Modelling temporal dependence of realized variances with vines

Claudia Czado[b], Eugen Ivanov[a], Yarema Okhrin[a,*]

[a] Department of Statistics, Faculty of Business and Economics, University of Augsburg, Germany
[b] Center for Mathematical Sciences, Technische Universität München, Germany

1. Introduction

The concept of using realized measures (*RM*s) to measure (co)volatility, has largely spread since the availability and accessibility of high-frequency data in finance improved. First models have arisen in the continuous time framework, leading to fundamental results and methods. However, practical issues, like microstructure noise, asynchronicity of data, stress on short and long term forecasting, etc. lead to a shift from stochastic to statistical modeling of *RM*s. Additionally to the very general assumptions of a multivariate diffusion process underlying the various estimation methods, the key advantage of *RM*s is that they are observable time series. Therefore, in most applied empirical work, the main task lies in modeling and forecasting.

By making volatility directly measurable, standard time series models can be applied to time series of *RM*s. In the univariate case, fractionally integrated ARMA (Granger and Joyeux, 1980) or Heterogeneous Autoregressive (HAR) processes (Corsi, 2009) are most commonly suggested to be used on logarithmic transformations, capturing long-memory dependence and allowing for a standard autoregressive structure. Sometimes models, see e.g. Corsi et al. (2008), are further extended, to include a GARCH component. The ARFIMA framework benefits from the availability of standard methods of forecasting, see, e.g. Beran (1994), whereas HAR models are a convenient and easy to estimate way of including volatility measured over different time horizons and account for multifractal scaling (Corsi, 2009). Based on the model, long-horizon forecasts can be obtained directly or iterated from conditional one-day volatility predictions, if the error term is Gaussian. The Mixed

* Corresponding author.
E-mail address: yarema.okhrin@wiwi.uni-augsburg.de (Y. Okhrin).

Data Sampling (MIDAS) suggested in Ghysels et al. (2004) offers an alternative way of constraining the parameters of a higher-order autoregressive process. This model also possesses long memory with only a few parameters to estimate.

The problem of the standard models is the imposed linearity, i.e. the current level is modelled as a linear function of its past values. In this paper we propose an alternative approach that helps to incorporate non-linear causal dependence into the predictive models for *RM*s. We consider a regression based on copulas, where the regression equation is derived from the joint multivariate distribution of dependent and explanatory variables. This joint distribution is modeled using vine copulas, which offer a flexible framework that precisely captures the inherited dependence in our data. To the best of our knowledge, Darsow et al. (1992) were the first to consider copulas for modeling time series. They derived a copula version of a univariate first-order Markov process with the help of conditional independence. Ibragimov (2009) generalized their approach to higher order univariate Markov processes as well as non-Markov processes. He also introduced new classes of copulas for modeling univariate time series. Copula-based stationary time series can still be estimated in a semiparametric framework, i.e. marginal distributions are fitted non-parametrically and copulas - parametrically with maximum likelihood estimation (MLE). The semiparametric approach for time series is described in Chen and Fan (2006b). Recently, Brechmann and Czado (2014), Beare and Seo (2015) and Smith (2015) independently developed vine-based models for stationary multivariate time series. To capture the cross-sectional dependence, Brechmann and Czado (2014) employ connected C-vines in the first tree, Smith (2015) considers a high-dimensional D-vine for all time points and Beare and Seo (2015) - connected D-vines in the first tree. Further, the R-vine structures proposed in Brechmann and Czado (2014); Beare and Seo (2015) assume the existence of a central variable, whose temporal dependence is explicitly modeled. In contrast, Smith (2015) suggested to model both the temporal and cross-sectional dependence within a single D-vine, such that no variable is emphasized. Also Smith and Vahey (2016) considered an application of a Gaussian vine for forecasting asymmetric densities of macroeconomic (low frequency) time series. Further multivariate approaches can be found in Fink et al. (2017); Brechmann et al. (2018); Simard and Remillard (2015). Sokolinskiy and van Dijk (2011) estimate bivariate copulas semiparametrically on univariate time series and their results advocate the use of copulas with upper-tail dependence.

In this paper we suggest two vine copula based methods for modeling and forecasting time series of a given *RM*. The first approach is based on C-vine regression, where regressors are linked directly to the current value. The second approach takes a traditional time series perspective by linking regressors sequentially within a D-vine structure. Both methods can be extended to the multivariate case, for example similar to Brechmann and Czado (2014); Beare and Seo (2015) or Smith (2015). Vine regressions are applied separately to information sets as in HAR and MIDAS, which serve as linear benchmarks. Additionally, bivariate copulas of consecutive observations are taken as non-linear benchmarks. Within an extensive empirical study all models are estimated on 13 time series of log bi-power variations. The full sample estimation results are used to assess the flexibility of the model and the detected non-linearity in the dependence. We pay particular attention to the extensions this approach allows compared to benchmark models. Furthermore, within a moving windows approach we address both the in-sample and out-of-sample performance, whereas the latter is evaluated on one-step-ahead forecasts using innovative performance measures. The results show that the vine-based approach significantly improves performance over linear benchmarks and one specific vine is a clear favorite. Further, we show that the inclusion of information larger than the last day improves the performance.

The paper is structured as follows. The next section provides basic information on *RM*s and the most popular modeling techniques. Section 3 contains a general introduction to pair-copula constructions and provides details on the vine-based regression. The estimation and forecasting is the subject of Section 4. The results of an extensive empirical study are summarized in Section 5. Section 6 concludes.

2. Modelling realized measures

Availability of high-frequency financial data stimulated a quickly expanding field of financial econometrics which deals with modeling the latent risk process of intraday price movements through realized measures (*RM*s). All of these measures are designed to estimate integrated variance (*IV*) of intraday price movements. In the ideal case, when the price process is not contaminated with microstructure noise, *IV* can be consistently estimated by (*RV*), which was the first and is by far the most well-known *RM*:

$$RV_t = \sum_{m=1}^{M} r_{t,m}^2,$$

where $r_{t,m}$ is the intraday return observed at the subinterval m on day t. When M goes to infinity, *RV* converges uniformly in probability to (*QV*) of the price process, which equals *IV* under no noise assumption. However, there is a strong empirical evidence for the presence of jumps and noise in observed intraday prices. Since *RV* treats jumps and noise as a part of price process it diverges or delivers biased estimates of *IV*. In order to estimate *IV* consistently given the intrinsic properties of the data, different types of realized measures capable of excluding market frictions have been developed. A simple and comprehensible measure, which isolates the variation of the pure price process is the (*BPV*). It is calculated as the sum of products of absolute values of two consecutive returns

$$BPV_t = \sum_{m=2}^{M} |r_{t,m}||r_{t,m-1}|.$$

Barndorff-Nielsen and Shephard (2004) showed that *BPV* converges uniformly in probability to the *IV*, assuming that the probability of a jump in two small consecutive intervals is negligible and that noise is not autocorrelated. For a very thorough overview of modeling high-frequency data we refer to Aït-Sahalia and Jacod (2014) and for a good overview of different realized measures to Liu et al. (2015). As discussed in the latter reference, *BPV* is one of the most efficient *RM*s and the best one available to us. Thus, we consider only the logarithm of *BPV*, although all models presented later can be applied to any other *RM*.

One of the most distinctive properties of *RM*s is the high persistency. From the economic perspective there are several reasons for this finding. It can be justified by the mixture of distributions hypothesis of Tauchen and Pitts (1983). Bollerslev and Jubinski (1999) argue that if trading volume and returns are driven by the same information flow process, then both time series should share the same long range dependence. The autocorrelation in this case depends on the fractional differencing parameter of the underlying information flow. An alternative explanation is based on the aggregation of weakly dependent linear processes. The changes in the volatility are driven by numerous financial phenomena and thus, following Zaffaroni (2004), induce long memory in the volatility. As a consequence, methods capable of capturing strong memory are preferred for modeling and forecasting purposes.

Researchers attempted to go beyond the standard AR(FI)MA processes and developed several models, which can be interpreted as constrained versions of high-order AR processes. We now briefly introduce the two most popular time series models, namely the HAR model of Corsi (2009) and the MIDAS model of Ghysels et al. (2004), which were developed for *RV* but are applicable to other *RM*s. Normally, the logarithmic transformation is applied, in order to ensure positive forecasts and to reduce the noise from outliers. Both models are constrained AR(20) processes and thus remain linear in the past realized volatilities.

HAR

The Heterogeneous Auto-Regressive (HAR) model was first proposed in Corsi (2009) and emerged as one of the most popular approaches. The model was inspired by the Heterogeneous Market Hypothesis, which tries to explain the dynamics of volatility using the activity of heterogeneous market participants. The specific feature the author exploited to construct HAR was the time horizon or dealing frequency. The basic idea is that market makers which have very high dealing frequency and pension funds with low dealing frequency differently affect the volatility. The model is defined as a linear regression between information of different time horizons

$$RM_t = \phi + \phi^{(d)} \cdot RM_{t-1} + \phi^{(w)} \cdot \frac{1}{4}\sum_{k=2}^{5} RM_{t-k} + \phi^{(m)} \cdot \frac{1}{15}\sum_{k=6}^{20} RM_{t-k} + \varepsilon_t,$$

where ϕ is the intercept, $\phi^{(d)}$ is the regression coefficient for the previous day observation, $\phi^{(w)}$ and $\phi^{(m)}$ - coefficients for the proxies of weekly and monthly information, respectively. Thus, HAR simplifies the full *AR*(20) by classifying past information into three groups. The model is estimated by OLS. Further, a multivariate extension was considered in Chiriac and Voev (2011).

MIDAS

MIDAS is a linear, reduced-form regression which was proposed in Ghysels et al. (2004) for the case, if regressors have higher frequency compared to the dependent variable. This model possesses long memory with only a handful of parameters to estimate. For our purposes we state the model as follows

$$RM_t = \beta_0 + \sum_{k=1}^{20} b(k,\theta) RM_{t-k} + \varepsilon_t,$$

where $b(k, \theta)$ are functional regression coefficients calculated at lag k with parameter vector θ. Further we choose the normalized exponential Almon lag MIDAS function for $b(k, \theta)$

$$b(k,\theta) = \delta \frac{\exp\left(\lambda_1(k+1) + \lambda_2(k+1)^2\right)}{\sum_{s=0}^{K} \exp\left(\lambda_1(s+1) + \lambda_2(k+1)^2\right)},$$

where $\theta = \{\delta, \lambda_1, \lambda_2\}$. The advantage of this model over HAR are the smoothly decreasing coefficients.

Note that both models are actually high-order autoregressive processes with specific restrictions on parameters. This simplification allows for dimension reduction, computational effectiveness and an economic interpretation (at least for the HAR model). However, both models are linear and impose linear and constant impact of past values on the current one. In the light of some stylized facts about volatility, for example clustering, this is a very strong and potentially misleading assumption.

