{20}$$

The hit vector $\mathbf{h}$ is defined as follows:

$$\mathbf{h} = \begin{cases} h_i = 1-\alpha & \text{if} \quad r_i < VaR_i \\ h_i = -\alpha & \text{else} \end{cases}. \tag{21}$$

The expected value of the hits equals 0. $\mathbf{X}$ is the matrix of independent variables (i.e. 1 for intercept, VaRs, and lagged hits $\mathbf{h}$).
Basically, the DQ test is a regression of the hit variable $\mathbf{h}$ on the independent variables VaR and lagged hits $\mathbf{h}$. The test statistic DQ is the product of the hit vector and the estimated hit vector, divided by $\alpha\cdot(1-\alpha)$. DQ is χ^2-distributed with degrees of freedom equalling the number of coefficients used in the DQ test. As we used the default setting, *df* is equal to 6 (intercept, VaRs, and lagged hits $\mathbf{h}$ with lag= $1, 2, 3, 4$. This way, more lags are considered than by the independence test, which tests for lag 1 only. Optionally, lagged VaRs and lagged returns can be included additionally and the intercept can be suppressed.

2 Empirical Analysis

2.1 Data and Methods

We took a sample of daily prices from 2000-01-01 to 2016-12-01 (trading days only) for the following assets which represent different German asset classes (Data from Bloomberg, Quantities vary due to different number of trading days):

- MDAX and CDAX represent the German capital market (4302 daily prices each).
- Euro/USD Dollar exchange rate represents the currency market (4414 daily prices).
- Gold represents the market for precious metals (4412 daily prices).
- Brent oil, wheat, and corn represent the commodity market (4335, 4263, and 4263 daily prices).

In the selection of assets, we are following Haugom et al (2016) to cover a wide area of assets showing different distributional properties. Additionally, the regulatory requirements for market risk measure using VaR to cover, if appropriate, interest rate risk, credit spread risk, equity price risk, foreign exchange risk, and commodity price risk are met as well. Equity risk is represented by MDAX and CDAX and foreign exchange risk by the EUR/USD exchange rate. Gold is a commodity, but serves as financial backup as well. Among the metals it has the highest trading volume and frequency. Pure commodity risk is represented by the latter three, namely Brent oil, wheat and corn. Brent oil from the Northern Sea is beside Western Texas Intermediate one of the two references for light sweet crude oil and therefore, for the other crude oils as well as their prices are linked to the reference prices. Wheat and corn are among the most important varieties of cereals, both are cultivated in Germany as well. All prices are daily closing prices, MDAX and CDAX prices in EUR, Gold spot price in USD, Brent oil, wheat, and corn future prices in USD.

Continuous daily returns ($\log p_t - \log p_{t-1}$) were calculated. Table 1 presents the main descriptive statistics for the daily returns. As expected, mean and median are around zero for all assets. Minimum returns are quite large for Brent oil and even more extreme for Corn. Brent oil, wheat, closely followed by corn show the strongest variation of data. The series are quite symmetric except of corn which is slightly left-skewed and shows a strong excess kurtosis. Additionally, gold, MDAX, and CDAX show a higher excess kurtosis. Figure 1 provides a time series graph of all returns examined in our study.

Table 1: Descriptive statistics for the the daily continuous returns of the assets under investigation, data from 2000-01-01 to 2016-12-01 (trading days only).

Asset	N	Mean	Median	Min	Max	SD	IQR	Skewness	Excess Kurtosis
MDAX	4301	0.0004	0.0010	-0.0906	0.1130	0.0131	0.0127	-0.2839	5.2356
CDAX	4301	0.0001	0.0008	-0.0755	0.1064	0.0144	0.0143	-0.0860	3.9660
EUR_USD	4413	0.0000	0.0001	-0.0252	0.0345	0.0063	0.0072	0.0484	1.4427
Gold	4411	0.0003	0.0004	-0.0951	0.1025	0.0113	0.0114	-0.2224	6.0104
Brent oil	4334	0.0002	0.0005	-0.1444	0.1271	0.0223	0.0234	-0.1175	2.8504
Wheat	4262	0.0001	0.0000	-0.0997	0.0879	0.0201	0.0240	0.1311	1.8174
Corn	4262	0.0001	0.0000	-0.2686	0.1276	0.0187	0.0200	-0.6169	12.1974

