Modelling Yield Curves and Economic Factors using Vine and Factor Copula Models

Imrane Tola

Thesis for the attainment of the academic degree

Master of Science (M.Sc.)

at the Department of Mathematics of the Technical University of Munich.

Study programme
Mathematical Finance and Actuarial Science

Examiner
Prof. Claudia Czado Ph.D.

Submitted
Munich, 29.01.2024

Declaration of Originality

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Munich, 29.01.2024

Imrane Tola

Imrane Tola

Abstract

This thesis focuses on modelling and forecasting the yield curves and economic factors using vine, factor copula models and factor models. To remove serial dependencies, time series models were employed. The autoregressive moving average model was initially considered, which was then extended by the GARCH model with skew Student's t innovation. The standardized residuals were then formed and transformed into copula data using the probability integral transform. The dependencies were then modelled using the vine copula models, the regular, canonical and drawable vine copula, the one- and two-factor copula model, the Student's t copula and finally the one- and two-factor Gaussian model. The resulting models were compared to determine the best one, which was then used to simulate the forecasts. The modelling and forecasting were performed on two case studies. The first case study considered German yield curves with maturities of 1, 5, 10, 15 and 20 years. The second case study considered the same yield curves as in the first case and in addition the German inflation rate. The vine copula models produced the most accurate results when modelling and forecasting yield curves and the inflation rate.

Keywords: Modelling, forecasting, copula model, regular vine, canonical vine, drawable vine, factor copula model, factor Gaussian model, yield curves, economic factors, inflation rate.

Contents

List of Figures

List of Tables

Acronyms

ACF autocorrelation function

AIC Akaike's Information Criterion

AR autoregressive

ARMA autoregressive moving average

BBA British Bankers' Association

BIC Bayesian Information Criterion

cdf cumulative distribution function

CRPS Continuous Ranked Probability Score

C-vine canonical vine

D-vine drawable vine

GARCH generalized autoregressive conditional heteroskedasticity

GDP Gross Domestic Product

IS Interval Score

LIBOR London Interbank Offered Rate

MA moving average

PACF partial autocorrelation function

pdf probability density function

R-vine regular vine

SARMA seasonal autoregressive moving average

1 Introduction

The importance of copulas in the financial and insurance sectors has been steadily growing, driven by their critical role in several key areas: risk management, financial product valuation and portfolio optimization. The primary challenge in these sectors is modelling the dependencies between risk factors or assets. Conventional methods often fail to capture the complexity of these relationships, which is where copulas come in. Copulas offer a significant advantage by effectively visualizing non-linear dependencies. They enable the decomposition of a multivariate distribution into its marginal distributions and its dependency structure. This separation is crucial for understanding and managing the intricate interrelations in financial markets.
Although the multivariate normal distribution has been commonly used to model dependencies, it is increasingly recognized that it has limitations, particularly it's inability to account for asymmetries and heavy tails in data distributions. This is a crucial shortfall, as financial markets often exhibit such characteristics. To address these limitations, vine copula models have been introduced. These models use different bivariate copula families to increase flexibility, allowing for a more accurate representation of complex dependencies.
Another innovative approach is the factor copula model. Distinctive for its ability to handle asymmetries and heavy tails, this model does not rely on the assumption of normally distributed data. This model is particularly relevant in scenarios where dependencies in observable variables are driven by multiple unobserved (latent) variables. In finance, latent variables often correspond to underlying economic factors that are not directly observable but have a significant impact on financial instruments and markets.

The aim of this thesis is to apply these advanced copula models - vine copulas and factor copulas - to model the dependence structure between yield curves and an economic factor. Yield curves are a fundamental tool in financial analysis as they represent the interest rates of bonds with different maturities. It is important to understand their relationship with key economic factors for market analysis and prediction. Additionally to modelling, future movements of these curves and factors are also predicted.

This Master's thesis is divided into two parts, each contributing to a comprehensive understanding of copula models and their applications. We commence with an exploration of the economic factors central to our analysis: yield curves and inflation rate. These two variables form the foundation upon which we build our models and insights. The mathematical background is the cornerstone of our work. Within this section, we delve into univariate parametric distributions and explore various time series models that underpin the subsequent application of copulas. Furthermore, we introduce statistical tests and model selection criteria to ensure the robustness of our analysis. Copula models take centre stage in this thesis, with a detailed examination of vine copulas, factor copulas and the Gaussian factor model. To demonstrate the practicality and effectiveness of copula models, we present two comprehensive case studies. The first case study focuses on modelling yield curves, while the second extends our analysis to encompass the interplay between yield curves and inflation.

These case studies serve as real-world applications, illustrating how copula models can be employed to analyse and forecast financial variables in a complex and dynamic environment.

2 Economic Factors

Economic factors have an impact on not only individual companies but also the economy as a whole. Economic factors encompass several indicators such as interest rates, exchange rates, inflation rates, unemployment rates, tax rates, law and policies, supply and demand, and Gross Domestic Product (GDP). Some of these economic factors are interconnected, like how interest rates affect inflation rates, which then impact unemployment rates. This thesis will focus on interest rates and inflation rates, as well as their interdependencies. These factors are now defined in detail, referring to the book by Hull (2015), Kwok (2008), Brigo and Mercurio (2007) and Conrad (2017).

2.1 Yield curves

An interest rate establishes the amount of money that a borrower promises to give to a lender. Various types of interest rates are determined according to the circumstances. Treasury rates, London Interbank Offered Rate (LIBOR), Fed Funds rates and the repo interest rate are distinguished.

Treasury Rates are the interest rates received by an investor for investments in Treasury Bills and Treasury Bonds. The government uses this type of interest rate to borrow in their own currency. A US Treasury rate denotes the interest rate at which the US government can borrow dollars. Since the government does not default on its obligations the Treasury rates can be considered completely risk free.

The **London Interbank Offered Rate** is a non-securitized short-term lending rate between banks. LIBOR rates are quoted for ten currencies and 15 different periods per trading day. These interest rates are published by the British Bankers' Association (BBA). However, prior to publication, the BBA conducted a survey among various banks to determine the interest rate at which the banks could obtain a loan. The upper and lower quartiles were disregarded, the remaining interest rates were averaged to determine the LIBOR for the day.

Federal Funds Rates are non-securitized credit interest rates. In the USA, a financial institution must have a certain amount of reserves on deposit at the US Federal Reserve. These institutions may have an excess or shortfall of funds at the Federal Reserve. If there is a shortfall, the financial institution must make up for it.This compensation then leads to overnight lending transactions, employing the UK Federal Funds Rate as the interest rate.

Collateralised loan interest rates are referred to as **repo rates**. A repurchase agreement, commonly known as a repo, is a contractual agreement where a financial institution sells its securities and repurchases them at a later stage for a higher value. The institution receives a loan where the interest rate, also known as the repo rate, is the difference between the sale price and the repurchase price. The repurchase agreement typically entails minimal default risk. If the securities owner fails to uphold the agreement, the company retains the securities. Alternatively, in case of a lending company's failure to honour the agreement, the holder receives the lender's funds.

As we have presented various types of interest rates and will discuss yield curves (alternative: interest rate curves) in this paper, we will now define yield curves.

The **yield curve** is the graphical representation of the so-called yield to maturity, where the yield to maturity represents the total interest rate that an investor will receive if the purchased bond is held until maturity. Mathematically, the yield curve may be defined as follows.

Definition 2.1 *(Yield curve) The yield curve at time t represents the graph of the function*

$$T \mapsto P(t, T), T > t$$

in which the function $P(t, T)$ is the zero coupon bond. The yield curve is a T decreasing function starting from $P(t, t) = 1$ due to the positive interest rates.

The yield curve is also referred to as term structure of discount factors.

2.2 Inflation rate

Inflation refers to the continuing rise in the price level. This means that an average of many prices must rise and this increase should not decline, i.e. the increase should be permanent. This is intended to exclude constant fluctuations in individual prices. To determine the **inflation rate**, the sustainable change in the general price level must be calculated.

Part I

Mathematical Background

Mathematical Background

This chapter outlines fundamental definitions important for time series analysis and copula modelling. The essential univariate distributions are elucidated, followed by clarifications of time series and copula models.

3 Parametric Univariate Distribution

As previously stated, we will provide a comprehensive explanation of significant univariate distributions. In this part for univariate distributions, we will use uppercase letters to denote random variables X and lowercase letters for the realizations x, i.e. $X = x$. We will use F and f to denote the probability mass function and density and these exist as we consider absolutely continuous or discrete distributions. So let us introduce some important continuous distributions. The following definitions are based on the book by Czado (2019) and Rigby et al. (2019).

Definition 3.1 *(Uniform distribution) An absolutely continuous random variable X is uniformly distributed in the interval $[a, b]$ with the density*

$$f(x) := \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & \textit{else} \end{cases} . \tag{3.1}$$

The uniformly distributed random variable is then denoted by $X \sim U(a, b)$.

Definition 3.2 *(Normal distribution) An absolutely continuous random variable X is normally distributed with mean $\mu \in \mathbb{R}$, variance $\sigma^2 > 0$ and density*

$$f(x; \mu, \sigma^2) := \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{3.2}$$

The normally distributed random variable X is then denoted by $X \sim N(\mu, \sigma^2)$. The standard normal distribution is given by $X \sim N(0, 1)$ and we denote the density and distribution function as $\phi(\cdot)$ and $\Phi(\cdot)$, respectively.

Definition 3.3 *(Student's t distribution) An absolutely continuous random variable X is Student's t distributed with mean $\mu \in \mathbb{R}$, scale parameter $\sigma^2 > 0$, degree of freedom $\nu > 0$ and density*

$$f_\nu(x; \mu, \sigma^2) := \frac{\gamma\left(\frac{\nu+1}{2}\right)}{\gamma\left(\frac{\nu}{2}\right)\sqrt{(2\pi\nu)}\sigma} \left\{1 + \left(\frac{x-\mu}{\sigma}\right)^2 \frac{1}{\nu}\right\}^{-\frac{\nu+1}{2}}. \tag{3.3}$$

The Student's t distributed random variable is then denoted by $X \sim t_\nu(\mu, \sigma^2)$. The density for the Student's t distribution in the d-dimensional case for a random vector $X = (X_1, ..., X_d)$ with $\nu > 0$ degrees of freedom, expectation vector $\boldsymbol{\mu} \in \mathbb{R}^d$, and scale parameter matrix $\boldsymbol{\Sigma} = cor(X_i, X_j) = (\rho_{ij})_{i,j=1,...,d}$ is given by

$$f_t(x_1, ..., x_d; \nu; \boldsymbol{\mu}; \Sigma) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{\frac{d}{2}}} |\boldsymbol{\Sigma}|^{-1} \left\{1 + \frac{1}{\nu}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}^{-\frac{\nu+d}{2}}.$$

The multivariate Student's t distribution is abbreviated by $t_d(\nu, \boldsymbol{\mu}, \boldsymbol{\Sigma})$.

The definitions of the following distributions are based on Rigby et al. (2019).

Definition 3.4 *(Skew t type 4 distribution) A random variable X is skew t type 4 (short ST4(μ, σ, ν, τ)) distributed with density*

$$f(x; \mu, \sigma, \nu, \tau) = \begin{cases} \frac{c}{\sigma}\left[1 + \frac{z^2}{\nu}\right]^{-\frac{-(\nu+1)}{2}} & \text{if } x < \mu \\ \frac{c}{\sigma}\left[1 + \frac{z^2}{\tau}\right]^{-\frac{-(\tau+1)}{2}} & \text{if } x \geq \mu \end{cases},$$

where $x, \mu \in \mathbb{R}, \sigma > 0, \nu > 0, \tau > 0, z = \frac{x-\mu}{\sigma}$ and $c = 2\left[\nu^{\frac{1}{2}} B\left(\frac{1}{2}, \frac{\nu}{2}\right) + \tau^{\frac{1}{2}} B\left(\frac{1}{2}, \frac{\tau}{2}\right)\right]^{-1}$.

The first moment, variance, kurtosis, and skewness of a skew t type 4 distribution can be calculated as follows:

$$\mathrm{E}(X) = \mu + \sigma E(Z) = \mu + \sigma c\left[\frac{\tau}{\tau-1} - \frac{\nu}{\nu-1}\right]$$

for $\nu > 1, \tau > 1$ where $c = [\nu^{\frac{1}{2}} B\left(\frac{1}{2}, \frac{1}{2}\right) + \tau^{\frac{1}{2}} B\left(\frac{1}{2}, \frac{\tau}{2}\right)]^{-1}$.

$$\mathrm{Var}(X) = \sigma^2 Var(Z) = \sigma^2\left(E(Z^2) - E(Z)^2\right)$$

where $E(Z^2) = \frac{c\tau^{\frac{3}{2}} B\left(\frac{1}{2}, \frac{\tau}{2}\right)}{2(\tau-2)} + \frac{c\nu^{\frac{3}{2}} B\left(\frac{1}{2}, \frac{\nu}{2}\right)}{2(\nu-2)}$.

$$\mathrm{Kurt}(X) = 3 + 3c\left[\frac{\tau^{\frac{1}{2}} B\left(\frac{1}{2}, \frac{\tau}{2}\right)}{\tau-4} + \frac{\nu^{\frac{1}{2}} B\left(\frac{1}{2}, \frac{\nu}{2}\right)}{\nu-4}\right]$$

for $\nu > 4$ and $\tau > 4$.

$$\mathrm{Skew}(X) = 2c\left[\frac{\tau^2}{(\tau-1)(\tau-3)} - \frac{\nu^2}{(\nu-1)(\tau-3)}\right]$$

for $\nu > 3$ and $\tau > 3$.

Definition 3.5 *(Logistic distribution) An absolutely continuous random variable X is logistic (LO(μ, σ)) distributed with mean μ and scale σ. Its density is given by*

$$f(x; \mu, \sigma) = \frac{1}{\sigma}\left\{exp\left[-\frac{x-\mu}{\sigma}\right]\right\}\left\{1 + exp\left[-\frac{x-\mu}{\sigma}\right]\right\}^{-2}. \tag{3.4}$$

The first moment and variance of a logistic distribution can be determined as follows

$$\mathrm{E}(X) = \mu \text{ and } \mathrm{Var}(X) = \frac{\pi^2\sigma^2}{3}.$$

Kurtosis and skewness are already known and are provided by

$$\mathrm{Kurt}(X) = 1.2 \text{ and } \mathrm{Skew}(X) = 0.$$

Definition 3.6 *(Generalized t distribution) An random variable X is generalized t (GT(μ, σ, ν, τ)) distributed with density*

$$f(x;\mu,\sigma,\nu,\tau) = \tau \left\{ 2\sigma\nu^{\frac{1}{\tau}} B\left(\frac{1}{\tau},\nu\right) \left[1 + \frac{|z|^\tau}{\nu}\right]^{\nu+\frac{1}{\tau}} \right\}^{-1}, \tag{3.5}$$

where $x, \mu \in \mathbb{R}, \sigma > 0, \nu > 0, \tau > 0, z = \frac{x-\mu}{\sigma}$. The generalized t distribution is symmetric about $x = \mu$.

Expectation of a type generalized t distribution is given by

$$\mathrm{E}(X) = \begin{cases} \mu & \nu\tau > 1 \\ \text{undefined} & \nu\tau \leq 1 \end{cases}.$$

The variance, kurtosis and skewness of a generalized t distribution are calculated in the following way

$$\mathrm{Var}(X) = \begin{cases} \frac{\sigma^2\nu^{\frac{2}{\tau}} B(3\tau^{-1},\nu-2\tau^{-1})}{B(\tau^{-1},\nu)} & \nu\tau > 2 \\ \infty & \nu\tau \leq 2 \end{cases},$$

$$\mathrm{Kurt}(X) = \begin{cases} \frac{\sigma^4\nu^{\frac{4}{\tau}} B(5\tau^{-1},\nu-4\tau^{-1})}{B(\tau^{-1},\nu)} & \nu\tau > 4 \\ \infty & \nu\tau \leq 4 \end{cases},$$

$$\mathrm{Skew}(X) = \begin{cases} 0 & \nu\tau > 3 \\ \text{undefined} & \nu\tau \leq 3 \end{cases}.$$

Definition 3.7 *(Skew normal type 2 distribution) The skew normal type 2 distribution is denoted by $SN2(\mu, \sigma, \nu)$ and its probability density function is expressed as follows*

$$f_X(x|\mu,\sigma,\nu) = \begin{cases} \frac{c}{\sigma} exp\left[-\frac{1}{2}(\nu z)^2\right] & x < \mu \\ \frac{c}{\sigma} exp\left[-\frac{1}{2}\left(\frac{z}{\nu}\right)^2\right] & x \geq \mu \end{cases},$$

where $x \in (-\infty, \infty), \mu \in (-\infty, \infty), \sigma > 0, \nu > 0, z = \frac{x-\mu}{\sigma}$ and $c = \frac{\sqrt{2\nu}}{\sqrt{\pi}(1+\nu^2)}$.

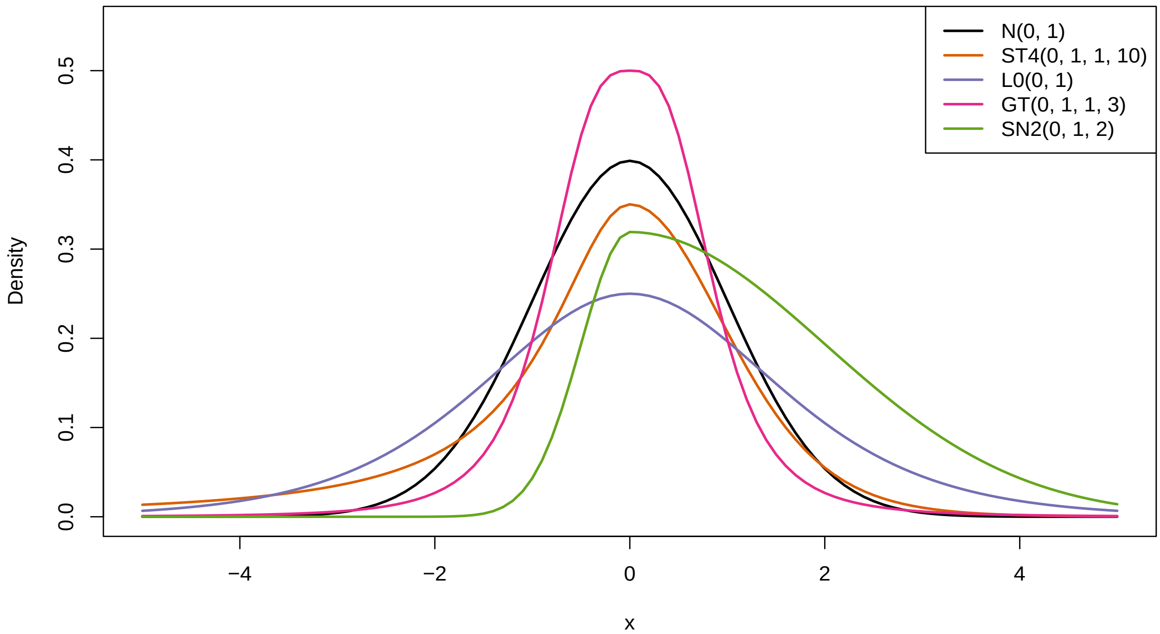

Figure 3.1 Graphical representation of the generalized t distribution, the logistic distribution, the skew t type 4 distribution, skew normal type 2 distribution and the standard normal distribution.

The expected value, variance, kurtosis and skewness of the skew normal type 2 distribution can be determined using the following functions

$$E(X) = \mu + \sigma E(Z) = \mu + \sigma \frac{\sqrt{2}}{\sqrt{\pi}}(\nu - \nu^{-1}),$$

$$\mathrm{Var}(X) = \sigma^2 \mathrm{Var}(Z) = \sigma^2 \left[(\nu^2 + \nu^{-2} - 1) - \frac{2}{\pi}(\nu^1 - \nu^{-1})^2 \right],$$

$$\mathrm{Kurt}(X) = \sigma^3 \left[\frac{2\sqrt{2}(\nu^4 - \nu^{-4})}{\sqrt{\pi}(\nu + \nu^{-1})} - 3\mathrm{Var}(Z)E(Z) - [E(Z)]^3 \right],$$

$$\mathrm{Skew}(X) = \sigma^4 \left[\frac{3(\nu^5 + \nu^{-5})}{(\nu + \nu^{-1})} - 4\frac{2\sqrt{2}(\nu^4 - \nu^{-4})}{\sqrt{\pi}(\nu + \nu^{-1})} E(Z) + 6\mathrm{Var}(Z)[E(Z)]^2 + 3[E(Z)]^4 \right],$$

where $Z = \frac{X-\mu}{\sigma}$.

Figure 3.1 shows the densities of the skew t type 4, logistic, skew normal type 2 and generalized t distributions compared to the standard normal distribution.

So far, we have only considered univariate distributions. In comparison, the multivariate distribution considers the behavior of several random variables simultaneously. In multivariate distribution, we distinguish between marginal, joint and conditional distribution. The densities and the distribution function for these three groups have been summarized in the Table 3.1.

Definition 3.8 *(Marginal, joint and conditional distributions) Let* $X = (X_1, ..., X_d)^T$ *be a d-dimensional random vector. For the marginal, joint and conditional densities and distributions the following notation is used.*

Table 3.1 Marginal, joint and conditional densities and distributions

	density function	distribution function
marginal	$f_i(x_i), i \in \{1, ..., d\}$	$F_i(x_i), i \in \{1, ..., d\}$
joint	$f(x_1, ..., x_d)$	$F(x_1, ..., x_d)$
conditional	$f_{i\|k}(x_i\|x_k), i \neq k$	$F_{i\|k}(x_i\|x_k), i \neq k$

Another significant definition is presented below and is crucial if one wants to transform random variables with an underlying distribution into the standard uniform distribution.

Definition 3.9 *(Probability integral transform) Let* $X \sim F$ *be a continuous random variable and* x *the realization of* X*, then the transformation* $u := F(x)$ *is called the probability integral transform (short PIT) at* x*.*

Remark 3.1 *(Distribution of the probability integral transform)*
Assume $X \sim F$ *then* $U := F(X)$ *is uniformly distributed in interval* $[0, 1]$*. This can be shown by*

$$P(U \leq u) = P(F(x) \leq u) = P(X \leq F^{-1}(u)) = F(F^{-1}(u)) = u, \quad \forall u \in [0, 1].$$

where F^{-1} *is the quantile function.*

4 Time Series Models

In this section, a general setup is given. The second section deals with statistical time series models and describes the most important ones in detail. The following definitions are based on the book by Shumway and Stoffer (2017).

4.1 Set up

Definition 4.1 *(Time series)* ***Time series*** *contain data that have been observed and collected over a certain period of time. Consider the time series as a stochastic process*

$$\{X_t\}, t = 0, \pm 1, \pm 2, ...$$

or as a sequence of random variables

$$X_1, X_2, X_3, ...$$

The observed values of a stochastic process are called realization and are denoted by $x_t, t = 1, 2, \ldots$.

An example for a stochastic process would be the so called white noise.

Definition 4.2 *(White noise) A* ***white noise*** *process W_t is a collection of uncorrelated random variables with $E(W_t) = 0$ and $Var(W_t) = \sigma_w^2 \in (0, \infty)$. We denote the white noise process by $W_t \sim wn(0, \sigma_w^2)$. Is W_t independet and identically normally distributed with mean 0 and variance σ_w^2, so we call this process* ***Gaussian white noise*** *and we will denote this by $W_t \overset{i.i.d}{\sim} N(0, \sigma_w^2)$.*

Time series analysis is based on the idea that the realizations of a stochastic process depend on each other at different points in time. This relationship can be described by a model or simply by covariance or correlation. Since one does not have only one realization, one has accordingly not only one covariance or correlation, but a whole sequence of covariances or correlations, which one calls autocovariance function or autocorrelation function.

Definition 4.3 *(Autocovariance function) The* ***autocovariance function*** *measures the linear dependence between two points on a time series at different times and is defined by*

$$\gamma(s,t) = cov(X_s, X_t) = E[(X_s - \mu_s)(X_t - \mu_t)] \quad \forall s, t, \tag{4.1}$$

where μ_s, μ_t are the mean for the time series at time s and t. For $s = t$ we get the variance

$$\gamma(t,t) = E[(X_t - \mu_t)^2] = var(X_t)$$

and if $\gamma(s,t) = 0$, X_s and X_t are not linearly related but there must be some dependence structure between the time series.

Definition 4.4 *(Autocorrelation function) With the* ***autocorrelation function (ACF)****, the linear predictability of a time series at time t is determined by considering the time series at time s. The autocorrelation function is defined as follows*

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}}. \tag{4.2}$$

The autocorrelation function $\rho(s,t)$ take values between -1 and 1.

Definition 4.5 *(Partial autocorrelation function) Let $\{X_t, t = 1, ..., T\}$ be a stationary time series. The* ***partial autocorrelation function (PACF)*** *is the correlation between the time series X_{t+h} and X_t and is denoted by $\phi_{hh}, h = 1, 2, ...$ with*

$$\begin{aligned} \phi_{11} &= cor(X_{t+1}, X_t) = \rho(1,1) \\ \phi_{hh} &= cor(X_{t+h} - \hat{X}_{t+h}, X_t - \hat{X}_t) \end{aligned}$$

where $\hat{X}_{t+h}$ is linearly regressed on $X_{t+h-1}, ..., X_{t+1}$. $\hat{X}_{t+h}$ is then denoted by

$$\hat{X}_{t+h} = \beta_1 X_{t+h-1} + \beta_2 X_{t+h-2} + ... + \beta_{h-1} X_{t+1}$$

and $\hat{X}_t$ is the linear regression on $X_{t+1}, ..., X_{t+h-1}$. Linear regression of $\hat{X}_t$ is then given by

$$\hat{X}_t = \beta_1 X_{t+1} + \beta_2 X_{t+2} + ... + \beta_{h-1} X_{t+h-1}.$$

Until now, no assumptions have been made about the behavior of time series. It might well be the case that some kind of regularity is present in the time series. This regularity is introduced with the following definition.

Definition 4.6 *(Stationarity) A* ***strictly stationary*** *time series is one where the probabilistic behavior of every collection of values $\{X_{t_1}, X_{t_2}, ..., X_{t_k}\}$ is identical to that of the time shifted set $\{X_{t_1+h}, X_{t_2+h}, ..., X_{t_k+h}\}$. So we get*

$$P(X_{t_1} \leq c_1, ..., X_{t_k} \leq c_k) = P(X_{t_1+h} \leq c_1, ..., X_{t_k+h} \leq c_k) \tag{4.3}$$

with time points $t_1, ..., t_k$, constants $c_1, ..., c_k$ and time shifts $h = 0, \pm1, \pm2, ... \forall k = 1, 2, ...$

This version of the definition is too strong for some applications, so a weak version of stationarity is needed. We present this in the following.

Definition 4.7 *(Weak Stationarity) A* ***weakly stationary*** *time series $\{X_t, t = 1, ..., T\}$ is a process with finite variance, for which holds*

1. *the mean value function μ_t is constant and does not depend on time t,*
2. *the autocovariation function $\gamma(s,t)$ depends on time s and t only through their difference $|s-t|$.*

Now that important definitions of time series analysis have been introduced, we proceed to examine time series models that attempt to explain complex processes using simpler ones, such as white noise. The autoregressive, moving average and autoregressive moving average models are potential candidates, and we delve into them in greater detail below. The ensuing definitions are derived from the book by Shumway and Stoffer (2017) and Brockwell and Davis (2016).

4.2 Statistical time series models

In this section, statistical time series models are defined. We will begin by examining the autoregressive model, the moving average model and the autoregressive moving average model, which combines the previous two models. Generalized autoregressive conditionally heteroskedastic models are also discussed in detail.

4.2.1 Autoregressive model

Definition 4.8 *(Autoregressive model) An autoregressive model of order p (short AR(p)) for a stationary time series $\{X_t, t = 1, ..., T\}$ is given by*

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + W_t, \tag{4.4}$$

where $W_t \overset{i.i.d.}{\sim} wn(0, \sigma_w^2)$ is a ***white noise*** *with zero mean and constant variance σ_w^2 and $\phi_1, ..., \phi_p$ constant model parameters with $\phi_p \neq 0$. Is the mean μ of time series not equal to zero, then replace the time series X_t with $X_t - \mu$ and one gets*

$$X_t = \alpha + \sum_{i=1}^{p} \phi_i X_{t-i} + W_t,$$

where $\alpha = \mu(1 - \phi_1 - ... - \phi_p)$.

For a better understanding of AR models, let's consider the model of order one AR(1) with the form $X_t = \phi X_{t-1} + W_t$. Figure 4.1 shows a path of an autoregressive process of order one with parameter $\phi = 0.6$. Iterating backwards in time, the autoregressive model with order one can be represented as a linear process, provided $|\phi| < 1$ and $\sup_t \operatorname{var}(X_t) < \infty$.

$$\begin{aligned} X_t &= \phi X_{t-1} + W_t \\ &= \phi\left(\phi X_{t-2} + W_{t-1}\right) + W_t \\ &\;\vdots \\ &= \sum_{j=0}^{\infty} \phi^j W_{t-j} \end{aligned} \tag{4.5}$$

Using (4.5) we now can calculate the mean

$$E(X_t) = \sum_{j=0}^{\infty} \phi^j E(W_{t-j}) = 0,$$

the autocovariance

$$\begin{aligned} \gamma(t+h, t) &= \operatorname{cov}(X_{t+h}, X_t) \\ &= E\left[\left(\sum_{j=0}^{\infty} \phi^j W_{t+h-j}\right)\left(\sum_{k=0}^{\infty} \phi^k W_{t-k}\right)\right] \\ &= \sigma_w^2 \sum_{j=0}^{\infty} \phi^{h+j} \phi^j \\ &= \frac{\sigma_w^2 \phi^h}{1 - \phi^2}, \; h \geq 0 \end{aligned}$$

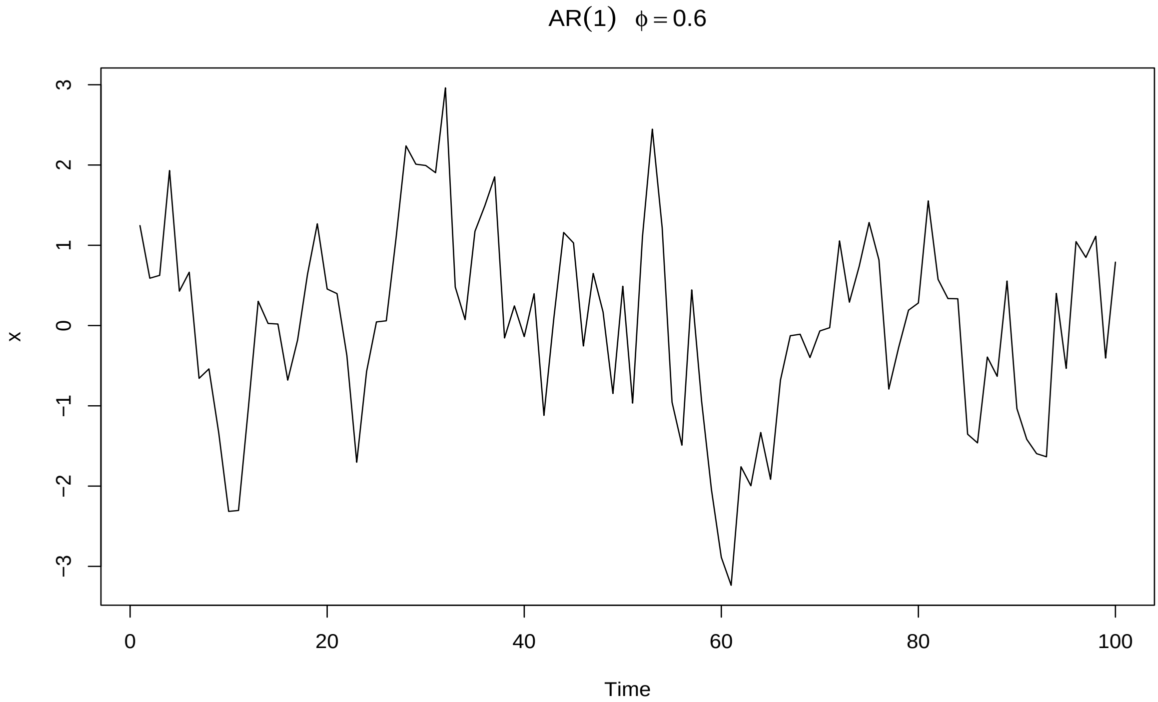

Figure 4.1 A realized path of an autoregressive model of order one AR(1) (Source: Based on Shumway and Stoffer (2017))

and the ACF of the AR(1) process

$$\rho(t+h,t) = \frac{\gamma(t+h,t)}{\gamma(t,t)} = \phi^h,\ h \geq 0.$$

4.2.2 Moving average model (MA)

Definition 4.9 *(Moving average models) The moving average (MA) model of order q for a stationary time series $\{X_t, t = 1, ..., T\}$ is defined by*

$$X_t = W_t + \sum_{i=1}^{q} \theta_i W_{t-i} \tag{4.6}$$

with white noise $W_t \overset{i.i.d.}{\sim} wn(0, \sigma_w^2)$ and constant model parameters $\theta_1, ..., \theta_q$ with $\theta_q \neq 0$. We abbreviate the moving average model of order q by MA(q).

Let's consider the first order model MA(1) to get a better insight. The first order moving average model can be represented by the following formula: $X_t = W_t + \theta W_{t-1}$. We can then

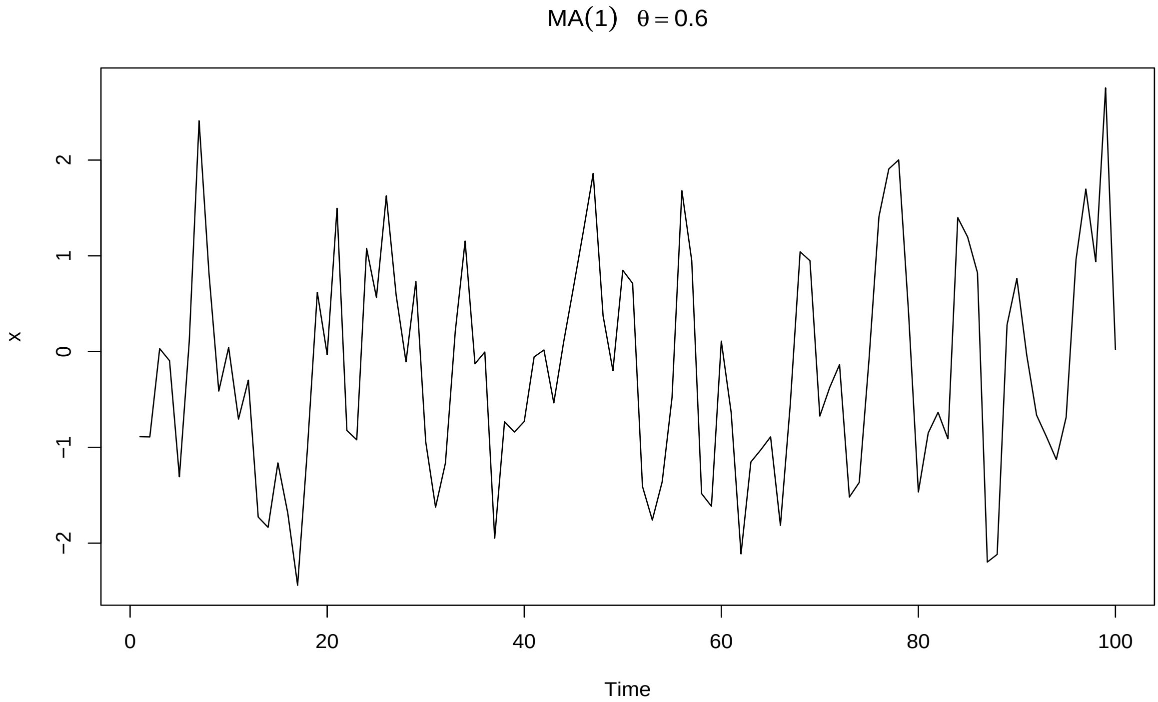

Figure 4.2 A realized path of an moving average model of order one MA(1) with $\theta = 0.6$. (Source: Based on Shumway and Stoffer (2017))

use this representation to calculate the moments of the time series.

$$E(X_t) = E(W_t) + \theta E(W_{t-1}) = 0,$$

$$\begin{aligned}\gamma(t+h,t) &= \text{cov}(X_{t+h}, X_t) \\ &= \begin{cases} (1+\theta^2)\sigma_w^2 & h = 0, \\ \theta\sigma_w^2 & h = 1, \\ 0 & h > 1. \end{cases}\end{aligned}$$

The autocorrelation function is given by

$$\rho(t+h,t) = \begin{cases} \frac{\theta}{(1+\theta^2)} & h = 1, \\ 0 & h > 1. \end{cases}$$

A realized path of the MA model of order one with $\theta = 0.6$ is shown in Figure 4.2.

Combining an autoregressive and moving average model then one gets an autoregressive moving average model with order p and q. We abbreviate the autoregressive moving average model by ARMA(p,q).

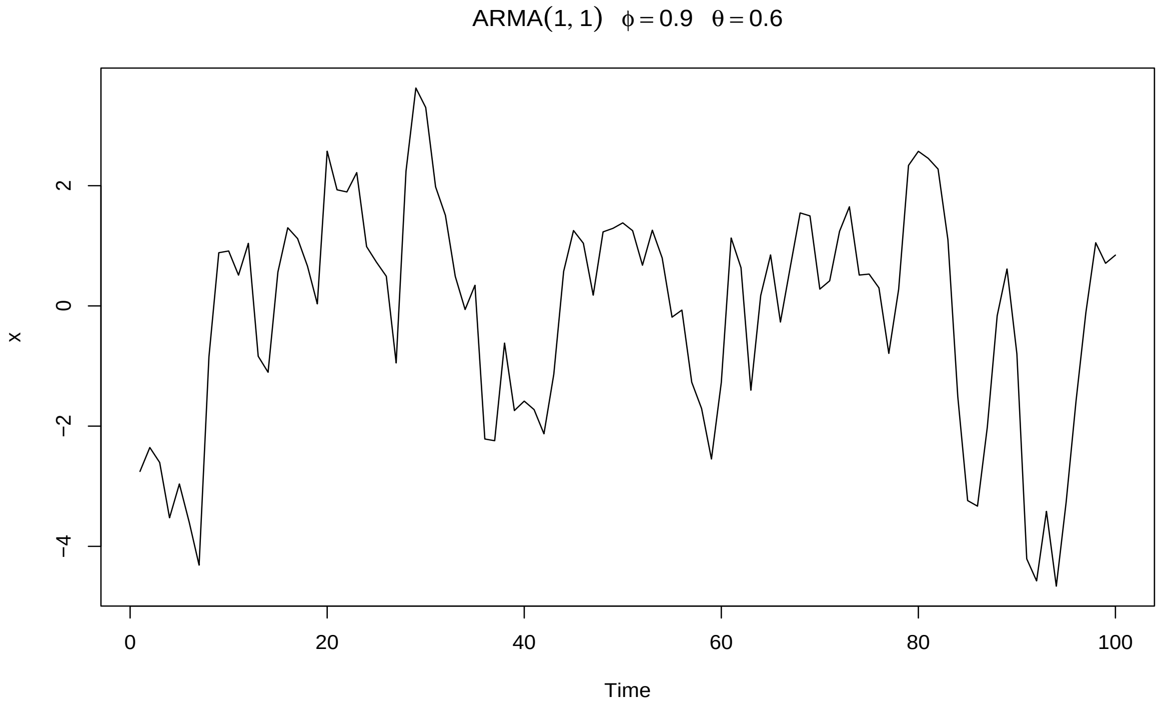

Figure 4.3 A realized path of an autoregressie moving average model of order $p = 1$ and $q = 1$ ARMA(1, 1) with $\phi = 0.9, \theta = 0.6$. (Source: Based on Shumway and Stoffer (2017))

4.2.3 Autoregressive moving average model (ARMA)

Definition 4.10 *(Autoregressive moving average model) An autoregressive moving average (ARMA) model of order p and q for a stationary time series $\{X_t, t = 1, ..., T\}$ is given by*

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + W_t + \sum_{j=1}^{q} \theta_j W_{t-j} \tag{4.7}$$

with the white noise $W_t \sim wn(0, \sigma_w^2)$ and parameters $\phi_p \neq 0$ and $\theta_q \neq 0$. Note again, if the mean is not zero, we replace the time series X_t by $X_t - \mu$ and we obtain

$$X_t = \alpha + \sum_{i=1}^{p} \phi_i X_{t-i} + W_t + \sum_{j=1}^{q} \theta_j W_{t-j}$$

where α is defined by $\mu(1 - \phi_1 - ... - \phi_p)$.

Choosing order p equal to zero, one obtains a moving average model with order q. Taking q equal to zero, one receives an autoregressive model with order p. An autoregressive moving average model with order $p = 1$ and $q = 1$ and parameters $\phi = 0.9$ and $\theta = 0.6$ is shown in Figure 4.3.

Order selection AR and MA model

In order to fit an ARMA model, the orders must first be determined. The orders can be obtained using the autocorrelation function and the partial autocorrelation function. The order p is given if the ACF tails off and the PACF cuts off after lag p. If the ACF cuts off after lag q and PACF tails off, the order q is obtained. For orders p and q in the ARMA model, ACF and PACF must tail off. The behavior of the ACF and PACF is summarized in Table 4.1.

Table 4.1 Illustration the behavior of the ACF and PACF for ARMA models. (Source: Shumway and Stoffer (2017)[page 99])

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

In Figure 4.4 the order q can be read from the ACF and p from the PACF plot. In AR(1) the ACF tails off and PACF is close to zero for lag greater than one. In MA(1) the ACF and PACF is close to zero for lag greater than one. The ARMA tails off in the ACF and PACF is close to zero for lag greater than one.

4.2.4 Multiplicative seasonal autoregressive moving average model (SARMA)

For some time series, dependence on the past is often strongest at multiples of the underlying seasonal lag. The presented ARMA model must therefore be modified. In the following, an autoregressive and moving average model that identify with the seasonal lags s is introduced.