Fig. 1 provides empirical evidence for this problem based on *BPV* of the S&P500 index. It presents the dynamics of 99% confidence intervals for the coefficients (subplot title) of three different HAR models, estimated for low, middle and high level of volatility, respectively. Thus, for each trading day between 2005 and 2016, an empirical cdf is estimated on 1000 most recent observations. Next, this data is split into "lower" and "upper" part, containing observations belonging to the lower and upper quartiles, respectively, and "middle" part, which contains the remaining half of the data. Finally, within

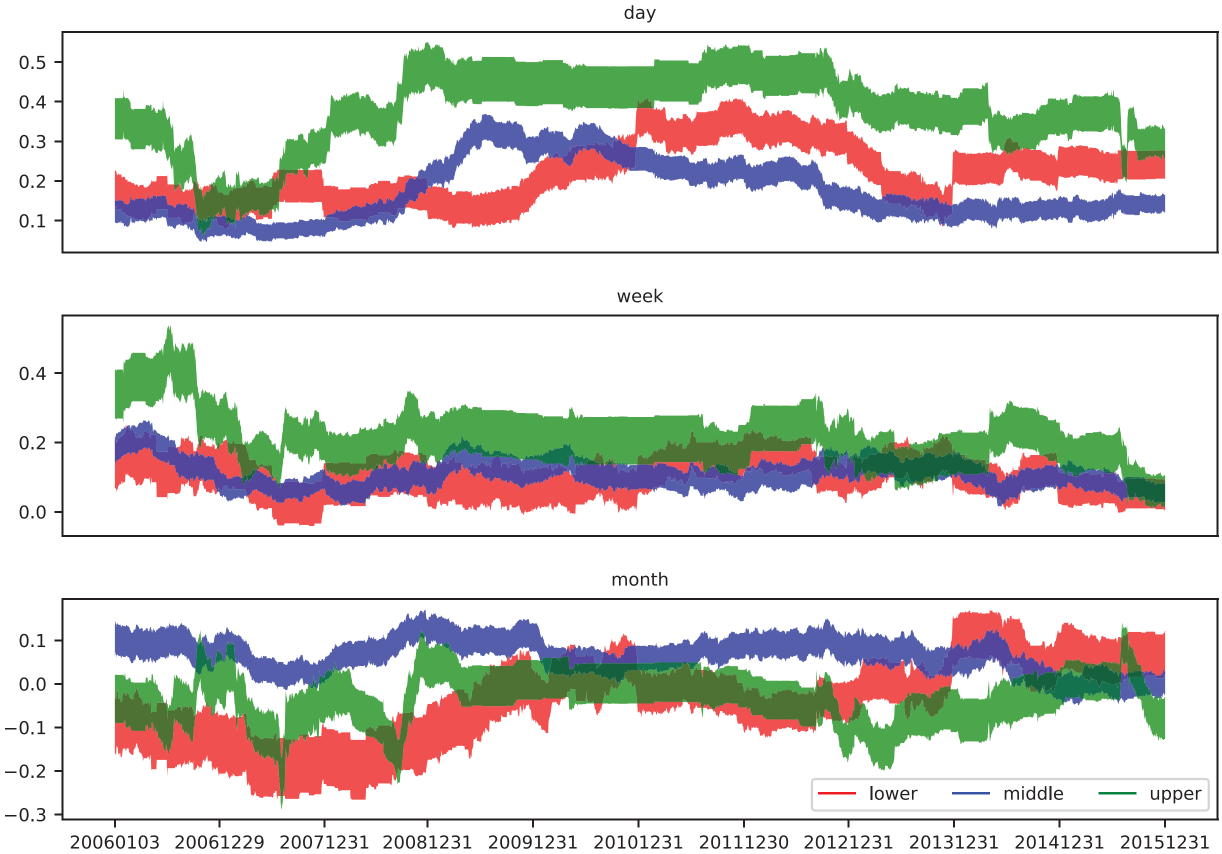


Fig. 1. 99% confidence intervals for coefficients (subplot title) of three HAR models, estimated for upper (green), lower (red) and middle (blue) levels of *BPV* of the S&P500.

each group HAR is fitted on the allocated observations as dependent variable and corresponding daily, weekly and monthly information as regressors. The differences between models' parameters are considered statistically significant if the corresponding confidence intervals don't overlap. All three plots in Fig. 1 indicate such differences consistently over time among the three models. This suggests heterogeneous levels of dependency and illustrates a form of volatility clustering, since the regression for the upper quartile has the largest coefficients for day and week effects. The lower plot, with monthly information as regressor, indicates a comparably less extreme difference in coefficients. Furthermore, there is a clear variation over time, indicating that the impact is not static. The steep increase in 2008 is due to the outburst of the most recent financial crisis in the U.S., the abrupt decrease in 2012 can be caused by the third round of quantitative easing.

To account for these intrinsic properties of *RMs* we suggest two models capable of capturing very heterogenous forms of non-linearity in temporal dependence. The suggested models use vine copulas for the temporal dependence and can be seen as an extension to the cross-sectional modeling of multivariate distributions with copulas. The alternative setups elaborated in this paper are based on C-vine and D-vine regressions. To keep the same informational contents in the suggested models compared to the linear benchmarks we consider the following information sets

$$RM_t, RM_{t-1}, \frac{1}{4}\sum_{k=2}^{5} RM_{t-k}, \frac{1}{15}\sum_{k=6}^{20} RM_{t-k}, \tag{1}$$

$$RM_t, RM_{t-1}, RM_{t-2}, \dots, RM_{t-20}, \tag{2}$$

where the specification (1) resembles the informational contents of the HAR model and (2) reflects the MIDAS or more generally the $AR(20)$ process.

3. Vine copulas for temporal dependence of realized volatilities

In this section we present a copula-based approach for modeling temporal dependence of *RM*s. Let $(x_1, \dots, x_d)$ be the realizations of the random vector $(X_1, \dots, X_d)$, with joint cdf $F(x_1, \dots, x_d)$, joint density $f(x_1, \dots, x_d)$, marginal cdfs

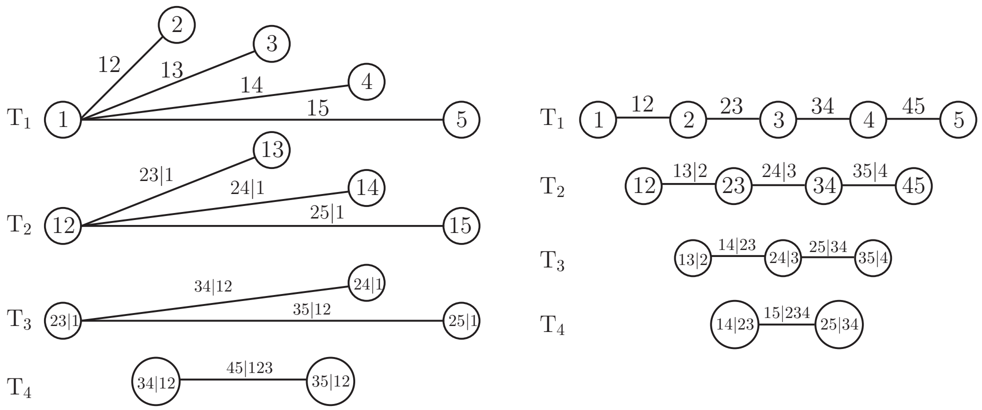


Fig. 2. C-vine (left)and D-vine (right) representation for $d = 5$.

$F_1(x_1), \ldots, F_d(x_d)$ and marginal densities $f_1(x_1), \ldots, f_d(x_d)$. In the specific cases of HAR and MIDAS, the random vectors correspond to those in (1) or (2), respectively. We now briefly introduce regular vines (R-vines).

Any joint continuous distribution can be factored as a product of the corresponding copula and marginal distributions using Sklar's theorem. This approach is particularly flexible for constructing new distributions, since margins and dependence structure, which is dictated by copula, can be selected independently. Very attractive as it may seem, this method has a serious practical limitation. Though, there is a multitude of well-studied bivariate copulas, most of their high-dimensional extensions are controlled by several parameters. Pair copula constructions (PCCs) were designed to solve the problem of scarcity of flexible high dimensional copulas. The idea is to construct a multivariate distribution by defining (conditional) copulas for pairs of variables. It is possible due to the decomposition of multivariate distributions which is unique up to a relabeling of the variables:

$$f(x_1, \ldots, x_d) = \prod_{i=2}^{d} f(x_i \mid x_1, \ldots, x_{i-1}) \cdot f_1(x_1).$$

This idea of constructing multivariate dependence from bivariate blocks goes back to Joe (1996). His motivation was to construct a flexible class of multivariate distributions with $d(d-1)/2$ parameters for d variables and properties that the multivariate normal distribution does not have. The particular construction described in Joe (1996) will be later called drawable vine (D-vine). The $d-$dimensional likelihood can be expressed without integrals as a product of d univariate marginal densities and of $d(d-1)/2$ conditional and unconditional bivariate copulas (see Theorem 4.2 in Kurowicka and Cooke, 2006).

In order to categorize different decompositions and to select right pairs a graphical "interface", called regular vine (R-vine), was introduced by Bedford and Cooke (2002), who also derived the density of a PCC. For a recent review of financial applications using R-vines we refer to Aas (2016) and a more detailed introduction to R-vines is contained in Stöber and Czado (2017).

Typical examples of these structures are canonical vines (C-vines) and D-vines, which are illustrated in Fig. 2 (see Czado, 2010; Aas et al., 2009). Every tree of a C-vine is defined by a root node. The degree of a node is defined as the number of nodes which the current node is connected to. The root node of every tree T_i has therefore the degree $d-i$, $i \in \{1, \ldots, d-1\}$. A D-vine is also solely defined through its first tree, where each node has a degree of at most 2.

Conditional copulas in an R-vine depend on conditioned values not only through their arguments, thus allowing statistical applications only for a subclass of elliptical distributions and d-dimensional Clayton copula (see Stöber et al., 2013). Aas et al. (2009) were the first to consider R-vines with arbitrary copulas by making what is now commonly named *simplifying assumption*, i.e. that copulas depend on the values they are conditioned on only through their arguments. They showed that this always results in valid multivariate distributions and copulas. An R-vine copula is specified by assigning a (conditional) pair copula (with parameters) to each edge of R-vine structure (see Fig. 2).

For purposes of statistical inference a matrix representation of R-vines was proposed in Morales-Napoles (2008) and further developed in Dißmann et al. (2013), who also provided methods to compute log likelihood of an R-vine. In order to computationally specify a $d-$dimensional R-vine a total of 4 lower (or upper) triangular $d \times d$ matrices are necessary: matrix with R-vine structure, copula families, first and second parameters of the bivariate copulas. The R-vine structure can be selected manually (our case) or estimated treewise by maximizing the sum of absolute values of some dependence measure (usually Kendall's τ) for each edge of a tree. Next, a copula family is selected for each edge by estimating parameters of all possible families and choosing the one with the smallest AIC (Akaike, 1973). This procedure is often executed iteratively starting from the first tree (see Czado et al., 2012) in order to provide starting values for a full MLE.

Table 1
Tail dependence coefficients for copula families considered in the empirical application.

	G	S	F	C	J	Gu	BB1	BB6	BB7
λ_L	0	$2t_{\nu+1}\left(-\frac{\sqrt{\nu+1}\sqrt{(1-\theta)}}{\sqrt{1+\theta}}\right)$	0	$2^{-\frac{1}{\theta}}$	0	0	$2^{-1/(\theta\delta)}$	0	$2^{-1/\delta}$
λ_U	0	$2t_{\nu+1}\left(-\frac{\sqrt{\nu+1}\sqrt{(1-\theta)}}{\sqrt{1+\theta}}\right)$	0	0	0	$2-2^{1/\theta}$	$2-2^{1/\delta}$	$2-2^{1/(\theta\delta)}$	$2-2^{1/\theta}$

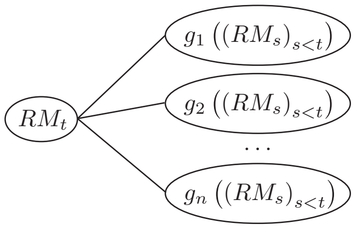


Fig. 3. The first tree of C-vine regression for a univariate time series.

Tail dependence is particularly important from the perspective of copulas. The conditional probability of extreme events for copula C of (X_1, X_2) is measured by

$$\lambda_L = \lim_{v\to 0^+} P(X_1 < F_1^{-1}(v) \mid X_2 < F_2^{-1}(v)) = \lim_{v\to 0^+} \frac{C(v,v)}{v},$$
$$\lambda_U = \lim_{v\to 1^-} P(X_1 > F_1^{-1}(v) \mid X_2 > F_2^{-1}(v)) = \lim_{v\to 1^-} \frac{1-2v+C(v,v)}{1-v},$$

where λ_L and λ_U are the lower and upper tail dependence coefficients, respectively (Nelsen, 2013). If these limits exist and are nonzero, then variables have either lower or upper tail dependence. The coefficients of tail dependence for the copula families considered during empirical application are given in Table 1, whereas θ, δ and ν stand for copula parameters of the appropriate family. Gumbel (Gu) and Clayton (C) copulas show asymmetric dependence in tails, whereas the probability function of extreme events in the case of Student's t (S) is equal in both tails. Gauss (G), Frank (F) and Joe (J) copula show independence in tails. BB1, BB6 and BB7 introduce additional flexibility into modeling tail dependence. For a precise definition of these bivariate copula families see Joe (2014).

3.1. Copula-based regression for realized measures

The basic construction of popular models for *RMs* is a linear regression between current state and some (aggregated) historical information. The simplicity of the linear approach implies a constant response rate to changes in regressors (past information), i.e. the response is independent of the level of regressors. A natural extension is a model which can react non-linearly to changes in past values or their dependence structure. For this purpose we now introduce C-vine and D-vine regressions for univariate time series.