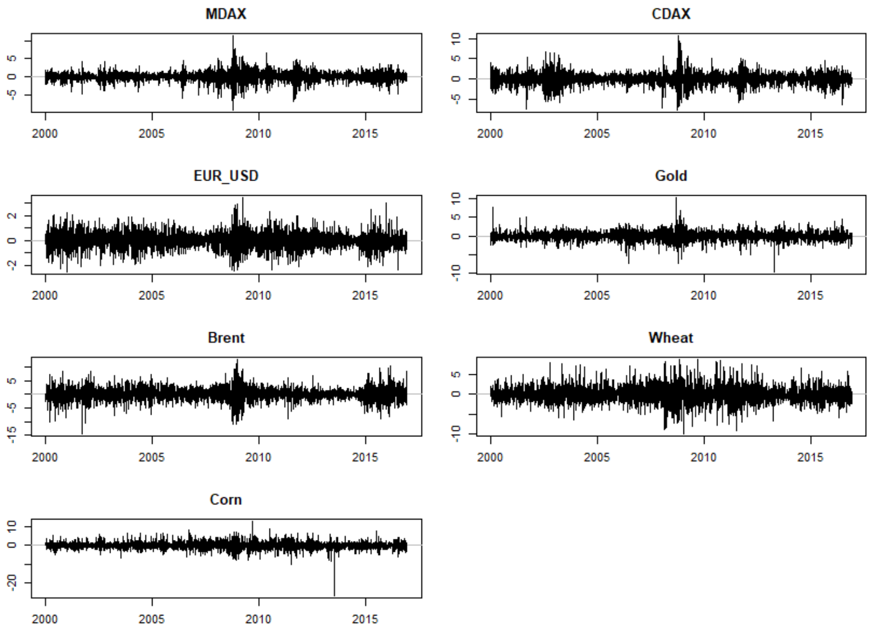

Figure 1: Time series graph of daily continuous returns (in %) of the assets under investigation, data from 2000-01-01 to 2016-12-01 (trading days only). Axes: x = trading days, y = continuous returns in %..

The distributional properties vary over time as Figure 2 shows. While the skewness of most of the assets varies between −1 and +0.5, corn has shown an extreme behaviour since mid of 2013. Since then, the corn distribution has been strongly left-skewed caused by the high negative return on 2013-07-15 (see Figure1).[2] This was due to very good weather conditions in mid 2013 in the US corn belt (continuous rain during the optimum growth period) which let the future price drop by nearly 24 % on one single day (Hirtzer, 2013).

[2] 1000-day rolling excess kurtosis of corn shows the same behaviour, not exceeding 3 until mid of July 2003, it increases until the end of the investigation period to more than 55.