Definition 4.11 *(Multiplicative seasonal autoregressive moving average model) An* ***pure seasonal autoregressive moving average (SARMA) model*** *(short ARMA$(P, Q)_s$) for a stationary time series $\{X_t, t = 1, ..., T\}$ is defined by*

$$\Phi_P(B^s)X_t = \Theta_Q(B^s)W_t \tag{4.8}$$

with the operators

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2(B^{2s}) - ... - \Phi_P(B^{2P})$$

and

$$\Theta_Q(B^s) = 1 + \Theta_1 B^s + \Theta_2(B^{2s}) + ... + \Theta_Q(B^{2Q}),$$

where B is the backshift operator and is defined by

$$BX_t = X_{t-1}.$$

If we consider the power of two, the backshift operator is then given by $B^2X_t = B(BX_t) = BX_{t-1} = X_{t-2}$. In general the backshift-opertor for the k-th power has the following form

$$B^k X_t = X_{t-k}.$$

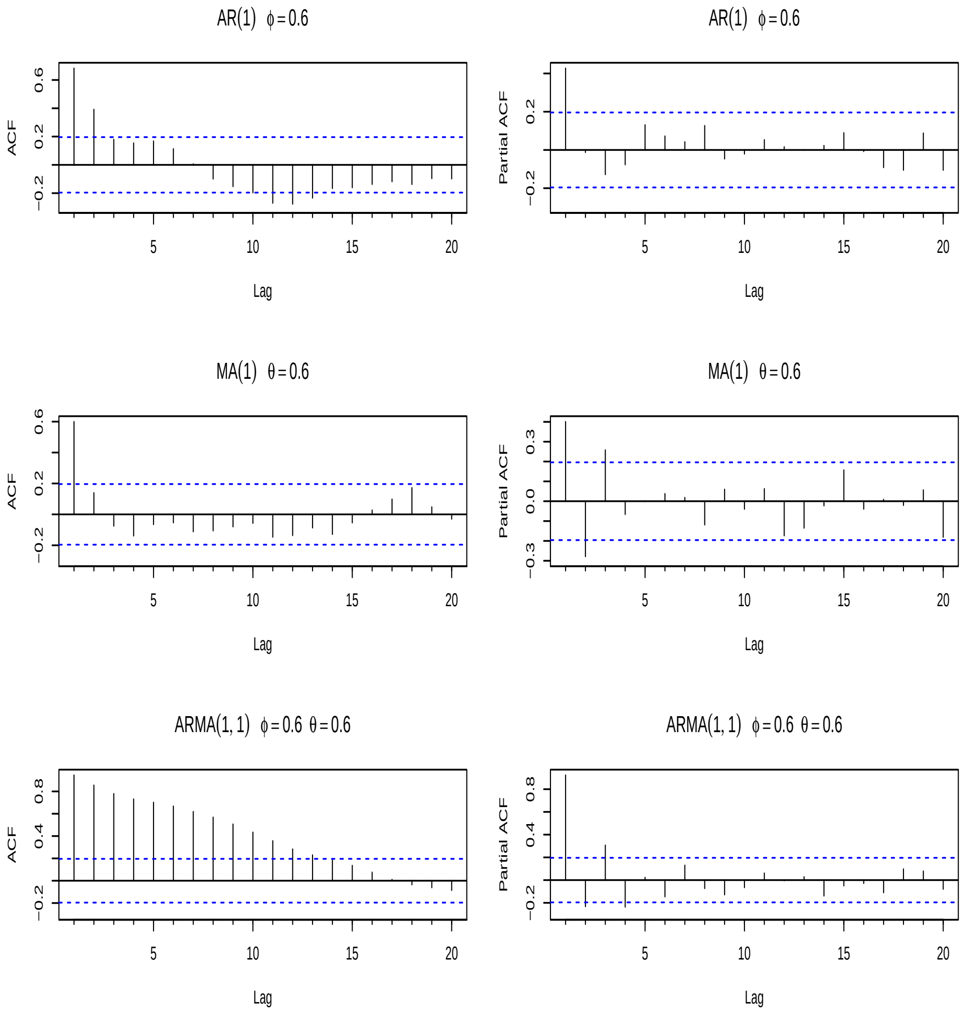

Figure 4.4 Illustration of the behavior of the ACF and PACF for AR(1), MA(1) and ARMA(1, 1).

Generally, it is possible to combine the seasonal and non-seasonal operators into a ***multiplicative seasonal autoregressive moving average model*** *and one gets an* $ARMA(p,q) \times (P,Q)_s$ *with the model formula*

$$\Phi_P(B^s)\phi(B)X_t = \Theta_Q(B^s)\theta(B)W_t. \tag{4.9}$$

Order selection SARMA model

As with the ARMA models, the order must also be determined for the pure SARMA models. We have already determined the orders for the non-seasonal part in Table 4.1. So it remains to specify the seasonal orders. We obtain the order P if the ACF tails off at lags ks and PACF cuts off after lag Ps and further we get the order Q if ACF cuts off after lag Qs and PACF tails off at lags ks. If ACF tails off and PACF tails off at lags ks, we get order Q and P. Again, the results have been summarized in Table 4.2.

Table 4.2 Illustration of the behavior of the ACF and PACF for SARMA models. (Source: Shumway and Stoffer (2017)[page 148])

	$\text{AR}(P)_s$	$\text{MA}(Q)_s$	$\text{ARMA}(P,Q)_s$
ACF*	Tails off at lags ks, k = 1, 2, ...	Cuts off after lag Qs	Tails off at lags ks
PACF*	Cuts off after lag Ps	Tails off at lags $ks, k = 1, 2, ...$	Tails off at lags ks

$*$ The value at non-seasonal lags $h \neq ks$, for $k = 1, 2, ...$ are zero

Figure 4.5 depicts the autocorrelation function and partial autocorrelation function plots for the seasonal AR(1)

$$\begin{aligned}(1 - \Phi_1 B^{12})X_t &= W_t \\ X_t &= \Phi_1 X_{t-12} + W_t,\end{aligned}$$

MA(1)

$$\begin{aligned}X_t &= (1 + \Theta_1 B^{12})W_t \\ X_t &= W_t + \Theta_1 W_{t-12},\end{aligned}$$

and ARMA(1,1)

$$\begin{aligned}(1 - \Phi_1 B^{12})X_t &= (1 + \Theta_1 B^{12})W_t \\ X_t &= \Phi_1 X_{t-12} + W_t + \Theta_1 W_{t-12}\end{aligned}$$

models. We have chosen a seasonal period of s = 12. For the seasonal AR(1), the ACF tails off after lag 12 and the PACF cuts off after lag 12. In contrast, for the seasonal MA(1) model, the ACF cuts off after lag 12 and the PACF tails off after lag 12. The seasonal ARMA(1, 1) model shows that the ACF and PACF approach zero for lag greater than 12.

So far in the ARMA models, we have assumed the conditional variance to be constant. However, this assumption is not always accurate. Therefore, we need a model that takes the variability into account. Such a model has been introduced by Engle and later extended by Bollerslev, which is described in the next definition.

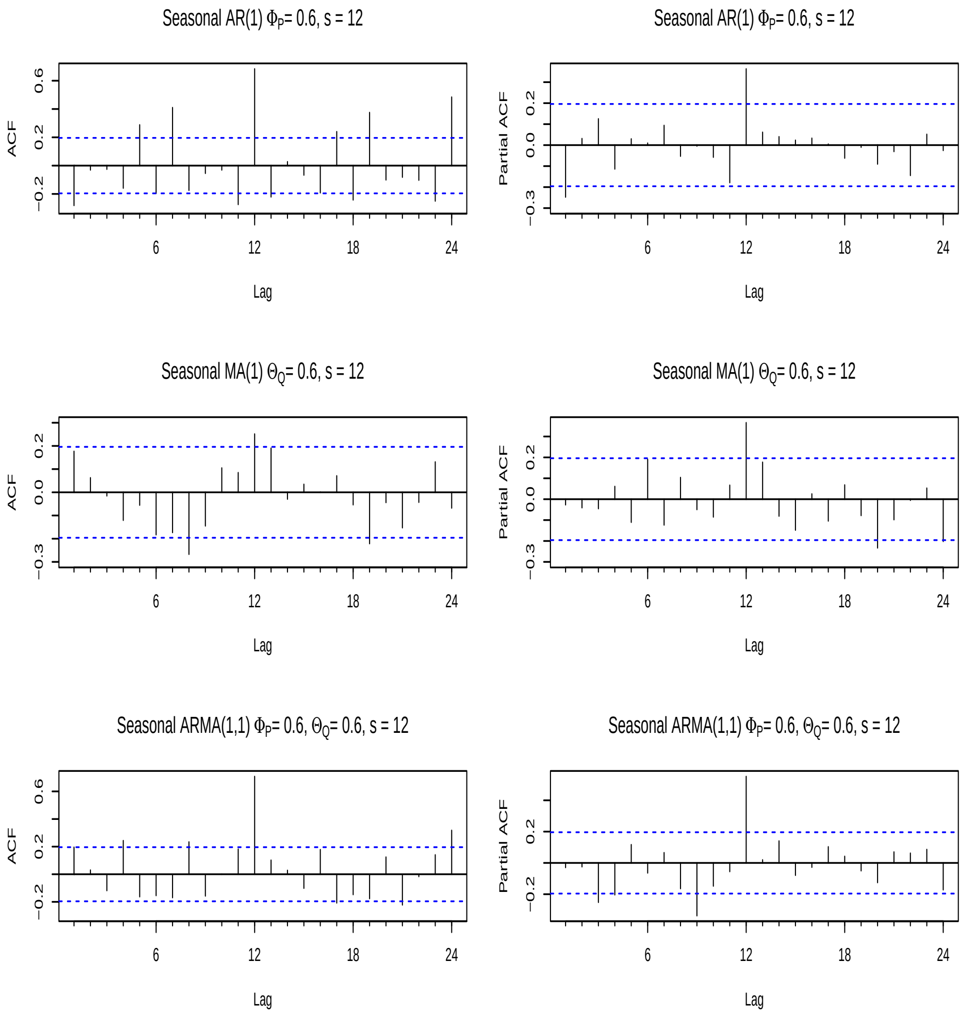

Figure 4.5 Illustration of the behavior of the ACF and PACF for an seasonal AR(1), MA(1) and ARMA(1, 1) with $s = 12$.

4.2.5 Generalized autoregressive conditionally heteroscedastic model (GARCH)

Definition 4.12 *(Generalized autoregressive conditionally heteroscedastic models) In the generalized autoregressive conditional heteroskedasticity (GARCH) model, the conditional variance of the disturbance term is assumed to follow an autoregressive moving average ARMA(p, q) process. The GARCH(P, Q) model can therefore be defined as follows*

$$\begin{aligned} W_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{P} \alpha_i W_{t-i}^2 + \sum_{j=1}^{Q} \beta_j \sigma_{t-j}^2 \end{aligned} \tag{4.10}$$

with white noise $\epsilon_t \overset{i.i.d.}{\sim} wn(0,1)$*, parameters* $\alpha_0, ..., \alpha_p, \beta_1, ..., \beta_q$ *where* $\alpha_p \neq 0$ *and* $\beta_q \neq 0$*.*

Combining the above described models we obtain the ARMA-GARCH model with orders p, q and P, Q.

4.2.6 ARMA-GARCH model

Definition 4.13 *(ARMA-GARCH models) The ARMA-GARCH model of order* p, q, P *and* Q *for a stationary time series* $\{X_t, t = 1, ..., T\}$ *has the following representation*

$$\begin{aligned} ARMA(p,q) : X_t &= \sum_{i=1}^{p} \phi_i X_{t-i} + W_t + \sum_{j=1}^{q} \theta_j W_{t-j} \\ GARCH(P,Q) : W_t &= \sigma_t \epsilon_t, \sigma_t^2 = \alpha_0 + \sum_{k=1}^{P} \alpha_k W_{t-k}^2 + \sum_{l=1}^{Q} \beta_l \sigma_{t-l}^2 \\ ARMA(p,\, q) - GARCH(P,Q) : X_t &= \sum_{i=1}^{p} \phi_i X_{t-i} + \sqrt{\alpha_0 + \sum_{k=1}^{P} \alpha_k W_{t-k}^2 + \sum_{l=1}^{Q} \beta_l \sigma_{t-l}^2} \cdot \epsilon_t \\ &+ \sum_{j=1}^{q} \theta_j \sqrt{\alpha_0 + \sum_{k=1}^{P} \alpha_k W_{(t-j)-k}^2 + \sum_{l=1}^{Q} \beta_l \sigma_{(t-j)-l}^2} \cdot \epsilon_{t-j} \end{aligned} \tag{4.11}$$

with $\varepsilon_t \overset{i.i.d.}{\sim} wn(0,1)$ *and model parameters* $\phi_1, ..., \phi_p, \theta_1, ..., \theta_q$ *and* $\alpha_0, ..., \alpha_p, \beta_1, ..., \beta_q$ *with* $\phi_p \neq 0, \theta_q \neq 0, \alpha_p \neq 0$ *and* $\beta_q \neq 0$*.*

4.3 Parameter estimation

The models presented consist of parameters such as $\phi_1, \ldots, \phi_p$, $\theta_1, \ldots, \theta_q$, and σ_w^2, which need to be estimated. The model parameters may be estimated through either maximum likelihood or least squares techniques. In the following, the parameter estimation for the AR(1) model is derived with the maximum likelihood method, whereas the parameter estimation for the ARMA (p, q) model is only briefly described.
The estimation of parameters for the AR(1) model is now derived. An AR(1) model is defined by

$$X_t = \mu + \phi(X_{t-1} - \mu) + W_t,$$

where W_t is independently and identically normally distributed with a mean of zero and a variance of σ_w^2 and $|\phi| < 1$.The likelihood function for the observations $x_1, \ldots, x_n$ is then given by

$$L(\mu, \phi, \sigma_w^2) = f(x_1, x_2, \ldots, x_n; \mu, \phi, \sigma_w^2).$$

For the special case AR(1), the likelihood function is given by

$$L(\mu, \phi, \sigma_w^2) = f(x_1)f(x_2|x_1)\cdots f(x_n|x_{n-1}) = f(x_1)\prod_{t=2}^{n} f(x_t|x_{t-1}). \tag{4.12}$$

Since $X_t|X_{t-1}$ follows a normal distribution and has a mean $\mu + \phi(x_{t-1} - \mu)$ and variance σ_w^2, the density of x_t given x_{t-1} can be written by

$$f(x_t|x_{t-1}) = f_w((x_t - \mu) - \phi(x_{t-1} - \mu))$$

utilizing f_w, which is the density of the white noise $W_t \sim N(0, \sigma_w^2)$. The likelihood function expressed in Equation (4.12) can be written as

$$L(\mu, \phi, \sigma_w^2) = f(x_1)\prod_{t=2}^{n} f_w((x_t - \mu) - \phi(x_{t-1} - \mu)).$$

To determine the distribution of X_1, the following representation

$$X_1 = \mu + \sum_{j=0}^{\infty} \phi^j W_{1-j} \tag{4.13}$$

must be considered. According to the Equation (4.13), X_1 follows a normal distribution with mean μ and variance $\frac{\sigma_w^2}{(1-\phi^2)}$. Finally, the likelihood function is given by

$$L(\mu, \phi, \sigma_w^2) = \frac{1}{\sqrt{(2\pi\sigma_w^2)^n}}\sqrt{1-\phi^2}\exp^{\left(-\frac{S(\mu,\phi)}{2\sigma_w^2}\right)}, \tag{4.14}$$

where $S(\mu, \phi)$, the unconditional sum of squares, is defined by

$$S(\mu, \phi) = (1 - \phi^2)(x_1 - \mu)^2 + \sum_{t=2}^{n}((x_t - \mu) - \phi(x_{t-1} - \mu))^2.$$

To obtain the parameters, it is necessary to take the logarithm of the Equation (4.14). The partial derivative with respect to σ_w^2 is performed and set to zero. The maximum likelihood estimator for σ_w^2 is then given by

$$\hat{\sigma}_w^2 = \frac{S(\hat{\mu}, \hat{\phi})}{n}, \tag{4.15}$$

where $\hat{\mu}, \hat{\phi}$ are the maximum likelihood estimators of μ and ϕ. To estimate the parameters of μ and ϕ, the function $S(\mu, \phi)$ needs to be minimized. By replacing n with $n-2$ in the Equation (4.15), an unconditional least squares estimate of σ_w^2 can be obtained. After replacing σ_w^2 with $\hat{\sigma}_w^2$ in the log-likelihood function, the estimators $\hat{\mu}$ and $\hat{\phi}$ can be obtained through minimizing the equation

$$l(\mu, \phi) = \log\left(\frac{S(\mu, \phi)}{n}\right) - \frac{1}{n}\log(1-\phi^2),$$

where $l(\mu, \phi) \propto -2\log L(\mu, \phi, \hat{\sigma}_w^2)$.

To estimate the parameters for the ARMA(p, q) model, the likelihood function needs to be established first. The ARMA model has $p+q+1$ parameters, defined by $\beta := (\mu, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)^t$. The likelihood function is given by

$$L(\beta, \sigma_w^2) = \prod_{t=1}^{n} f(x_t | x_{t-1}, \ldots, x_1).$$

The conditional distribution of $X_t | X_{t-1}, \cdots, X_1$ is the normal distribution with mean x_t^{t-1} and variance P_t^{t-1}, where $P_t^{t-1} = \gamma(t,t) \prod_{j=1}^{t-1}(1-\phi_{jj}^2)$. Now we can follow the same steps as the AR(1) model and derive the parameter σ_w^2 through maximization and the estimator for β through minimization.

To estimate the parameters of the GARCH model using the maximum likelihood method, the steps previously outlined for the AR(1) model must be replicated.

5 Statistical Tests

The purpose of this chapter is to define some statistical hypothesis tests which will be used later. These are tests that check the sample for the presence of a normal distribution, for autocorrelation coefficients and for ARCH effects. The Jarque-Bera and the Shapiro-Wilk test are used to test the sample for normal distribution. The autocorrelation coefficients are checked using the Ljung-Box test. Finally, we test for the presence of ARCH effects on the given data using the LM ARCH test. The definitions presented are based on the books by Shapiro (1964), Neusser (2011), Tsay (2009) and the papers by Engle (1982), Catani and Ahlgren (2017), GEL and CHEN (2012) and Lee (1991).

Definition 5.1 *(Jarque-Bera Test) The* ***Jarque-Bera Test*** *is carried out to assess whether the provided i.i.d sample* $(x_i)_{i=1,\dots,n}$ *from a random variable X follows a normal distribution.*

Hypothesis: $H_0 : X \sim N(\mu, \sigma^2)$ *vs.* $H_1 : X \nsim N(\mu, \sigma^2)$

Test statistic: $JB = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right) \overset{\text{under } H_0}{\sim} \chi^2_2$,

where n is the number of observations,

$$S = \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \bar{X})^3}{\left(\frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \bar{X})^2\right)^{\frac{3}{2}}}$$

is the estimated random skewness and

$$K = \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \bar{X})^4}{\left(\frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \bar{X})^2\right)^{2}}$$

the estimated random kurtosis.

Rejection rule: *If* $JB > \chi^2_{1-\alpha,2}$*, the null hypothesis is rejected, where* $\chi^2_{1-\alpha,2}$ *is the* $(1-\alpha)$*-quantile of the* χ^2_2 *distribution.*

Definition 5.2 *(Shapiro-Wilk Test) The* ***Shapiro-Wilk Test*** *determines if the sample data* $(x_i)_{i=1,\dots,n}$ *is compatible with a normal distribution with unknown mean* μ *and unknown variance* σ^2.

Hypothesis: $H_0 : X \sim N(\mu, \sigma^2)$ *vs.* $H_1 : X \nsim N(\mu, \sigma^2)$

Test statistic: $W = \frac{b^2}{S^2} = \frac{b^2}{\sum\limits_{i=1}^{n}(X_i - \bar{X})^2}$.

For even sample size $n = 2k$, then we define $b = \sum_{i=1}^{k} a_{n-i+1}(x_{n-i+1} - x_i)$. If $n = 2k+1$ is odd, then $b = a_n(x_n - x_1) + \ldots + a_{k+2}(x_{k+2} - x_k)$. Coefficients a_i are approximated and given in tables in Shapiro (1964)[page 603 f].

Rejection rule: *If the value of W exceeds a critical threshold $W_{critical}$, the null hypothesis will not be rejected. Approximations of $W_{critical}$ can be found in tables provided on page 605 of Shapiro (1964).*

Definition 5.3 *(Ljung-Box Test) The **Ljung-Box Test** is used to test whether the correlation coefficients differ significantly from zero.*

Hypothesis: *$H_0 : \rho(1) = \ldots = \rho(N) = 0$ vs. $H_1 : \rho(1) = \ldots = \rho(N) \neq 0$, for $N = 1, 2, \ldots$, where $\rho(l)$ is the autocorrelation of lag $l, l = 1, ..., N$ and for a weakly stationary time series $\{X_t, t = 1, ..., T\}$ is defined by*

$$\rho(l) = \frac{cov(X_t, X_{t-l})}{\sqrt{var(X_t)var(X_{t-l})}}.$$

Test statistic: $Q = T(T+2) \sum_{h=1}^{N} \frac{\hat{\rho}^2(h)}{T-h} \overset{under\ H_0}{\sim} \chi^2_N$

with T the sample size, $\hat{\rho}(h)$ the estimated autocorrelation at lag h and N the number of autocorrelations to test.

Rejection rule: *We reject the null hypothesis at level α if test statistic $Q > \chi^2_{\alpha,N}$, where $\chi^2_{\alpha,N}$ is the α-quantile of the χ^2_N distribution.*

Definition 5.4 *(LM ARCH Test) The **LM ARCH Test** studies the presence of ARCH effects in the data.*

Hypothesis: *$H_0 : \alpha_1 = ... = \alpha_p = \beta_1 = ... = \beta_q = 0$ vs. H_1 : At least one $\alpha_i \neq 0$ and $\beta_i \neq 0$, $i = 1, ..., p$, where $\alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q$ represent the constant parameters of the GARCH model.*

Test statistic: $Q = T \cdot R^2 \overset{under\ H_0}{\sim} \chi^2_p$,

where T is the number of observations and $R^2 := \mathbf{f^0}' z(z'z)^{-1} z' \mathbf{f^0} / \mathbf{f^0}' \mathbf{f^0}$ the squared multiple correlation, where

$$z = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ e_p^2 & e_{p+1}^2 & \ldots & e_{T-1}^2 \\ \vdots & \vdots & \ldots & \vdots \\ e_p^2 & e_{p+1}^2 \ldots & \ldots & e_{T-1}^2 \end{pmatrix},$$

e_t the residuals and $\mathbf{f^0}' = \left(\frac{e_{p+1}^2}{\hat{\sigma}^2} - 1, \frac{e_{p+2}^2}{\hat{\sigma}^2} - 1, ..., \frac{e_T^2}{\hat{\sigma}^2} - 1 \right)$, where $\hat{\sigma}^2 = \frac{1}{T-m} \sum_{t=m+1}^{T} e_t^2$ is the maximum likelihood estimate of σ^2 from the GARCH model, with $m = max(p, q)$, where p and q are the orders from the GARCH model.

Rejection rule: *We reject null hypothesis at level α if $Q > \chi^2_{\alpha,p}$,where $\chi^2_{\alpha,p}$ is the α-quantile of the χ^2_p distribution.*

6 Model Selection Criterion

In the next chapter, several models are presented from which the best model must be determined. The selection of the best model requires the use of model selection criteria such as Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC). The log-likelihood function, which is needed to calculate AIC and BIC, is defined below. The following definitions are based on the book by Georgii (2015), Fahrmeir et al. (2013) and Rossi (2018).

Definition 6.1 *(Log-likelihood function) Let $f(x;\boldsymbol{\theta})$ be a probability model with parameter space Θ.For a sample $x_1,\dots,x_d$ the* ***likelihood function*** *for a model $\mathcal{M}$ is given by the joint probability density function (pdf) of the random vector $(X_1,\dots,X_d)$ and is denoted by*

$$L_{\mathcal{M}}(\boldsymbol{\theta}) = f(x_1,\dots,x_n;\boldsymbol{\theta}).$$

For identically, independent random variables, the likelihood function is given by

$$L_{\mathcal{M}}(\boldsymbol{\theta}) = \prod_{i=1}^{n} f_i(x_i;\boldsymbol{\theta}),$$

where f_i is the probability density function of X_i. The ***log-likelihood function*** *is the logarithm of the likelihood function and is represented by the following equation*

$$l_{\mathcal{M}}(\boldsymbol{\theta}) = log(L_{\mathcal{M}}(\boldsymbol{\theta})). \tag{6.1}$$

Let's now define the model selection criteria of AIC and BIC. These specific metrics measure the goodness of fit of a model, where the model with the smallest AIC or BIC is considered the best fit.

Definition 6.2 *(Akaike's Information Criterion (AIC)) The* ***Akaike's Information Criterion*** *for a model $\mathcal{M}$ is defined by*

$$AIC_{\mathcal{M}} = -2l_{\mathcal{M}}(\hat{\boldsymbol{\theta}}) \pm 2k, \tag{6.2}$$

where k is the number of model parameters, $l_{\mathcal{M}}$ the log-likelihood function for the data at hand and a model $\mathcal{M}$ and $\hat{\boldsymbol{\theta}}$ maximum likelihood estimate of $\boldsymbol{\theta}$ based on $x_1,\dots,x_d$.

Definition 6.3 *(Bayesian Information Criterion (BIC)) The* ***Bayesian Information Criterion*** *is defined by*

$$BIC_{\mathcal{M}} = -2l_{\mathcal{M}}(\hat{\boldsymbol{\theta}}) \pm log(n)k, \tag{6.3}$$

where n is the sample size and k is the number of model parameters, $l_{\mathcal{M}}$ the log-likelihood function for the data at hand and a model $\mathcal{M}$ and $\hat{\boldsymbol{\theta}}$ maximum likelihood estimate of $\boldsymbol{\theta}$ based on $x_1,\dots,x_d$.

7 Copula

This section provides a definition of copula and presents different copula classes. Additionally, it explains important copula families and models, such as vine copula and factor copula models. Algorithms that are necessary for the simulation are also introduced and explained. The following definitions are based on the book by Czado (2019).

Definition 7.1 *(Copula) A multivariate distribution function C on the d-dimensional hypercube $[0,1]^d$ with uniformly distributed marginals, is a d-dimensional copula. The copula density for the absolutely continuous copula is dentoed by c and is determined by partial differentiation*

$$c(u_1, ..., u_d) = \frac{\partial^d}{\partial u_1 \cdots \partial u_d} C(u_1, ..., u_d) \quad \forall u \text{ in } [0,1]^d. \tag{7.1}$$

The following theorem allows to combine any marginal distribution with a copula to form new multivariate distribution functions.

Theorem 7.1 *(Sklar's theorem) The joint distribution function F of a d-dimensional random vector X can be expressed by*

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)) \tag{7.2}$$

with marginal distribution functions $F_i, i = 1, ..., d$ and copula C. The corresponding density is then given by

$$f(x_1, ..., x_d) = c(F_1(x_1), ..., F_d(x_d)) f_1(x_1) \cdots f_d(x_d), \tag{7.3}$$

where c is the copula density. Is the distribution function absolutely continuous, then the copula is unique. Considering the inverse, the copula and its density can be expressed as follows

$$C(u_1, ..., u_d) = F(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)), \tag{7.4}$$

$$c(u_1, ..., u_d) = \frac{f(F_1^{-1}(u_1), ..., F_d^{-1}(u_d))}{f_1(F_1^{-1}(u_1)) \cdots f_d(F_d^{-1}(u_d))}. \tag{7.5}$$

The following definition is needed if one wants to transform the given data into copula data. For example, original data can be transformed to copula data (u-scale) by applying PIT from the Definition 3.9. This data can in turn be converted to standard normal margins (z-scale). The transformed data can also be converted back to original data (x-scale). The scales described here are now defined in more detail.

Definition 7.2 *(Variable scales) The scales described above are defined by*

u-scale: *Copula scale (U_1, U_2) where $U_i := F_i(X_i)$ and copula density $c(u_1, u_2)$,*

x-scale: *Original scale (X_1, X_2) with density $f(x_1, x_2)$,*

z-scale: *Marginal normalized scale (Z_1, Z_2) where $Z_i := \Phi^{-1}(U_i) = \Phi^{-1}(F_i(X_i))$ for $i = 1, 2$ with density $g(z_1, z_2) = c(\Phi(z_1), \Phi(z_2))\phi(z_1)\phi(z_2)$.*

So far, the multivariate distribution has been presented in the terms of a copula and its marginal distribution. In the next part, the bivariate conditional distribution is presented using the marginal distributions and a copula.

Definition 7.3 *(Bivariate conditional distribution and density) The bivariate conditional distribution function and density is defined by*

$$F_{1|2}(x_1|x_2) = \frac{\partial}{\partial u_2} C_{12}(F_1(x_1), u_2) \mid_{u_2 = F_2(x_2)} \tag{7.6}$$
$$=: \frac{\partial}{\partial F_2(x_2)} C_{12}(F_1(x_1), F_2(x_2)),$$
$$f_{1|2}(x_1|x_2) = c_{12}(F_1(x_1), F_2(x_2)) f_2(x_2). \tag{7.7}$$

Copulas can be constructed by three classes, elliptical, Archimedian and the extreme-value copulas. The bivariate elliptical copulas are described first, the Archimedian copulas and finally extreme-value copulas are presented.

Bivariate elliptical copulas

The class of elliptical copulas includes the multivariate Gaussian copula, as well as the multivariate Student's t copula. In this paper, for the Gaussian copula, only the bivariate case is considered, while for the Student's t copula, the bivariate and multivariate cases are presented.

Definition 7.4 *(Bivariate Gaussian copula) Using a bivariate normal distribution with zero mean vector, unit variances and correlation ρ one gets the* ***bivariate Gaussian copula***

$$C(u_1, u_2; \rho) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho), \tag{7.8}$$

where Φ is the distribution function of standard normal distribution and $\Phi_2(\cdot, \cdot; \rho)$ is the normal distribution with zero means, unit variances and correlation ρ. The bivariate Gaussian copula density is given by

$$c(u_1, u_2; \rho) = \frac{1}{\phi(x_1)\phi(x_2)} \frac{1}{\sqrt{1-\rho^2}} exp \left\{ -\frac{\rho(x_1^2 + x_2^2) - 2\phi x_1 x_2}{2(1-\rho^2)} \right\}, \tag{7.9}$$

where x_1, x_2 is defined by $x_1 := \Phi^{-1}(u_1)$ and $x_2 := \Phi^{-1}(u_2)$.

Definition 7.5 *(Bivariate Student's t copula) For the* ***bivariate Student's t copula*** *we use a bivariate Student's t distribution with ν degrees of freedom, zero mean vectors and correlation ρ and get*

$$C(u_1, u_2; \nu, \rho) = \int_0^{u_1}\int_0^{u_2} \frac{t(T_\nu^{-1}(v_1), T_\nu^{-1}(v_2); \nu, \rho)}{t(T_\nu^{-1}(v_1))t(T_\nu^{-1}(v_2))} dv_1 dv_2 \tag{7.10}$$

$$= \int_{-\infty}^{b_1}\int_{-\infty}^{b_2} t(x_1, x_2; \nu, \rho) dx_1 dx_2, \tag{7.11}$$

where $b_1 := T_\nu^{-1}(u_1)$, $b_2 := T_\nu^{-1}(u_2)$ *and* $\frac{t(T_\nu^{-1}(v_1), T_\nu^{-1}(v_2); \nu, \rho)}{t(T_\nu^{-1}(v_1))t(T_\nu^{-1}(v_2))}$ *the bivariate Student's t copula density.*

Expanding the bivariate case to several variables results in obtaining the multivariate Student's t copula.

Definition 7.6 *(Multivariate Student's t copula) The multivariate Student's t copula is derived from the multivariate Student's t distribution and defined accordingly*

$$C(u_1, ..., u_d; \Sigma, \nu) = T_{R,\nu}(T_\nu^{-1}(u_1), ..., T_\nu^{-1}(u_d)) \tag{7.12}$$

where $T_{\Sigma,\nu}$ *the multivariate Student's t distribution with scale parameter matrix* $\Sigma \in [-1, 1]^{d\times d}$ *and the degrees of freedom* $\nu > 0$.

Next, the second class of copulas, known as Archimedean copulas, will be examined. First, Archimedean copulas will be defined and the families used in this thesis will be presented in a table.

Archimedean copulas

As before, bivariate Archimedean copulas are considered only.

Definition 7.7 *(Bivariate Archimedean copulas) Let* Ω *be the set of all continuous, strictly monotone decreasing, and convex functions* $\varphi : I \to [0, \infty]$ *with* $\varphi(1) = 0$. *Let* $\varphi \in \Omega$, *then*

$$C(u_1, u_2) = \varphi^{[-1]}(\varphi(u_1) + \varphi(u_2)) \tag{7.13}$$

is a copula with generator φ *and pseudo-inverse* $\varphi^{[-1]}$. *The generator* φ *is called strict if* $\varphi(0) = \infty$ *and the pseudo-inverse of* φ *is defined by* $\varphi^{[-1]} : [0, \infty] \to [0, 1]$ *with*

$$\varphi^{[-1]}(t) := \begin{cases} \varphi^{-1}(t) & t \in [0, \varphi(0)] \\ 0 & t \in [\varphi(0), \infty] \end{cases}$$

The most important Archimedean copulas are now summarized in Table 7.1.
Finally, the extreme-value copula constitutes the last class of copulas.

Table 7.1 Bivariate Archimedean copulas

Copula family	$C(u_1,u_2)$	Parameters
Clayton	$(u_1^{-\delta}+u_2^{-\delta}-1)^{-\frac{1}{\delta}}$	$0<\delta<\infty$
Gumbel	$exp(-\left[(-ln(u_1))^{\delta}+(-ln(u_2))^{\delta}\right]^{\frac{1}{\delta}})$	$\delta \geq 1$
Frank	$-\frac{1}{\delta}ln\left(\frac{1}{1-e^{-\delta}}\left((1-e^{-\delta})-(1-e^{-\delta u_1})(1-e^{-\delta u_2})\right)\right)$	$\mathbb{R}\backslash\{0\}$
Joe	$1-\left((1-u_1)^{\delta}+(1-u_2)^{\delta}-(1-u_1)^{\delta}(1-u_2)^{\delta}\right)^{\frac{1}{\delta}}$	$\delta \geq 1$

Extreme-value copulas

In this part, the behavior of extreme events and their dependence is studied. As before, we will only consider the bivariate case.

Definition 7.8 *(Bivariate extreme-value copula) A bivariate copula C is called an extreme-value copula if there is a bivariate copula C_X such that for $n \to \infty$ we have*

$$\left(C_X(u_1^{\frac{1}{n}},u_2^{\frac{1}{n}})\right)^n \to C(u_1,u_2) \quad \forall (u_1,u_2) \in [0,1]^2, \tag{7.14}$$

where C_X is in the domain of attraction of copula C.

Furthermore, the extreme-value copula can only be defined for the bivariate case in terms of Pickand's dependence function and its characterization is given in the following definition.

Definition 7.9 *(Characterization of bivariate extreme-value copulas in terms of the Pickands dependence function) A bivariate copula C is an extreme value copula if, and only if, the following holds*

$$C(u_1,u_2) = exp\left\{[ln(u_1)+ln(u_2)]\, A\left(\frac{ln(u_2)}{ln(u_1u_2)}\right)\right\}, \tag{7.15}$$

where $A:[0,1] \to \left\{\frac{1}{2},1\right\}$ is the Pickands dependence function. Function A is convex and satisfies $A \in [max\{1-t,t\},1]$ for all $t \in [0,1]$.

As only a single extreme-value copula family is utilised in this study, we shall only present this family: the Extended Joe family, also referred to as the BB8 family.

Definition 7.10 *(Extended Joe (BB8)) The representation of the Pickands dependence function of the Extended Joe copula is provided below.*

$$A(t) = \left[t^{\theta}+(1-t)^{\theta}-\left((1-t)^{-\theta\delta}+t^{-\theta\delta}\right)^{-\frac{1}{\delta}}\right]^{\frac{1}{\theta}}. \tag{7.16}$$

The bivariate Extenden Joe copula with the parameters $\theta \geq 1$ and $\delta > 0$ is given by equation (7.15), *where the Pickands dependence function is replaced by the equation* (7.16).

Before studying the next class of copulas, the vine copulas, we need to introduce some essential notation.

A pair copula decomposition of a multivariate distribution is when the copulas associated with conditional distributions can depend on the specific value of the underlying conditioning variable. For example $c_{13;2}(\cdot,\cdot;x_2)$ means that $c_{13;2}(\cdot,\cdot;x_2)$ depends on x_2. In the following, this dependence is often assumed to be disregarded. The following definition, the so-called simplifying assumption, takes this case into account. We specify the simplifying assumption for the three-dimensional case.

Definition 7.11 *(Simplifying assumption) The simplifying assumption of a three-dimensional pair copula is fulfilled if the following condition holds for any $x_2 \in \mathbb{R}$.*

$$c_{13;2}(u_1,u_2;x_2) = c_{13;2}(u_1,u_2) \text{ for } u_1,u_2 \in [0,1],$$

where $c_{13;2}(\cdot,\cdot;x_2)$ depends on x_2.

Furthermore, this section presents in detail the construction of joint parametric density as well as parametric copulas and copulas associated with bivariate conditional distributions.

Definition 7.12 *(Pair copula construction of a joint parametric density) A three-dimensional density with a parameter vector $\Theta = (\theta_{12},\theta_{23},\theta_{12;3})$ can be specified by using a parametric pair copula construction and is defined by*

$$\begin{aligned} f(x_1,x_2,x_3;\Theta) = {} & c_{13;2}(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2);\theta_{13;2}) \times c_{23}(F_2(x_2),F_3(x_3);\theta_{23}) \\ & \times c_{12}(F_1(x_1),F_2(x_2);\theta_{12}) f_3(x_3) f_2(x_2) f_1(x_1), \end{aligned} \tag{7.17}$$

where $c_{13;2}(\cdot,\cdot;\theta_{13;2}), c_{12}(\cdot,\cdot;\theta_{12}), c_{23}(\cdot,\cdot;\theta_{23})$ are arbitrary parametric bivariate copula densities.

Definition 7.13 *(Pair copula construction of a parametric copula) A three-dimensional parametric copula family with a parameter vector $\theta = (\theta_{12},\theta_{23},\theta_{12;3})$ can be defined in the following way*

$$c(u_1,u_2,u_3;\theta) = c_{13;2}(C_{1|2}(u_1|u_2),C_{3|2}(u_3|u_2);\theta_{13;2}) \times c_{23}(u_2,u_3;\theta_{23}) \times c_{12}(u_1,u_2;\theta_{12}) \tag{7.18}$$

with conditional distribution functions of U_1 given $U_2 = u_2$ denoted by $C_{1|2}(\cdot|u_2)$ and U_3 given $U_2 = u_2$ as indicated by $C_{3|2}(\cdot|u_2)$.

Definition 7.14 *(Copulas associated with bivariate conditional distributions) Given a set of random variables $(X_1,\dots,X_d)$, and a set D containing indices $\{1,...,d\}$ excluding i and j, the copula associated with the bivariate conditional distribution for (X_i,X_j) given $X_D = x_D$ is denoted by $C_{ij;D}(\cdot,\cdot;x_D)$.*
The conditional distribution function of (U_i,U_j) given $U_D = u_D$ is denoted by $C_{ij|D}(\cdot,\cdot;u_D)$ with bivariate density function $c_{ij|D}(\cdot,\cdot;u_D)$.
For indices i,j, where $i<j$ and $D := \{i_1,i_2,...,i_k\}$ with $i_1 < ... < i_k$ the following abbreviation is used

$$c_{i,j;D} := c_{i,j;D}(F_{i|D}(x_i|x_D),F_{j|D}(x_j|x_D);x_D).$$

Now that important definitions have been introduced, in the next section we will have a closer look at the classes of vine copulas.

7.1 Vine copula

The vine copula decomposes a multivariate density into products of conditional densities. A set $(X_1, \ldots, X_d)$ of random variables with joint distribution $F_{1,\ldots,d}$ and density $f_{1,\ldots,d}$. The decomposition is then given by

$$\begin{aligned} f_{1,\ldots,d}(x_1, \ldots, x_d) &= f_{d|1,\ldots,d-1}(x_d|x_1, \ldots, x_{d-1}) f_{1,\ldots,d-1}(x_1, \ldots, x_{d-1}) \\ &= \left[\prod_{t=2}^{d} f_{t|1,\ldots,t-1}(x_t|x_1, ..., x_{t-1}) \right] \cdot f_1(x_1). \end{aligned} \tag{7.19}$$

The joint density for the case $d = 3$ is now described in detail. Assuming the presence of random variables X_1, X_2 and X_3, the joint density is expressed as follows

$$f_{1,2,3}(x_1, x_2, x_3) = f_{3|12}(x_3|x_1, x_2) f_{2|1}(x_2|x_1) f_1(x_1). \tag{7.20}$$

To ascertain the component $f_{3|12}(x_3|x_1, x_2)$, the bivariate conditional density $f_{13|2}(x_1, x_3|x_2)$ must be taken into account. This density has $F_{1|2}(x_1|x_2) f_{1|2}(x_1|x_2)$ and $F_{3|2}(x_3|x_2) f_{3|2}(x_3|x_2)$ as marginal distributions with corresponding copula density $c_{13;2}(x_1, x_3; x_2)$, which belongs to the conditional distribution of (X_1, X_3) given $X_2 = x_2$. The conditional density of (X_1, X_3) given $X_2 = x_2$, can be determined with the following equation

$$f_{13|2}(x_1, x_3|x_2) = c_{13;2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2); x_2) f_{1|2}(x_1|x_2) f_{3|2}(x_3|x_2).$$

Finally, we can provide the first component from the Equation (7.20).

$$f_{3|12}(x_3|x_1, x_2) = c_{13;2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2); x_2) f_{3|2}(x_3|x_2)$$

The conditional density for X_2 given $X_1 = x_1$ and X_3 given $X_2 = x_2$, is then given by

$$\begin{aligned} f_{2|1}(x_2|x_1) &= c_{12}(F_1(x_1), F_2(x_2)) f_2(x_2) \\ f_{3|2}(x_3|x_2) &= c_{23}(F_2(x_2), F_3(x_3)) f_3(x_3). \end{aligned}$$

Equation (7.20) can now be completed and is represented as follows

$$\begin{aligned} f(x_1, x_2, x_3) = {} & c_{13;2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2); x_2) \times c_{23}(F_2(x_2), F_3(x_3)) \\ & \times c_{12}(F_1(x_1), F_2(x_2)) f_3(x_3) f_2(x_2) f_1(x_1). \end{aligned}$$

Now that we have examined the details of the case for d = 3, we will introduce three classes of decompositions, namely regular, canonical and drawable vine copulas.