C-vine regression

As a general structure for C-vine regression we propose to set the central node of the first tree to the current state RM_t and other nodes to some compressions (information filters) of $\{RM_s\}_{s<t}$, which are treated as classic regressors. The first tree of the considered C-vine is illustrated in Fig. 3, where $g_k(\{RM_s\}_{s<t})$, $k = 1, \dots, n$ are some potentially overlapping but distinct information filters.

However, in this particular R-vine structure, second and all trees above contain edges that condition past on relatively more recent information. This is the case, for example for the second tree which contains pairs $(g_l(\{RM_s\}_{s<t}), g_m(\{RM_s\}_{s<t})|RM_t)$, $l \neq m$. Copulas are traditionally applied to capture cross-sectional, non-sequential dependencies, thus such a conflict with the sequential logic of time series modeling is inevitable but does not lead to problems with statistical modeling. Conditioning on the most recent values does not hinder forecasting, since only RM_t is unknown at $t-1$ and is predicted by utilizing the associated multivariate conditional density. Additionally, due to descending estimation accuracy for further trees, such central positioning of most recent observations remains advantageous from an inference perspective.

We further constrain the discussion to information filters as in HAR and MIDAS with the corresponding order. In this case, the central nodes of further trees are the most recent information available at that level. As an example, consider a C-vine illustrated in Fig. 4 with the information set (1), which we further refer to as HAR-C-vine. It is a 4-dimensional C-vine, where the central node in the first tree T_1 is RM_t (most recent observation) and other nodes contain previous day, weekly and monthly information. As mentioned, nodes of further trees condition past information (RM_s, $s<t$) on the value we

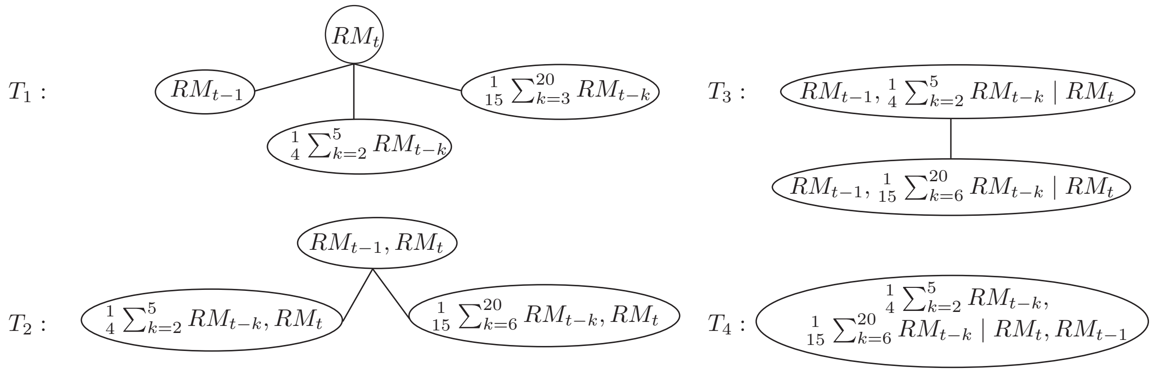


Fig. 4. Vine structure of HAR-C-vine.

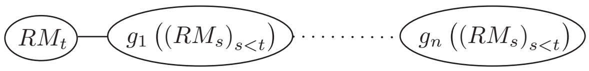


Fig. 5. The first tree of D-vine regression for a univariate time series.

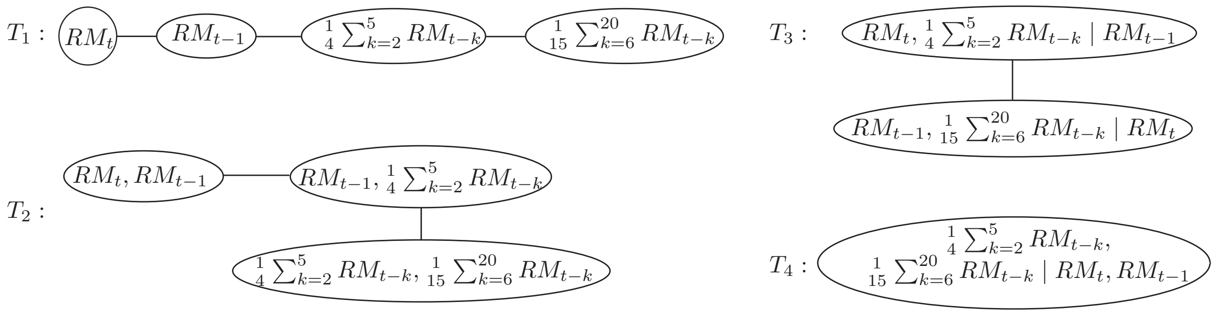


Fig. 6. Vine structure of a HAR-D-vine.

want to forecast, RM_t. The approach appears to be economically infeasible at the first look, but remaining trees guarantee a precise modeling of the overall dependence and are thus justified from the statistical perspective. We use the notation MIDAS canonical vine (MIDAS-C-vine) for a C-vine regression with information set as in MIDAS.

D-vine regression

A D-vine is a special case of an R-vine when each node is connected to at most two other nodes. The first tree of the proposed D-vine regression is illustrated in Fig. 5, where $g_k(\{RM_s\}_{s<t})$, $k = 1, \ldots, n$ are again some probably overlapping information filters.

As an example we consider D-vine with variables and their order as in HAR. The structure of this vine model is presented in Fig. 6. Conditioning on future information occurs in the third and fourth trees. As already discussed for C-vine regression, this fact follows from this particular statistical model. We introduce a HAR drawable vine (HAR-D-vine) and MIDAS drawable vine (MIDAS-D-vine) similar to C-vine regression.

The two proposed R-vine regressions are substantially different. The first tree of C-vine models dependencies between the response variable and all regressors directly, while further trees contain conditional copulas measuring the dependence between pairs of regressors given response variable. This is in contrast with D-vine where the dependencies between pairs of regressors are captured in the first tree and conditional dependencies between regressors and response variable in the following trees. Another crucial difference is that the conditional density $f(x_1|x_2, \ldots, x_d)$ - which is used for forecasting - is given by an analytic expression for D-vine, while it must be numerically approximated for C-vine. It is important to stress that if all bivariate copulas are Gaussian, both C-vine and D-vine have parameters equal to partial correlations. This fact resembles the interpretation of the last coefficient in an autoregressive process. The latter equals the partial correlation between the current and the lagged value of the process with the impact of intermediate observations being removed.

Bivariate copula model

Additionally to the linear benchmarks above, we consider a bivariate copula as a non-linear benchmark for the C-vine and D-vine regressions. We assume that all relevant information about RM_t is contained in RM_{t-1} and consequently only the joint distribution of this pair is of relevance for our purpose. We model this distribution by a bivariate copula. Thus, the essential feature of this non-linear benchmark is that we disregard a lot of past information, but model the remaining information in a non-linear way. Note that this method is nested in both vine regressions. An application of bivariate copulas to financial data can be found in Sokolinskiy and van Dijk (2011).

4. Estimation and forecasting

To estimate any copula we have to select the marginal distributions F_i, $i = 1, \dots, d$ first. Since the observations are not independent and identically distributed (*iid*) the selection of an appropriate distribution should be handled with care. As pointed out in Noh et al. (2013) a fully parametric estimator is probably biased, if either copula or margins are missspecified. Thus, we follow Kraus and Czado (2017) and choose a semiparametric approach, which was proven to be asymptotically normal also for not *iid* data (see Chen and Fan, 2006a; Chen and Fan, 2006b). Here we combine the parametric families of copulas with non-parametric estimates of margins, given by the kernel smoothing estimator

$$\hat{F}_i(x) = \frac{1}{T}\sum_{t=1}^{T} K\left(\frac{x - x_{it}}{h}\right), \tag{3}$$

where K is a proper kernel function and bandwith h is selected automatically. The nonparametric estimate in (3) is used to determine empirical copula data $\hat{u}_{it} = \hat{F}_i(x_{it})$.

Fitting a C-vine or D-vine regression requires a choice of a bivariate copula family for each node and subsequent estimation of the corresponding parameters. To estimate vine regression on $(\hat{u}_{1t}, \dots, \hat{u}_{dt})_{t=1,\dots,T}$ we use treewise estimation which sequentially selects copula families for every node of each tree. With such method it is not guaranteed that the best possible model fit is achieved, since every tree is examined separately. However, as mentioned in Dißmann et al. (2013) such a sequential approach is justified by the fact that the first tree often has the greatest influence on the model fit. In order to select copulas we first apply the independence test for Kendall's τ discussed in Genest and Favre (2007). If independence cannot be rejected, then the independence copula is chosen for the given pair. Else, the optimal copula family for each node is selected using AIC (Akaike, 1973), which compensates for number of parameters less strictly than BIC. The superiority of AIC as a model selection tool compared to goodness-of-fit tests was advocated by Manner (2007); Brechmann et al. (2012). The selection procedure calculates AIC for each candidate copula family and chooses the one with the smallest AIC. Here the parameters are estimated using MLE. The set of possible families consists of copulas in Table 1 along with their rotations.

Conditional copulas of an R-vine, i.e. copulas for the second tree and further, can be estimated using a specific recursion. For this purpose let C_I be the copula and c_I the copula density associated with random vector $\boldsymbol{X}_I = (X_{i_1}, \dots, X_{i_n})$ for some index set $I = (i_1, \dots, i_n)$. Further, let $C_{I;\,D}$ denote the copula associated with the conditional distribution of $\boldsymbol{X}_I$ given $\boldsymbol{X}_D = \boldsymbol{x}_D$ for some index set D. In the case of $I = (i, j)$, the copula $C_{I;D} = C_{ij;D}$ is the copula associated with the bivariate distribution of (X_i, X_j) given $\boldsymbol{X}_D = \boldsymbol{x}_D$. For $j \in D$ and $D_{-j} = D \backslash \{j\}$ we define

$$F_{i|D}(x_i|\boldsymbol{x}_D) = h_{i|j,D_{-j}}(F_{i|D_{-j}}(x_i|\boldsymbol{x}_{D_{-j}})|F_{j|D_{-j}}(x_j|\boldsymbol{x}_{D_{-j}})), \tag{4}$$

where the h-function is the derivative of the (conditional) copula

$$h_{i|j,D_{-j}}(u \mid v) = \frac{\partial C_{ij;D_{-j}}(u, v)}{\partial v}. \tag{5}$$

Here copula $C_{ij;D_{-j}}$ is evaluated at the pair

$$(F_{i|D_{-j}}(x_i|\boldsymbol{x}_{D_{-j}}), F_{j|D_{-j}}(x_j|\boldsymbol{x}_{D_{-j}})).$$

Next, further two indices, say $k, m \in D_{-j}$, are taken to determine $F_{i|D_{-j}}(x_i|\boldsymbol{x}_{D_{-j}})$ and $F_{j|D_{-j}}(x_j|\boldsymbol{x}_{D_{-j}})$ similar to (4). At some point this way of proceeding reaches the first tree. The required index we take out of the conditioning set is determined by the R-vine structure.

For example, selected copulas of the first tree of C-vine in Fig. 2 are used to calculate pseudo observations, i.e. transformed variables

$$h_{i|1}(u_i \mid u_1) = \frac{\partial \hat{C}_{i,1}(u_i, u_1)}{\partial u_1},$$

for $i = 2, 3, 4, 5$. Next, copulas $C_{23;\,1}$, $C_{24;\,1}$ and $C_{25;\,1}$ are selected based on the following pairs of pseudo observations:

$$(h_{2|1}(u_2 \mid u_1),\ h_{3|1}(u_3 \mid u_1)), (h_{2|1}(u_2 \mid u_1),\ h_{4|1}(u_4 \mid u_1)) \text{ and } (h_{2|1}(u_2 \mid u_1),\ h_{5|1}(u_5 \mid u_1)).$$

In the next iteration step copula $C_{34;\,12}$ of the third tree is selected based on

$$(h_{3|1,2}(h_{3|1}(u_3 \mid u_1) \mid h_{2|1}(u_2 \mid u_1)),\ \ h_{4|1,2}(h_{4|1}(u_4 \mid u_1) \mid h_{2|1}(u_2 \mid u_1))),$$

and $C_{35;\,12}$ on

$$(h_{3|1,2}(h_{3|1}(u_3 \mid u_1) \mid h_{2|1}(u_2 \mid u_1)),\ \ h_{5|1,2}(h_{5|1}(u_5 \mid u_1) \mid h_{2|1}(u_2 \mid u_1))).$$

In this way all conditional copulas up to the last tree are selected recursively.