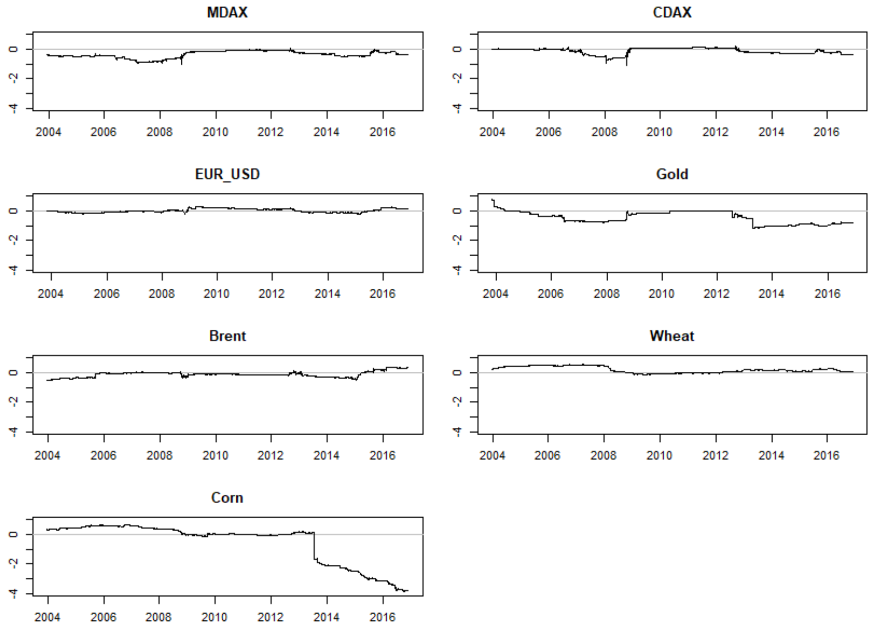

Figure 2: Rolling skewness of the daily continuous returns of the assets under investigation. A rolling window of 1000 days is applied for the calculation of skewness. Sample data are from 2000-01-01 to 2016-12-01 (trading days only), therefore, the rolling skewness plots start by end of 2003. Axes: x = trading days, y = skewness.

For each asset we used 1000 observations to calculate the one day ahead forecast of the VaR in a rolling window approach. For HAR-QREQ only 980 observations were used as 20 observations are required for estimation of monthly volatility. For back testing we compared the computed VaRs with the observed returns during the same time frame. To consider different model characteristics we calculated the VaRs for three different significance levels $\alpha = 0.01, 0.025, 0.05$. All back tests were evaluated at a 5 % significance level. We used R version 3.4.1 (R Core Team, 2017) to carry out the statistical analysis employing fGarch version 3010.82.1 (Wuertz et al, 2016), quantreg version 5.33 (Koenker, 2017), quantileVaR version 1.0 (Veka, 2013), and Rcpp version 0.12.12 (Eddelbuettel and François, 2011).

2.2 Results

In respect to the various assets, the models perform differently. As an example, Figure 3 compares RiskMetrics and HAR-QREG on a two year sample of MDAX. Compared to the RiskMetrics approach HAR-QREG follows the observed returns much more closely, resulting in overall better performance.

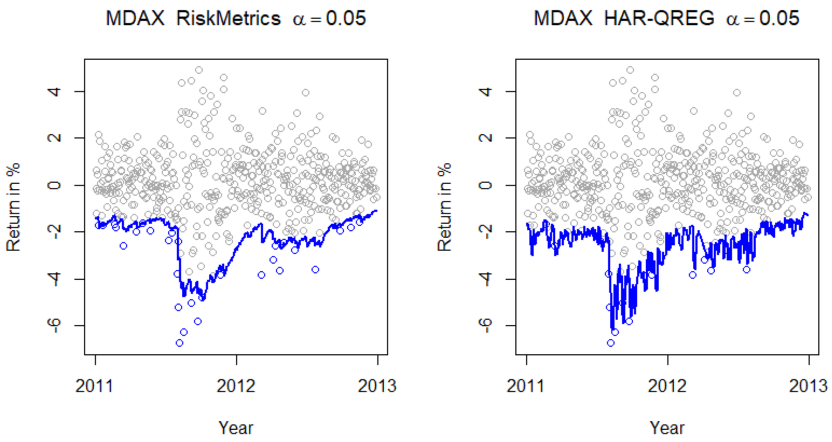

Figure 3: Comparison of RiskMetrics and HAR-QREG applied to a 2-year sample of MDAX returns at 5 % significance level. HAR-QREG (right) follows the observed returns more closely.

Table 2 shows the results of the four different back tests applied to the various VaR estimation methods and assets. By the quantity of tests successfully passed, one can directly recognize that CAViaR-Adapt and Historical Simulation failed in most of the applications. RiskMetrics performed somewhat better, but only if applied to commodities and with higher significance levels. GARCH worked reasonably well, especially the EUR/USD exchange rate could be simulated perfectly under all conditions. The remaining three CAViaR approaches and HAR-QREQ showed the best performance. Nonetheless, no clear differentian can be made between the latter four from the detailed results.