Before introducing regular vine copulas, we provide theoretical graph background definitions and the definition of regular vine (R-vine) distribution. Subsequently, we present the definition of regular vine copula.

Definition 7.15 *(Graph theoretic background)*

1. ***Graph*** *is a pair* $G = (N, E)$ *of sets such that* $E \subseteq \{\{x, y\} : x, y \in N\}$
2. *Elements of E are referred to as* ***edges*** *of the graph G, while those of N are* ***nodes****.*
3. *The number of neighbors of a node* $v \in N$ *is the* ***degree*** *of v, denoted by* $d(v)$*.*

Definition 7.16 *(Regular vine tree sequence) The set of trees* $V = (T_1, \ldots, T_{d-1})$ *is a regular vine tree sequence on d elements if the following conditions are satisfied:*

1. $\forall j = 1, \ldots, d-1$*, a tree* $T_j = (N_j, E_j)$ *is connected.*
2. *A tree* T_1 *has a node set* $N_1 = (1, \ldots, d)$ *and a edge set* E_1*.*
3. T_j *for* $j \geq 2$ *is a tree with node set* $N_j = E_{j-1}$ *and edge set* E_j*.*
4. *For* $j = 2, \ldots, d-1$ *and* $\{a, b\} \in E_j$ *it must hold that* $|a \cap b| = 1$*.*

An example of an R-vine tree is displayed in Figure 7.1. The graph contains five trees. Considering Tree 1, the nodes are represented by $N = \{1, 2, 3, 4, 5, 6\}$, and edges by $E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}, \{4, 6\}\}$. The degree for node 4 in Tree 1 is indicated by $d(4) = 3$. The R-vine density of the R-vine tree structure in Figure 7.1 is given by

$$\begin{aligned} f_{123456}(x_1, x_2, x_3, x_4, x_5, x_6) =& c_{16;2345} \cdot c_{26;345} \cdot c_{15;234} \cdot c_{36;45} \cdot c_{25;34} \cdot c_{14;23} \cdot c_{56;4} \cdot c_{35;4} \cdot c_{24;3} \\ & \cdot c_{13;2} \cdot c_{54} \cdot c_{46} \cdot c_{34} \cdot c_{23} \cdot c_{12} \cdot f_6 \cdot f_5 \cdot f_4 \cdot f_3 \cdot f_2 \cdot f_1, \end{aligned}$$

where $f_{123456}(x_1, x_2, x_3, x_4, x_5, x_6)$ is the joint density, $f_j(x_j), j = 1, ..., 6$ the marginal densities and $c_{C_{e,a}, C_{e,b}; D_e}$ the corresponding pair copula densities.

Definition 7.17 *(Complete union and conditioned sets) The definition of the complete union of an edge e, denoted by the notation* A_e*, is as follows*

$$A_e = \{j \in N_1 | \exists e_1 \in E_1, \ldots, e_{i-1} \in E_{i-1} \text{ such that } j \in e_1 \in \ldots \in e_{i-1} \in e\}$$

for any edge $e \in E_i$*.*

The conditioning set D_e *of an edge* $e = \{a, b\}$ *defined by*

$$D_e := A_a \cap A_b.$$

Conditioning set $C_{e,a} := A_a \backslash D_e, C_{e,b} := A_b \backslash D_e$ *and* $C_e := C_{e,a} \cup C_{e,b}$ *are given. Edge* $e = (C_{e,a}, C_{e,b}; D_e)$ *is abbreviated by* $e = (e_a, e_b; D_e)$*.*

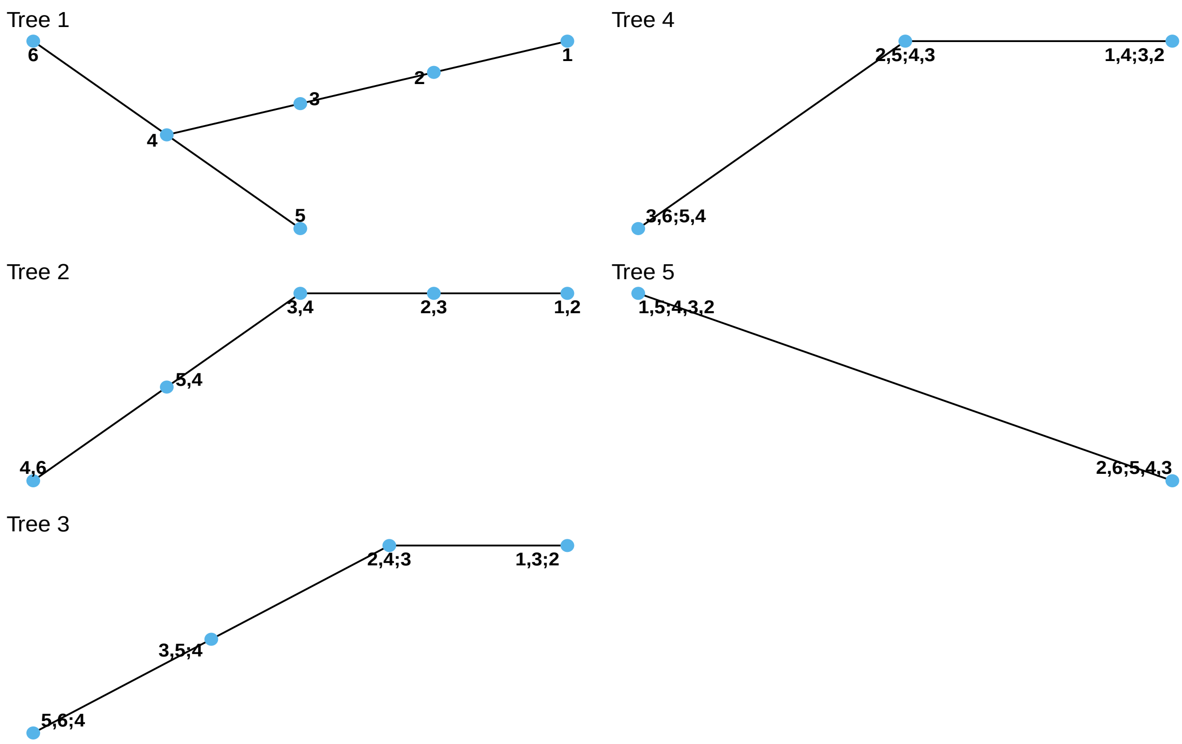

Figure 7.1 R-vine tree including the yield curve for maturities of 1, 5, 10, 15 and 20 year(s), alongside the inflation rate.

Definition 7.18 *(Regular vine distribution) If one can specify a triplet $(\mathcal{F}, V, \mathcal{B})$ so that the conditions below are fulfilled, then the joint distribution F for the d-dimensional random vector $X = (X_1, \ldots, X_d)$ has a regular vine distribution.*

1. *$\mathcal{F} = (F_1, \ldots, F_d)$ a vector of marginal distribution functions of the random variable $X_i, i = 1, \ldots, d$.*
2. *V is a R-vine tree sequence with d elements.*
3. *A set $\mathcal{B} = \{C_e | e \in E_i, i = 1, \ldots, d-1\}$ with C_e being a symmetric bivariate copula and E_i is the set of edges of tree T_i in the R-vine tree sequence V for $i = 1, \ldots, d-1$.*
4. *For each $e \in E_i, i = 1, \ldots, d-1$ and $e = \{a, b\}$, C_e is the copula associated with the conditional distribution of $X_{C_{e,a}}$ and $X_{C_{e,b}}$ given $X_{D_e} = x_{D_e}$.*

Definition 7.19 *(Existence regular vine distribution) The probability density function of a unique d-dimensional distribution F can be determined if the set $(\mathcal{F}, V, \mathcal{B})$ satisfies condition (1) - (3) from Definition 7.18. The density can then be obtained.*

$$f_{1,\ldots,d}(x_1, \ldots, x_d) = f_1(x_1) \cdots f_d(x_d) \cdot \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{C_{e,a}C_{e,b};D_e}(F_{C_{e,a}|D_e}(x_{C_{e,a}}|x_{D_e}), F_{C_{e,b}|D_e}(x_{C_{e,b}}|x_{D_e})), \tag{7.21}$$

where $e \in E_i, i = 1, \ldots, d-1$ with $e = \{a, b\}$. The distribution function of $X_{C_{e,a}}$ and $X_{C_{e,b}}$ given $X_{D_e} = x_{D_e}$ is defined by

$$F_{C_{e,a}C_{e,b}|D_e}(x_{C_{e,a}}, x_{C_{e,b}}|x_{D_e}) = C_e(F_{C_{e,a}|D_e}(x_{C_{e,a}}|x_{D_e}), F_{C_{e,b}|D_e}(x_{C_{e,b}}|x_{D_e})).$$

Definition 7.20 *(Regular vine (R-vine) copula) A regular vine copula is a regular vine distribution, where all margins are uniformly distributed on $[0, 1]$.*

Now consider the canonical vine (C-vine) copula, which is a subclass of the regular vine copula. Compared to the regular vine copula, the C-vine copula has a star shape. The definition of the tree sequence and the density of the canonical vine copula are described below.

Definition 7.21 *(C-vine tree sequence) A regular vine tree sequence $V = (T_1, ..., T_{d-1})$ is a canonical vine tree sequence if in each Tree T_i there exist one node $n \in N_i$ such that $|\{e \in E_i | n \in e\}| = d - i$. A node of this type is also referred to as the root node of the tree T_i.*

Definition 7.22 *(Canonical vine (C-vine) density) The joint density $f_{1,\ldots,d}$ is decomposed by*

$$f_{1,\ldots,d}(x_1, \ldots, x_d) = \left[\prod_{j=1}^{d-1}\prod_{i=1}^{d-j} c_{j,(j+i);1,\ldots,j-1}\right] \cdot \left[\prod_{k=1}^{d} f_k(x_k)\right] \tag{7.22}$$

and we call the decomposition, a canonical vine distribution.

A C-vine tree including the yield curve with maturities of 1, 5, 10, 15 and 20 years and the inflation rate is shown in Figure 7.2. Considering Figure 7.2, in Tree 1 the root node is represented by node 6. The cardinality of the edges in this tree is then equal to 5, since it is a C-vine on $d = 6$ elements, where i is equal to 1 for the first tree.

The last subclass of regular vine copulas, the drawable vine copula, will now be defined. First the tree structure will be outlined, followed by the density of the drawable vine (D-vine) copula. Compared to the C-vine tree structure, the drawable vine copula has a path as a tree structure. Mathematically this is defined as follows

Definition 7.23 *(D-vine tree sequence) A regular vine tree sequence $V = (T_1, ..., T_{d-1})$ is a drawable vine tree sequence if for each node $n \in N_i$ we have $|\{e \in E_i | n \in e\}| \leq 2$.*

The density of the D-vine copula is hence defined.

Definition 7.24 *(Drawable vine (D-vine) density) A drawable vine is given by*

$$f_{1,\ldots,d}(x_1, \ldots, x_d) = \left[\prod_{j=1}^{d-1}\prod_{i=1}^{d-j} c_{i,(i+j);(i+1),\ldots,(i+j-1)}\right] \cdot \left[\prod_{k=1}^{d} f_k(x_k)\right], \tag{7.23}$$

where the joint density $f_{1,\ldots,d}$ is decomposed into products of marginal densities f_k and pair-copula densities $c_{i,j|D}(\cdot, \cdot; x_D)$, where D is a set of indices from $\{1, \ldots, d\}$.

Figure 7.3 displays an example of a D-vine tree structure. Focusing on Tree 1, we can see that the cardinality of the edge is less than or equal to two. If we observe the edge (1, 2), the cardinality in this case is equal to one. Similarly, upon analysing the edge (3, 4), we obtain a cardinality equal to 2.

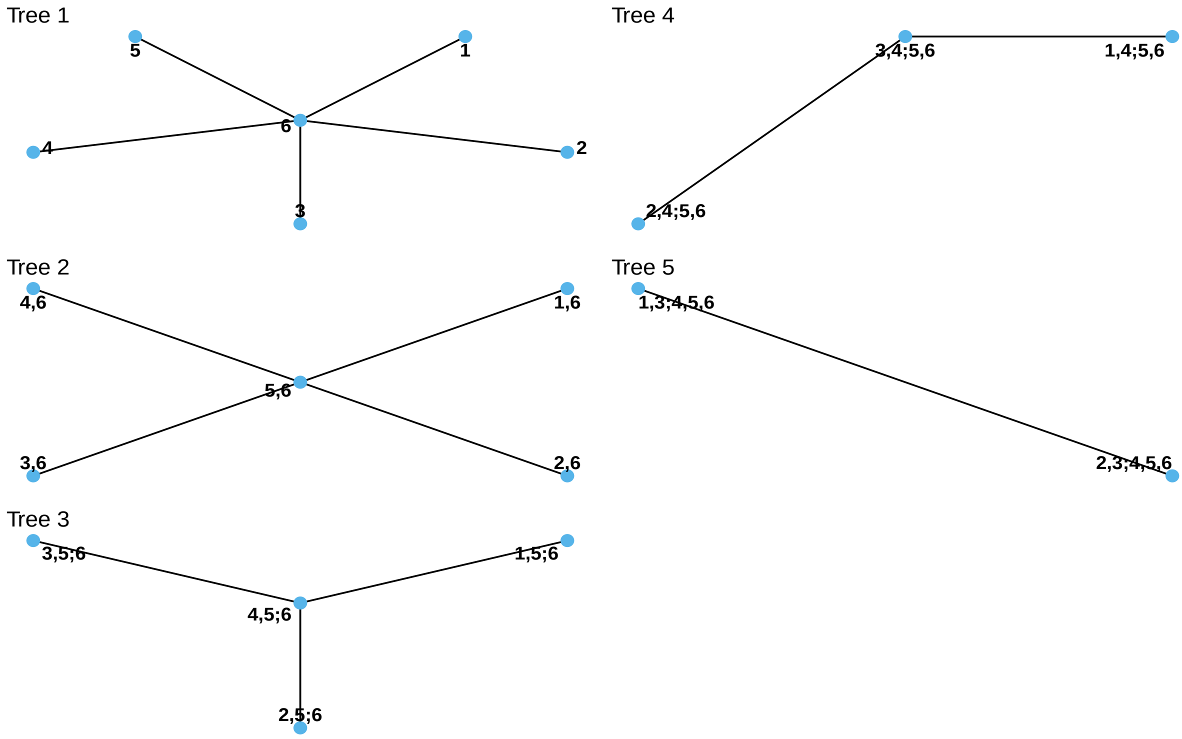

Figure 7.2 Graphical representation of the C-vine including the yield curve for maturities of 1, 5, 10, 15 and 20 year(s), alongside the inflation rate.

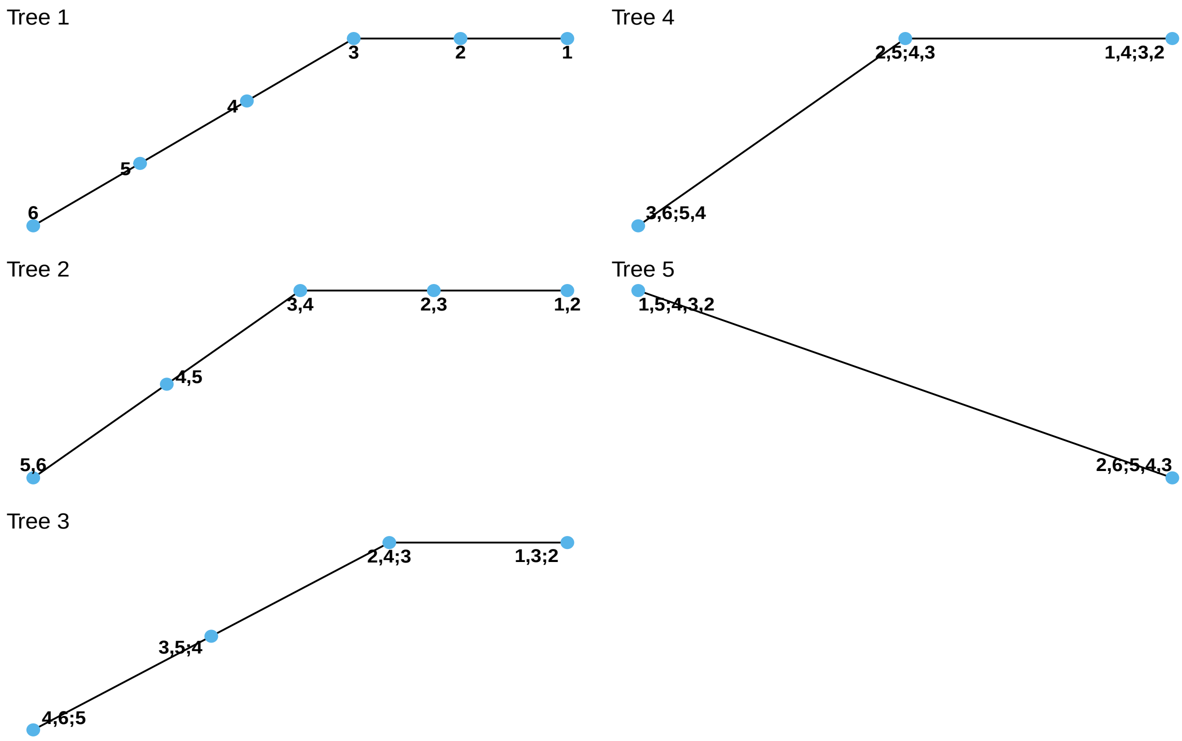

Figure 7.3 Graphical representation of the D-vine including the yield curve for maturities of 1, 5, 10, 15 and 20 year(s), alongside the inflation rate.

The simplifying assumption applies not only for general pair copulas, but also for the C- and D-vine copulas. This assumption holds for C- and D-vine copulas if the following definition is met.

Definition 7.25 *(Simplifying assumption for C- and D-vines) Simplifying assumption for C- and D-vines is satisfied if*

$$c_{ij,D}(F_{i|D}(x_i|x_D), F_{j|D}(x_j|x_D); x_D) = c_{ij,D}(F_{i|D}(x_i|x_D), F_{j|D}(x_j|x_D))$$

holds for all x_D and i, j and D containing indices $\{1, ..., d\}$ are chosen to occur in (7.22) *and* (7.23) *the corresponding C- and D-vine distribution is then simplified.*

In addition to the vines, truncated vines are also a method for consideration. For the truncated vine models, the tree structure, pair copulas and parameters remain the same for the first n trees; for the remaining trees T_{n+1} to T_{d-1}, the independence copula is chosen for each pair copula. Let now consider the mathematical definition of the truncated R-vine copula, based on the paper by Brechmann and Joe (2015) and Brechmann et al. (2012).

Definition 7.26 *(Truncated R-vine copula) Let a random vector with uniform margins be given by $\mathbf{U} = (U_1, ..., U_d)'$ and the truncation level by $n \in \{1, ..., d\}$. The random vector $\mathbf{U}$ is then distributed according to the d-dimensional n-truncated R-vine copula $C_{C_{e,a},C_{e,b};D_e}$ if $C_{C_{e,a},C_{e,b};D_e}$ is a d-dimensional R-vine with*

$$C_{C_{e,a},C_{e,b};D_e} = \Pi \quad \forall e \in E_i, i = n+1, ..., d-1,$$

where Π represents the bivariate independence copula.

Another copula model, known as the factor copula model, is an alternative to the truncated vine copula. In the next section, we will define the one- and two-factor copula models as special cases.

7.2 Factor copula

In this section, we will examine additional copula models, specifically, factor copula models. The **factor copula model** links observable and unobservable (latent) variables. The dependence of observable variables is explained by a few unobservable variables. The advantage of the model is that it is easy to interpret and it can deal with tail dependencies and asymmetries, which is not the case with normal copula models.

On the basis of the paper by Krupskii and Joe (2013) and Krupskii and Joe (2015), we will focus in the following on the definition of the one- and two-factor copula model.

Definition 7.27 *(One-factor copula model) The factor copula model with one latent variable $p = 1$ is defined by*

$$C(u_1, \dots, u_d) = \int_0^1 \prod_{j=1}^d F_{j|V}(u_j|v)dv = \int_0^1 \prod_{j=1}^d C_{j|V}(u_j|v)dv, \tag{7.24}$$

where $U_j, j = 1, ..., d$ and the latent random variable V are $U(0, 1)$ distributed with $U_j|V$ being independent of $U_{j'}|V$ for $j \neq j'$. $C_{j|V}$ represents the partial derivative of the joint cumulative distribution function (cdf) copula of (U_j, V), mathematically given by $F_{j|V} = C_{j|V} = \frac{\partial C_{j,V}(u_j,v)}{\partial v}$. The joint cdf is denoted by $C_{j,V}$. The differentiation of equation (7.24) leads to the copula density

$$c(u_1, \dots, u_d) = \frac{\partial^d C(u_1, \dots, u_d)}{\partial u_1 \cdots \partial u_d} = \int_0^1 \prod_{j=1}^d c_{j,V}(u_j, v)dv,$$

where $c_{j,v}(u_j, v)$ is the joint copula density of (U_j, V).

If an additional latent parameter is included in the factor copula model, it results in the formation of the two-factor copula model.

Definition 7.28 *(Two-factor copula model) The copula model, with two latent variables $p = 2$, is defined by*

$$C(u_1, \dots, u_d) = \int_0^1 \int_0^1 \prod_{j=1}^d F_{j|V_1,V_2}(u_j|v_1, v_2)dv_1 dv_2 = \int_0^1 \int_0^1 \prod_{j=1}^d C_{j|V_2;V_1}(C_{j|V_1}(u_j|v_1)|v_2)dv_1 dv_2. \tag{7.25}$$

where $C_{j|V_1}$ the partial derivative of the joint cdf copula of (U_j, V_1), $C_{j,V_2;V_1}$ the copula for $F_{j|V_1} = F_{U_j|V_1}$ and $F_{V_2|V_1}$. For $C_{j,V_2;V_1}$ we use the simplifying assumption , i. e. $C_{j,V_2;V_1}(u_j, v_2; v_1) = C_{j,V_2;V_1}(u_j, v_2)$. Since $F_{V_2|V_1}$ is the $U(0, 1)$ cdf, one can assume that V_2 is independent of V_1 which implies

$$\begin{aligned}
F_{j|V_1,V_2}(u|v_1, v_2) &= P(U_j \leq u|V_1 = v_1, V_2 = v_2) \\
&= \frac{\partial}{\partial v_2} P(U_j \leq u, V_2 \leq v_2|V_1 = v_1) \\
&= \frac{\partial}{\partial v_2} C_{jV_2;V_1}(C_{j|V_1}(u|v_1), v_2) \\
&= C_{j|V_2;V_1}(C_{j|V_1}(u|v_1)|v_2).
\end{aligned}$$

The density function for the two-factor copula model is defined by

$$c(u_1,\ldots,u_d) = \int_0^1 \int_0^1 \prod_{j=1}^{d} c_{j,V_2;V_1}(C_{j|V_1}(u_j|v_1), v_2) \cdot c_{j,V_1}(u_j, v_1) dv_1 dv_2.$$

We now examine the special case of the factor copula model. We provide the definitions of the Gaussian one-factor and two-factor copulas.

Definition 7.29 *(Gaussian one-factor copula model) In the **Gaussian one-factor copula model**, the bivariate normal copula with correlation α_{j1} for $j = 1, \ldots, d$ is referred to as $C_{j,V}$ and is defined as such*

$$C_{j,V}(u,v) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \alpha_{j1}),$$

where Φ represents the standard normal cumulative distribution function and $\Phi_2(\cdot;\rho)$ denotes the bivariate normal cumulative distribution function with correlation ρ. The conditional distribution for U_j given V, can be expressed as follows

$$F_{j|V}(u|v) = \Phi\left(\frac{\Phi^{-1}(u) - \alpha_{j1}\Phi^{-1}(v)}{\sqrt{1-\alpha_{j1}^2}}\right).$$

Extending the Gaussian one-factor copula model with a latent variable yields the Gaussian two-factor copula model.

Definition 7.30 *(Gaussian two-factor copula model) Let C_{j,V_1} and $C_{j,V_2;V_1}$ be the bivariate copulas with correlation α_{j1} and $\gamma_j = \frac{\alpha_{j2}}{\sqrt{1-\alpha_{j1}^2}}$ for $j = 1, \ldots, d$, where α_{j2} is the correlation of $Z_j = \Phi(U_j)$ and $W_2 = \Phi(V_2)$ such that the independence of V_1, V_2 implies that γ_j is the partial correlation of Z_j and W_2 given $W_1 = \Phi(V_1)$. The conditional distribution of the bivariate copula is given by*

$$C_{j|V_2;V_1}(C_{j|V_1}(u|v_1)|v_2) = \Phi\left(\frac{\Phi^{-1}(u) - \alpha_{j1}\Phi^{-1}(v_1) - \gamma_j(1-\alpha_{j1}^2)^{\frac{1}{2}}\Phi^{-1}(v_2)}{\sqrt{(1-\alpha_{j1}^2)(1-\gamma_j^2)}}\right).$$

The factor copula model can be extended for $p > 2$. This extension represents the p-factor multivariate normal distribution, which possesses a correlation matrix Σ having a p-factor structure. This is described in the next section and is referred to as the Gaussian p-factor model.

7.3 Gaussian p-factor

In this section, we present the Gaussian factor model, a multivariate normal model that exhibits a structured correlation matrix $\Sigma = (\rho_{ij})_{i,j=1,\ldots,d}$ with $O(d)$ dependence parameters. The definition of the Gaussian factor model is presented below, referencing the book by Everitt and Hothorn (2011) and Akaike (1987).

Definition 7.31 *(p-factor models) Let $\mathbf{X} = (X_1, X_2, \ldots, X_d)$ be a set of random variables which is assumed to be linked to p unobserved latent random variables $f_1, \ldots, f_p$, where $p < d$ by a regression model of*

$$\begin{aligned} X_1 &= \lambda_{11} f_1 + \lambda_{12} f_2 + \ldots + \lambda_{1p} f_p + \varepsilon_1, \\ X_2 &= \lambda_{21} f_1 + \lambda_{22} f_2 + \ldots + \lambda_{2p} f_p + \varepsilon_2, \\ &\vdots \\ X_d &= \lambda_{d1} f_1 + \lambda_{d2} f_2 + \ldots + \lambda_{dp} f_p + \varepsilon_d. \end{aligned} \tag{7.26}$$

Error terms are given by $\varepsilon_1, \ldots, \varepsilon_d$ with mean zero and variance $\psi_i, i = 1, ..., d$. $\varepsilon_1, \ldots, \varepsilon_d$ are uncorrelated with each other and the random factors $f_1, \ldots, f_p$. Since the factors $f_1, \ldots, f_p$ are unobserved, their location can be fixed and scales are arbitrary and one assumes that they occur with mean zero and standard deviation one. Regression coefficients λ_js are alternatively referred to as factor loadings and are the correlations of the observed variables and the factors. It is also assumed that the latent factors are uncorrelated with one another.

Considering the assumption that the latent factors are independent of one another and of the error terms, the variance of X_i is given by

$$\sigma_i^2 = \sum_{j=1}^{p} \lambda_{ij}^2 + \psi_i.$$

The covariance matrix of the variables is given by

$$\Sigma = \Lambda \Lambda^T + \Psi$$

with $\Psi = diag(\psi_i)$ and Λ matrix with the factor loadings.

If $p = 1$, the factor model includes a single latent factor. This is known as the Gaussian one-factor model and is defined by

$$\begin{aligned} X_1 &= \lambda_{11} f_1 + \varepsilon_1, \\ X_2 &= \lambda_{21} f_1 + \varepsilon_2, \\ &\vdots \\ X_d &= \lambda_{d1} f_1 + \varepsilon_d, \end{aligned}$$

where the latent factor f_1 is standard normally distributed and the error terms $\varepsilon_i, i = 1, ..., d$ are normally distributed with mean zero and variance ψ_i. If one include another latent factor, we get the Gaussian two-factor model, which is given by

$$\begin{aligned} X_1 &= \lambda_{11} f_1 + \lambda_{12} f_2 + \varepsilon_1, \\ X_2 &= \lambda_{21} f_1 + \lambda_{22} f_2 + \varepsilon_2, \\ &\vdots \\ X_d &= \lambda_{d1} f_1 + \lambda_{d2} f_2 + \varepsilon_d. \end{aligned}$$

Before using the above model, it is necessary to estimate the model parameters, which includes the factor loadings Λ, the entries in the covariance matrix Σ and the variances Ψ. Usually, the covariance matrix Σ is estimated by the sample covariance matrix S from $(\mathbf{X_1}, ..., \mathbf{X_d})$. The remaining parameters to be estimated are then only the factor loadings and the variances.

Given the estimations for the factor loadings, the entries ψ_i can be computed by

$$\hat{\psi}_i = s_i^2 - \sum_{j=1}^{p} \hat{\lambda}_{ij}^2, i = 1, ..., d,$$

where s_i^2 are the sample variances. Now two methods are available to estimate these parameters: principal factor analysis and maximum likelihood factor analysis. In the subsequent section, we will elaborate on the maximum likelihood factor analysis.

Maximum likelihood factor analysis

In the context of maximum likelihood factor analysis, the initial assumption is that the data comply with a multivariate normal distribution. Under this premise, and given that the factor analysis model remains valid, the resulting likelihood function L can be expressed as $-\frac{1}{2}nF$ plus a function of the observations only. Here, F is defined as follows:

$$F = \ln|\Lambda\Lambda^T + \Psi| + \text{trace}(S|\Lambda\Lambda^T + \Psi|^{-1}) - \ln|S| - d, \tag{7.27}$$

where S is the estimator of the covariance matrix Σ. If S equals $\Lambda\Lambda^T + \Psi$, then function F is zero. The estimators for the factor loadings and variances are determined through minimization of the function F with respect to these parameters.

Model selection criterion

Additionally, model selection criteria should be considered in order to identify the most optimal model. The model selection criteria are Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC). Using function $\frac{1}{2}nF$ plus a function of the observations as described above, the log likelihood l at the MLE estimate is given by

$$\hat{l} = l(\hat{\Lambda}, \hat{\Psi}; X) = -\frac{1}{2}n \left[\ln|\hat{\Lambda}\hat{\Lambda}^T + \hat{\Psi}| + \text{trace}(S|\hat{\Lambda}\hat{\Lambda}^T + \hat{\Psi}|^{-1}) \right].$$

We can now apply the log-likelihood function to determine the AIC and BIC.

$$\begin{aligned} AIC &= -2\hat{l} + [2d(p+1) - p(p-1)] \\ BIC &= -2\hat{l} + \left[d(p+1) - \frac{p(p-1)}{2} \right] \end{aligned}$$

The best model is then the one with the lowest AIC and BIC.

7.4 Simulation from vine copula models

In this section, different algorithms for simulation from the copula models are presented.The algorithm for simulating from the R-, C-, and D-vine copula is outlined below. To begin with, it is necessary to describe the definition of h-functions. The algorithms and definitions presented in this section are based on the book by Czado (2019).

Definition 7.32 *(h-functions of bivariate parametric copulas) For a bivariate copula $C_{ij}(u_i, u_j; \theta_{ij})$ with parameter θ_{ij}, the h-functions are defined as follows:*

$$h_{i|j}(u_i|u_j; \theta_{ij}) := \frac{\partial}{\partial u_j} C_{ij}(u_i, u_j; \theta_{ij}) \tag{7.28}$$

$$h_{j|i}(u_j|u_i; \theta_{ij}) := \frac{\partial}{\partial u_i} C_{ij}(u_i, u_j; \theta_{ij}) \tag{7.29}$$

For the the parametric pair copula $C_{e_a,e_b;D_e}(w_1, w_2; \theta_{e_a,e_b;D_e})$ in a simplified regular vine corresponding to the edge $e_a, e_b; D_e$ the following notation is introduced

$$h_{e_a|e_b;D_e}(w_1|w_2; \theta_{e_a,e_b;D_e}) := \frac{\partial}{\partial w_2} C_{e_a,e_b;D_e}(w_1, w_2; \theta_{e_a,e_b;D_e}) \tag{7.30}$$

$$h_{e_b|e_a;D_e}(w_2|w_1; \theta_{e_a,e_b;D_e}) := \frac{\partial}{\partial w_1} C_{e_a,e_b;D_e}(w_1, w_2; \theta_{e_a,e_b;D_e}) \tag{7.31}$$

Before presenting the algorithms, necessary notations must be introduced. The vector $(i, ..., j)$ will be abbreviated as $i : j$, and $(u_i, ..., u_j)$ as $u_{i:j}$. Additionally, the copula parameter θ_{ij} is excluded in the notation, and $h_{i|j}(u_i|u_j; \theta_{ij})$ is substituted with $h_{i|j}(u_i|u_j)$.

Considering the simulation from a C-vine copula, the $d \times d$ - dimensional R-vine structure matrix M has the following structure:

$$M = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & \cdots \\ & 2 & 2 & 2 & 2 & \cdots \\ & & 3 & 3 & 3 & \cdots \\ & & & 4 & 4 & \cdots \\ & & & & 5 & \cdots \\ & & & & & \cdots \end{pmatrix}.$$

The univariate conditional distribution function $C_{i|1:i-k}(u_i|\mathbf{u}_{1:i-k})$ is provided by

$$C_{i|1:i-k}(u_i|\mathbf{u}_{1:i-k}) = \frac{\partial C_{i,i-k;1:i-k-1}(C_{i|1:i-k-1}(u_i|\mathbf{u}_{1:i-k-1}), C_{i-k|1:i-k-1}(u_{i-k}|\mathbf{u}_{1:i-k-1}))}{\partial C_{i-k|1:i-k-1}(u_{i-k}|\mathbf{u}_{1:i-k-1})}$$

for $i = 1, \ldots, d$ and $k = 1, \ldots, d$, the copula parameters are contained in a $d \times d$-dimensional strict upper triangular matrix, Θ, with the entries $\eta_{ik} = \theta_{ik;1:i-1}$ for $i < k = 2, \ldots, d$. Θ has the following representation

$$\Theta = \begin{pmatrix} - & \theta_{1,2} & \theta_{1,3} & \theta_{1,4} & \cdots \\ - & - & \theta_{2,3;1} & \theta_{2,4;1} & \cdots \\ - & - & - & \theta_{3,4;1,2} & \cdots \\ - & - & - & - & \cdots \end{pmatrix}.$$

The copula parameter matrix Θ is employed to compute the elements of the upper triangular $d \times d$-dimensional matrix V of conditional distribution functions with entries $\eta_{ik}, i \le k = 1, \dots, d$.

$$V = \begin{pmatrix} u_1 & u_2 & u_3 & u_4 & \cdots \\ & C_{2|1}(u_2|u_1) & C_{3|1}(u_3|u_1) & C_{4|1}(u_4|u_1) & \cdots \\ & & C_{3|1,2}(u_3|u_1,u_2) & C_{4|1,2}(u_4|u_1,u_2) & \cdots \\ & & & C_{4|1,2,3}(u_4|u_1,u_2,u_3) & \cdots \\ & & & & \cdots \end{pmatrix}$$

The procedure for simulating from the C-vine copula is described in Algorithm 1.

Algorithm 1 Sampling from a C-vine copula

1: **Inputs:**
   Θ is a strict upper triangular matrix of copula parameters with entries $\eta_{ki} = \theta_{ki;1:k-1}$ for $k < i, i = 1, \cdots, d$ and $k = 1, \cdots, d$ for a d-dimensional C-vine
2: Sample $w_i \overset{\text{i.i.d}}{\sim} U[0,1], i = 1, \dots, d$
3: $v_{1,1} := w_1$
4: **for** i = 2, ..., d **do**
5:     $v_{i,i} := w_i$
6:     **for** $k = i-1, ..., 1$ **do**
7:         $v_{k,i} := h^{-1}_{i|k;1:k-1}(v_{k+1,i}|v_{k,k}, \eta_{k,i})$
8:     **end for**
9: **end for**
10: **Return** $u_i := v_{1,i}, i = 1, \dots, d$

Let us consider the simulation from a D-vine copula. First, we present the R-vine structure matrix M.

$$M = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 & \cdots \\ & 2 & 1 & 2 & 3 & \cdots \\ & & 3 & 1 & 2 & \cdots \\ & & & 4 & 1 & \cdots \\ & & & & 5 & \cdots \\ & & & & & \cdots \end{pmatrix}$$

Then, we define the conditional distribution function $C_{i|k:i-1}(u_i|\mathbf{u}_{k:i-1})$

$$C_{i|k:i-1}(u_i|\mathbf{u}_{k:i-1}) = \frac{\partial C_{i,k;k+1:i-1}(C_{i|k+1:i-1}(u_i|\mathbf{u}_{k+1:i-1}), C_{k|k+1:i-1}(u_k|\mathbf{u}_{k+1:i-1}))}{\partial C_{k|k+1:i-1}(u_k|\mathbf{u}_{k+1:i-1})} \tag{7.32}$$

with $i = 3, \dots, d$ and $k = 2, \dots, i-1$, where $i > k$.
The second argument in Equation (7.32) must be determined, requiring an additional step. Similar to simulating from the C-vine copula, the $d \times d$ upper triangular matrix V is essential.

The definition of this matrix is

$$V = \begin{pmatrix} u_1 & u_2 & u_3 & u_4 & \cdots \\ & C_{2|1}(u_2|u_1) & C_{3|2}(u_3|u_2) & C_{4|3}(u_4|u_3) & \cdots \\ & & C_{3|2,1}(u_3|u_2,u_1) & C_{4|3,2}(u_4|u_3,u_2) & \cdots \\ & & & C_{4|3,2,1}(u_4|u_3,u_2,u_1) & \cdots \\ & & & & \cdots \end{pmatrix}.$$

As previously described, an additional upper triangular matrix V^2 must be incorporated. The matrix is specified by

$$V^2 = \begin{pmatrix} u_1 & u_2 & u_3 & u_4 & \cdots \\ & C_{1|2}(u_1|u_2) & C_{2|3}(u_2|u_3) & C_{3|4}(u_3|u_4) & \cdots \\ & & C_{1|2,3}(u_1|u_2,u_3) & C_{2|4,3}(u_2|u_4,u_3) & \cdots \\ & & & C_{1|4,3,2}(u_1|u_4,u_3,u_2) & \cdots \\ & & & & \cdots \end{pmatrix}.$$

The entries of matrices V and V^2 must also be identified in this context, which is determined with the copula parameter matrix Θ.

$$\Theta = \begin{pmatrix} - & \theta_{1,2} & \theta_{2,3} & \theta_{3,4} & \cdots \\ - & - & \theta_{3,1;2} & \theta_{4,2;3} & \cdots \\ - & - & - & \theta_{4,1;3,2} & \cdots \\ - & - & - & - & \cdots \end{pmatrix}.$$

The algorithm for simulating from the D-vine copula is explained in more detail in the Algorithm 2.

Algorithm 2 Sampling from a D-vine copula

```
Inputs:
    Θ is a strict upper triangular matrix of copula parameters with entries
    η_ki = θ_{ki;1:k−1} for k < i, i = 1, ··· , d and k = 1, ··· , d for a
    d-dimensional D-vine
Sample w_i ~(i.i.d) U[0,1], i = 1, ..., d
v_{1,1} := w_1, v^2_{1,1} := w_1
for i = 2, ..., d do
    v_{i,i} := w_i
    for k = i − 1, ..., 1 do
        v_{k,i} := h^{−1}_{i|i−k;i−k+1:i−1}(v_{k+1,i} | v^2_{k,i−1}, η_{k,i})
        if i < d then
            v^2_{k+1,i} := h_{i−k|i;i−k+1:i−1}(v^2_{k,i−1} | v_{k,i}, η_{k,i});
        end if
    end for
    v^2_{1,i} := v_{1,i}
end for
Return u_i := v_{1,i}, i = 1, ..., d
```

Finally, we will discuss the algorithm used to simulate the R-vine copula. The R-vine matrix M is defined as $M = (m_{i,j})_{i,j=1,\dots,d}$. The matrix Θ includes the copula parameters and is provided by

$$\Theta = \begin{pmatrix} - & \theta_{m_{1,2},2} & \theta_{m_{1,3},3} & \theta_{m_{1,4},4} & \cdots \\ - & - & \theta_{m_{2,3},3;m_{1,3}} & \theta_{m_{2,4},4;m_{1,4}} & \cdots \\ - & - & - & \theta_{m_{3,4},4;m_{2,4},m_{1,4}} & \cdots \\ - & - & - & - & \cdots \end{pmatrix}.$$

The conditional distribution function C is defined as follows:

$$C(u_{m_{i,i}}|u_{m_{k,i}}, u_{m_{k-1,i}}, \dots, u_{m_{1,i}}) = \frac{\partial C_{m_{i,i},m_{k,i};m_{k-1,i},\dots,m_{1,i}}(C(u_{m_{i,i}}|u_{m_{k-1,i}}, \dots, u_{m_{1,i}}), C(u_{m_{i,i}}|u_{m_{k,i}}, u_{m_{k-1,i}}, \dots, u_{m_{1,i}}))}{\partial C(u_{m_{k,i}}|u_{m_{k-1,i}}, \dots, u_{m_{1,i}})}.$$

where $i = 3, \dots, d$ and $k = 2, \dots, i-1$. The representations of matrices V and V^2 are as follows:

$$V = \begin{pmatrix} u_1 & u_2 & u_3 & u_4 & \cdots \\ & C(u_2|u_{m_{1,2}}) & C(u_3|u_{m_{1,3}}) & C(u_4|u_{m_{1,4}}) & \cdots \\ & & C(u_3|u_{m_{1,3}}, u_{m_{2,3}}) & C(u_4|u_{m_{1,4}}, u_{m_{2,4}}) & \cdots \\ & & & C(u_4|u_{m_{1,4}}, u_{m_{2,4}}, u_{m_{3,4}}) & \cdots \\ & & & & \cdots \end{pmatrix},$$

$$V^2 = \begin{pmatrix} u_1 & u_2 & u_3 & u_4 & \cdots \\ & C(u_{m_{1,2}}|u_2) & C(u_{m_{1,3}}|u_3) & C(u_{m_{1,4}}|u_4) & \cdots \\ & & C(u_{m_{2,3}}|u_{m_{1,3}}, u_3) & C(u_{m_{2,4}}|u_{m_{1,4}}, u_4) & \cdots \\ & & & C(u_{m_{3,4}}|u_{m_{1,4}}, u_{m_{2,4}}, u_4) & \cdots \\ & & & & \cdots \end{pmatrix}.$$

To simulate from the R-vine copula, another matrix $\mathcal{M} := (\tilde{m}_{k,i}), k \leq i$ needs to be introduced, which is defined as

$$\tilde{m}_{k,i} := \max \{m_{k,i}, m_{k-1,i}, \dots, m_{1,i}\}.$$

Now that all the significant variables have been introduced and defined, the simulation algorithm of the R-Vine copula is explained.