Lastly, we use a full maximum likelihood (ML) procedure to estimate vine parameters, where starting values are obtained from treewise estimation. Given observations
$\{x_{i,t}\}_{i=1,\dots,d;t=1,\dots,T}$ of the random vector $(X_1, \dots, X_d)$ the log-likelihood for C-vine is calculated as

$$\sum_{j=1}^{d-1}\sum_{i=1}^{n-j}\sum_{t=1}^{T}\log\left[c_{j(j+i);1,\dots,j-1}\left(F_{j|1,\dots,j-1}\left(x_{j,t}|x_{1,t},\dots,x_{j-1,t}\right),\ F_{j+i|1,\dots,j-1}(x_{j+i,t}|x_{1,t},\dots,x_{j-1,t})\right)\right],$$

and for D-vine as

$$\sum_{j=1}^{d-1}\sum_{i=1}^{n-j}\sum_{t=1}^{T}\log\left[c_{i(i+j);i+1,\dots,i+j-1}\left(F_{i|i+1,\dots,i+j-1}\left(x_{i,t}|x_{i+1,t},\dots,x_{i+j-1,t}\right),\ F_{i+j|1,\dots,j-1}(x_{i+j,t}|x_{i+1,t},\dots,x_{i+j-1,t})\right)\right].$$

Since R-vines are mostly fitted to *iid* data, developed family selection procedures rely on this fact. To simplify the estimation we neglect the residual autocorrelation in the data and fit families, as if the data were *iid*. Due to the high persistence of *BPV*, we include all corresponding past information in HAR-C-vine, HAR-D-vine, MIDAS-C-vine and MIDAS-D-vine, as well as preserve the order of the variables.

Forecasting is the key objective in modeling of *RM*s. Due to the asymmetric nature of the data and susceptibility to outliers, one-step-ahead forecasts are computed as the median of the underlying conditional distribution. Additionally, 95% quantiles are calculated as upper bounds of forecast intervals. Taking quantiles as forecasts is natural not only within the copula framework but also for linear time series models in financial econometrics. As mentioned earlier, the distribution of the response variable conditioned on regressors is directly accessible in case of D-vine but must be numerically approximated in case of C-vine. Now we provide computational details for these two forecasting methods based on C-vine and D-vine regression respectively.

Forecasting with C-vine

We only discuss the forecasting procedure for HAR-C-vine, since it can be easily extended to MIDAS-C-vine. For this purpose, short-hand notation x_1, ..., x_4 is used for the four variables as in (1) and $u_i = F_i(x_i)$ - for copula data. The joint density of the first tree of HAR-C-vine can be decomposed as follows:

$$\begin{aligned} f(x_1,\dots,x_4) = {} & f_1(x_1)\cdot f_2(x_2)\cdot f_3(x_3)\cdot f_4(x_4) && \text{margins}\\ & \cdot c_{12}(u_1,u_2)\cdot c_{13}(u_1,u_3)\cdot c_{14}(u_1,u_4) && \text{1st tree}\\ & \cdot c_{23;1}(F_{2|1}(u_2|u_1),F_{3|1}(u_3|u_1))\cdot c_{24;1}(F_{2|1}(u_2|u_1),F_{4|1}(u_4|u_1)) && \text{2nd tree}\\ & \cdot c_{34;12}(F_{3|12}(u_3|u_1,u_2),F_{4|12}(u_4|u_1,u_2)) && \text{3rd tree}\end{aligned}$$

For a probabilistic model, such as the one based on vines, we first calculate the conditional distribution of forecasts for RM_t, i.e.

$$RM_t \mid RM_{t-1},\ \frac{1}{4}\sum_{k=2}^{5}RM_{t-k},\ \frac{1}{15}\sum_{k=6}^{20}RM_{t-k}. \tag{6}$$

This distribution can be expressed analytically as follows:

$$f_{1|234}(x_1|x_2,x_3,x_4) = \frac{f_{1234}(x_1,x_2,x_3,x_4)}{f_{234}(x_2,x_3,x_4)} = \frac{c_{1234}(u_1,u_2,u_3,u_4)}{c_{234}(u_2,u_3,u_4)}f_1(x_1). \tag{7}$$

Next we calculate the conditional density (7) on equidistant grid with step $\Delta = 10^{-4}$ on (0, 1), i.e. for each $u = \frac{j}{10000}$, $j = 1,\dots,9999$ calculate for specified values u_2, u_3, u_4

$$\frac{c_{1234}(u,u_2,u_3,u_4)}{c_{234}(u_2,u_3,u_4)}\cdot f_1\left(F_1^{-1}(u)\right). \tag{8}$$

The conditional density can now be used to calculate the conditional median and 95% quantile for RM_t, using a corresponding numerical approximation based on (8).

Forecasting with D-vine

In case of D-vine regression we employ the procedure described in Kraus and Czado (2017). Using the previously established notation, the conditional forecast for X_1 given $X_2,\dots,X_d$ is taken from the conditional quantile function for α:

$$F^{-1}_{X_1|X_2,\dots,X_d}(\alpha \mid x_2,\dots,x_d) = F^{-1}_{X_1}(C^{-1}_{1|2,\dots,d}(\alpha \mid u_2,\dots,u_d)), \tag{9}$$

where again $u_i = F_i(x_i)$. Using recursion (4) the conditional quantile function $C^{-1}_{1|2,\dots,d}$ can be expressed in terms of nested inverse h-functions. For example, in case of a D-vine on X_1, X_2, X_3, X_4

$$C_{1|2,3,4}(v|u_2,u_3,u_4) = h_{1|2,3,4}\{C_{1|2,3}(v|u_2,u_3)|C_{4|2,3}(u_4|u_2,u_3)\}$$
$$= h_{1|2,3,4}\Big\{h_{1|2,3}\big[C_{1|2}(v|u_2)|C_{3|2}(u_3|u_2)\big]\Big|h_{4|2,3}\big[C_{4|3}(u_4|u_3)|C_{2|3}(u_2|u_3)\big]\Big\}$$
$$= h_{1|2,3,4}\Big\{h_{1|2,3}\big[h_{1|2}(v|u_2)|h_{3|2}(u_3|u_2))\big]\Big|h_{4|2,3}\big[h_{4|3}(u_4|u_3)|h_{2|3}(u_2|u_3)\big]\Big\}.$$

Inversion of this conditional distribution function yields

$$C^{-1}_{1|2,3,4}(\alpha|u_2,u_3,u_4) = h^{-1}_{1|2}\Big(h^{-1}_{1|2,3}\Big[h^{-1}_{1|2,3,4}\Big\{\alpha|h_{4|2,3}\big[h_{4|3}(u_4|u_3)|h_{2|3}(u_2|u_3)\big]\Big\}\Big|h_{3|2}(u_3|u_2)\Big]\Big|u_2\Big).$$

The inverse function (9) gives explicit forecasts for any quantile of RM_t, whereas we calculate median and 95% quantile.

Forecasting with bivariate copula

Calculating forecasts from bivariate copula is done analogously to (9) using

$$F^{-1}_{X_1|X_2}(\alpha|X_2) = F^{-1}_{X_1}(C^{-1}_{1|2}(\alpha|u_2)).$$

Again median and 95% quantile are extracted.

Measures of comparison

We assess the goodness of forecasts from several perspectives. First, point forecasts are compared using a specific suitable loss function. Next, the quality of forecast distributions is quantified and, finally, the forecast intervals - in our case 95% quantiles - are examined.

As discussed in Patton (2011), the ranking of forecasts $\hat{y}_t$ based on imperfect volatility proxies $\hat{\sigma}_t^2$ - in our case *BPV* - depends on the choice of loss function $L(\hat{\sigma}_t^2, \hat{y}_t)$. Thus, this feasible ranking is generally different from the infeasible one, which is based on the true values of volatility process σ_t^2, i.e. $L(\sigma_t^2, \hat{y}_t)$. Further, Patton (2011) defines a class of loss functions robust to noise in the volatility proxy. The ranking of forecasts using this class of functions is the same under the true volatility σ_t^2 and the unbiased proxy $\hat{\sigma}_t^2$. The most popular loss functions nested in this class are mean squared error (MSE) and QLIKE:

$$\text{MSE}: \left(\hat{\sigma}_t^2 - \hat{y}_t\right)^2,$$
$$\text{QLIKE}: \log \hat{y}_t + \frac{\hat{\sigma}_t^2}{\hat{y}_t},$$

whereas the latter is less sensitive to extreme observations and levels of volatility. Further discussions regarding QLIKE can be found in Bollerslev et al. (1994).

As proposed in Patton (2011), we compare the predictive ability of models with Diebold-Mariano-West (DMW) test (Diebold and Mariano, 1995; West, 1996) using QLIKE, since the moment conditions required for DMW test under QLIKE are considerably weaker than for the MSE (Patton 2006). Further, Patton and Sheppard (2009) found that the power of DMW test using QLIKE is higher than for MSE. Given two forecasts $\hat{y}_{1t}$ and $\hat{y}_{2t}$, forecast error is defined as

$$e_{it} = \hat{y}_{it} - y_t, \quad i = 1, 2,$$

where y_t is the observed value. Under some loss function L the null hypotheses is of equal predictive accuracy against two alternatives

$$H_0: \mathbb{E}[L(e_{1t}) - L(e_{2t})] = 0,$$
$$H_1: \mathbb{E}[L(e_{1t}) - L(e_{2t})] > 0,$$
$$H_2: \mathbb{E}[L(e_{1t}) - L(e_{2t})] < 0.$$

In case of H_1 the second forecast is considered better, in case of H_2 - the first one.

Due to the probabilistic nature of the forecasts we further compare the fit of estimated conditional distributions to the observed values. As argued in Gneiting et al. (2007, 2004) the goal of probabilistic forecasting is to maximize the concentration of the forecasting distribution subject to consistency with the distribution of the observed values. A popular instrument for such comparison are scoring rules. This numerical score functions provide goodness-of-fit measure for evaluating and comparing probabilistic forecasts. Let $\mathcal{F}$ be a convex class of probability measures on a sample space Ω. Consider an observed random variable Y with (unknown) distribution $G \in \mathcal{F}$ and realization y. The purpose of modeling procedure for Y is calculating forecasts, given as random variable X with a (known) distribution $F \in \mathcal{F}$. A scoring rule $S(F, y)$ assigns a numerical score, based on the distribution of forecasts F and observed value y. The value $S(F, y)$ can be interpreted as distance between realization y and forecasting distribution F. Further, we denote the expectation of the score $s(F, Y)$ as

$$S(F, G) = \mathbb{E}_G[s(F, Y)] = \int S(F, y)dG(y).$$

A scoring rule is called strictly proper if for all distributions $F, G \in \mathcal{F}$ the following property is fulfilled:

$$S(G, G) \leq S(F, G),$$

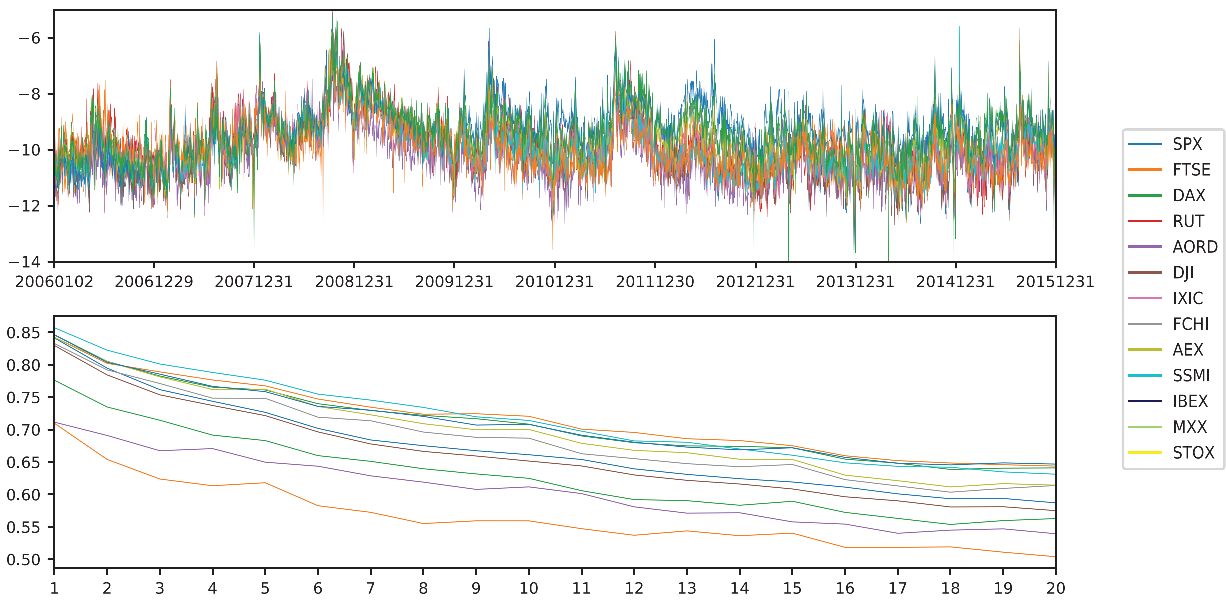


Fig. 7. Time series of the log *BPV* of all equity indices considered (upper plot) and autocorrelation function for up to 20 lags (lower plot) estimated on full sample.