Table 2: Detailed back testing results of all eight VaR estimation methods each applied to seven different assets. Four different back tests were performed, the tests sucessfully passed are marked with X.

	VaR Model	1.00 %	U	I	C	D	2.50 %	U	I	C	D	5.00 %	U	I	C	D [b]
C-SAV	MDAX	1.03 %	X	X	X		2.30 %	X	X	X	X	4.79 %	X			X
	CDAX	0.94 %	X	X	X	X	2.54 %	X	X	X	X	4.94 %	X	X	X	
	EUR/USD	1.38 %		X	X		2.90 %	X	X	X		5.13 %	X	X	X	X
	Gold	0.97 %	X	X	X	X	2.43 %	X	X	X	X	5.16 %	X	X	X	X
	Brent	0.66 %		X	X	X	2.19 %	X	X	X	X	4.80 %	X	X	X	X
	Wheat	1.13 %	X	X	X		2.51 %	X	X	X	X	5.30 %	X	X	X	X
	Corn	0.98 %	X		X	X	2.67 %	X		X	X	4.90 %	X			
C-AS	MDAX	1.12 %	X	X	X	X	2.61 %	X	X	X	X	5.06 %	X	X	X	X
	CDAX	1.12 %	X	X	X	X	2.97 %	X	X	X		5.03 %	X	X	X	X
	EUR/USD	1.52 %		X			2.78 %	X	X	X		5.04 %	X	X	X	X
	Gold	0.79 %	X	X	X	X	2.64 %	X	X	X	X	5.10 %	X	X	X	X
	Brent	0.72 %	X	X	X	X	1.89 %		X	X	X	4.50 %	X	X	X	X
	Wheat	1.38 %		X	X		2.79 %	X	X	X	X	5.82 %		X	X	X
	Corn	0.86 %	X	X	X	X	2.27 %	X		X	X	5.00 %	X			
C-IG	MDAX	0.91 %	X	X	X	X	2.21 %	X	X	X	X	4.73 %	X		X	X
	CDAX	1.00 %	X	X	X	X	2.70 %	X	X	X	X	5.03 %	X	X	X	
	EUR/USD	1.29 %	X	X	X	X	2.70 %	X	X	X	X	5.13 %	X	X	X	X
	Gold	0.94 %	X	X	X	X	2.55 %	X	X	X	X	5.19 %	X	X	X	X
	Brent	0.69 %	X	X	X	X	2.16 %	X	X	X	X	4.68 %	X	X	X	X
	Wheat	1.07 %	X	X	X	X	2.91 %	X	X	X		5.52 %	X	X	X	X
	Corn	0.89 %	X		X		2.58 %	X	X	X	X	4.81 %	X			
C-Adapt	MDAX	1.73 %					3.54 %					6.21 %				
	CDAX	1.67 %					3.79 %					6.81 %				
	EUR/USD	1.67 %		X			3.16 %					5.60 %	X			
	Gold	1.17 %	X	X	X		2.46 %	X	X	X		4.87 %	X		X	
	Brent	1.80 %		X			4.02 %		X			6.57 %				
	Wheat	1.44 %		X			3.10 %		X			6.01 %				
	Corn	1.29 %	X				3.16 %					5.79 %				
HS	MDAX	1.27 %	X	X	X		3.03 %	X				5.76 %	X			
	CDAX	0.97 %	X		X		2.51 %	X				5.00 %	X			
	EUR/USD	1.17 %	X	X	X		2.75 %	X				5.16 %	X			
	Gold	1.08 %	X	X	X		2.67 %	X	X	X		5.36 %	X	X	X	
	Brent	1.47 %					3.03 %	X				5.49 %	X			
	Wheat	1.35 %	X				3.00 %	X	X	X		5.49 %	X			
	Corn	1.62 %		X			3.31 %					5.64 %	X			
RiskMetrics	MDAX	2.21 %		X			3.48 %					5.66 %	X			
	CDAX	2.15 %		X			3.97 %		X			5.84 %		X		
	EUR/USD	1.49 %		X			3.22 %		X			5.86 %		X		
	Gold	2.20 %		X			3.61 %		X			5.89 %		X	X	
	Brent	1.77 %		X			3.27 %		X			5.46 %	X	X	X	X
	Wheat	1.13 %	X	X	X		2.48 %	X	X	X		4.96 %	X	X	X	X
	Corn	1.50 %					2.64 %	X	X	X	X	5.03 %	X	X	X	X