Algorithm 3 Sampling from a R-vine copula

Inputs:
Θ is a strict upper triangular matrix of copula parameters with entries $\eta_{ki} = \theta_{ki;1:k-1}$ for $k < i, i = 1, \cdots, d$ and $k = 1, \cdots, d$ for a d-dimensional D-vine
Sample $w_i \overset{\text{i.i.d}}{\sim} U[0,1], i = 1, \ldots, d$
$v_{1,1} := w_1, v^2_{1,1} := w_1$
for i = 2, ..., d **do**
  $v_{i,i} := w_i$
  for $k = i-1, ..., 1$ **do**
    if $m_{k,i} = \tilde{m}_{k,i}$ **then**
      $v_{k,i} := h^{-1}_{m_{ii}|m_{ki};m_{1i},\ldots,m_{k-1,i}}(v_{k+1,i}|v_{k,\tilde{m}_{k,i}}, \eta_{k,i})$
    else$v_{k,i} := h^{-1}_{m_{ii}|m_{ki};m_{1i},\ldots,m_{k-1,i}}(v_{k+1,i}|v^2_{k,\tilde{m}_{k,i}}, \eta_{k,i})$
    end if
    if $i < d$ **then**
      if $m_{k,i} = \tilde{m}_{k,i}$ **then**
        $v^2_{k+1,i} := h^{-1}_{m_{ki}|m_{ii};m_{1i},\ldots,m_{k-1,i}}(v_{k,\tilde{m}_{k,i}}|v_{k,i}, \eta_{k,i});$
      else
        $v^2_{k+1,i} := h^{-1}_{m_{ki}|m_{ii};m_{1i},\ldots,m_{k-1,i}}(v^2_{k,\tilde{m}_{k,i}}|v_{k,i}, \eta_{k,i});$
      end if
    end if
  end for
end for
Return $u_i := v_{1,i}, i = 1, \ldots, d$

7.5 Simulation from factor copula models

As this paper generates predictions using both vine copula models and factor copula models, it is necessary to explain the simulation algorithm for the one- and two-factor models. The algorithms below are based on the book by Czado (2019) and the paper by Kadhem and Nikoloulopoulos (2021).

To simulate the u data using the one-factor copula model, the latent factor v must first be simulated from the $U(0,1)$ distribution. In the second step, independent, identical, uniformly distributed random numbers $w_j, j = 1, ..., d$ are generated. The u data is obtained by using the inverse of the h function, with the simulated latent factor v and the generated random numbers w_j as function arguments. The simulation procedure using the one-factor copula model is summarised in Algorithm 4.

Algorithm 4 Sampling from a one-factor copula

1: Sample $v \sim U(0,1)$
2: Sample $w_j \overset{\text{i.i.d}}{\sim} U[0,1], j = 1, \ldots, d$
3: Then $\quad u_j = C^{-1}_{j|v}(w_j|v), j = 1, ..., d$

The procedure for sampling from the two-factor copula is summarised in Algorithm 5. To simulate the two-factor copula, first the latent factors v_1 and v_2 must be simulated from the $U(0,1)$ distribution. Next, the random numbers w_j are simulated similarly to the one-factor copula model. Then, calculate $C_{j|v_1}(w_j|v_1), j = 1, \ldots, d$ using the first latent factor v_1 and the random numbers w_j as arguments to determine the first argument of the h-function. Finally, the inverse of the h-function, $C_{j|v_1}$ and the second latent factor v_2 are used to generate the u-data.

Algorithm 5 Sampling from a two-factor copula

1: Sample $v_1 \sim U(0,1)$
2: Sample $v_2 \sim U(0,1)$
3: Sample $w_j \overset{\text{i.i.d}}{\sim} U[0,1], j = 1, \ldots, d$
4: Then $\quad rb_j = C_{j|v_1}(w_j|v_1), j = 1, \ldots, d$
5: $\qquad\quad u_j = C^{-1}_{j|v_2}(rb_j|v_2), j = 1, \ldots, d$

The samples simulated in this section are used to determine forecasts, which then need to be checked for accuracy. This paper will focus on two scoring rules for evaluating predictions: The continuous ranked probability score and the interval score. These scoring rules will be defined in the next section.

7.6 Scoring rule

This section presents the scoring rules that will be used to evaluate the forecasts. The continuous ranked probability score and the interval score are considered and are presented below. The definitions are based on the paper by Ferro et al. (2008), Broecker (2012) and Gneiting and Raftery (2004).

Definition 7.33 *(Continuous ranked probability score) The Continuous Ranked Probability Score (CRPS) is a measure that compares a single observation y with the cumulative distribution function F of the forecasts. The mathematical definition of CRPS is as follows*

$$CRPS(F,y) = \int_{\mathbb{R}} (F(x) - \mathbb{1}_{\{x \geq y\}})^2 dx.$$

A CRPS value of zero indicates an accurate prediction, while a value of one indicates an inaccurate prediction.

If the forecasts are given in intervals, the CRPS should not be used; there is a more appropriate method, known as the interval score.

Definition 7.34 *(Interval score) The definition of the Interval Score (IS) is as follows*

$$IS_\alpha(F,y) = (u-l) + \frac{2}{\alpha}(l-y)\mathbb{1}_{\{y<l\}} + \frac{2}{\alpha}(y-u)\mathbb{1}_{\{y>u\}},$$

where y are the observations, l and u are the lower and upper endpoints of the $(1-\alpha)\cdot 100\%$ prediction interval. The terms l is the $\frac{\alpha}{2}$ and u the $1-\frac{\alpha}{2}$ quantile of F.

Observations below the lower endpoint l are penalised in the term $\frac{2}{\alpha}(l-y)\mathbb{1}_{\{y<l\}}$, while the penalty for observations above the upper endpoint u is considered in the term $\frac{2}{\alpha}(y-u)\mathbb{1}_{\{y>u\}}$.

Part II

Case study

8 Yield Curves

This chapter describes the procedure for obtaining copula data using a flowchart. The copula models and factor models used are also summarised in a flowchart. Subsequently, the data are described. The results of modelling the dependencies and forecasting using copula and factor models are presented.

The statistical software RStudio was used for modelling and prediction, with the following packages: fGarch from Wuertz et al. (2022), FactorCopula from Kadhem and Nikoloulopoulos (2023), gamlss.dist from Stasinopoulos and Rigby (2022), rvinecopulib from Nagler and Vatter (2023), VineCopula from Nagler et al. (2023).

8.1 Procedure description

The following section presents a detailed description of the process for determining the copula data and copula model, which is summarized in Flowchart 8.1 and 8.2.

To create the copula data to be used for estimation, an exploratory data analysis is initially conducted on the provided data. In a first step the serial dependence of each variable has to be removed to create necessary copula data consisting of independent rows. The ACF and PACF of each variable are taken into consideration. If they indicate the presence of trends and seasonal effects, said trends and seasonal effects are eliminated. The subsequent phase involves investigating if a seasonal autoregressive moving average model (SARMA) is needed. If the standardized residuals of the SARMA model display autocorrelation in the ACF and PACF plots that is not close to zero, then the ACF and PACF are examined for the squared standardized residuals to see if the GARCH effects are needed. If these autocorrelation in the ACF and PACF are high, GARCH effects are incorporated. Subsequently, a SARMA-GARCH model needs to be fitted, in which the standardized residuals are also taken into account. The standardized residuals distribution is assessed using a histogram to check if it aligns with the assumed innovation F. The probability integral transformation based on it is used to create the copula data. However, if the residuals distribution is not congruent with distribution F, an appropriate distribution must be re-determined. In that event, copula data is then created from the identified probability integral transform using the fitted distribution.

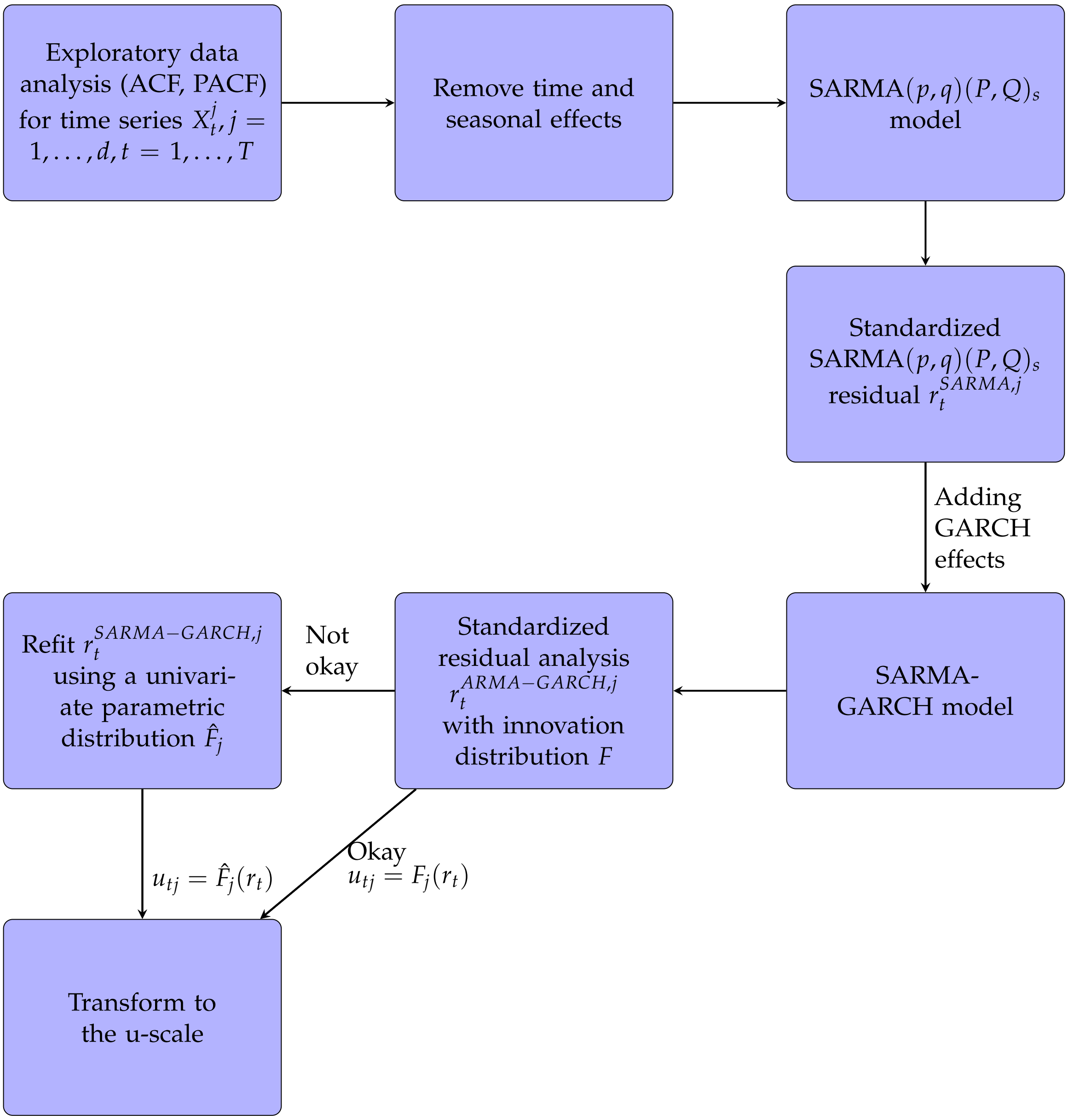

Figure 8.1 Flowchart for obtaining the copula data $\mathcal{U} = (u_{tj}, t = 1, \ldots, T, j = 1, \ldots, d)$

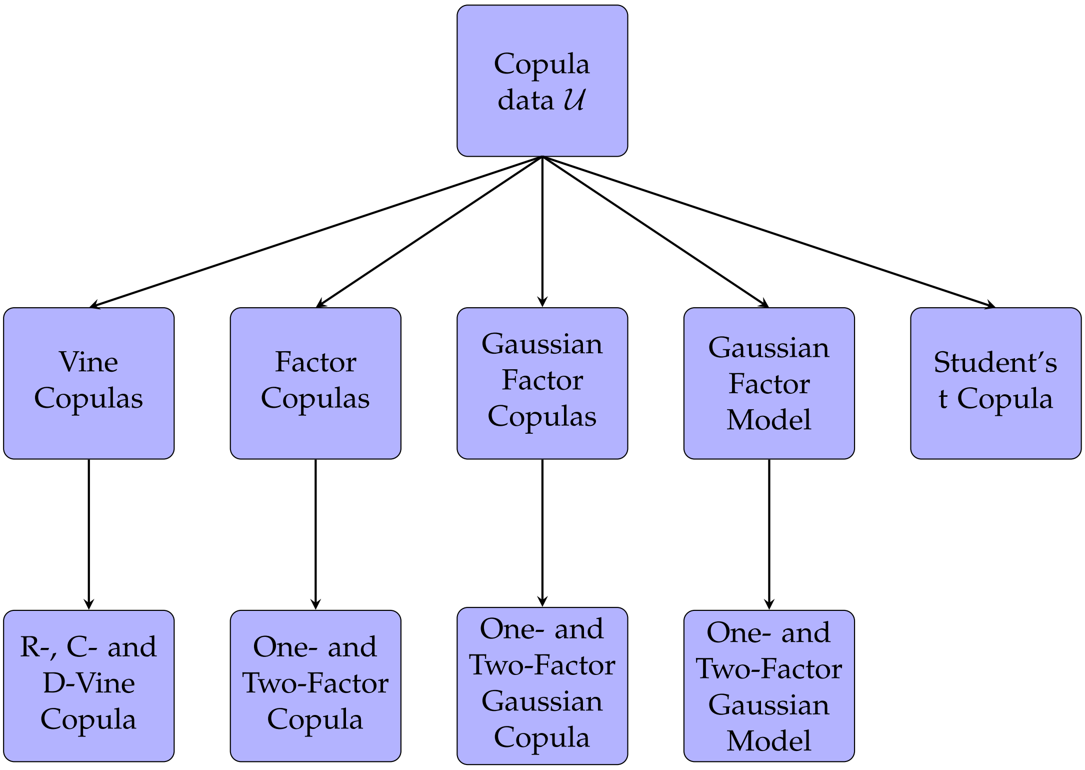

Figure 8.2 Copula models studied using the copula data $\mathcal{U}$

Having created the copula data, we need to find a suitable model to detect dependency structures. We will therefore explore different models, starting with R-, C-, and D-Vine copula models. Additionally, we will consider the one- and two-factor copula model, as well as a specific instance of the Gaussian one- and two-factor copula model. Further to this, we will also compare to the multivariate Student's t copula model and the Gaussian factor model.

8.2 Data description

Monthly German yield curves with maturities of 1, 5, 10, 15 and 20 years, covering the period from 1986-06-01 to 2019-12-01 (403 observations) were selected to analyse the dependency. The yield curves are shown in Figure 8.3.
For the analysis, the following abbreviations are used for the yield curve with different maturities: **M1** for the 1-year maturity, **M5** for the 5-year maturity, **M10** for the 10-year maturity, **M15** for the 15-year maturity and **M20** for the 20-year maturity.

The Figure shows a trend in the yield curves for the various maturities. To prepare the data for time series analysis, the trend must be removed by differentiating the yield curves

$$\Delta X_t^j = X_t^j - X_{t-1}^j, j = 1, ..., d, t = 1, ..., T,$$

where $X_t^j = Mj_t, j = 1, 5, 10, 15, 20$. Differenced yield curves are shown in Figure 8.4. After removing the trend, an autoregressive moving average model is fitted. The results are presented in the next section.

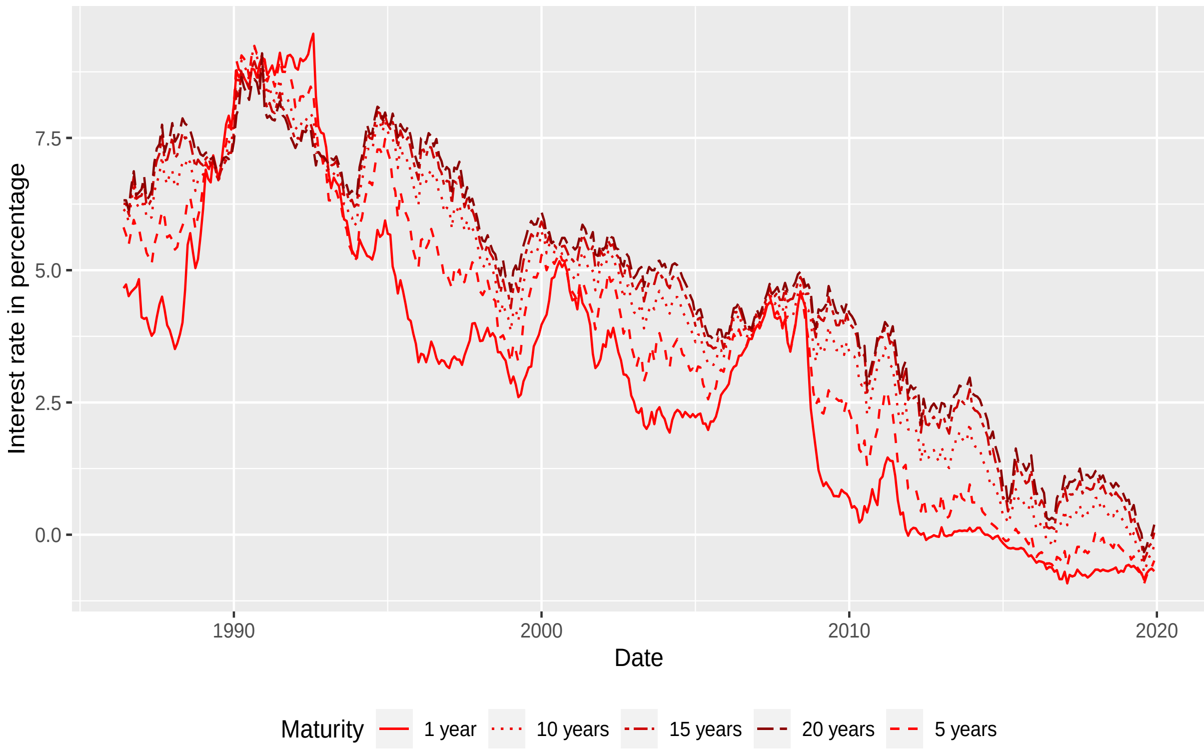

Figure 8.3 Monthly yield curves for maturities of 1, 5, 10, 15 and 20 year(s) with observation number $T = 403$

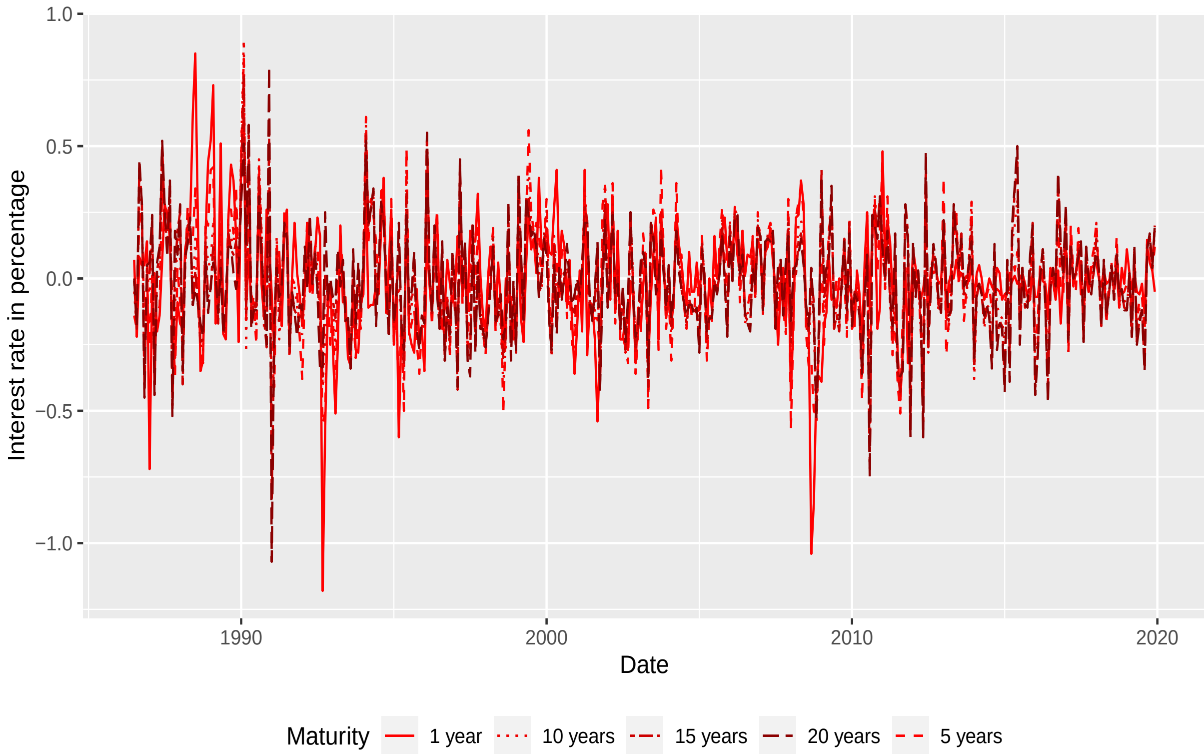

Figure 8.4 Differenced yield curves for maturities of 1, 5, 10, 15 and 20 year(s) with observation number $T = 403$

8.3 Marginal ARMA fitting

Autoregressive moving average models for differenced yield curves are fitted below. Table 8.1 presents the ARMA model formulas for the differenced yield curves with maturities of 1, 5, 10, 15 and 20 years.

Table 8.1 Fitted ARMA model formula

ARMA model formula	W_t
$M1_t = -0.013 + 0.325M1_{t-1} + W_t$	$W_t \sim N(0, 0.040)$
$M5_t = -0.016 + W_t + 0.173W_{t-1}$	$W_t \sim N(0, 0.044)$
$M10_t = -0.016 - 0.054M10_{t-1} - 0.817M10_{t-2} + 0.156M10_{t-3} + W_t + 0.113W_{t-1} + 0.796W_{t-2}$	$W_t \sim N(0, 0.040)$
$M15_t = -0.015 + W_t$	$W_t \sim N(0, 0.042)$
$M20_t = -0.015 - 0.149M20_{t-1} - 0.962M20_{t-2} - 0.02M20_{t-3} + W_t + 0.083W_{t-1} + 0.946W_{t-2}$	$W_t \sim N(0, 0.043)$

Starting with the differenced yield curves with a maturity of one year, an autoregressive model of order one was fitted. The AR(1) model formula is given in the table above, where $M1_{t-1}$ is the differenced yield curve with maturity one year at time $t-1$ with coefficient $\phi_1 = 0.325$, W_t the white noise process and the intercept -0.013. Next a moving average model with an order of one was fitted to the differenced yield curve with a maturity of five years. As with the one-year maturity yield curve, the formula for the MA(1) model can be found in Table 8.1. The ARMA(3, 2) model was used to fit the differenced yield curve with a maturity of 10 years. In the model formula, $M20_{t-1}, M20_{t-2}, M20_{t-3}$ are the differenced yield curves at time $t-1, t-2$ and $t-3$ and W_{t-1}, W_{t-2} the white noise at time $t-1$ and $t-2$. The differenced yield curve with a maturity of 15 years was only fitted by the white noise process and the intercept. The autoregressive moving average model with orders $p = 3$ and $q = 2$ was fitted for the differenced yield curve with a maturity of 20 years. The terms $M20_{t-1}, M20_{t-2}, M20_{t-3}$ are the differenced yield curves with maturity 20 years at time points $t-1, t-2$ and $t-3$ with coefficients $\phi_1 = -0.149, \phi_2 = -0.962, \phi_3 = -0.02$. The white noise for the time points $t-1$ and $t-2$ is then denoted by W_{t-1} and W_{t-2}, respectively, with coefficients $\theta_1 = 0.083$ and $\theta_2 = 0.946$.

After fitting the ARMA model, it is necessary to check if the standardized residuals follow the white noise model. This can be done by examining the ACF and PACF of the standardized residuals, as well as the ACF and PACF of the squared standardized residuals. The ACF and PACF of the standardized residuals, as shown in Appendix A in Figure A.1, reveal a few significant lags. Additionally, the ACF and PACF of the squared standardized residuals in Figure A.2 indicate a need for the GARCH model, as more lags are significant.

Table 8.2 ARMA for the yield curves with maturity of 1, 5, 10, 15 and 20 year(s)

	Dependent variable:				
	M1	M5	M10	M15	M20
ϕ_1	0.325***		−0.054		−0.149**
	(0.047)		(0.096)		(0.064)
ϕ_2			−0.817***		−0.962***
			(0.093)		(0.064)
ϕ_3			0.156***		−0.020
			(0.054)		(0.057)
θ_1		0.173***	0.113		0.083**
		(0.050)	(0.086)		(0.040)
θ_2			0.796***		0.946***
			(0.105)		(0.070)
intercept	−0.013	−0.016	−0.016	−0.015	−0.015
	(0.015)	(0.012)	(0.011)	(0.010)	(0.010)
Observations	402	402	402	402	402
Log Likelihood	74.683	56.203	77.880	66.038	63.009
σ^2	0.040	0.044	0.040	0.042	0.043
Akaike Inf. Crit.	−143.366	−106.407	−141.760	−128.075	−112.018
Note:					*p<0.1; **p<0.05; ***p<0.01

8.4 Marginal ARMA-GARCH with skew Student's t innovation

This section presents the results of the ARMA-GARCH model with skew Student's t innovation. Table 8.3 shows the formulas for the fitted ARMA-GARCH model and Table 8.4 summarises the results of the ARMA-GARCH model.

Table 8.3 Fitted ARMA-GARCH model formula

ARMA-GARCH model formula
$M1_t = -0.008 + 0.210M1_{t-1} + \sqrt{0.223W_{t-1}^2 + 0.808\sigma_{t-1}^2}\varepsilon_t$
$M5_t = -0.012 + \sqrt{0.077W_{t-1}^2 + 0.916\sigma_{t-1}^2}\varepsilon_t + 0.119\sqrt{0.077W_{t-2}^2 + 0.916\sigma_{t-2}^2}\varepsilon_{t-1}$
$M10_t = -0.029 - 0.034M10_{t-1} - 0.803M10_{t-2} + 0.151M10_{t-3} + \sqrt{0.005 + 0.048W_{t-1}^2 + 0.84\sigma_{t-1}^2}\varepsilon_t + 0.1\sqrt{0.005 + 0.048W_{t-2}^2 + 0.84\sigma_{t-2}^2}\varepsilon_{t-1} + 0.814\sqrt{0.005 + 0.048W_{t-3}^2 + 0.84\sigma_{t-3}^2}\varepsilon_{t-2}$
$M15_t = -0.020 + \sqrt{0.005 + 0.092W_{t-1}^2 + 0.8\sigma_{t-1}^2}\varepsilon_t$
$M20_t = -0.035 - 0.064M20_{t-1} - 0.853M20_{t-2} + 0.051M20_{t-3} + \sqrt{0.03 + 0.272W_{t-1}^2}\varepsilon_t + 0.05\sqrt{0.03 + 0.272W_{t-2}^2}\varepsilon_{t-1} + 0.88\sqrt{0.03 + 0.272W_{t-3}^2}\varepsilon_{t-2}$

Let this section start by looking at the ARMA$(1,0)$-GARCH$(1,1)$ model for the differenced yield curve with a maturity of one year, which is given in Table 8.3, where $M1_{t-1}$ the differenced yield curve with maturity one year at time $t-1$, σ^2_{t-1} the conditional variance and W_{t-1} the process from Equation (4.10). We will now examine the tests for the standardized residuals, which can be found in Appendix Section A.
The Jarque-Bera and Shapiro-Wilk tests indicate that the current standardized residuals do not conform to a normal distribution. As the p-value is less than 5%, the null hypothesis can be rejected. The Ljung box test indicates that the correlation coefficients are zero, while the LM ARCH test shows that no ARCH effects are included.

The differenced yield curve with a maturity of five years was modelled using the ARMA$(0,1)$-GARCH$(1,1)$ model. The standardized residuals tests, Jarque-Bera and Shapiro-Wilk, show that the standardized residuals follow a normal distribution. The Ljung-Box and LM ARCH tests indicate that there are no significant correlation coefficients for any lags and no ARCH effects present.

The 10-year differenced yield curve was modelled using the ARMA$(3,2)$-GARCH$(1,1)$ model. Examining the residual test, it is evident that the standardized residuals do not follow a normal distribution, as indicated by the Jarque-Bera and Shapiro-Wilk tests. Similar to previous cases, the Ljung-Box and LM ARCH tests also indicate the presence of correlation coefficients of zero for all lags and no ARCH effects.

Additionally, the GARCH$(1,1)$ model of the differenced yield curve with a maturity of 15 years is considered. The model is defined by the equation in Table 8.3. The Jarque-Bera test indicates that the standardized residuals do not follow a normal distribution. Since the p-value for the remaining residual tests is greater than $\alpha = 5\%$, the null hypothesis cannot be rejected. According to the Shapiro-Wilk test, this suggests that the standardized residuals follow a normal distribution. The correlation coefficients for all lags are zero and the LM ARCH test finally proves the non-existence of ARCH effects.

The ARMA$(3,2)$-GARCH$(1,1)$ model was used to model the differenced yield curve with a maturity of 20 years. The Jarque-Bera and Shapiro-Wilk tests indicate that the standardized residuals do not follow a normal distribution. As before, the Ljung-Box test indicates a correlation coefficient of zero for all lags. Furthermore, the LM ARCH test indicates no ARCH effects.

The ACF and PACF for the standardized and squared standardized residuals were also taken into account for the fitted ARMA-GARCH models. The ACF and PACF are shown in Figure A.3 and Figure A.4 in the Appendix. The plots show that there are no significant correlation coefficients for most lags.

Table 8.4 ARMA-GARCH model for the yield curves with maturity of 1, 5, 10, 15 and 20 year(s)

	Dependent variable:				
	M1	M5	M10	M15	M20
μ	-0.008	-0.012	-0.029	-0.020*	-0.035
	(0.006)	(0.011)	(0.019)	(0.010)	(0.018)
ϕ_1	0.210***		-0.034		-0.064
	(0.054)		(0.072)		(0.059)
ϕ_2			-0.803***		-0.853***
			(0.063)		(0.039)
ϕ_3			0.151**		0.051
			(0.053)		(0.053)
θ_1		0.119*	0.100		0.050
		(0.051)	(0.054)		(0.029)
θ_2			0.814***		0.884***
			(0.071)		(0.040)
α_0	0.000	0.000	0.005	0.005	0.031**
	(0.000)	(0.000)	(0.005)	(0.003)	(0.010)
α_1	0.223**	0.077*	0.048	0.092*	0.272**
	(0.075)	(0.036)	(0.035)	(0.047)	(0.105)
β_1	0.808***	0.916***	0.839***	0.801***	0.000
	(0.050)	(0.041)	(0.140)	(0.093)	(0.252)
skew	0.954***	1.175***	1.116***	1.027***	0.945***
	(0.075)	(0.086)	(0.104)	(0.082)	(0.075)
df	5.590***	10**	10**	10**	7.16**
	(1.395)	(3.345)	(3.455)	(3.735)	(2.300)
Observations	402	402	402	402	402
Log Likelihood	0.381	0.167	0.217	0.188	0.221
Akaike Inf. Crit.	−0.726	−0.299	−0.379	−0.346	−0.386
Note:			*p<0.1; **p<0.05; ***p<0.01		

After modelling the marginal dependencies with ARMA-GARCH, the standardized residuals can be transformed into copula data using the skew Student's t innovation. This transformation requires the probability integral transform given in (3.9). However, since the skew Student's t distribution is not suitable as an innovation, a more appropriate distribution must be determined. The table below presents the univariate parametric distribution used for the transformation to the u-scale.

Table 8.5 Univariate fitted parametric distribution $\hat{F}_j$ to the standardized residuals of ARMA-GARCH model

Variable	Distribution $\hat{F}$
M1	Skew t type 4
M5	Skew normal type 2
M10	Generalized t
M15	Generalized t
M20	Logistic

Figure 8.5 displays QQ-plots comparing the skew Student's t distribution with the chosen univariate parametric distributions. The plots indicate that the selected distributions are only slightly more appropriate than the skew Student's t distribution. The standardized residuals were transformed into copula data using the PIT and the selected distributions. The pairs plot of the copula data is displayed in Figure 8.6 with the estimated Kendall's tau values above, contour plots below and histograms on the diagonal. From the pair plot we can see that Kendall's tau is above 50% for the following pairs: M1 and M5, M5 and M10, M10 and M15, M15 and M20, M5 and M10, M10 and M20.

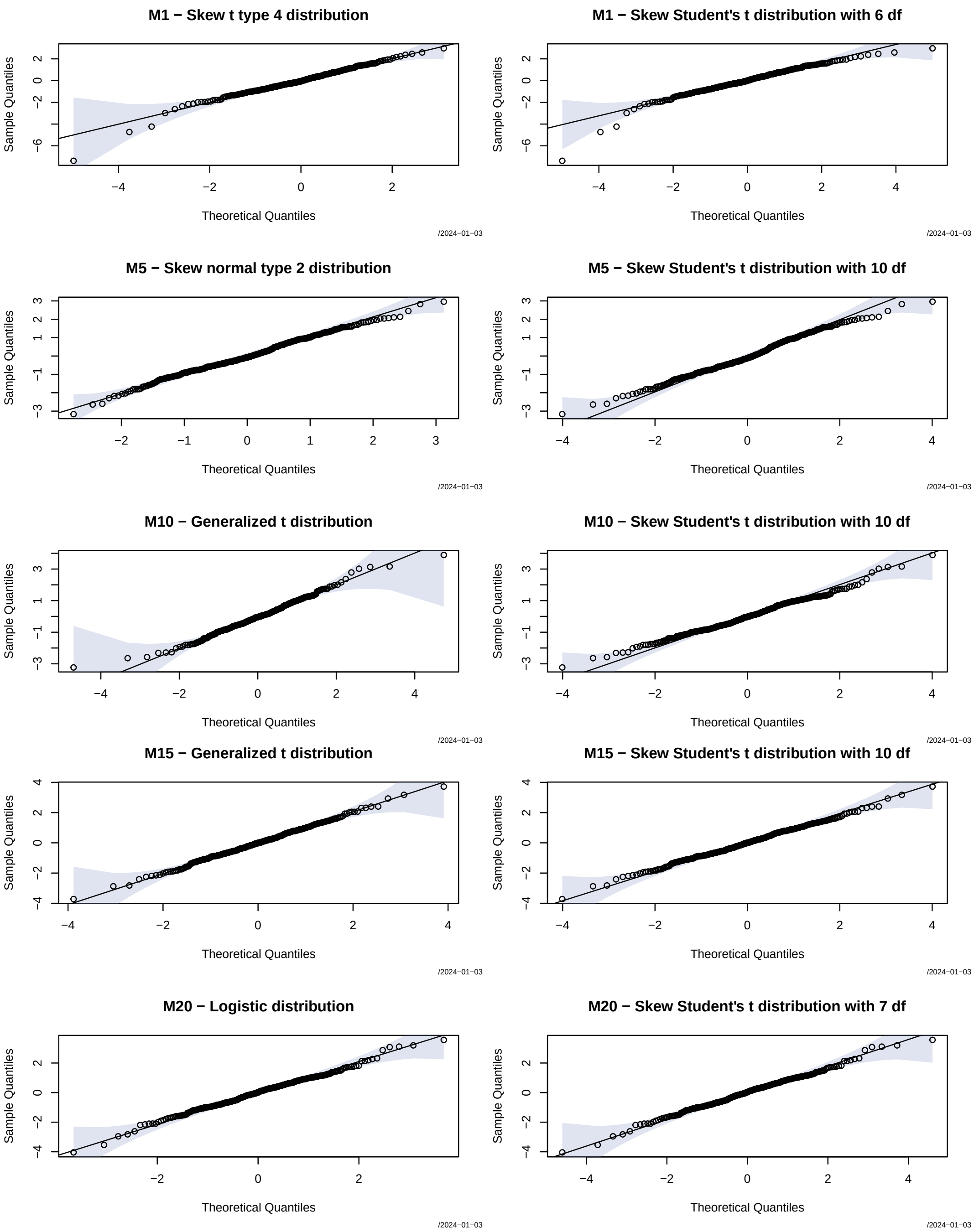

Figure 8.5 QQ-plots for the standardized residuals with the skew Student's t (right column) and the refit distributions (left column).

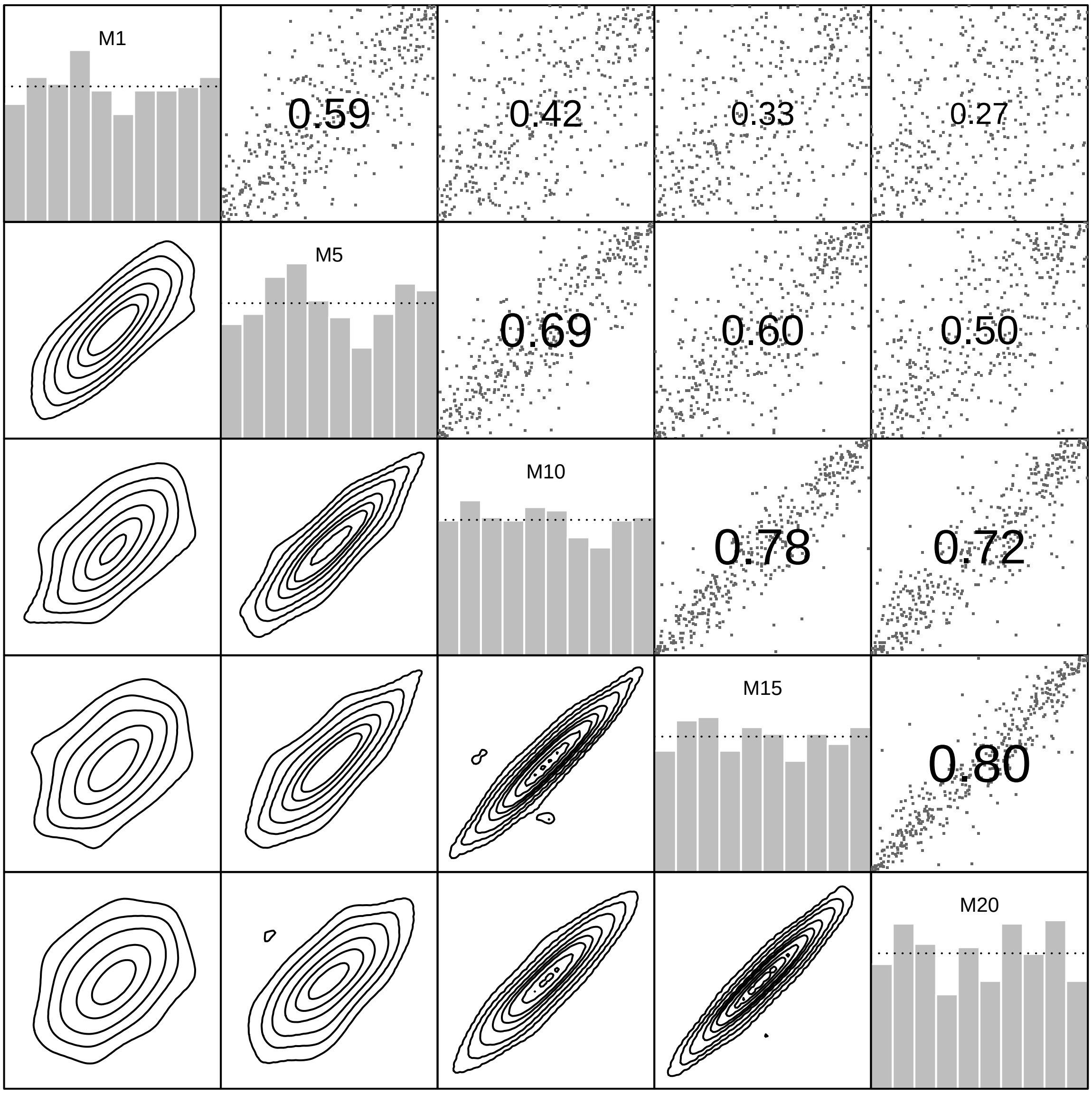

Figure 8.6 Pairs plot of the copula data created for the differenced yield curves

8.5 Vine copula fitting

This section presents the fitting of various vine copula models. First, the yield curve is modelled using the regular vine copula. The modelling results are summarized in Table 8.6. The structure of the R-vine indicates that it corresponds to the D-vine. For the first tree, the copula pair families consist of the Gaussian and Student's t copulas. The dependency measure tau exceeds 50% for all pair copulas in tree one. Trees 2 to 4 have as pair copula families the Frank, Gaussian, Gumbel, Student's t and the independence copula. The dependency measure τ for Tree 2 to 4 ranges from -0.34 to 0.