with equality if and only if $F = G$. Thus, under a strictly proper scoring rule the best mean score is achieved in case the forecast distribution F equals the true distribution G. Consequently, strictly proper scoring rules can be used for MLE over mean score. One of the most popular strictly proper scoring rules is Continuous Ranked Probability Score (CRPS)

$$\mathrm{CRPS}(F, y) = \int_{-\infty}^{\infty} (F(x) - \mathbf{1}_{x \geq y})^2 dx. \tag{10}$$

CRPS measures the distance between forecast distribution F and a point mass on observed value y. This scoring rule is strictly proper with respect to the Borel probability measures with finite first moment (Gneiting and Raftery (2007)). It follows from (10), that for deterministic models CRPS is the absolute error $|x - y|$. Lower values of CRPS are considered to indicate better performance.

In order to compare interval forecasts, we use as in Brechmann and Czado (2014) the mean interval score (MIS) by Gneiting and Raftery (2007) for $\alpha = 0.05$:

$$\mathrm{MIS}_{\alpha}(\boldsymbol{l}, \boldsymbol{u}; \boldsymbol{x}) = \frac{1}{T}\sum_{t=1}^{T}\left[(u_t - l_t) + \frac{2}{\alpha}(l_t - x_t)\mathbf{1}_{x_t < l_t} + \frac{2}{\alpha}(x_t - u_t)\mathbf{1}_{u_t < x_t}\right],$$

where T is the number of out-of-sample predictions, $\boldsymbol{x} = (x_1, \ldots, x_T)'$ is the vector of true values, $\boldsymbol{l} = (l_1, \ldots, l_T)'$ and $\boldsymbol{u} = (u_1, \ldots, u_T)'$ are the vectors of lower and upper bounds of $100(1 - \alpha)\%$ confidence intervals. The quantities $\frac{2}{\alpha}(l_t - x_t)\mathbf{1}_{x_t < l_t}$ and $\frac{2}{\alpha}(x_t - u_t)\mathbf{1}_{u_t < x_t}$ are penalization terms for not covering the true value. Lower values of MIS_{α} are preferred.

5. Application

The purpose of the empirical application is to study the differences among the two common linear frameworks, bivariate copulas and the proposed non-linear methods for modeling *RMs*. The performance of HAR, HAR-C-vine, HAR-D-vine, MIDAS, MIDAS-C-vine, MIDAS-D-vine and bivariate copulas is compared on the time series of log *BPV* of 13 main world equity indices: S&P500 (SPX), FTSE, DAX (DAX), Russell 2000 (RUT), ASX All Ordinaries (AORD), DJIA, NASDAQ 100 (IXIC), CAC 40 (FCHI), AEX, SMI, IBEX 35, MIPC (MXX), EuroStoxx50 (STOX). The data - as provided by the Oxford-Man database - contains time series which start mainly on the first trading day of 2000 and end on the last trading day of 2015. We define the full sample as the period between the beginning of 2006 and the end of 2015 and constrain our analysis to this specific time interval, since it provides a suitable number of historical observations and shock events.

Fig. 7 serves the purpose of basic visual analysis of the considered time series over the full sample. The upper plot in Fig. 7 illustrates all equity indices and the lower one - the associated autocorrelation functions. The upper plot reveals changing dependency between time series, which increases strongly during the recent financial crisis. That is confirmed by the narrowing distances between time series during this period and the elevated levels of volatility. The autocorrelation function (ACF) in the lower plot matches findings reported in other studies, namely the strong persistence of variance. This results in ACF of at least 50% even at lag 20. On the average, the most volatile market index according to *BPV* is the German

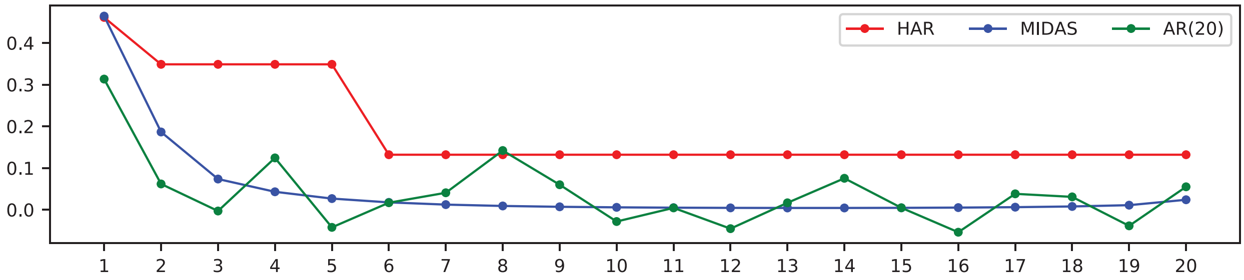


Fig. 8. Coefficients of HAR, MIDAS and AR(20) averaged over all indices, estimated on the full sample.

DAX and least volatile - Australian AORD. The three US indices, Dow Jones, S&P500 and Russell 2000 show quite similar moments, despite having considerably different amount of constituents. The technology heavy Nasdaq 100, on the other hand, demonstrated much higher levels of volatility.

To draw robust conclusions from the study, we estimate and analyze all models for each index on the full sample and within moving windows approach. Thereby, all models are re-estimated daily on the most recent 1000 trading days, whereas the first window of data ends on the last trading day of 2005 and the last window - on the second to last trading day in 2015. Thus, within moving windows approach we run about 2500 daily re-estimations for each model and each index. The in-sample fit of the models is compared both on the full sample and moving windows using criteria presented later. We measure the out-of-sample performance of models within moving windows approach only and use for this purpose one-step-ahead forecasts and forecasting distributions as described in Section 4.

In-sample comparison on full sample

For each time series all models along with an unconstrained AR(20) are estimated on the full sample. As discussed earlier, HAR is a constrained AR(20) regression, with coefficients approximated by a step function, which is constant for lags 2 to 5 and 6 to 20. MIDAS, on the other hand, approximates coefficients with an exponentially decaying function with two parameters. The averages of the estimated coefficients for each of the three linear models over all time series are presented in Fig. 8. The parameters of the full AR(20) are comparatively volatile, which could indicate that HAR and MIDAS might be too restrictive.

Coefficients of AR(20), HAR and MIDAS are not directly comparable to the model specifications of vines or bivariate copulas. To achieve comparability across all models, we calculate marginal effects, defined as sensitivities of the forecast for response variable with respect to each explanatory variable. Marginal effects are simply regression coefficients in case of linear models. For copula based models in our case, marginal effects are defined as the derivatives of median of the underlying conditional distribution with respect to all conditioned values.

Using notation as in (7), we restrict the discussion of the computational details of the procedure to HAR-based models. Given the conditional distribution at some point (u_2, u_3, u_4), forecast y is extracted as its median. In order to evaluate the marginal effect some regressor has on the forecast, we change one of the explanatory variables by 0.01, leaving other intact and recompute the forecast. For example, set $\hat{u}_2 = u_2 + 0.01$, leave u_3, u_4 unchanged and compute a new median forecast $\hat{y}$. The marginal effect of x_2 is then approximated as the change in the conditional forecast divided by the change in the explanatory variable, i.e.

$$\frac{y - \hat{y}}{F_2^{-1}(u_2 + .01) - F_2^{-1}(u_2)}. \tag{11}$$

By changing only one of the regressors we can assess the influence of this specific variable on the forecast. In this example, u_2 corresponds to the previous day level of *BPV*. The same procedure is repeated for the remaining regressors and is easily expanded to MIDAS vines.

Empirically, sensitivities (11) could be estimated for all possible points on a multivariate grid of $[0, 1]^d$, whereas d is 3 and 20 in case of HAR and MIDAS vines, respectively. Due to obvious computational infeasibility of this approach for a discretized version of interval [0,1] we constrain computations to 500 randomly drawn vectors.

The marginal effects for HAR-C-vine, HAR-D-vine and bivariate copulas are presented in Fig. 9. First, 500 random vectors of three dimensional copula data were drawn, then sensitivities for the three models were calculated separately for each market index on the same copula data. Panels (a) and (b) in Fig. 9 contain the sensitivities of HAR-C-vine and HAR-D-vine, respectively, with regard to daily (upper plot), weekly (middle) and monthly information. Panel (c) illustrates marginal effects for the bivariate copula model. Fig. 9 reveals that, as expected, the marginal effects are not constant - although to a different degree - for all market indexes. Due to the centristic design of C-vine, weekly and monthly information within HAR-C-vine have more pronounced marginal effects compared to HAR-D-vine. On the other hand, the latter has generally higher sensitivity to shocks in daily information compared to information further in the past. The sensitivities for the bivariate copula model are mostly between those of HAR-C-vine and HAR-D-vine. Notice, that the difference in marginal effects of