Table 2: Detailed back testing results of all eight VaR estimation methods each applied to seven different assets. Four different back tests were performed, the tests sucessfully passed are marked with X.

	VaR Model	1.00 %	U	I	C	D	2.50 %	U	I	C	D	5.00 %	U	I	C	D [b]
GARCH	MDAX	1.88 %		X			3.33 %		X			5.66 %	X	X		
	CDAX	1.85 %		X			3.51 %		X			5.63 %	X	X	X	
	EUR/USD	1.29 %	X	X	X	X	2.87 %	X	X	X	X	5.07 %	X	X	X	X
	Gold	1.96 %		X			3.22 %		X		X	5.07 %	X	X	X	X
	Brent	1.29 %	X	X	X		2.43 %	X	X	X	X	4.71 %	X	X	X	X
	Wheat	0.92 %	X	X	X		2.33 %	X	X	X		4.47 %	X	X	X	X
	Corn	1.44 %		X			2.42 %	X	X	X	X	4.05 %				
HAR-QREG	MDAX	1.18 %	X	X	X		2.67 %	X	X	X	X	4.70 %	X			
	CDAX	1.18 %	X	X	X	X	2.76 %	X	X	X	X	5.24 %	X	X	X	
	EUR/USD	1.11 %	X	X	X	X	2.75 %	X				5.13 %	X	X	X	X
	Gold	1.17 %	X	X	X	X	2.58 %	X	X	X	X	5.34 %	X	X	X	X
	Brent	1.05 %	X	X	X		2.58 %	X	X	X	X	5.31 %	X	X	X	X
	Wheat	1.13 %	X	X	X	X	2.60 %	X	X	X		5.30 %	X	X	X	X
	Corn	1.32 %	X	X	X		2.79 %	X	X	X	X	5.12 %	X			

[b] In header: 1.0, 2.5, 5.0 % – VaR level, U – Unconditional coverage test, I – Independence test, C – Conditional coverage test, D – DQ test,
In table: percentage of hits in results, X – test successfully passed

To evaluate the overall performance of a VaR model, we counted the number of back tests that successfully passed. For each VaR level (1 %, 2.5 %, 5 %) four back tests were performed (UC, Independence, CC, DQ) resulting in 12 tests per asset. Seven assets were tested, therefore, it sums up to a total of 84 test per VaR estimation model. Table 3 shows the overall performance of the different models.

Table 3: Number of back tests successfully passed for the different VaR estimation methods. Maximum is 84, equaling 7 assets times 12 tests each. Three CAViaR and the HAR-QREG methods show clearly the best results.

CAViaR-SAV	CAViaR-AS	CAViaR-IG	CAViaR-Adapt	HS	RM*	GARCH	HAR-QREG
70	71	76	15	32	36	54	70

Note: *RM = RiskMetrics

CAViaR-IG achieved the best overall results (76 of 84 test successfully passed) closely followed by CAViaR-AS (71), CAViaR-SAV (70), and HAR-QREG

(70). In the middle field GARCH (54) was at the upper end, RiskMetrics (36) and Historical Simulation (32) at the lower end. CAViaR-Adapt showed the worst performance with only 15 of 84 tests successfully passed. Regarding the different asset classes the various VaR estimation models performed differently:

- MDAX and CDAX were best represented by CAViaR-AS (12 resp. 11) and CAViaR-IG (both 11).
- EUR/USD exchange rate was best represented by the GARCH(1,1) and CAViaR-IG models (both 12).
- Gold was modeled well by four models: CAViaR-AS, CAViaR-IG, CAViaR-SAV, and HAR-QREG (all 12).
- Brent Oil could be modelled by CAViaR-AS (11), CAViaR-IG (12), CAViaR-SAV (11), GARCH (11), and HAR-QREG (11).
- Wheat could be represented well by CAViaR-IG, CAViaR-SAV, and HAR-QREG (all 11), followed by GARCH and RiskMetrics (both 10).
- Corn was difficult to model: None of the models convinced in total, the best results were given by CAViaR-AS (8), HAR-QREG (8), and RiskMetrics (8), followed by CAViaR-IG (7) and CAViaR-SAV (7).

In terms of computational time (see Table 4) Historical Simulation and RiskMetrics were the fastest models, but closely followed by HAR-QREG which gives a far better average modeling performance. All other methods (GARCH and the CAViaR variants) needed much more computation time (up to 80 times compared to HAR-QREG).

Table 4: Computational time of the various VaR estimation models relative to the fastest one (Historical Simulation).

CAViaR-SAV	CAViaR-AS	CAViaR-IG	CAViaR-Adapt	HS	RM*	GARCH	HAR-QREG
363.3	2753.0	663.2	765.7	1.0	1.2	205.2	8.3

Note: *RM = RiskMetrics

3 Discussion

All CAViaR models (except Adaptive) successfully achieved precise modelling over a wide range of assets – at the expense of high computational time. HAR-QREG, which is comparatively easy and fast in computation, seems to be a good alternative. Both approaches do not assume normal distribution of the returns, consider stylized facts of financial time series, and use quantile regression techniques to model VaR. Overall, Historical Simulation, CAViaR Adaptive, and RiskMetrics did not satisfactorily model the VaR of the selected German asset classes in the years 2000–2016. As a result, and from an applied practitioner point of view, HAR-QREG seems to be a good choice for the selected assets during the tested time period.

Nevertheless, all methods failed to model corn returns. As discussed in data section, there was one extreme price drop in 2013 which heavily changed the distributional properties of the corn time series. Hence, all methods were sensitive to extreme values. In Table 5 back testing results for corn are presented when excluding years 2013–2016. Three methods perfectly passed all 12 tests successfully, another two at least 10 of 12.

Table 5: Detailed back testing results for corn, dates from 2001-01-01 to **2012-12-31**, trading days only. Last column shows total of test successfully passed.

VaR Model	1.00 %	U	I	C	D	2.50 %	U	I	C	D	5.00 %	U	I	C	D	Total
C-SAV	1.19 %	X	X	X	X	2.81 %	X	X	X	X	5.19 %	X	X	X	X	12
C-AS	0.88 %	X	X	X	X	2.42 %	X	X	X	X	5.32 %	X	X	X	X	12
C-IG	0.97 %	X	X	X	X	2.77 %	X	X	X	X	5.06 %	X		X		10
C-Adapt	1.28 %	X	X	X	X	3.17 %	X				5.63 %	X				6
HS	1.89 %		X			3.74 %					6.33 %					1
RiskMetrics	1.32 %	X	X	X	X	2.46 %	X	X	X	X	5.03 %	X	X	X	X	12
GARCH	1.36 %	X	X	Y	X	2.59 %	X	X	X	X	4.49 %	X				9
HAR-QREG	1.54 %		X		X	2.95 %	X	X	X	X	5.50 %	X	X	X	X	10

In future research it should be analysed whether it is possible to derive features of the time series which may enable modelling the choice of an appropriate VaR estimation method for one asset at hand. Also extreme value approaches could be an interesting addition to the methods investigated so far.

Acknowledgements We would like to thank Simone Manganelli and Steinar Veka for their support and the reviewers for their valuable remarks.