Table 8.6 Fitted R-vine copula for the differenced yield curves

tree	conditioned	conditioning	family	rotation	parameters	tau
1	1, 2		gaussian	0	0.77	0.564
1	2, 3		gaussian	0	0.87	0.674
1	3, 4		t	0	0.94, 3.80	0.776
1	4, 5		t	0	0.95, 3.36	0.789
2	1, 3	2	gumbel	90	1.3	-0.213
2	2, 4	3	t	0	-0.21, 5.51	-0.131
2	3, 5	4	t	0	-0.03, 3.14	-0.019
3	1, 4	3, 2	gaussian	0	-0.23	-0.149
3	2, 5	4, 3	frank	0	-3.4	-0.337
4	1, 5	4, 3, 2	indep	0		0.000

Now, let's take a look at the so-called canonical vine copula with order 14253. The outcomes are presented in Table 8.7. C-vine copula models typically exhibit a star structure, with the yield curve at 10 years maturity being the central point in the first tree. This point is then connected to the other variables. The pair copulas families for the first tree are then the Frank, Student's t and Gumbel copulas. The second tree has the yield curve with a 20-year maturity as the central point, and it uses the Frank and Gaussian copulas as pair copulas. The third tree has the yield curve with a 5-year maturity as the central point, in which the Gaussian and Student's t copulas have been selected as pair copulas. The dependency measure τ ranges from -0.31 to 0.78 for all trees.

Table 8.7 Fitted C-vine copula for the differenced yield curves

tree	conditioned	conditioning	family	rotation	parameters	tau
1	1, 3		frank	0	4.5	0.42
1	4, 3		t	0	0.94, 3.80	0.78
1	2, 3		gaussian	0	0.87	0.67
1	5, 3		t	0	0.9, 6.1	0.71
2	1, 5	3	frank	0	-2.7	-0.28
2	4, 5	3	frank	0	6.4	0.53
2	2, 5	3	gaussian	0	-0.47	-0.31
3	1, 2	5, 3	gaussian	0	0.61	0.41
3	4, 2	5, 3	t	0	0.22, 10.37	0.14
4	1, 4	2, 5, 3	gaussian	0	-0.2	-0.13

The last model is the drawable copula. The first tree in the D-vine copula has a structure of order 45321. The pair copulas chosen were the Gaussian and Student's t copulas. The dependence measure τ in trees 2 to 4 ranges from -0.31 to 0.79, and the chosen pair copulas are the Clayton, Gaussian, Gumbel and Student's t copulas. Table 8.8 presents the results.

Table 8.8 Fitted D-vine copula for the differenced yield curves

tree	conditioned	conditioning	family	rotation	parameters	tau
1	4, 5		t	0	0.95, 3.36	0.789
1	5, 3		t	0	0.9, 6.1	0.711
1	3, 2		gaussian	0	0.87	0.674
1	2, 1		gaussian	0	0.77	0.564
2	4, 3	5	t	0	0.67, 5.79	0.468
2	5, 2	3	gaussian	0	-0.47	-0.314
2	3, 1	2	gumbel	270	1.3	-0.213
3	4, 2	3, 5	t	0	0.22, 11.31	0.142
3	5, 1	2, 3	clayton	270	0.13	-0.062
4	4, 1	2, 3, 5	gaussian	0	-0.24	-0.156

8.6 Factor copula models on the copula scale

The following section discusses factor copula models, including the one- and two-factor models.

8.6.1 One-factor arbitrary copula

The one-factor model is now examined in detail. Table 8.9 display the families, parameters, and tau. The dependency measure τ between the latent variable and the variables M1, M5, M10, M15 and M20 is illustrated in Figure 8.7.

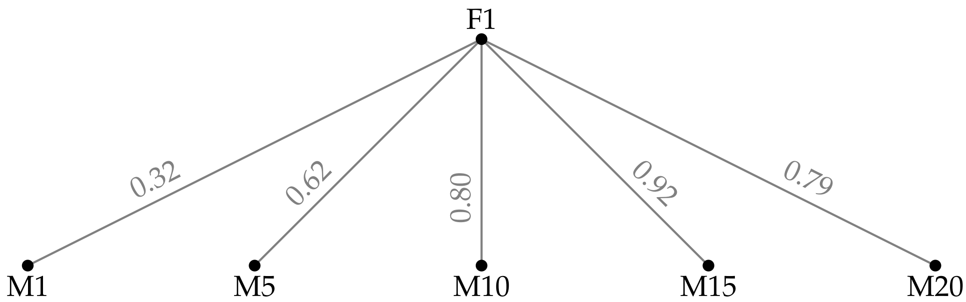

Figure 8.7 Fitted one-factor copula for the differenced yield curves with the given dependence measure τ

The reflected Gumbel copula was chosen as the pair copula for the yield curve with a maturity of 1 year and the latent factor, with a dependency measure of 32%. For the yield curve with a maturity of 5 years, the Frank copula was selected as the pair copula, and for the yield curve with maturities of 10 and 20 years, the Student's t copula with three degrees of freedom was chosen. The yield curve with a maturity of 15 years was modelled using the Gaussian copula as the pair copula. The latent factor explained 62% of the dependency for M5, 80% for M10, 92% for M15, and 79% for M20.

Table 8.9 Fitted one-factor copula for the differenced yield curves

variables	family	parameters	tau
M1	rgum	1.46	0.32
M5	frk	8.35	0.62
M10	bvt3	0.95	0.80
M15	bvn	0.99	0.92
M20	bvt3	0.95	0.79

8.6.2 Two-factor arbitrary copula

We will now discuss the fitted two-factor copula model. The pair copulas selected for this model are the same as those used in the one-factor copula model. Table 8.10 presents the families, parameters and tau. In the one-factor copula model, the dependency measure for yield curves with maturities of 10, 15 and 20 years exceeds 50%. The tau for the two-factor model exceeds 50% for the maturities of 1, 5, 10 and 15 years.

Table 8.10 Fitted two-factor copula for the differenced yield curves

variables	factor	family	parameters	tau
M1	1	rgum	1.15	0.13
M5	1	frk	4.30	0.41
M10	1	bvt3	0.84	0.64
M15	1	bvn	0.95	0.80
M20	1	bvt3	0.96	0.82
M1	2	rgum	2.51	0.60
M5	2	frk	29.03	0.87
M10	2	bvt3	0.89	0.69
M15	2	bvn	0.89	0.70
M20	2	bvt3	0.46	0.31

Table 8.11 summarises the abbreviations for the pair copula families.

Table 8.11 Abbreviations for pair copula families

Abbreviation	Pair copula family
rgum	reflected Gumbel
frk	Frank
bvt3	Student’s t with three degrees of freedom
bvn	Gaussian

The previous models selected arbitrary pair copula families. In the following, Gaussian copulas are selected as pair copula families.

8.6.3 One-factor Gaussian copula

The one-factor Gaussian copula model is defined by Equation (8.1), where $C_{i|V}(u_i|v), i = M1, M5, M10, M15, M20$ is the partial derivative of the joint copula $C_{iV}(u_i, v)$ with respect to v.

$$C(u_{M1}, u_{M5}, u_{M10}, u_{M15}, u_{M20}) = \int_0^1 C_{M1|v}(u_{M1}|v) C_{M5|v}(u_{M5}|v) C_{M10|v}(u_{M10}|v) \cdot C_{M15|v}(u_{M15}|v) C_{M20|v}(u_{M20}|v) dv. \tag{8.1}$$

Now, we provide a more detailed representation of $C_{M1|v}(u_{M1}|v)$. First, the bivariate Gaussian copula is presented.

$$C(u_{M1}, v; \rho) = \Phi_2(\Phi^{-1}(u_{M1}), \Phi^{-1}(v); \rho)$$

Now the partial derivative of the copula $C_{M1,v}(u_{M1}, v; \rho)$ with respect to v has to be computed.

$$\begin{aligned}
C_{M1|v}(u_{M1}|v) &= \frac{d}{dv} C_{M1,v}(u_{M1}, v; \rho) \\
&= \frac{d}{dv} \Phi_2(\Phi^{-1}(u_{M1}), \Phi^{-1}(v); \rho) \\
&= \Phi\left(\frac{\left[\Phi^{-1}(u_{M1}) - \rho_{M1,1}\Phi^{-1}(v)\right]}{\sqrt{1-\rho_{M1,1}^2}}\right)
\end{aligned}$$

The remaining terms can be expressed similarly. To obtain the one-factor Gaussian copula, multiply the conditional copula distribution functions and integrate them.
Now let's consider the results for the one-factor Gaussian copula, which are summarised in Table 8.12. Similar to the one-factor copula model with arbitrary copulas, the yield curve with a maturity of 1 year has a dependency lower than 50%, while the remaining variables have a higher dependency measure.

Table 8.12 Fitted one-factor Gaussian copula for the differenced yield curves

variables	family	parameters	tau
M1	bvn	0.54	0.37
M5	bvn	0.85	0.64
M10	bvn	0.97	0.84
M15	bvn	0.97	0.83
M20	bvn	0.93	0.76

8.6.4 Two-factor Gaussian copula

In comparison to the one-factor copula model, the two-factor copula model includes an additional latent factor. Mathematically, the two-factor copula model can be represented by the following equation.

$$\begin{aligned}
C(u_{M1}, u_{M5}, u_{M10}, u_{M15}, u_{M20}) = \int_0^1 \int_0^1 & C_{M1|V_2;V_1}(C_{M1|V_1}(u_{M1}|v_1)|v_2) \\
& \cdot C_{M5|V_2;V_1}(C_{M5|V_1}(u_{M5}|v_1)|v_2) \\
& \cdot C_{M10|V_2;V_1}(C_{M10|V_1}(u_{M10}|v_1)|v_2) \\
& \cdot C_{M15|V_2;V_1}(C_{M15|V_1}(u_{M15}|v_1)|v_2) \\
& \cdot C_{M20|V_2;V_1}(C_{M20|V_1}(u_{M20}|v_1)|v_2) dv_1 dv_2
\end{aligned}$$

The inner term $C_{j|V_2;V_1}(C_{j|V_1}(u_j|v_1)|v_2), j = M1, M5, M10, M15, M20$ is determined by

$$C_{j|V_2;V_1}(C_{u_j|V_1}(u_j|v_1)|v_2) = \frac{\partial}{\partial v_2} C_{jV_2;V_1}(C_{u_j|V_1}(u_j|v_1), v_2).$$

For the Gaussian copula, we obtain

$$C_{u_j|V_2;V_1}(C_{u_j|v_1}|v_2) = \Phi\left(\frac{\Phi^{-1}(u_j) - \alpha_{j1}\Phi^{-1}(v_1) - \gamma_j(1-\alpha_{j1}^2)^{\frac{1}{2}}\Phi^{-1}(v_2)}{\sqrt{(1-\alpha_{j1}^2)(1-\gamma_j^2)}}\right)$$

with correlation α_{j1} for C_{j,V_1}. The partial correalation $\gamma_j = \frac{\alpha_{j2}}{\sqrt{1-\alpha_{j1}^2}}$ for $(\Phi(U_j), \Phi(V_2))$ given $\Phi(V_1)$, where α_{j2} is the correlation of $\Phi(U_j)$ and $\Phi(V_2)$. As with the one-factor model, multiply the conditional copulas first and then integrate to obtain the two-factor copula model.

Consider now the results for the two-factor copula model, which are summarised in Table 8.13. In this case, the Gaussian copula was used as the pair copula family for all variables. The estimated parameter values are shown in the fourth column, and the dependence measure is shown in the last column. The dependence measure ranges from 11% to 82% for the first factor and from 46% to 96% for the second factor.

Table 8.13 Fitted two-factor Gaussian copula for the differenced yield curves

variables	factor	family	parameters	tau
M1	1	bvn	0.18	0.11
M5	1	bvn	0.50	0.33
M10	1	bvn	0.82	0.61
M15	1	bvn	0.91	0.72
M20	1	bvn	0.96	0.82
M1	2	bvn	0.81	0.60
M5	2	bvn	1.00	0.96
M10	2	bvn	0.91	0.73
M15	2	bvn	0.85	0.64
M20	2	bvn	0.66	0.46

8.7 Gaussian factor models on z-scale

This section presents the results for the one and two factor models, starting with the one factor model. The factor models require data on the z-scale defined in 7.2. This data is obtained by using the quantile of the standard normal distribution. Now consider the case where only one factor is included.

8.7.1 One-factor Gaussian model

The fitted factor model with one latent factor f_1 is given by

$$\begin{aligned} M1 &= 0.56 f_1 + u_{M1} \\ M5 &= 0.86 f_1 + u_{M5} \\ M10 &= 0.98 f_1 + u_{M10} \\ M15 &= 0.95 f_1 + u_{M15} \\ M20 &= 0.91 f_1 + u_{M20}, \end{aligned}$$

where $u_i, i = M1, ..., M20$ are normally distributed random variables with mean zero and variance $\psi_i = \text{diag}(0.68, 0.25, 0.03, 0.10, 0.17)$ and

$$\Sigma = \begin{pmatrix} 1.00 & & & & \\ 0.49 & 1.00 & & & \\ 0.56 & 0.85 & 1.00 & & \\ 0.54 & 0.82 & 0.93 & 1.00 & \\ 0.51 & 0.79 & 0.90 & 0.86 & 1.00 \end{pmatrix}$$

the sample correlation matrix of the observed data for variables **M1**, **M5**, **M10**, **M15**, **M20**.

As the data has been transformed on the z-scale, the results of the factorial model cannot be compared with the other models directly. In order to be able to compare them, the results must be transformed to the u-scale. This was done using the density transformation theorem.

8.7.2 Two-factor Gaussian model

Adding another latent factor, f_2, the fitted two-factor Gaussian model is given by the following equations

$$\begin{aligned} M1 &= 0.19f_1 + 0.79f_2 + u_{M1} \\ M5 &= 0.51f_1 + 0.86f_2 + u_{M5} \\ M10 &= 0.81f_1 + 0.54f_2 + u_{M10} \\ M15 &= 0.89f_1 + 0.39f_2 + u_{M15} \\ M20 &= 0.95f_1 + 0.22f_2 + u_{M20}. \end{aligned}$$

The variance for u_i is given by

$$\psi = \text{diag}(0.34, 0.01, 0.06, 0.05, 0.04)$$

and the sample correlation matrix of the observed data for variables **M1**, **M5**, ..., **M20** by

$$\Sigma = \begin{pmatrix} 1.00 & & & & \\ 0.77 & 1.00 & & & \\ 0.58 & 0.87 & 1.00 & & \\ 0.48 & 0.78 & 0.93 & 1.00 & \\ 0.36 & 0.67 & 0.89 & 0.94 & 1.00 \end{pmatrix}.$$

8.8 Multivariate t copula in 5 dimensions

This section fits the Student's t copula in five dimensions. The Student's t copula(Σ, ν) requires the multidimensional Student's t distribution with the parameters, ν the degrees of freedom and Σ the correlation matrix. For the fit, we did not specify the degrees of freedom, but had them estimated. The estimated degrees of freedom are $\nu = 8$ and the correlation coefficients are given in the following matrix.

$$\Sigma = \begin{pmatrix} 1.00 & 0.77 & 0.57 & 0.46 & 0.38 \\ 0.77 & 1.00 & 0.88 & 0.81 & 0.70 \\ 0.57 & 0.88 & 1.00 & 0.95 & 0.90 \\ 0.46 & 0.81 & 0.95 & 1.00 & 0.95 \\ 0.38 & 0.70 & 0.90 & 0.95 & 1.00 \end{pmatrix}$$

The R package copula from Hofert et al. (2023), Jun Yan (2007), Ivan Kojadinovic and Jun Yan (2010) and Marius Hofert and Martin Mächler (2011) was used to fit the multivariate Student's t copula.

So far we have fitted various vine copulas, factor copula models, the multivariate Student's t copula and factor models. Now the best model has to be determined. This is done by comparing the different models. The selection criteria used for the comparison are AIC, BIC and log-likelihood. The results of the comparison are presented in detail in the next section.

8.9 Model comparison

Table 8.14 compares copula models and factor models using selection criteria AIC, BIC, and log-likelihood. The drawable vine copula model with order 45321 has the smallest AIC and BIC compared to the other models, making it the best model among the models studied. The AIC and BIC for the R-vine, C-vine, two-factor copula, two-factor Gaussian copula, two-factor Gaussian model and Student's t copula are closely behind the D-vine copula.

Table 8.14 Model comparison

Models	AIC	BIC	logLik	Number of parameters
RVine = DVine (BIC) with order 12345	-2942.45	-2890.50	1484.23	13
CVine (BIC) with order 14253	-2935.74	-2883.79	1480.87	13
DVine (BIC) with order 45321	-2951.29	-2895.34	1489.64	14
One-factor copula	-2231.26	-2203.29	1122.63	7
Two-factor copula	-2812.71	-2756.76	1420.35	14
One-factor Gaussian copula	-2145.10	-2125.12	1077.55	5
Two-factor Gaussian copula	-2740.29	-2700.33	1380.15	10
One-factor Gaussian model transformed to u-scale	-2142.26	-2102.29	1081.13	15
Two-factor Gaussian model transformed to u-scale	-2742.83	-2702.87	1381.42	20
t copula	-2938.77	-2894.81	1480.39	11

8.10 Forecasting

This section presents a forecast of the yield curve for the next three months. The process involves several steps, beginning with predicting time $T = 404$.

1. Forecasted mean and standard deviations
   For the variables **M1**, **M5**, ..., **M20**, the mean $\hat{\mu}_{(T+1)}$ and standard deviation $\hat{\sigma}_{(T+1)}$ were forecasted for the time $T + 1 = 404$. The forecasted parameters for the different maturities are presented in Table 8.15.

Table 8.15 Mean and standard deviation forecast for $T + 1 = 404$

	$\hat{\mu}_{T+1}$	$\hat{\sigma}_{T+1}$
M1	-0.018	0.085
M5	0.003	0.139
M10	0.016	0.188
M15	-0.020	0.193
M20	-0.007	0.204

2. Sampling from a copula model
   Using the fitted R-, C- and D-vine copula and the one- and two-factor copula models, $n = 1000$ samples are generated in this step.
   $$(u^n_{T+1,1}, ..., u^n_{T+1,d}) \sim C, \quad n = 1, ..., 1000, T + 1 = 404.$$

3. Transform simulated u-data into standardized residuals
   The u-data generated in the second step are transformed into the standardized residuals. These residuals are determined by the inverse of the univariate distribution function.
   $$u_{(T+1)j} = \hat{F}_j(\hat{r}^n_{(T+1)j}) \Leftrightarrow \hat{r}^n_{(T+1)j} = \hat{F}_j^{-1}(u^n_{(T+1)j}), j = M1, ..., M20, T + 1 = 404.$$

4. Forecast of differenced data
   Using the estimated forecasts from step one and the results from step three, we obtain predictions for the differenced data.
   $$\widehat{\Delta Y}^n_{(T+1)j} = \hat{\mu}_{(T+1)j} + \hat{\sigma}_{(T+1)j} \cdot \hat{r}^n_{(T+1)j}$$
   where we use $\Delta Y_{(T+1)j}$ instead of $\Delta M1_{(T+1)j}, ..., \Delta M20_{(T+1)j}$ with $n = 1, ..., 1000, j = M1, ..., M20$.

5. Forecasting the yield curve
   To obtain a forecast on the original scale, the differenced data step must be reversed.
   $$\widehat{Y}_{(T+1)j} = \widehat{\Delta Y}_{(T+1)j} + Y_{Tj}$$

To predict the yield curves for time $T + 2 = 405$, the given data needs to be shifted by one month, starting at $t = 2, ..., 404$. The time series analysis, ARMA-GARCH model, and modelling using vine copula and factor copula models must be performed again on the new data set. To determine the predicted yield curve for $T + 2 = 405$, repeat steps 1 to 5

Table 8.16 Mean and standard deviation forecast for $T+2=405$

	$\hat{\mu}_{T+2}$	$\hat{\sigma}_{T+2}$
M1	0.005	0.084
M5	-0.036	0.145
M10	-0.046	0.195
M15	-0.021	0.199
M20	-0.020	0.216

as previously instructed. The parameters estimated $\hat{\mu}_{(T+2)}$ and $\hat{\sigma}_{(T+2)}$ in the first step are summarized in Table 8.16. For the time $T+3=406$, we do the same procedure. We form the new data set and start at time $t=3$ and go to $t=405$. The time series analysis and modelling were performed with the data and finally the yield curve was predicted for $T+3$. The estimated parameters used for this are presented in Table 8.17.

Table 8.17 Mean and standard deviation forecast for $T+3=406$

	$\hat{\mu}_{T+3}$	$\hat{\sigma}_{T+3}$
M1	-0.025	0.089
M5	-0.029	0.143
M10	-0.028	0.193
M15	-0.022	0.201
M20	-0.018	0.209

Results from the forecasting using the R-, C- and D-vine copula models

The results obtained from the simulation are summarized in Table A.1 in Appendix A, including the observed value, the mode and the prediction interval. Figure 8.8 displays the values predicted using the R-vine copula. The figure contains the density, the 5% and 95% quantile of the predictions and the observed value. The figure presents the maturities M1, M5, M10, M15 and M20 in the rows. The columns show the times $T+1, T+2$ and $T+3$. For all maturities and times, the observed values are included in the forecasts and lie between the 5% and 95% quantiles. This indicates that the predicted values are accurate.

Figure 8.9 presents the forecasts using the C-vine copula. The figure has the same structure as above. The observed values lie between the two quantiles and we can conclude that the observed values were well covered by the predictions.

As with the two previous cases, Figure 8.10 presents the results of the prediction using the D-vine copula model. The predictions cover the observed value well, as the observed value is contained within the predicted values and lies between the two quantiles.

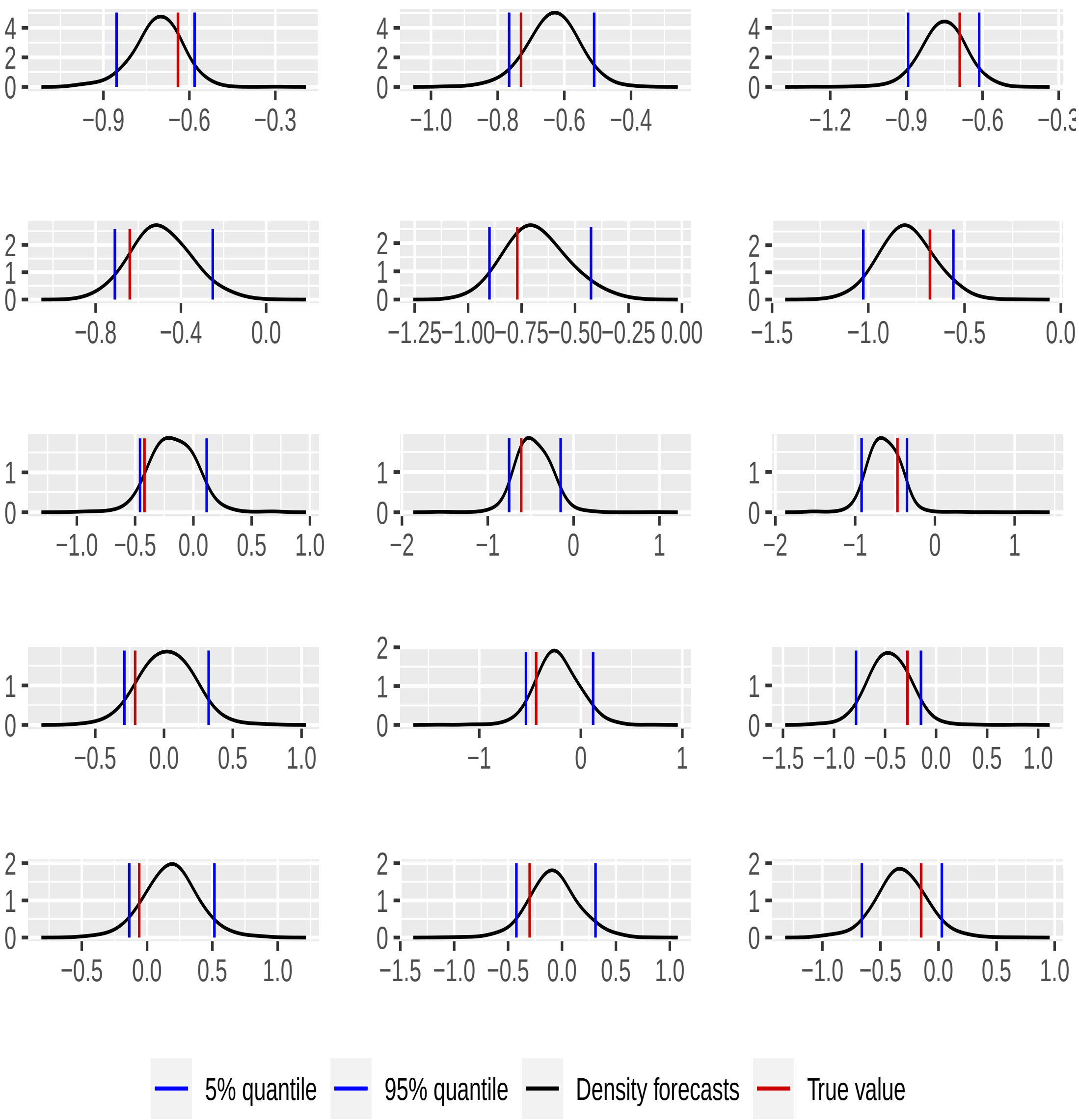

Figure 8.8 The R-vine copula is used to forecast yield curves. The maturities M1, M5, M10, M15, and M20 are presented in rows and the predicted times $T+1=404, T+2=405, T+3=406$ are presented in columns.

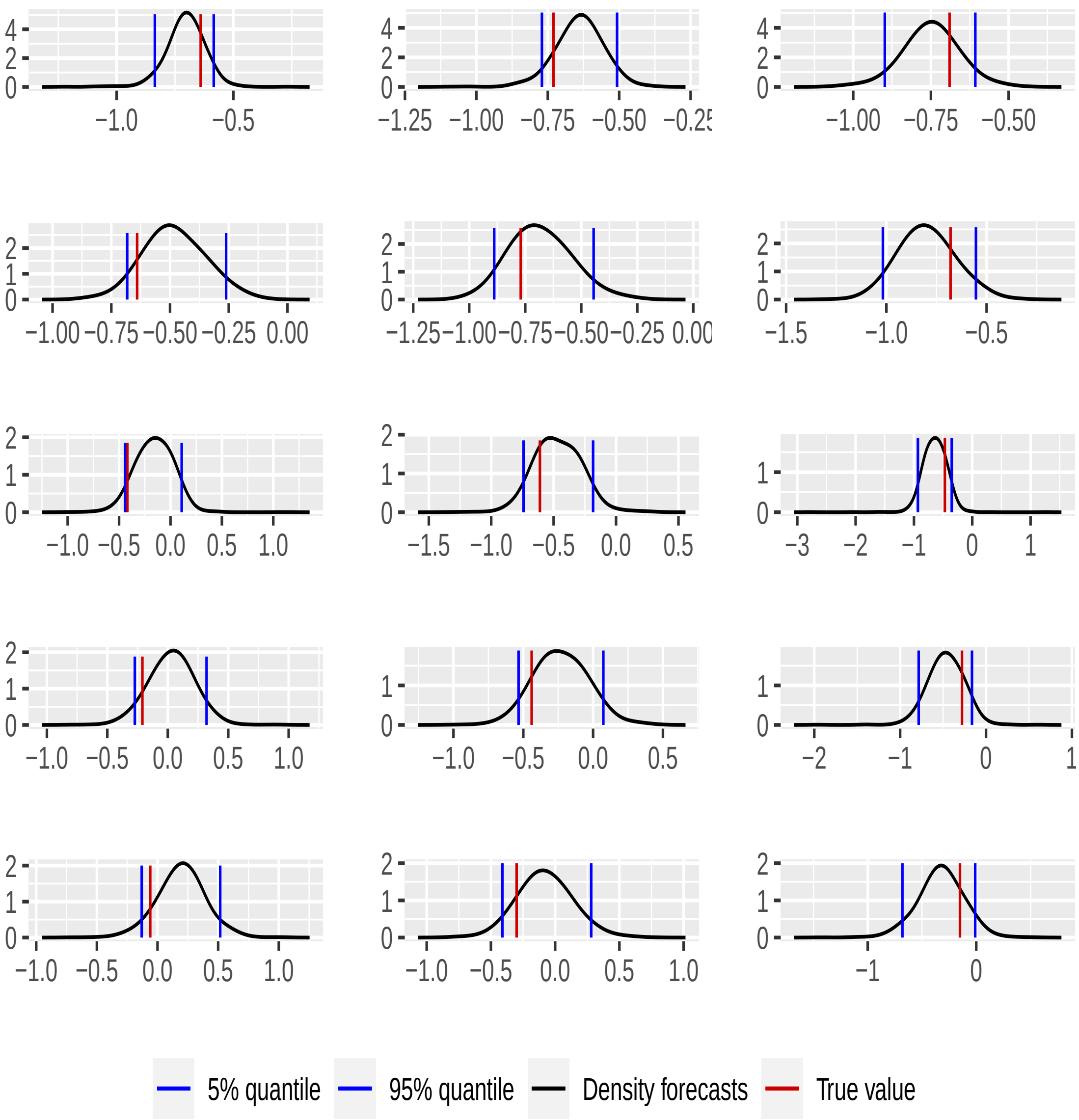

Figure 8.9 The C-vine copula is used to forecast yield curves. The maturities M1, M5, M10, M15, and M20 are presented in rows and the predicted times $T + 1 = 404, T + 2 = 405, T + 3 = 406$ are presented in columns.

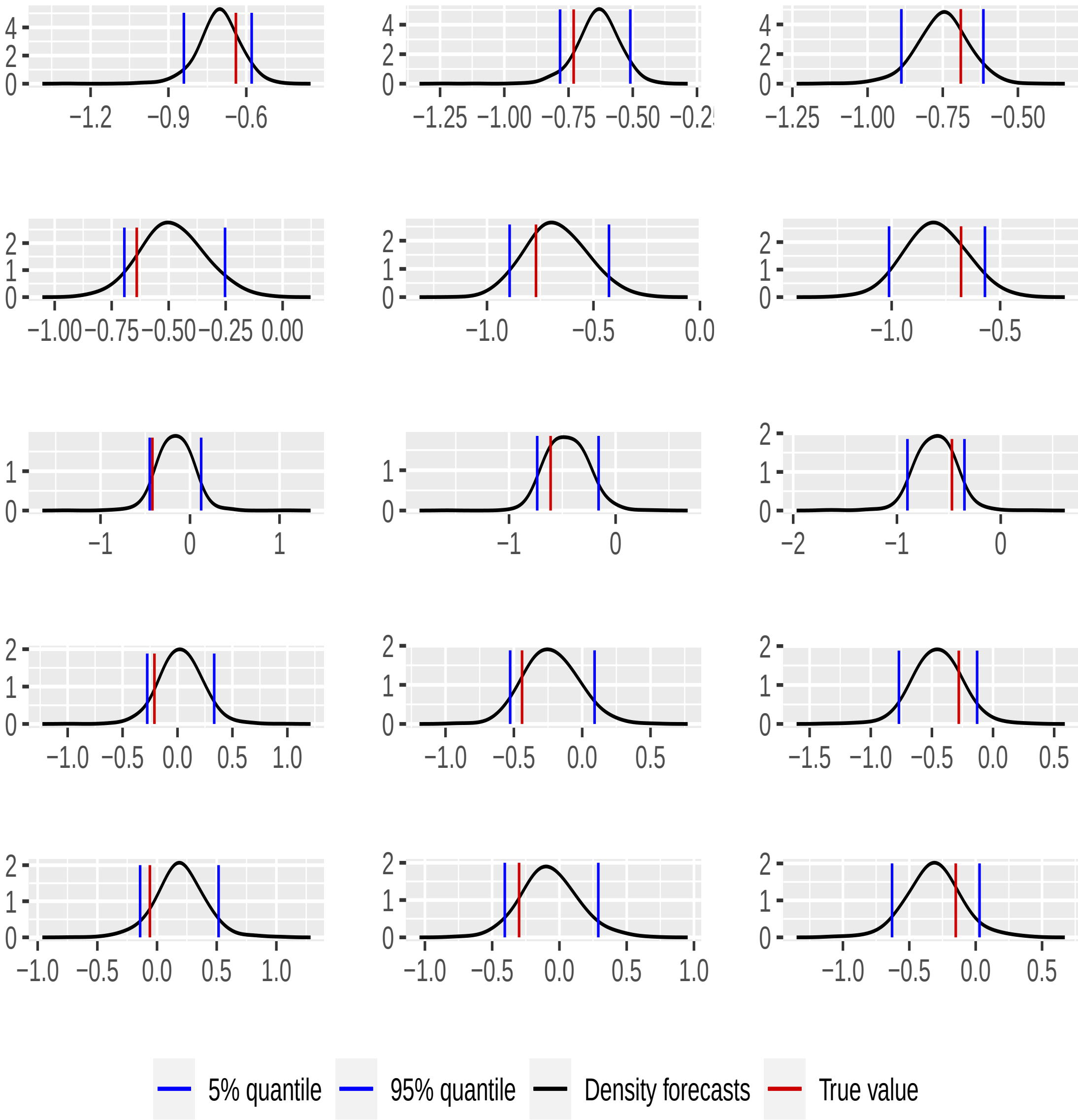

Figure 8.10 The D-vine copula is used to forecast yield curves. The maturities M1, M5, M10, M15, and M20 are presented in rows and the predicted times $T+1, T+2, T+3$ are presented in columns.

Results from the forecasting using the one- and two-factor copula models

Now consider the forecasting using the one and two factor copula model. Figures 8.11 and 8.12 display the prediction results, while Appendix A Table A.2 presents the simulation results, including the mode, predicted interval and observed value. As in the previous cases, for both the one-factor and two-factor copula models, the observed values are included in the predictions and lie between the two quantiles. This indicates that the true values are well covered by the predictions.

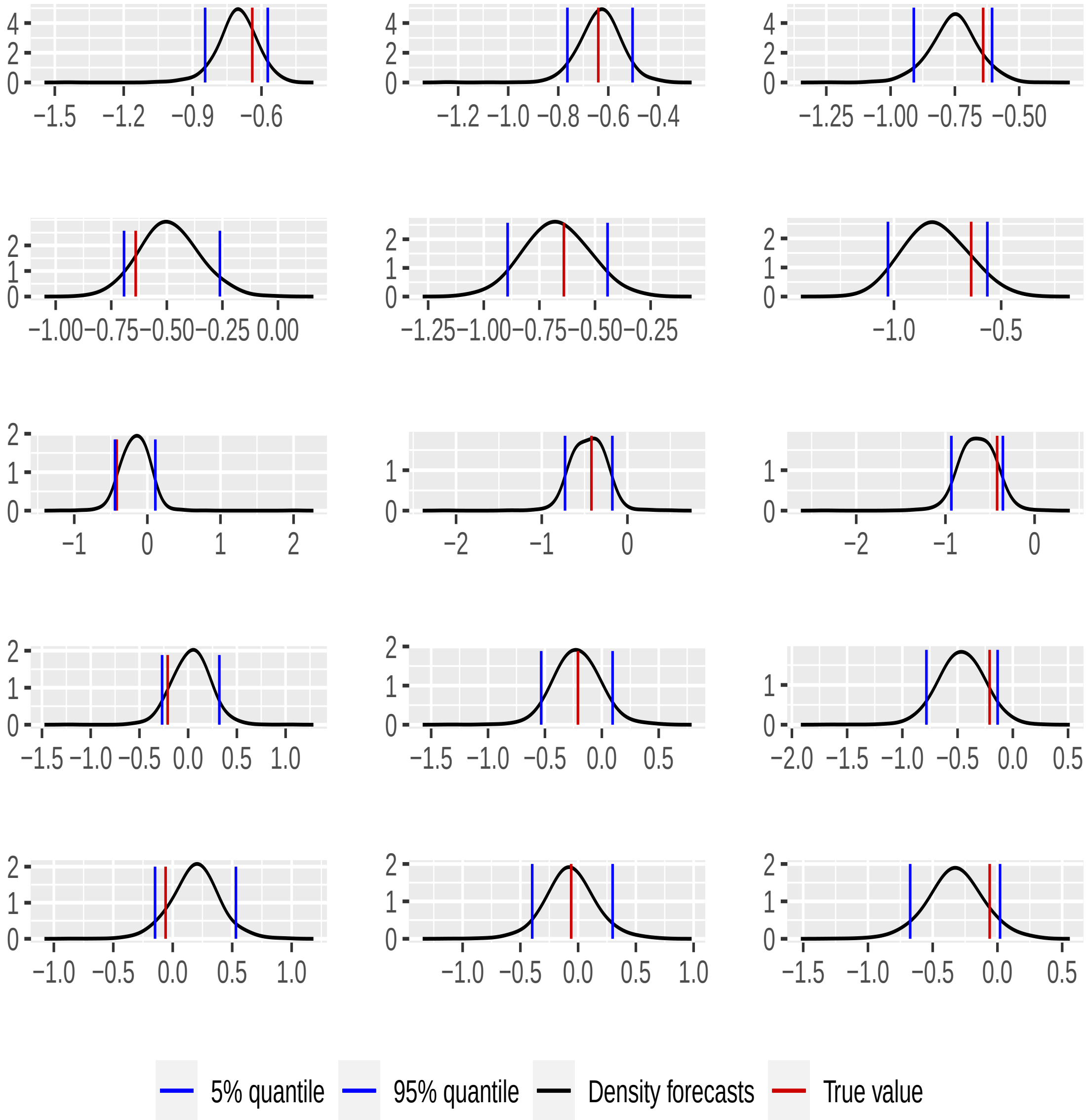

Figure 8.11 The one-factor copula is used to forecast yield curves. The maturities M1, M5, M10, M15, and M20 are presented in rows and the predicted times $T+1, T+2, T+3$ are presented in columns.

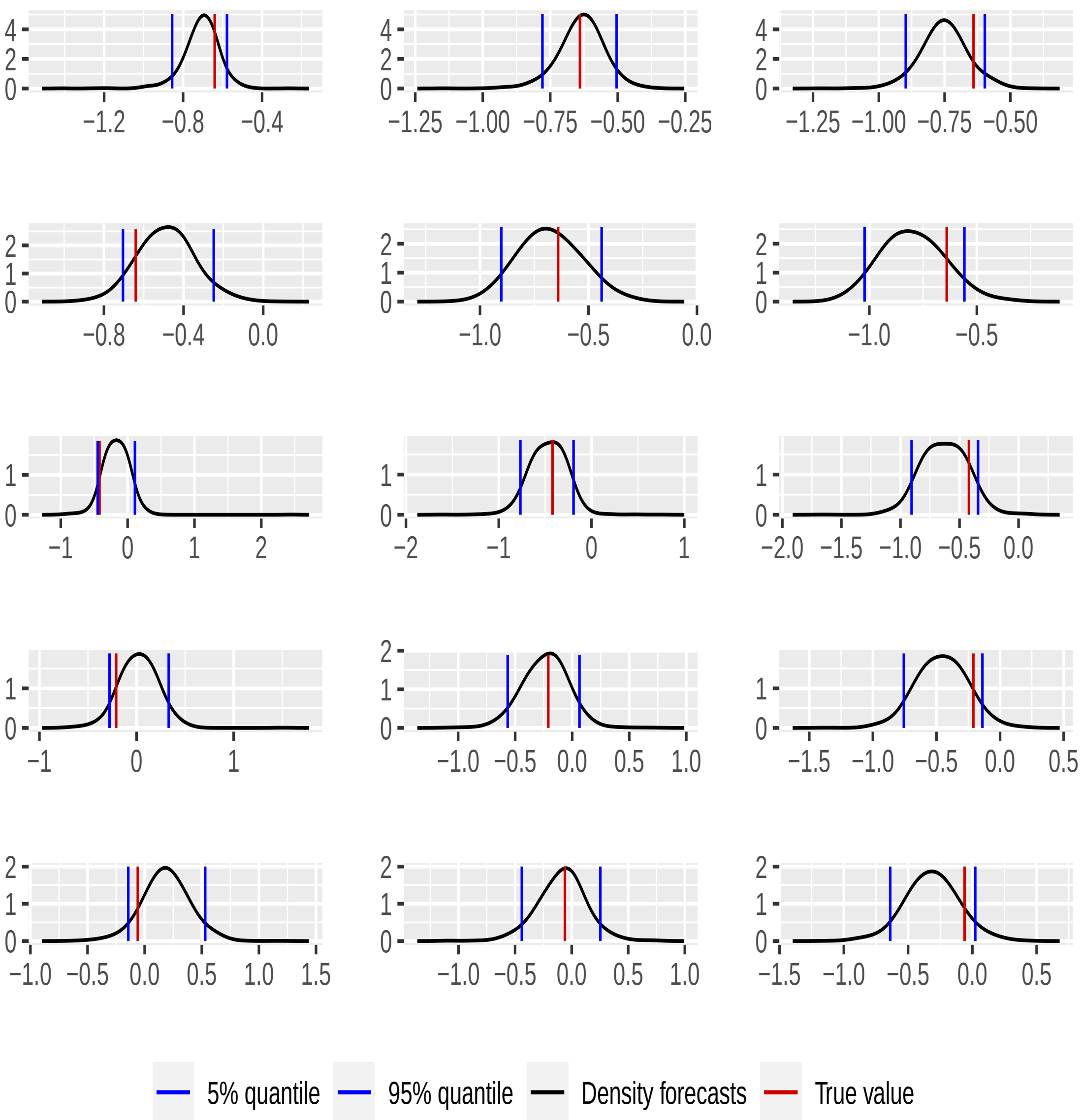

Figure 8.12 The two-factor copula is used to forecast yield curves. The maturities M1, M5, M10, M15, and M20 are presented in rows and the predicted times $T+1, T+2, T+3$ are presented in columns.

In order to determine the best model, a comparison of the results from the predictions is necessary. This is achieved through the use of the continuous ranked probability score (CRPS) and interval score (IS). The CRPS for all models and predicted time points are summarised in Table 8.18. The table indicates that all models have the same CRPS for the yield curve with a maturity of 1 year. For the remaining variables, several models provide the same CRPS in some cases. To identify the optimal model, we search for the one with the smallest CRPS. As multiple models give the same result, there is no definitive best model. To select a model regardless, we computed the sum of CRPS for each model. It is important to note that the variables are not independent, so the sums cannot simply be calculated. Nevertheless, we consider the sums and obtain the best model, the D-Vine copula model, which was also identified as the best model in Section 8.9.