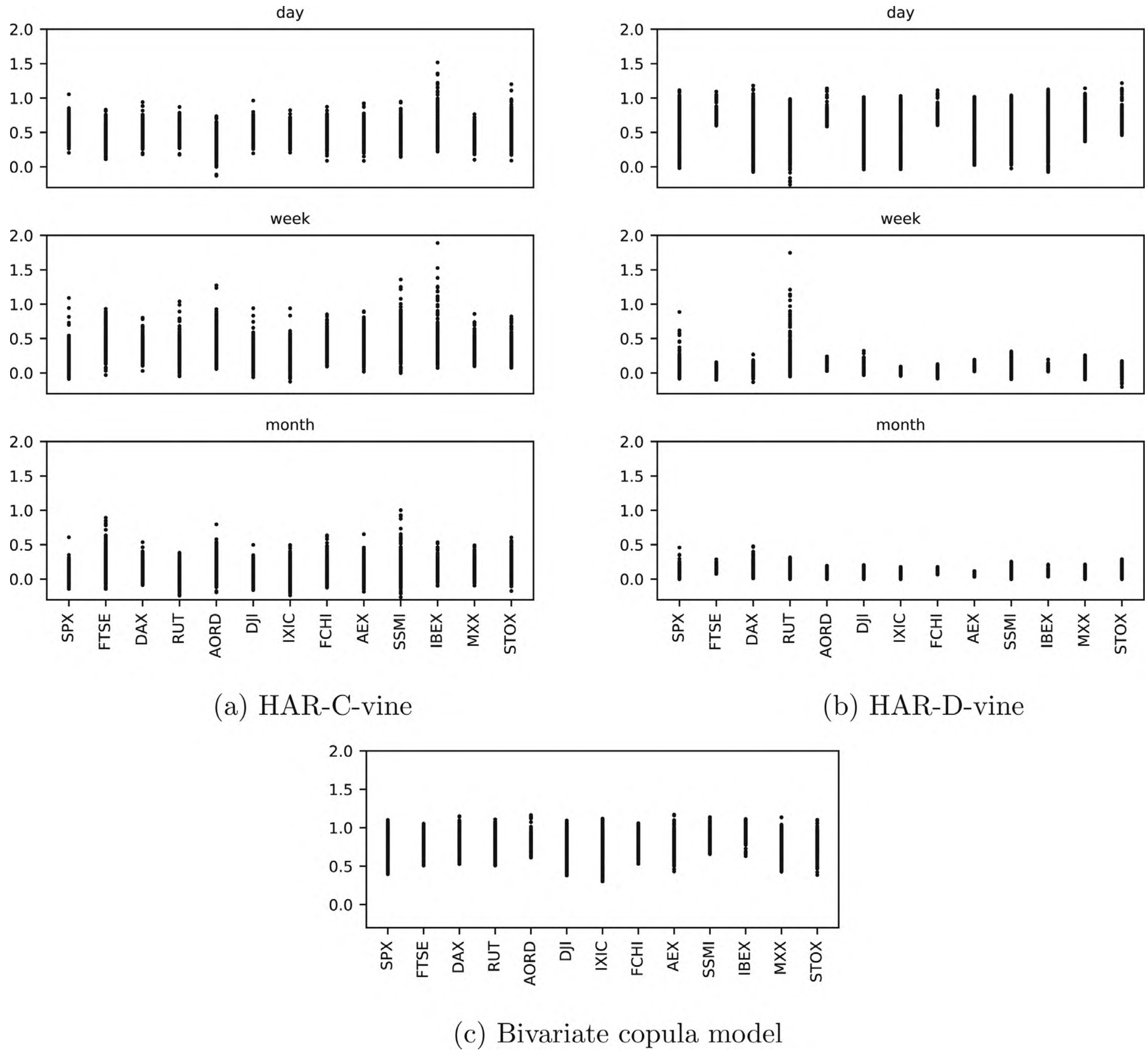


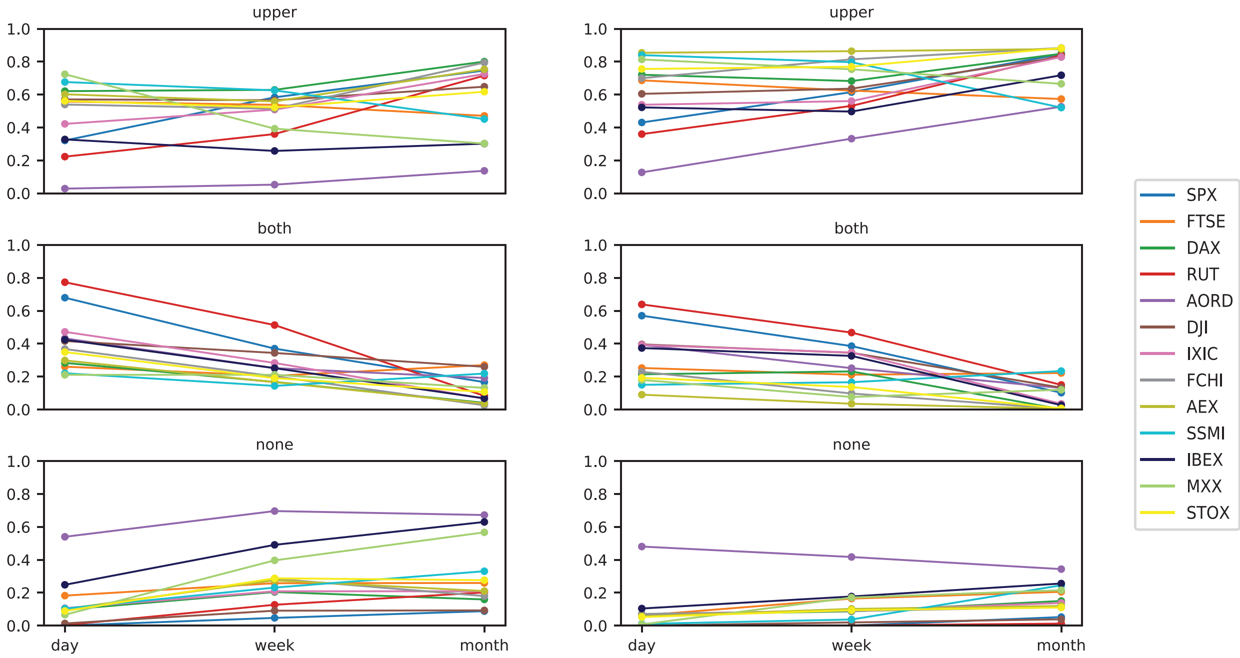
Fig. 9. (a), (b) The sensitivities of median forecast from HAR-C-vine and HAR-D-vine with respect to daily (top), weekly (middle) and monthly (bottom) variables and each market index. (c) The sensitivities of median forecast from bivariate copula model for each market index. Sensitivities for the three models are calculated on the same randomly drawn copula data.

vines and bivariate copula model is obvious, which might imply that taking merely the last observation for modeling is restrictive. The results for MIDAS are similar and are not presented here for space reasons. Again, most recent observations have the biggest impact on forecasting in case of MIDAS-D-vine, while the effect is spread across different lags for MIDAS-C-vine.

In-sample comparison on moving windows

We use the results gained within moving windows approach to measure the relevance of tail dependence and evaluate the dynamics of model specifications. To motivate the use of copulas we examine families selected by the estimation procedure for vines, whereas results for bivariate copula approach are very similar to copulas nested at an appropriate node of the first tree. We do not differentiate between individual copula families but instead classify them into three groups: with only upper, symmetric and no tail dependence. Copulas with upper tail dependence are regarded as capable of mimicking volatility clustering and are expected to dominate the family selection procedure.

The results of estimation procedure for copulas nested in the first tree of HAR-C-vine and HAR-D-vine are summarized in Fig. 10. Each dot stands for the percentage proportion of some copula group for each market index and variable, indicated



(a) HAR-C-vine (b) HAR-D-vine

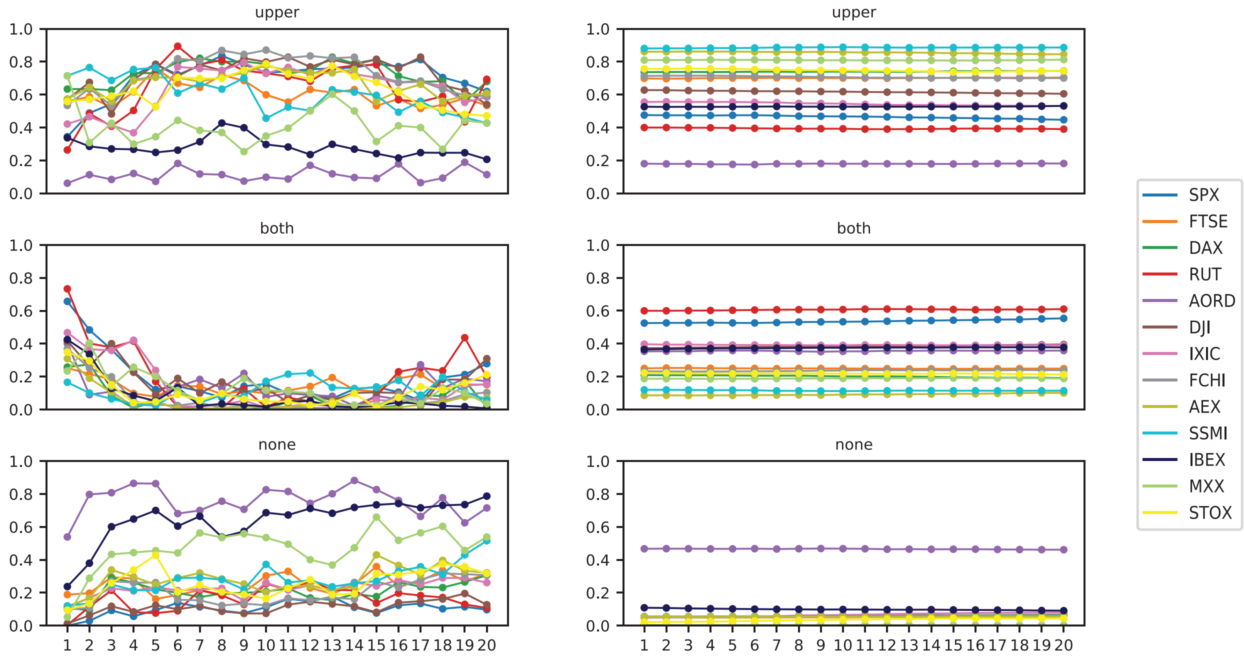
Fig. 10. The proportion of copula families with upper (top), symmetric (middle) and absent (bottom) tail dependence in the first tree of HAR-C-vine (a) and HAR-D-vine (b), estimated on the moving windows of length 1000. Each dot represents a percentage of copulas selected for the corresponding variable (x-axis), market index (color) and copula group (subplot title).

by the corresponding subplot title, line color and tick on the x-axis at the bottom, respectively. Both upper plots show the percentage of selected copulas with only upper tail dependence, middle - with symmetric tail dependence and both lower plots - without any. Notice, that none of the selected families had merely lower tail dependence. Copulas selection for SPX and RUT was highly skewed to t-copula, which has symmetrical dependence in tails, for all three information sets and both vine regressions. AORD, IBEX and MXX stand out with comparatively high proportion of families without tail dependence in case of HAR-C-vine. For the remaining market indexes copulas were mostly with upper tail dependence. Notice, that the percentage of families with only upper tail dependence generally increases with the increasing time distance. This can suggest volatility clustering, which is present for several weeks. The results for the bivariate approach are very similar to those for “day” variable.

The results of selection procedure for copulas nested in the first tree of MIDAS-C-vine and MIDAS-D-vine are summarized in Fig. 11. As was the case for HAR-C-vine, estimation procedure again preferred copula families without tail dependence for AORD, IBEX and MXX in case of MIDAS-C-vine. For the remaining indices, most recent lags have a noticable proportion of copulas with symmetric tail dependence, other lags are dominated by copulas with upper tail dependence. Since the first tree of MIDAS-D-vine has only pairs RM_t, RM_{t-1} all copulas are equal and the results for bivariate approach are very similar. RUT and SPX have approximately the same proportion of upper tail and symmetric copulas, whereas, for AORD in half of cases copulas without tail dependence were selected.

Next we characterize the dynamics of copula specifications estimated within moving windows approach. Parameters of different copula families generally have distinct value spans and interpretations, thus constraining compatibility to copulas of similar type only. A comparison measure which is uniform and essential for all copulas is Kendall’s τ. It can be calculated empirically from data or as a coefficient implied by the copula specification. In fact, implied τ’s calculated from a vine can be different from the empirical ones, since the parameters are estimated jointly. The actually observed difference are, however, negligible. Dynamics of Kendall’s τ implied by copulas in the first tree of the HAR-C-vine are presented in panel (a) of Fig. 12. The results for MIDAS-C-vine, HAR-D-vine, MIDAS-D-vine and bivariate copulas were also calculated but are omitted since they are comparable to those of HAR-C-vine. The results demonstrate considerable time-variability of estimated HAR-C-vine through time and over all equity indices and that the strength of dependence is comparable for all three information sets and almost all time series. The time span till the end of 2007 - just before the recent financial crisis - is characterized by rapidly decreasing dependence.

Finally, we analyse the coefficients of HAR and MIDAS estimated within moving windows approach. Results for HAR and MIDAS are presented in the panels (b) and (c) of Fig. 12, respectively. Since all variables of both linear models are on the same scale, estimated model coefficients indicate comparable strength of dependence on regressors and their dynamics char-



(a) MIDAS-C-vine (b) MIDAS-D-vine

Fig. 11. The proportion of copula families with upper, symmetric and absent tail dependence in MIDAS-C-vine (a) and MIDAS-D-vine (b) estimated within moving windows approach. Each dot represents a percentage of copulas selected for the corresponding variable (x-axis), market index (color) and copula group (subplot title).

acterize the variability of model specifications over the full sample. Results for HAR reveal dynamics similar to - although less expressed than - that of Kendall's τ implied by HAR-C-vine for all market indexes over the full sample. Also, a strong diminishing effect of monthly information is apparent, although previous day and weekly data has a quite strong influence. The outliers for both upper plots are AORD and MXX. The results for MIDAS show, that the coefficients drop steeply already starting at lag 2. This contradicts the strong persistence for aggregated weekly information seen in panel (b) of Fig. 12, although is primarily a property of the exponentially decaying function.

Out-of-sample comparison of forecasts

Forecasting performance is evaluated using the three methods presented in Section 4 based on one-step-ahead point forecasts and conditional distributions for each model specification estimated within moving windows approach. Forecasts for HAR and MIDAS are extracted as median and 95% quantile of the underlying conditional distribution, which is assumed to be Gaussian.