References

Allen DE, Singh AK, Powell R (2012) A Gourmet's delight: CAViaR and the Australian stock market. Applied Economics Letters 19(15):1493–1498, DOI 10.1080/13504851.2011.636017

Allen L, Boudoukh J, Saunders A (2004) Understanding Market, Credit, and Operational Risk: The Value at Risk Approach. Blackwell Publishing Ltd, Oxford, UK. ISBN: 978-1-405142-26-7

Angelidis T, Benos A, Degiannakis S (2004) The use of GARCH models in VaR estimation. Statistical Methodology 1(1-2):105–128, DOI 10.1016/j.stamet.2004.08.004

Bao Y, Lee T, Saltoglu B (2006) Evaluating predictive performance of value at risk models in emerging markets: a reality check. Journal of Forecasting 25(2):101–128, DOI 10.1002/for.977

Bilandi MM, Kudła J (2016) Comparing the Precision of Different Methods of Estimating VaR With a Focus on EVT. ACRN Oxford Journal of Finance and Risk Perspectives 5(1):109–147, URL http://www.acrn-journals.eu/resources/jofrp0501f.pdf

Chambers DR, Kelly MA, Lu Q (2014) Understanding the estimation risks of value at risk. Journal of Alternative Investments 16(3):64–85, Institutional Investor Journals Umbrella, DOI 10.3905/jai.2013.16.3.064

Christoffersen PF (1998) Evaluating Interval Forecasts. International Economic Review 39(4):841–862, DOI 10.2307/2527341

Danielsson J (2001) Financial Risk Forecasting. John Wiley & Sons Ltd, Chichester, West Sussex, UK, DOI 978-0470669433

Dimitrakopoulos DN, Kavussanos MG, Spyrou SI (2010) Value at risk models for volatile emerging markets equity portfolios. Quarterly Review of Economics and Finance 50(4):515–526, Board of Trustees of the University of Illinois, DOI 10.1016/j.qref.2010.06.006

Eddelbuettel D, François R (2011) Rcpp: Seamless R and C++ Integration. Journal of Statistical Software 40(8):1–18, DOI 10.18637/jss.v040.i08, URL http://www.jstatsoft.org/v40/i08/

Engle RF, Manganelli S (2004) CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. Journal of Business and Economic Statistics 22(4):367–381, DOI 10.1198/073500104000000370

Hannemann R, Schneider A, Weigl T (2013) Mindestanforderungen an das Risikomanagement (MaRisk) – Kommentar unter Berücksichtigung der Instituts-Vergütungsverordnung (Insituts-VergV). Schäffer-Poeschel, Stuttgart. ISBN: 978-3-791033-07-5

Hansen PR, Lunde A (2005) A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? Journal of Applied Econometrics 20(7):873–889, DOI 10.1002/jae.800

Hartz C, Mittnik S, Paolella M (2006) Accurate value-at-risk forecasting based on the normal-GARCH model. Computational Statistics and Data Analysis 51(4):2295–2312, DOI 10.1016/j.csda.2006.09.017

Haugom E, Ray R, Ullrich CJ, Veka S, Westgaard S (2016) A parsimonious quantile regression model to forecast day-ahead Value-at-Risk. Finance Research Letters 16(2):196–207, DOI 10.1016/j.frl.2015.12.006

Hirtzer M (2013) Rains preceded record drop in U.S. corn, soy cash prices, URL https://www.reuters.com/article/usa-grains-cash/rains-preceded-record-drop-

J.P.Morgan, Reuters (1996) RiskMetrics – Technical Document. Technical report, Morgan Guaranty Trust Company, Reuters Ltd

Koenker R (2017) quantreg: Quantile Regression. URL https://CRAN.R-project.org/package=quantreg, R package version 5.33

Kupiec PH (1995) Techniques for verifying the accuracy of risk measurement models. Journal of Derivatives 3(2):75–84, DOI 10.3905/jod.1995.407942

R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/

Veka S (2013) quantileVaR: Quantile regression for estimating value at risk and backtesting., R package version 1.0

Wuertz D, Chalabi Y, Miklovic M, Boudt C, Chausse P, et al (2016) fGarch: Rmetrics - Autoregressive Conditional Heteroskedastic Modelling. URL https://cran.r-project.org/web/packages/fGarch/fGarch.pdf, r package version 3010.82.1