Table 8.18 Continuous ranked probability score for the forecasted yield curves

	$T+1=404$					$T+2=405$					$T+3=406$				
Model	RVine	CVine	DVine	1F	2F	RVine	CVine	DVine	1F	2F	RVine	CVine	DVine	1F	2F
M1	0.04	0.04	0.04	0.04	0.04	0.06	0.06	0.06	0.06	0.06	0.04	0.04	0.03	0.04	0.04
M5	0.09	0.09	0.09	0.09	0.10	0.05	0.06	0.05	0.06	0.05	0.07	0.07	0.07	0.07	0.08
M10	0.16	0.16	0.15	0.16	0.17	0.08	0.09	0.08	0.09	0.09	0.10	0.10	0.09	0.09	0.10
M15	0.14	0.15	0.14	0.14	0.16	0.13	0.13	0.13	0.13	0.13	0.11	0.11	0.10	0.11	0.11
M20	0.15	0.16	0.15	0.15	0.17	0.13	0.13	0.13	0.14	0.13	0.10	0.10	0.10	0.10	0.11
$\sum$	0.58	0.60	0.57	0.58	0.64	0.45	0.47	0.45	0.48	0.46	0.42	0.42	0.39	0.41	0.44

Now lets consider the interval score in Table 8.19 where we can see that several models have the same IS. To determine the best model, we select the one with the lowest interval score. It is important to note that the variables are dependent, so their sum should not be calculated. However, to select a model, we determined that the one with the smallest sum is the best. The D-vine copula model is considered the best model for time $T+1$, while the one-factor copula model is the best for time $T+2$, and the R-vine copula model is the best for time $T+3$. As the D-vine copula was chosen as the best model for the CRPS and also produced low IS in some cases, it is considered to be the most suitable model.

Table 8.19 Interval score for the forecasted yield curves

	$T+1=404$					$T+2=405$					$T+3=406$				
Model	RVine	CVine	DVine	1F	2F	RVine	CVine	DVine	1F	2F	RVine	CVine	DVine	1F	2F
M1	0.28	0.28	0.27	0.27	0.27	0.33	0.27	0.28	0.27	0.26	0.29	0.27	0.28	0.28	0.27
M5	0.45	0.46	0.43	0.45	0.45	0.53	0.47	0.48	0.48	0.47	0.45	0.46	0.45	0.45	0.44
M10	0.54	0.58	0.56	0.57	0.56	0.68	0.57	0.60	0.55	0.56	0.54	0.55	0.56	0.58	0.57
M15	0.61	0.62	0.59	0.62	0.61	0.75	0.62	0.66	0.60	0.60	0.60	0.63	0.63	0.62	0.63
M20	0.64	0.68	0.66	0.69	0.67	0.87	0.70	0.71	0.64	0.68	0.64	0.69	0.67	0.67	0.68
$\sum$	2.52	2.62	2.51	2.60	2.56	3.16	2.63	2.73	2.54	2.57	2.52	2.60	2.59	2.60	2.59

9 Yield Curves and Inflation Rate

This chapter presents the second case study, which includes an economic factor in the analysis. The chapter begins with a description of the data and presents the results of the modelling and forecasting in the following sections.

9.1 Data description

The second case study includes the German inflation rate as an economic factor in addition to the yield curves from the first case. The yield curves take into account maturities of 1, 5, 10, 15 and 20 years. Both the inflation rate and yield curves are provided for the period 1990-10-01 to 2019-12-01 (351 observations). Technical abbreviations are used for the analysis, with the same abbreviations for yield curves as in Chapter 8. The inflation rate is abbreviated as **Inf** in the following. Figure 9.1 displays the time series for the yield curves and inflation rate.

Figure 9.1 Monthly yield curves for maturities of 1, 5, 10, 15 and 20 year(s) and inflation rate with observation number $T = 351$

The figure shows a trend in the data which needs to be removed. This was removed by differentiation. Figure 9.2 presents the differenced data.

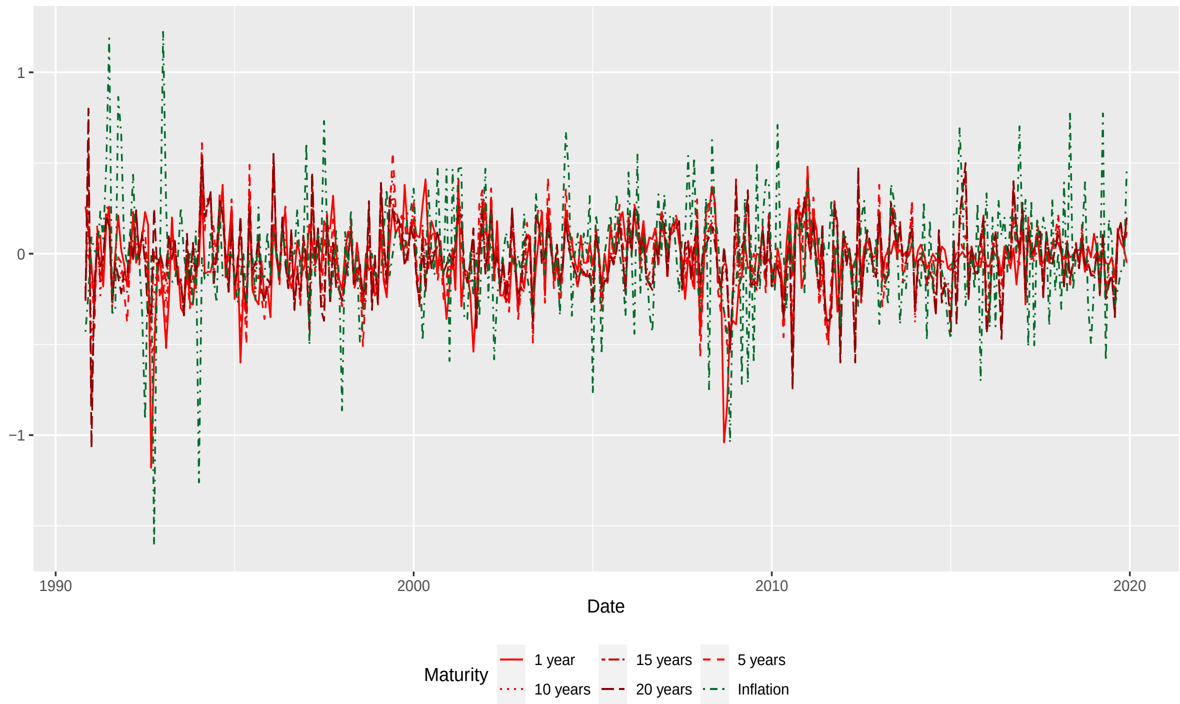

Figure 9.2 Differenced yield curves for maturities of 1, 5, 10, 15 and 20 year(s) and inflation rate with observation number $T = 351$

For modelling and forecasting, the procedure is the same as for the first case study. The results are presented and not explained in detail in this case.

9.2 Marginal ARMA fitting

This section models the serial dependencies using the ARMA model. Table 9.1 presents the ARMA models with their respective orders.

Table 9.1 ARMA models for the yield curves and inflation rate

Variable	Model
M1	$\text{ARMA}(1,0)$
M5	$\text{ARMA}(0,1)$
M10	$\text{ARMA}(3,2)$
M15	$\text{ARMA}(3,0)$
M20	$\text{ARMA}(3,0)$
Inf	$\text{ARMA}(0,2) \times (2,1)_{s=12}$

The differenced inflation rate was modelled using a seasonal ARMA model. The non-seasonal part has an order of $p = 0$ and $q = 2$, while the seasonal part has an order of $P = 2$, $Q = 1$ and frequency $s = 12$. The seasonal ARMA model is determined using Definition 4.9, which yields the following representation.

$$(1 - \Phi_1 B^{12} - \Phi_2 B^{24})\text{Inf} = (1 + \Theta_1 B^{12})(1 + \theta_1 B + \theta_2 B^2)W_t \tag{9.1}$$

The solution to Equation (9.1) provides the final seasonal ARMA model.

$$\begin{aligned}
\text{Inf}_t &= \Phi_1 \text{Inf}_{t-12} + \Phi_2 \text{Inf}_{t-24} + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \Theta_1 W_{t-12} + \theta_1 \Theta_1 W_{t-13} \\
&\quad + \theta_2 \Theta_1 W_{t-14} \\
&= -0.007 - 0.069\text{Inf}_{t-12} - 0.057\text{Inf}_{t-24} + W_t - 0.029W_{t-1} + 0.080W_{t-2} \\
&\quad - 0.593W_{t-12} + 0.069 \cdot 0.593W_{t-13} + 0.057 \cdot 0.593W_{t-14}
\end{aligned}$$

The remaining variables can be modelled using ARMA models, similar to the first case study. The model formulas for these variables are presented in Table 9.2.

Table 9.2 ARMA model formulas for the differenced yield curves and inflation rate

ARMA model formula	W_t
$M1_t = -0.026 + 0.313M1_{t-1} + W_t$	$W_t \sim N(0, 0.033)$
$M5_t = -0.027 + W_t + 0.150W_{t-1}$	$W_t \sim N(0, 0.041)$
$M10_t = -0.026 - 0.044M10_{t-1} - 0.788M10_{t-2} + 0.144M10_{t-3} + W_t + 0.105W_{t-1} + 0.758W_{t-2}$	$W_t \sim N(0, 0.036)$
$M15_t = -0.025 - 0.024M15_{t-1} - 0.034M15_{t-2} + 0.131M15_{t-3} + W_t$	$W_t \sim N(0, 0.038)$
$M20_t = -0.024 - 0.082M20_{t-1} - 0.021M20_{t-2} + 0.115M20_{t-3} + W_t$	$W_t \sim N(0, 0.041)$
$\text{Inf}_t = -0.007 - 0.069\text{Inf}_{t-12} - 0.057\text{Inf}_{t-24} + W_t - 0.029W_{t-1} + 0.080W_{t-2}$ $- 0.593W_{t-12} + 0.069 \cdot 0.593W_{t-13} + 0.057 \cdot 0.593W_{t-14}$	$W_t \sim N(0, 0.072)$

To determine whether the GARCH model is necessary, the standardized residuals are examined. This is done by analysing the ACF and PACF of the standardized residuals and the squared standardized residuals. The ACF and PACF plots (Figure B.1) show some significant lags and the ACF and PACF for the squared standardized residuals (Figure B.2) show increased significant lags, confirming the need for the GARCH model.

Table 9.3 ARMA for yield curves with maturity of 1, 5, 10, 15, 20 year(s) and inflation rate

	Dependent variable:					
	M1	M5	M10	M15	M20	Inf
ϕ_1	0.313***		−0.044	−0.024	−0.082	
	(0.051)		(0.146)	(0.053)	(0.053)	
ϕ_2			−0.788***	−0.034	−0.021	
			(0.226)	(0.054)	(0.054)	
ϕ_3			0.144**	0.131**	0.115**	
			(0.068)	(0.055)	(0.057)	
θ_1		0.150***	0.105			−0.029
		(0.053)	(0.137)			(0.054)
θ_2			0.758***			0.080
			(0.248)			(0.053)
intercept	−0.026*	−0.027**	−0.026**	−0.025**	−0.024**	−0.007
	(0.014)	(0.012)	(0.011)	(0.011)	(0.011)	(0.006)
Φ_1						−0.069
						(0.137)
Φ_2						−0.057
						(0.099)
Θ_1						−0.593***
						(0.131)
Observations	350	350	350	350	350	350
Log Likelihood	100.382	61.154	85.593	76.672	63.199	−40.762
σ^2	0.033	0.041	0.036	0.038	0.041	0.072
Akaike Inf. Crit.	−194.764	−116.309	−157.186	−143.344	−116.398	95.524

Note: *p<0.1; **p<0.05; ***p<0.01

9.3 Marginal ARMA-GARCH fitting

This section extends the ARMA model from the previous section by incorporating the GARCH model. Similarly to the first case, the ARMA-GARCH models for the yield curve with maturities of 1, 5, 10, 15 and 20 years can be determined. As the inflation rate requires a seasonal ARMA model, a two-step approach is needed for the GARCH model. First, the seasonal ARMA model is fitted and the standardized residuals are determined. These are then used as a new dataset for the GARCH model, taking only the order of the GARCH model into account. Table 9.4 presents the results of the ARMA-GARCH fit. The table is not explained in detail.

Table 9.4 ARMA-GARCH model for the yield curves with maturity of 1, 5, 10, 15, 20 year(s) and inflation rate

	Dependent variable:					
	M1	M5	M10	M15	M20	Inf
μ	-0.009	-0.020	-0.040*	-0.021*	-0.022*	-0.006
	(0.006)	(0.011)	(0.020)	(0.010)	(0.010)	(0.196)
ϕ_1	0.195***		-0.127	0.034	0.006	
	(0.058)		(0.151)	(0.053)	(0.052)	
ϕ_2			-0.617***	-0.034	-0.034	
			(0.142)	(0.052)	(0.051)	
ϕ_3			0.131*	0.117*	0.108*	
			(0.064)	(0.051)	(0.048)	
θ_1		0.098*	0.209			
		(0.056)	(0.147)			
θ_2			0.600***			
			(0.142)			
α_0	0.000	0.001	0.002	0.003	0.004	13.275***
	(0.000)	(0.001)	(0.002)	(0.002)	(0.003)	(1.813)
α_1	0.255**	0.083*	0.030	0.061	0.118*	0.053
	(0.083)	(0.038)	(0.024)	(0.042)	(0.059)	(0.070)
β_1	0.781***	0.907***	0.930***	0.873***	0.777***	
	(0.055)	(0.043)	(0.488)	(0.090)	(0.113)	
skew	0.949***	1.121***	1.104***	1.002***	0.899***	1.046***
	(0.077)	(0.086)	(0.100)	(0.091)	(0.087)	(0.079)
df	5.938***	10**	10**	10**	10**	5.34***
	(1.717)	(3.907)	(3.467)	(3.548)	(3.766)	(1.482)
Observations	350	350	350	350	350	350
Log Likelihood	0.478	0.200	0.259	0.268	0.266	-2.696
Akaike Inf. Crit.	−0.917	−0.360	−0.455	−0.484	−0.480	5.421
Note:			*p<0.1; **p<0.05; ***p<0.01			

To determine if the standardized residuals follow a white noise process, the ACF and PACF plots must also be considered, which are presented in Appendix B for both the standardized and squared standardized residuals.

As in the first case study, the skew Student's t innovation do not sufficiently cover the standardized residuals. Therefore, it is necessary to determine more appropriate univariate parametric distributions. For the yield curve with a maturity of one year, the skew t type 4 distribution was chosen. For maturities of 5, 10 and 15 years, the normal distribution was selected and for 20 years, the generalized t distribution. The logistic distribution was found to be more suitable for the inflation rate than the selected innovation. Figure B.5 compares QQ-plots of the selected distributions and the innovation.
The standardized residuals were converted to copula data with the selected distributions and the pairs plot is displayed in Figure 9.3 with the estimated Kendall's tau values above, contour plots below and histograms on the diagonal.

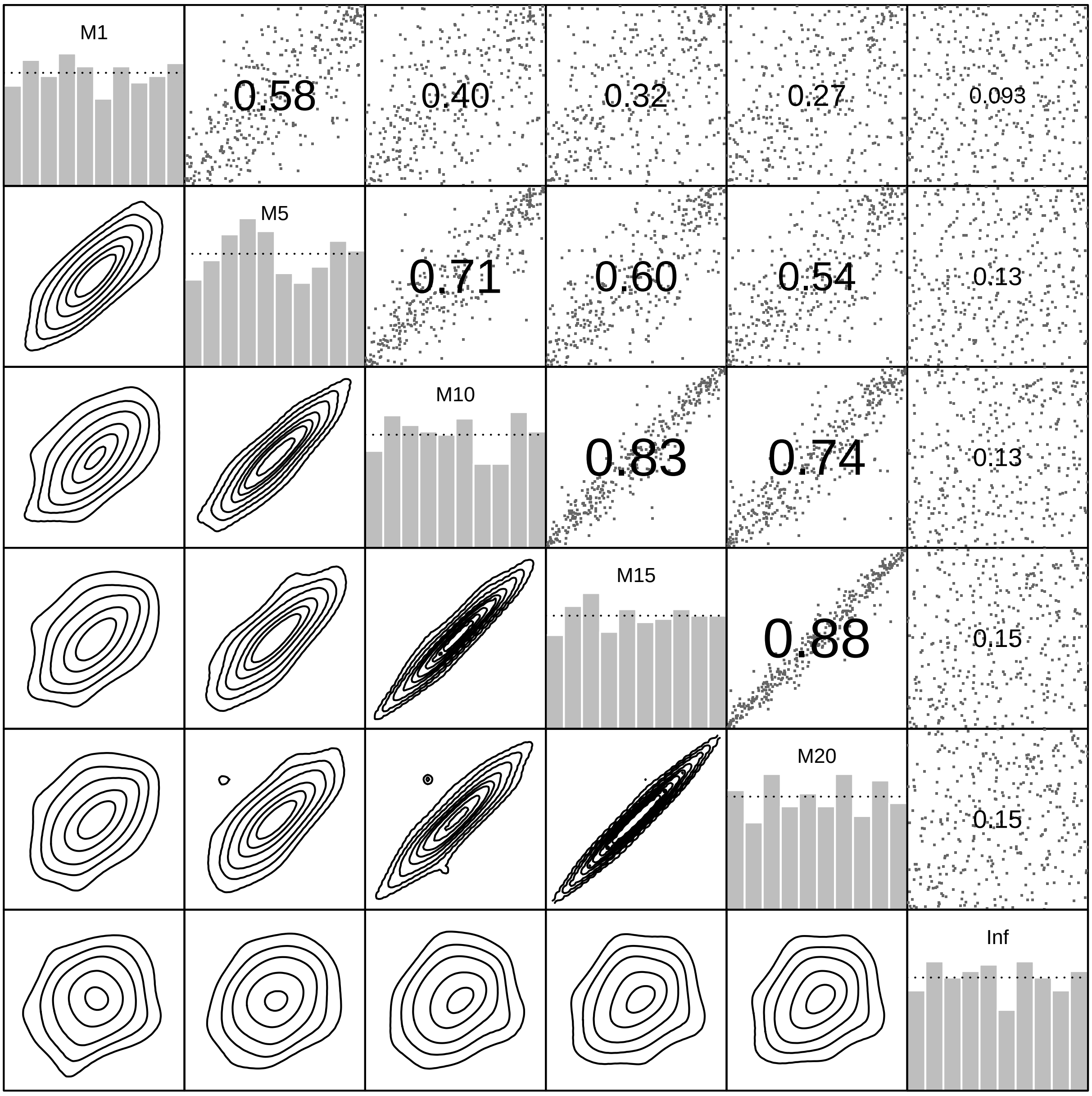

Figure 9.3 Pairs copula for the yield curves for maturities of 1, 5, 10, 15 and 20 year(s) and inflation rate

The pair plot shows that the dependency measure τ between the variable Inf and the yield curves is less than 20%, while it is greater for the other cases.

9.4 Vine copula fitting

This section presents the fitted R-, C- and D-vine copula models. As the dependence measure tau was close to or equal to zero for several trees, these trees were truncated and the two-truncated R-vine, the three-truncated C-vine and the D-vine are additionally considered.

Regular vine copula model

The yield curves and inflation rate were modelled using the regular vine copula model. The results are summarised in Table 9.5. The R-vine has the structure of a D-vine, as shown by the tree structure for the first tree. The first tree uses Gaussian, Student's t, Gumbel and BB7 as pair copulas. The dependency measure tau is non-zero for trees one and two and close to or equal to zero for trees three to five. In the next case, the trees with tau equal to zero were truncated.

Table 9.5 Fitted R-vine copula model for the differenced yield curves and inflation rate

tree	conditioned	conditioning	family	rotation	parameters	tau
1	1, 2		gaussian	0	0.77	0.562
1	2, 3		gaussian	0	0.89	0.693
1	3, 4		bb7	0	5.3, 5.3	0.789
1	4, 5		t	0	0.98, 5.00	0.878
1	5, 6		gumbel	180	1.2	0.136
2	1, 3	2	gaussian	0	-0.46	-0.305
2	2, 4	3	frank	0	-4.8	-0.447
2	3, 5	4	frank	0	-4.6	-0.433
2	4, 6	5	indep	0		0.000
3	1, 4	3, 2	indep	0		0.000
3	2, 5	4, 3	t	0	0.059, 4.263	0.038
3	3, 6	5, 4	indep	0		0.000
4	1, 5	4, 3, 2	indep	0		0.000
4	2, 6	5, 4, 3	indep	0		0.000
5	1, 6	5, 4, 3, 2	indep	0		0.000

The study considers the two-truncated R-vine copula model and Table 9.6 presents a summary of the results. Furthermore, the tree plots for the R-vine are shown in Figure 9.4.

Table 9.6 Fitted two-truncated regular vine copula model for the differenced yield curves and inflation rate

tree	conditioned	conditioning	family	rotation	parameters	tau
1	1, 2		gaussian	0	0.77	0.56
1	2, 3		gaussian	0	0.89	0.69
1	3, 4		bb7	0	5.3, 5.3	0.79
1	4, 5		t	0	0.98, 5.00	0.88
1	5, 6		gumbel	180	1.2	0.14
2	1, 3	2	gaussian	0	-0.46	-0.30
2	2, 4	3	frank	0	-4.8	-0.45
2	3, 5	4	frank	0	-4.6	-0.43
2	4, 6	5	indep	0		0.00

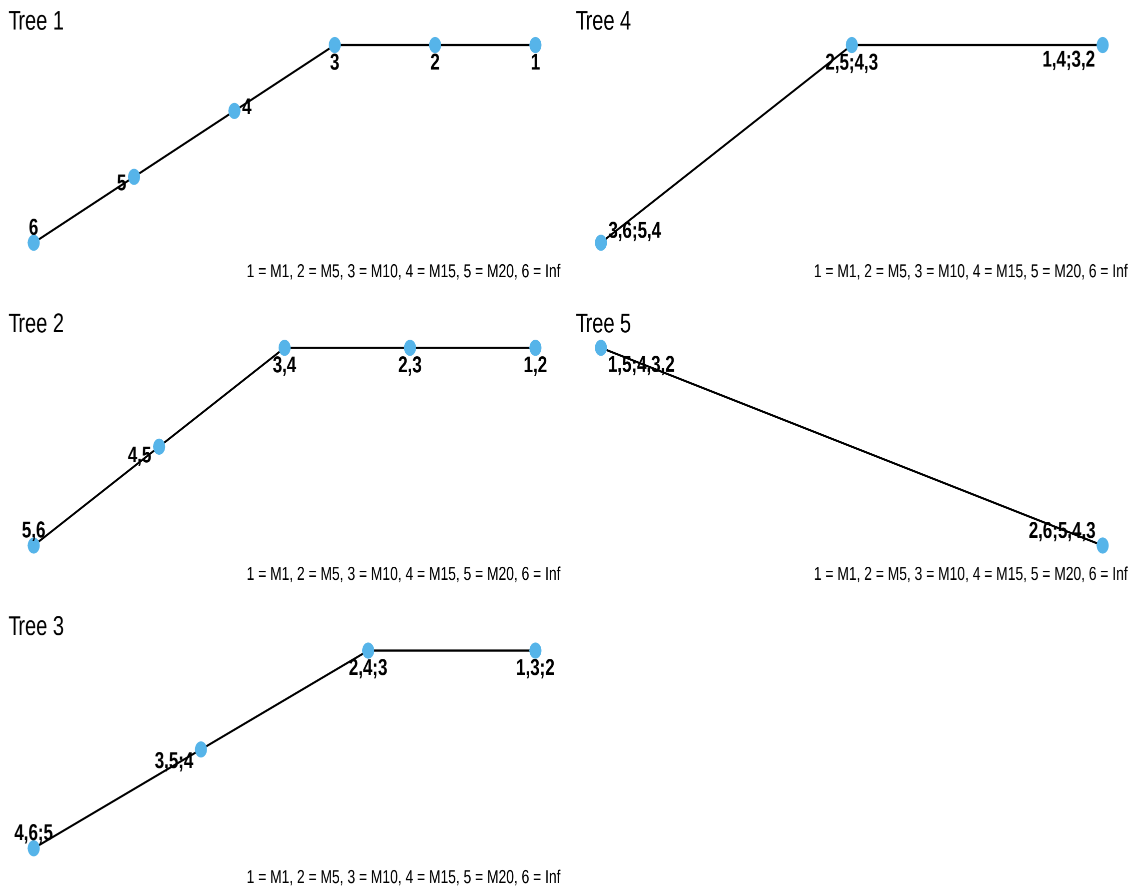

Figure 9.4 Fitted R-vine copula tree plot for the differenced yield curves and inflation rate

Canonical vine copula model

Table 9.7 presents the results of the canonical vine copula model fit. The first tree's root node is variable 3, which represents the yield curve with a maturity of 10 years. Trees 4 and 5 have a dependence measure close to or equal to zero, and therefore, they are truncated. The resulting model is displayed in Table 9.8.

Table 9.7 Fitted C-vine copula model for the differenced yield curves and inflation rate

tree	conditioned	conditioning	family	rotation	parameters	tau
1	1, 3		frank	0	4.2	0.399
1	6, 3		clayton	0	0.26	0.115
1	5, 3		t	0	0.91, 5.20	0.734
1	2, 3		gaussian	0	0.89	0.693
1	4, 3		bb7	0	5.3, 5.3	0.789
2	1, 4	3	frank	0	-2.8	-0.293
2	6, 4	3	indep	0		0.000
2	5, 4	3	frank	0	13	0.730
2	2, 4	3	frank	0	-4.8	-0.447
3	1, 2	4, 3	gaussian	0	0.63	0.438
3	6, 2	4, 3	gumbel	0	1.1	0.056
3	5, 2	4, 3	indep	0		0.000
4	1, 5	2, 4, 3	joe	270	1.1	-0.032
4	6, 5	2, 4, 3	indep	0		0.000
5	1, 6	5, 2, 4, 3	indep	0		0.000

Table 9.8 Fitted three-truncated canonical vine copula model for the differenced yield curves and inflation rate

tree	conditioned	conditioning	family	rotation	parameters	tau
1	1, 3		frank	0	4.2	0.399
1	6, 3		clayton	0	0.26	0.115
1	5, 3		t	0	0.91, 5.20	0.734
1	2, 3		gaussian	0	0.89	0.693
1	4, 3		bb7	0	5.3, 5.3	0.789
2	1, 4	3	frank	0	-2.8	-0.293
2	6, 4	3	indep	0		0.000
2	5, 4	3	frank	0	13	0.730
2	2, 4	3	frank	0	-4.8	-0.447
3	1, 2	4, 3	gaussian	0	0.63	0.438
3	6, 2	4, 3	gumbel	0	1.1	0.056
3	5, 2	4, 3	indep	0		0.000

Drawable vine copula model

The drawable vine copula model in Table 9.9 shows dependencies between -0.37 and 0.88 for trees 1 to 3. For the first tree, the Gaussian and Student's t copulas were selected as the pair copulas in the D-vine model.

Table 9.9 Fitted D-vine copula model for the differenced yield curves and inflation rate

tree	conditioned	conditioning	family	rotation	parameters	tau
1	4, 5		t	0	0.98, 5.00	0.878
1	5, 3		t	0	0.91, 5.20	0.734
1	3, 2		gaussian	0	0.89	0.693
1	2, 1		gaussian	0	0.77	0.562
1	1, 6		gaussian	0	0.15	0.094
2	4, 3	5	t	0	0.85, 4.51	0.641
2	5, 2	3	bb8	270	2.64, 0.89	-0.369
2	3, 1	2	gaussian	0	-0.46	-0.305
2	2, 6	1	clayton	0	0.22	0.101
3	4, 2	3, 5	t	0	-0.34, 5.21	-0.220
3	5, 1	2, 3	joe	90	1.1	-0.033
3	3, 6	1, 2	indep	0		0.000
4	4, 1	2, 3, 5	indep	0		0.000
4	5, 6	1, 2, 3	gumbel	180	1.1	0.071
5	4, 6	1, 2, 3, 5	indep	0		0.000

Additionally, the three-truncated D-vine model was considered due to a tau equal or close to zero for trees 4 and 5.

Table 9.10 Fitted three-truncated drawable vine copula model for the differenced yield curves and inflation rate

tree	conditioned	conditioning	family	rotation	parameters	tau
1	4, 5		t	0	0.98, 5.00	0.878
1	5, 3		t	0	0.91, 5.20	0.734
1	3, 2		gaussian	0	0.89	0.693
1	2, 1		gaussian	0	0.77	0.562
1	1, 6		gaussian	0	0.15	0.094
2	4, 3	5	t	0	0.85, 4.51	0.641
2	5, 2	3	bb8	270	2.64, 0.89	-0.369
2	3, 1	2	gaussian	0	-0.46	-0.305
2	2, 6	1	clayton	0	0.22	0.101
3	4, 2	3, 5	t	0	-0.34, 5.21	-0.220
3	5, 1	2, 3	joe	90	1.1	-0.033
3	3, 6	1, 2	indep	0		0.000

9.5 Factor copula models on the copula scale

This section fits the factor copula models, focusing on the one- and two-factor copula models with arbitrary pair copulas and the Gaussian pair copula.

9.5.1 One-factor arbitrary copula

For the yield curve with a maturity of 1 year, the Student's t copula with five degrees of freedom was chosen as the pair copula in the one-factor copula model. The Student's t copula with three degrees of freedom was used as the pair copula for the yield curve with maturities of 5, 10 and 20 years, the Gaussian copula for the maturity of 15 years and the reflected Gumbel copula for the inflation rate. Upon examining the dependency measure, it is evident that the inflation rate exhibits the lowest dependency measure when compared to the other variables.

Table 9.11 Fitted one-factor copula model with arbitrary pair copulas for the differenced yield curves and inflation rate

variables	family	parameters	tau
M1	bvt5	0.47	0.31
M5	bvt3	0.78	0.57
M10	bvt3	0.96	0.81
M15	bvn	1.00	0.95
M20	bvt3	0.98	0.87
Inf	rgum	1.17	0.14

9.5.2 Two-factor arbitrary copula

The two-factor copula model utilises the Student's t, Frank, and Gaussian copulas as pair copulas. In this instance, the inflation rate exhibits the lowest dependence measure.

Table 9.12 Fitted two-factor copula model with arbitrary pair copulas for the differenced yield curves and inflation rate

variables	family	parameters	tau
M1	frk	1.99	0.21
M5	bvt5	0.67	0.47
M10	bvt3	0.89	0.70
M15	bvn	0.98	0.87
M20	bvt3	0.98	0.87
Inf	bvt5	0.22	0.14
M1	bvn	0.80	0.59
M5	bvn	0.97	0.84
M10	bvt3	0.93	0.77
M15	bvt3	0.97	0.84
M20	frk	4.60	0.43
Inf	bvn	0.07	0.05

9.5.3 One-factor Gaussian copula

Consider the one-factor copula model with the Gaussian copula as the pair copula. Table 9.13 summarizes the results of the one-factor copula model. The dependence measure for the inflation rate is the same as that of the one-factor copula model with arbitrary pair copulas.

Table 9.13 Fitted one-factor copula model with Gaussian pair copulas for the differenced yield curves and inflation rate

variables	family	parameters	tau
M1	bvn	0.46	0.30
M5	bvn	0.80	0.59
M10	bvn	0.96	0.82
M15	bvn	1.00	0.94
M20	bvn	0.98	0.87
Inf	bvn	0.22	0.14

9.5.4 Two-factor Gaussian copula

The two-factor copula model also contains the Gaussian copula as a pair copula. The results are consistent with previous cases. The inflation rate and yield curve with a one-year maturity exhibit the smallest dependency measure for the first factor. For the second factor, only the inflation rate shows a small dependency measure.

Table 9.14 Fitted two-factor copula model with Gaussian pair copulas for the differenced yield curves and inflation rate

variables	family	parameters	tau
M1	bvn	0.27	0.17
M5	bvn	0.63	0.43
M10	bvn	0.87	0.68
M15	bvn	0.97	0.83
M20	bvn	0.97	0.85
Inf	bvn	0.20	0.13
M1	bvn	0.82	0.62
M5	bvn	0.97	0.84
M10	bvn	0.92	0.75
M15	bvn	0.98	0.88
M20	bvn	0.66	0.46
Inf	bvn	0.10	0.06

9.6 Gaussian factor models on the z-scale

The results of the factor Gaussian model, both one-factor and two-factor, on the z-scale are presented below.

9.6.1 One-factor Gaussian model

The one-factor Gaussian model contains one latent factor and is given by

$$\begin{aligned}
\mathrm{M1} &= 0.45 f_1 + u_{\mathrm{M1}} \\
\mathrm{M5} &= 0.79 f_1 + u_{\mathrm{M5}} \\
\mathrm{M10} &= 0.96 f_1 + u_{\mathrm{M10}} \\
\mathrm{M15} &= 1.00 f_1 + u_{\mathrm{M15}} \\
\mathrm{M20} &= 0.98 f_1 + u_{\mathrm{M20}} \\
\mathrm{Inf} &= 0.21 f_1 + u_{\mathrm{Inf}},
\end{aligned}$$

where $u_i, i = \mathrm{M1}, \mathrm{M5}, \mathrm{M10}, \mathrm{M15}, \mathrm{M20}, \mathrm{Inf}$ are normally distributed with variance $\psi = \mathrm{diag}(0.80, 0.37, 0.08, 0.01, 0.04, 0.96)$ and

$$\Sigma = \begin{pmatrix}
1.00 & & & & & \\
0.36 & 1.00 & & & & \\
0.43 & 0.76 & 1.00 & & & \\
0.45 & 0.79 & 0.96 & 1.00 & & \\
0.44 & 0.78 & 0.94 & 0.98 & 1.00 & \\
0.09 & 0.16 & 0.20 & 0.21 & 0.20 & 1.00
\end{pmatrix}$$

represents the sample correlation matrix for the variables **M1**, **M5**, .., **Inf**.

9.6.2 Two-factor Gaussian model

The two-factor model incorporates an additional factor, f_2, alongside the latent factor, f_1. The model is defined by the following equations.

$$\begin{aligned}
\mathrm{M1} &= 0.19 f_1 + 0.80 f_2 + u_{\mathrm{M1}} \\
\mathrm{M5} &= 0.55 f_1 + 0.83 f_2 + u_{\mathrm{M5}} \\
\mathrm{M10} &= 0.84 f_1 + 0.51 f_2 + u_{\mathrm{M10}} \\
\mathrm{M15} &= 0.94 f_1 + 0.33 f_2 + u_{\mathrm{M15}} \\
\mathrm{M20} &= 0.96 f_1 + 0.23 f_2 + u_{\mathrm{M20}} \\
\mathrm{Inf} &= 0.18 f_1 + 0.13 f_2 + u_{\mathrm{Inf}}
\end{aligned}$$

The sample correlation matrix for variables **M1** to **M20** and **Inf** is represented by

$$\Sigma = \begin{pmatrix}
1.00 & & & & & \\
0.77 & 1.00 & & & & \\
0.57 & 0.89 & 1.00 & & & \\
0.44 & 0.79 & 0.96 & 1.00 & & \\
0.37 & 0.72 & 0.92 & 0.98 & 1.00 & \\
0.13 & 0.20 & 0.21 & 0.21 & 0.20 & 1.00
\end{pmatrix}.$$

9.7 Multivariate t copula in 6 dimensions

This section presents another copula model, which fits the Student's t copula in six dimensions. The estimated correlation coefficients are presented in the matrix below

$$\Sigma = \begin{pmatrix} 1.00 & 0.77 & 0.56 & 0.44 & 0.39 & 0.13 \\ 0.77 & 1.00 & 0.90 & 0.80 & 0.74 & 0.19 \\ 0.56 & 0.90 & 1.00 & 0.97 & 0.92 & 021 \\ 0.44 & 0.80 & 0.97 & 1.00 & 0.98 & 0.23 \\ 0.39 & 0.74 & 0.92 & 0.98 & 1.00 & 0.24 \\ 0.13 & 0.19 & 0.21 & 0.23 & 0.24 & 1.00 \end{pmatrix}$$

and the degrees of freedom are $\nu = 9.5$.

The following section compares the models that have been fitted so far.

9.8 Model comparison

The previously fitted models are compared to determine the best model. The selection criteria used were AIC, BIC and log-likelihood.

Table 9.15 Model comparison

Models	AIC	BIC	logLik	Number of parameters
RVine (BIC)	-3455.30	-3409.01	1739.65	12
RVine two truncated	-3423.77	-3385.19	1721.88	10
CVine (BIC) with order 165243	-3316.69	-3266.54	1671.35	13
CVine three truncated	-3312.61	-3266.32	1668.31	12
DVine (BIC) with order 453216	-3353.28	-3287.70	1693.64	17
DVine two truncated	-3347.01	-3285.29	1689.51	16
One-Factor Copula	-2392.80	-2365.79	1203.40	10
Two-Factor Copula	-3109.77	-3055.76	1568.89	18
Gaussian One-Factor Copula Model	-2342.22	-2322.93	1176.11	6
Gaussian Two-Factor Copula Model	-3005.97	-2967.39	1512.99	12
Gaussian One-Factor Model transformed to u-scale	-2377.14	-2330.85	1200.57	21
Gaussian Two-Factor Model transformed to u-scale	-3036.62	-2990.32	1530.31	27
t Copula	-3327.07	-3265.34	1679.53	16

The table indicates that the R-vine copula model is the best among those considered. The R vine model is closely followed by the two truncated R-vine and D-vine copula models.

9.9 Forecasting

The same approach as in the first case (Section 8.10) is used for the forecasting. The next three months were also predicted for this case. The parameters required for the forecast are given in Appendix Section B. Now we present the results from the prediction using the R-vine copula model.

Figure 9.5 displays the density, as well as the 5% and 95% quantiles of the predicted 1000 samples. The variables M1, M5, M10, M15, M20 and Inf are shown in the rows, while the time points $T + 1 = 352, T + 2 = 353$ and $T + 3 = 354$ are shown in the columns. The graphs show clearly that the observed value is included in the generated samples. Furthermore, this value lies between the two quantiles. In general, the forecasts cover the observed values well.

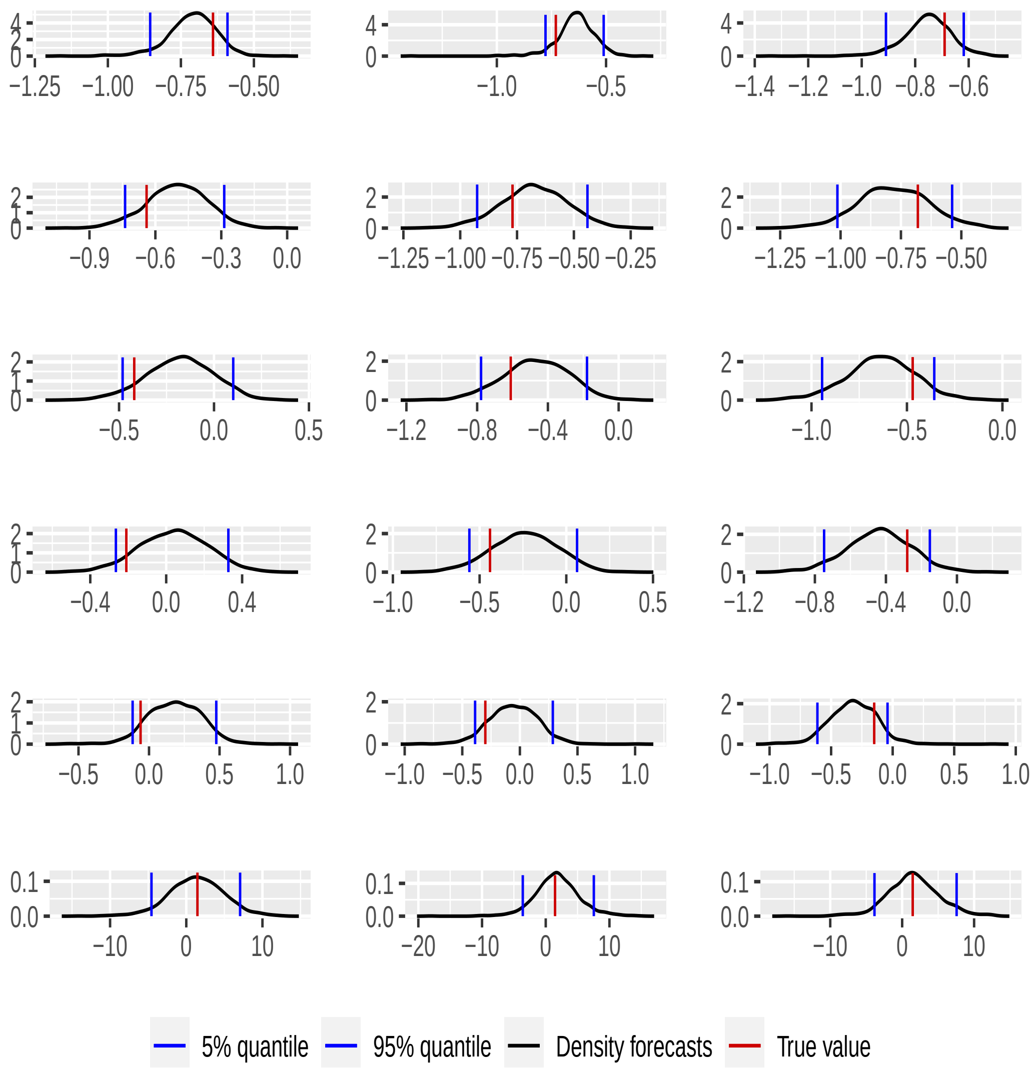

Figure 9.5 The R-vine copula is used to forecast yield curves and inflation rate. The maturities M1, M5, M10, M15, M20 and Inf are presented in rows and the predicted times $T + 1 = 352, T + 2 = 353, T + 3 = 354$ are presented in columns.

In addition to using R-vines for prediction, C- and D-vines, as well as one- and two-factor copula models, were also employed. These models produced comparable results to the R-vine copula model. The prediction graphs can be found in Appendix B. In addition to the graphical interpretation, the score rules were also considered. The results for the CRPS are summarized in Table 9.16 and for the score interval in Table 9.17.