Results of Diebold-Mariano-White test are summarized in Table 2. Each vine regression (full name at the top) is tested against all other alternatives (column heading). p-values larger than 0.95 are marked in bold and indicate that forecasts from a vine are significantly better than those from an alternative model. Values less than 0.05 are marked in italics and mean the opposite. The key conclusion is that in none of these cases was HAR or MIDAS significantly better than vines, whereas, in contrast, C- and D-vines were systematically superior to both linear benchmarks. HAR-C-vine was significantly better than HAR in 8 and MIDAS in 7 cases, whereas HAR-D-vine showed better results for 9 equity indexes. Compared to each other, HAR-D-vine provided better forecasts in 4 cases. However, HAR vines seemed inferior to the bivariate copula models at point forecasts. As further comparison show, bivariate copulas have a crucial disadvantage due to disregarding past information. Overall, MIDAS-C-vine showed the best performance among all considered models and is a clear favorite. Forecasts from this vine model significantly dominated those from HAR in 11, MIDAS - 10, HAR-C-vine - 8, HAR-D-vine - 2, MIDAS-D-vine - 6 and the bivariate copula model - 5 cases. Thus, in contrast to HAR, MIDAS seems to profit more from C-vine than from D-vine structure. Compared to bivariate copula approach, MIDAS-C-vine delivered the best results, whereas benefits of additional information were less pronounced for other vines.

Next, average CRPS over all days in the full sample are calculated and the results are presented in Table 3a, whereas the best value for each equity index is marked in bold. Smaller values of the criterion indicate that the forecasts are on average more precise and have a smaller variation. Forecasts from HAR and MIDAS are treated as deterministic, thus, CRPS of both models is absolute forecast error according to (10). For probabilistic forecasts from vine-based models the score is

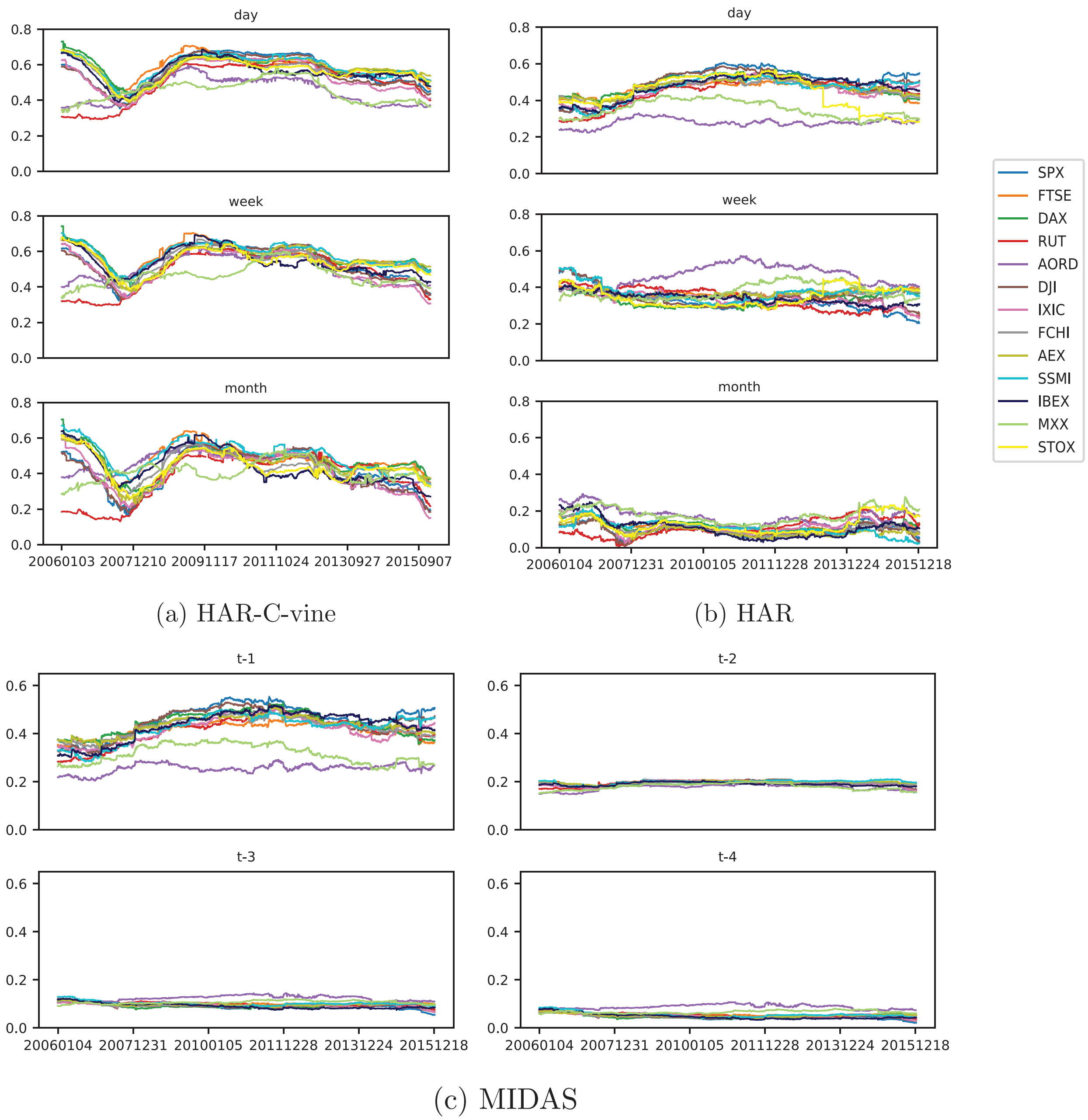


Fig. 12. Kendall τ's as implied by the 1st tree of HAR-C-vine (a), coefficients of HAR (b) and first four parameters of MIDAS (c) estimated within the moving windows approach. Market index is indicated by color.

computed based on estimated conditional CDF and the observed value. HAR-C-vine has consistently lower values of CRPS, but its superiority over HAR-D-vine, MIDAS-C-vine and bivariate copulas is very minor. Surprisingly, however, MIDAS-D-vine is the worst model within this comparison. This observation can be explained by higher variation of the forecasts produced by MIDAS-D-vine. To the best of our knowledge there is currently no test available to test the significance of the differences.

MIS and coverage probabilities are summarized in Tables 3b and 3c, respectively. Coverage probability is computed as the percentage of all observations, which were greater than the calculated 95% quantiles. Smaller values of MIS and coverage probabilities closer to 5% indicate a better forecasting performance. Since in case of volatility only the upper bound matters, MIS was calculated with the lower bound equal to zero. The dominance of bivariate copulas with respect to MIS is evident and explainable through coverage ratios, which are almost zero in all cases. Due to the design of MIS, outliers are substantially stricter penalized than wider intervals, thus, these two measures can be analyzed only jointly. We mark the best values in bold, excluding the results for bivariate copulas from ranking. Thus far HAR-C-vine is the best one with regard

Table 2
The results of DMW test for all indexes and pairs of estimated models. H means HAR, HC - HAR-C-vine, HD - HAR-D-vine, M - MIDAS, MC - MIDAS-C-vine, MD - MIDAS-D-vine, C - bivariate copula. Values greater than 0.95 are marked bold, whereas, values lower than.05 - italic.

	HAR-C-vine							HAR-D-vine					
	H	HD	M	MC	MD	C		H	HC	M	MC	MD	C
SPX	**0.998**	0.000	0.054	0.000	0.001	**1.000**	SPX	**1.000**	**1.000**	**1.000**	0.648	0.905	**1.000**
FTSE	**1.000**	0.000	**1.000**	0.001	0.676	0.000	FTSE	**1.000**	**1.000**	**1.000**	0.211	0.902	0.005
DAX	**1.000**	0.876	**1.000**	0.000	0.005	0.000	DAX	0.136	0.124	0.138	0.070	0.094	0.098
RUT	0.170	0.140	0.158	0.129	0.155	0.139	RUT	**1.000**	0.860	**0.999**	0.087	**0.995**	0.325
AORD	**1.000**	0.722	**1.000**	0.249	**0.992**	0.000	AORD	**0.982**	0.278	**0.979**	0.098	0.755	0.000
DJI	0.152	0.092	0.106	0.097	0.099	0.102	DJI	**0.986**	0.908	**0.999**	0.694	0.730	**0.997**
IXIC	**0.996**	0.000	**1.000**	0.035	0.000	0.000	IXIC	**1.000**	**1.000**	**1.000**	0.507	0.703	0.377
FCHI	**0.982**	0.919	**0.997**	0.003	0.000	0.000	FCHI	0.093	0.081	0.097	0.057	0.050	0.052
AEX	0.850	0.002	0.318	0.000	0.810	0.000	AEX	**1.000**	**0.998**	**0.997**	0.014	0.841	0.184
SSMI	**1.000**	0.070	**0.984**	0.000	0.004	0.021	SSMI	**0.997**	0.930	**0.972**	0.000	0.146	0.004
IBEX	**1.000**	0.733	**1.000**	0.000	0.812	0.002	IBEX	0.462	0.267	0.494	0.080	0.336	0.139
MXX	0.325	0.109	0.262	0.506	0.078	0.030	MXX	**0.963**	0.891	**0.997**	0.789	0.856	0.000
STOX	0.158	0.838	0.156	0.914	0.156	0.156	STOX	0.159	0.162	0.159	0.273	0.159	0.159
	MIDAS-C-vine							MIDAS-D-vine					
	H	HC	HD	M	MD	C		H	HC	HD	M	MC	C
SPX	**1.000**	**1.000**	0.352	**1.000**	0.816	**1.000**	SPX	**1.000**	**0.999**	0.095	**1.000**	0.184	**1.000**
FTSE	**1.000**	**0.999**	0.789	**1.000**	**0.950**	0.542	FTSE	0.520	0.324	0.098	0.538	0.050	0.049
DAX	**1.000**	**1.000**	0.930	**1.000**	**0.986**	**1.000**	DAX	**1.000**	**0.995**	0.906	**1.000**	0.014	0.634
RUT	**1.000**	0.871	0.913	**1.000**	**1.000**	0.880	RUT	**1.000**	0.845	0.005	0.792	0.000	0.000
AORD	**1.000**	0.751	0.902	**1.000**	**0.999**	0.000	AORD	**1.000**	0.008	0.245	**1.000**	0.001	0.000
DJI	**1.000**	0.903	0.306	0.584	0.418	0.513	DJI	**0.952**	0.901	0.270	0.915	0.582	0.791
IXIC	**0.992**	**0.965**	0.493	**0.993**	0.592	0.464	IXIC	**1.000**	**1.000**	0.297	**1.000**	0.408	0.241
FCHI	**1.000**	**0.997**	0.943	**1.000**	0.244	0.320	FCHI	**1.000**	**1.000**	**0.950**	**1.000**	0.756	0.628
AEX	**1.000**	**1.000**	**0.986**	**1.000**	0.864	**0.972**	AEX	0.199	0.190	0.159	0.185	0.136	0.152
SSMI	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**0.997**	SSMI	**1.000**	**0.996**	0.854	**0.995**	0.000	0.153
IBEX	**1.000**	**1.000**	0.920	**1.000**	**1.000**	**0.953**	IBEX	0.945	0.188	0.664	**0.977**	0.000	0.001
MXX	0.418	0.494	0.211	0.355	0.260	0.117	MXX	**1.000**	0.922	0.144	0.704	0.740	0.013
STOX	0.079	0.086	0.727	0.079	0.079	0.079	STOX	0.855	0.844	0.841	0.935	0.921	0.107

to MIS and HAR-D-vine - to coverage probabilities. Particularly interesting is the robustness of the coverage probabilities to the choice of the index.

The discrepancies in ranking models are explainable through differences among considered measures. Typically one evaluates the quality of forecasts only with loss functions. The measures we consider in this paper have a much wider scope. DMW is a test for equal predictive ability and compares the significance in the difference between losses. On the other hand, MIS quantifies the quality of forecast intervals, although the size of punitive term has crucial impact. And finally, CRPS is designed to compare different forecast densities. Thus, the discrepancy in the conclusions stem from different objectives of the performance measures. Since it is not as yet possible to test whether differences of MIS and CRPS are statistically significant, we put more weight on the results of DMW test and consider MIDAS-C-vine as the best performing model. It also shows that the linear models (MIDAS and HAR) are inferior to the copula based models. Further using more information than the last day improves the performance.

6. Summary

In this paper we tackle the problem of modeling and forecasting the time dynamics of realized variance time series. The popular models are linear in nature and fail to mimic the non-linearities in the temporal dependence. The two most popular approaches in the literature are HAR and MIDAS models, which are restricted versions of an AR(20) process. The first model, HAR, models the realized variance as a linear combination of the last day, average of the last week and the average of the last month realized variances. MIDAS imposes a slightly more complex structure, with the coefficients being determined by specific beta functions. The models suggested in this paper extends these models using pair copula constructions. The one suggested model captures the time dependence using a C-vine regression. This implies that we link the current value of the process directly to each of the explanatory variables. The second model uses a D-vine approach and it links the variables sequentially using the natural time ordering of the data. In both cases the explanatory variables are selected as in the popular benchmarks HAR and MIDAS. Within an extensive empirical study we show that the models are successful in mimicking the dynamics of volatilities and clearly outperform the alternatives in the out-of-sample forecasting.

Table 3
The results of forecasts comparison, with calculated CRPS in subtable (a), MIS in (b) and percentage of observed values, not in the 95% forecasting interval, in subtable (c). The smallest values are marked in bold. H means HAR, HC - HAR-C-vine, HD - HAR-D-vine, M - MIDAS, MC - MIDAS-C-vine, MD - MIDAS-D-vine, C - bivariate copula.

	H	HC	HD	M	MC	MD	C
SPX	0.45	**0.31**	0.33	0.44	0.33	0.59	0.32
FTSE	0.41	**0.27**	0.29	0.38	0.30	0.56	0.30
DAX	0.41	**0.29**	0.30	0.40	0.31	0.59	0.31
RUT	0.46	**0.32**	0.34	0.45	0.34	0.52	0.33
AORD	0.47	**0.33**	0.34	0.45	0.35	0.50	0.36
DJI	0.44	**0.31**	0.32	0.44	0.33	0.57	0.32
IXIC	0.44	**0.30**	0.31	0.43	0.32	0.53	0.31
FCHI	0.46	**0.28**	0.29	0.39	0.30	0.55	0.30
AEX	0.42	**0.28**	0.30	0.40	0.31	0.57	0.30
SSMI	0.40	**0.24**	0.25	0.34	0.27	0.50	0.26
IBEX	0.45	**0.27**	0.28	0.38	0.29	0.52	0.28
MXX	0.44	**0.30**	0.31	0.42	0.32	0.47	0.33
STOX	0.51	**0.30**	0.31	0.41	0.32	0.56	0.31

(a) CRPS averaged for every model and every time series over the OOS. The lowest value for each index is marked bold.

	H	HC	HD	M	MC	MD	C
SPX	47.8	50.0	**47.5**	48.8	65.1	47.9	20.1
FTSE	**32.7**	34.0	34.8	32.9	50.0	40.9	16.4
DAX	58.9	**55.0**	63.9	58.4	81.5	77.6	29.1
RUT	**54.5**	59.4	61.3	54.9	72.6	55.0	25.1
AORD	27.9	29.7	30.5	27.9	36.3	**26.4**	14.4
DJI	**47.1**	53.0	48.0	49.1	62.3	48.9	18.7
IXIC	45.1	**40.9**	42.5	44.9	53.3	44.1	19.5
FCHI	58.0	**56.1**	63.0	57.9	70.5	62.6	28.0
AEX	47.0	**44.7**	47.1	47.3	53.1	49.4	22.7
SSMI	41.4	**40.6**	42.4	41.2	55.0	43.9	16.3
IBEX	66.5	**64.2**	71.8	67.1	79.0	69.4	34.1
MXX	34.1	36.1	38.7	**33.2**	41.7	35.0	17.2
STOX	65.3	**63.2**	64.4	66.0	80.3	68.7	29.8

(b)MIS $\times 10^5$ calculated for all models over the whole OOS.

	H	HC	HD	M	MC	MD	C
SPX	6.8	6.2	**5.3**	6.7	8.0	6.6	0.0
FTSE	5.4	5.8	**5.4**	5.4	9.7	7.2	0.1
DAX	5.2	5.6	**5.1**	5.2	10.4	7.2	0.0
RUT	6.5	6.6	**5.9**	5.9	9.1	6.3	0.0
AORD	6.4	7.2	8.1	6.4	10.3	**5.8**	0.0
DJI	6.6	6.3	**5.1**	6.7	8.1	6.9	0.0
IXIC	6.2	6.4	**5.8**	6.4	9.1	7.4	0.0
FCHI	**5.4**	6.1	5.8	5.4	9.5	7.8	0.0
AEX	5.2	5.9	**4.9**	5.4	8.2	7.7	0.0
SSMI	**5.6**	5.9	6.5	5.8	10.9	6.8	0.0
IBEX	5.9	6.6	**5.6**	6.0	9.5	7.5	0.1
MXX	**5.7**	5.8	7.3	5.7	9.0	6.3	0.4
STOX	4.9	6.4	5.7	**5.0**	10.1	7.5	0.1

(c) Percentage of true values outside of forecasting intervals over the whole OOS

Acknowledgments

This research was partly supported by the German Science Foundation (DFG): projects CZ 86/4 and OK 103/5 “Vine copula base modeling and forecasting of multivariate realized volatility time series”.

References

Aas, K., 2016. Pair-copula constructions for financial applications: a review. Econometrics 4 (4), 43.
Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insur.: Math. Econ. 44 (2), 182–198.
Aït-Sahalia, Y., Jacod, J., 2014. High-Frequency Financial Econometrics. Princeton University Press.
Akaike, H., 1973. Information theory and an extension of the likelihood ratio principle. In: Petrov,B. N., (Ed.), Proceedings of the Second International Symposium of Information Theory. Akademiai Kiado, Budapest, 257–281.
Barndorff-Nielsen, O.E., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps. J. Financ. Econom. 2 (1), 1–37.
Beare, B., Seo, J., 2015. Vine copula specifications for stationary multivariate Markov chains. J. Time Ser. Anal. 36 (2), 228–246.
Bedford, T., Cooke, R., 2002. Vines: a new graphical model for dependent random variables. Ann. Stat. 30 (40), 1031–1068.
Beran, J., 1994. Statistics for Long-Memory Processes, 61. CRC Press.
Bollerslev, T., Engle, R.F., Nelson, D.B., 1994. ARCH models 4, 2959–3038.
Bollerslev, T., Jubinski, D., 1999. Equity trading volume and volatility: latent information arrivals and common long-run dependencies. J. Bus. Econ. Stat. 17 (1), 9–21.
Brechmann, E.C., Czado, C., 2014. COPAR - multivariate time series modeling using the copula autoregressive model. Appl. Stoch. Models Bus. Ind. 31 (4), 495–514.
Brechmann, E.C., Czado, C., Aas, K., 2012. Truncated regular vines in high dimensions with application to financial data. Can. J. Stat. 40 (1), 68–85.
Brechmann, E.C., Heiden, M., Okhrin, Y., 2018. A multivariate volatility vine copula model. Econ. Rev. 37 (4), 281–308.
Chen, X., Fan, Y., 2006. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. J. Econom. 135 (1), 125–154.
Chen, X., Fan, Y., 2006. Estimation of copula-based semiparametric time series models. J. Econom. 130 (2), 307–335.
Chiriac, R., Voev, V., 2011. Modelling and forecasting multivariate realized volatility. J. Appl. Econom. 26 (6), 922–947.
Corsi, F., 2009. A simple approximate long-memory model of realized volatility. J. Financ. Econom. 7 (2), 174–196.
Corsi, F., Mittnik, S., Pigorsch, C., Pigorsch, U., 2008. The volatility of realized volatility. Econom. Rev. 27 (1–3), 46–78.
Czado, C., 2010. Pair-copula constructions of multivariate copulas. In: Copula Theory and its Applications. Springer, pp. 93–109.
Czado, C., Brechmann, E.C., Gruber, L., 2012. Selection of vine copulas. In: Copulae in Mathematical and Quantitative Finance. Springer, pp. 17–37.
Darsow, W.F., Nguyen, B., Olsen, E.T., 1992. Copulas and Markov processes. Illinois J. Math. 36 (4), 600–642.
Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. J. Bus. Econ. Stat. 13 (3), 253–263.

Dißmann, J., Brechmann, E.C., Czado, C., Kurowicka, D., 2013. Selecting and estimating regular vine copulae and application to financial returns. Comput. Stat. Data Anal. 59, 52–69.
Fink, H., Klimova, Y., Czado, C., Stöber, J., 2017. Regime switching vine copula models for global equity and volatility indices. Econometrics 5 (1).
Genest, C., Favre, A.-C., 2007. Everything you wanted to know about copula modeling but were afraid to ask. J. Hydrol. Eng. 12 (4), 15–29.
Ghysels, E., Santa-Clara, P., Valkanov, R., 2004. The MIDAS Touch: Mixed Data Sampling Regression Models. UCLA: Finance.
Gneiting, T., Balabdaoui, F., Raftery, A.E., 2007. Probabilistic forecasts, calibration and sharpness. J. Royal Stat. Soc.: Ser. B (Stat. Methodol.) 69 (2), 243–268.
Gneiting, T., Raftery, A., Balabdaoui, F., Westveld, A., 2004. Verifying probabilistic forecasts: calibration and sharpness. In: Proceedings of the 17th Conference on Probability and Statistics in the Atmospheric Sciences, 2. American Meteorological Society, Seattle, WA. Preprints
Gneiting, T., Raftery, A.E., 2007. Strictly proper scoring rules, prediction, and estimation. J. Am. Stat. Assoc. 102 (477), 359–378.
Granger, C.W., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. J. Time Ser. Anal. 1 (1), 15–29.
Ibragimov, R., 2009. Copula-based characterizations for higher order Markov processes. Econom. Theory 25 (03), 819–846.
Joe, H., 1996. Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. Lect. Notes-Monogr. Ser. 28, 120–141.
Joe, H., 2014. Dependence Modeling with Copulas. Chapman & Hall/CRC.
Kraus, D., Czado, C., 2017. D-vine copula based quantile regression. Comput. Stat. Data Anal. 110, 1–18.
Kurowicka, D., Cooke, R., 2006. Uncertainty Analysis with High Dimensional Dependence Modelling. John Wiley & Sons.
Liu, L.Y., Patton, A.J., Sheppard, K., 2015. Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. J. Econom. 187 (1), 293–311.
Manner, H., 2007. Estimation and model selection of copulas with an application to exchange rates. Research Memorandum. METEOR.
Morales-Napoles, O., 2008. Bayesian beliefs nets and vines in aviation safety and other applications. Technical University of Delft (Ph.D. thesis).
Nelsen, R., 2013. An Introduction to Copulas, Springer Science & Business Media, 139.
Noh, H., Ghouch, A.E., Bouezmarni, T., 2013. Copula-based regression estimation and inference. J. Am. Stat. Assoc. 108 (502), 676–688.
Patton, A.J., 2011. Volatility forecast comparison using imperfect volatility proxies. J. Econom. 160 (1), 246–256.
Patton, A.J., Sheppard, K., 2009. Evaluating volatility and correlation forecasts. In: Handbook of Financial Time Series. With a foreword by Robert Engle. Springer, Berlin, pp. 801–838.
Simard, C., Remillard, B., 2015. Forecasting time series with multivariate copulas. Depend. Model. 3 (1).
Smith, M.S., 2015. Copula modelling of dependence in multivariate time series. Int. J. Forecast. 31 (3), 815–833.
Smith, M.S., Vahey, S.P., 2016. Asymmetric forecast densities for US macroeconomic variables from a Gaussian copula model of cross-sectional and serial dependence. J. Bus. Econ. Stat. 34 (3), 416–434.
Sokolinskiy, O., van Dijk, D., 2011. Forecasting Volatility with Copula-Based Time Series Models. Tinbergen Institute Discussion Papers. 11-125/4
Stöber, J., Czado, C., 2017. Pair copula constructions. In: Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications. World Scientific Publishing, Singapore, pp. 185–230.
Stöber, J., Joe, H., Czado, C., 2013. Simplified pair copula constructions - limitations and extensions. J. Multivar. Anal. 119, 101–118.
Tauchen, G., Pitts, M., 1983. The price variability-volume relationship on speculative markets. Econometrica 51 (2), 485–505.
West, K., 1996. Asymptotic inference about predictive ability. Econometrica 64, 1067–1084.
Zaffaroni, P., 2004. Contemporaneous aggregation of linear dynamic models in large economies. J. Econom. 120 (1), 75–102.