Since several models sometimes produced the same CRPS, no clear best model could be determined. As in the first case study, the sum was calculated for each model to determine the best one. It is important to note that the variables are not independent, so the sum cannot be simply calculated. The R-vine copula model was the best predictor for the first month, whereas the C-vine copula model was the best predictor for the second and third month.

Table 9.16 Continuous ranked probability score for the forecasted yield curves and inflation rate

	$T+1=352$					$T+2=353$					$T+3=354$				
Model	RVine	CVine	DVine	1F	2F	RVine	CVine	DVine	1F	2F	RVine	CVine	DVine	1F	2F
M1	0.04	0.04	0.04	0.04	0.04	0.06	0.06	0.06	0.06	0.06	0.04	0.04	0.04	0.04	0.04
M5	0.09	0.09	0.08	0.08	0.08	0.05	0.05	0.06	0.05	0.05	0.06	0.06	0.07	0.06	0.06
M10	0.16	0.16	0.16	0.16	0.15	0.08	0.08	0.08	0.08	0.08	0.10	0.11	0.11	0.10	0.11
M15	0.17	0.17	0.16	0.16	0.16	0.12	0.12	0.12	0.12	0.12	0.09	0.10	0.10	0.09	0.10
M20	0.16	0.16	0.16	0.16	0.15	0.15	0.12	0.13	0.13	0.12	0.10	0.11	0.11	0.10	0.11
Inf	0.73	0.76	0.74	0.77	0.74	0.79	0.76	0.77	0.75	0.77	0.79	0.73	0.79	0.78	0.75
$\sum$	1.19	1.38	1.34	1.37	1.32	1.25	1.19	1.22	1.19	1.20	1.18	1.15	1.22	1.17	1.17

Similar to CRPS, IS also indicated that the C-vine copula model was the best predictor for the second and third month. The two-factor copula model forecasted the first month best.

Table 9.17 Interval score for the forecasted yield curves and inflation rate

	$T+1=352$					$T+2=353$					$T+3=354$				
Model	RVine	CVine	DVine	1F	2F	RVine	CVine	DVine	1F	2F	RVine	CVine	DVine	1F	2F
M1	0.26	0.27	0.26	0.28	0.26	0.26	0.26	0.27	0.28	0.27	0.29	0.28	0.29	0.28	0.30
M5	0.43	0.45	0.44	0.41	0.44	0.46	0.44	0.45	0.48	0.46	0.44	0.45	0.46	0.44	0.46
M10	0.54	0.59	0.55	0.54	0.55	0.56	0.55	0.59	0.60	0.55	0.58	0.60	0.59	0.57	0.60
M15	0.58	0.61	0.57	0.57	0.57	0.58	0.56	0.61	0.62	0.57	0.60	0.60	0.60	0.61	0.63
M20	0.57	0.60	0.58	0.57	0.56	0.63	0.61	0.64	0.66	0.62	0.61	0.59	0.59	0.60	0.64
Inf	11.25	11.58	11.51	11.81	11.10	12.48	10.73	12.05	11.66	11.98	12.00	11.02	12.07	11.58	11.34
$\sum$	13.63	14.10	13.91	14.18	13.48	14.97	13.15	14.61	14.30	14.45	14.52	13.54	14.60	14.08	13.97

As the CRPS has identified the C-vine copula model as the best model and IS has used the canonical vine copula model for the last two months and the two-factor copula model for the first month, and the C-vine is only just behind, we identify the C-vine copula model as the best model.

10 Conclusion

The goal of this thesis was the modelling of the dependencies and the forecasting of yield curves and economic factors using vine and factor copula models. Two case studies were conducted: the first involved modelling and predicting German yield curves, while the second involved modelling and forecasting yield curves and the German inflation rate. Prior to modelling, the data had to be transformed to the copula scale. This required removing serial dependencies using the ARMA-GARCH model. The standardized residuals were derived from the ARMA-GARCH model and transformed to copula data using the probability integral transform and the selected univariate distributions. The dependence structure was modelled using various copulas, including R-, C-, and D-vine copulas, one- and two-factor copulas, the Student's t copula and the one- and two-factor Gaussian model. Additionally, truncated vine copula models were considered in the second case study. The best model was then selected using the selection criteria AIC, BIC and log likelihood. The D-vine copula was found to be the best model for the first study case, while the R-vine copula was the best model for the second study case. The vine copula and factor copula models were used in the next step to generate 1000 samples for the next three months, $T+1, T+2$ and $T+3$. The prediction procedure involved several steps. Firstly, the parameters, $\hat{\mu}$ and $\hat{\sigma}$, were estimated from the ARMA-GARCH model for the three time points. The yield curves and inflation rate predictions were determined using the estimated parameters and the transformed samples. To predict time points $T+2$ and $T+3$, the yield curve and inflation rate data were shifted forward by one and two months respectively and used for modelling and forecasting. After predicting all three months, the observed values were compared with the predicted samples. Graphs and tables presenting the results were provided. The observed values for the next three months were well covered by the predicted samples, as they fell between the 5% and 95% quantile. Finally, the best model was determined by CRPS and IS, which identified the D-vine copula model for the first case study and the C-vine copula model for the second case study.

A Appendix - Yield Curves

This appendix presents the results of the time series analysis and modelling of dependencies. The section is organised as follows: first, relevant results and graphs from the time series analysis are presented, followed by the results from the modelling of dependencies.

Marginal ARMA fitting

This section presents the autocorrelation function (ACF) and the partial autocorrelation function (PACF), as well as the squared ACF and PACF of the standardized residuals from the ARMA model discussed in Section 8.3.

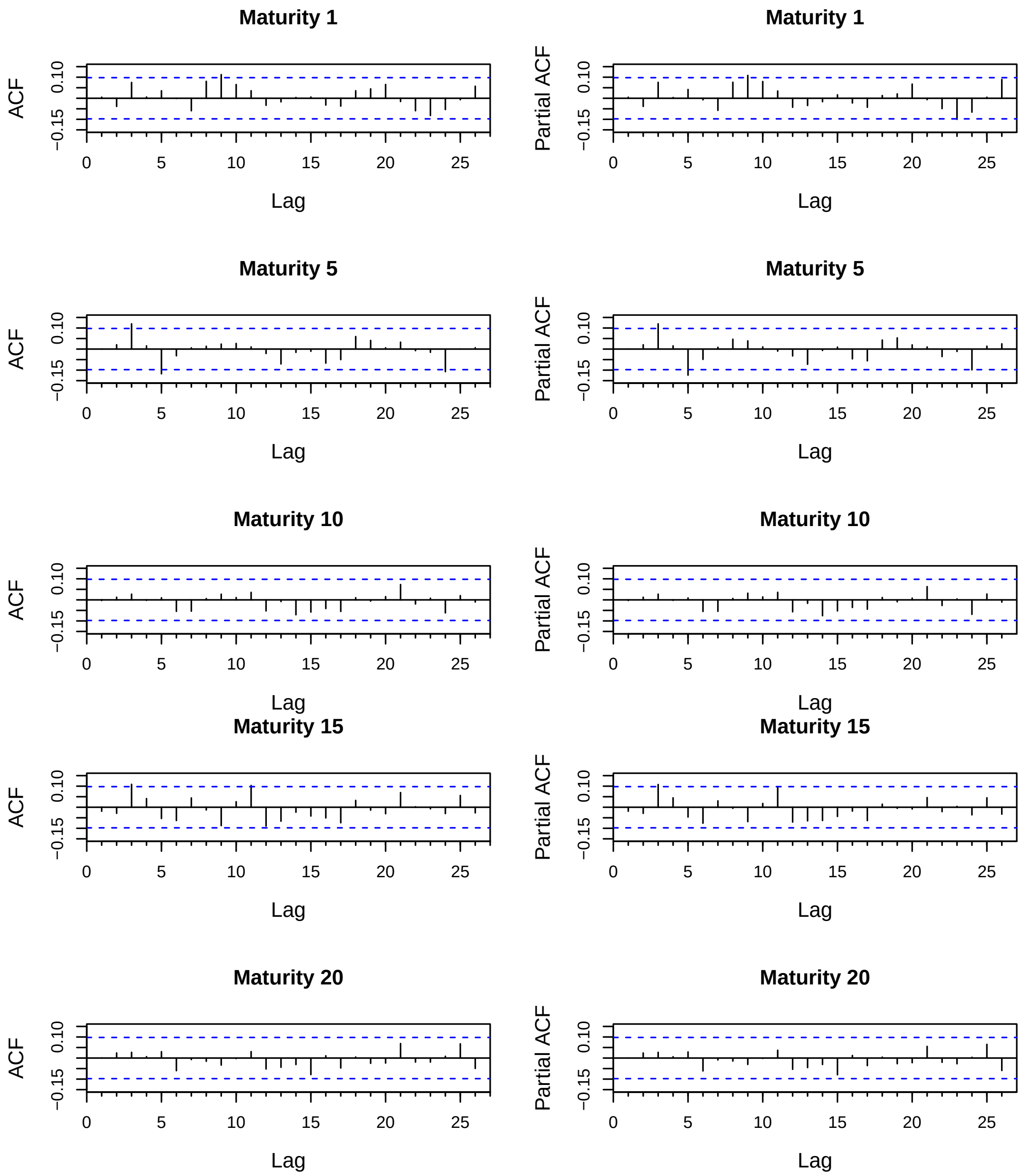

Figure A.1 Estimated ACF and PACF of the standardized residuals obtained from the ARMA models fitted in Table 8.2.

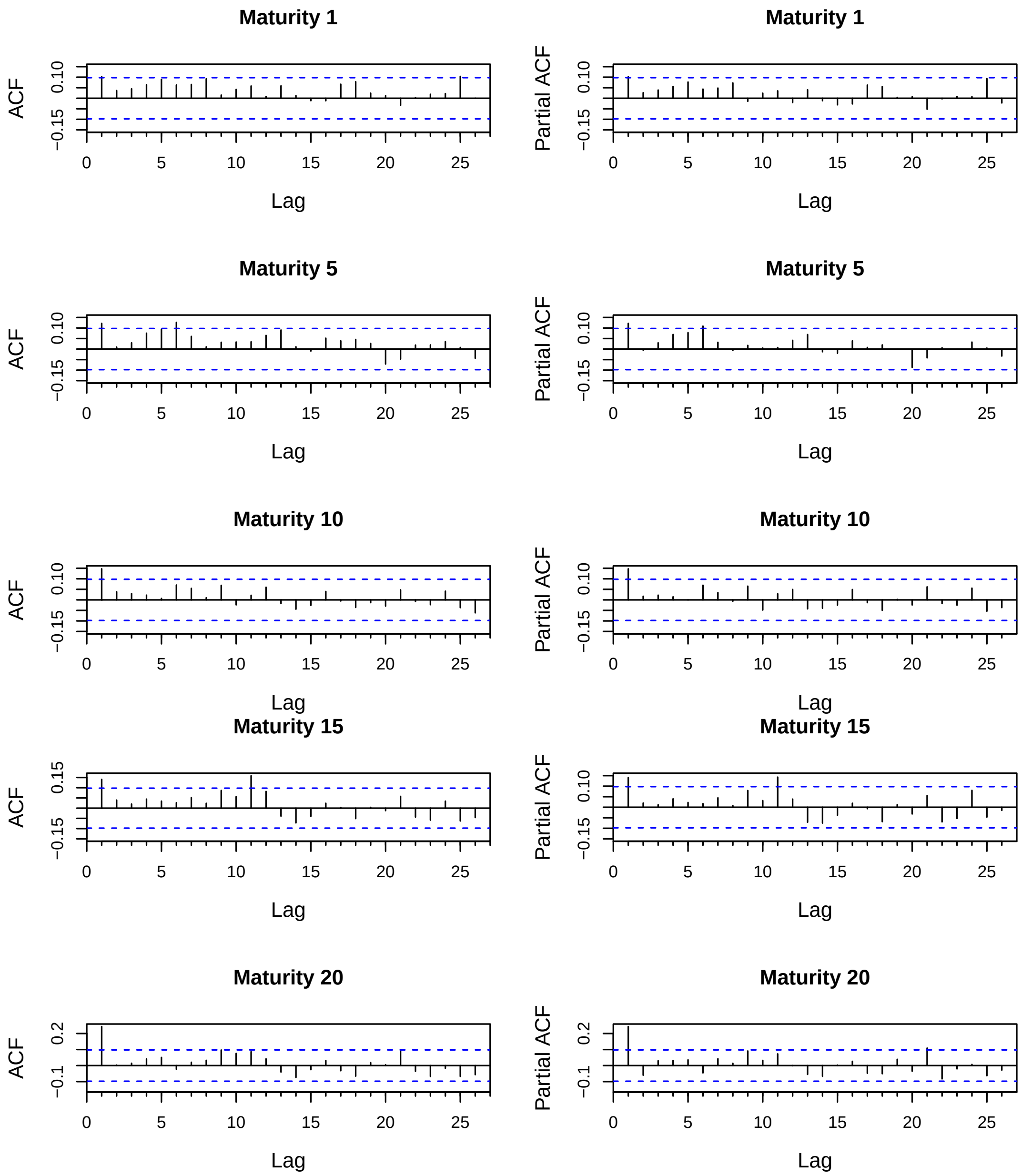

Figure A.2 Estimated ACF and PACF of the squared standardized residuals obtained from the ARMA models fitted in Table 8.2.

Marginal ARMA-GARCH fitting

This section presents the results of the ARMA-GARCH model. The fitted ARMA-GARCH models are presented first, followed by the ACF and PACF of the standardized residuals and the squared standardized residuals.

1. Yield curve with maturity 1 year

   ARMA(1,0)-GARCH(1,1) model

```
## Mean and Variance Equation:
##  data ~ arma(1, 0) + garch(1, 1)
##
## Conditional Distribution:
##  sstd
##
## Std. Errors:
##  based on Hessian
##
## Error Analysis:
##          Estimate  Std. Error  t value Pr(>|t|)
## mu     -0.0075582   0.0057966   -1.304  0.19227
## ar1     0.2102283   0.0536721    3.917 8.97e-05 ***
## omega   0.0002301   0.0002057    1.118  0.26338
## alpha1  0.2229789   0.0753065    2.961  0.00307 **
## beta1   0.8083600   0.0500170   16.162  < 2e-16 ***
## skew    0.9536537   0.0754969   12.632  < 2e-16 ***
## shape   5.5895554   1.3946832    4.008 6.13e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
##  153.0226    normalized:  0.3806533
##
##
## Standardised Residuals Tests:
##                                 Statistic p-Value
##  Jarque-Bera Test   R    Chi^2  1011.54   0
##  Shapiro-Wilk Test  R    W      0.932716  1.726717e-12
##  Ljung-Box Test     R    Q(10)  14.70929  0.1430254
##  Ljung-Box Test     R    Q(15)  16.00547  0.3816884
##  Ljung-Box Test     R    Q(20)  20.40283  0.4329977
##  Ljung-Box Test     R^2  Q(10)  1.524733  0.9988562
##  Ljung-Box Test     R^2  Q(15)  2.963248  0.9996274
##  Ljung-Box Test     R^2  Q(20)  4.819588  0.9997919
##  LM Arch Test       R    TR^2   2.172042  0.9990931
##
## Information Criterion Statistics:
##        AIC        BIC        SIC       HQIC
## -0.7264807 -0.6568907 -0.7270734 -0.6989276
```

2. Yield curve with maturity 5 years

   ARMA(0,1)-GARCH(1,1) model

```
##
```

```
## Mean and Variance Equation:
##  data ~ arma(0, 1) + garch(1, 1)
##
## Conditional Distribution:
##  sstd
##
## Std. Errors:
##  based on Hessian
##
## Error Analysis:
##          Estimate  Std. Error  t value Pr(>|t|)
## mu     -0.0119483   0.0107102   -1.116  0.26459
## ma1     0.1188595   0.0512155    2.321  0.02030 *
## omega   0.0005371   0.0007630    0.704  0.48147
## alpha1  0.0771739   0.0363891    2.121  0.03394 *
## beta1   0.9157478   0.0414612   22.087  < 2e-16 ***
## skew    1.1752702   0.0857164   13.711  < 2e-16 ***
## shape  10.0000000   3.3449091    2.990  0.00279 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
##  67.19564    normalized:  0.1671533
##
##
## Standardised Residuals Tests:
##                                 Statistic p-Value
##  Jarque-Bera Test   R    Chi^2  0.7973731 0.6712011
##  Shapiro-Wilk Test  R    W      0.9944849 0.1575729
##  Ljung-Box Test     R    Q(10)  12.01499  0.2840545
##  Ljung-Box Test     R    Q(15)  14.24569  0.5069828
##  Ljung-Box Test     R    Q(20)  20.282    0.4404179
##  Ljung-Box Test     R^2  Q(10)  5.185914  0.8784176
##  Ljung-Box Test     R^2  Q(15)  7.723698  0.9343816
##  Ljung-Box Test     R^2  Q(20)  12.05613  0.9141306
##  LM Arch Test       R    TR^2   4.865108  0.9623332
##
## Information Criterion Statistics:
##        AIC        BIC        SIC       HQIC
## -0.2994808 -0.2298908 -0.3000735 -0.2719277
```

3. Yield curve with maturity 10 years

   ARMA$(3,2)$-GARCH$(1,1)$ model

```
## Mean and Variance Equation:
##  data ~ arma(3, 2) + garch(1, 1)
##
## Conditional Distribution:
##  sstd
##
## Std. Errors:
##  based on Hessian
##
## Error Analysis:
```

```
##         Estimate  Std. Error  t value Pr(>|t|)
## mu     -0.028815    0.018846   -1.529  0.12627
## ar1    -0.034495    0.071671   -0.481  0.63031
## ar2    -0.802648    0.062526  -12.837  < 2e-16 ***
## ar3     0.150640    0.052794    2.853  0.00433 **
## ma1     0.100255    0.054459    1.841  0.06563 .
## ma2     0.814237    0.071389   11.406  < 2e-16 ***
## omega   0.004577    0.004820    0.950  0.34226
## alpha1  0.047681    0.035427    1.346  0.17833
## beta1   0.839091    0.140216    5.984 2.17e-09 ***
## skew    1.115602    0.104046   10.722  < 2e-16 ***
## shape  10.000000    3.454722    2.895  0.00380 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
##  87.13325    normalized:  0.2167494
##
##
## Standardised Residuals Tests:
##                                 Statistic p-Value
##  Jarque-Bera Test   R    Chi^2  11.48007  0.003214648
##  Shapiro-Wilk Test  R    W      0.990448  0.01039652
##  Ljung-Box Test     R    Q(10)  3.231746  0.9754322
##  Ljung-Box Test     R    Q(15)  8.141291  0.9179932
##  Ljung-Box Test     R    Q(20)  10.1318   0.9657185
##  Ljung-Box Test     R^2  Q(10)  7.323457  0.6945963
##  Ljung-Box Test     R^2  Q(15)  10.23191  0.8049041
##  Ljung-Box Test     R^2  Q(20)  12.47884  0.8986096
##  LM Arch Test       R    TR^2   8.65872   0.7317707
##
## Information Criterion Statistics:
##        AIC        BIC        SIC       HQIC
## -0.3787724 -0.2694167 -0.3802174 -0.3354746
```

4. Yield curve with maturity 15 years

GARCH(1, 1) model

```
## Mean and Variance Equation:
##  data ~ garch(1, 1)
##
## Conditional Distribution:
##  sstd
##
## Std. Errors:
##  based on Hessian
##
## Error Analysis:
##         Estimate  Std. Error  t value Pr(>|t|)
## mu     -0.019681    0.009938   -1.980  0.04765 *
## omega   0.004600    0.002944    1.563  0.11812
## alpha1  0.091807    0.046745    1.964  0.04953 *
## beta1   0.800997    0.093474    8.569  < 2e-16 ***
## skew    1.027406    0.082172   12.503  < 2e-16 ***
```

```
## shape  10.000000    3.734823    2.678  0.00742 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
##  75.54121    normalized:  0.1879135
##
##
## Standardised Residuals Tests:
##                                 Statistic p-Value
##  Jarque-Bera Test   R    Chi^2  8.793161  0.0123194
##  Shapiro-Wilk Test  R    W      0.9944189 0.1508464
##  Ljung-Box Test     R    Q(10)  10.06641  0.4346861
##  Ljung-Box Test     R    Q(15)  19.60005  0.1878177
##  Ljung-Box Test     R    Q(20)  24.05066  0.2401863
##  Ljung-Box Test     R^2  Q(10)  2.488234  0.9910461
##  Ljung-Box Test     R^2  Q(15)  11.54151  0.713349
##  Ljung-Box Test     R^2  Q(20)  13.48361  0.8556843
##  LM Arch Test       R    TR^2   6.659633  0.8792628
##
## Information Criterion Statistics:
##       AIC        BIC        SIC       HQIC
## -0.3459762 -0.2863276 -0.3464130 -0.3223592
```

5. Yield curve with maturity 20 years

   ARMA(3,2)-GARCH(1,1) model

```
## Mean and Variance Equation:
##  data ~ arma(3, 2) + garch(1, 1)
##
## Conditional Distribution:
##  sstd
##
## Std. Errors:
##  based on Hessian
##
## Error Analysis:
##          Estimate  Std. Error  t value Pr(>|t|)
## mu     -3.548e-02   1.829e-02   -1.940  0.05238 .
## ar1    -6.409e-02   5.946e-02   -1.078  0.28113
## ar2    -8.527e-01   3.946e-02  -21.608  < 2e-16 ***
## ar3     5.110e-02   5.265e-02    0.971  0.33170
## ma1     4.952e-02   2.883e-02    1.718  0.08585 .
## ma2     8.838e-01   3.971e-02   22.255  < 2e-16 ***
## omega   3.082e-02   1.049e-02    2.937  0.00331 **
## alpha1  2.715e-01   1.053e-01    2.578  0.00995 **
## beta1   1.000e-08   2.519e-01    0.000  1.00000
## skew    9.449e-01   7.549e-02   12.518  < 2e-16 ***
## shape   7.160e+00   2.299e+00    3.115  0.00184 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
##  88.66988    normalized:  0.2205718
```

```
##
##
## Standardised Residuals Tests:
##                                  Statistic p-Value
##  Jarque-Bera Test   R    Chi^2  32.31334  9.621611e-08
##  Shapiro-Wilk Test  R    W      0.9854238 0.000453954
##  Ljung-Box Test     R    Q(10)  5.136804  0.8818539
##  Ljung-Box Test     R    Q(15)  9.463042  0.8520899
##  Ljung-Box Test     R    Q(20)  11.11462  0.9431956
##  Ljung-Box Test     R^2  Q(10)  4.31216   0.9321681
##  Ljung-Box Test     R^2  Q(15)  7.394249  0.9457725
##  Ljung-Box Test     R^2  Q(20)  11.34611  0.9367482
##  LM Arch Test       R    TR^2   3.93698   0.9845469
##
## Information Criterion Statistics:
##        AIC        BIC        SIC       HQIC
## -0.3864173 -0.2770617 -0.3878623 -0.3431196
```

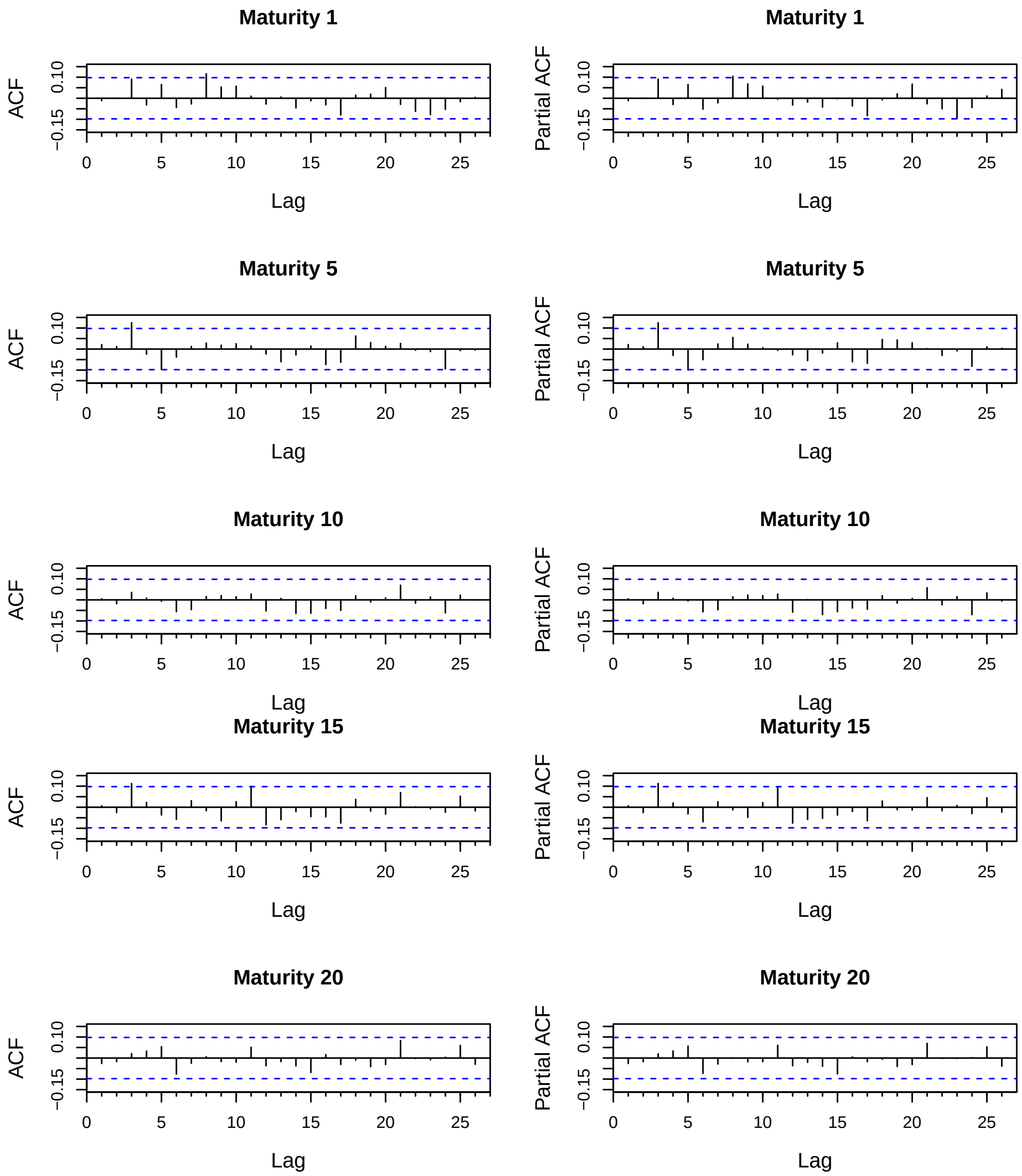

Figure A.3 Estimated ACF and PACF of the standardized residuals obtained from the ARMA-GARCH models fitted in Table 8.4.

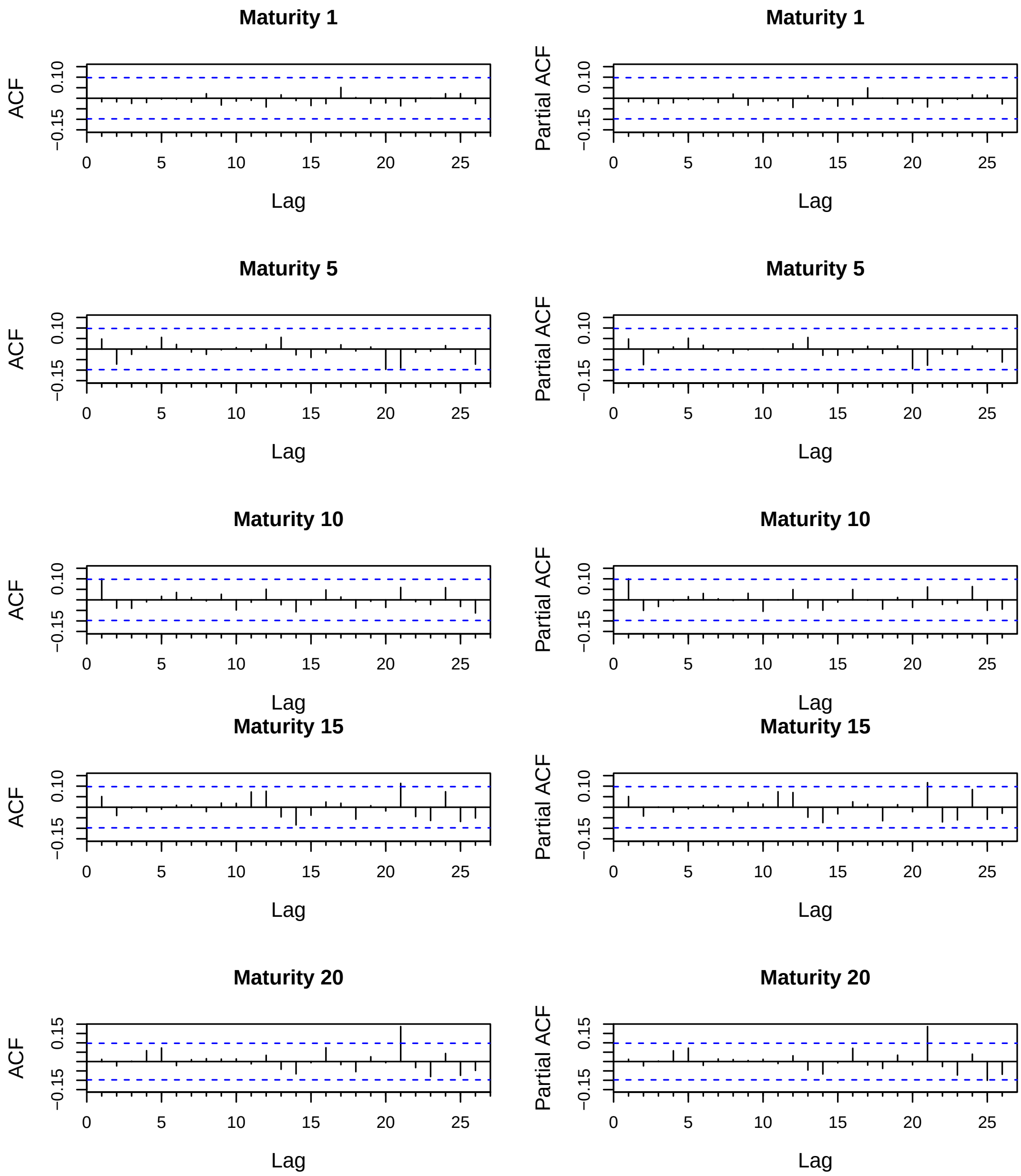

Figure A.4 Estimated ACF and PACF of the squared standardized residuals obtained from the ARMA-GARCH models fitted in Table 8.4.

Forecast

The following tables show the density mode, the prediction interval and the observed value for the R-, C- and D-vine copula and the one- and two-factor copula models.

Table A.1 Density mode and prediction interval of forecasted yield curves using R-, C- and D-Vine copula models and the observed value

Variable	Copula model		$T = 404$	$T = 405$	$T = 406$
		Obs. value	-0.64	-0.73	-0.69
M1	R-vine	Mode	-0.71	-0.62	-0.75
		Prediction interval	[-1.17, -0.47]	[-1.16, -0.39]	[-1.54, -0.46]
	C-vine	Mode	-0.69	-0.65	-0.74
		Prediction interval	[-1.2, -0.47]	[-1.11, -0.34]	[-1.17, -0.49]
	D-vine	Mode	-0.7	-0.63	-0.74
		Prediction interval	[-1.12, -0.39]	[-1.17, -0.36]	[-1.51, -0.41]
		Obs. value	-0.64	-0.77	-0.68
M5	R-vine	Mode	-0.49	-0.74	-0.81
		Prediction interval	[-0.93, -0.11]	[-1.06, -0.22]	[-1.29, -0.39]
	C-vine	Mode	-0.54	-0.65	-0.83
		Prediction interval	[-0.87, 0.04]	[-1.11, -0.24]	[-1.18, -0.36]
	D-vine	Mode	-0.47	-0.69	-0.8
		Prediction interval	[-0.93, -0.01]	[-1.1, -0.23]	[-1.2, -0.3]
		Obs. value	-0.42	-0.61	-0.47
M10	R-vine	Mode	-0.29	-0.55	-0.76
		Prediction interval	[-1.8, 0.66]	[-1.57, 0.33]	[-2.2, 0.41]
	C-vine	Mode	-0.11	-0.34	-0.68
		Prediction interval	[-0.78, 1.02]	[-1.53, 0.26]	[-1.53, 0.08]
	D-vine	Mode	-0.12	-0.5	-0.48
		Prediction interval	[-1.45, 1.3]	[-1.64, 0.4]	[-1.43, 0.15]
		Obs. value	-0.21	-0.44	-0.28
M15	R-vine	Mode	0.02	-0.31	-0.54
		Prediction interval	[-1.12, 0.76]	[-1.08, 0.46]	[-1.56, 0.34]
	C-vine	Mode	0.08	-0.21	-0.44
		Prediction interval	[-0.67, 0.7]	[-0.99, 0.43]	[-1.45, 0.22]
	D-vine	Mode	0.04	-0.18	-0.53
		Prediction interval	[-0.81, 0.82]	[-0.97, 0.64]	[-1.22, 0.47]
		Obs. value	-0.06	-0.30	-0.15
M20	R-vine	Mode	0.2	-0.08	-0.38
		Prediction interval	[-0.82, 0.93]	[-0.93, 0.68]	[-1.21, 0.47]
	C-vine	Mode	0.24	-0.09	-0.32
		Prediction interval	[-0.59, 0.82]	[-0.95, 0.64]	[-1.14, 0.3]
	D-vine	Mode	0.2	-0.11	-0.35
		Prediction interval	[-0.56, 0.99]	[-0.81, 0.75]	[-1.07, 0.53]

Table A.2 Density mode and prediction interval of forecasted yield curves using one- and two-factor copula models and the observed value

Variable	Copula model		$T = 404$	$T = 405$	$T = 406$
		Obs. value	-0.64	-0.73	-0.69
M1	One-factor	Mode	-0.69	-0.65	-0.73
		Prediction interval	[-1.32, -0.31]	[-1.03, -0.3]	[-1.26, -0.45]
	Two-factor	Mode	0.02	-0.17	-0.49
		Prediction interval	[-0.74, 0.77]	[-0.97, 1]	[-1.29, 0.36]
		Obs. value	-0.64	-0.77	-0.68
M5	One-factor	Mode	-0.52	-0.72	-0.82
		Prediction interval	[-0.87, -0.03]	[-1.18, -0.2]	[-1.12, -0.34]
	Two-factor	Mode	-0.5	-0.68	-0.82
		Prediction interval	[-0.88, -0.07]	[-1.15, -0.28]	[-1.25, -0.4]
		Obs. value	-0.42	-0.61	-0.47
M10	One-factor	Mode	-0.19	-0.56	-0.63
		Prediction interval	[-2.28, 1.53]	[-1.22, 0.66]	[-1.94, 0.88]
	Two-factor	Mode	-0.16	-0.56	-0.73
		Prediction interval	[-1.05, 0.73]	[-1.25, 1.43]	[-1.52, 0.32]
		Obs. value	-0.21	-0.44	-0.28
M15	One-factor	Mode	0.05	-0.27	-0.48
		Prediction interval	[-1.43, 1.28]	[-0.98, 0.63]	[-1.19, 0.88]
	Two-factor	Mode	0.02	-0.17	-0.49
		Prediction interval	[-0.74, 0.77]	[-0.97, 1]	[-1.29, 0.36]
		Obs. value	-0.06	-0.30	-0.15
M20	One-factor	Mode	0.21	-0.09	-0.38
		Prediction interval	[-0.8, 1.05]	[-0.83, 0.67]	[-1.16, 0.67]
	Two-factor	Mode	0.19	-0.07	-0.32
		Prediction interval	[-0.68, 0.88]	[-0.87, 0.94]	[-1.08, 0.59]

B Appendix - Yield Curves and Inflation Rate

This appendix section presents the results of modelling and forecasting yield curves and inflation rate.

Marginal ARMA fitting

The ARMA model results are presented in this section. The autocorrelation function and partial autocorrelation function of the standardized residuals and squared standardized residuals are displayed.

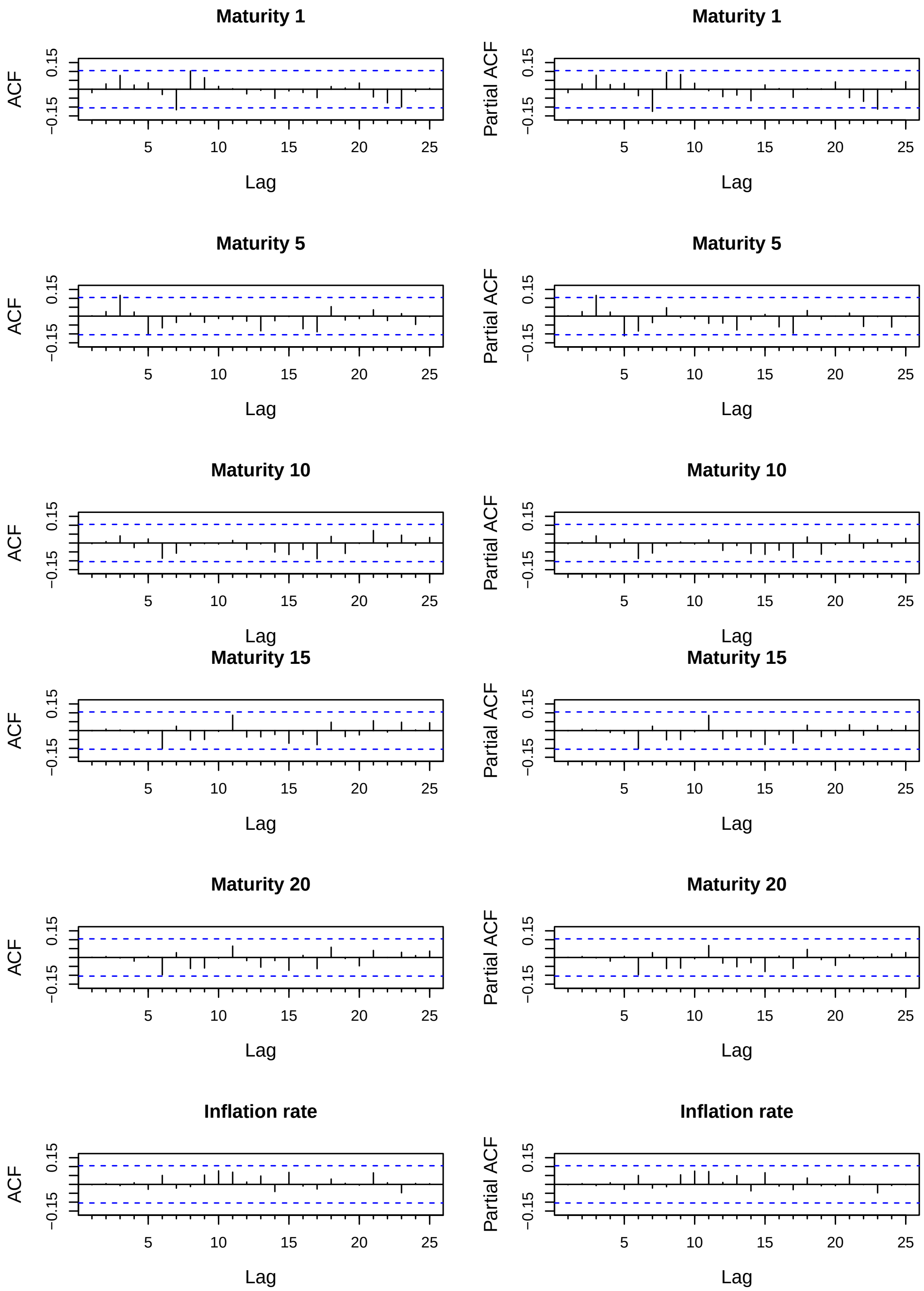

Figure B.1 Estimated ACF and PACF of the standardized residuals obtained from the ARMA models fitted in Table 9.3.

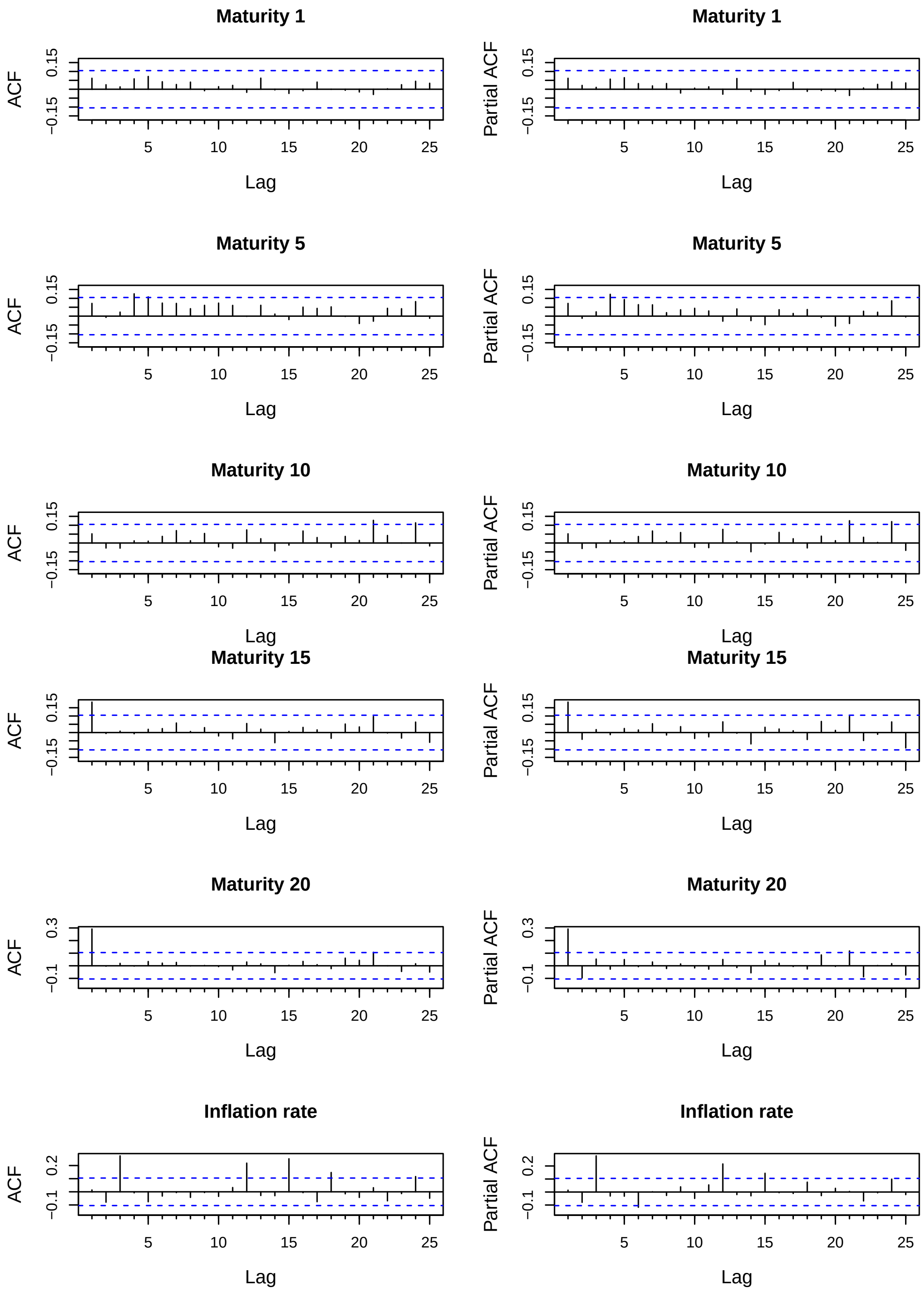

Figure B.2 Estimated ACF and PACF of the squared standardized residuals obtained from the ARMA models fitted in Table 9.3.

Marginal ARMA-GARCH fitting

The ACF and PACF plots of the standardized residuals, the squared standardized residuals and the QQ-plots with the skewed Student's t distributions and the selected univariate parametric distributions. The ARMA-GARCH model was used to form the standardized residuals.

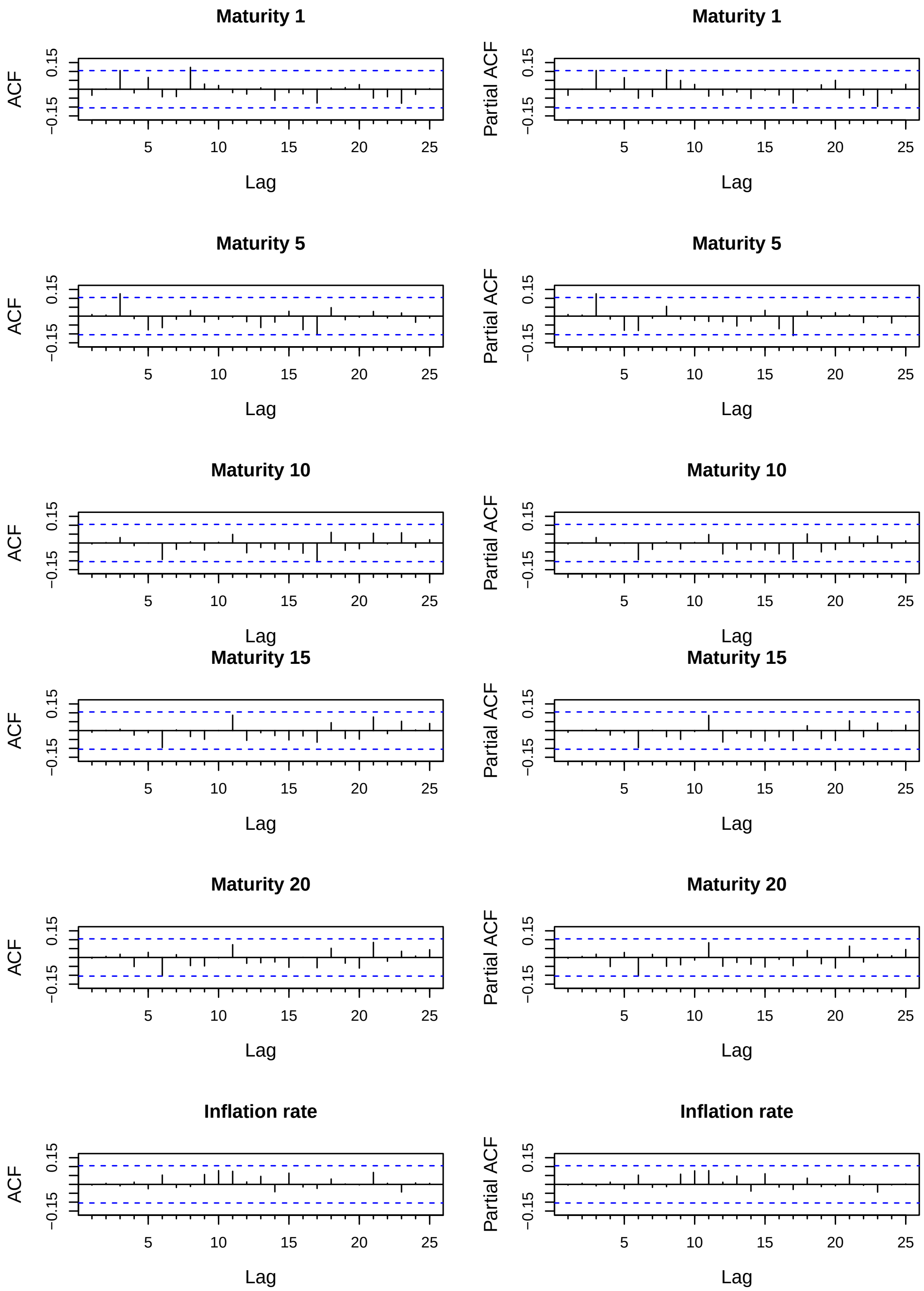

Figure B.3 Estimated ACF and PACF of the standardized residuals obtained from the ARMA-GARCH models.

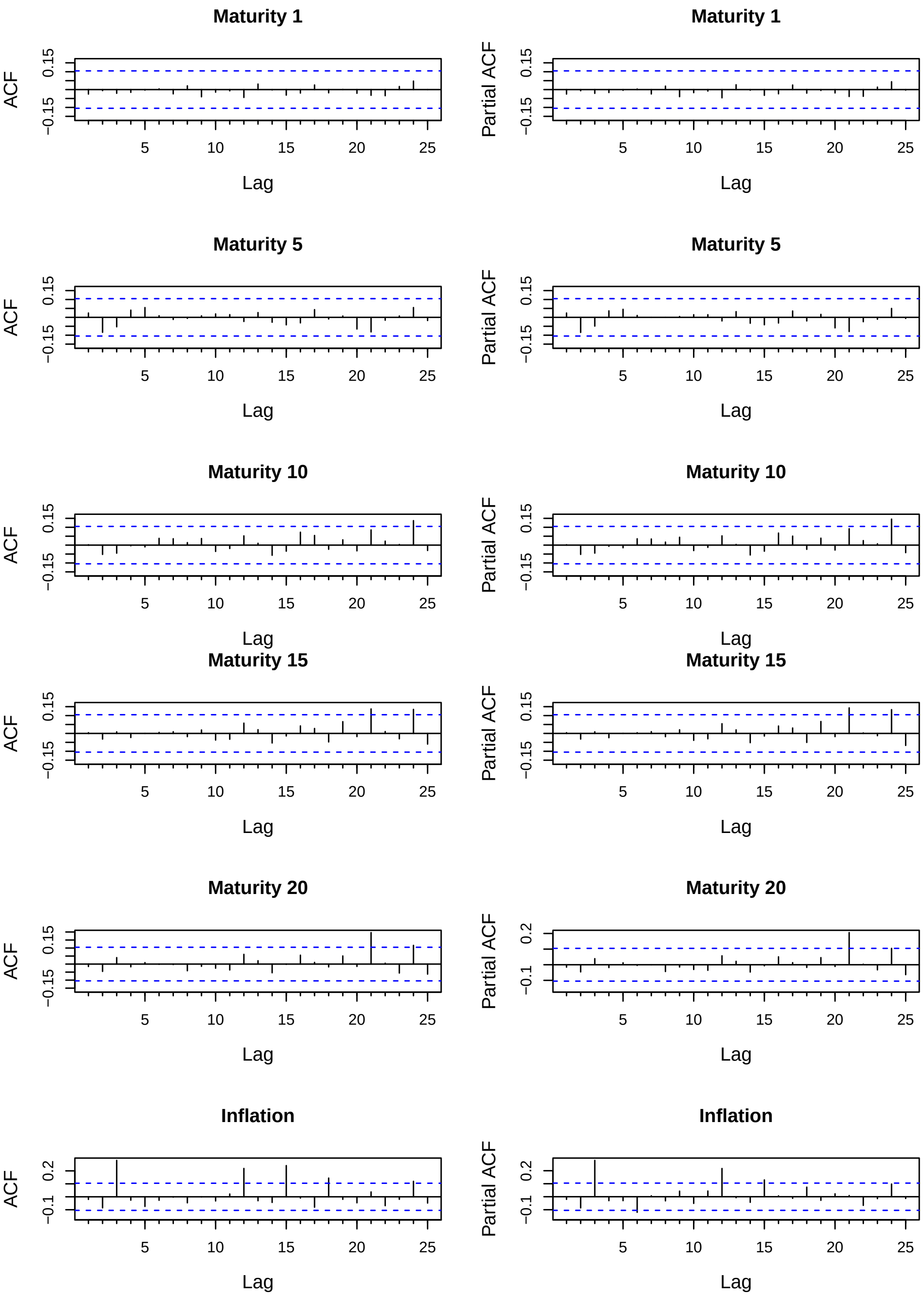

Figure B.4 Estimated ACF and PACF of the squared standardized residuals obtained from the ARMA-GARCH models.

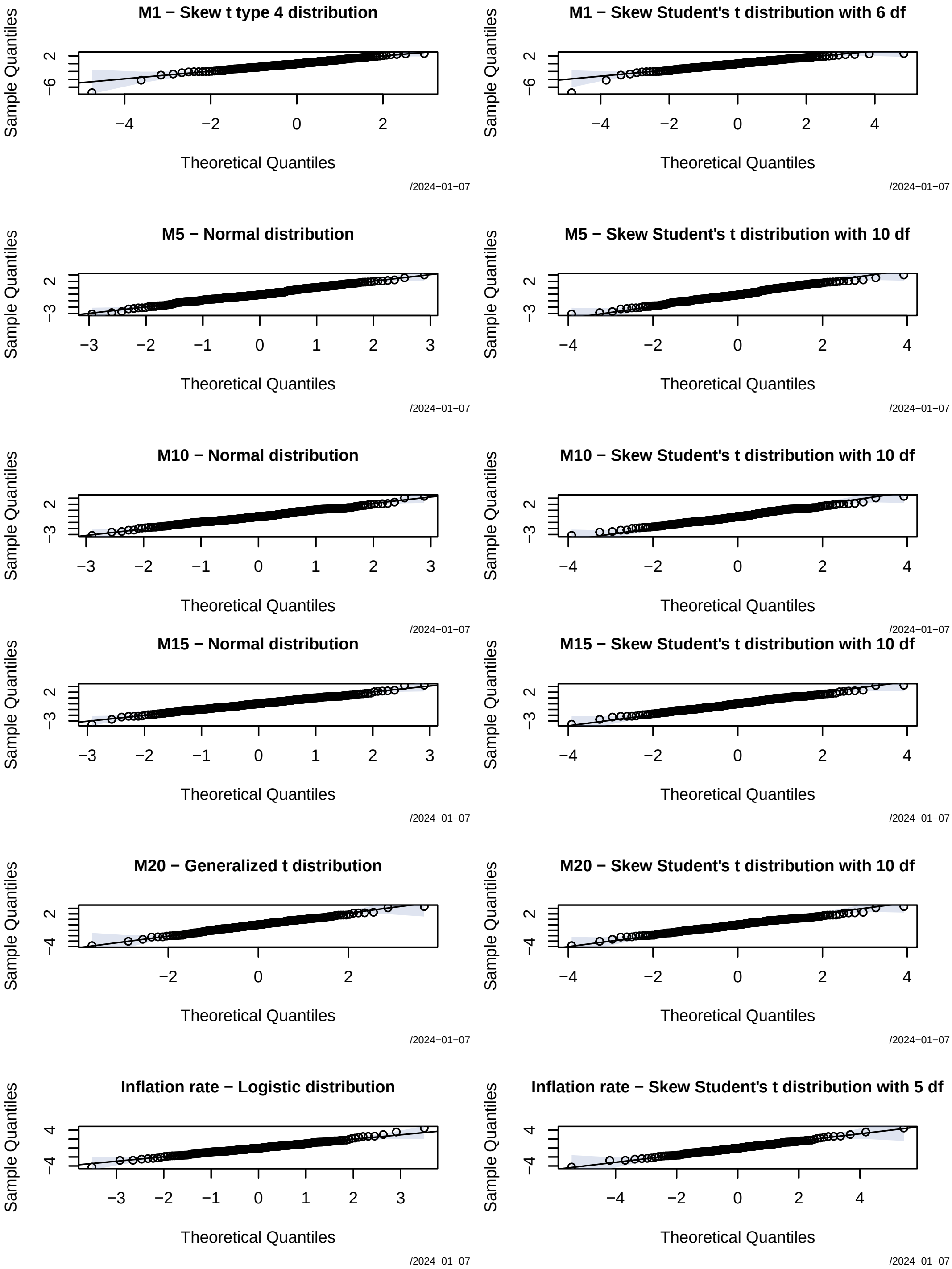

Figure B.5 QQ-Plots with skew Student's t distribution and selected univariate parametric distributions.

Forecasting

The estimated parameters obtained from the ARMA-GARCH model for the time points $T+1=352, T+2=353$ and $T+3=354$ are shown in the tables below.

Table B.1 Parameter estimation

		$T+1$	$T+2$	$T+3$
M1	$\hat{\mu}$	-0.019	0.002	-0.025
	$\hat{\sigma}$	0.085	0.084	0.089
M5	$\hat{\mu}$	-0.006	-0.038	-0.013
	$\hat{\sigma}$	0.139	0.143	0.142
M10	$\hat{\mu}$	0.014	-0.052	-0.032
	$\hat{\sigma}$	0.178	0.182	0.183
M15	$\hat{\mu}$	0.003	-0.031	-0.003
	$\hat{\sigma}$	0.182	0.184	0.187
M20	$\hat{\mu}$	-0.003	-0.024	-0.024
	$\hat{\sigma}$	0.190	0.207	0.196
Inf	$\hat{\mu}$	-0.006	-0.040	-0.029
	$\hat{\sigma}$	3.707	3.672	3.654

Forecasting using the C-vine copula model

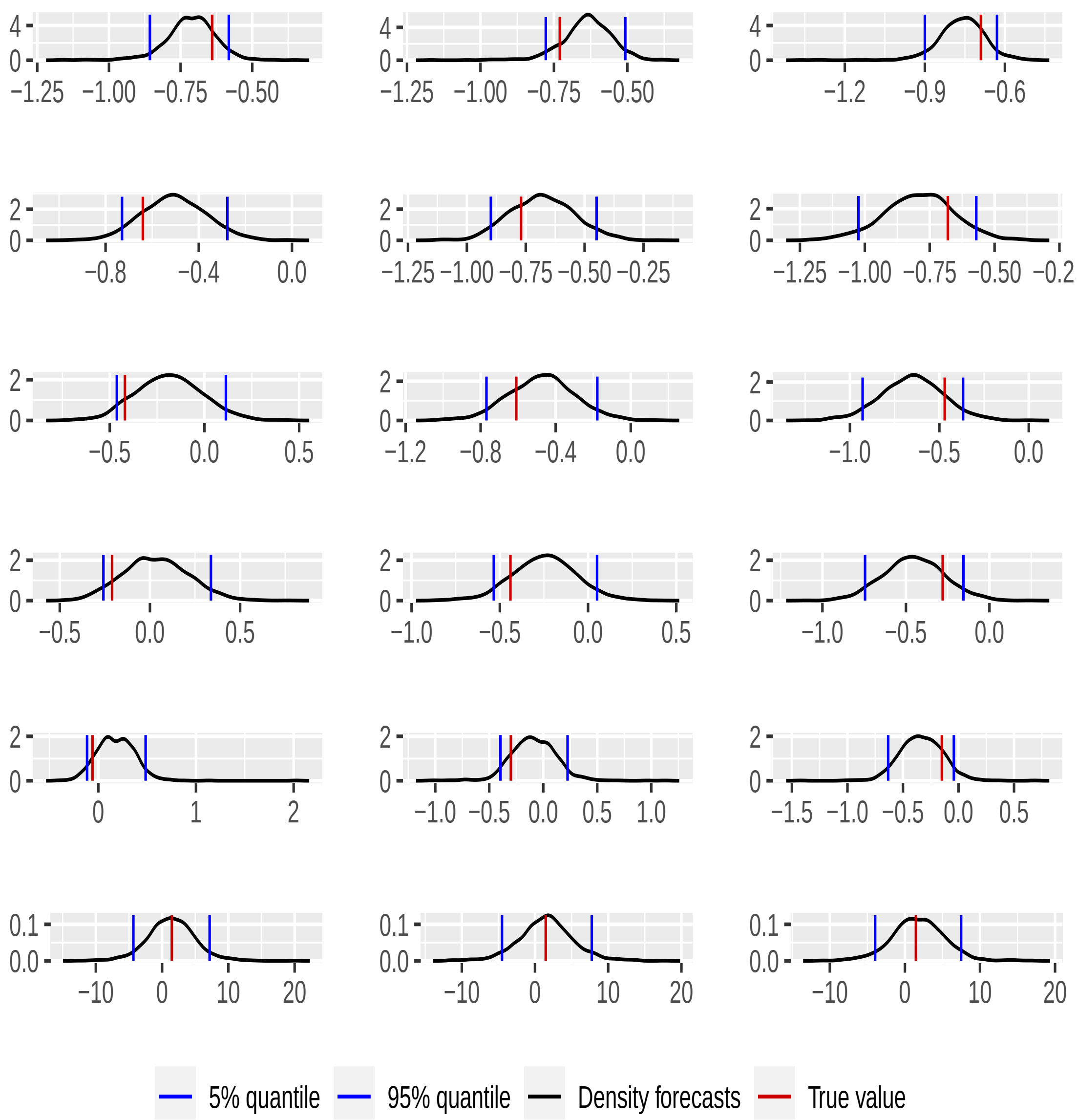

Figure B.6 The C-vine copula is used to forecast yield curves and inflation rate. The maturities M1, M5, M10, M15, M20 and Inf are presented in rows and the predicted times $T+1 = 352, T+2 = 353, T+3 = 354$ are presented in columns.

Forecasting using the D-vine copula model

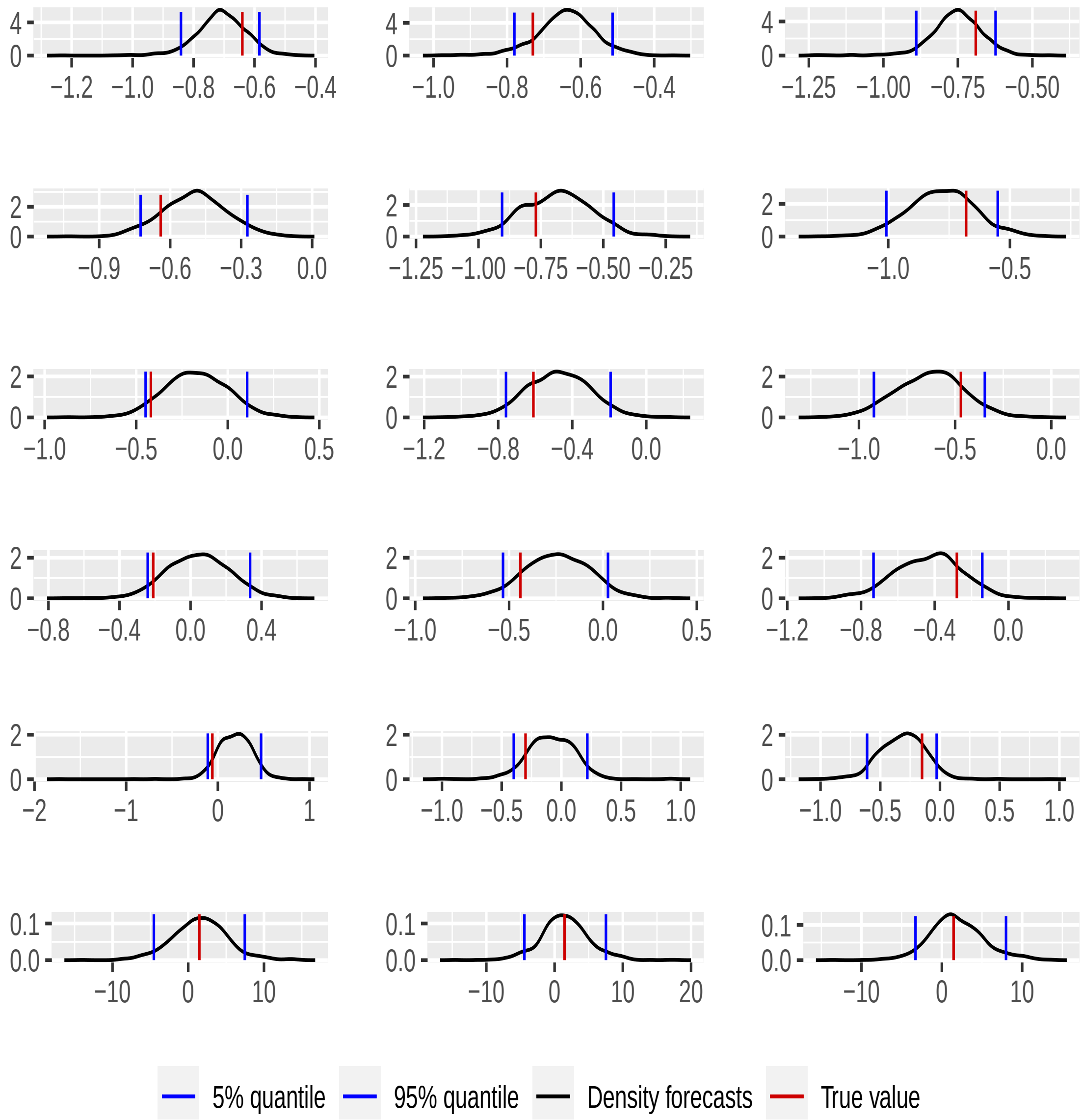

Figure B.7 The D-vine copula is used to forecast yield curves and inflation rate. The maturities M1, M5, M10, M15, M20 and Inf are presented in rows and the predicted times $T + 1 = 352, T + 2 = 353, T + 3 = 354$ are presented in columns.

Forecasting using the one-factor copula model

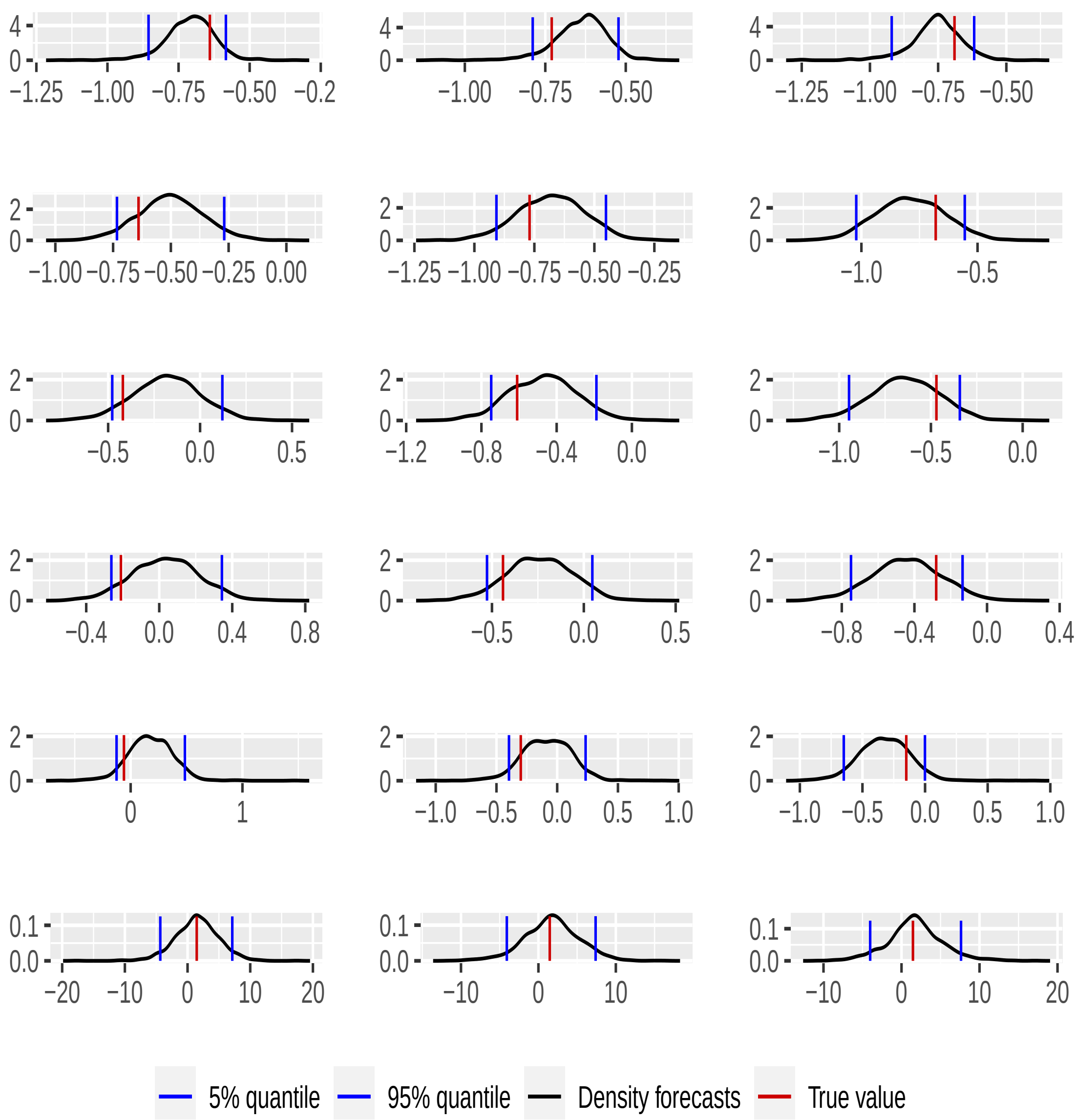

Figure B.8 The one-factor copula is used to forecast yield curves and inflation rate. The maturities M1, M5, M10, M15, M20 and Inf are presented in rows and the predicted times $T+1 = 352, T+2 = 353, T+3 = 354$ are presented in columns.

Forecasting using the two-factor copula model

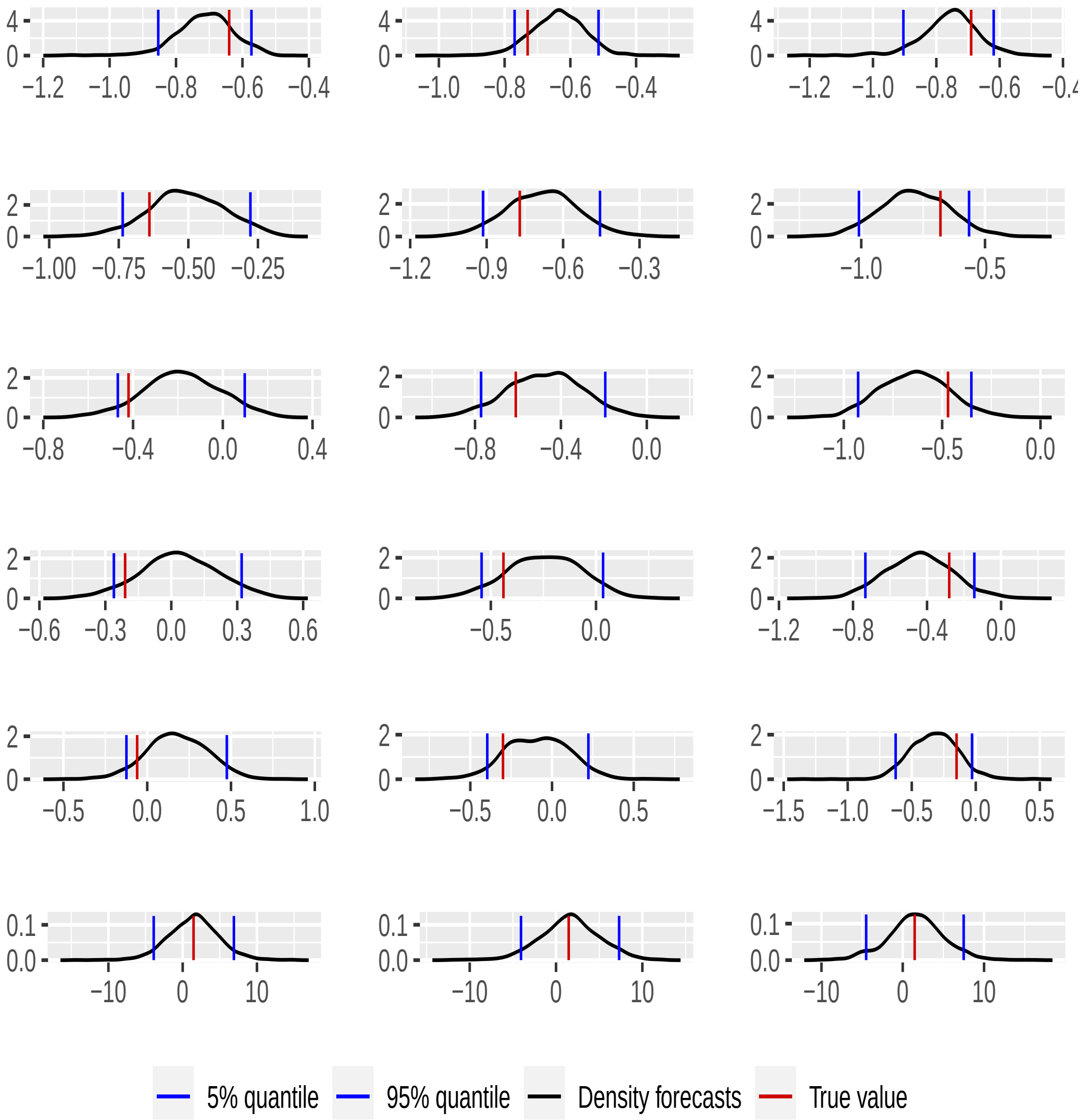

Figure B.9 The two-factor copula is used to forecast yield curves and inflation rate. The maturities M1, M5, M10, M15, M20 and Inf are presented in rows and the predicted times $T+1 = 352, T+2 = 353, T+3 = 354$ are presented in columns.

The following tables show the density mode, the prediction interval and the observed value for the R-, C- and D-vine copula and the one- and two-factor copula models.

Table B.2 Density mode and prediction interval of forecasted yield curves and inflation rate using R-, C- and D-Vine copula models and the observed value

Variable	Copula model		$T = 352$	$T = 353$	$T = 354$
		Obs. value	-0.64	-0.73	-0.69
M1	R-vine	Mode	-0.71	-0.64	-0.74
		Prediction interval	[-1.15, -0.43]	[-0.99, -0.36]	[-1.27, -0.49]
	C-vine	Mode	-0.69	-0.63	-0.73
		Prediction interval	[-1.13, -0.51]	[-1.02, -0.37]	[-1.63, -0.42]
	D-vine	Mode	-0.69	-0.64	-0.76
		Prediction interval	[-1.14, -0.47]	[-0.98, -0.31]	[-1.1, -0.42]
		Obs. value	-0.64	-0.77	-0.68
M5	R-vine	Mode	-0.48	-0.69	-0.83
		Prediction interval	[-0.97, 0.05]	[-1.07, -0.15]	[-1.24, -0.38]
	C-vine	Mode	-0.49	-0.68	-0.78
		Prediction interval	[-0.99, -0.06]	[-1.11, -0.22]	[-1.18, -0.3]
	D-vine	Mode	-0.51	-0.71	-0.83
		Prediction interval	[-0.93, -0.07]	[-1.11, -0.2]	[-1.26, -0.38]
		Obs. value	-0.42	-0.61	-0.47
M10	R-vine	Mode	-0.21	-0.52	-0.6
		Prediction interval	[-0.79, 0.51]	[-0.97, 0.16]	[-1.27, -0.07]
	C-vine	Mode	-0.22	-0.47	-0.66
		Prediction interval	[-0.82, 0.38]	[-1.01, 0.1]	[-1.24, -0.14]
	D-vine	Mode	-0.22	-0.53	-0.71
		Prediction interval	[-0.71, 0.35]	[-1, 0.16]	[-1.31, -0.07]
		Obs. value	-0.21	-0.44	-0.28
M15	R-vine	Mode	-0.01	-0.26	-0.37
		Prediction interval	[-0.56, 0.74]	[-0.77, 0.37]	[-1.05, 0.17]
	C-vine	Mode	0	-0.26	-0.47
		Prediction interval	[-0.58, 0.62]	[-0.76, 0.29]	[-1.04, 0.09]
	D-vine	Mode	0.03	-0.28	-0.49
		Prediction interval	[-0.47, 0.53]	[-0.77, 0.35]	[-1.08, 0.16]
		Obs. value	-0.06	-0.30	-0.15
M20	R-vine	Mode	0.13	-0.09	-0.23
		Prediction interval	[-0.78, 1.62]	[-0.95, 0.83]	[-1.24, 0.63]
	C-vine	Mode	0.24	-0.04	-0.37
		Prediction interval	[-0.82, 0.87]	[-0.77, 0.6]	[-1.13, 0.43]
	D-vine	Mode	0.1	-0.01	-0.4
		Prediction interval	[-0.47, 0.86]	[-0.84, 1.04]	[-1.36, 0.78]
		Obs. value	1.74	1.65	1.47
Inf	R-vine	Mode	1.78	2.05	2.29
		Prediction interval	[-11.71, 12.94]	[-14.52, 21.22]	[-14.24, 12.52]
	C-vine	Mode	1.49	1.81	1.58
		Prediction interval	[-13.4, 16.05]	[-14.05, 16.9]	[-11.34, 16.57]
	D-vine	Mode	1.77	1.73	1.55
		Prediction interval	[-9.89, 15.7]	[-12.92, 15.17]	[-15.02, 13.31]

Table B.3 Density mode and prediction interval of forecasted yield curves and inflation rate using one- and two-factor copula models and the observed value

Variable	Copula model		$T = 352$	$T = 353$	$T = 354$
		Obs. value	-0.64	-0.73	-0.69
M1	One-factor	Mode	-0.72	-0.64	-0.76
		Prediction interval	[-1.23, -0.38]	[-1.04, -0.36]	[-1.29, -0.47]
	Two-factor	Mode	0	-0.29	-0.42
		Prediction interval	[-0.55, 0.59]	[-0.89, 0.33]	[-1.03, 0.04]
		Obs. value	-0.64	-0.77	-0.68
M5	One-factor	Mode	-0.48	-0.69	-0.78
		Prediction interval	[-0.93, 0]	[-1.11, -0.23]	[-1.3, -0.28]
	Two-factor	Mode	-0.51	-0.72	-0.84
		Prediction interval	[-0.97, -0.11]	[-1.17, -0.21]	[-1.3, -0.37]
		Obs. value	-0.42	-0.61	-0.47
M10	One-factor	Mode	-0.2	-0.44	-0.63
		Prediction interval	[-0.68, 0.49]	[-1.06, 0.08]	[-1.22, 0.04]
	Two-factor	Mode	-0.21	-0.51	-0.68
		Prediction interval	[-0.75, 0.35]	[-1.1, 0.07]	[-1.22, -0.15]
		Obs. value	-0.21	-0.44	-0.28
M15	One-factor	Mode	0.01	-0.24	-0.4
		Prediction interval	[-0.48, 0.72]	[-0.88, 0.36]	[-1.02, 0.24]
	Two-factor	Mode	0	-0.29	-0.42
		Prediction interval	[-0.55, 0.59]	[-0.89, 0.33]	[-1.03, 0.04]
		Obs. value	-0.06	-0.30	-0.15
M20	One-factor	Mode	0.15	-0.03	-0.27
		Prediction interval	[-0.43, 1.78]	[-1.06, 0.94]	[-1.14, 1.13]
	Two-factor	Mode	0.13	-0.17	-0.32
		Prediction interval	[-0.7, 0.94]	[-1.34, 0.86]	[-1.15, 0.24]
		Obs. value	1.74	1.65	1.47
Inf	One-factor	Mode	2.49	1.64	2.32
		Prediction interval	[-12.85, 20.09]	[-10.54, 15.55]	[-14.86, 18.41]
	Two-factor	Mode	0.96	1.07	2.22
		Prediction interval	[-11.94, 16.89]	[-14.29, 21.19]	[-10.56, 16.28]

Bibliography

Akaike, H. (1987). Factor analysis and aic - psychometrika.

Brechmann, E. C., C. Czado, and K. Aas (2012, Mar). Truncated regular vines in high dimensions with application to financial data. *40*(1).

Brechmann, E. C. and H. Joe (2015, Jun). Truncation of vine copulas using fit indices. *138*.

Brigo, D. and F. Mercurio (2007). *Interest Rate Models - Theory and Practice With Smile, Inflation and Credit*. Berlin Heidelberg: Springer Science and Business Media.

Brockwell, P. J. and R. A. Davis (2016). *Introduction to Time Series and Forecasting*. Springer Texts in Statistics. Springer International Publishing.

Broecker, J. (2012, Jul). Evaluating raw ensembles with the continuous ranked probability score. *138*(667).

Catani, P. and N. Ahlgren (2017, Jan). Combined lagrange multiplier test for arch in vector autoregressive models. *1*.

Conrad, C. A. (2017). *Angewandte Makrooekonomie - Eine praxisbezogene Einfuehrung*. Berlin Heidelberg New York: Springer-Verlag.

Czado, C. (2019). *Analyzing Dependent Data with Vine Copulas A Practical Guide With R*. Berlin, Heidelberg: Springer.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica 50*(4), 987–1007.

Everitt, B. and T. Hothorn (2011). *An Introduction to Applied Multivariate Analysis with R*. Berlin-Heidelberg: Springer New York.

Fahrmeir, L., T. Kneib, S. Lang, and B. Marx (2013). *Regression - Models, Methods and Applications*. Berlin Heidelberg: Springer Science and Business Media.

Ferro, C. A. T., D. S. Richardson, and A. P. Weigel (2008, Mar). On the effect of ensemble size on the discrete and continuous ranked probability scores. *15*(1).

GEL, Y. R. and B. CHEN (2012). Robust lagrange multiplier test for detecting arch/garch effect using permutation and bootstrap. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique 40*(3), 405–426.

Georgii, H.-O. (2015). *Stochastik - Einfuehrung in die Wahrscheinlichkeitstheorie und Statistik*. Berlin: Walter de Gruyter GmbH and Co KG.

Gneiting, T. and A. E. Raftery (2004, Sep). Strictly proper scoring rules, prediction, and estimation. *102*(477).

Hofert, M., I. Kojadinovic, M. Maechler, and J. Yan (2023). *copula: Multivariate Dependence with Copulas*. R package version 1.1-2.

Hull, J. (2015). *Optionen, Futures und andere Derivate*. Muenchen: Pearson.

Ivan Kojadinovic and Jun Yan (2010). Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software 34*(9), 1–20.

Jun Yan (2007). Enjoy the joy of copulas: With a package copula. *Journal of Statistical Software 21*(4), 1–21.

Kadhem, S. H. and A. K. Nikoloulopoulos (2021, Mar). Factor copula models for mixed data. *74*(3).

Kadhem, S. H. and A. K. Nikoloulopoulos (2023). *FactorCopula: Factor, Bi-Factor, Second-Order and Factor Tree Copula Models*. R package version 0.9.3.

Krupskii, P. and H. Joe (2013, September). Factor copula models for multivariate data. *120*.

Krupskii, P. and H. Joe (2015, Jun). Structured factor copula models. *138*.

Kwok, Y.-K. (2008). *Mathematical Models of Financial Derivatives*. Berlin Heidelberg: Springer Science and Business Media.

Lee, J. H. (1991). A lagrange multiplier test for garch models. *Economics Letters 37*(3), 265–271.

Marius Hofert and Martin Mächler (2011). Nested archimedean copulas meet R: The nacopula package. *Journal of Statistical Software 39*(9), 1–20.

Nagler, T., U. Schepsmeier, J. Stoeber, E. C. Brechmann, B. Graeler, and T. Erhardt (2023). *VineCopula: Statistical Inference of Vine Copulas*. R package version 2.4.5.

Nagler, T. and T. Vatter (2023). *rvinecopulib: High Performance Algorithms for Vine Copula Modeling*. R package version 0.6.3.1.1.

Neusser, K. (2011). *Zeitreihenanalyse in den Wirtschaftswissenschaften*. Berlin Heidelberg: Springer Science Business Media.

Rigby, R. A., M. D. Stasinopoulos, G. Z. Heller, and F. D. Bastiani (2019). *Distributions for Modeling Location, Scale, and Shape Using GAMLSS in R*. Boca Raton, Fla: CRC Press.

Rossi, R. J. (2018). *Mathematical Statistics - An Introduction to Likelihood Based Inference*. New York: John Wiley Sons.

Shapiro, S. S. (1964). *An Analysis of Variance Test for Normality Complete Samples*. Freuburg i.B.

Shumway, R. H. and D. S. Stoffer (2017). *Time Series Analysis and Its Applications With R Examples*. Berlin, Heidelberg: Springer.

Stasinopoulos, M. and R. Rigby (2022). *gamlss.dist: Distributions for Generalized Additive Models for Location Scale and Shape*. R package version 6.0-5.

Tsay, R. S. (2009). *ANALYSIS OF FINANCIAL TIME SERIES, 2ND ED -*. New Delhi: Wiley India Pvt. Limited.

Wuertz, D., Y. Chalabi, T. Setz, and M. Maechler (2022). *fGarch: Rmetrics - Autoregressive Conditional Heteroskedastic Modelling*. R package version 4022.89.